

# STATISTIQUE ET ANALYSE DES DONNÉES

J. P. VILA

## Un logiciel d'ajustement de modèles non linéaires

*Statistique et analyse des données*, tome 2, n° 1 (1977), p. 102-118

[http://www.numdam.org/item?id=SAD\\_1977\\_\\_2\\_1\\_102\\_0](http://www.numdam.org/item?id=SAD_1977__2_1_102_0)

© Association pour la statistique et ses utilisations, 1977, tous droits réservés.

L'accès aux archives de la revue « Statistique et analyse des données » implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme  
Numérisation de documents anciens mathématiques  
<http://www.numdam.org/>

## UN LOGICIEL D'AJUSTEMENT DE MODELES NON LINEAIRES

-----

M. L. P.

Maximum Likelihood Program

J.P. VILA  
Laboratoire de Biométrie du C.N.R.Z.  
78350 JOUY EN JOSAS

Résumé :

*Le logiciel M.L.P. offre à l'utilisateur un langage simple lui permettant de commander l'ajustement à ses données du modèle de son choix. Celui-ci pourra être éventuellement défini s'il ne fait pas déjà partie des modèles standardisés reconnus par le programme de traitement : la plupart des modèles non linéaires et de distributions de probabilités classiques. L'ajustement du modèle aux données s'effectue de façon transparente à l'utilisateur, par la méthode des moindres carrés ou la méthode du maximum de vraisemblance. La convergence des algorithmes d'optimisation utilisés est favorisée par la prise en compte de l'aspect statistique du problème.*

Mots clés.

Maximum de vraisemblance  
Modèle  
Modèle non linéaire  
Moindres carrés  
Optimisation  
Régression non linéaire

## I- INTRODUCTION -

Le logiciel M.L.P. a été conçu par G.J.S. ROSS (1970). Il permet l'ajustement de modèles non linéaires et de distributions de probabilités, par la méthode des moindres carrés ou la méthode du maximum de vraisemblance. Chaque ajustement est mené en tenant compte :

- 1 - des propriétés analytiques et statistiques du modèle
- 2 - de la structure des données
- 3 - des relations entre modèle et données.

Ceci, afin de simplifier le problème et conduire à une solution admissible directement, ou indirectement après transformation du modèle ou des données. Cette démarche, systématisée pour tous les ajustements possibles a été intégrée dans M.L.P. de façon transparente à l'utilisateur, de manière à ce que :

- 1 - l'utilisateur n'ait qu'à fournir les données et spécifier son modèle il n'a pas à fournir de valeurs initiales aux paramètres. Il n'a pas non plus à connaître le principe de l'ajustement.
- 2 - la convergence soit rapide et les résultats aussi précis que possible.
- 3 - toute anomalie dans les données ou incompatibilité avec le modèle soit signalée immédiatement.
- 4 - les cas vraiment insolubles soient détectés aussitôt que possible et ceux qui peuvent être "repris" après transformation, le soient effectivement.
- 5 - en cas d'existence de solution multiple, la solution souhaitable puisse être atteinte.
- 6 - toute analyse statistique appropriée et utile soit fournie, pour juger de la validité des résultats.

## II- CARACTERISTIQUES GENERALES ET FACILITES OFFERTES .-

II-1. Le logiciel M.L.P. offre un langage simple à l'utilisateur qui a le choix entre différents modèles "standardisés", mais qui peut aussi définir rapidement le sien :

II-1.1. Les modèles standardisés disponibles sont une cinquantaine répartis en classes dont les plus importantes concernent les courbes, les distributions de probabilités discrètes et continues, les réponses par tout ou rien (analyse des probits):

Chaque modèle a été considéré de manière particulière de façon à éliminer toutes ses difficultés spécifiques lors d'un ajustement. Les principes des traitements effectués sont quelquefois généraux (réduction des dimensions, transformation de paramètres, politique d'ajustements 'emboîtés'), quelquefois spécifiques aux modèles. Mais tout traitement est la résultante d'un raisonnement analytique et statistique a priori, et d'une expérience pratique des auteurs. C'est ce qui explique son efficacité dans la plupart des cas.

II.1.1.1. Courbes :  $y = f(x, A, B, C, \dots) + e$

les erreurs  $e$  peuvent être normales, log-normales. Elles peuvent être pondérées

a - Courbes exponentielles

b - Courbes de "croissance"

. Courbes logistiques

. Courbes de Gompertz

c - Modèles à compartiments

d - Autres : polynômes orthogonaux, fractions rationnelles....

II.1.1.2. Distributions de probabilité.

Il est possible d'ajuster à des distributions de fréquences observées.

a - des distributions discrètes :

de Poisson, de Neyman type A, géométrique, logarithmique, binomiale négative.

b - des distributions continues :

normale, combinaison de lois normales, log-normale, exponentielle, gamma, chi-deux, béta, de Fisher, de Weibull.

II.1.1.3. Régression linéaire multiple.

II.1.1.4. Analyse de probit.

Un ajustement linéaire de la transformée probit ou logit d'une variable pourcentage peut-être effectué en fonction d'une ou deux variables indépendantes dans différents types de problèmes (Estimation des coefficients de la droite de régression des probits, analyses de mélanges, problème de Wadley, etc...)

II.1.1.5. Autres types d'ajustements possibles.

Ils sont relatifs à des modèles plus spécialisés.

II.1.1.5.1. Modèles de génétique : estimation de linkage

II.1.1.5.2. Modèle de dilution : dénombrement biologique

II.1.1.5.3. Modèle de transition entre états dans un système fermé ou ouvert : ajustement de paramètres déterminant les éléments d'une matrice de transition stationnaire ou non.

#### II.1.2. Ajustements de modèles spécifiques :

S'il n'est pas satisfait par les modèles standardisés reconnus par le programme de traitement, l'utilisateur peut définir son propre modèle, au moyen :

- 1 - de combinaisons linéaires de variables intermédiaires de travail
- 2 - d'options contrôlant le calcul de la vraisemblance compte tenu du modèle et de la distribution supposée de sa partie aléatoire :
  - a. normale ou log-normale pour courbes et surfaces
  - b. poissonnienne pour des dénombrements
  - c. binomiale pour des variables pourcentages
  - d. multinomiale pour des distributions
  - e. au choix de l'utilisateur qui programmera directement sa fonction de vraisemblance.

Les principes statistiques et numériques qui assurent l'efficacité des ajustements des modèles standardisés sont également appliqués dans les ajustements de ces modèles spécifiques.

### III- L'UTILISATION EFFICACE DES METHODES DE MINIMISATION DANS M.L.P. .-

La méthode de minimisation utilisée est une méthode à convergence quadratique de type Newton modifié (Box (1969)). Elle est utilisée pour minimiser par rapport aux paramètres

- . une somme de carrés d'écarts pondérés, ou
- . l'opposé du logarithme de la fonction de vraisemblance calculée sur les données en fonction du modèle théorique et de la loi supposée de sa partie aléatoire.

Mais certains modes d'utilisation de cette procédure dans M.L.P., associés à certaines opérations sur le modèle, favorisent cette minimisation d'une manière décisive :

#### III-1. Minimisation hiérarchique : réduction des dimensions des espaces paramétriques de recherche.

Dans certains types de modèles, M.L.P. traite les paramètres selon une hiérarchie, telle que des valeurs optimales pour les paramètres de rangs inférieurs puissent être obtenues aisément, pour toutes valeurs des paramètres de rangs supérieurs.

Il est quelquefois possible d'obtenir pour les paramètres de rangs inférieurs à partir des équations normales, des solutions analytiques fonctions des paramètres de plus hauts rangs. Les premiers peuvent alors être éliminés dans la fonction critère: La minimisation s'effectuera dans un espace de dimension plus faible. De plus les paramètres éliminés ne nécessitent plus de valeurs initiales. Ainsi les paramètres linéaires seront-ils éliminés en priorité. Ces idées avaient déjà été exploitées par Richards (1961) et depuis par Lawton et Sylvestre (1971).

Quand une élimination analytique de paramètres n'est plus possible, la procédure de minimisation est utilisée de manière récursive dans une hiérarchie de sous espaces paramétriques complémentaires.

III-2. Transformation des paramètres afin de réduire les variances et corrélations de leurs estimateurs et d'obtenir une approximation quadratique suffisante de la fonction critère dans un voisinage du minimum : création de paramètres "stables" selon ROSS (1970).

Dans le cas des modèles non linéaires, il n'existe pas de transformation linéaire orthogonale assurant les objectifs désirés dans un domaine assez vaste pour englober toutes les estimations initiales plausibles des paramètres. Il faut donc essayer des transformations non-linéaires et si possible biunivoques pour pouvoir revenir facilement aux paramètres initiaux du modèle. Le choix de la transformation sera d'autant plus aisé que l'on pourra trouver des paramètres variant très peu dans la région de "bon" ajustement de l'espace paramétrique : d'où leur qualification de stable. De plus étant fonction des valeurs ajustées, ils pourront souvent être initialisés automatiquement par ces mêmes fonctions.

La première démarche opératoire consiste à examiner les espérances de certaines statistiques qui, fonction des paramètres du modèle, peuvent elles-mêmes être des paramètres potentiels. Ainsi, dans les ajustements de distributions, les moments d'ordres inférieurs sont assez stables (sauf si la queue de distribution est exceptionnellement longue). Dans le cas d'ajustement de courbe, un modèle à  $p$  paramètres, peut-être défini par les espérances en  $p$  points différents. Quand ces points sont régulièrement répartis sur l'intervalle des données, les intercorrélations sont généralement faibles, et si la courbe est un polynôme de degré  $p-1$ , il existe même des ensembles de  $p$  points pour lesquels, toutes les intercorrélations entre valeurs ajustées sont nulles. Bien entendu, cette recherche de paramètres stables est combinée quand c'est possible avec la démarche vue précédemment de réduction des dimensions de l'espace paramétrique : ainsi on recherchera de nouveaux paramètres

stables caractéristiques seulement de la forme de la courbe, et non de son échelle et de sa position caractérisées par les paramètres linéaires qui auront été éliminés. Pour cela, il est possible d'utiliser les différences premières ou secondes des ordonnées, ou leurs rapports, ou d'autres transformations (ROSS (1970)) suivant les modèles.

### III-3. Minimisation séquentielle : ajustements emboîtés.

Le modèle décrit ou appelé par l'utilisateur est quelquefois remplacé dans le corps des programmes du système par une suite de modèles emboîtés plus simples. Leurs ajustements successifs amènent à de bonnes estimations initiales pour le premier modèle. Par exemple, un paramètre "stable" déterminé pourra être fixé à une constante convenable estimée. Et un ajustement préliminaire dans un espace de dimension réduite sera effectué. Cette méthode a été en particulier appliquée dans les modules de M.L.P. relatifs aux ajustements de distributions, car les premiers moments sont très "stables", et ont été pour cette raison utilisés.

ROSS (1970) décrit une application à l'ajustement d'une distribution combinaison linéaire inconnue de deux distributions normales ;

le modèle à ajuster

$$f(x) = \alpha N(\mu_1, \sigma^2) + (1 - \alpha) N(\mu_2, \sigma^2)$$

est remplacé par les trois modèles successifs

Modèle 1 :  $\mu_1 = \mu_2$

Modèle 2 :  $\alpha = 0.5$  ,  $\mu_1 \neq \mu_2$

Modèle 3 :  $\alpha \neq 0.5$  ,  $\mu_1 \neq \mu_2$

A chaque étape, des valeurs initiales pour les paramètres sont obtenues en fixant la moyenne et la variance de la distribution combinée à des valeurs dérivées de l'étape précédente. Il est très facile de plus, d'obtenir un test de validité d'ajustement pour chaque modèle.

## IV- TECHNIQUE OPERATOIRE .-

IV-1. D'une manière très concise, l'utilisateur

1. introduit ses données
2. spécifie numériquement ou mnémoniquement par une seule directive le modèle à ajuster et des contraintes éventuelles
3. commande l'ajustement par une seconde directive.

Toutes ces opérations peuvent être exécutées en mode différé ou en mode conversationnel.

Exemple : ajustement à une courbe de croissance logistique

```

      X                               Y
    _____                     _____
DATA  1 2 3 4 5 6 7  3.5 5.8 8.2 10.1 11.3 11.7 13.3 ;
CMODEL = LOGISTIC; FIT CURVE

```

IV.2. Dans tous les cas, les sorties minimum sont constituées

- 1 - des estimations des paramètres du modèle
- 2 - des erreurs-types et des corrélations des estimateurs de ces paramètres
- 3 - des valeurs estimées de la variable à expliquer
- 4 - des résidus
- 5 - de la valeur atteinte par la fonction critère.

Ce résultat peut conduire à un test de validité du modèle.

Ses sorties peuvent être complétées par diverses informations statistiques, numériques, et des graphiques.

Pour faciliter à l'utilisateur sa recherche de paramètres les plus "stables" quand il ajuste son propre modèle, il est possible de définir des fonctions de paramètres qui seront calculées avec leurs variances et leur corrélations. L'utilisateur peut également faire calculer en un réseau de points de son choix dans l'espace paramétrique, les valeurs de la fonction critère. Il peut obtenir les représentations graphiques des courbes de niveaux correspondantes, dans les différents plans paramétriques.

IV-3. Des diagnostics d'erreurs ont été prévus pour aider aux corrections de la syntaxe et à l'amélioration du modèle quand l'ajustement est défectueux

#### V- LE PRODUIT PROGRAMME .-

Le logiciel M.L.P. est un interpréteur écrit en langage fortran, d'un encombrement de 150 KØ environ. Les différents sous programmes sont disposés selon une structure arborescente à deux niveaux.

Le produit est diffusé par le département de statistique de Rothamsted Experimental Station, Herts, Grande-Bretagne. En France M.L.P. a été étudié et généré dans le cadre d'une A.T.P. au laboratoire de Biométrie du C.N.R.Z.-I.N.R.A. (Jouy-en-Josas). Il fonctionne en mode différé sur l'ordinateur IBM 370/I45 de ce centre et sur l'ordinateur IBM 370/I68 du C.I.R.C.E. (Orsay). Une version interactive devrait être générée sous peu dans ce dernier centre.

Des documents techniques et utilitaires sont disponibles au laboratoire de Biométrie du C.N.R.Z.

VI- EXEMPLES .-

VI-1. Ajustement d'une double exponentielle en faisant appel au modèle standardisé :

$$Y = A + B.\exp(-K.x) + C.\exp(-L.x)$$

VI-2. Ajustements de distributions de Poisson, binomiale négative, et de Neyman type A à des mêmes fréquences observées.

VII- BIBLIOGRAPHIE .-

- BOX M.J., DAVIES D. and SWANN W.H. (1969) : Non linear optimisation techniques.  
I.C.I. Monograph n° 5. Edinburgh : Oliver & Boyd.
- LAWTON W.H., and SYLVESTRE E.A. (1971) : Elimination of linear parameters in non-linear regression. Technometrics, 13, 461-467.
- RICHARDS F.S.G. (1961) A method of maximum likelihood estimation.  
J.R. Statistical Soc. B, 23, 469-476.
- ROSS G.J.S. (1970) The efficient use of function minimization in non-linear maximum likelihood estimation.  
J.R. Statistical. Soc., C, 19, 205-221.
- ROSS G.J.S. (1975) : Simple non-linear modelling for the General user.  
Bulletin of the International Statistical Institute.  
Proceedings of the 40<sup>th</sup> session - Volume XLVI. Book 2 -  
Invited papers - p. 585 - 593 WARSAW 1975.

oooOooo

VIII- REMERCIEMENTS .-

=====

Je remercie Monsieur Jean-Pierre PAGES (laboratoire de statistiques et d'études Economiques et Sociales - C.E.A.) pour ses judicieuses suggestions.

oooOooo

ANNEXE :- EXEMPLES .-1. Ajustement d'une double exponentielle :

$$y = A + B \cdot \text{EXP}(-Kx) + C \cdot \text{EXP}(-Lx)$$

CAPTION (5) AJUSTEMENT D'UNE DOUBLE EXPONENTIELLE AVEC IMPRESSION DE LA SURFACE CRITERE - DONNEES DE G.J.S. ROSS ;

DATA 1 2 3 4 5 6 7 8 9 IO5 35 24 20 I7 I4 I5 II 9 ;

OPT	2	2	0	0	4	I	I
	modèle	axe des x asymptote vers la droite	x	moindres carrés ordinaires	décimales en sortie	erreur résiduelle	erreurs types des ajustées
	0	0	0	I	;		
	pas de pente calculée	pas d'ajuste- ment par poly- gônes orthogonaux		graphique			
FIT CURVE;DISPLAY		1	2	100	1000		
		rang de l'ajustement concerné	projection dans le plan des deux para- mètres non linéaires	unité d'impression des valeurs	seuil provo- quant l'impres- sion d'un X		
	0.0	0.0	2.0	0.0	2.0	1.0	;
	point S-0 du reseau		point S-E		point N-W		
DISPLAY	I	2	5	IOO	.80	.OI6	.9I
DISPLAY	I	2	I	20	.80	.OI6	.9I
					.OI6	.80	.2I ;
							END

Remarque

Les deux dernières projections sont des agrandissements de la zone solution entourée sur la lère projection.

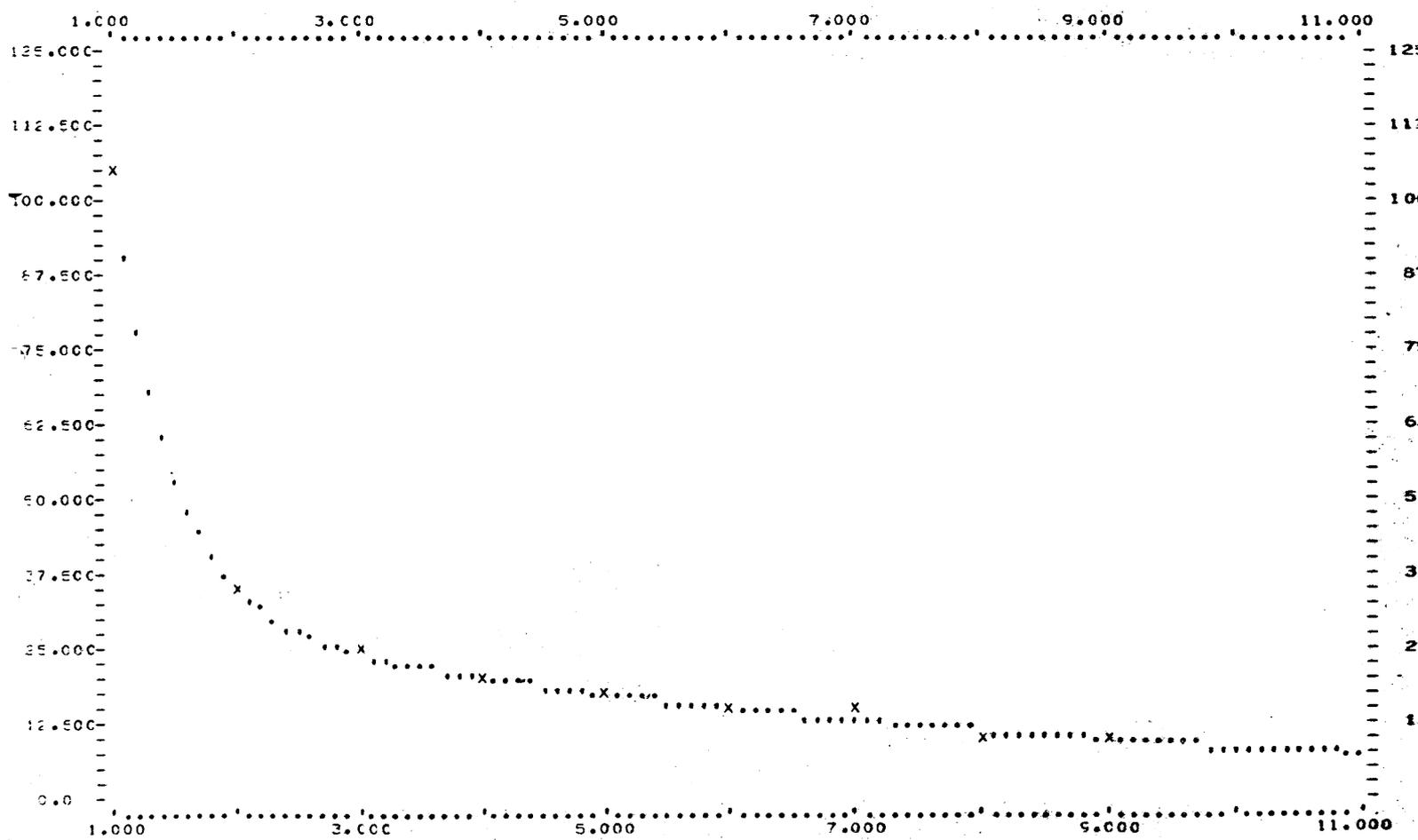
AJUSTEMENT D'UNE DOUBLE EXPONENTIELLE  
 AVEC IMPRESSIONS DE LA SURFACE CRITERE -  
 DONNEES DE G.J.S. RC 35

$Y = A + B*R**X + C*S**X$   $R=EXP(-K)$ ,  $S=EXP(-L)$   
 WITH ASYMPTOTE  $Y = 0.0$

	PARAMETER	S.E.	CORRELATIONS			
1	R	0.86690	0.01391	1.0000		
2	S	0.11495	0.02474	0.7128	1.0000	
3	B	35.20181	3.16776	-0.9528	-0.8033	1.0000
4	C	647.96045	124.95264	-0.6480	-0.9923	0.7385
FIX	A	0.0				1.0000
	K	0.14283	0.01604			
	L	2.16327	0.21523			

	X	Y	F(Y)	WTD.RES.	S.E.
1	1.0000	105.0000	104.9990	0.0010	1.0714
2	2.0000	35.0000	35.0166	-0.0166	1.0599
3	3.0000	24.0000	23.9180	0.0820	0.7331
4	4.0000	20.0000	19.9946	0.0054	0.6353
5	5.0000	17.0000	17.2484	-0.2484	0.4744
6	6.0000	14.0000	14.9429	-0.9429	0.4373
7	7.0000	15.0000	12.9529	2.0471	0.4936
8	8.0000	11.0000	11.2288	-0.2288	0.5683
9	9.0000	9.0000	9.7343	-0.7343	0.6298

RESIDUAL	M.S.	1.147959	D.F.	5
RESIDUAL	S.S.	5.739796		

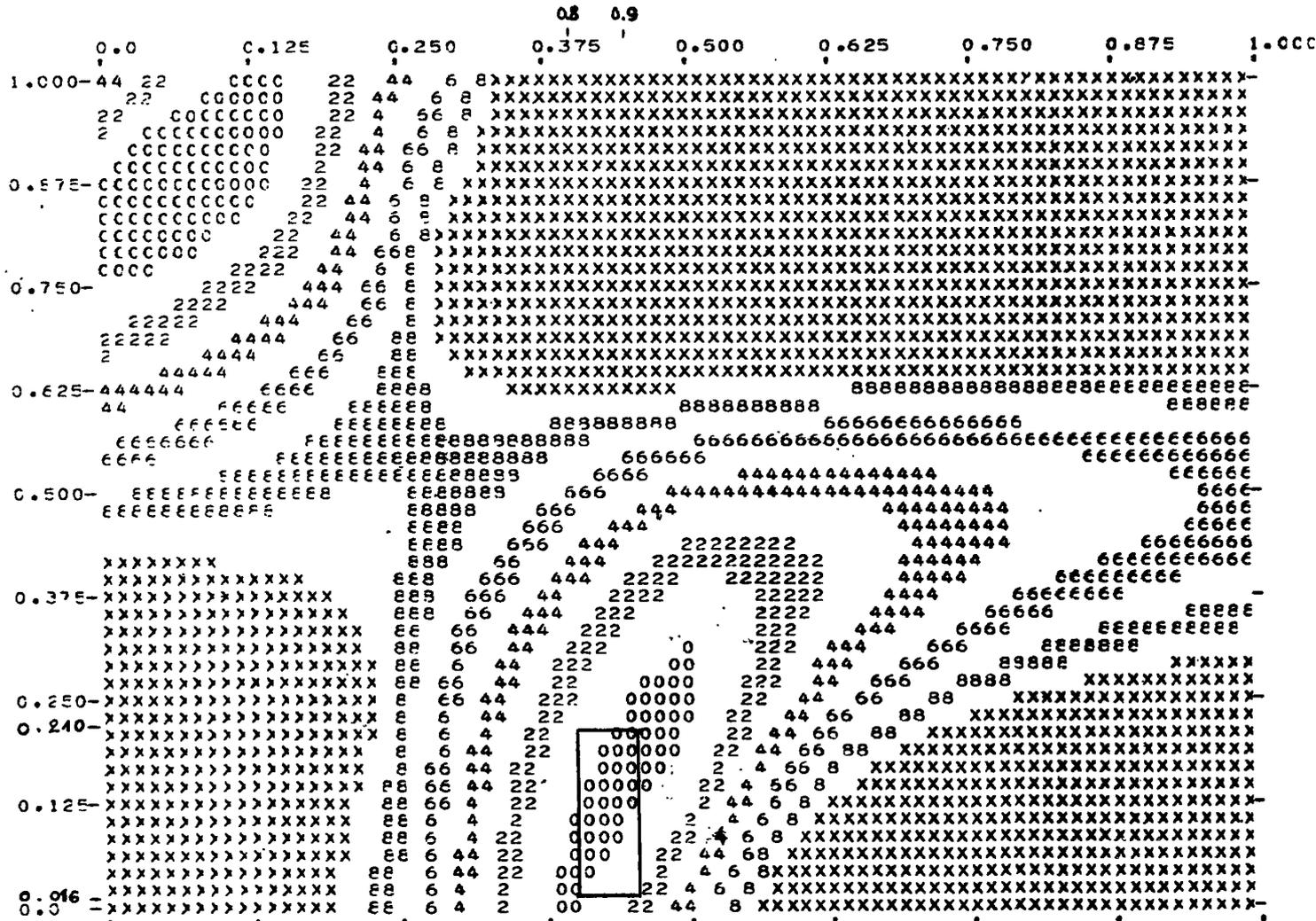


Y = A + B\*R\*\*X + C\*S\*\*X R=EXP(-K), S=EXP(-L)  
WITH ASYMPTOTE Y= C.O

CONTOUR PLOT OF FUNCTION IN UNITS OF 100.0000

PAR 1 SW SE NW NE  
PAR 2 0.0 2.00000 0.0 1.00000 1.00000

Table with 9 columns of numerical values representing function values at various coordinates.

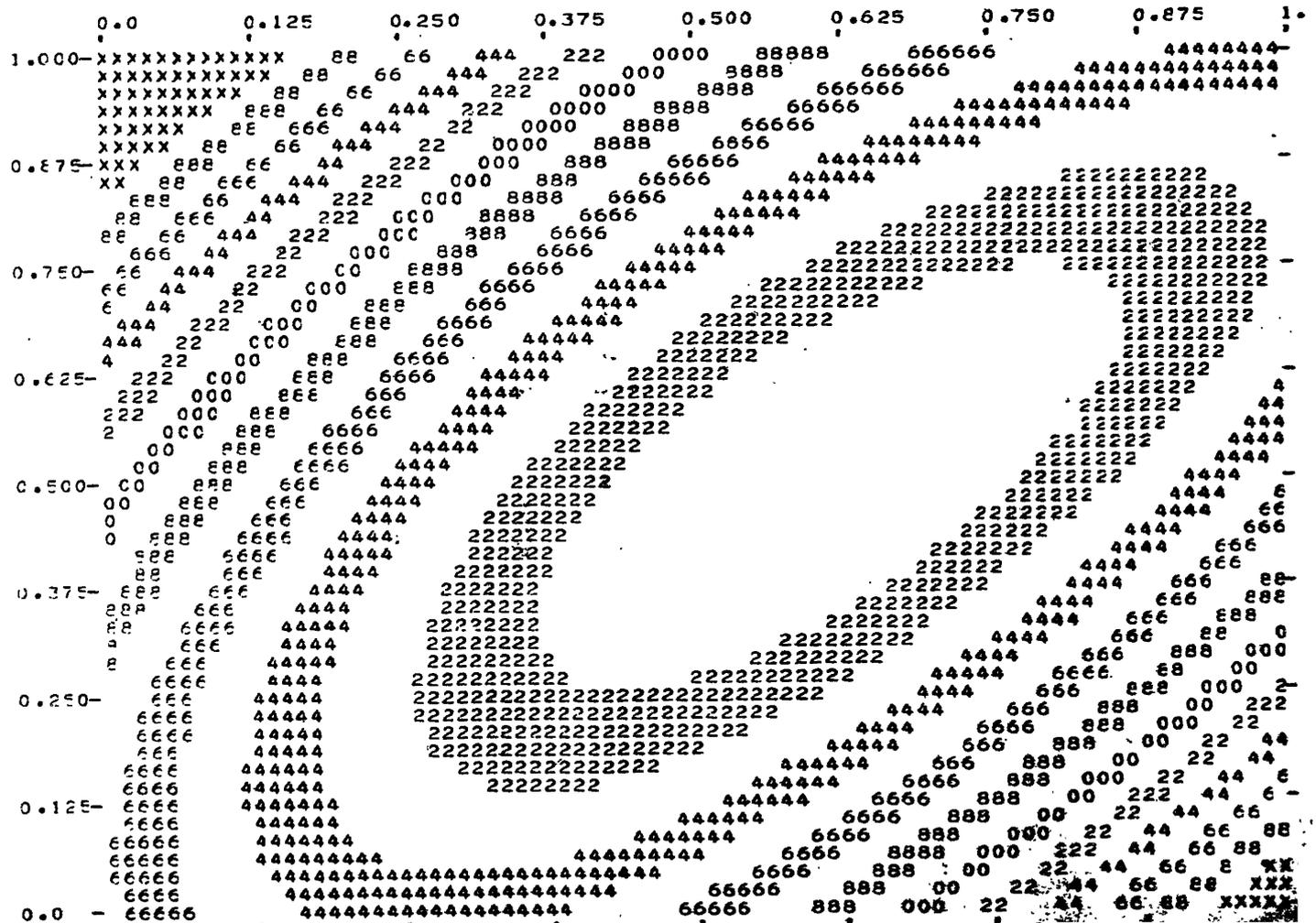


Y = A + B\*R\*\*X + C\*S\*\*X R=EXP(-K) ,S=EXP(-L)  
 WITH ASYMPTOTE Y= 0.0

CONTOUR PLOT OF FUNCTION IN UNITS OF 5.0000

PAR 1	SW 0.80000	SE 0.91000	NW 0.80000	NE 0.91000
PAR 2	0.01600	0.01600	0.21000	0.21000

26.1119	21.3168	17.0479	13.3536	10.2780	7.8587	6.1264	5.1031	4.6022
21.3168	16.7540	12.7715	9.4216	6.7492	4.7925	3.5815	3.1371	3.4700
17.2234	12.9354	9.2865	6.3288	4.1082	2.6631	2.0228	2.2064	3.2224
13.2228	9.8560	6.5861	4.0668	2.3448	1.4585	1.4361	2.2952	4.0412
11.1122	7.5098	4.6630	2.6267	1.4484	1.1662	1.8073	3.3872	5.9081
9.0866	5.8904	3.5091	1.9990	1.4078	1.7735	3.1219	5.4661	8.8052
7.7403	4.5909	3.1163	2.1739	2.2119	3.2676	5.3654	8.6159	12.7144
7.0471	4.8038	3.4757	3.1413	3.8491	5.6356	8.5235	12.5204	17.6181
7.0606	5.3215	4.5766	4.8913	6.3081	8.8646	12.5817	17.4640	23.4952

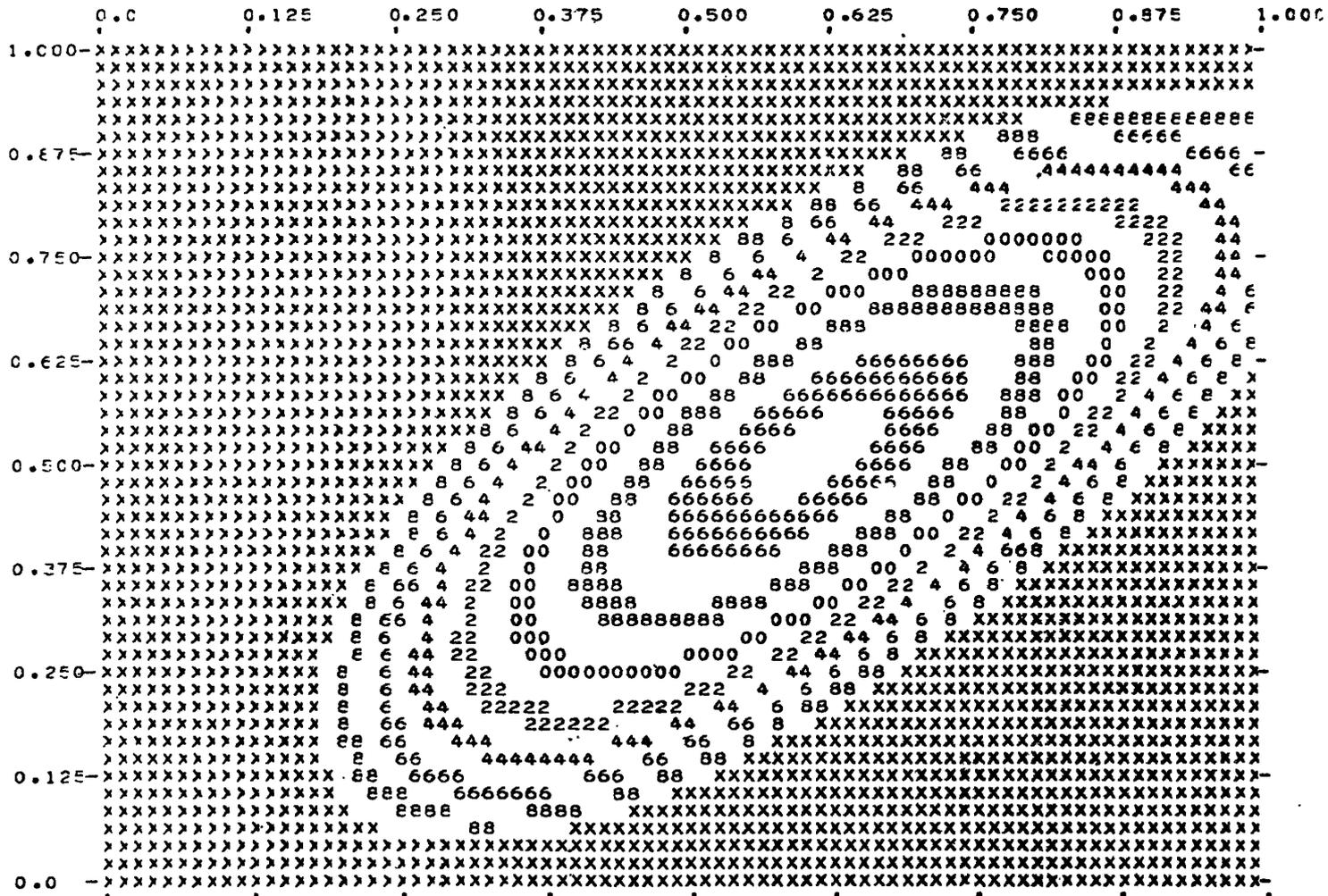


Y = A + P\*R\*\*X + C\*S\*\*X R=EXP(-K) ,S=EXP(-L)  
WITH ASYMPTOTE Y= C.O

CONTOUR PLOT OF FUNCTION IN UNITS OF 1.0000

PAR 1 SW SF NW NE  
PAR 2 C.ECCCC C.G1CCCC 0.80000 0.91000  
PAR 2 0.C160C C.G160C 0.21000 0.21000

Table with 9 columns of numerical values ranging from 130.5594 to 117.4555.



2. Ajustements de distributions discrètes.

CAPTION (4) AJUSTEMENTS DE DISTRIBUTIONS DE POISSON, BINOMIALE NEGATIVE,  
DE NEYMAN TYPE A, A DES MEMES FREQUENCES OBSERVEES ;

DATA 12 19 22 16 10 9 9 4 6 2 / 1 1 2 1

Fréquences observées des valeurs 0,1,2,3,4,5,6,7,8,9      Fréquences observées des valeurs 10,11, 12,13. Elles seront cumulées.

EXTRA 17 26 37 ;

Valeurs de queue de distribution

OPTION	0	1 4 5	0	0	1	;
	les données sont des fréquences	codes des trois distributions à ajuster	il n'y a pas de quatrième distribu- tion	la table des fré- quences observées et calcu- lées sera imprimée	les histo- grammes des fréquences seront imprimés	

FIT DISTRIBUTION ;

AJUSTEMENTS DE DISTRIBUTIONS DE POISSON,  
BINOMIALE NEGATIVE, DE NEYMAN TYPE A,  
A DES MEMES FREQUENCES OBSERVEES

NO OF SAMPLES	117.	SAMPLE MEAN	4.12
SAMPLE VARIANCE	23.32	SAMPLE SKEWNESS	3.75
POISSON INDEX	1.13	N.B. INDEX	4.04

## FIT POISSON

CHI-SQUARED = 93.19 CN 10 D.F.

	PARAMETER	S.E.
1	M	4.11966
		0.13963

## FIT NEGATIVE BINOMIAL

CHI-SQUARED = 5.34 CN 9 D.F.

	PARAMETER	S.E.	CORRELATIONS
1	MEAN	4.11966	1.0000
2	VAR	12.94681	0.7699 1.0000

## DEFINING PARAMETERS

	PARAMETER	S.F.	CORRELATIONS
1	M	4.11966	1.0000
2	K	1.92266	-0.1681 1.0000
	X	0.98375	0.00417

## FIT NEYMAN TYPE A

CHI-SQUARED = 11.33 CN 9 D.F.

	PARAMETER	S.E.	CORRELATIONS
1	MEAN	3.80170	1.0000
2	VAR	9.06864	0.7010 1.0000

## DEFINING PARAMETERS

	PARAMETER	S.E.	CORRELATIONS
1	M1	2.74408	1.0000
2	M2	1.38542	-0.9035 1.0000

## TABLE OF OBSERVED AND FITTED VALUES

	NO. OBS.	NO. EXP.	N.B.	NY.A
0	12.	1.9	12.9	15.0
1	19.	7.8	17.0	14.2
2	22.	16.1	16.9	16.6
3	16.	22.2	15.1	16.1
4	10.	22.8	12.6	14.1
5	9.	18.8	10.2	11.6
6	9.	12.9	8.0	9.0
7	4.	7.6	6.2	6.7
8	6.	3.9	4.7	4.7
9	2.	1.8	3.5	3.3
10+	5.	1.1	7.1	5.0
14+	3.	0.0	2.7	0.8
TAIL VALUES	17.	26.	37.	

HISTOGRAM OF OBSERVED AND FITTED VALUES

0 OBSERVED VALUES  
 1 POISSON FIT  
 4 NEGATIVE BINOMIAL FIT  
 5 NEYMAN TYPE A FIT

