REVUE DE STATISTIQUE APPLIQUÉE

J. REYROLLE

Vieillissement et maintenance d'un système fortement modulaire. Estimation et factorisation des paramètres d'un modèle de Weibull

Revue de statistique appliquée, tome 27, nº 3 (1979), p. 5-14

<http://www.numdam.org/item?id=RSA_1979__27_3_5_0>

VIEILLISSEMENT ET MAINTENANCE D'UN SYSTÈME FORTEMENT MODULAIRE ESTIMATION ET FACTORISATION DES PARAMÈTRES D'UN MODÈLE DE WEIBULL*

D'après une étude réalisée sur le fichier "rails" SNCF

par J. REYROLLE

(ex Conseiller scientifique au service "Recherche")

RESUME

La littérature technique abonde sur la fiabilité de systèmes dont les pannes surviennent avec un taux d'incidence constant. Cette hypothèse nie le vieillissement. Elle rejette l'origine des pannes à l'extérieur du système. Ceci devrait nous rendre plus circonspects quant aux valeurs M.T.B.F.(1) trouvées dans les banques de données.

Nous nous plaçons ici dans une hypothèse d'usure accélérée, en choisissant le modèle de WEIBULL pour représenter la croissance du taux d'incidence. La mission particulière du système étudié exige l'intégrité de chaines formées d'éléments semblables et nombreux en série. Le paramètre d'échelle de ces éléments dépend de divers facteurs sur lesquels on attend des choix techniques. Le paramètre de forme est commun et fort. L'organisation de la maintenance impose alors une limite au-delà de laquelle il devient plus économique de remplacer une portion du système que de continuer à l'entretenir. Cette limite simplifie l'expression des statistiques exhaustives relatives au paramètre d'échelle. Le présent article développe quelques conséquences de cette simplification quant aux problèmes de jugement sur échantillon rencontrés dans les choix techniques ; puis une factorisation du paramètre par régression pondérée. Enfin il propose un moyen de constituer une banque de données susceptibles d'éviter l'écueil habituel de colinéarité et ses conséquences sur l'interprétation des résultats.

PLAN

I – Système étudié
II – Modèle fiabiliste
III – Objectif de l'étude
IV – Problèmes de jugement sur échantillon relatifs au paramètre d'échelle
V – Modélisation du paramètre d'échelle
VI – Généralisation : cas ou le paramètre de forme est inconnu
VII – Quelques difficultés relatives à la définition des modalités pour certains facteurs
VIII – Plan d'observation visant une utilisation optimale de l'information
IX – Note bibliographique succincte

(*) Cet article a fait l'objet d'un exposé aux Journées Nationales de la qualité le 28.3.79.

(1) M.T.B.F. "Moyenne des Temps de Bon Fonctionnement" ou "Mean Time Between Failures" (terminologie anglo-saxonne).

Mots-Clés : **Weibull, Maintenance, Modélisation**

I – SYSTEME ETUDIE

Le système peut être un réseau de communication ou d'alimentation, voire un échangeur tubulaire, etc. Sa continuité doit être assurée pour qu'il puisse accomplir sa mission. Sur le graphe qui en donne l'image, chaque arc est constitué d'éléments nombreux, identiques quant à leur fonction et à peu près quant à leur constitution, assemblés bout à bout : nous les appellerons "maillons". Ils sont sujets à une usure liée à l'accomplissement de leur mission. Ils sont remplaçables, soit isolément pour avarie individuelle, soit en groupe pour parer à leur vieillissement. Nous appellerons "tronçon isochrone" un tel groupe de maillons successifs mis en service simultanément au cours d'une opération de renouvellement, et rigoureusement identiques quant à leur constitution.

Le découpage en maillons peut n'être que fictif, le maillon correspondant alors à l'unité de longueur, ou mieux au quantum de remplacement en cas d'avarie individuelle.

II – MODELE FIABILISTE

L'avarie d'un maillon est sans incidence sur l'état des maillons voisins, mais la discontinuité qu'elle engendre pénalise le service rendu par le réseau, du moins sur les chemins empruntant l'arc dont un maillon au moins est avarié.

Chaque maillon subit un taux instantané d'avarie croissant comme une fonction puissance du "flot" écoulé depuis sa pose sur l'arc auquel appartient le maillon. C'est-à-dire que le flot écoulé avant avarie suit une loi de WEIBULL dont nous désignerons les paramètres, suivant les notations habituelles :

β = paramètre de forme, *commun à tous les maillons*

η = paramètre d'échelle, variable (et nous poserons $\lambda = \eta^{-\beta}$)

$F(t) = 1 - \exp[-(t/\eta)^{\beta}] = 1 - \exp[-\lambda t^{\beta}]$ = fonction de répartition des flots t écoulés avant avarie (1)

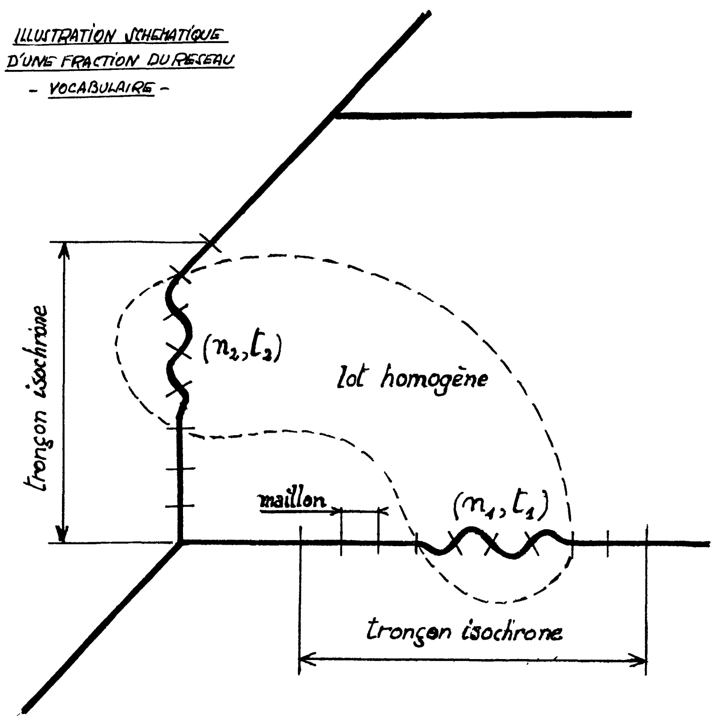
$h(t) = [1 - F(t)]^{-1} \, d[F(t)]/dt = \lambda \beta t^{\beta - 1}$ = taux instantané d'avarie.

Le paramètre de forme β peut être connu par une étude antérieure, ou inconnu (voir paragraphe 6 : généralisation). Il est seulement supposé supérieur à 2, de sorte que le taux instantané d'avarie h(t) croit rapidement. Il s'ensuit que tout maillon ayant supporté un certain flot est affecté d'un taux instantané d'avarie trop important pour rester compatible avec des opérations de maintenance raisonnable : il y a lieu de renouveler tout le tronçon isochrone.

Pratiquement ce renouvellement a lieu dès que le taux de remplacement individuel atteint pour le tronçon une valeur de l'ordre de 0,10. Cette valeur est suffisamment faible pour qu'on puisse négliger la probabilité concernant l'occurrence de 2 avaries successives sur le même maillon.

---

(1) On introduit souvent dans le modèle de WEIBULL un 3^{e} paramètre t_0, dit de positon. C'est-à-dire que t est remplacé par $(t - t_0)$ dans les expressions de f et h, qui sont nulles pour tout $t < t_0$. Outre que l'échantillon nécessaire à l'estimation simultanée de 3 paramètres devient considérable, il faut remarquer qu'un tel modèle implique une modification du processus physique de vieillissement : inexistant pour $t < t_0$, il ne se mettrait en œuvre que pour $t > t_0$.



III – OBJECTIFS DE L'ETUDE

L'homogénéité de constitution et de mission d'un tel réseau est généralement toute relative, de sorte que le paramètre d'échelle η est très variable suivant le maillon considéré.

L'hétérogénéité résulte de facteurs qu'on peut grossièrement répartir comme ci-dessous :

a) constitution : plusieurs modalités qualitatives (matière, fournisseur, . . .) ;

b) mission : facteurs liés à la nature et à l'intensité du flot, ils caractérisent son agressivité (difficiles à quantifier si le flot varie dans le temps) ;

c) environnement : nombreux facteurs susceptibles de moduler plus ou moins l'agressivité du flot (profil de l'arc, ambiance, . . .) ;

On désire d'abord permettre des choix techniques relatifs au groupe (a), entièrement maîtrisable, et au groupe (b) qui ne l'est guère, du moins à court terme, étant admis qu'on subit entièrement la variabilité présentée par les facteurs (c).

Ensuite, on voudrait comparer différentes politiques de maintenance (curative et préventive, et de renouvellement (son échéance étant modulable).

N.B. – Sur un tronçon isochrone tel que défini au paragraphe I, les facteurs de constitution et de mission sont constants. Les maillons ne diffèrent objectivement que par les modalités des facteurs d'environnement, dont l'influence sur le paramètre d'échelle est moindre que celle des autres facteurs. Il arrive qu'on dispose de fichiers d'avaries très lourds où sont consignés de nombreux facteurs "d'environnement", dont le renseignement, d'utilité douteuse, alourdit les opérations de maintenance. Un objectif subsidiaire est de préciser l'information strictement utile.

IV – PROBLEMES DE JUGEMENT SUR ECHANTILLON RELATIFS AU PARAMETRE D'ECHELLE

Sur un tronçon isochrone, considérons le sous-ensemble de maillons présentant les mêmes modalités pour les facteurs d'environnement, donc pour tous les facteurs (III N.B.).

A l'instant d'observation, trois données essentielles caractérisent ce sous-ensemble :

n = nombre de maillons posés à l'origine (au renouvellement)

t = flot écoulé depuis l'origine

k = nombre de maillons déposés pour avaries individuelles

Nous avons affaire à un "essai tronqué avec remplacement". Mais du fait que k/n est petit ($< 0,10$) il est aisé de montrer que pratiquement :

- l'occurence de deux avaries successives sur le même maillon est impossible ;
- k suit une loi de POISSON de paramètre $n\lambda t^\beta$,
- le triplet n t k résume toute l'information relative à λ.

C'est-à-dire que les instants d'avarie n'interviennent pratiquement pas, comme s'il s'agissait d'une loi exponentielle. Plus précisément, l'estimation du maximum de vraisemblance est pratiquement dans ce cas :

$$\hat{\lambda} = \frac{k}{n t^\beta} \qquad \text{avec}$$

$$E(\hat{\lambda}) = \lambda \quad \text{(sans biais)} \qquad \text{Var } \hat{\lambda} = \frac{\lambda}{n t^\beta}$$

On peut développer sur cet estimateur les méthodes classiques du jugement sur échantillon, dans le but par exemple de comparer deux types de maillons (deux modalités du facteur a) ou de tester l'apport d'un nouveau constituant, ou de donner un intervalle de confiance aux opérations de maintenance à prévoir dans la période couvrant la tranche de flot $[t , t + \Delta t]$. Une généralisation des abaques de M. CAVE permet en particulier (voir annexe) de résoudre le problème de dimension relatif à ce plan de contrôle, qu'on détermine ici par le produit $n t^\beta$.

S'il s'avère que le tronçon isochrone n'est pas assez important, compte tenu de l'obligation de dépose (troncature) dès que le taux d'avarie crée des problèmes insolubles pour la maintenance, on peut augmenter la quantité d'information en réunissant plusieurs sous-ensembles ℓ de maillons présentant la même modalité, mais répartis sur plusieurs tronçons isochrones dont les dates de pose, donc les flots à la troncature, peuvent différer.

On montre encore dans ce cas que l'estimation linéaire sans biais de moindre variance est :

$$\lambda = \frac{\sum_{\ell} k_{\ell}}{\sum_{\ell} n_{\ell} t_{\ell}^{\beta}}$$

avec

$$E(\hat{\lambda}) = \lambda \qquad \operatorname{Var} \hat{\lambda} = \frac{\lambda}{\sum_{\ell} n_{\ell} t_{\ell}^{\beta}}$$

Cet estimateur permet de faire porter les comparaisons précédentes sur des ensembles de maillons plus importants. Si $\sum_{\ell} k_{\ell}$ est suffisant ($\hat{>}$ 30 par ex.) pour chacun des deux types de maillons I et II, le critère $\frac{\hat{\lambda}_I - \hat{\lambda}_{II}}{\sqrt{\operatorname{Var} \hat{\lambda}_I + \operatorname{Var} \hat{\lambda}_{II}}}$ suit pratiquement une loi normale réduite dans l'hypothèse nulle ou les deux paramètres sont égaux.

On estimera les variances par :

$$\operatorname{Var} \lambda = \frac{\lambda}{\sum_{\ell} n_{\ell} t_{\ell}^{\beta}} \# \frac{\sum_{\ell} k_{\ell}}{\left(\sum_{\ell} n_{\ell} t_{\ell}^{\beta}\right)^2}$$

Ce critère permet de séparer les facteurs les plus importants quant au paramètre d'échelle, par exemple à l'aide d'une procédure de segmentation portant sur un sous-ensemble très important de maillons.

V – MODELISATION DU PARAMETRE D'ECHELLE

Ayant sélectionné par une étude préalable un nombre limité de facteurs qui semblent importants quant au paramètre d'échelle, il peut être intéressant de rechercher un modèle explicatif de ce paramètre si l'on dispose d'un échantillon suffisant. La logique et l'exigence de simplicité s'accordent souvent sur un modèle multiplicatif de la forme :

$$\lambda_r^* = K \prod_{ij} (a_{ij})^{x_{ijr}}$$

r = indice d'un lot homogène de maillon

a_{ij} = effet de la modalité j du facteur A_i.

x_{ijr} = variable indicatrice de cette modalité sur le lot d'indice r.

Ce modèle se linéarise en :

$$\operatorname{Log} \lambda_r^* = \operatorname{Log} K + \sum_{ij} x_{ijr} \operatorname{Log} a_{ij}$$

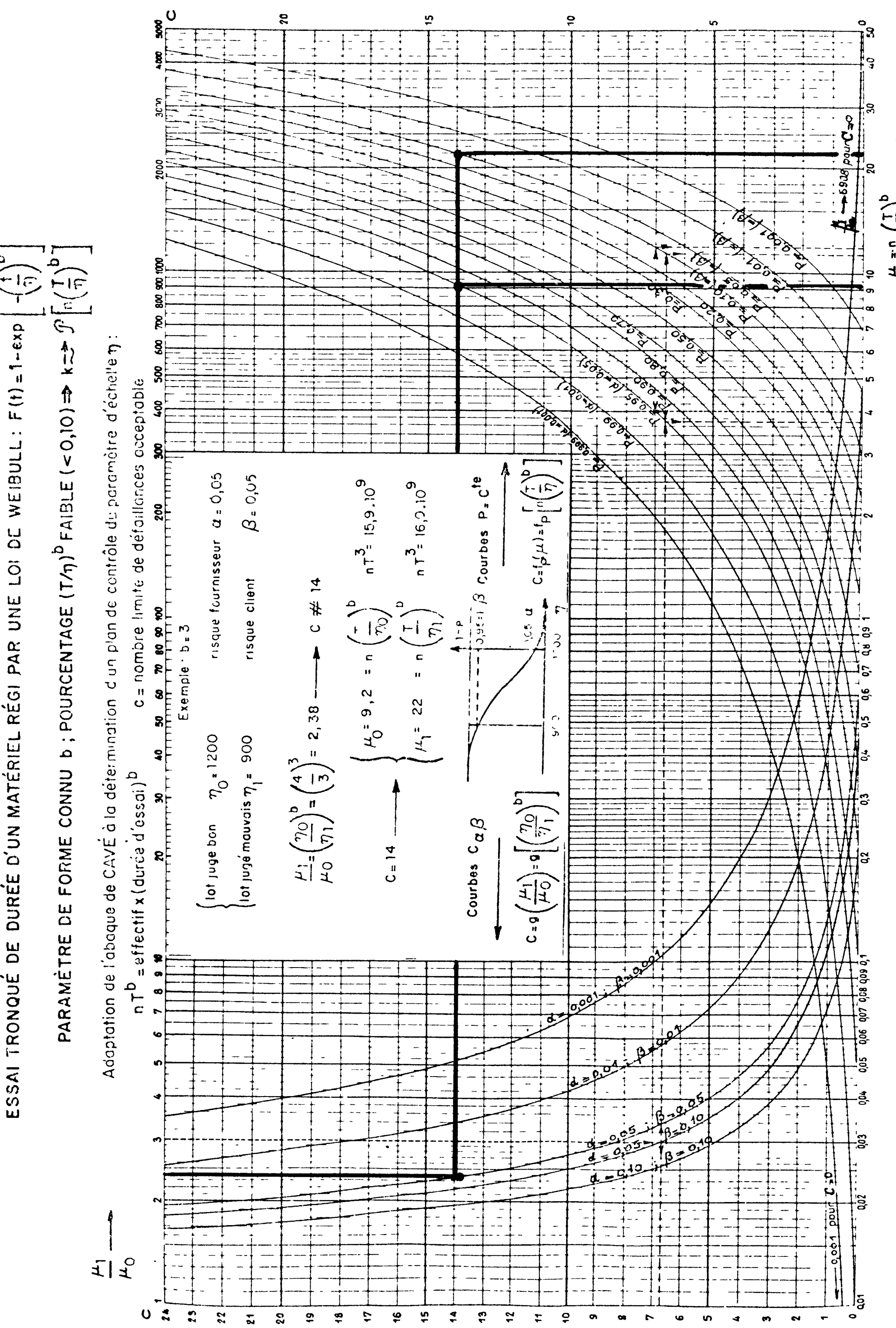

ESSAI TRONQUÉ DE DURÉE D'UN MATÉRIEL RÉGI PAR UNE LOI DE WEIBULL : F(t) = 1 - exp [-(t/η)^b]
PARAMÈTRE DE FORME CONNU b ; POURCENTAGE (T/η)^b FAIBLE (< 0,10) ⇒ k ⇄ f [n(T/η)^b]
Adaptation de l'abaque de CAVÉ à la détermination d'un plan de contrôle du paramètre d'échelle η :
n T^b = effectif x (durée d'essai)^b
c = nombre limite de défaillances acceptable
Exemple : b = 3
lot jugé bon η0 = 1200
risque fournisseur α = 0,05
lot jugé mauvais η1 = 900
risque client β = 0,05
μ1/μ0 = (η0/η1)^b = (4/3)^3 = 2,38 → C # 14
C = 14
μ0 = 9,2 = n(T/η0)^b
n T^3 = 15,9.10^9
μ1 = 22 = n(T/η1)^b
n T^3 = 16,0.10^9
Courbes Cαβ
C = g(μ1/μ0) = g[(η0/η1)^b]
Courbes P = C^te
C = f(μ) = fP[n(T/η)^b]
μ1/μ0
μ = n(T/η)^b
C

Figure 2

Les effets a_{ij} s'obtiendront par une régression qui revient, dans l'espace des variables, à projeter le vecteur correspondant à la variable à expliquer sur l'hyperplan qu'engendrent les vecteurs correspondant aux variables explicatives (les indicatrices X_{ijr}).

Il convient ici de noter qu'on ne dispose pas de valeurs "observées" pour λ_r mais seulement d'estimation $\hat{\lambda}_r$ calculées comme indiqué au paragraphe IV (c'est k_r qui est observé), après réunion en un même "lot homogène" de tous les maillons présentant une même combinaison r de modalités pour l'ensemble des facteurs explicatifs. Toutes ces estimations sont sans biais, globalement indépendantes, mais de variances inégales.

Pour obtenir des estimations $\hat{a}_{ij}$ qui jouissent d'une certaine optimalité (centrées, variance minimale), le théorème de GAUSS-MARKOV nous conduit à pondérer chaque valeur Log λ_r par un coefficient inversement proportionnel à sa variance σ_r^2. Ceci revient à définir l'orthogonalité, dans la projection envisagée, par la matrice diagonale $[\theta_{rr'}/\sigma_r^2]$, inverse de la matrice des covariances.

Le paragraphe IV donne :

$$\text{Var}\,\hat{\lambda}_r = \frac{\lambda_r}{\sum_{\ell} n_\ell t_\ell^\beta}$$

d'où l'on tire :

$$\sigma_r^2 = \text{Var}\,(\text{Log}\,\lambda_r) = \frac{1}{\lambda_r^2}\,\frac{\lambda_r}{\sum_{\ell} n_\ell t_\ell^\beta} = \frac{1}{\lambda_r \sum_{\ell} n_\ell t_\ell^\beta}$$

que l'on estimera par

$$\sigma_r^2 = \frac{1}{\lambda_r \sum_{\ell} n_\ell t_\ell^\beta} = \frac{1}{kr}$$

La matrice diagonale de pondération, définissant l'orthogonalité, sera donc : $[\theta_{rr'}\, k_r]$.

La réunion en un même lot de tous les maillons présentant une même combinaison r de modalités, quel que soit le tronçon d'origine, conduit à l'estimation unique λ_r pour cette combinaison. Ainsi la variance de cette estimation est-elle réduite, et le risque minimisé de trouver zéro pour le nombre d'avaries donc pour λ_r ; enfin la distribution des résidus est normalisée ce qui facilite l'interprétation classique des résultats par le test de FISHER.

Remarque sur le système de pondération.

Il valorise l'information portée par les lots de faible importance mais correspondant à la conjonction de circonstances très défavorables. Il élimine par contre des lots de taille moyenne dont le faible taux d'avarie... et le hasard conduisent à un nombre nul d'avarie. Pour nous prémunir contre cet effet, on peut remplacer le coefficient de pondération k_r par $\sum_{\ell} n_\ell t_\ell^\beta$ qui demeure, quelque soit le paramètre d'échelle, proportionnel à l'espérance de k_r. Il faut toutefois dans ce cas, à cause de la transformation logarithmique, fixer un seuil pour le total $\sum_{\ell} n_{\ell} t_{\ell}^{\beta}$ de sorte que soient éliminés tous les lots trop petits, dont en particulier ceux pour lesquels $k_r = 0$.

Les résultats des deux ajustements sont plus complémentaires que contradictoires : ils apportent sur le rôle des différents facteurs des éclairages différents.

Un autre intérêt de la pondération par nt^{β} apparaîtra en VIII.

VI – GENERALISATION DE LA METHODE AU CAS OU LE PARAMETRE DE FORME β EST INCONNU

Si tous les maillons d'un même lot homogène r présentent la même valeur t_r pour le flot cumulé depuis la pose, le modèle précédent s'écrit, après linéarisation

$$\text{Log } k_r^* = \text{Log } K + \text{Log } n_r + \beta \text{ Log } t_r + \sum_{ij} x_{ijr} \text{ Log } a_{ij}$$

Sous cette forme on voit que β peut être introduit comme un paramètre inconnu que l'on estimera par la régression, pourvu qu'il soit admis que sa valeur est indépendante des modalités prises sur les facteurs explicatifs(1). Cette hypothèse est le plus souvent très réaliste(2). La nouvelle version du modèle dispense alors de l'investigation préalable nécessaire à connaître la valeur de β. On peut même dire, compte tenu du caractère très général de la loi de WEIBULL, qu'il suffit de savoir que le taux instantané d'avarie croit rapidement avec le flot écoulé. Le système de pondération reste le même, c'est-à-dire que la matrice définissant l'orthogonalité est toujours : $[\theta_{rr'}, k_r]$.

La seule contrainte est d'avoir à travailler avec un plus grand nombre d'observations, chacune correspondant à un "sous-lot" de même valeur pour le flot t, issu en général d'un même tronçon isochrone de pose, et conduisant à des variances plus grandes donc à un risque accru de nullité pour l'estimation $\hat{\lambda}$.

VII – QUELQUES DIFFICULTES RELATIVES A LA DEFINITION DES MODALITES POUR CERTAINS FACTEURS

Il y a contradiction entre :

– finesse de la partition, nécessaire pour éclairer le rôle des facteurs :

– importance des lots homogènes, souhaitable pour diminuer la variance de $\hat{\lambda}_r$ et le risque de nullité, et normaliser la distribution des résidus.

---

(1) Qualité des estimateurs. S'il est vrai que $\hat{\lambda} = k/nt^{\beta}$ soit toujours l'estimateur du maximum de vraisemblance, on peut s'interroger sur l'estimateur $\hat{\beta}$ fourni par la minimisation de la somme des carrés des résidus. Du moins, dans la condition où l'approximation normale est valable pour la distribution de Log k (et alors le biais introduit par la transformation logarithmique est minime), peut-on admettre l'équivalence de la méthode des moindres carrés (pondérés) avec celle du maximum de vraisemblance.

(2) Si η dépend étroitement de l'agressivité du flot et de l'environnement et de la qualité des maillons, il semble bien que β soit une constante spécifique de la mission du système.

En particulier on perdra de l'information chaque fois qu'un facteur se présente normalement comme une variable continue. Dans ce cas, il s'agira moins d'estimer l'effet d'une modalité que de tester le caractère significatif de l'effet. Encore faut-il disposer d'une bonne partition.

Le paragraphe III évoque la difficulté propre aux facteurs "de mission". Dans le cas, par exemple, de courants de trafic techniquement hétérogènes, il est intéressant de mettre en lumière les différences d'agressivité entre ces divers courants. Leurs effets cumulatifs sur l'émergence des avaries rendent cette séparation malaisée. Un procédé consiste à répartir les tronçons isochrones en classes présentant des similarités quant à la nature des flots écoulés depuis la pose jusqu'à la troncature. Une analyse factorielle des correspondances entre tronçons isochrones et nature de trafic parait indiquée pour engendrer des classes de similarité. Si le test de FISHER révèle alors comme significatif le facteur "nature de trafic", il reste naturellement à découvrir les courants responsables des différences entre les modalités ainsi constituées.

Enfin on a vu en VI que l'estimation par régression du paramètre de forme β exige une partition en "sous-lots", correspondant à des classes homogènes pour le flot cumulté t, qui augmente la dispersion des estimations $\hat{\lambda}$ et le risque de nullité. Pour s'affranchir de la contrainte issue des observations $k = 0$, HAROLD J. BALABAN et KENT HASPERT proposent une méthode (voir note bibliographique). Elle consiste à faire une première régression ne tenant pas compte des sous-lots pour lesquels $k = 0$. Puis à calculer l'espérance du flot t relatif à la première défaillance sur les tronçons où $k = 0$ en prenant pour λ la valeur issue de cette régression. Enfin à réitérer la régression en prenant en compte tous les lots.

VIII – PLAN D'OBSERVATION VISANT UNE UTILISATION OPTIMALE DE L'INFORMATION

De telles études s'appuient le plus souvent sur la disponibilité de fichiers très lourds, constitués sans but bien défini : ça peut toujours servir à quelque chose... et puis l'ordinateur est là... Le personnel chargé de la maintenance est d'autant moins motivé pour la collecte de renseignements minutieux qu'il n'en voit jamais les retombées sinon comme sanction d'une franchise trop zélée.

Le raffinement du traitement multidimentionnel mis en œuvre est souvent sans commune mesure avec la qualité de tels fichiers, dont le taux d'erreur croît avec le volume et la complexité. Ces données "tout-venant" débouchent alors sur des résultats médiocres, d'autant plus ambigus que les facteurs proposés à l'explication sont naturellement plus corrélés entre eux. Une analyse "pas à pas", solution chère, sert souvent d'alibi au refus de la mise en cause de ces fichiers. Il semblerait du moins préférable de rechercher au préalable les composantes principales de ces facteurs ; encore faudrait-il qu'elles soient facilement interprétables, ce qui n'est pas souvent le cas.

Mais la remarque sur le système de pondération (paragraphe V) nous suggère *un plan de collecte des données* dont l'analyse échappe pour l'essentiel aux inconvénients signalés. Le produit $n_r t_r^{\beta}$, proportionnel à l'espérance du nombre k_r d'avaries, mesure en quelque sorte, avant toute observation, la quantité d'information apportée par le lot r. Après élimination de facteurs manifestement neutres, sur l'analyse grossière du fichier disponible, on constituera un plan d'échantillonnage. Il inclura des tronçons donnant, pour chacune des combinaisons r de modalités, un produit $n_r t_r^\beta$ très voisin d'une même valeur fixé à priori, et compatible avec la précision souhaitée... et le temps d'obtention. Si β est inconnu il suffit de disposer d'une première approximation. On peut aussi, après choix des tronçons isochrones, faire en sorte que l'information soit stockée jusqu'à dépose collective des tronçons, moment où elle est la plus riche compte tenu de la croissance du taux instantané d'avarie. Ainsi pourra-t-on tout à la fois alléger la collecte et aboutir à des résultats facilement interprétables.

IX – NOTE BIBLIOGRAPHIQUE SUCCINCTE

Cette approche a été mise en oeuvre sur le fichier "rails" de la S.N.C.F. en 1974. Par ailleurs, HAROLD S. BALABAN et KENT HASPERT [1] développent, dans le cas d'un modèle de WEIBULL où seul le paramètre d'échelle est inconnu, une approche plus générale en ce sens qu'elle n'exige pas de limite pour le taux de dépose k/n. S'appuyant exlcusivement sur une statistique de rangs, elle perd cependant une partie de l'information (1). Deux cas sont envisagés : avec et sans remplacement des défaillants.

J.C. LIGERON et R. GOARIN [2] exposent une factorisation du taux de défaillance, dans le cas $\beta = 1$ (loi de survie exponentielle) par analyse des correspondances multiples ; puis ils modélisent sur les facteurs retenus par différentes méthodes adaptées à la forme du modèle ; enfin ils proposent une procédure bayesienne d'amélioration progressive de leurs estimations.

REFERENCES

[1] HAROLD S. BALABAN et KENT HASPERT – Estimating WEIBULL parameters for a general class of devices from limited failures data. *I.E.E. Transactions on Reliability,* Vol. 21, n° 2, Mai 1972.

[2] LIGERON J.C. et GOARIN R. – Modélisation multidimensionnelle des taux de défaillance des composants électroniques. *R.S.A.*, Vol. XXII, n° 3, 1974.

(1) La perte d'information liée au remplacement des valeurs t_ℓ par leurs rangs peut être notable, du moins si k/n est grand.