APPLICATION DE LA DESCRIPTION STATISTIQUE DES SERIES CHRONOLOGIQUES A L'ETUDE DES TRAJECTOIRES DE SOINS

Journée médicale MODULAD Genève, 22 juin 1991.

Laurence TRICOT

Conservatoire National des Arts et Métiers - Département de mathématiques
292 rue Saint Martin 75141 PARIS CEDEX 05

RESUME
En présence d'une série temporelle non modélisable, il s'agit de réduire l'information en classifiant les éléments de la série compte tenu du temps. On classifie d'abord les intervalles de temps entre deux éléments contigus, puis les éléments eux-mêmes. Les classes obtenues sont appelées épisodes.
Une application sur les contacts d'un groupe de patients suicidaires avec les services psychiatriques de Genève, permet de déterminer des épisodes de soins et de situer les tentatives de suicide par rapport à ceux-ci.

Mots clés : données médicales, épisode, partitionnement, série temporelle, trajectoire.

INTRODUCTION

Un **épisode**, dans le Littré, est "une action incidente liée à une action principale" dans un récit, par exemple. Un épisode est donc une partie d'un processus historique, partie suffisamment homogène pour être discernée en tant que telle. Dans le domaine médical et plus particulièrement psychiatique dont sont issues les données analysées à la fin de cet article, on parle d'"épisode de soins", ou, ce qui n'est pas la même chose, d'"épisode de maladie" ; dans tous les cas il s'agit d'une série d'événements associés entre eux dans le temps. Lorsqu'on parle de **plusieurs épisodes**, chez un patient, on exprime la survenue de plusieurs séries d'événements, séparées par des périodes plus ou moins longues de non-présence d'événements. Chaque début d'épisode à l'exception du premier est significative de reprise, de récidive.

figure 1.

Si l'on observe sur un graphique les événements concernant un individu **(fig1.)**, par exemple la série de ses consultations durant trois ans, les épisodes n'apparaissent pas toujours de manière évidente. Il s'agit bien d'un problème de réduction de l'information par classification de données ordonnées. Des méthodes statistiques adéquates permettront d'effectuer ce genre de regroupement : une fois les épisodes définis, le médecin est en mesure de répondre à des question telles que :

- les soins dispensés l'ont-ils toujours été au même rythme?
- y a-t-il eu récidive ou reprise de soins (selon le type de données)?
- comment se situent dans le temps des événements d'un autre type que ceux qui ont été analysés, par rapport aux épisodes déjà construits ? Par exemple : comment se situent les tentatives de suicide (ou la survenue de tel diagnostic, ou la survenue de tel "événement de vie") observées chez un patient au long cours, par rapport aux soins qui lui ont été dispensés? Au début d'un nouvel épisode? En cours d'épisode?

Un autre exemple, en psychiatrie infantile : à la suite d'une rencontre enregistrée en vidéo, on relève, après avoir partagé la période d'observation en intervalles de temps égaux de l'ordre du dixième de seconde, la direction des regards de l'enfant, de la mère et du médecin durant chaque intervalle de temps, pour aboutir aux données suivantes : qui regarde qui, durant chaque intervalle ? Les "épisodes" de communication par le regard entre l'enfant et la mère, par exemple, peuvent avoir lieu à des moments bien précis de la séance, ou encore, se réduire à un épisode unique avec absence d'effet de l'intervention du médecin, etc.

Une analyse sur le premier thème cité, celui des consultations, sera effectuée dans la seconde partie de cet article, la première partie étant consacrée aux méthodes statistiques permettant de formaliser la notion d'épisode.

1. METHODES

1. 1. Formalisation de la notion d'épisode

1.1.1. série temporelle

On étudie l'évolution d'un phénomène dans le temps. La définition d'une série temporelle relative à ce phénomène sera donnée par deux ensembles et une application :

- un ensemble de **dates**
- un ensemble d'**états** du phénomène étudié
- une **application** notée x du premier ensemble dans le second.

L'ensemble des dates sera inclus dans celui des entiers relatifs.

La série temporelle est une suite, finie ou non de **couples (date,état)** que nous appelerons **éléments**: nous pouvons choisir de noter ainsi la série, en toute généralité : [4]

$$(1) \quad \begin{pmatrix} \dots & -t & -t+1 & \dots & -1 & -2 & 0 & 1 & 2 & \dots & t-1 & t & \dots \\ \dots & x_{-t} & x_{-t+1} & \dots & x_{-2} & x_{-1} & x_0 & x_1 & x_2 & \dots & x_{t-1} & x_t & \dots \end{pmatrix}$$

x_t étant l'état du phénomène à la date t.

Dans le problème traité ici, l'ensemble des dates sera réduit à celui des entiers compris entre 0 et un entier positif T fixé : l'intervalle [0,T] sera appelé **période d'observation.**

Par ailleurs, l'ensemble des états d'un phénomène sera réduit celui des entiers naturels 0, 1, 2, ... : ***nombre de survenues d'un événement*** (par exemple, une consultation dans un service donné) durant une unité de temps.

La donnée de la suite des dates

(2) $(t_1 \ldots t_n)$ $t_1 \geq 0$ $t_n \leq T$

où l'événement est survenu, n étant le nombre d'occurences, suffit pour caractériser la série temporelle (1) concernant un individu avec la convention suivante :
Si plusieurs événements sont survenus durant une unité de temps, nous aurons autant d'éléments consécutifs de la suite (2) ayant la même valeur.

D'autre part, si n est supérieur à 1, seul cas intéressant et toujours supposé par la suite, on notera

$$
\begin{aligned}
\theta_1 &= t_2 - t_1 \\
\theta_2 &= t_3 - t_2 \\
&\ldots \\
\theta_j &= t_{j+1} - t_j \\
&\ldots \\
\theta_{n-1} &= t_n - t_{n-1}
\end{aligned}
$$

les intervalles de temps positifs ou nuls entre les différentes occurences.

On considérera la suite d'intervalles :

(3) $(\theta_1 \quad \theta_2 \quad \ldots \quad \theta_{n-1})$

au paragraphe suivant.

Remarques :

- les entiers T et n dépendent en général de l'individu, ce qui n'apparaît pas ici pour ne pas alourdir les notations.
- l'occurence de plusieurs événements durant la même unité de temps se traduit par la nullité d'un ou plusieurs θ_j.

1.1.2. épisode

Un **épisode** sera une sous-suite de (1) formée d'éléments consécutifs, donc défini à partir d'une sous-suite de (2).

Le calcul d'un épisode se fera en trois étapes :

***1ère étape** : partitionnement de la suite (2) des dates à partir d'un **seuil s** fixé.* Ce paramètre est une commande dont dispose l'observateur qui connaît le contexte dont sont issues les données. Le résultat de cette première étape est une partition en **c classes** ou sous-suites d'éléments contigus de la suite (2) ; les classes sont formées ainsi : t_{j+1} appartient à la même classe que t_j si la différence $\theta_j = t_{j+1} - t_j$ est inférieure ou égale à s. $j = 1, \ldots, n-1$

Exemple 1 : l'expérience médicale dans un centre de soins a montré qu'une consultation demandée plus de trois mois après la précédente est le reflet d'une rechute plutôt que d'une continuité dans la pathologie. Or on dispose pour un patient des données suivantes, les dates étant comptées en nombres de jours après une origine des temps :

(966 967 976 976 984 4413 4450)

ici n = 7.

On en déduit la suite (3) :

(1 9 0 8 3429 37)

pour s = 90 jours on observe un élément de la suite ci-dessus supérieur à 90 : $\theta_5 = 3429$.

1ère classe de dates : (966 967 976 976 984)

2ème classe (4413 4450)

par conséquent : c = 2 (nombre de classes).

***2ème étape** : en vue d'affiner le découpage précédent, on cherche une partition optimale, non plus de la suite des dates, mais de celle des intervalles*, en **c + 1 classes au plus,** c étant le nombre de classes obtenu à la 1ère étape, conséquence de la donnée de s. On choisit c + 1 plutôt que c pour rendre possible la seconde étape même si la première phase a fourni une classe unique.

Il s'agit dans cette seconde étape d'une classification d'objets ordonnés [2], les intervalles. Ce partitionnement optimal est fourni par **l'algorithme de Fisher**, permettant de classifier un nombre fini d'objets ordonnés en un nombre de classes inférieur ou égal à un seuil donné.

La famille de partition optimales ne forme pas une hiérarchie.

Le choix du nombre de classes optimales est à faire à l'aide d'un test F si les données le permettent. "décision par test statistique", ou en considérant comment évolue l'optimum en fonction du nombre de classes, "décision empirique". On notera

$$k \qquad\qquad (k \leq c + 1)$$

un nombre de classes, optimales selon un critère donné.

Critère

Un diamètre de classe est défini [2] : par exemple la somme des carrés des écarts à leur moyenne des éléments de la classe ; le critère à optimiser sera alors la somme des diamètres des k classes, c'est-à-dire la *somme des carrés intra-classes, dont le minimum pour k fixé sera noté* S_k.

S_1 est donc la somme des carrés totale.

décision par test statistique [2]

Dans l'hypothèse d'une densité gaussienne des observations à l'intérieur d'une classe, la matrice de variance étant constante d'une classe à l'autre, la maximisation de la vraisemblance de l'échantillon par rapport aux espérances mathématiques des classes revient à minimiser la somme des carrés intra-classes. Le carré moyen

$$\frac{S_k}{n - k}$$

suit sous ces hypothèses un χ^2 à n - k degrés de liberté. D'autre part, supposons que S_{k+1}, somme des carrés intra-classes optimale pour k + 1 classes, est obtenue en partageant une classe C de la partition en k classes, en deux classes C1 et C2 ; la différence $S_k - S_{k+1}$ se réduit au membre de gauche de l'égalité ci-dessous :

somme des carrés des écarts à la moyenne de la classe C	-	s. des c des écarts à la m. de la classe C1	-	s. des c. des écarts à la m. de la classe C2	=	somme des carrés entre les classes C1 et C2

Le membre de droite de l'égalité ne fait intervenir que les moyennes de C1, de C2, et la moyenne générale. Donc $S_k - S_{k+1}$ suit un χ^2 à 1 degré de liberté, elle est indépendante de S_{k+1} [3] et l'expression

$$(4) \qquad (n - k - 1)\left(\frac{S_k}{S_{k+1}} - 1\right) = \frac{S_k - S_{k+1}/1}{S_{k+1}/(n - k - 1)}$$

suit par conséquent un $F_{1,n-k-1}$.

Mais ceci n'est vrai que sous des conditions de loi gaussienne à l'intérieur des classes, d'égalité des variances des classes, et du caractère hiérarchique [1] de la famille de partitions à partir de k classes, hiérarchie qui signifie des classes bien séparées. Lorsque le nombre k de classes est petit, en l'absence du caractère hiérarchique S_k pourra être beaucoup plus élevé que S_{k+1} et (4) aura tendance à être élevé. Lorsque k est proche de n, (4) se rapprochera de la valeur d'un F. *On choisira le plus petit k pour lequel la valeur de la statistique (4) appartient à l'ensemble des valeurs probables d'un F.*

décision empirique

Les données ne se prêtent pas toujours aux tests F. C'est le cas de celles qui seront analysées en 2ème partie.

Il s'agit, pour choisir k, de considérer les accroissements de la somme des carrés intra-classes minimale lorsqu'on diminue d'une unité le nombre de classes :

(5) $S_c - S_{c+1},\ S_{c-1} - S_c,\ \ldots\ S_2 - S_3,\ S_1 - S_2.$

Ces accroissements sont présents dans la statistique (4), où l'on les rapporte à la somme des carrés de plus grand indice. Lorsqu'on observe la suite des accroissements, tant que les classes ne sont pas homogènes, les sommes de carrés S_j ont de fortes variations. On choisira le plus petit k tel que les accroissements $S_j - S_{j+1}$ soient du même ordre de grandeur pour tout j supérieur ou égal à k : par exemple, on peut choisir le plus petit k tel que

$$S_j - S_{j+1} < 5\%\ S_1$$

Exemple 2 : la suite (2) des dates en nombres de jours est, pour un individu donné :

2012	2016	2020	2026	2040	2042	2048	2061	2083	2096
2103	2124	2146	2147	2174	2195	2218	2232	2237	2265
2351	2362	2453	3468	4229	4576	5000			

ici n = 27.

La suite (3) des intervalles est donc :

4 4 6 14 2 6 13 22 13 7 21 22 1 27 21 23 14 5
28 86 11 91 1015 761 347 424

1ère étape : soit s = 90 jours. On trouve alors c = 6 épisodes correspondant aux traits apparaissant ci-dessous dans la suite des dates :

2012	2016	2020	2026	2040	2042	2048
2061	2083	2096	2103	2124	2146	2147
2174	2195	2218	2232	2237	2265	2351
2362 |	2453 |	3468 |	4229 |	4576 |	5000	

Soit un épisode de 22 événements suivi de 5 épisodes réduits à 1 événement.

2ème étape : l'algorithme de Fisher donne les résultats suivants pour un découpage optimal en 7, 6, 5, 4, 3, 2, 1 classes de la suite des intervalles :

nombre de classes	somme des carrés des écarts	accroissement	accroissement relatif*	commentaire
7	4248.0			
6	5452.8	1204.8	1‰	
5	8417.3	2964.5	1‰	
4	14727.5	6310.2	4‰	
3	46985.5	32258.0	2%	classes (4...91) (1015 761) (347 424)
2	299491.7	252506.2	16%	réunion de (1015 761) et de (347 424)
1	1586744.0	1287252.3	81%	

* accroissement divisé par la somme des carrés totale S_1.

On choisit k = 3 puisque $(S_3 - S_4) / S_1$ est inférieur à 5% et que pour $k < 3$ les accroissements relatifs deviennent supérieurs à $5\% S_1$. Ce qui sépare la suite (3) en trois sous-suites : (4 ... 91), (1015 761), (347 424).

3 ème étape : *partitionnement final de la suite (2) des dates* à partir de celui de la suite (3) des intervalles en k classes.

Dans le cas particulier où k = 1, la suite (1) constitue un épisode unique.
A partir de k = 2, la suite des dates sera partagée de la façon suivante :

Règle de construction d'un épisode

soit (θ_j .. θ_{j+h}) une classe d'intervalles (réduite à un élément si h = 0) ; on a :

$j \geq 1 \qquad j + h \leq n - 1$ et $\qquad \theta_j = t_{j+1} - t_j$

$$\vdots$$

$$\theta_{j+h} = t_{j+h+1} - t_{j+h}$$

Il correspond à la suite (θ_j .... θ_{j+h}) une sous-suite de (1) appelée **épisode** contenant :

- t_j si $j = 1$ ou si t_j est plus proche de t_{j+1} que de t_{j-1}
- t_{j+h+1} si $j+h+1 = n$ ou si t_{j+h+1} est plus proche de t_{j+h} que de t_{j+h+2}
- tous les éléments intermédiaires s'il y en a.

Exemple

A partir des trois classes d'intervalles obtenues dans l'exemple 2, on obtient les épisodes suivants :

classe d'intervalles	épisode	commentaire
(4.....91) θ_1 θ_{22}	$\begin{pmatrix} 2012 \ldots 2453 \\ t_1 \quad t_{23} \end{pmatrix}$	j = 1, h = 21 t_1 = 2012 appartient au premier épisode d'après la règle de construction. t_{23} = 2453 est plus proche de t_{22} = 2362 que de t_{24} = 3468.
(1015 761) θ_{23} θ_{24}	$\begin{pmatrix} 3468 \\ t_{24} \end{pmatrix}$	j = 23, h = 1 t_{23} a déjà été classé. t_{25} = 4229 est plus proche de t_{26} = 4576 que de t_{24} = 3468.
(347 424) θ_{25} θ_{26}	$\begin{pmatrix} 4229 \ldots 5000 \\ t_{25} \quad t_{27} \end{pmatrix}$	j=25, h = 1 t_{25} a été exclu de l'épisode précédent et appartient donc à celui-ci. t_{27} appartient au dernier épisode par construction.

Ainsi, les cinq traits tracés à l'intérieur de la suite des dates pour délimiter les épisodes à la première étape, se réduisent à deux, ceux qui isolent la date 3468. On peut observer que l'épisode qui précède l'événement survenu à le date 3468 et l'épisode qui suit cet événement n'ont pas la même structure interne : le premier a beaucoup d'événements très rapprochés, le dernier est constitué d'événements très distants dans le temps.

***4ème et dernière étape** : la prise en compte de la fin de la période d'observation.*
Pour un individu, les dates sont observées durant une période d'observation [0, T]. Le dernier épisode calculé peut donc être **tronqué**. Pour décider s'il en est ainsi, on introduit dans les calculs, dès la première étape, un **événement fictif** à la date T. Ainsi, la suite (2) est complétée par $t_{n+1} = T$.

Après la 3ème étape, deux cas peuvent se produire pour le dernier épisode :

- ou celui-ci est constitué par un élément unique : l'événement fictif à la date T, date très éloignée de t_n
- ou il contient plusieurs éléments, ce qui signifie que la différence $T - t_n$ est du même ordre que $t_n - t_{n-1}$.

Dans le premier cas, on supprimera le dernier épisode, fictif, de la description de la série temporelle, dont le dernier événement est daté à t_n.
Dans le second cas, on considérera que le dernier épisode est tronqué.

Exemple : dans l'exemple 2, supposons T = 5000 : cette date constitue une classe à elle seule. Donc l'avant-dernier épisode (constitué lui aussi d'un seul événement) n'est pas tronqué. Il l'aurait été si le résultat final avait réuni les dates 4576 et 5000.

1.2. La nouvelle version du programme EPIDOR

La nouvelle version du programme effectue la première étape de la construction des épisodes, ainsi qu'une partie des seconde et quatrième étapes :

- dans la première étape, le seuil s est un paramètre de commande
- dans la seconde étape sont programmés
  - le calcul des S_k, pour k = 1, ..., c+1
  - l'expression (4), valeur du F, pour k = 1, ..., c ;
- dans la quatrième étape le programme introduit automatiquement un événement fictif à la date T, paramètre de commande.

Le reste, c'est-à-dire

- le choix final du nombre d'épisodes au vu du comportement des S_k (2ème étape)
- la partition de la suite des dates (3ème étape)
- la décision sur le caractère tronqué du dernier épisode d'après le contenu de celui-ci,

sont à faire manuellement en considérant les résultats.

L'ancienne version d'EPIDOR est dans la bibliothèque MODULAD et fait l'objet d'un article dans la Revue ([5]).

1.3. Description d'une série temporelle

1.3.1. Trajectoire

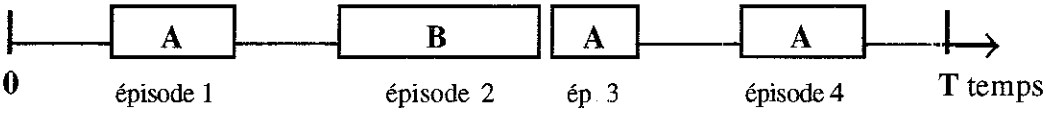
Une trajectoire est ici le résultat d'une réduction de l'information : une description "par épisodes" de la série des événements concernant un individu.

Les épisodes construits auront chacun leur structure interne ; à chaque épisode on peut attribuer un type, à partir de sa représentation par différentes variables telles que
- nombre d'événements
- intervalle moyen entre 2 événements contigus
- durée totale
- concentration des événements dans le temps

etc.
De telles variables permettent d'effectuer une typologie par l'analyse des données de l'ensemble des épisodes construits pour un échantillon d'individus.

Conséquence : une fois obtenus des types d'épisodes, la **trajectoire** d'un individu pourra être représentée par exemple ainsi (**fig.2**):

figure 2.



Cette réduction de l'information facilite la comparaison des individus entre eux.

1.3.2. Ruptures dans la trajectoire

Considérons le cas particulier où, dans la phase de partitionnement des intervalles par l'algorithme de Fisher, on a isolé un intervalle formant une classe à lui tout seul. ce résultat est dû au caractère "hors norme" d'un tel intervalle par rapport à ceux des classes contiguës. "Hors norme" peut vouloir dire "beaucoup plus petit" (premier cas) ou "beaucoup plus grand" (second cas).

D'après le mode de construction des épisodes à la 3ème étape, le premier cas implique que les deux dates définissant l'intervalle isolé se retrouvent ensemble, sans être associées à d'autres, pour donner lieu à un épisode de deux éléments proches dans le temps. Le second cas implique que les deux dates vont s'associer respectivement à des dates antérieure et postérieure, pour former des épisodes séparés par un grand intervalle de temps. C'est dans ce cas que nous parlerons de **rupture dans la trajectoire**.

Exemples :

- Sur la **fig.2.** on observe 4 épisodes, et 2 ruptures : entre les épisodes 1 et 2, et entre les épisodes 3 et 4.
- Dans l'exemple 1 (§1.1.2) la 2ème étape donne trois classes d'intervalles :

  (1 9 0 8) (3429) (37)

La 3ème étape confirme le résultat de la 1ère étape :

	$(S_{j-1} - S_j) / S_1$
S_4 =48.7	0.0‰
S_3 =65.0	59.1%
S_2 =5752897.0	40.9%
S_1 =9736513.3	

L'avant-dernier intervalle 3429 forme une classe à lui tout seul dans la partition de la suite (3), et il y a rupture entre les deux épisodes caractérisés par les classes de dates (966 967 976 976 984) et (4413 4450).

2. APPLICATION A DES DONNEES MEDICALES

2.1. Les données

352 patients ont été répertoriés dans les Institutions Universitaires de Psychiatrie de Genève comme ayant fait au moins une tentative de suicide en 1989. Précisons qu'à Genève, un patient vu en urgence pour une tentative de suicide est obligatoirement vu par un psychiatre. Le nombre cité est donc une assez bonne estimation des patients suicidaires à Genève adultes et non retraités cette année-là. Par rapport au nombre d'habitants, il s'agit d'un taux inférieur à 1/1000.

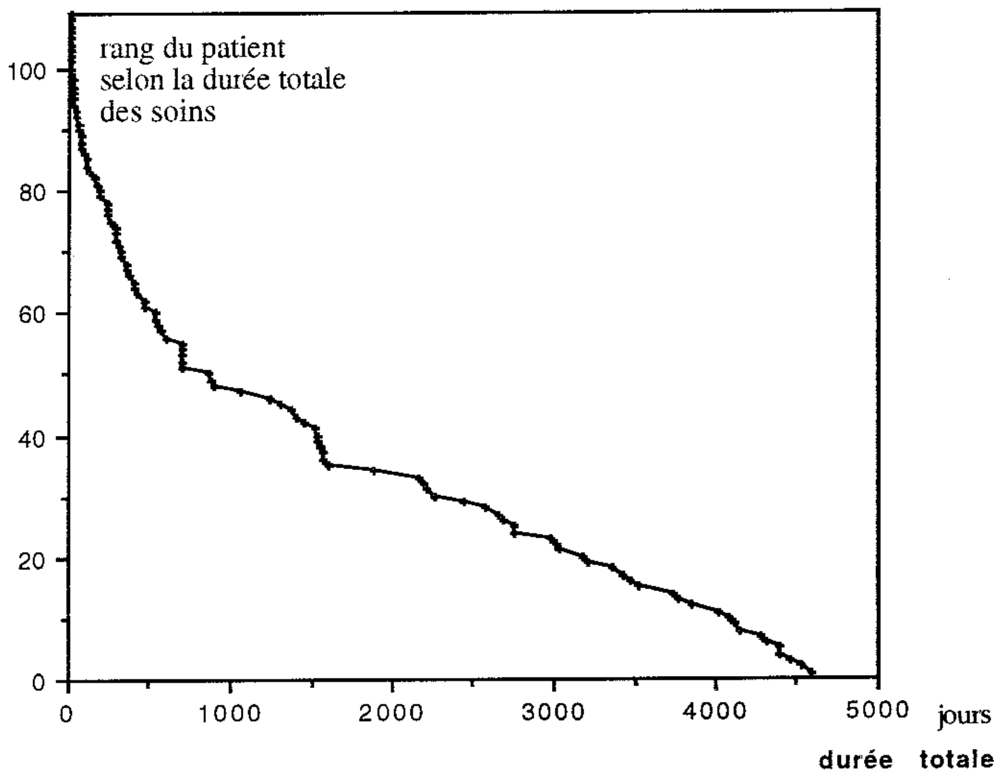
Un échantillon de 176 patients (un sur deux) a été retenu par tirage systématique.

Les contacts de ces patients avec les Institutions ont été observés durant 5000 jours (entre 13 et 14 ans), jusqu'au 10 septembre 1990.

Sur cette très longue période d'observation, on observe qu'un certain nombre de patients ont été "vus", mais non suivis, du moins en institution : *pour 38% des patients la durée totale des soins, c'est-à-dire le nombre de jours qui séparent le premier contact du dernier, est inférieurs à 5 jours.*
Ce pourcentage est classique dans les études de fonctionnement de la psychiatrie genevoise depuis 1978.

Les patients de l'ensemble complémentaire sont au nombre de **109**. Le **diagramme 1.** présente la durée totale des soins dispensés à chacun d'entre eux, les patients étant ordonnés selon cette durée totale. On voit qu'une part importante des patients fait l'objet de soins durables, sans qu'il soit parlé ici de quantité ni d'intensité.

diagramme 1.



Trois lieux de soins ont été retenus :

- le centre de consultations ambulatoires
- le centre de thérapies brèves (C.T.B.)
- la clinique psychiatrique.

Pour chaque patient, un événement est caractérisé par :

- sa date (unité de temps : le jour)
- son lieu.

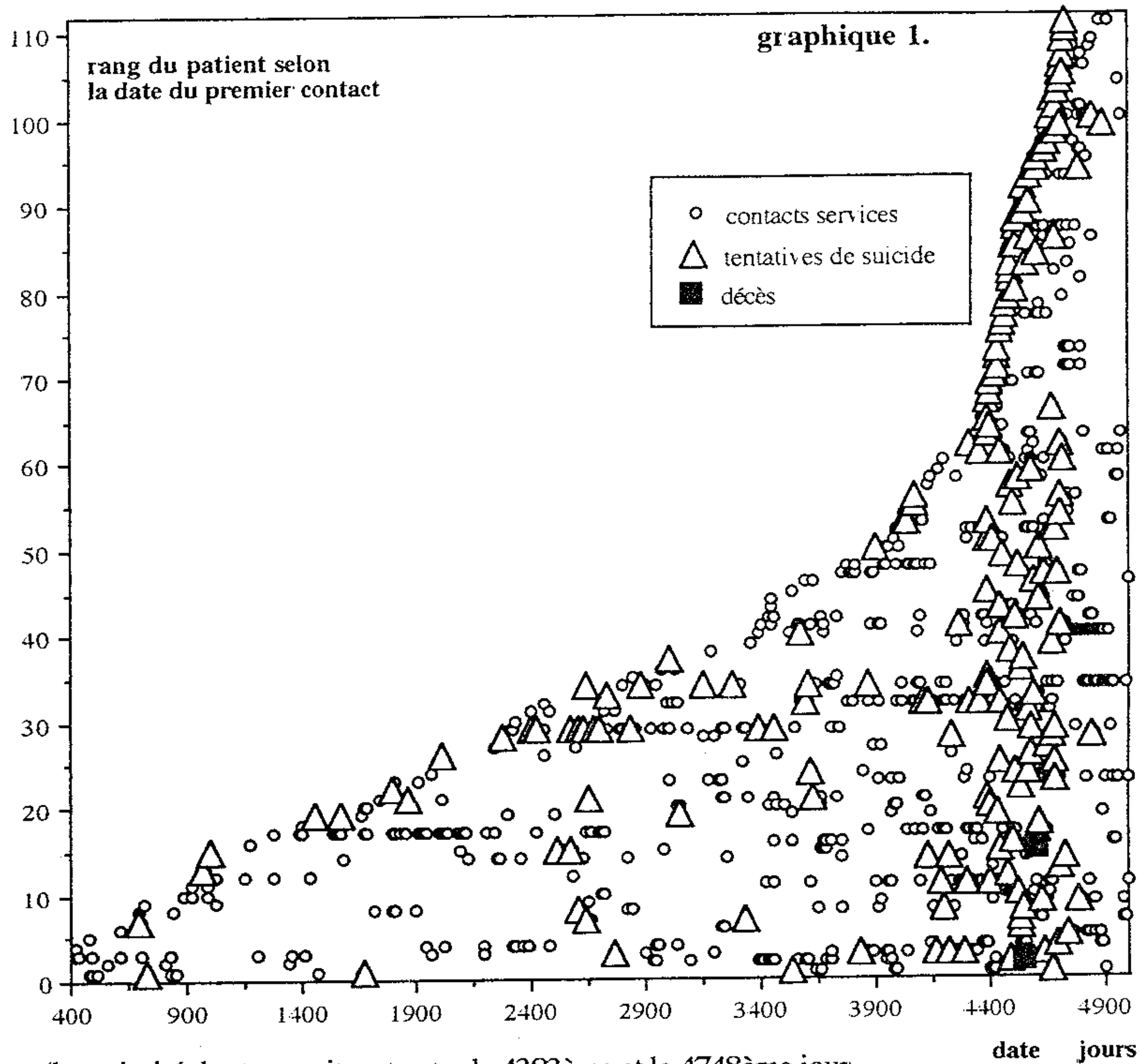
2.2. Tentatives de suicide et épisodes de soins

Un patient étant donné, on relève pour chaque unité de temps, comme il a été dit au §1.1.1, le nombre de contacts dans un lieu de soins : en général, on observe 0 ou 1 contact par jour. Les différents lieux sont traités séparément :

- pour les consultations ambulatoires, le seuil s de la première étape de construction des épisodes (§ 1.1.2) a été fixé à 90 jours ;
- pour le C.T.B., où les patients sont vus presque quotidiennement sur une courte période, le seuil s a été fixé à 7 jours ;
- pour la clinique psychiatrique, où les séjours sont continus, les méthodes décrites dans la première partie n'ont pas été utilisées : en effet, l'épisode de soins est ici clairement délimité par une admission et par une sortie.

Le **graphique 1.** montre l'ensemble des événements qui délimitent les épisodes dans les trois lieux de soins (voir §2.1.), les tentatives de suicide et les décès, relatifs aux 109 premiers patients lorsque ceux-ci sont ordonnés d'après la date du début des soins. Un tel classement n'a pas d'autre raison que la lisibilité du graphique. Au numéro d'ordre d'un individu lu en ordonnée, correspondent les débuts et fins d'épisodes calculés pour cet individu : il s'agit donc du résultat de l'analyse décrite au §1.

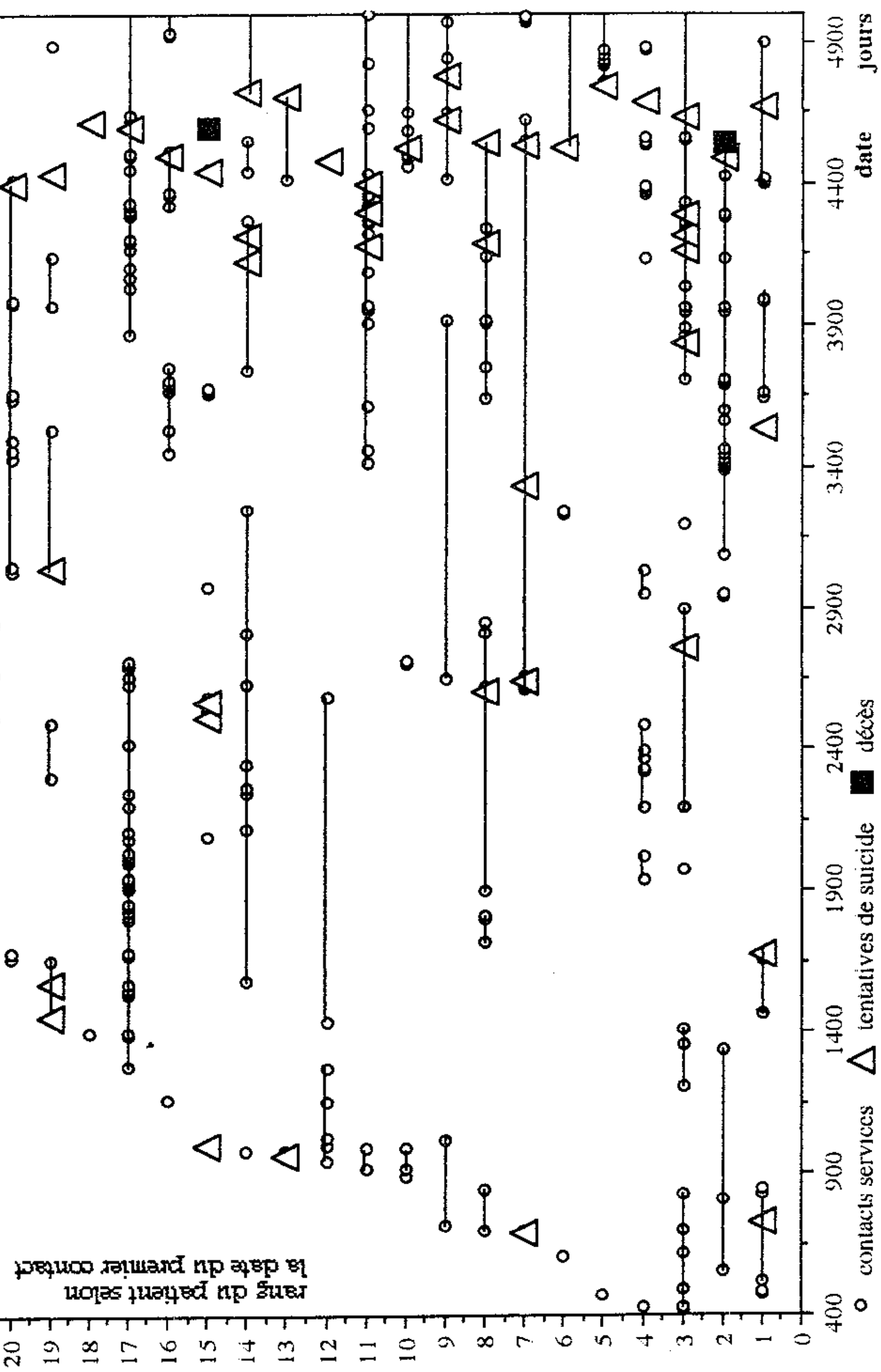
Remarque :
Les trois lieux de soins ne sont pas différenciés sur ce graphique, ni sur le **graphique 2.** plus bas. Mais ils l'ont été dans le calcul des épisodes : un épisode ne concerne qu'un lieu.



(la majorité des t.s. se situent entre le 4383ème et le 4748ème jour
à cause du mode de sélection des patients : voir§ 2.1.)

Pour un graphique plus précis, on représentera chaque épisode par des traits horizontaux reliant son début et sa fin. Le **graphique 2.** montre le résultat pour les 20 premiers patients : par exemple, on observe pour le patient n° 1 quatre épisodes de soins

Une fois définis les épisodes, il est simple de situer par rapport à ceux-ci les tentatives de suicide. On voit sur le **graphique 2.** que le patient n°8 a fait une tentative au cours graphique 2.


rang du patient selon
la date du premier contact
20
19
18
17
16
15
14
13
12
11
10
9
8
7
6
5
4
3
2
1
0
400
900
1400
1900
2400
2900
3400
3900
4400
4900
date
jours
contacts services
tentatives de suicide
décès

du troisième épisode, et deux autres au cours du quatrième, ces tentatives ayant eu lieu vers la fin de sa prise en charge observée.

Les 109 patients sélectionnés se répartissent comme suit d'après le nombre d'épisodes :

Nombre d'épisodes*	nombre de patients	%
0	7	6.5
1	30	27.5
2	21	19.0
de 3 à 5	19	17.5
de 6 à 8	15	14.0
de 9 à 11	2	2.0
de 12 à 14	5	4.5
15 et plus	10	9.0
total	109	100.0

*non compris les épisodes réduits à un événement unique. C'est pourquoi il existe des patients présentant 0 épisode.

D'autre part, on observe 176 tentatives de suicides chez ces 109 patients, réparties comme suit :

situation de la tentative de suicide	nombre de tentatives de suicide	%
événement isolé	39	22.0
au début d'un épisode	66	37.5
durant un épisode	54	31.0
à la fin d'un épisode	17*	9.5
total	176	100.0

* dont 13 à la fin de l'ensemble des soins observés.

Lorsque la tentative de suicide est un événement isolé, on peut penser que souvent le patient est suivi en milieu privé.

Au vu des résultats du dernier tableau, il serait nécessaire de faire une étude clinique des cas, relativement nombreux, où la tentative a lieu pendant ou à la fin d'un épisode : il s'agit de 40% des tentatives, et de 45 patients, soit 41% des 109 cas suivis.

OUVRAGES CITES

[1] CHANDON J.L., PINSON S. Analyse typologique. Masson 1981.

[2] HARTIGAN J.A. Clustering Algorithms. Wiley & Sons 1974.

[3] SAPORTA G. Probabilités, analyse des données et statistique. Technip 1990.

[4] TRICOT L. Sur la notion de distance entre opérateurs. Application à quelques problèmes d'identification des modèles à retards échelonnés. Thèse de doctorat. Peter Lang. Berne 1980.

[5] TRICOT L. Homogénéisation de processus observés : les programmes EPIDOR, EPICLA et EPIMAT. La Revue de Modulad n° 2, pp 1-12, 1988.