TP SdF N° 20

La loi de Weibull

Ce TP porte sur la loi de Weibull et ses techniques d'ajustement.

1 – Caractéristiques de la loi de Weibull

Présenter les principales caractéristiques de la loi de Weibull.

2 – Ajustement d'un modèle de Weibull

Expliquer les différentes techniques d'ajustement (graphiques et par calcul) de la loi de Weibull à 2 et 3 paramètres en soulignant leurs limitations éventuelles.

Réaliser un ajustement à partir du recueil de données censurées suivant :

Durées avant panne :

3510	3183	8408	4808	5587	4591	5198	267	5164	5234	7261	5 575	6021	6483
1978	1767	4367	2320	4223	6477	5212	3129	8908	1810	4410	4 724	3840	2744
4356	3955	5222	7959	4940	3507	5693	1110	4573	1746	2996	4826	3415	4562

Durées sans panne :

3777	5171	8693	7091	958	3866	3694

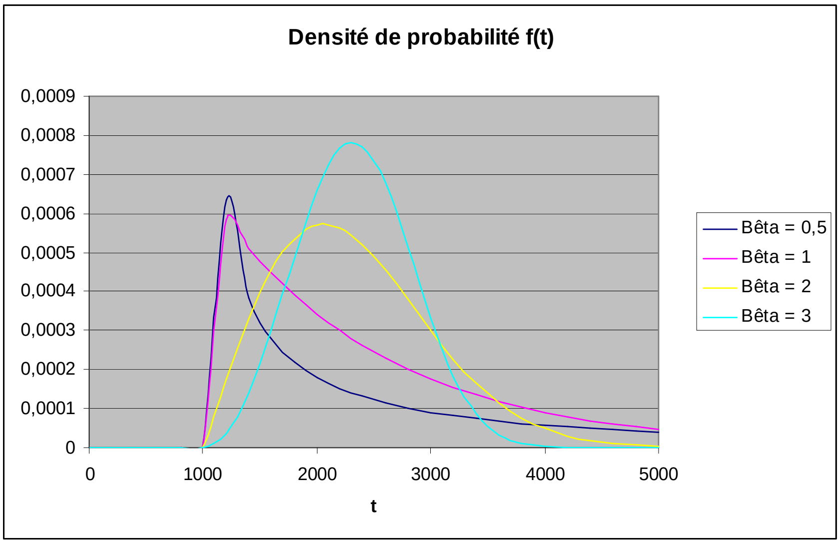
3 – Utilisation de la loi de Weibull

Conclure sur l'utilisation de la loi de Weibull et l'emploi des diverses techniques d'ajustement.

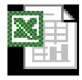
1 – La loi de Weibull

Proposée par l'ingénieur et mathématicien suédois Ernst Hjalmar Waloddi Weibull[1] (1887-1979), la loi de Weibull est une loi de probabilité à 3 paramètres qui est très utilisée pour modéliser la durée de vie des produits en raison de sa grande flexibilité. Elle est caractérisée par :

Sa densité de probabilité : $\mathbf{f(t) = \beta(t-\gamma)^{\beta-1}/\sigma^{\beta} \exp(-[(t-\gamma)/\sigma]^{\beta})}$ avec $\beta > 0$, $\sigma > 0$ et $t > \gamma$

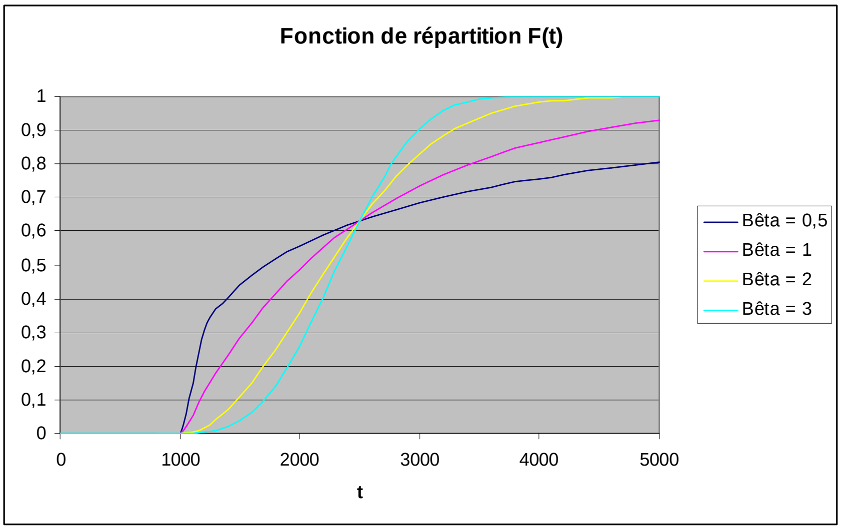


Ouverture du fichier correspondant par double clic :



Feuille de calcul Microsoft Excel

Et sa fonction de répartition : $\mathbf{F(t) = 1- \exp(-[(t-\gamma)/\sigma]^{\beta})}$



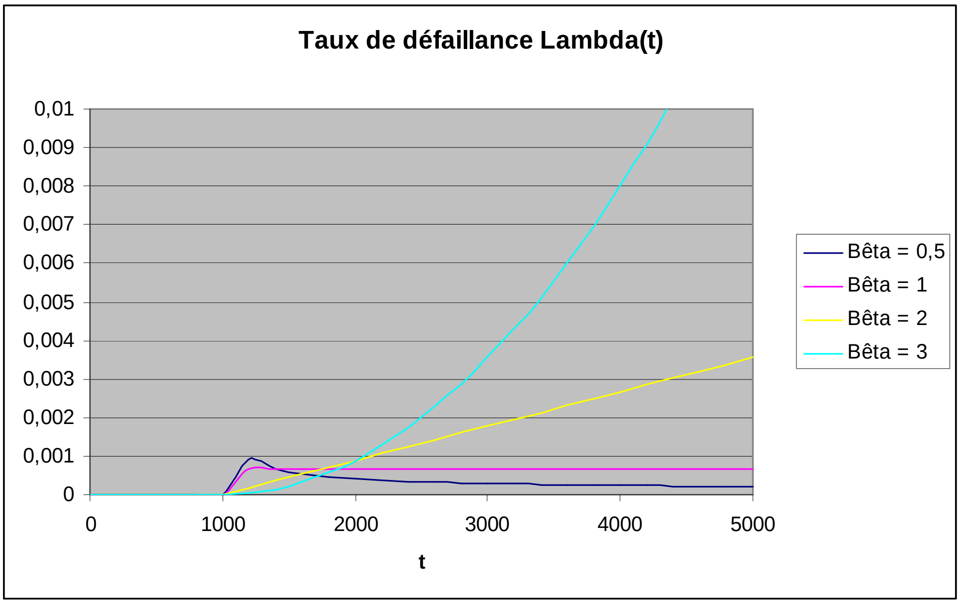
β (0,5 à 3) est le paramètre de forme, σ (1500) le paramètre d'échelle et γ (1000) le paramètre de localisation de la distribution par rapport à l'origine.

Pour la valeur de $t = \gamma + \sigma$, on observe un quantile caractéristique à 63 % quelle que soit la valeur du paramètre β ($F(t) = 1- \exp(-[(t-g)/\sigma]^{b}) = 1-1/e = 0{,}632120559$). Ce taux de mortalité de 63 % de la population, indépendant de β, est notamment considéré pour définir le taux d'accélération dans le

[1] Wallodi Weibull, "A Statistical Distribution Function of Wide Applicability", *ASME Journal Of Applied Mechanics Paper*, 1951. The hallmark paper of Weibull analysis.

cadre d'essais accélérés (rapport entre les durées correspondant à ce quantile avec ou sans conditions de stress renforcées : $\sigma = F_s\, \sigma_s$ si $\gamma = 0$).

Le taux de défaillance de la loi de Weibull est $\lambda(t) = \beta(t-\gamma)^{\beta-1}/\sigma^{\beta}$, puisque $\lambda(t) = f(t) / (1-F(t))$ par définition.



Le taux de défaillance est croissant si $\beta > 1$, constant si $\beta = 1$, décroissant si $\beta < 1$. La loi exponentielle est une loi de Weibull de paramètres $\beta = 1$, $\sigma = 1/\lambda$ et $\gamma = 0$.

La moyenne et la variance de loi de Weibull s'expriment de la manière suivante :

$$\text{Moyenne : } \gamma + \sigma\, \Gamma\, [(1+\beta)/\beta] \qquad \text{Variance : } \sigma^2\, [\, \Gamma(1+2/\beta) - \Gamma^2(1+1/\beta)]$$

à partir de la fonction gamma : $\Gamma(\beta)=\int_0^{\infty} x^{\beta-1}\, e^{-x}dx$ ou $(\beta-1)\,!$ pour des valeurs entières.

2 – Ajustement d'un modèle de Weibull

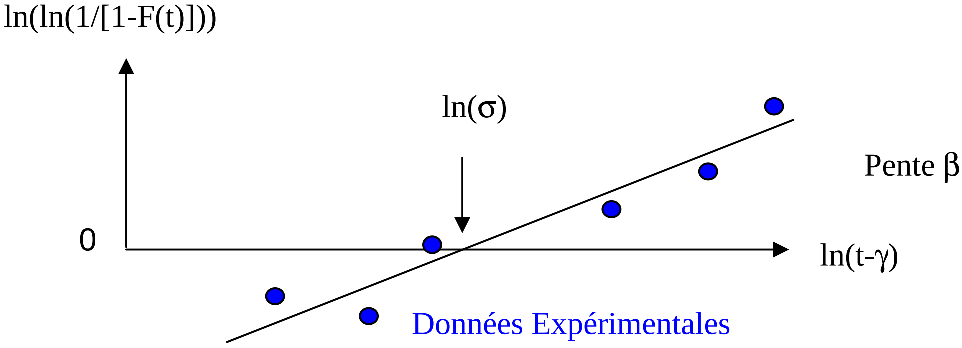
L'ajustement consiste à trouver les paramètres d'une fonction mathématique afin de la faire correspondre au mieux à une courbe expérimentale.

2.1 – Ajustement graphique

L'ajustement graphique consiste à effectuer un changement de variables judicieux permettant de ramener l'ajustement à une simple régression linéaire, ce que permet la loi de Weibull :

$$F(t) = 1-\exp(-((t-\gamma)/\sigma)^{\beta}) \Rightarrow 1-F(t) = \exp(-((t-\gamma)/\sigma)^{\beta}) \Rightarrow -\ln[1-F(t)] = \ln(1/[1-F(t)]) = ((t-\gamma)/\sigma)^{\beta}$$

$$\text{D'où : } \ln(\ln(1/[1-F(t)])) = \beta \ln(t-\gamma) - \beta \ln(\sigma)$$



La pente de la droite est β et la droite coupe l'axe des abscisses à la valeur $\ln(\sigma)$. L'origine nulle de l'axe des ordonnées correspond à nouveau au quantile caractéristique de 63 % :

$$\ln(\ln(1/[1-F(t)])) = 0 \Rightarrow \ln(1/[1-F(t)]) = 1 \Rightarrow 1/[1-F(t)] = e \Rightarrow F(t) = 1-1/e = 0{,}632120559$$

Deux difficultés subsistent cependant :

- Définir la valeur des quantiles $F(t_i)$ correspondant aux données expérimentales,
- Estimer la valeur de γ pour lequel l'ajustement est linaire.

Plusieurs méthodes peuvent être utilisées pour surmonter la première difficulté :

- Remplacer $1/1-F(t_i) = 1/R(t_i)$ par $(n+1)/(n+1-i)$ avec i la ième panne parmi n équipements (1 est ajouté aux deux termes du quotient pour éviter la valeur nulle au dénominateur),
- Calculer les valeurs des $F(t_i)$ à 50 % (borne de l'intervalle de confiance unilatéral à 50 % de la loi binomiale calculée par résolution de l'équation $\sum_0^i C_i^N F(t_i)^i (1- F(t_i))^{N-i} = 50$ %)
- Utiliser l'approximation proposée par A. Benard [2] : $F(t_i) = (i-0{,}3)/(n+0{,}4)$
- Calculer les valeurs des $F(t_i)$ par la méthode de Kaplan-Meier, notamment dans le cas de données censurées : $1-F(ti) = \Pi_{k=1}^{i} (n_k-d_k)/n_k$ avec n_k le nombre d'équipements encore en vie juste avant t_k et d_k le nombre de défaillances à t_k.

Dans le cas d'un nombre de données réduites, ces méthodes peuvent donner des résultats assez différents comme le montre le tableau suivant :

ti	i	(i+1)/(N+1)	Kaplan-Meyer	Benard	$F(ti)_{50\%}$
16	1	0,2857	0,1667	0,1094	0,1091
34	2	0,4286	0,3333	0,2656	0,2644
53	3	0,5714	0,5000	0,4219	0,4214
75	4	0,7143	0,6667	0,5781	0,5786
93	5	0,8571	0,8333	0,7344	0,7356
120	6	1,0000	1,0000	0,8906	0,8909

Feuille de calcul Microsoft Excel

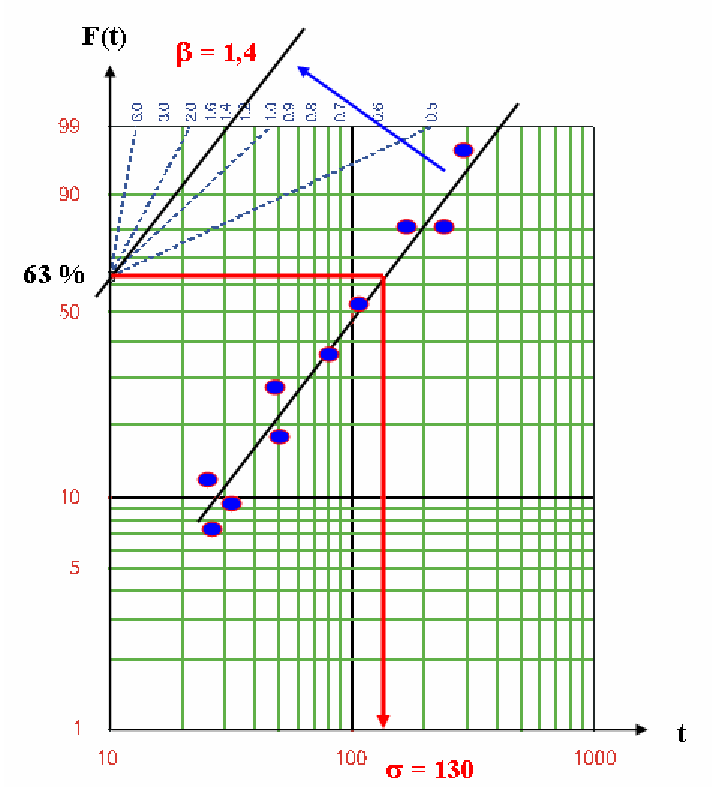
La deuxième difficulté disparaît si $\gamma = 0$ (ajustement d'une loi de Weibull à 2 paramètres) bien que ce cas exclut la majorité des phénomènes réels pour lesquels les défaillances n'apparaissent qu'après le franchissement d'un certain seuil d'usure.

2.2 – Utilisation du papier de Weibull

Le papier de Weibull utilise le changement de variables indiqué précédemment tout en gardant des échelles de graduation en t et F(t). La valeur σ correspond à l'abscisse du point d'ordonnée 63 % et une sorte de rapporteur d'angle placé en haut à gauche du papier permet de lire directement la valeur β, correspondant à la pente de la droite, en traçant une parallèle à cette dernière.

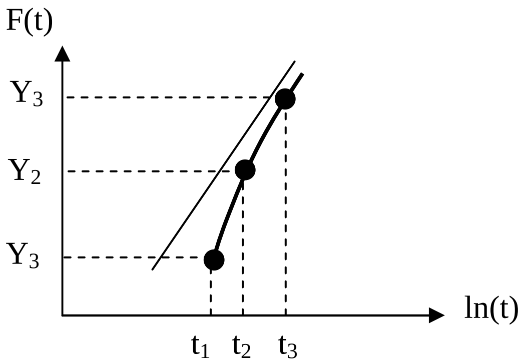
---

[2] Benard, A. and Bos-Levenbach, E. C. (1953): Het uitzetten van waarnemingen op waarschijnlijkheids-papier. *Statistica Neerlandica*, Vol. 7 pp. 163-173. English translation by Schop, R. (2001): The Plotting of Observations on Probability Paper. Report SP 30 of the Statistical Department of the Mathematics Centrum, Amsterdam.



Papier de Weibull

Le papier de Weibull ne permet d'ajuster que des lois à 2 paramètres et la droite se transforme en courbe si γ est différent de zéro. La valeur γ peut cependant être estimée préalablement à partir de 3 points d'une telle courbe par la méthode proposée par J.David[3], avant de procéder à l'ajustement proprement dit avec des points d'abscisse t_i-γ.



Si les points correspondent à une loi de Weibull, on peur écrire la condition de linéarité :

$$(Y_3-Y_2)/(\ln(t_3-\gamma)-\ln(t_2-\gamma)) = (Y_2-Y_1)/(\ln(t_2-\gamma)-\ln(t_1-\gamma))$$

De plus, si les 3 points sont choisis tels que Y_3-Y_2 = Y_2-Y_1 on obtient :

$$\ln(t_3-\gamma)-\ln(t_2-\gamma) = \ln(t_2-\gamma)-\ln(t_1-\gamma) \Rightarrow (t_3-\gamma)/(t_2-\gamma) = (t_2-\gamma)/(t_1-\gamma) \Rightarrow (t_3-\gamma)(t_1-\gamma) = (t_2-\gamma)^2$$

$$\text{Soit : } \gamma = (t_2^2-t_1t_3)/(2t_2-(t_1+t_3))$$

[3] J. David, Détermination sans tâtonnement du coefficient de γ de la loi de Weibull, *Revue de statistique appliquée,* tome 23, n°3 (1975), p. 81-85.

2.3 – Résolution mathématique de la régression linéaire

Calculée par la méthode des moindres carrés (Min Σ_1^n (yi – axi -b)^2), les coefficients de la droite y = ax + b ont pour expression :

$$a = cov(x,y) / V(x) \quad et \quad b = Y- aX$$

avec : $cov(x,y) = 1/n\ \Sigma_1^n\ (xi-X)(yi-Y)$ $V(x) = 1/n\ \Sigma_1^n\ (xi-X)^2$ $X = 1/n\ \Sigma_1^n\ xi$ $Y = 1/n\ \Sigma_1^n\ yi$

L'ajustement peut être réalisé par résolution mathématique de la régression linéaire en conservant toutefois les limitations relatives à l'estimation de la valeur des quantiles et du paramètre γ. La méthode du maximum de vraisemblance permet de surmonter ces difficultés.

2.4 – Ajustement par la méthode du maximum de vraisemblance

La méthode du maximum de vraisemblance (maximum likelihood) consiste à rechercher le modèle théorique qui donne la densité de probabilité maximale pour les données expérimentales, soit la valeur des paramètres qui maximise le produit $\boldsymbol{\Pi_1^n}$ **f(ti)** pour n durées de fonctionnement d'équipement avant panne.

Dans le cas de données de retour d'expérience, généralement censurées à droite, on multiplie ce produit par la probabilité de non-apparition de panne au moment de la censure pour chacune des durées correspondantes, soit $\boldsymbol{\Pi_1^n\ f(ti) * \Pi_1^m\ (1\text{-}F(tj))}$ dans le cas de m données censurées.

Si les données sont nombreuses, il est préférable de maximiser la somme des logarithmes des différents termes du produit, dont certains ont des valeurs très proches de zéro, pour éviter des problèmes de calcul numérique.

L'optimisation est cependant délicate car la vraisemblance comprend plusieurs maxima locaux :

- Certains outils tentent de résoudre le système d'équations résultant de l'annulation des dérivées partielles de la vraisemblance en Bêta et Sigma, en adoptant une recherche itérative suivant le paramètre Gamma.

- L'ajustement réalisé par l'outil SIMCAB est basé sur la méthode locale du Simplexe (algorithme de Nelder-Mead) menée à partir d'un point initial judicieusement choisi.

- L'ajustement peut être également réalisé de manière encore plus efficace par l'outil GENCAB qui couple le simplexe (méthode locale) aux algorithmes génétiques (méthode globale). L'application demandée est traitée, ci-après, au moyen de cet outil.

Ajustement d’un modèle de Weibull

Bêta :	2,81		Σln(f(Ti)) :	-376,74016	*Pannes*
Sigma :	6042		Σln(1-F(Tj))	-7,53121318	*Censures*
Gamma :	-630		Σ :	-384,271373	

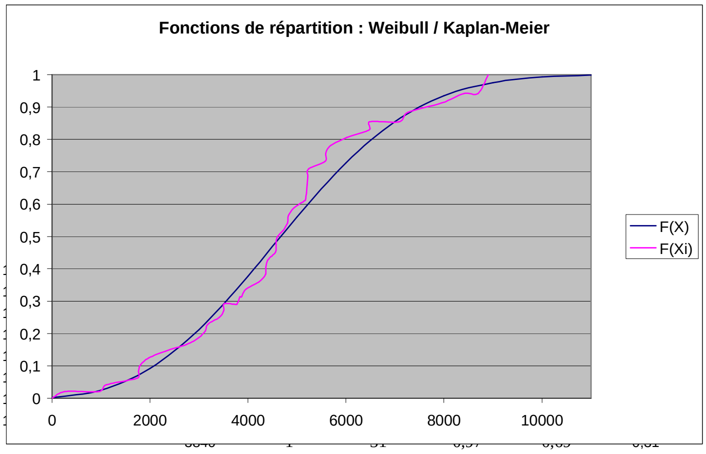
Vraisemblance

Durées avant panne

T_i	ln(f(T_i))
3510	-8,70305252
3183	-8,78072698
8408	-10,04576
4808	-8,60734347
5587	-8,70473809
4591	-8,60064812
5198	-8,64177397
267	-11,1321123
5164	-8,63762946
5234	-8,64640016
7261	-9,30756057
5575	-8,7023677
6021	-8,80887668
6483	-8,9594479

Durées avant censure

T_j	ln(1-F(T_j))
3777	-0,41189769
5171	-0,89193654
8693	-3,38530076
7091	-1,99253653
958	-0,02336972
3866	-0,43571209
3694	-0,39045984



Feuille de calcul Microsoft Excel

Suite à l’ajustement, différents tests statistiques (khi-2, Kolmogorov-Smirnov…) peuvent être employés avant d’accepter ou de rejeter le modèle théorique. Ces tests consistent à évaluer de différentes manières l’écart entre les fonctions de répartition du modèle théorique et des données expérimentales puis d’en déduire une probabilité d’erreur via les tables correspondantes.

3 – Conclusion sur l’utilisation de la loi de Weibull et l’emploi des diverses techniques d’ajustement

En raison de sa flexibilité, la loi de Weibull est très utile pour modéliser la fiabilité d’équipements divers à partir d’un retour d’expérience.

L’ajustement graphique ne présente plus vraiment d’intérêt car les méthodes automatisées, basées sur la méthode du maximum de vraisemblance, se révèlent beaucoup plus précises, dans la mesure où elles possèdent une réelle capacité de s’affranchir des divers optima locaux.

Cependant la représentation graphique garde tout son intérêt car elle donne une image visuelle de la qualité de l’ajustement réalisé. En outre, elle permet d’identifier d’éventuels modes de défaillance différents matérialisés par plusieurs portions de droite sur le graphique (dans sa version 9 l’outil SIMCAB propose une telle représentation après ajustement et différents tests statistiques réalisés par rapport au seuil de 5 % de risque de première espèce).