



Les répartitions en classes et quelques unes de leurs applications [Partie 1]

Philippe Sentis

► To cite this version:

Philippe Sentis. Les répartitions en classes et quelques unes de leurs applications [Partie 1]. Annales de l'ISUP, 1958, VII (1), pp.15-91. hal-04094608

HAL Id: hal-04094608

<https://hal.science/hal-04094608>

Submitted on 11 May 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

INTRODUCTION

LES RÉPARTITIONS EN CLASSES ET QUELQUES UNES DE LEURS APPLICATIONS

par

M. Philippe SENTIS

LES SEPARATIONS EN CLASSES
ET QUELQUES UNES
DE LEURS APPLICATIONS

par

M. PHILIPPE BERTRAND

INTRODUCTION

La théorie des variables aléatoires est bien connue lorsqu'on possède une infinité dénombrable de variables aléatoires indépendantes. C'est ainsi que l'on sait former des séries de variables aléatoires indépendantes. Habituellement quand on ne se trouve pas dans ce cas, on cherche à s'y ramener. Toutefois il a semblé à l'auteur du présent mémoire qu'une étude directe n'était peut-être pas impossible et le travail suivant essaye de fournir quelques méthodes nouvelles d'approche du cas général. Dans le premier chapitre on verra que toutes les variables dont les valeurs sont fixées par un même événement peuvent s'exprimer comme des fonctions certaines d'un ensemble de variables dont on peut se fixer arbitrairement les valeurs. On verra au quatrième chapitre que si des variables initiales se prêtent à une étude statistique, on peut les exprimer comme des fonctions certaines de variables aléatoires indépendantes.

Le deuxième chapitre précise en effet ce qu'on peut entendre par variables qui se prêtent à une étude statistique. Ce sont toutes celles dont on peut répartir les valeurs en classes. Si l'on sait définir plusieurs répartitions en classes sur l'ensemble des valeurs de la variable, on dit que cet ensemble possède une structure d'ensemble classifiée.

Les ensembles classifiés possèdent des propriétés qui font l'objet des chapitres 3 et 4. Ces propriétés sont intéressantes au point de vue du calcul des propriétés car elles permettent de définir des fonctionnelles σ - additives et π - multiplicatives, généralisant aussi les notions de série et de produit infini à des ensembles d'indices qui peuvent n'être pas dénombrables. La généralisation des indices par les intégrales et les mesures était déjà connue mais ne permettait pas d'intégrer les fonctions à valeurs aléatoires. La présente étude, en introduisant une nouvelle espèce de multiplication d'une variable aléatoire par un nombre, permet de définir l'intégrale d'une fonction à valeurs aléatoires indépendantes et la dérivée d'un processus de Wiener-Lévy. La généralisation des produits infinis était connue comme exponentielle

d'une intégrale. Celle qui est définie ici permet d'étudier directement des produits connus pour des matrices ou des noyaux intégraux par exemple.

On verra aux chapitres 5 et 6 que les méthodes exposées précédemment permettent de nombreuses applications au calcul des probabilités et à la statistique mathématique telles que la généralisation des équations de Kolmogoroff et la solution d'un problème pratique qui se pose dans certaines papeteries.

A ce sujet je remercie M. Samuel-Lajeunesse, Directeur de la C. E. N. P. A. qui m'a autorisé à faire quelques expériences dans l'une de ses usines.

Je remercie M. le Professeur Darmois qui m'a aidé de ses conseils au cours de ce travail et a bien voulu présider le Jury ainsi que MM. Dugué et Pisot qui ont bien voulu en faire partie.

CHAPITRE I

SUR LES RELATIONS ENTRE LES VARIABLES ALÉATOIRES DÉTERMINÉES PAR UN MÊME ÉVÈNEMENT

Le travail suivant a pour objet l'étude des différents aspects possibles d'un même évènement et les relations entre ces aspects.

Etant donné un évènement, nous pouvons toujours trouver dans les circonstances qui l'accompagnent, un grand nombre d'aspects différents, on pourrait même dire une infinité d'aspects différents.

Si nous considérons le jet d'un dé nous pouvons nous intéresser au nombre de points portés par la face supérieure, mais également au nombre de points portés par une des faces latérales, celles par exemple que l'on voit successivement du Nord et de l'Est. On pourrait également s'intéresser au gain de la partie si ce jet de dés est celui qui doit en décider et à la fortune des joueurs telle qu'elle en résultera. On peut également s'intéresser à la position du centre de gravité du dé définie par ses coordonnées. Si l'on considère l'une quelconque de ces variables que nous venons d'envisager et si l'on considère le comportement des autres variables en fonction de celle-ci différents cas peuvent se présenter. Prenons pour fixer les idées le nombre de points portés par la face supérieure comme variable principale. Ce sera une variable pouvant prendre l'une des six valeurs entières numérotées de 1 à 6. La donnée de cette variable déterminera complètement le gain de la partie et par conséquent la fortune des joueurs à la fin de celle-ci. Par contre, il est bien évident que les coordonnées du centre de gravité peuvent être choisies arbitrairement quel que soit le nombre de points porté par la face supérieure.

Les premières variables dépendent d'une façon certaine de la variable aléatoire principale. Les dernières n'en dépendent certainement pas. Nous dirons qu'elles en sont faiblement indépendantes. Mais

il existe des variables qui n'entrent ni dans l'un ni dans l'autre de ces deux cas : par exemple, le nombre des points porté par une face latérale ; dans ce cas nous dirons que ces variables dépendent stochastiquement de la variable principale. La variable Y définie par le nombre de points porté par une face latérale est un nombre entier compris entre 1 et 6 bornes incluses ; mais quand le nombre X des points porté par la face supérieure est fixé Y ne peut prendre ni cette valeur ni la valeur correspondant au nombre des points porté par la face opposée.

Pour un $dé$ donné, le nombre des points porté par la face opposée une face donnée, est donné. Nous admettrons pour simplifier l'exposé que les nombres des points portés par deux faces opposées diffèrent de trois unités. Dans ces conditions la différence entre Y et X sera un nombre différent de 0 et de 3 ; autrement dit $|Y - X|$ vaut 1, 2, 4 ou 5 ce que l'on peut écrire également :

$$Y = X + Z \pmod{6}$$

Z étant pris arbitrairement parmi l'un des nombres 1, 2, 4 ou 5.

Ainsi Y est une fonction certaine de la variable X et d'une variable Z faiblement indépendante de X . Nous allons démontrer qu'il n'existe pas d'autre mode de relation entre deux variables liées à un même événement que ceux que nous venons d'envisager.

Remarquons que l'exemple traité ci-dessus ne suppose nullement que les variables soient probabilisées (il peut très bien s'agir d'un $dé$ pipé de façon inconnue).

Nous exprimerons que les Y_j et X_i sont liés à un même événement en posant que si les Y_j peuvent prendre leur valeur dans des ensembles N_j pour tout système de valeurs Y_j assignées aux Y_j on peut trouver au moins un système de valeurs x_i prises dans des ensembles N_i que l'on puisse assigner aux X_i et réciproquement. Nous dirons que les X_i et Y_j appartiennent à un même champ. Cette relation est manifestement réflexive ; de plus, elle est transitive : si les X_i et les Y_j appartiennent à un même champ et si les Y_j et les Z_k appartiennent à un même champ les X_i et les Z_k appartiennent à un même champ. En effet, les X_i étant donnés on peut trouver les Y_j donc les Z_k et réciproquement.

SYSTEME DE VARIABLES :

Nous dirons que Y appartient au système S des variables X_i si Y appartient au champ des X_i et si les valeurs x_i des X_i étant données la valeur y de Y est fixée d'une façon unique, il existe des systèmes de variables. En effet, on peut toujours considérer le système constitué par

la seule variable X et les fonctions certaines de cette variable. Nous appellerons (X) un tel système. Nous dirons qu'un ensemble F de variable Y_j appartient au système S si pour tout Y_j appartenant à l'ensemble F , Y_j appartient au système S . Tout Z appartenant au système S' des Y appartiendra au système S .

Nous dirons que S' est un sous-système de S si tout Z appartenant au système S' appartient au système S . En particulier tous les (X) sont des sous-systèmes de S .

Par définition nous dirons que les systèmes S et S' sont équivalents si tout X_i appartenant au système S appartient au système S' et si tout X , appartenant au système S' appartient au système S - autrement dit chacun d'eux est un sous-système de l'autre il est évident qu'il est possible de trouver deux définitions du système telles que les systèmes correspondant à ces deux définitions soient équivalents il suffit de définir par exemple un des systèmes S par les variables fixées de façon unique quand les valeurs d'un ensemble donné E de variables sont données et le système S' par les variables fixées de façon unique quand sont données les valeurs des variables d'un ensemble E' qui ne diffèrent de E que par la substitution à une (ou plusieurs) des variables d'une variable qui lui correspond bi-univoquement (leur correspondent).

Nous dirons que Y est faiblement indépendant du système S si les valeurs x_1 des X_1 étant données, la valeur y de Y peut être choisie arbitrairement dans un ensemble M donné quelles que soient les valeurs x_1 des X_1 . En particulier, si Y est constant, il est faiblement indépendant de tout S .

Si Y est faiblement indépendant du système S il est faiblement indépendant de tout sous-système S' de S et en particulier de tous les X .

THEOREME I - 1 :

Si Y est faiblement indépendant de la variable X_1 et si X appartient au système (Y) X est constant.

En effet, supposons que X puisse prendre deux valeurs distinctes x_1 et x_2 , chacune de ces valeurs étant données, il est possible de choisir Y dans le même intervalle. Soit y une valeur déterminée de cet intervalle ; il est possible de choisir pour X , x_1 ou x_2 ce qui est contraire à l'hypothèse que X appartient au système Y .

THEOREME I - 2 :

Si Y est faiblement indépendant de la variable X et si X appartient

au champ de (Y) sans appartenir au système (Y) X est faiblement indépendant de Y.

En effet, supposons que X puisse prendre ses valeurs dans un ensemble M, pour toute valeur x de M, Y prend ces valeurs dans un ensemble M' indépendant de x, par conséquent pour toute valeur y de Y prise dans l'ensemble M' il sera possible de fixer à X n'importe quelle valeur de l'ensemble M. Par conséquent la valeur y de Y étant fixée, on pourra assigner à X n'importe quelle valeur choisie dans un ensemble M_y identique à M et par conséquent indépendant de y ; par suite Y sera algébriquement indépendant de Y.

On remarquera que pour des variables appartenant au champ des Y les hypothèses des deux théorèmes précédents sont contradictoires et par conséquent nous avons épuisé tous les cas possibles donc. Si Y appartient au champ des X et n'est pas faiblement indépendant du système (X), nous dirons qu'il est lié au système (X). Il résulte du théorème 2 que si Y est lié à X, X est lié à Y ; on dira que X et Y sont liés. Si Y est lié au système S des X_i quand les valeurs x_i des X_i sont données, la valeur y des Y peut être choisie arbitrairement dans un ensemble M_x qui sera fonction des valeurs x_i assignées aux X_i .

THEOREME I - 3 :

Si Y appartient au champ des X_i il appartient au système des X_i et d'une variable Z faiblement indépendante des X .

Si Y est indépendant des X_i on peut prendre pour variable Z la variable Y elle-même. Si Y est lié aux X_i on peut prendre pour variable telle qu'on puisse définir une application de l'ensemble M_x sur l'espace des échantillons de Z quelles que soient les valeurs x (on pourra prendre par exemple l'espace produit de tous les M_x ce qui démontre l'existence d'au moins une telle variable. Il va de soi que cette variable ne sera pas la plus simple que l'on puisse choisir. A ce propos, si l'on se reporte à l'exemple traité précédemment on voit que l'on a pu prendre un ensemble formé de 4 points seulement alors que l'espace-produit en eût compté 4^5).

REMARQUES :

1/ - les X étant fixés Z est déterminé par Y.

2/ - Si l'ensemble M_x se réduit à un élément quel que soit X, Y est une fonction certaine des X. C'est le cas où Y non seulement est lié aux X mais encore appartient au système des X. Si X et Y sont liés, ou bien Y appartient au système des X ou bien il ne lui appartient pas ; si Y appartient au système des X c'est une fonction certaine des X. Si de plus,

X appartient au système des Y, à tout x correspond un y et réciproquement ; il y a entre X et Y une correspondance bi-univoque. Nous dirons que X et Y sont liés bi-univoquement.

Si Y n'appartient pas au système des X, ou bien X est une fonction certaine de Y ou il n'appartient pas au système des Y. Dans ce cas, nous dirons que X et Y sont liés stochastiquement.

En résumé, si X et Y appartiennent à un même champ, nous sommes dans l'un des cinq cas suivants :

- 1/ - X et Y sont faiblement indépendants
- 2/ - X et Y sont stochastiquement liés
- 3/ - X et Y sont liés bi-univoquement
- 4/ - X est fonction certaine de Y
- 5/ - Y est fonction certaine de X.

Toutes les variables qui appartiennent à un même champ ne forment peut-être pas un ensemble ; ou du moins je n'ai pas démontré qu'elles en formaient un (1) - Mais dans la suite de ce travail, je ne considérerai que des ensembles de variables qui appartiennent à un même champ.

Dans ce cas, il sera possible d'établir le théorème suivant :

THEOREME I - 4 :

Etant donné un ensemble E satisfaisant à l'axiome de Zermelo de variables Y_j appartenant à un même champ, il est possible de trouver un ensemble B de variables X_i faiblement indépendantes au système desquelles appartiennent les Y_j .

DEFINITION :

L'ensemble B sera appelé la base du système. Soit B_k un ensemble de variables aléatoires faiblement indépendantes et E_k l'ensemble des variables de E qui appartiennent au système de base B_k .

Soit Y_k une variable de E qui n'appartient pas à E_k . Ou bien Y_k est faiblement indépendant de B_k et on pose $X_k = Y_k$, ou bien Y_k est lié à B_k et il existe une variable X_k faiblement indépendante de B_k (puisque Y_k n'appartient pas au système de base B_k) telle que Y_k appartienne au système de base B_k X_k . Les variables qui constituent l'ensemble

(1) J'ai cependant utilisé l'expression "appartenir à" car elle ne présente aucune ambiguïté.

B_k X_k sont faiblement indépendantes ; Donc, l'on peut poser :

$B_{k+1} = B_k \cup X_k$ et en déduire un ensemble E_{k+1} . L'ensemble E_{k+1} contient l'ensemble E_k ; nous pouvons donc déterminer une suite strictement croissante d'ensembles E_k auxquels correspondent des ensembles B_k formant également une suite strictement croissante. D'après le théorème de Zermelo, la suite E_k possède un élément maximal identique à E auquel correspondra pour la suite B_k un élément maximal que l'on posera identique à B .

Soit une variable Y appartenant à un système S de variable X et soit x_i un ensemble de variables indépendantes ; nous pouvons toujours considérer cet ensemble de variables comme une variable unique définie sur le produit des ensembles sur lesquels sont définis les x_i de sorte que tout ce qui a été dit sur la correspondance entre y et une variable x est encore valable pour celle entre Y et un ensemble x_i de variables indépendantes.

Dans la suite nous continuerons à parler de correspondance entre Y et X pour simplifier mais il est clair que la théorie sera encore valable si on remplace X par un système de variables indépendantes.

On appellera sous-base d'une base donnée tout ensemble de variables appartenant à la base. Si Y est faiblement indépendant d'un système S de base $B_1 \subset B$ il est faiblement indépendant de tout sous-système de base $B_2 \subset B_1$; donc, si Y est faiblement indépendant de B_1 et B_2 il est faiblement indépendant de $B_1 \cap B_2$. Si Y appartient au système de base $B_1 \subset B$ il appartiendra à tout système de base $B_2 \supset B_1$, $B_2 \subset B$. Donc, si Y appartient aux systèmes de bases B_1 et B_2 il appartient au système de base $B_1 \cup B_2$.

THEOREME I - 5 :

Soit 2 sous-bases B_1 et B_2 complémentaires, si Y est fonction des $X_i \in B_1$ il est faiblement indépendant des $X_j \in B_2$.

En effet, si les x_i étant donnés Y est déterminé, on peut fixer arbitrairement les valeurs des x_j et celle de Y , car si nous nous donnons la valeur y de Y , il existe un système de valeurs $X_i = x_i$ $X_j = x_j$ des X_i et des X_j pour lesquelles Y prend la valeur y ; Y étant déterminé par la donnée des seuls x_i on aura $Y = y$ pour $X_i = x_i$ $X_j = x'_j$ quels que soient les x'_j , donne en particulier si ce sont les valeurs fixées à l'avance pour les x_j . La réciproque de ce théorème n'est pas vraie : Si Y est indépendant des X_j sans prendre la même valeur pour deux systèmes arbitraires de valeurs attribuées aux X_i quand les x_i sont fixés... On ne peut rien dire de la relation entre Y et les X_i .

Si Y est fonction des X_i nous dirons qu'il est fortement indépendant des X_j . Pour que Y soit fortement indépendant des X_j il faut et il suffit que y prenne la même valeur quelles que soient les valeurs x_i attribuées aux X_j pour un même système de valeurs x_i attribuée aux X_i . Si Y est fortement indépendant des X_j et si $B_k \subset B_j$, Y est fortement indépendant des X_k . En effet, Y appartient au système des X_i en posant $B = C B_k$ puisque $B_i \subset B_1$.

Les variables appartenant à la base sont fortement indépendantes entre elles. En effet, X_i appartient à tout système de base B_1 telle que $X_i \in B_1$ donc est fortement indépendant de tout B_2 tel que $X_i \notin B_2$ et en particulier de X_j tel que $X_j \neq X_i$.

THEOREME I - 6 :

Si Y est fortement indépendant des éléments de B_1 et de B_2 il est fortement indépendant des éléments $B_1 \cup B_2$. En effet, posons :

$$B_1 = B_1 \cap B_2$$

$$B_j = B_1 \cap C B_2$$

$$B_k = C B_1 \cap B_2$$

$$B_1 = C B_1 \cap C B_2$$

Etant donné $Y = y$ il existe au moins un système de valeurs:

$$X_i = x_i$$

$$X_j = x_j$$

$$X_k = x_k$$

$$X_1 = x$$

des X qui lui correspond.

Y étant indépendant des éléments de B_2 on a la même valeur $Y = y$ pour $X_i = x'_i$

$$X_j = x_j$$

$$X_k = x'_k$$

$$X_1 = x_1$$

quels que soient les x'_i et les x'_k ; mais puisque Y est indépendant des éléments de B_1 , on a la même valeur $Y = y$ pour $X_i = x'_i$

$$X_j = x'_j$$

$$X_k = x_k$$

$$X = x$$

quels que soient les x'_j .

On peut donc se fixer arbitrairement y et les valeurs des X_i des X_j et des X_k , deux systèmes quelconques x_i, x_j, x_k et x'_i, x'_j, x'_k de valeurs des X_i, X_j, X_k donnant la même valeur y à Y pour un même système de valeurs x des X . Par suite

Y est fortement indépendant de :

$$B_1 \cup B_2 = C(CB_1 \cap CB_2)$$

Donc, si Y appartient aux systèmes de bases B_1 et B_2 , il appartient au système de base $B_1 \cap B_2$.

Soit H_t un ensemble de variables X_j fortement indépendantes de Y et soit $H = \bigcup H_t$, la réunion étant étendue à tous les H_t possibles. Tout ensemble de variables X_j sera inclus dans H et toute variable fortement indépendante de Y sera élément de H . Y appartiendra à un système de base $P = C H$ et il n'existera pas de système de base $Q \subset P$ et Q différent de P possédant la même propriété. En effet, s'il en existait un, Y serait fortement indépendant du système de base CQ donc $CQ \subset H$ et $P \subset Q$, donc $P = Q$.

P s'appellera la base minimum dont dépend Y .

Nous dirons que Y_1 est fortement indépendant de Y_2 si quelles que soient les valeurs x_j des X_j dont dépend Y_2 , Y_1 prend la même valeur pour un système donné x_i des X_i dont dépend Y_1 . Ceci implique les bases minimum dont dépendent Y_1 et Y_2 soient disjointes. Réciproquement, si les plus petites bases dont dépendent Y_1 et Y_2 sont disjointes quelles que soient les valeurs x_i des X_i dont dépend Y_1 la valeur y_2 de Y_2 sera déterminée de façon unique par la donnée x_j des X_j dont dépend Y_2 ; et Y_2 sera fortement indépendant de Y_1 : on dira que Y_1 et Y_2 sont fortement indépendants.

OPÉRATIONS SUR LES VARIÉTÉS APPARTENANT A UN MÊME CHAMP

I - CORRELATION ENTRE DEUX VARIABLES -

Si nous considérons Y stochastiquement lié à X , à tout x nous

avons fait correspondre pour y un ensemble M_x de valeurs possibles que nous appellerons ensemble des échantillons de Y pour x et nous dirons que cet ensemble est corréllé à x ce que nous noterons $M_x = \gamma(x)$.

Si X peut prendre des valeurs dans un ensemble A , Y peut prendre ses valeurs dans un ensemble M qui est la réunion des M_x pour toutes les valeurs de $x \in A$, ensemble que nous noterons $\gamma(A) = \bigcup_{x \in A} M_x$ nous dirons que (A) est corréllé à A . La relation entre les ensembles A de valeurs de X et les ensembles $\gamma(A)$ de valeur de Y s'appellera la corrélation de X à Y .

Si nous nous disons que les variables X et Y sont corréllées. Si de plus X est stochastiquement lié à Y à tout y il correspond un ensemble de M_y de valeurs de x telles que l'on puisse avoir $X = x$; $Y = y$ ensemble que nous appellerons ensemble estimateur de x pour y j'aurais donc une corrélation de Y à X que je noterais γ^{-1} si je note γ la corrélation de X en Y . Si $y \in M_x$ je pourrais avoir simultanément $X = x$ $Y = y$ donc $x \in M_y$ ce qui montre que

$$x \in \gamma^{-1}(\gamma(x))$$

de façon plus générale si $x \in A$ il existe $y \in \gamma(A)$ tel que l'on puisse avoir $X = x$ $Y = y$ donc

$$A \subset \gamma^{-1}(\gamma(A))$$

de la même façon si $y \in B$ il existe $x \in \gamma^{-1}(B)$ tel que l'on puisse avoir $X = x$ $Y = y$ donc

$$B \subset \gamma(\gamma^{-1}(B))$$

Etudions maintenant les propriétés des corréllations ; on a évidemment

$$\begin{aligned} \gamma(\bigcup A_i) &= \bigcup \gamma(A_i) \\ A_1 \subset A_2 &\Rightarrow \gamma(A_1) \subset \gamma(A_2) \end{aligned}$$

Pour étudier $\gamma(\bigcap A_i)$ posons $\bigcap A_i = B$ j'aurais $B \subset A_i$ quel que soit i par conséquent $\gamma(B) \subset \gamma(A_i)$ quel que soit i . Donc $\gamma(B) \subset \bigcap \gamma(A_i)$ et par suite

$$\gamma(\bigcap A_i) \subset \bigcap \gamma(A_i)$$

Si j'appelle L l'ensemble de toutes les valeurs possibles pour X , M l'ensemble de toutes les valeurs possibles pour Y j'ai : $\gamma(L) = M$. De $L = A \cup CA$ je tire $\gamma(A) \cup \gamma(CA) = M$ donc

$$C \gamma(CA) \subset \gamma(A)$$

Si j'ai l'égalité $C \gamma(CA) = \gamma(A)$ quel que soit A cela prouve que je ne

peux pas trouver $X_1 \in A$ $X_2 \in CA$ tels qu'il leur corresponde une même valeur pour y ; donc A étant quelconque, à tout y il ne correspond qu'un x et x est une fonction certaine de y . Réciproquement si x est une fonction certaine de y , à deux x différents correspondent deux y différents, $\gamma(A)$ et $\gamma(CA)$ ne pourront avoir de points communs donc $\gamma(CA) \subset C \gamma(A)$ $\gamma(A) \subset C \gamma(CA)$ et en vertu de la relation ci-dessus qui est encore vraie

$$C \gamma(CA) = \gamma(A)$$

Cette relation entraîne en posant $CA_1 = B_1$ et en portant dans la formule

$$\gamma(\bigcup CA_1) = \bigcup \gamma(CA_1)$$

on a

$$C \gamma(\bigcap A_1) = \bigcup C \gamma(A_1)$$

$$\gamma(\bigcap A_1) = \bigcap \gamma(A_1)$$

réciproquement si quels que soient les A_1 on a :

$$\gamma(\bigcap A_1) = \bigcap \gamma(A_1)$$

en prenant A et CA pour A_1 on a :

$$\gamma(\emptyset) = \gamma(A \cap CA) = \gamma(A) \cap \gamma(CA)$$

si de plus

$$\gamma(\emptyset) = \emptyset$$

On a

$$\gamma(A) \subset C \gamma(CA)$$

donc

$$\gamma(A) = C \gamma(CA)$$

Il en résulte trois conditions :

$$C \gamma(A) = \gamma(CA) \text{ quel que soit } A ;$$

$$\bigcap \gamma(A_1) = \gamma(\bigcap A_1) \text{ quels que soient les } A_1 ; \gamma(\emptyset) = \emptyset \supset$$

X est fonction de Y ;
sont équivalentes.

Si Y est faiblement indépendant d'une ou plusieurs variables appartenant à la plus petite base dont dépende S nous dirons que Y leur est corréllé au sens large.

Si on cherche à définir une corrélation entre une variable Y et

une variable X qui lui corréllée au sens large on a un ensemble des échantillons de Y qui est identique à M ; d'où $\gamma(A) = M$ comme quand Y est strictement indépendant de X . Réciproquement si étant donné Y et une variable X (ou un ensemble X_1 de variables) appartenant à la plus petite base dont dépende Y , la corrélation de Y à X vérifie la relation $\gamma(x) = M$, Y est faiblement indépendant de X . En effet pour tout x il est possible de choisir $Y = y \in M$.

Comme précédemment pour simplifier l'écriture et le langage, nous remplacerons un ensemble de variables indépendantes par une variable unique de sorte qu'une répartition de la plus petite base dont dépende Y en 2 ensembles disjoints pourra s'écrire X_1, X_2 et l'on aura $Y = f(X_1, X_2)$ soit γ la corrélation de Y à X_1 nous pourrons écrire que $\gamma(x_1)$ est l'ensemble de toutes les valeurs $Y = f(x_1, x_2)$ ou x_1 est fixe et x_2 prend toutes les valeurs de l'ensemble des valeurs possibles pour X_2 .

Nous avons donc une application de l'ensemble L_2 sur un sous-ensemble de l'ensemble M . Y est indépendant de X_1 si quel que soit X_1 cette application est une application de L_2 sur l'ensemble lui-même.

Supposons maintenant qu'au lieu de connaître y pour tout élément de L_2 nous ne connaissions y que pour les éléments d'un ensemble M de L_2 qui sera dit valeur approchée de L_2 .

Pour tout x_1 nous pourrons définir l'ensemble des valeurs de $f(x_1, x_2)$, pour x_2 (A que nous appellerons $\gamma_M(x_1)$). Nous pourrons définir comme précédemment.

$$\gamma_M(A) = \bigcup_{x_1 \in A} \gamma_M(x_1)$$

γ_M sera dite corrélation approchée à la précision M de Y à X_1 .

Les théorèmes

$$A_1 \subset A_2 \implies \gamma_M(A_1) \subset \gamma_M(A_2)$$

$$\gamma_M(\bigcup A_i) = \bigcup \gamma_M(A_i)$$

$$\gamma_M(\bigcap A_i) \subset \bigcap \gamma_M(A_i)$$

$$C \gamma_M(CA \subset \gamma_M(A))$$

se démontreront de la même façon que les théorèmes correspondants pour γ . De plus si je considère différentes corrélations approchées à M_i près et les ensembles $\gamma_{M_i}(A)$ correspondants à un même ensemble

A dans ces différentes corrélations j'ai les mêmes ensembles que ceux que j'obtiendrai dans la corrélation approchée à A près de Y à X_2 pour les ensembles M_i j'aurais donc

$$M_1 \subset M_2 \implies \gamma_{M_1}(A) \subset \gamma_{M_2}(A)$$

$$\gamma_{\bigcup M_i}(A) = \bigcup \gamma_{M_i}(A)$$

$$\gamma_{\bigcap M_i}(A) \subset \bigcap \gamma_{M_i}(A)$$

$$C \gamma_M(CA) \subset \gamma_M(A)$$

Si quel que soit M $\gamma_M(A) = \gamma_M(B)$ quels que soient A et B, Y est strictement indépendant de X_1 . En effet si cette propriété a lieu elle entraîne $\gamma_A(M) = \gamma_B(M)$ quel que soit M A et B donc $f(x_1, x_2) = f(x'_1, x_2)$ quels que soient x_1, x'_1 et x_2 . On peut se fixer arbitrairement y et x'_1 , en effet y étant donné il existe un système de valeur x_1, x_2 auquel correspond $y = f(x_1, x_2)$ mais $f(x_1, x_2) = f(x'_1, x_2)$ quel que soit x'_1, x_2 étant fixé $f(x'_1, x_2)$ ne dépend pas de x'_1 .

Considérons maintenant 2 variables quelconques Y_1 et Y_2 appartenant à un même champ. Si elles peuvent être définies en fonction de 2 systèmes de bases disjointes B_1 et B_2 Y_1 et Y_2 sont indépendants. Sinon Y_1 et Y_2 appartiendront à $B = B_1 \cup B_2$. Soit A_1 et A_2 les plus petites bases de δ dont dépendent Y_1 et Y_2 . $A_1 \cup A_2$ peut être vide comme dans le cas précédent et Y_1 et Y_2 seront indépendants. Dans le cas contraire on posera :

$$D_1 = A_1 \cap C A_2$$

$$D_2 = A_2 \cap C A_1$$

$$D_3 = A_1 \cap A_2$$

Y_1 dépendra de D_1 et D_3 ; Y_2 dépendra de D_2 et D_3 .

Soit γ_1 la corrélation de D_3 à Y_1 et

γ_2 la corrélation de D_3 à Y_2

Si l'on se fixe la valeur y_1 de Y_1 il sera possible de se fixer les valeurs d_1 de D_1 et d_3 de D_3 à condition que $d_3 \in \gamma_1^{-1}(y_1)$ (d_1 étant alors déterminé par d_3). Mais d_3 étant fixé, Y_2 peut être alors choisi dans $\gamma_2(d_3)$ à condition de prendre convenablement d_2 . Puisque l'on peut se fixer simultanément d_1 et d_2 on pourra choisir Y_2 arbitrairement dans l'ensemble défini par

$$\gamma_2(\gamma_1^{-1}(y_1))$$

On peut donc définir une corrélation de Y_1 à Y_2 par :

$$r(y_1) = \gamma_2(\gamma_1^{-1}(y_1))$$

et par :

$$r(A) = \bigcup_{y_1 \in A} \gamma_2(\gamma_1^{-1}(y_1))$$

de la même façon, on peut définir la corrélation de Y_2 à Y_1 par :

$$r^{-1}(B) = \bigcup_{y_2 \in B} \gamma_1(\gamma_2^{-1}(y_2))$$

II - SUPREMUM ET INFIMUM D'UN ENSEMBLE DE VARIABLES :

Si nous considérons un ensemble G de variables prenant leurs valeurs dans un même ensemble E ordonné, nous pouvons définir une relation d'ordre sur G en posant que $Y_1 \leq Y_2$ si pour tout système de valeurs x_1 attribuées aux X_1 on a : $Y_1 \leq Y_2$. La relation ainsi définie est bien une relation d'ordre ; en effet :

1/ Si l'on a, à la fois $Y_1 \leq Y_2$

$$Y_2 \leq Y_1$$

c'est que pour tout système de valeurs x_1 attribuées aux X_1 on a, à la fois, $y_1 \leq y_2$ et $y_2 \leq y_1$ donc $y_1 = y_2$ et par suite, les deux variables sont identiques :

$$Y_1 = Y_2$$

2/ - Si l'on a, à la fois, $Y_1 < Y_2$ et $Y_2 < Y_3$, c'est que pour tout système de valeurs x_1 attribuées aux X_1 :

$$y_1 < y_2 \quad y_2 < y_3 \quad \text{donc : } y_1 < y_3 ; \text{ et par suite on a}$$

$$Y_1 < Y_3$$

Si E est un treillis, l'ensemble des variables qui appartiennent à un même système et qui prennent leur valeur dans E est un treillis. En effet, étant donné un sous-ensemble G de ces variables, on peut toujours définir un supremum et un infimum des variables de G . On posera :

$Z_1 = \text{Sup}(Y_j)$ et la variable qui pour tout système de valeurs x_1 attribuées aux X_1 prend la valeur $z_1 = \text{Sup}(y_j)$.

$Z_2 = \text{Inf}(Y_j)$ et la variable qui pour tout système de valeurs x_1 attribuées aux X_1 prend la valeur $z_2 = \text{Inf}(y_j)$.

En effet, toute variable Z qui est supérieure à tous les y_j a, pour

un système de valeurs x_i attribuées aux X_i , une valeur z supérieure ou égale à tous les y_j donc supérieure ou égale à : $\text{Sup}(y_j)$ et par conséquent on a : $z \geq z_1$ quels que soient les x_i . Donc $Z \geq Z_1$.

De même, si Z est inférieure ou égale à tous les Y_j , elle a pour tout système de valeurs x_i attribuées aux X_i , une valeur z inférieure ou égale à tous les y_j ; donc inférieure ou égale à $\text{Inf}(y_j)$ et par conséquent on a : $z \leq z_2$ quels que soient les x_i ; donc $Z \leq Z_2$.

REMARQUES -

1/ - Etant données les variables Y_j quelconques, il n'y a généralement pas de relations d'ordre entre elles, car l'ensemble des variables appartenant à un même système n'est pas totalement ordonné. Par suite, $\text{Sup}(Y_j)$ et $\text{Inf}(Y_j)$ ne sont généralement pas égaux à l'un des Y_j .

2/ - Quelles que soient les variables Y_j , on a la relation :

$\text{Inf}(Y_j) \leq Y_k \leq \text{Sup } Y_j$; Y_k étant l'un quelconque des Y_j puisque quels que soient les x_i on a :

$$\text{Inf}(y_j) \leq y_k \leq \text{Sup}(y_j)$$

Si E est un treillis distributif les variables aléatoires appartenant à un même système et prenant leur valeur dans E forment un treillis distributif.

En effet, on a quels que soient les x_i :

$$\text{Sup}[y_j, \text{Inf}(y_k, y_1)] = \text{Inf}[\text{Sup}(y_j, y_k), \text{Sup}(y_j, y_1)]$$

Donc :

$$\text{Sup}[Y_j, \text{Inf}(Y_k, Y_1)] = \text{Inf}[\text{Sup}(Y_j, Y_k), \text{Sup}(Y_j, Y_1)]$$

CHAPITRE II

ÉTUDE DES ENSEMBLES CLASSIFIÉS

Nous avons vu que l'étude des variables aléatoires déterminées par un même évènement se ramène à celle des fonctions y d'une variable x qui décrit un ensemble E .

Quand un statisticien étudie une variable X à valeurs dans F au moyen d'un système de valeurs de Y il les répartit en un nombre fini ou une infinité dénombrable de classes qui sont des ensembles disjoints Q_i d'éléments de F dont la réunion recouvre F .

Si nous considérons les ensembles P_i d'éléments de E dont l'image est élément de Q_i deux tests ensembles sont disjoints car s'ils avaient un point commun son image appartiendrait à $2Q_i$ ce qui est impossible puisque les Q_i sont disjoints.

Comme il y a une correspondance bi-univoque entre les Q_i et les P_i il y a un nombre fini ou une infinité dénombrable de P_i - les P_i forment donc une répartition en classes des éléments de E . Nous savons donc faire une répartition en classe π des éléments de E . Si nous étudions plusieurs variables $Y_1 \dots Y_n \dots$ ou si nous étudions de plusieurs façons différentes la même variable Y nous obtiendrons plusieurs répartitions en classes $\pi_1 \dots \pi_n$ de l'ensemble E .

Nous nous proposons d'étudier la structure d'un ensemble E sur lequel nous savons déterminer une ou plusieurs répartitions en classe.

THEOREME II - 1

Les répartitions en classes d'un même ensemble forment un treillis.

Nous pouvons définir une relation sur l'ensemble des répartitions en classes d'un ensemble E par la convention suivante : " π est plus fin que π' si toute classe de π est une union finie ou dénombrable de classes de π' ".

Nous pouvons également définir cette relation par une autre convention : " π est plus fin que π' si toute classe de π est incluse dans une classe de π' ". Ces deux conventions sont équivalentes.

1/ - Si toute classe de π est une réunion finie ou dénombrable de classes de π' toute classe P_i de π est incluse dans une classe de π' . Soit x un point de P_i la classe de π' dont x est élément, $P_i \subset P'_j$ car si P_i ne faisait pas partie de la famille de classes de π dont la réunion est P'_j x ne serait pas élément de P'_j .

2/ - Si toute classe de π est incluse dans une classe de π' toute classe de π' est une réunion finie ou dénombrable de classes de π . Toute classe P'_j de π' contient au moins une classe de π car si x est élément de P'_j la classe de P_x de π dont x est élément, est incluse dans une classe de π' qui est confondue avec P'_j puisque son intersection avec P'_j n'est pas vide. $P_x \subset P'_j$

La réunion des P_x pour tous les éléments de P'_j contient les x donc :

$$\bigcup_{x \in P'_j} P_x \supset P'_j$$

Mais

$$\bigcup_{x \in P'_j} P_x \subset P'_j$$

Donc :

$$\bigcup_{x \in P'_j} P_x = P'_j$$

La réunion est finie ou dénombrable car sinon l'ensemble des classes de π aurait une puissance supérieure ou dénombrable.

Ces conventions définissent bien une relation d'ordre car elles sont transitives. (En effet toute union finie ou dénombrable d'unions finies ou dénombrables est encore une union finie ou dénombrable) et réflexives. (En effet soit P_i une classe de π elle est incluse dans une classe P_k de π donc :

$$P_i \subset P_j \subset P_k$$

$$P_i \cap P_k \neq \emptyset$$

donc $P_1 = P_k \quad R_k \subset P_j \subset P_1 \quad P_1 = P_j$

les classes de π et π' sont les mêmes).

Etant données deux répartitions en classes π_1 et π_2 de E il existe une classe π' formée par les intersections de toutes les classes de π_1 avec toutes les classes de π_2 . π' est plus fin que π_1 et que π_2 puisque toute classe de π' est incluse dans une classe de π_1 et dans une classe de π_2 .

Toute répartition en classes π plus fine que π_1 et que π_2 est plus fine que π' ; en effet toute classe de π_1 et toute classe de π_2 est une union finie ou dénombrable de classes de π . L'intersection d'une classe de π_1 et d'une classe de π_2 est l'intersection de réunions dénombrables donc la réunion dénombrable d'intersections d'éléments de π deux à deux, ces intersections sont soit vides, soit confondues avec un élément de π donc tout élément de π' est une union finie ou dénombrable d'éléments de π . Nous avons donc montré l'existence de la moins fine des répartitions plus fines que π_1 et de π_2 . π' s'appellera le supremum de π_1 et de π_2 . La condition nécessaire et suffisante pour que π_1 soit le supremum de π_1 et de π_2 est que π_1 soit plus fin que π_2 .

Il existe des répartitions en classes moins fines que π_1 et que π_2 . La répartition triviale (tous les éléments dans une seule classe) est moins fine que toute autre. Supposons qu'il en existe plusieurs; elles forment un ensemble G. Si π_1'' et π_2'' sont des éléments de G, le supremum π''' de π_2'' est aussi un élément de G. En effet si $A_1 \in \pi_1''$ et $A_2 \in \pi_2''$ sont des unions finies ou dénombrables de classes de π_1 et de π_2 . $A_1 \cap A_2$ en est également une donc tous les éléments de π''' sont des répartitions en classe de π_1 et de π_2 . Je dirai que π_p est un élément maximal de G s'il n'existe aucun π_q plus fin que π_p et appartenant à G.

S'il existe un élément maximal, il n'en existe qu'un. En effet s'il en existait deux π_p et π_q , $\text{Sup}(\pi_p, \pi_q)$ ferait partie de F et serait plus fin que π_p et π_q ce qui est contraire à l'hypothèse car il ne peut être confondu avec les deux que si π_p et π_q sont confondus.

Nous allons montrer l'existence d'au moins un élément maximal. S'il n'y en avait pas c'est que pour tout élément π_p on pourrait trouver un élément π_q plus fin que π_p il y aurait donc une suite ordonnée π_i de répartitions en classes; considérons les classes P_{ix} des π_i qui contiennent le point x: ce sont les réunions finies ou dénombrables de classes de π_1 .

Si π_1 est plus fin que π_{jx} P_j est union finie ou dénombrable de classes de π_1 parmi lesquelles se trouve P_{ix} on a donc $P_{ix} \subset P_{jx}$. Les P_i sont des ordonnés par rapport à l'inclusion.

Considérons l'intersection R_x de tous les P_{1x} elle n'est pas vide car elle contient au moins la classe de P_1 dont x est élément. Considérons R_x et R'_x correspondant à deux éléments distincts x et x' . Ou bien pour tout π_1 R_x et R'_x sont inclus dans un même P_1 et sont donc définis par la même intersection donc confondus.

Ou bien il existe un π_1 pour lequel les classes qui contiennent R_x et R'_x sont disjointes et alors $R_x \cap R'_x = \emptyset$.

Considérons l'ensemble des R_x distincts. C'est une famille de sous-ensembles disjoints de E dont la réunion recouvre E . De plus à tout R_x je peux faire correspondre une classe de π_1 qui lui est incluse. Toutes ces classes sont distinctes puisque disjointes comme contenues dans des ensembles disjoints. Donc la puissance de l'ensemble des R_x est inférieure ou égale à la puissance de l'ensemble des classes de π_1 . L'ensemble des R_x est donc fini ou dénombrable.

Nous avons donc ainsi défini une répartition de E en classes. Cette répartition en classes $\bar{\pi}$ est moins fine que P_1 . En effet la classe C_x de π_1 dont x est élément, est incluse dans toutes les classes P_{1x} car si C_x ne faisait pas partie de la famille de classes π_1 dont la réunion définit P_{1x} , P_{1x} ne contiendrait pas le point x . Donc C_x est inclus dans l'intersection R_x des P_{1x} $C_x \subset R_x$.

De la même façon on démontrerait que $\bar{\pi}$ est moins fin que π_2 donc $\bar{\pi}$ est un élément de G .

$\bar{\pi}$ est plus fin que tout π_1 . En effet toute classe R_x est contenue dans la classe P_{1x} de π_1 .

Il existe donc un élément $\bar{\pi}$ de G plus fin que tous les π_1 . Puisque $\bar{\pi}$ n'est pas maximal il existe $\bar{\pi}_2$ plus fin que $\bar{\pi}$ et ainsi de suite on peut donc définir une nouvelle suite ordonnée $\bar{\pi}_1$ d'éléments de G à partir de laquelle on peut définir $\bar{\pi}$ et continuer ainsi transfiniment. L'ensemble T des éléments π obtenu par cette méthode est formée d'éléments de G . Il a donc une puissance inférieure ou égale à celle de G . S'il n'a pas de dernier élément sa puissance est supérieure à celle de tout ordinal transfini ce qui est contradictoire.

Nous avons donc montré l'existence d'un élément maximal π'' de G . Il n'existe pas d'éléments de G plus fin que π'' donc tout élément de G est moins fin que π'' car s'il en existait un, π , ne serait pas confondu avec π et π serait un élément de G plus fin que π'' .

π'' s'appellera l'Infremum de π_1 et π_2

$$\pi'' = \text{Inf} (\pi_1 \pi_2)$$

La condition nécessaire et suffisante pour que :

$\pi_1 = \text{Inf}(\pi_1, \pi_2)$ est que π_1 soit moins fin que π_2 .

CLASSIFICATIONS

Nous nous intéresserons aux familles ordonnées et dénombrables de répartition en classes de E. Nous appellerons classification de E une telle famille et nous dirons qu'elle détermine sur E une structure d'ensemble classifié.

Soit une partie M de E formée par une réunion finie ou dénombrable de classe de π_1 . π_1 et la famille des π_j plus fins que π_1 déterminent sur M des répartitions en classes π'_j en effet M est union finie ou dénombrable de classes des π_j que l'on peut considérer comme des classes des π'_j . D'autre part la relation d'ordre entre les π'_j est la même que la relation d'ordre entre les π_j , en effet si tout P_k élément de π_j plus fin que π_j est inclus dans P_j , élément de π_j ; mais P_k est élément de π'_k et P_j de π'_j donc toute classe de π'_k est incluse dans une classe de π'_j .

M sera donc également pourvue d'une structure d'ensemble classifié. Nous dirons que M est une partie classifiée de E si M est union finie ou dénombrable de classes appartenant à l'un quelconque des π . Plus généralement nous appellerons parties C - classifiées tout élément de la Σ - algèbre de Boole engendrée par les classes appartenant aux répartitions en classes de la classification C et nous désignerons par S(C) cette Σ - algèbre de Boole. Quand il n'y aura pas ambiguïté nous dirons partie classifiée et S. Si nous considérons les classifications C_1 et C_2 déterminées :

C_1 par les répartitions $\pi_1 \pi_2 \dots \pi_i \dots$

C_2 par les répartitions $\pi_2 \dots \pi_i \dots$

Nous avons : $S(C_1) \equiv S(C_2)$, en effet :

Considérons une union finie ou dénombrable de P_{1j} ($1 \leq i$) appartenant aux répartitions $\pi_1, \pi_2 \dots \pi_i$ c'est une union finie ou dénombrable de P_{1j} et de P_{ij} ($2 \leq i$).

En remplaçant les P_{1j} par des unions finies ou dénombrables de

P_{2j} nous avons une réunion finie ou dénombrable de P_{1j} ($2 \leq i$).

Réciproquement une union finie ou dénombrable de P_{1j} ($i \leq 2$) est une union finie ou dénombrable de P_{1j} ($i \leq 1$).

D'une façon plus générale on peut supprimer de la famille π_1 un ensemble fini de π_1 sans modifier $S(C)$. De même on peut rajouter des répartitions moins fines que la première répartition.

Les propriétés les plus intéressantes des classifications ne concernent que les parties classifiées qui sont les éléments de $S(C)$. Nous sommes amenés à considérer comme équivalentes les classifications C_j qui engendrent la même $\sum$ algèbre $S(C_j)$. Nous dirons que les C_j engendrent la même classification.

Quand nous nous intéresserons à la famille des π_1 au moyen de laquelle nous avons défini la classification C nous la désignerons sous le nom de base de C de sorte que compte tenu de la convention précédente nous pouvons dire que différentes bases engendrent la même classification.

Soit une famille dénombrable de parties classifiées disjointes et recouvrant E . Elle forme une répartition en classe $\pi(E)$ si π n'appartient pas à la base qui nous a servi à définir C nous pouvons trouver une autre base de C dont π soit élément.

Il suffit par exemple de prendre :

$$\pi, \text{Sup}(\pi_1, \pi), \dots, \text{Sup}(\pi_1, \pi), \dots$$

qui forme une famille dénombrable de répartitions en classes de plus en plus fines.

La classification C' définie par cette base est équivalente à C .

1/ - Toutes les classes de cette nouvelle classification sont des intersections de classes de π qui sont des parties classifiées de $S(C)$ avec des classes de π_1 qui sont aussi des éléments de $S(C)$ donc la $\sum$ -algèbre de Boole $S(C')$ est engendrée par une famille d'éléments de $S(C)$: $S(C') \subset S(C)$.

2/ - Toute classe de C est réunion des intersections de cette classe avec les éléments de π or, toutes ces intersections, sont éléments de $S(C')$ donc toute classe de C est élément de $S(C')$. $S(C)$ est engendré par une famille d'éléments de $S(C')$ et par conséquent $S(C) \subset S(C')$ donc : $S(C) = S(C')$.

En résumé toute famille dénombrable de parties classifiées d'une classification C de E , disjointe et recouvrant E est une répartition en classes appartenant à une base d'une classification de E équivalente à C . Plus brièvement nous dirons que c'est une répartition associée à C .

THEOREME II - 2 :

L'ensemble des classifications de E est un treillis.

Nous pouvons définir une relation d'ordre sur l'ensemble des classifications de E en posant que C' est plus fin que C'' si $S(C'') \subset S(C')$ et réciproquement. Nous avons bien une relation d'ordre.

1/ - Si C'' est plus fin que C''' $S(C''') \subset S(C'') \subset S(C')$

Donc, C' est plus fin que C'''

2/ - Si C' est plus fin que C'' et C'' plus fin que C' , on a :

$$S(C'') \subset S(C')$$

$$S(C') \subset S(C'')$$

$$S(C') = S(C'')$$

et les classifications $S(C')$ et $S(C'')$ sont équivalentes. Si C' et C'' sont définis par deux familles de répartitions π_i' et π_i'' nous pouvons poser :

$$\bar{\pi} = \text{Sup} (\pi_i', \pi_i'')$$

$$\underline{\pi}_i = \text{Inf} (\pi_i', \pi_i'')$$

La classification $\bar{C}$ définie par les $\bar{\pi}$ est le Supremum de C' et C'' . En effet, toute union finie ou dénombrable de classe de π_i' est union finie ou dénombrable de classe de π_i donc appartient à $S(\bar{C})$. Il en résulte que tout élément de $S(\bar{C}')$ appartient à $S(\bar{C})$; $S(C') \subset S(\bar{C})$

De même : $S(C'') \subset S(\bar{C})$; donc :

$$S(C') \cup S(C'') \subset S(\bar{C})$$

D'autre part, supposons qu'une classification C soit plus fine que C' et C'' . Elle contient toutes les classes de C' et C'' comme parties classifiées ; donc toutes les intersections des classes de π_i' avec celles de π_i'' c'est-à-dire toutes les classes de $\bar{\pi}_i$ et par suite, tous les éléments de $S(\bar{C})$. Elle est donc plus fine que $\bar{C}$; $\bar{C}$ est donc la moins fine des classifications plus fines que C' et C'' .

Nous avons donc défini le supremum de C' et C'' : de la même façon nous verrions que $\underline{C}$ est la plus fine des classifications moins fines que C' et C'' et par conséquent définit l'infimum de C' et C'' .

REMARQUES :

1/ - La démonstration précédente peut s'étendre par récurrence et on peut définir le supremum et l'infimum d'une famille finie de classifications. Mais on ne peut pas toujours définir le sup. d'une famille infinie (même dénombrable) de classifications. On peut toujours poser :

$$\bar{C}_n = \text{Sup } (C_i) \\ (i \leq n)$$

Les $\bar{C}_n$ forment une famille de classifications de plus en plus fines. Elles ne définissent un supremum que si les $\bar{C}_n$ sont équivalents à partir d'un certain rang.

2/ - La famille des classes d'une classification qui contiennent un point donné est ordonnée par rapport à l'inclusion. Soit P_i la classe de π_i qui contient x , P_j la classe de π_j qui contient x - Si π_i est plus fin que π_j toute classe de π_i est incluse dans une classe de π_j donc P_i est incluse dans une classe P'_j de π_j mais comme il n'y a qu'une classe de π_j qui contient x $P'_j = P_j$ $P_i \subset P_j$.

Si nous désignons par P_x l'intersection de toutes les classes de C qui contiennent x ; les R_x forment une famille d'ensembles disjoints qui recouvrent E . Mais cette famille n'est une répartition en classe que si elle est dénombrable. Dans ce cas cette répartition est plus fine que toutes les répartitions qui définissent la classification c'est-à-dire que la base de la classification possède un élément plus fin que tous les autres (ce qui sera réalisé si toutes les répartitions sont identiques à partir d'un certain rang). Dans le cas contraire la famille R_x n'est pas dénombrable mais a la puissance du continu.

Un ensemble E sera dit séparable par rapport à C si étant donnés deux éléments x et y de E il est toujours possible de trouver deux classes P_x et P_y disjointes telles que :

$$x \in P_x$$

$$y \in P_y$$

donc : $x \notin P_y$

$$y \notin P_x$$

Si E est séparable par rapport à C l'intersection des classes de C qui contiennent x ne contient que x car si elles contenaient y il ne serait pas possible de trouver P_x tel que : $y \notin P_x$.

Si E est séparable par rapport à C il a la puissance du continu car l'ensemble des R_x qui coïncide avec l'ensemble des éléments de E a la puissance du continu.

ORDINATION D'UN ENSEMBLE CLASSIFIE :

Soit un ensemble E pourvu d'une classification C définie au moyen des répartitions en classes : $\pi_0, \pi_1, \dots, \pi_i$ la répartition π_0 ne comporte qu'une classe P_0 identique à E. La répartition π_1 comporte des classes que nous noterons P_{1j} .

La répartition π_2 comporte des classes que nous noterons P_{2jk} avec $P_{2jk} \subset P_{1j}$.

Nous définirons ainsi de proche en proche les classes $P_{ijk\dots uv}$ avec : $P_{ijk\dots uv} \subset P_{(i-1)jk\dots u}$. Nous pourrions simplifier et unifier la notation en complétant à partir du dernier indice avec des zéros (étant entendu que ces zéros sont inférieurs à tous les indices jk etc ... et en supprimant l'indice i qui est toujours égal au nombre d'indices non nuls qui le suivent.

Dans ces conditions deux classes quelconques s'écriront $P_{j_1 k_1 \dots}$... $P_{j_2 k_2 \dots}$ les indices étant tous nuls à partir d'un certain rang, jamais nuls avant.

Si $j_1 < j_2$ nous dirons que $P_{j_1 \dots} < P_{j_2 \dots}$.

Si $j_1 > j_2$ nous dirons que $P_{j_1 \dots} > P_{j_2 \dots}$.

Si $j_1 = j_2$ soit l le premier indice pour lequel on ait : $l_1 \neq l_2$, on dira que $P_{j_1 \dots l_1 \dots} < P_{j_1 \dots l_2 \dots}$.

Si $l_1 < l_2$ et que $P_{j_1 \dots l_1 \dots} > P_{j_2 \dots l_2 \dots}$ si $l_1 > l_2$.

La relation ainsi définie est manifestement totale car on a toujours $P_{j_1 \dots} \leq P_{j_2 \dots}$ sauf si $l_1 = l_2$ quel que soit l'indice l auquel cas les deux classés sont identiques.

Elle est transitive : supposons $P_{j_1 \dots} < P_{j_2 \dots}, P_{j_2 \dots} < P_{j_3 \dots}$.

On a pour premier indice $j_1 \leq j_2 \leq j_3$ donc : $j_1 \leq j_3$. Si les trois indices ne sont pas égaux on a donc : $j_1 < j_3$.

Donc : $P_{j_1} \dots < P_{j_3} \dots$

S'ils sont égaux, soit l_1 le premier indice à partir duquel les trois indices ne sont pas égaux :

$$l_1 \leq l_2 \quad l_2 \leq l_3 \quad \text{entraîne} \quad l_1 \leq l_3$$

mais l'égalité des trois indices étant exclue $l_1 < l_3$

donc : $P_{j_1} \dots l_1 \dots P_{j_3} \dots l_3$.

Nous avons donc défini une relation d'ordre sur l'ensemble des classes de la classification. Nous pouvons étendre cette relation d'ordre aux ensembles R_x . On pourra définir R_x par une suite infinie d'indices non nuls et définir la relation d'ordre par la même convention que précédemment. Cette relation d'ordre est stricte : si l'on n'a pas identiquement l'égalité des indices correspondant mais dans ce cas les R_x seraient définis par la même famille de classes incluses les unes dans les autres et coïncideraient. Si la classification sépare E les R_x coïncident avec les points de E et la relation d'ordre est étendue aux points de E .

CLASSIFICATION D'UN ENSEMBLE TOTALEMENT ORDONNE :

Si nous choisissons un ensemble fini ou dénombrable N d'éléments x_n de l'ensemble ordonné E , nous pouvons considérer que les x_n sont ordonnés par la relation d'ordre sur E . Nous pouvons définir une répartition en classes à partir des points x . Soit x un point de E , si $x = x_1$ élément de N , nous dirons que x appartient à la classe P . Si $x \notin E$ il est possible de définir une coupure sur N en plaçant dans l'ensemble A les $x_n < x$ et dans l'ensemble B les $x_n > x$. Si B n'est pas vide il a un premier élément x_i et nous dirons que x appartient à la classe P_i . Si B est vide nous dirons que x appartient à la dernière classe (ou $x \notin P_i$).

Tout x appartient à une classe et une seule, les classes sont en nombre fini ou en infinité dénombrable (comme N) nous avons donc défini une répartition en classes π_1 .

Remarque :

Les classes sont de la forme $[a < x \leq b]$. Soit E_i l'ensemble des points de E qui appartiennent à la classe P_i . Si nous choisissons un ensemble fini ou dénombrable N_i de points de P_i nous définissons sur E_i une répartition en classes qui nous détermine des classes P_{ij} . Si nous faisons de même sur tous les ensembles E_i nous aurons une répartition en classes π_2 de E qui est plus fine que E .

Si nous continuons ainsi, de proche en proche nous définissons des répartitions π_n formant une famille dénombrable de répartition de plus en plus fine. Nous avons ainsi défini une classification C de E. A partir de cette classification nous pouvons définir une relation d'ordre sur les classes de cette classification qui s'étendra aux éléments de E si cette classification sépare E. Nous allons comparer l'ordination initiale de E et celle que l'on peut déduire d'une telle classification.

Soient deux points x et y de E, il existe au moins une répartition en classes pour laquelle x et y appartiennent à une même classe. La répartition π_0 qui ne comporte qu'une seule classe en est une. Il existe des répartitions en classes π_j pour lesquelles x et y appartiennent à des classes différentes car sinon on ne pourrait pas trouver de classes P_x et P_y disjointes telles que $x \in P_x$, $y \in P_y$.

Définissons une coupure sur l'ensemble des indices en plaçant dans la première classe les i pour lesquels les classes de π_i qui contiennent x et y sont confondus et dans la deuxième classe les i pour lesquels les classes de π_i qui contiennent x et y sont disjointes. Soit m, m + 1 le saut défini par la coupure, soit P_m la classe de π_m qui contient x et y, $P_{(m+1),1}$, $P_{(m+1),1'}$ les classes de π_{m+1} qui contiennent x et y. Supposons $l < l'$ pour fixer les idées, donc : $l \leq l' - 1$.

Nous aurons : $x < x_l$ $x_{l'-1} < y$

Or, l'ordination des l étant la même que celle des x_l , $x_l \leq x_{l'-1}$;

Donc : $x \leq x_l$ $x_{l'-1} \leq y$ donc : $x < y$

L'ordination déduite de la classification est la même que l'ordination initiale. Nous dirons que la classification est compatible avec l'ordination de E. Si une classification est compatible avec la relation d'ordre sur E, toutes les classes sont de la forme : $a < x \leq b$.

Si les classes P_i et P_j contiennent x_i et x_j , $i < j$ sera équivalent à $x_i < x_j$. Si l'on a : $x_i < x_j$, on aura : $x_i' < x_j'$ quels que soient $x_i' \in P_i$ $x_j' \in P_j$.

THEOREME II - 3 :

Si deux classifications sont compatibles avec la relation d'ordre sur E et si elles séparent E, elles sont équivalentes.

Montrons que toute classe de l'une est une partie classifiée de l'autre. Soit $\alpha < x \leq \beta$ une classe K de C'. Soit $a_1 \leq x \leq b_1$ une classe de π_1 contenue dans K. Soit $c_1 < x \leq d_1$ la classe de π_1 qui contient β ;

nous appellerons R_i la réunion de toutes les classes $a_i < x \leq b_i$, si la classe $c_i < x \leq d_i$. Nous poserons :

$$R = \bigcup R_i$$

$$S = \bigcap S_i$$

Tout point x_0 de K différent de β appartient à un R_i . En effet, si E est séparé par R il existe un P_{ix} et un $P_{i\alpha}$ disjoints, tels que $x \in P_{ix}$, $\alpha \in P_{i\alpha}$. De même, il existe un P_{jx} et un $P_{j\beta}$ disjoints ; nous prendrons celle des deux classes qui est contenue dans l'autre si elles sont distinctes. Soit P_{kx} cette classe ; elle est de la forme $a_k < x \leq b_k$ avec $\alpha \leq a_k$ et $b_k < \beta$; donc, elle est incluse dans l'ensemble K . Donc tout point de K à l'exception peut-être de β appartient à R mais β appartient à tous les S_i donc c'est le seul point de S , par conséquent K est contenu dans $R \cup S$.

Tout point qui n'appartient pas à K n'appartient à aucun des R_i donc il ne peut appartenir à R . Comme S ne contient que β il ne peut non plus lui appartenir ; donc :

$$K = R \cup S$$

Comme R et S sont des parties classifiées K l'est également.

CLASSIFICATION D'UN ENSEMBLE PARTIELLEMENT ORDONNE :

Soit E un ensemble partiellement ordonné. Nous désignerons par $S(x_3, \dots)$ l'ensemble des éléments de E qui sont inférieurs ou égaux à x , par $S(\cdot ; x)$ le complémentaire de cet ensemble, par $S(x ; y)$ l'intersection $S(x ; \cdot) \cap S(\cdot ; x)$, par $S(\dots x_i, \dots, y_j \dots)$ l'intersection de tous les $S(x_i ; \cdot)$ et de tous les $S(\cdot ; y_j)$.

Etant donnée une famille finie F d'éléments de E si l'on considère l'ensemble $\sum$ des $S(\dots x_i \dots ; \dots y_j \dots)$ pour tous les systèmes de valeurs possibles de x_i et y_j choisis dans F il est possible d'en déduire une répartition en classes où toutes les classes seront de la forme $S(u_K ; \dots v_{K_j})$ avec : $v_{K_j} < u_K$. On peut poser : $u_K = \inf(x_1 \dots)$

$$v_{K_j} = \inf(u_K ; y_j)$$

Tout point Z de E appartient à un $S(\dots x_i \dots ; \dots y_j \dots)$ non vide (ne serait-ce que celui pour lesquels les x_i sont les éléments de F supérieurs ou égaux à Z et les y_j ceux qui ne le sont pas).

Si nous désignons par $\sum_Z$ tout ensemble $S(\dots u_K \dots ; \dots v_{K_j} \dots)$ qui contient Z et par S_Z l'intersection de tous les $\sum_Z$. S_Z est encore de la

même forme ; les ensembles S_Z sont en nombre fini puisque obtenus par des intersections d'un nombre fini d'ensembles ; ils recouvrent E puisque tout Z appartient au moins à un Σ_Z donc à un S_Z . De plus les S_Z distincts sont disjoints car s'ils avaient un point commun Z_0 ils seraient définis par les mêmes Σ_{Z_0} et seraient confondus.

Si nous donnons des familles F_i incluses les unes dans les autres nous avons des répartitions en classe de plus en plus fines de E. Nous avons donc déduit des familles F_i et de la relation d'ordre sur E une classification C de E dans laquelle toutes les classes sont des ensembles de la forme $S(u_k ; \dots v_{k_1})$

Si dans une classification de E toutes les classes sont de cette forme nous dirons qu'elle est compatible avec la relation d'ordre sur E.

THEOREME II - 4 :

Toutes les classifications qui sont compatibles avec la structure d'ordre de E et qui séparent E sont équivalentes. Il nous suffira de montrer que tout ensemble $M = S(u; \dots v_k \dots)$ est une partie classifiée pour n'importe quelle classification compatible avec la structure d'ordre sur E et qui sépare E.

Soit π_1 une répartition en classes de E nous appellerons :

A_{im} les classes qui n'ont aucun point commun avec M

B_{im} les classes qui ont au moins un point $\in$ M et au moins un point $\in$ d'un $S(v_k ; \dots)$

C_{ip} les classes qui sont incluses dans M.

D_{iq} toutes les autres classes.

Posons : $A_i = \bigcup A_{im}$

$B_i = \bigcup B_{im}$

$C_i = \bigcup C_{ip}$

$D_i = \bigcup D_{iq}$

Nous avons :

$$E = A_i \cup B_i \cup C_i \cup D_i$$

$$C_i \subset M \subset B_i \cup C_i \cup D_i$$

A_1, B_1, C_1, D_1 sont des parties classifiées de E . Nous avons :

$$C(A_1 \cup C_1) = B_1 \cup D_1$$

$$C \cup (A_1 \cup C_1) = \cap (B_1 \cup D_1)$$

et en posant : $\cup A_1 = A \quad \cup C_1 = C$

$$C(A \cup C) = \cap (B_1 \cup D_1)$$

Si nous posons : $\cap B_1 = B$ on a : $B \subset \cap (B_1 \cup D_1)$

et nous pouvons poser : $\cap (B_1 \cup D_1) = B \cup D$

D étant un ensemble disjoint de B . Il est également disjoint de A et de C puisque contenu dans $C(A \cup C)$. B est disjoint de A et C et A et C sont disjoints puisque :

$$A = \cup A_1$$

$$A_1 \subset CM$$

$$A \subset CM$$

$$C = \cup C_1$$

$$C_1 \subset M$$

$$C \subset M$$

On a donc : $E = A \cup B \cup C \cup D$

où A, B, C, D sont disjoints deux à deux. A, B, C sont des parties classifiées donc D aussi. Montrons que $A \cup B = CM$

$$M \cap (A \cup B) = (M \cap A) \cup (M \cap B)$$

Tout point de A est élément d'un A_1 inclus dans CM donc est élément de CM et par suite :

$$A \cap M = \emptyset$$

Supposons qu'il existe un point x de M qui soit point de B . C est un élément de tous les B_1 . Il existe donc :

$S(u' ; \dots v'_k \dots)$ tel que $x_0 \in S(u ; \dots v_k \dots) \quad x_0 \in S(u' ; \dots v'_k \dots)$.

$S(x_0, \dots v'_k \dots)$ n'est pas vide ni réduit à x car s'il était vide ou réduit à x_0 nous aurions : $x_0 \leq v'_k$

$$x_0 \notin S(u' ; \dots v'_k \dots)$$

$S(x_0 ; \dots v'_k \dots v_k \dots)$ n'est pas vide ni réduit à x_0 car pour qu'il le soit il faudrait qu'il en soit de même pour $S(x_0 ; \dots v_k \dots)$ ou que :

$$S(x_0 ; \dots v'_k \dots) \subset C S(x_0 ; \dots v_k \dots)$$

$$\text{Or, } C[S(x_0 ; \dots v_k \dots)] = S(. ; x_0) \cup [\cup S(v_k ; .)]$$

$$\text{et : } S(x_0 ; \dots v'_k \dots) \cap S(. ; x_0) = \emptyset - S(x_0 ; \dots v'_k \dots) \cap S(v_k ; .)$$

$$S(x_0 ; \dots v'_k \dots) \cap S(v_k ; .) \text{ impliquerait } x_0 \in S(u ; \dots v_k \dots)$$

Mais si $S(x_0 ; \dots v_k \dots)$ était vide cela prouverait que tout $x \leq x_0$ est $\leq v_k$ donc : $x_0 \leq v_k$ et $x_0 \notin S(u ; \dots v_k \dots)$.

Il existe donc un point $x_1 \in S(x_0 ; \dots v'_k \dots v_k \dots)$. Si C sépare E il est possible de trouver un $S(u'' ; \dots v''_1 \dots)$ qui contienne x_0 et tel que tous les $S(v''_1 ; .)$ contiennent x_1 . Désignons par N l'ensemble des points de E qui ne sont pas supérieurs à x_1 . $x_0 \notin N$; par N_1 l'intersection de N avec la classe de π_1 qui contient x_0 . L'intersection de tous les N_i est vide car sinon il y aurait au moins un point distinct de x_0 qui appartiendrait à toutes les classes qui contiennent x et C ne séparerait pas E . Il y a donc au moins un N qui est vide et la classe correspondante est contenue dans le complémentaire de N . Donc tous les points de cette classe sont supérieurs à x_1 ; comme x_1 n'est inférieur à aucun des v_k , aucun des points de cette classe n'est inférieur à un v_k . Par conséquent x_0 appartient à un C_1 et ne peut donc appartenir au B_1 correspondant. Il n'appartient donc pas à B , donc $B \cap M = \emptyset$; donc : $A \cup B = CM$

$$\text{et par suite : } M = C \cup D$$

or, C et D sont des parties classifiées donc M aussi.

CONCLUSION -

Les exemples de classification que nous venons de traiter montrent que la notion d'ensemble classifié est extrêmement générale et peut s'appliquer à de nombreux ensembles, par exemple l'ensemble des nombres réels qui est un ensemble totalement ordonné. Nous pouvons remarquer que si nous définissons sur cet ensemble une topologie par la famille d'ouverts formés par les ensembles $a < x < b$, tous les ouverts et tous les fermés de cette topologie sont des parties classifiées dans la classification qui sépare E et qui est compatible avec la structure d'ordre. (Je dis la classification puisque toutes les classifications jouissant de ces propriétés que nous pourrions définir sont équivalentes). Nous pouvons dire que l'ensemble des réels est un ensemble topologique classifié.

On peut penser que de nombreuses propriétés de l'ensemble des réels sont des propriétés générales des ensembles classifiés. En particulier, il semble que la théorie de la mesure et de l'intégration est une théorie qui se rattache non à la théorie des espaces topologiques mais à celle des ensembles classifiés dans la suite de ce travail, nous aurons besoin d'étendre la théorie de la mesure et de l'intégration à des

ensembles qui ne sont pas pourvus d'une topologie mais qui sont classifiés. S'il semble souhaitable de construire une théorie générale des ensembles classifiés qui me paraît devoir être extrêmement riche d'applications et de conséquences (j'ai d'ailleurs l'intention de revenir dans des travaux ultérieurs sur cette question), je me contenterai dans le présent travail de développer les points de cette théorie qui me seront utiles par la suite, en me limitant à signaler au passage quelques sujets de recherche qui me semblent particulièrement intéressants.

CORRELATIONS ENTRE ENSEMBLES CLASSIFIES :

Soit deux ensembles E_1 et E_2 , soit une corrélation γ entre E_1 et E_2 , on peut se poser les problèmes suivants :

1/ - Soit une classification C_1 de E_1 . Déterminer les corrélations γ telles que les ensembles $\gamma(A_1)$ correspondant à toutes les parties classifiées de E_1 dans une classification C_2 que l'on déterminera également.

2/ - Soit une classification C_1 de E_1 et une classification C_2 de E_2 , déterminer des corrélations γ qui font correspondre les parties classifiées de E_1 avec celles de E_2 .

3/ - Si E_2 est un ensemble totalement ordonné, soit une fonction f qui fait correspondre à tout élément x de E_1 un élément y de E_2 ; on peut définir une corrélation π_f par : $\pi_f(x)$ est l'ensemble des Z tels que $Z < y$. Déterminer toutes les fonctions f telles que γ_f soit une des corrélations définies au 1/ - ou au 2/ -.

Corrélations intérieures :

Soit une corrélation γ . Nous appellerons corrélation intérieure δ relative à γ la fonctionnelle définie par :

$$\delta(A) = \bigcap_{x \in A} \gamma(x)$$

Cette fonctionnelle vérifie les relations :

$$\delta(A \cup B) = \delta(A) \cup \delta(B)$$

$$A \subset B \implies \delta(B) \subset \delta(A)$$

$$\delta(A) \cup \delta(B) \subset \delta(A \cap B)$$

Si $\delta(A) = \emptyset$ pour tout ensemble A non réduit à un point, x est fonction de y . En effet, supposons que $\gamma^{-1}(y) = A$ ne soit pas réduit à un point. Pour tout $x \in A$ on aura : $y \in \gamma(x)$ donc :

$$y \in \bigcap_{x \in A} \gamma(x) = \delta(A)$$

donc : $\delta(A)$ n'est pas vide.

Si pour toute partie classifiée A de E_1 $\delta(A)$ est une partie classifiée de E_2 la corrélation sera dite classifiante.

Soit une fonction f qui fait correspondre à tout élément x de E_1 un élément y d'un ensemble E_2 totalement ordonné si γ^{-1}_f est classifiante f sera dite classifiante.

Nous allons donner un exemple de fonctions classifiantes à valeurs dans R . Nous utiliserons dans cet exemple la topologie de R mais il semble qu'une étude directe ne soit pas impossible. En effet, la notion de fonction étagée est une notion qui ne fait pas intervenir la topologie de R . D'autre part, pour un ensemble ordonné, il est possible de définir une sorte de limite qui ne fait pas intervenir la topologie de R . Soit en effet les valeurs y_n des fonctions $f_n(x)$ pour une valeur donnée de x . S_n les ensembles définis par :

$$z < y_n$$

$$T_n = \bigcup_n S_n$$

$$T = \bigcap_1^\infty T_n$$

$$V_n = \bigcap_n S_n$$

$$V = \bigcup_1^\infty V_n$$

Si $T = V$ on a un ensemble de la forme : $\{z \mid z \leq y\}$

y peut être considéré comme la limite de y_n quand n tend vers l'infini et y est définie indépendamment de toute topologie sur R .

Si l'on détermine directement les fonctions classifiantes pour C_1 et C_2 on pourra se poser le problème suivant : déterminer les structures topologiques qu'il faut donner à E_1 et à E_2 pour que les fonctions continues de E_1 sur E_2 soient des fonctions classifiantes, ces topologies étant déterminées en fonction des classifications C_1 et C_2 .

L'exemple suivant montre que la topologie où les ouverts sont les ensembles $a < x < b$, sur l'ensemble des réels classifiés par la classification admettant comme partie $a < x \leq b$, répond à la question.

EXEMPLES DE FONCTIONS CLASSIFIANTES :

Soit une application de E dans un groupe distancié ; on appellera fonction étagée sur E toute fonction telle qu'il existe au moins une répartition en classes de E telle que $\text{Osc.}(f) = 0$ quel que soit la classe P_1 de la répartition. J'appelle $\text{Osc.}(f)$ le maximum de $|f(x') - f(x)|$ pour $x \in P_1, x' \in P_1$.

On appellera fonction ε -étagée sur E toute fonction telle qu'il existe au moins une répartition telle que $\text{Osc.}(f) < \varepsilon$ quel que soit la classe P_1 .

On appellera fonction semi-étagée sur E toute fonction qui est ε -étagée quel que soit ε .

THEOREME II - 5

Soit $f(x)$ une fonction semi-étagée sur E . Les ensembles de points satisfaisant à l'une des conditions

$$f(x) < \lambda, \quad f(x) = \lambda, \quad f(x) > \lambda, \quad f(x) > \lambda$$

sont des parties classifiées quel que soit λ .

En effet, si $f(x)$ est semi étagée pour toute valeur de ε , il est possible de trouver une répartition en classe telle que sur toute classe

$$P_1 : \max_{P_1} f - \min_{P_1} f < \sum$$

On peut classer les classes P_1 en trois catégories :

P_j pour lesquels $\text{Max}(f) < \lambda$

P_k pour lesquels $\text{Min}(f) \leq \lambda \leq \text{Max}(f)$

P_l pour lesquels $\lambda < \text{Min}(f)$

$$\text{Posons : } J_\varepsilon = \bigcup P_j \quad K_\varepsilon = \bigcup P_k \quad L_\varepsilon = \bigcup P_l$$

et $M =$ ensemble des x pour lesquels $f(x) < \lambda$

$$\text{J'ai : } J_\varepsilon \subset M \subset J_\varepsilon \cup K_\varepsilon$$

quel que soit ε . Je peux considérer une suite dénombrable de ε tendant vers zéro. J'aurai :

$$\bigcup_1^\infty J_\varepsilon \subset M \subset \left(\bigcup_1^\infty S_\varepsilon \right) \cup \left(\bigcap_1^\infty (S_\varepsilon \cap K_\varepsilon) \right)$$

Le dernier terme est égal à $\bigcup_1^\infty J_\varepsilon) \cap \bigcap_1^\infty K_\varepsilon$

ce que je pourrais écrire en posant :

$$\bigcup_1^\infty J_\varepsilon = J \quad \bigcap_1^\infty K_\varepsilon = K$$

$$J \subset M \subset J \cap K$$

Tous les points appartenant à K_ε sont tels que :

$$\lambda - \varepsilon \leq f(x) \leq \lambda + \varepsilon$$

Si cette propriété est vraie quel que soit ε , $f(x)$ ne peut être différent de λ et par conséquent pour tout $x \in K$, $f(x) = \lambda$ donc x ne peut appartenir à M ; donc $M = \emptyset$.

Et par suite, les J_ε étant des parties classifiées, M est une partie classifiée. Comme K et M sont disjoints

$$J \cap K = \emptyset$$

Pour tout point de L_ε , $f(x) > \lambda$ donc $f(x) > \lambda$, pour tout point de $L = \bigcup L_\varepsilon$, donc $K \cap L = \emptyset$. En résumé :

$$x \in J \iff f(x) < \lambda$$

$$x \in K \implies f(x) = \lambda$$

$$x \in L \implies f(x) > \lambda$$

Il en résulte que : $f(x) = \lambda \implies x \in K$

$$f(x) > \lambda \implies x \in L$$

D'où les ensembles de points pour lesquels on a l'une des cinq conditions

$$f(x) < \lambda$$

$$f(x) \leq \lambda$$

$$f(x) = \lambda$$

$$f(x) \geq \lambda$$

$$f(x) > \lambda$$

sont des parties classifiées quel que soit λ .

Réciproquement, si l'une des 4 conditions : "l'ensemble des x tels que $f(x) < \lambda$, $f(x) < \lambda$, $f(x) > \lambda$, $f(x) > \lambda$ est une partie classifiée quel que soit λ " est réalisée, la fonction est semi-étagée. Les trois autres le seront alors, d'après le théorème direct ainsi que la condition : "l'ensemble des points x tels que $f(x) = \lambda$ est une partie classifiée", condition qui n'admet pas de réciproque si l'ensemble des valeurs de $f(x)$ n'est pas dénombrable. (Si cet ensemble est dénombrable la condition entraîne manifestement que f est étagée). Supposons la condition $f(x) < \lambda$ réalisée, quel que soit λ et choisissons une infinité dénombrable de valeurs λ_i de $(-\infty < i < +\infty)$ telles que $\lambda_{i+1} - \lambda_i < \varepsilon$ ce qui est possible quel que soit ε ; soit P_i la partie classifiée de E constituée par l'ensemble des points tels que $f(x) < \lambda_i$.

$P_{i+1} \cap C P_i$ sera l'ensemble des points tels que :

$$\lambda \leq f(x) < \lambda_{i+1}$$

C'est une partie classifiée Q_i .

Les parties classifiées Q_i sont disjointes. De plus, si l'on cherche osc. f sur Q_i , on trouve :

$$|f(x_1) - f(x_2)| = \sup f(x_1), f(x_2) - \inf f(x_1), f(x_2)$$

$$\leq \lambda_{i+1} - \lambda_i$$

Donc :

$$\text{Osc. } f < \lambda_{i+1} - \lambda_i < \varepsilon$$

$$E = \bigcup_{-\infty}^{+\infty} (P_{i+1} \cap C P_i) = \bigcup Q_i$$

La réunion étant dénombrable on a déterminé une répartition en classes sur toute classe de laquelle :

$$\text{Osc. } f < \varepsilon$$

La démonstration est analogue pour les trois autres conditions. On démontre aisément que :

- 1/ - toute limite uniforme de fonctions semi-étagées est étagée.
- 2/ - toute suite de fonctions semi-étagées qui convergent uniformément sur un ensemble M tel que toute classe de E contient un point de M admet une limite uniforme.
- 3/ - toute fonction continue de fonctions semi-étagées est semi-étagée.

4/ - toute fonction semi-étagée peut être obtenue comme limite de fonctions étagées.

5/ - toute fonction étagée est une forme linéaire de fonctions caractéristiques d'ensemble.

PRODUIT D'ENSEMBLES CLASSIFIES

Soit les ensembles E_1 et E_2 pourvus des classifications C_1 et C_2 , il est possible de définir une classification C' sur l'ensemble produit $E_1 \times E_2$.

Soit π_{11} une répartition en classes appartenant à une base de la classification C_1 .

P_{11j} une classe appartenant à la répartition π_{11} .

π_{2i} une répartition en classe appartenant à une base de C_2 .

π_{2ik} une classe appartenant à la répartition π_{2i} .

Considérons les sous-ensembles Q_{ijk} de $E_1 \times E_2$ définis par :

$$Q_{ijk} = P_{11j} \times P_{2ik}$$

La famille formée par ces ensembles quand j et k décrivent des ensembles finis ou dénombrables est finie ou dénombrable. Elle recouvre $E_1 \times E_2$ puisque tout élément x, y de $E_1 \times E_2$ est tel que x est contenu dans un P_{11j} et y dans un P_{2ik} .

Les Q_{ijk} sont disjoints car ils diffèrent par l'un au moins des deux indices j et k et les classes correspondant aux deux valeurs de cet indice sont disjointes.

Nous avons donc une répartition en classe π'_i de $E_1 \times E_2$; les π'_i forment donc une famille dénombrable de répartitions de plus en plus fines. Posons pour fixer les idées $i < i'$:

$$\begin{aligned} P_{11j} &= \bigcup_{j'=1}^{j'=\infty} P_{11'jj'} \\ P_{2ik} &= \bigcup_{k'=1}^{k'=\infty} P_{21'kk'} \end{aligned}$$

$$\begin{aligned}
 Q_{ijk} &= \bigcup_{\substack{j'=1 \\ k'=1 \\ j''=1}}^{j'=\infty \\ k'=\infty \\ j''=\infty} P_{11jj'} \times P_{21kk'} \\
 &= \bigcup_{\substack{k'=1 \\ j'=1}}^{j'=\infty \\ k'=\infty} Q_{ijkj'k'}
 \end{aligned}$$

Nous avons défini une classification C' sur $E_1 \times E_2$. Quand nous remplaçons les bases de C_1 et de C_2 par des bases engendrant les mêmes classifications nous obtenons une autre base de la même classification C' . Désignons par R_1 un élément de $S(C_1)$; R_2 un élément de $S(C_2)$ par P_1 un élément d'une base de C_1 ; par P_2 un élément d'une base de C_2 .

$S(C')$ est engendré par $P_1 \times P_2$. Or, $P_1 \times P_2$ est un élément de $S(C_1) \times P_2$ mais $S(C_1) \times P_2$ est une σ -algèbre de Boole. Comme elle est engendrée par des éléments de $S(C')$ on a :

$$S(C_1) \times P_2 \subset S(C') \text{ donc :}$$

$$R_1 \times P_2 \in S(C')$$

$S(C')$ est engendré par les $R_1 \times P_2$.

Mais $R_1 \times S(C_2)$ est une σ -algèbre de Boole ; comme elle est engendrée par des éléments de $S(C')$, on a :

$$R_1 \times S(C_2) \subset S(C')$$

$$\text{donc : } R_1 \times R_2 \in S(C')$$

$S(C')$ est engendré par les $R_1 \times R_2$.

Mais $R_1 \times R_2$ est un élément quelconque de $S(C_1) \times S(C_2)$; donc $S(C')$ est engendré par les éléments de $S(C_1) \times S(C_2)$. Si nous remplaçons les bases de C_1 et de C_2 par des bases qui engendrent les mêmes $S(C_1)$ et $S(C_2)$, $S(C')$ sera remplacé par une σ -algèbre de Boole engendrée par les éléments du même ensemble $S(C_1) \times S(C_2)$ et qui par conséquent lui sera équivalente.

La classification C' que nous avons définie ne dépend donc que des classifications C_1 et C_2 (et non de leurs bases). On l'appellera le produit des classifications C_1 et C_2 .

En raisonnant par récurrence nous pouvons définir une classification sur le produit $E_1 \times E_2 \times \dots \times E_p$, à partir des classifications $C_1, C_2, \dots, C_p$, définies sur $E_1, E_2, \dots, E_p$. Il est manifeste que le produit ainsi défini est commutatif et associatif.

On peut remarquer que toute classe de C_1 étant de la forme UR_x (la réunion pouvant ne pas être dénombrable et R_x étant l'intersection des classes qui contiennent x) que toute classe de C_2 étant de la forme UR_y tout Q_{1j} est de la forme $UR_x \times R_y$. Plus généralement cette forme est celle de toute partie classifiée puisque celles-ci s'obtiennent par réunion ou intersection de cette forme. Celle-ci se trouve donc bien conservée puisque l'intersection de deux $R_x \times R_y$ est soit vide, soit $R_x \times R_y$.

On peut définir d'autres classifications sur $E_1 \times \dots \times E_p$ à partir d'autres répartitions en classes de $E_1 \times \dots \times E_p$. Si les répartitions en classes π_n sont associées à C' la classification C'' qu'elle détermine sera moins fine que C' puisque toute classe de C'' sera une partie classifiée de C' . On peut donc se proposer de déterminer toutes les répartitions associées à C' .

En effet, il existe les répartitions en classes π_n distinctes des π_1 que l'on peut définir à partir des Q_{1jk} puisque les classes de π_n qui sont des unions dénombrables R_m de Q_{1jk} , sont des parties classifiées de C' . Il en sera de même si les classes de π_n sont des intersections dénombrables de R_n .

Pour simplifier l'exposé nous nous placerons dans le cas d'un produit $E_1 \times E_2$, cas qui peut facilement être étendu au moyen d'un raisonnement par récurrence.

Supposons le problème résolu. Soit une classe P_n de π_n . C'est une partie classifiée. Il en est de même de : $P_n \cap (P_{1j} \times E_2)$, P_{1j} étant une classe quelconque d'une répartition en classe de E_1 . Si nous prenons une famille de P_{1j} ordonnée par rapport à l'inclusion $\cap P_{1j} = R_x$, on a donc :

$$\cap [P_n \cap (P_{1j} \cdot E_2)] = P_n \cap [(\cap P_{1j}) \cdot E_2] = P_n \cap (R_x \cdot E_2)$$

$P_n \cap (R_x \cdot E_2)$ est l'ensemble des points de P_n qui se projettent sur R_x . Comme c'est une partie classifiée, cet ensemble est de la forme : $UR_x \cdot R_y$. Il ne peut y figurer que le seul R_x et par conséquent on peut le mettre sous la forme :

$$P_n \cap (R_x \cdot E_2) = R_x \cdot \gamma_n(x).$$

$\gamma_n(x)$ est l'ensemble qui correspond au point x dans la corrélation dé-

finie par : $(x, y) \in P_n$ et R_x . $\gamma_n(x)$ est une partie classifiée de C' . Or, les éléments de $S(C')$ qui sont de la forme R_x . M forment une σ -algèbre de Boole qui coïncide avec R_x . $S(C_2)$ donc $\gamma_n(x) \in S(C_2)$ et $\gamma_n(x)$ est une partie classifiée de C_2 et ceci quel que soit x .

Si nous supposons de plus que la corrélation définie par P_n est compatible avec la relation d'ordre c'est-à-dire que $\gamma_n(x)$ est de la forme $f_2(x) < y \leq f_1(x)$.

On voit que les fonctions $f_1(x)$ et $f_2(x)$ sont constantes sur tout R_x et que les valeurs qu'elles peuvent prendre sont celles qui correspondent à des suprema de parties classifiées de C_2 .

Nous avons là des conditions nécessaires pour que f_1 et f_2 soient des fonctions classifiantes. Il serait intéressant de montrer que f_1 et f_2 sont bien des fonctions classifiantes. Nous nous contenterons de démontrer la réciproque à savoir que si f_1 et f_2 sont des fonctions classifiantes, l'ensemble P_n comme l'ensemble des points (x, y) tels que $y \leq f_1(x)$ est une partie classifiée de C' . Il nous suffira de le montrer pour l'ensemble des points (x, y) tels que $y \leq f_1(x)$. On en déduira la même propriété pour $y \leq f_2(x)$ et on l'étendra à P_n en prenant l'intersection avec le complémentaire.

Prenons une répartition en classe π_1 de E_2 défini par les points : $\dots y_{1j}, y_{1j+1}, \dots$ Soit :

A_{1j} l'ensemble des x tel que $f(x) \leq y_{1j}$

A_{1j+1} l'ensemble des x tel que $f(x) \leq y_{1j+1}$

$$A'_{1j} = A_{1j} \cap C A_{1j+1}$$

f_1 étant une fonction classifiante $A_{1j}, A_{1j+1}, A'_{1j}$ sont les parties classifiées de C_1 .

Soit B_{1j} la réunion des classes de π_1 contenues dans $\bigcap_{x \in A'_{1j}} \{f(x) < y\}$. B_{1j} n'est pas vide puisqu'il contient toujours au moins $\{y_{j+1} < y\}$. B_{1j} est une partie classifiée de C_2 . $D_1 = \bigcup_j A'_{1j} \times B_{1j}$ est une partie classifiée de $E_1 \times E_2$. $D_1 \subset C P_n$ puisque pour tout $(x, y) \in D_1$ on a $f(x) < y$. Montrons que tout point de $C P_n$ appartient à un D_1 .

En effet soit (x, y_0) un point de $C P_n$ nous avons : $f_1(x) < y_0$. On peut avoir $y_0 \in R_{f_1(x)}$ $f_1(x)$ étant le supremum d'une partie classifiée de C_2 il serait celui de $R_{f_1(x)}$ donc on aurait : $y_0 \leq f(x)$.

Il existe donc au moins un y_{1j} tel que :

$$f_1(x) \leq y_{1j}$$

$$y_{1j} < y_0$$

Soit a cet y_{1j} . Nous poserons $y_{1j+1} = b_1$ toutes les fois que l'on aura $y_{1j} = a$.

Ou bien il existe un $b_1 < y_0$ et dans ce cas D_1 contient y_0 .

Ou bien tous les b_1 sont supérieurs ou égaux à y_0 et dans ce cas l'intersection N de tous les $\{z \mid a < z \leq b_1\}$ contient également y_0 . Et il n'existe aucune partie classifiée dont le Supremum $\leq y_0$ et $> a$ donc aucun $f(x)$ n'est compris entre a et $b = \sup_{z \in N} z$.

Par conséquent tous les $f(x)$ sont égaux à a pour $x \in M$

(en posant $M = \{x \mid f(x) \leq a\} \cap C\{x \mid f(x) \leq b\}$) sont égaux à a .

Nous dirons que la classification C définie sur $E_1 \times E_2$ résulte des classifications $C(E_1)$ et $C(E_2)$.

Soit un espace fonctionnel F défini comme produit d'ensemble E_t , t étant un paramètre quelconque élément de l'ensemble G . Tout point X et F sera caractérisé par les valeurs de ces projections sur E_t ; ce qui revient à définir une fonction $x(t)$. Nous désignerons par $x_A(t)$ la fonction déduite de $x(t)$ en prenant les valeurs de $x(t)$ sur les éléments t de l'ensemble A - $x_A(t)$ définit la projection de M sur les sous-espaces vectoriels $F_A(t)$ produit des E_t pour tous les éléments t de l'ensemble A .

Soit une classification $C(F)$ de l'ensemble F . Soit P une classe de $C(F)$, soit M un élément F , M_A sa projection sur F_A , M_{C_A} sa projection sur F_{C_A} ;

$$M = M_A \cdot M_{C_A}$$

Nous pouvons définir une corrélation entre F_A et F_{C_A} en écrivant que : $M \in P$.

Si cette corrélation est classifiante quel que soit P c'est qu'il existe une classification $C(F_A)$ sur F_A et une classification de $C(F_{C_A})$ sur F_{C_A} , telles que $C(F)$ résulte de $C(F_A)$ et de $C(F_{C_A})$.

Supposons la propriété établie quelle que soit la partie classifiée A d'une classification qui sépare G . Soit $(P)_A$ la projection de P sur F_A ,

$(P_1)_A$ une classe de $C(F_A)$ telle que :

$$(P_1)_A \subset (P)_A$$

J'ai en projetant sur E_t :

$$\left((P_1)_A \right)_t \quad \left((P)_A \right)_t$$

Si je considère 2 classes $A_{j'}$, A_j , telles que :

$$A_{j'} \subset A_j$$

de la classification $C(G)$ j'ai : $(P_{11'})_{A_{j'}} \subset (P_1)_{A_j}$

$$\text{Donc : } \left((P_{11'})_{A_{j'}} \right)_t \subset \left((P_1)_{A_j} \right)_t$$

nous poserons pour simplifier les notations :

$$B_{j1} = \left((P_1)_{A_j} \right)_t$$

$$B_{j1j'1'} = \left((P_{11'})_{A_{j'}} \right)_t$$

et en continuant de proche en proche nous écrirons de la même façon :

$$B_{j1j'1j''1''} \text{ etc ...}$$

ce que nous écrirons :

B_I , I étant un ensemble d'indices d'après les relations obtenues précédemment entre les B_I nous aurons :

$$B_{j1} \supset B_{j1j'1'} \supset B_{j1j'1j''1''} \supset \dots$$

les B_I forment donc une suite d'ensembles qui décroît quand I croît ; or I croît quand on prend des répartitions en classes de G de plus en plus fines. A_p étant la classe de la répartition π_p de C_G qui contient t ; il y a une infinité dénombrable $I_n(p)$ d'ensembles d'indices possibles ce qui détermine une infinité dénombrable de $B_{I_n(p)}$. Les $B_{I_n(p)}$ définissent par leurs intersections non vides une répartition en classes π'_p sur E_t . Si nous considérons une répartition en classes π_q de $C(G)$ plus fine que π_p les $I_n(q)$ se déduiront des $I_n(p)$ par adjonction d'indices. Les $B_{I_n(q)}$ seront donc inclus dans les $B_{I_n(p)}$ et la répartition en classes π'_q sur E_t sera plus fine que la répartition π'_p .

Quand les A_p tendent vers $\{t\}$ les π'_p engendrent une classification sur E_t .

NOTA :

La classification engendrée peut très bien ne comporter qu'une seule classe si tous les π_p' n'en comportaient qu'une, parce que tous les $B_{I_n(p)}$ se réduiraient tous à E_t .

Toute classe P de $C(F)$ peut se mettre sous la forme :

$$\bigcup_1^{+\infty} \left(\bigcap_{A \in \pi_p} (P_i) \right)_A$$

puisque $C(F)$ résulte des $C(F_A)$.

Les $\bigcap_{A \in \pi_p} (P_i)_A$ forment une répartition en classes de P. Quand p augmente indéfiniment, ces répartitions engendrent une classification et les éléments limites de cette classification sont de la forme :

$$\bigcap_{t \in G} Q_y(t)$$

$Q_y(t)$ étant l'intersection de toutes les classes de $C(E_t)$ qui contiennent l'élément $y \in E_t$. Donc pour toute classe P de $C(F)$ nous avons :

$$P = \bigcup \left(\bigcap_{t \in G} Q_y(t) \right),$$

la réunion pouvant toutefois n'être pas dénombrable.

Nous dirons que $C(F)$ résulte des $C(E_t)$, $t \in G$.

Nous pouvons définir de la même façon la relation "résulte des" sur un sous-ensemble de F.

Il suffira pour cela de considérer que la relation " $C(F_A)$ résulte de $C(F_B)$ et de $C(F_A \cap C_B)$ " est vérifiée par toute les classes B de G qui sont des sous-classes de la classe A. La démonstration précédente est valable sans changement.

Les classes P d'une classification $C(F_A)$ pourront se mettre sous la forme :

$$P = \bigcup \left(\bigcap_{t \in A} Q_y(t) \right)$$

De sorte que nous pourrions dire :

$C(F_A)$ résulte des $C(E_t)$, $t \in A$.

La relation "résulte des" est associative, c'est-à-dire si

$$M = \bigcup A_i$$

(les A_i étant des parties classifiées de G) entraîne que $C(F_M)$ résulte des $C(F_{A_i})$ et si $C(F_{A_i})$ résulte des $C(E_t)$, $t \in A_i$ alors $C(F_M)$ résulte des

$C(E_t)$, $t \in \bigcup A_i = M$. Il suffira de prendre parmi les π_p qui servent à faire le raisonnement pour établir que $C(F_M)$ résulte des $C(E_t)$ $t \in M$ la répartition en classes de M déterminée par les A_i , (si la réunion est dénombrable) ou une suite de répartitions en classes dont les parties ultimes sont les A_i (si la réunion n'est pas dénombrable, cette réunion devant toutefois être associée à la classification qui sépare G).

En résumé nous pouvons énoncer le théorème suivant :

THEOREME II - 6 -

Si une classification de F résulte de la classification $C(F_L)$ et de la classification $C(F_M \cup F_N)$, elle résulte de la classification $C(F_L \cup F_M)$ et de la classification $C(F_N)$.

CHAPITRE III

FONCTIONNELLES ADDITIVES

Généralités concernant les fonctionnelles sur un ensemble classifié

Je dirai que j'ai défini une fonctionnelle sur un ensemble classifié E , à valeur dans F si à tout élément M d'une famille de parties classifiées de E , je sais faire correspondre un élément de F noté $\varphi(M)$.

Je dirai qu'une fonctionnelle φ' est plus fine qu'une fonctionnelle si toutes les fois que $\varphi(M)$ existe $\varphi'(M)$ existe et est égal à $\varphi(M)$.

Je dirai que j'ai défini une norme pour les valeurs de la fonctionnelle si j'ai défini une application de $\varphi(M)$ sur un ensemble ordonné A de telle sorte que à tout M_i je sais faire correspondre un $k_i = k(\varphi(M_i))$.

Je dirai que j'ai défini une borne pour les valeurs de la fonctionnelle sur une famille L d'ensemble M_i si je sais déterminer une fonction h_L des k_i satisfaisant à :

$$k_i \leq h_{M_i}$$

$$L_1 \subset L_2 \Rightarrow h_{L_1} \leq h_{L_2}$$

Il en résulte que :

1/ - Si nous avons une borne $h(\varphi)$ sur l'ensemble des parties classifiées pour lesquelles $\varphi(M)$ est défini, il y en a une h_L sur toute famille L et $h_L \leq h(\varphi)$.

En particulier pour tout k on a : $k_i \leq h(\varphi)$

Nous dirons que la fonctionnelle est bornée et nous appellerons $h(\varphi)$ la hauteur de φ .

2/ - Si φ est plus fin que φ' et si φ est borné φ' est borné et $h(\varphi') \leq h(\varphi)$.

En effet l'ensemble des parties classifiées sur lequel $\varphi'(M)$ est défini peut être considéré comme un sous-ensemble de celui des parties classifiées sur lequel $\varphi(M)$ est défini.

FONCTIONNELLES ADDITIVES -

Nous dirons qu'une fonctionnelle est additive si φ prend ses valeurs dans un demi-groupe additif

$$\varphi(A \cup B) = \varphi(A) + \varphi(B) \text{ pour } A \cap B = \emptyset$$

Nous dirons qu'une fonctionnelle est σ -additive si

$$\varphi\left(\bigcup_1^{+\infty} A_i\right) = \sum_1^{+\infty} \varphi(A_i)$$

avec $A_i \cap A_j = \emptyset$ pour $i \neq j$

Nous supposons désormais que nous connaissons une norme pour les éléments de F satisfaisant à :

1/ - $k(A) + k(B) \leq k(A + B) \leq k(A) + k(B)$ quel que soient les éléments de A et B de F .

2/ - si la série $\sum_1^{+\infty} k(A_i)$ est convergente, la série $\sum_1^{+\infty} A_i$ est associativement et commutativement convergente quels que soient les éléments A_i de F .

Définition d'une fonctionnelle σ -additive -

1/ - sur une répartition en classe π_i

Soit P_{ij} avec $j \in Z$ une classe de la répartition.

Nous poserons

$$\pi_i(P_{ij}) = A_{ij}$$

$$k(A_{ij}) = k_{ij}$$

Pour avoir une fonctionnelle σ -additive il suffira que $\sum_Z k_{ij}$ existe.

Nous poserons $\sum_{j \in Z} k_{ij} = h_i$, hauteur de φ_i .

En effet pour tout ensemble $M = \bigcup_{j \in Z, C \subset Z} P_{ij}$, on a une série convergente $\sum_{j \in Z'} k_{ij}$ et par conséquent $\sum_{j \in Z'} A_{ij}$ converge également.

Nous pourrions poser :

$$\varphi_i(M) = \sum_{j \in Z} \varphi_i(P_{ij})$$

Si nous avons mis M sous la forme $M = \bigcup M_1$ les M_1 étant disjoints et de la forme :

$$M_1 = \bigcup_{j \in Z_1} P_{1j}$$

nous devons avoir

$$Z = \bigcup Z_1$$

et la série $\sum_{j \in Z} \varphi(P_{1j})$ étant associativement convergente on a :

$$\sum_{j \in Z} \varphi(P_{1j}) = \sum_{\tau} \left(\sum_{j \in Z_1} \varphi_i(P_{1j}) \right) = \sum \varphi(M_1)$$

2/ - Sur une classification C , soit une suite de fonctionnelle φ_i de plus en plus fine définies sur les répartitions π_i d'une classification et bornées dans leur ensemble (c'est-à-dire que $h_i < H$ quel que soit i).

Nous pouvons définir une fonctionnelle Φ par : $\Phi(M) = \varphi_i(M)$ pour tout ensemble M pour lequel on peut trouver un φ_i tel que $\varphi_i(M)$ existe.

On aura alors $\varphi_{i'}(M) = \varphi_i(M)$ pour tout $i' \geq i$. La fonctionnelle $\Phi(M)$ est σ -additive si pour toute une suite M_n d'ensembles qui convergent vers $\emptyset$ on a : $\Phi(M_n) \rightarrow 0$.

Quelle que soit la façon dont on mette E sous la forme $E = \bigcup P_{ij}$ on aura : $\sum k_{ij} \leq h$.

En effet si on ordonne par rapport à i on peut écrire

$$E = \left(\bigcup_{j_1} P_{1j_1} \right) \cup \left(\bigcup_{j_2} P_{2j_2} \right) \cup \dots \cup \left(\bigcup_{j_n} P_{nj_n} \right) \cup \dots$$

$$\text{Nous avons : } \sum_{j_1} k_{1j_1} \leq h_1$$

mais en exprimant les P_{1j_1} en fonction des P_{2j_1} nous avons :

$$K_{1j_1} \leq \sum_1 k_{2j_1}$$

$$\text{or } \left(\bigcup_{j_1} P_{1j_1} \right) \cup \left(\bigcup_{j_2} P_{2j_2} \right) = \left(\bigcup_{j_1} P_{2j_1} \right) \cup \left(\bigcup_{j_2} P_{2j_2} \right)$$

$$\text{donc : } \sum_{j_1} \sum_1 k_{2j_1} + \sum_{j_2} K_{2j_2} \leq h_2$$

$$\text{donc : } \sum_{j_1} k_{1j_1} + \sum_{j_2} k_{2j_2} \leq h_2$$

On démontrerait de même par récurrence que :

$$\sum k_{1j_1} + \sum k_{2j_2} + \dots + \sum k_{nj_n} \leq h_n$$

et par conséquent

$$\sum_1^{\infty} k_{1j} \leq H$$

La série $\sum k_{ij}$ étant convergente la série de $\sum \varphi(P_{ij})$ l'est également et l'on peut poser :

$$\varphi(E) = \sum \varphi(P_{ij})$$

Si E peut être mis sous la forme $\cup E_m$ les E_m étant de la forme $\cup P_{ij}$ les séries $\sum k_{ij}$ correspondant aux E_m sont des sous-séries de la série $\sum k_{ij}$ correspondant à E. Elles convergent donc et par suite les séries $\sum \varphi(P_{ij})$ correspondantes convergent également et on peut les prendre pour définir les $\varphi(E_m)$.

De plus comme

$$k\varphi(E_m) \leq \sum_{P_{ij} \subset E_m} k(\varphi(P_{ij})) \leq \sum_{P_{ij} \subset E} k(\varphi(P_{ij})) \leq H$$

On a pour toutes les réunions M de classes de C :

$$k(\varphi(M)) \leq H$$

Montrons que par passage à la limite on obtient toujours des ensembles jouissant de cette propriété.

Soit une suite d'ensembles A_p jouissant de la propriété

$$k(\varphi(A_p)) \leq H$$

A_p possède donc une limite donc

$$A_{p+1} \cup B_{p+1} = A_p \cup D_p$$

$$\text{avec } A_{p+1} \cap B_{p+1} = \emptyset$$

$$A_p \cap D = \emptyset$$

$$B_{p+1} \cap D_p = \emptyset$$

$$B_{p+1} \longrightarrow \emptyset$$

$$D_p \longrightarrow \emptyset$$

$$\text{Nous avons : } \varphi(A_{p+1}) + \varphi(B_{p+1}) = \varphi(A_p) + \varphi(B_p)$$

$$k(\varphi(A_{p+1})) \leq k(\varphi(A_{p+1}) + (B_{p+1})) + k(\varphi(B_{p+1}))$$

$$\leq k(\varphi(A_p) + \varphi(D_p)) + k(\varphi(B_{p+1}))$$

$$\leq k(\varphi(A_p)) + k(\varphi(B_p)) + k(\varphi(B_{p+1}))$$

$$k(\varphi(A_p)) \leq H \text{ et : } k(\varphi(A_{p+1})) - k(\varphi(A_p)) \rightarrow 0$$

Donc $k \varphi(A_p)$ tend vers une limite $\leq H$.

On peut déterminer la valeur $\varphi(A)$ correspondant à cette limite en écrivant

$$\varphi(A_{p+1}) - \varphi(A_p) = \varphi(D_p) - \varphi(B_{p+1})$$

pour toutes les valeurs de p et en additionnant ce qui donne :

$$\varphi(A) - \varphi(A_1) = \sum_1^{+\infty} (\varphi(D_p) - \varphi(B_{p+1}))$$

Si l'ensemble A peut être obtenu de 2 façons différentes comme limites de 2 suites d'ensembles

A_n et A'_n on pourra poser :

$$A_n \cup B_n = A'_n \cup B'_n$$

avec :

$$A_n \cap B_n = \emptyset$$

$$A'_n \cap B'_n = \emptyset$$

$$B_n \cap B'_n = \emptyset$$

$$B_n \rightarrow \emptyset$$

$$B'_n \rightarrow \emptyset$$

et par conséquent $\varphi(A_n) + \varphi(B_n) = \varphi(A'_n) + \varphi(B'_n)$

$\varphi(B_n)$ et $\varphi(B'_n)$ tendant vers zéro les limites de $\varphi(A_n)$ et $\varphi(A'_n)$ sont égales.

$\varphi(A)$ est donc bien définie de façon unique et la fonctionnelle φ est bien σ -additive.

APPLICATIONS -

$1/$ - E est un ensemble de nombres réels. F est l'ensemble des nombres positifs.

$$k(a) = a \quad \sum a_i = 1$$

Dans ce cas φ est une distribution de probabilités $\sum a_i = h$ avec

$$\varphi(x_i < x < x_{i+1}) = x_{i+1} - x_i$$

E est borné.

Un cas particulier important est celui de la mesure de Lebesgue, quand l'ensemble E est un ensemble de réels de la forme $a < x \leq b$. Pour simplifier l'exposé nous prendrons pour ensemble E, l'ensemble $0 < x \leq 1$ l'extension se faisant sans difficultés.

Nous pouvons définir sur E les répartitions en classes

$$\pi_n \text{ par : } P_{nq} = \left\{ \frac{q}{n} < x \leq \frac{q+1}{n} \right\}.$$

Les répartitions π_n ne forment pas une suite de répartitions de plus en plus fines mais il est facile d'extraire une telle suite de la famille des π_n pour définir la classification C qui sépare E. Les répartitions qui n'appartiendront pas à cette suite sont manifestement associées à la classification.

Si nous définissons $\varphi(P_{nq})$ pour toute valeur de n et de q nous avons un excédent de valeurs de φ pour définir la fonctionnelle ; mais si nous démontrons que la valeur choisie pour P_{nq} est la même que celle que l'on obtient à partir des valeurs de $\varphi(P_{nq})$ choisies pour les classes de toutes les répartitions plus fines nous démontrons à la fois que les φ_i sont de plus en plus fines et que les valeurs excédentaires sont bien celles prises par la fonctionnelle telle qu'elle est définie au moyen des φ_i .

$\pi_{n'}$ est plus fin que π_n si n' est un multiple de n : $n' = \lambda n$ mais alors la classe P_{nq} est la réunion de λ classes $P_{n'q'}$.

Nous poserons :

$$\varphi(P_{nq}) = \frac{1}{n} \text{ pour } P = \left\{ \frac{q}{n} < x \leq \frac{q+1}{n} \right\}$$

Nous avons :

$$\varphi(P_{nq}) = \frac{1}{n}$$

$$\varphi(P_{n'q'}) = \frac{1}{n'}$$

$$\varphi\left(\bigcup_{P_{n'q'} \subset P_{nq}} P_{n'q'}\right) = \lambda \cdot \frac{1}{n'}$$

$$\text{Or } \frac{\lambda}{n'} = \frac{1}{n}$$

$$\text{Donc } \varphi\left(\bigcup_{P_{n'q'} \subset P_{nq}} P_{n'q'}\right) = \varphi(P_{nq}).$$

Pour tout ensemble $M : 0 < x \leq r$, r étant un nombre rationnel, on a : $\varphi(M) = r$.

Par passage à la limite on voit que pour tout ensemble de la forme : $0 < x \leq s$ on a : $\varphi(M) = s$.

Nous sommes donc ramenés à la théorie classique de Borel-Lebesgue. Nous dirons que nous avons défini la mesure de Lebesgue sur la classification de Borel.

- 2/ - E est un ensemble totalement ordonné
 F est l'ensemble des nombres réels.

$$k(a) = |a|$$

$$\varphi(x_i < x < x_{i+1}) = f(x_{i+1}) - f(x_i)$$

La fonctionnelle existe à condition que f soit une fonction à variation bornée.

C'est le cas classique de la fonctionnelle de Stieltjes. E est un ensemble partiellement ordonné.

Dans ce cas on se rapportera au travail d'André Revuz.

- 3/ - F est un espace vectoriel normé

$$k(a) = |a|.$$

Il est possible de trouver des fonctionnelles φ satisfaisant aux conditions du paragraphe précédent.

Je l'établirai par un exemple

Prenons pour E l'ensemble $0 < x \leq 1$, pour classification la classification C définie par les classes

$$P_{iq} : \frac{q}{2^i} < x \leq \frac{q+1}{2^i}$$

Pour $\varphi(P_{iq})$ le vecteur de module $\frac{1}{2^i} \prod_{n=1}^{n=i} \frac{1}{\cos \frac{\pi}{2^{n+1}}}$ et pour angle polaire $\pi(1 - \frac{q + \frac{1}{2}}{2^i})$

Il est facile de voir que la fonctionnelle φ satisfait aux conditions du paragraphe précédent et par conséquent elle est σ -additive.

4/ - F est un champ de variables aléatoires à moment du 1° ordre et 2° ordre fini.

$$k(X) = \sqrt{m^2 + \sigma^2} \quad m \text{ étant le moment du 1° ordre.}$$

σ étant le moment du 2° ordre réduit.

Nous savons que si $X = X_1 + X_2$ on a

$$m = m_1 + m_2; \quad \sigma = \sigma_1 + \sigma_2.$$

Nous avons donc : $k(X) \leq k(X_1) + k(X_2)$ et $k(X)$ est bien une norme.

De plus nous savons que si $\sum \sqrt{m_i^2 + \sigma_i^2}$ converge $\sum m_i$ et $\sum \sigma_i$ convergent et que $\sum X_i$ est associativement et commutativement convergente. Prenons pour ensemble E l'ensemble $0 < t \leq 1$. Nous définirons une classification comme dans le cas d'une mesure de Lebesgue et nous prendrons mêmes précautions pour les φ_n correspondant à des répartitions associées à C .

Nous poserons donc $\varphi(E) = X$.

Si X est divisible par n c'est-à-dire il est égal à la somme de n variables aléatoires indépendantes de même fonction de répartition X_n ce que nous écrirons $X = n * X_n$; $\frac{1}{n} * X = X_n$

$$\text{Nous poserons } \varphi(P_{nq}) = X_n \text{ pour } P_{nq} = \left\{ \frac{q}{n} < t \leq \frac{q+1}{n} \right\}$$

Nous pourrions donc définir φ sur tous les P_{nq} si X est divisible par tout n - Nous avons :

$$\varphi(P_{nq}) = X_n = \frac{1}{n} * X$$

$$\varphi(P_{n'q'}) = X_{n'} = \frac{1}{n'} * X$$

$$P_{n'q'} \bigcup_{C} P_{nq} = \lambda * \left(\frac{1}{n'} * X \right)$$

Or si nous formons $n * (\lambda * (\frac{1}{n'} * X))$ cela revient à additionner λn variables indépendantes égales à $X_{n'}$, donc à former $n' * X_{n'} = X$. Nous avons donc :

$$\lambda * X_{n'} = \frac{1}{n} * X = X_n$$

$$\varphi\left(\bigcup P_{n'q'}\right) = \varphi(P_{nq})$$

Il reste donc à démontrer que φ est borné et continu. Pour chaque φ_i on a : $h_i = \sum_q k_{iq}$

$$\text{Or} \quad \varphi(P_{iq}) = \frac{1}{n} * X$$

On devra donc avoir en désignant par m_{iq} et σ_{iq} les moments du 1er ordre et du 2ème ordre de $\varphi(P_{iq})$.

$$n m_{iq} = m ; n \sigma_{iq} = \sigma$$

donc

$$k_{iq} = \sqrt{\left(\frac{m}{n}\right)^2 + \left(\frac{\sigma}{n}\right)^2}$$

$$h_i = \sum_q k_{iq} = n \sqrt{\frac{m^2 + \sigma^2}{n^2}} = \sqrt{m^2 + \sigma^2}$$

h_i est constant donc H existe.

Si P_{iq} tend vers $\{t\}$ on a :

$$m_{iq} = \frac{m}{n} \rightarrow 0$$

$$\sigma_{iq} = \frac{\sigma}{n} \rightarrow 0 ; \text{ donc } \varphi(P_{iq}) \rightarrow 0.$$

Ceci est donc vrai à fortiori si $P_{iq} \rightarrow \emptyset$. Nous avons donc défini la fonctionnelle σ -additive pour toute partie classifiée de C.

Remarque :

Nous avons défini une opération $\frac{q}{n} * X$ pour laquelle $\frac{1}{n} * X$ se présente comme l'opération inverse de la somme de n variables aléatoires indépendantes de même distribution.

Si l'on considère cette opération comme une sorte de multiplication l'opération inverse peut être considérée comme une division et les expressions $\frac{q}{n} * X$ comme des fractions de X ou comme produit de X par une fraction.

Sur toute partie classifiée M de C dont la mesure de Lebesgue est $\frac{q}{n}$ on a :

$$\varphi(M) = \frac{q}{n} * X.$$

Par extension pour toute partie classifiée M dont la mesure de Lebesgue est un nombre réel t nous poserons $t * X = \varphi(M)$.

Cette définition est cohérente car si deux parties classifiées M' et M'' ont même mesure de Lebesgue t la fonctionnelle φ prend la même valeur pour ces deux ensembles.

En effet on peut les approcher par deux suites d'ensembles M' et M'' de mesures de Lebesgue r rationnelles et égales entre elles.

On a donc constamment $\varphi(M'_n) = r_n * X = \varphi(M''_n)$ donc à la limite : $\varphi(M') = \varphi(M'')$.

5/ - E est un ensemble totalement ou partiellement ordonné F est un espace vectoriel dont les composantes sont certaines ou aléatoires.

Ce cas réunit toutes les difficultés trouvées dans les cas précédents mais ne pose pas de problème nouveau.

ENSEMBLES EQUIVALENTS -

Soit m une fonctionnelle σ -additive sur un ensemble E, un sous ensemble A de E est m-négligeable si $m(B) = 0$ pour tout B inclus dans A. En particulier $m(A) = 0$; mais tout ensemble A tel que $m(A) = 0$ n'est pas nécessairement m-négligeable sauf si la fonctionnelle prend ses valeurs dans un ensemble F dont tous les éléments sont positifs ou nuls.

Si $m(A) = 0$ sans que $m(B) = 0$ pour tout $B \subset A$. On dira que A est irrégulier. Si A est n-négligeable, on écrira :

$A \xrightarrow{m} 0$. (on dira négligeable et on écrira $A \rightleftharpoons 0$ lorsqu'aucune confusion n'est à craindre).

Deux ensembles A et B sont équivalents s'ils ne diffèrent par un ensemble m-négligeable au plus m - c'est-à-dire si :

$$(A \cap C B) \rightleftharpoons 0 \quad B \cap C A \rightleftharpoons 0$$

Cette relation est symétrique et réflexive. Elle est de plus transitive en effet si l'on a en outre :

$$B \cap C D \rightleftharpoons 0 \quad D \cap C B \rightleftharpoons 0$$

On aura :

$$A \cap C D = [A \cap (B \cap C D)] \cup [(A \cap C B) \cap C D]$$

$$D \cap C A = [D \cap (B \cap C A)] \cup [(A \cap C B) \cap C D]$$

Donc :

$$A \cap C D = 0 \quad D \cap C A \Leftrightarrow 0$$

Une réunion finie d'ensembles négligeables est négligeable car tout sous-ensemble est une réunion finie d'ensembles de mesure nulle.

Une réunion dénombrable d'ensembles négligeables est aussi gligeable car la limite d'un ensemble de mesure nulle a une mesure nulle. De même toute intersection finie ou dénombrable d'ensembles de mesure nulle.

Théorème III - 1 :

Soit une famille Φ d'ensembles A_i non m - négligeable, si l'ensemble F des valeurs prises par la fonctionnelle est archimédien, il existe une famille G dénombrable d'ensembles B_j non m - négligeables, disjoints, tels que tout A_i soit m - équivalent à une réunion finie ou dénombrable de B_j . On dira que A_i est non m - négligeable si $m(A_i) \neq 0$.

Soit A_0 un des A_i tels que : $m(A_i) \geq h > 0$.

Formons la famille Φ_1 des $A_i \cap C A_0$.

S'il existe plusieurs éléments de Φ_1 pour lesquels la valeur de la fonctionnelle est supérieure ou égale à h ; choisissons en un A_1 et formons $A_1 \cap C A_0 \cap C A_1$ qui formeront une famille Φ_2 .

Continuons jusqu'à ce que nous soyons arrivés à une famille Φ_n qui ne contienne aucun ensemble pour lequel la valeur de la fonctionnelle est supérieure ou égale à h .

Nous y parviendrons en un nombre fini d'opérations en effet A_0, A_1, A_2 etc. sont tous disjoints et la valeur de m est supérieure ou égale à h . Donc :

$$m(A_0 \cup A_1 \cup \dots \cup A_{n-1}) \geq n h$$

or la valeur de la fonctionnelle pour tout ensemble est borné donc n est borné.

Nous avons donc formé une famille Φ' comportant un nombre fini d'ensembles disjoints pour lesquels la valeur de la fonctionnelle est $\geq h$ et une infinité d'ensembles pour lesquels la valeur de la fonctionnelle est $< h$.

Tout A_i est réunion d'un nombre fini d'éléments de Φ' . Nous pouvons recommencer avec les ensembles pour lesquels la valeur de la fonctionnelle est inférieure à h pour une autre valeur $k < h$.

Nous obtiendrons ainsi un nombre fini d'ensembles disjoints pour

lesquels la valeur de la fonctionnelle est comprise entre h et k .

En répétant le procédé pour une suite décroissante convergeant vers zéro, de valeurs positives $h, k, \dots$, puis pour une suite croissante et convergeant vers zéro, de valeurs négatives $h', k', \dots$, nous pouvons former une infinité dénombrable G d'ensembles disjoints B_j pour lesquels la valeur de la fonctionnelle n'est pas nulle. Tout A_i est donc une réunion d'une infinité dénombrable ou d'un nombre fini B_j et d'un ensemble C_k appartenant à une famille H d'ensembles pour lesquels la valeur de la fonctionnelle est nulle.

Les C_k sont négligeables. En effet supposons qu'ils ne le soient pas, c'est qu'il existerait des ensembles $D_1 \subset C_k$ pour lesquels $m(D_1) \neq 0$.

Il nous suffit pour éviter cela de procéder avec une famille F contenant avec tout ensemble non négligeable ses sous-ensembles non négligeables de sorte que les D_1 sont eux-mêmes une réunion de B_j .

Application :

Soit deux fonctionnelles m_1 et m_2 définies sur une même structure d'ensemble classifiée. Considérons la famille des ensembles A_i m_2 -négligeables qui ne sont pas m_1 -négligeables. Ils forment une famille F ils sont donc m_2 -équivalents aux réunions finies ou dénombrables G d'ensembles B_j m_1 -non négligeables, m_2 -négligeables, disjoints.

Posons : $M = \bigcup B_j$.

Tout B_j est un sous-ensemble de M . Donc tout $A_i \cap C M$ est m_2 -négligeable mais il est également m_1 -négligeable car sinon il serait m_2 -négligeable donc inclus dans M . Donc :

Théorème III - 2 :

Tout ensemble m_2 -négligeable, m_1 -non négligeable est m_2 -équivalent et m_1 -équivalent à un sous ensemble d'un ensemble m_2 -négligeable, m_1 -non négligeable.

REPARTITIONS REGULIERES -

On appellera répartition en classe m - régulière d'un ensemble E , une répartition en classe pour toute classe P_i de laquelle on a soit :

$$m(P_i) > 0, \text{ soit } P_i \xrightarrow[m]{} 0.$$

Théorème III-3 :

Etant donné une répartition en classes π d'un ensemble il est

toujours possible de trouver une répartition en classe m - régulière plus fine que π . Supposons que la répartition en classe ne soit pas régulière, c'est-à-dire qu'il existe au moins une classe P_i irrégulière dont il existe donc des sous-classes pour lesquelles la valeur de la fonctionnelle n'est pas nulle.

Les classes d'une répartition en classes de P_i sont des ensembles pour lesquels la valeur de la fonctionnelle est non-nulle ou nulle. Si toutes ces classes pour lesquelles la valeur de la fonctionnelle est nulle, sont négligeables, il suffira de remplacer P_i par cette répartition dans π pour avoir une répartition régulière. Sinon il y aura des sous-classes Q_j irrégulières.

On recommencera avec les classes Q_j comme avec les classes P_i et ainsi de suite.

Nous obtenons donc des répartitions de plus en plus fines formées de classes non- négligeables et irrégulières. Les classes irrégulières formant des suites ordonnées par rapport à l'inclusion. L'intersection de toutes les classes d'une telle suite est nécessairement négligeable car la valeur de la fonctionnelle pour cette intersection est nulle et il n'y a pas de sous-ensembles irréguliers car s'il y en avait un, il appartiendrait à la suite et contiendrait une intersection.

Dans la suite des répartitions successives, nous pouvons distinguer une partie régulière et une partie irrégulière. Nous appellerons E_1 l'intersection de toutes ces parties irrégulières. Le complémentaire de E_1 est la réunion d'ensembles M_i tels que $M_i \subset M_{i'}$, si $i < i'$. Nous pourrions prendre sur chacun des M_{i+1} intersections complémentaires de M_i la répartition en classe de cet ensemble par rapport à π_{i+1} . Nous avons donc une réunion dénombrable de réunions dénombrables de classes donc une répartition en classes de $C E_1$. E_1 est une partie classifiée. Ou bien E_1 est négligeable ou bien il ne l'est pas. S'il ne l'est pas il existe des sous-ensembles de E_1 pour lesquels la valeur de la fonctionnelle n'est pas nulle. A partir de ces sous-ensembles on détermine une répartition en classes de E_1 dont les classes sont des ensembles pour lesquels la valeur de la fonctionnelle n'est pas nulle ou bien est nulle. Ceux pour lesquels elle est nulle sont négligeables ou irréguliers. Nous pouvons donc recommencer à partir de E_1 la suite d'opérations que nous avons faite sur E . On obtiendra ainsi une partie classifiée E_2 à partir de E_1 de la même façon que l'on a obtenu E_1 à partir de E , $E_1 \subset C E_2$ étant réparti en classes dont aucune n'est irrégulière. Si E_2 est irrégulier on recommence et on obtient E_3 et ainsi de suite.

L'intersection $E_{(\omega)}$ de tous les E_i est négligeable car sur l'ensemble $E_{(\omega)+1}$ la fonctionnelle prend la même valeur que sur l'ensemble $E_{(\omega)}$.

$$m(E_\omega \subset C E_{\omega+1}) = 0.$$

quelle que soit la façon d'obtenir $E_{\omega+1}$ à partir de E_ω .

Donc pour tout sous-ensemble de E_ω , la valeur de la fonctionnelle est nulle. La répartition en classes de E obtenue en prenant les répartitions en classes de $E \subset C E_1 \dots$ de $E_1 \subset C E_{1+1} \dots$ et en complétant avec l'ensemble E_ω , ne contient que des ensembles non m - négligeables ou m - négligeables. Nous avons ainsi obtenu une répartition régulière.

Applications :

Si nous considérons la famille des classes d'une répartition régulière pour lesquelles la valeur de la fonctionnelle est positive nous obtenons une partition cellulaire (théorie par défaut).

Si nous considérons des réunions de ces classes avec des classes négligeables de façon à recouvrir E , la même classe négligeable pouvant être réunie à plusieurs autres classes non-négligeables, nous obtenons une partition cellulaire (théorie par excès).

Pour le développement de ces théories on se reportera aux travaux de Christian Pauc et K. Krickeberg.

OPERATIONS sur LES FONCTIONNELLES σ - ADDITIVES -

1/ - Addition

Si deux fonctionnelles σ - additives φ_1 et φ_2 sont définies sur une même famille de parties classifiées et prennent leur valeur dans un même groupe, la fonctionnelle $\varphi_3(A) = \varphi_1(A) + \varphi_2(A)$ est σ - additive.

En effet nous aurons :

$$\text{donc} \quad k(\varphi_3) \leq k(\varphi_1) + k(\varphi_2)$$

$$h(\varphi_3) \leq h(\varphi_1) + h(\varphi_2)$$

si

$$A \longrightarrow \emptyset \quad \varphi_1(A) \longrightarrow 0$$

$$\varphi_2(A) \longrightarrow 0$$

$$\text{donc} \quad \varphi_3(A) \longrightarrow 0$$

$$\sum (\varphi_1(A_i) + \varphi_2(A_i)) = \sum \varphi_1(A_i) + \sum \varphi_2(A_i)$$

$$\text{Lim. } (\varphi_1(A_i) + \varphi_2(A_i)) = \text{Lim } \varphi_1(A_i) + \text{Lim } \varphi_2(A_i)$$

2/ - Soustraction -

Il est clair que la différence de deux fonctionnelles est une fonctionnelle dans les mêmes conditions que pour l'addition.

3/ - Limites -

Si la suite des hauteurs h_n d'une suite de fonctionnelles μ_n tend vers zéro on dira que la suite des fonctionnelles converge vers la fonctionnelle nulle. La suite des valeurs des fonctionnelles pour une partie classifiée donné tend vers zéro car :

$$k(\mu_n(A)) < h_n \longrightarrow 0$$

Si la suite des hauteurs $h(\mu_n)$ d'une suite de fonctionnelles μ_n est telle que :

$$h(\mu_n) < H$$

$h(\mu_n - \mu_q) \longrightarrow 0$ quelle que soit la façon dont p et q augmentent indéfiniment, il existe une fonctionnelle μ telle que :

$\mu - \mu_n$ converge vers la fonctionnelle nulle. On dira que μ_n converge vers μ ou que μ est la limite de μ_n .

A toute partie classifiée A il correspond $\mu_n(A)$. Etant donné ε_p je peux déterminer q de façon que

$$k(\mu_p(A) - \mu_q(A)) < \varepsilon_p$$

et en prenant : $\varepsilon_q < \varepsilon_p - k(\mu_p(A) - \mu_q(A))$ l'intervalle

$$k(\mu_q(A)) - \varepsilon_q, k(\mu_q(A)) + \varepsilon_q \text{ sera entièrement contenu}$$

dans l'intervalle :

$$k(\mu_p(A)) - \varepsilon_p, k(\mu_p(A)) + \varepsilon_p.$$

Le filtre ainsi défini a un point adhérent et $\mu_n(A)$ converge vers une valeur que l'on appellera $\mu(A)$ quel que soit ε il est possible de trouver $\mu_N(A)$ tel que :

$$k(\mu_n(A) - \mu(A)) < \varepsilon$$

pour tout $n > N$.

Soit $B = \bigcup_1^{+\infty} A_i$ les A_i étant des parties classifiées disjointes.

Soit : $\mu(A_i)$ la valeur de la fonctionnelle μ pour A_i . Il est possible de choisir des ϵ_i de façon que $\sum \epsilon_i < \epsilon$ et des N_i de façon que :

$$k(\mu_{n_i}(A_i) - \mu(A_i)) < \epsilon \text{ pour } n_i > N_i.$$

posons :

$$S_p = \sum_1^p \mu(A_i)$$

$$B_p = \bigcup_1^p A_i$$

$$N_{(p)} = \sup_1^p N_i$$

Pour tout $n > N_{(p)}$ on aura :

$$k(\mu_n(A_i) - \mu(A_i)) < \epsilon_i$$

$$k\left(\sum_{i=1}^{i=p} \mu(A_i) - S_p\right) < \sum_{i=1}^{i=p} \epsilon_i < \epsilon$$

$$k(\mu_n(B_p) - S_p) < \epsilon.$$

Or $\mu_n(B_p)$ converge vers $\mu(B_p)$ et il est possible de choisir n de façon que :

$$k(\mu_n(B_p) - \mu(B_p)) < \eta$$

On aura donc :

$$k(\mu(B_p) - S_p) < \epsilon + \eta$$

quels que soient ϵ et η ; donc $\mu(B_p) = S_p$. Ceci étant vrai quel que soit p fini. Si p augmente indéfiniment, ...
 S_p converge vers une limite S en effet

$$\sum_{i=1}^{i=p} k[\mu(A_i)] < \sum_1^p \epsilon_i + \sum_1^p k[\mu_{n_i}(A_i)]$$

or

$$\sum_1^p k[\mu_{n_i}(A_i)] < h(\mu_{n_i}) < H$$

$$\sum_1^p k[\mu(A_i)] < H + \epsilon$$

qui est une borne indépendante de p donc $\sum \mu(A_i)$ est absolument convergente $\mu(B_p)$ converge donc vers S et l'on a :

$$\mu\left(\bigcup_1^{\infty} A_i\right) = \sum_1^{\infty} \mu(A_i)$$

4/ - Séries :

Théorème III - 5

Une forme linéaire de fonctionnelles σ -additives définit une fonctionnelle σ additive si la série $\sum |a_n| h(\mu_n)$ est convergente.

Posons :
$$\gamma_p = \sum_{n=1}^p a_n \mu_n$$

γ_p est une fonctionnelle - Sa hauteur $h(\gamma_p)$ est bornée car l'on a :

$$h(\gamma_p) \leq \sum_{n=1}^p |a_n| h(\mu_n)$$

en effet pour tout A

$$k(\gamma_p(A)) = k\left(\sum_{n=1}^p a_n \mu_n(A)\right) \leq \sum_{n=1}^p |a_n| k(\mu_n(A)) \leq \sum_{n=1}^p |a_n| h(\mu_n)$$

La suite γ_p est convergente si : $h(\gamma_p) < H$ quel que soit p $h(\gamma_p - \gamma_q) \rightarrow 0$ quels que soient p et q suffisamment grands.

Or si S converge vers une valeur H

$$h(\gamma_p) \leq \sum_{n=1}^p |a_n| h(\mu_n) < H \text{ quel que soit } p.$$

$$h(\gamma_p - \gamma_q) \leq \sum_{n=q}^p |a_n| h(\mu_n) \rightarrow 0$$

quand p et q augmentent indéfiniment en effet pour tout A

$$\begin{aligned} k(\gamma_p(A) - \gamma_q(A)) &= k\left(\sum_{n=q}^p a_n \mu_n(A)\right) \leq \sum_{n=q}^p |a_n| k(\mu_n(A)) \\ &\leq \sum_{n=q}^p |a_n| h(\mu_n) \end{aligned}$$

5/ - Produit d'une fonctionnelle par une fonction semi-étagée.

Soit une fonctionnelle semi-étagée f prenant ses valeurs y dans un groupe F muni d'une norme n(y) une fonctionnelle m prenant ses valeurs z dans un groupe distancié G. Soit une application de $F \times G$ dans H notée $y \times z = u$; H étant pourvu d'une norme notée |u| satisfaisant la relation $|y \times z| \leq n(y) k(z)$.

Nous supposons de plus que la multiplication $y \times z$ est distributive par rapport à l'addition dans G.

Il est possible de définir une fonctionnelle $f \times m$ prenant ses valeurs dans H, l'application $t = f \times m$ du produit de l'espace μ des fonctions f par l'espace θ des fonctionnelles m dans l'espace θ des fonctionnelles t étant distributive et continue par rapport à l'addition dans μ et continue par rapport à l'addition dans φ . Soit d'abord une fonction

f étagée. Pour toute classe P de C sur lequel $\text{Oscf} = 0$, c'est-à-dire $f = \lambda_P$, on prendra $t(p) = \lambda_P m(p)$. Toute classe Q de C est une réunion dénombrable de classes P ; en effet soit π une réunion en classes sur les classes de laquelle $\text{Oscf} = 0$. J'aurai :

$$E = \bigcup P_i \quad Q = E \cap Q \quad P_i \in \pi \\ = \bigcup (P_i \cap Q)$$

Or l'intersection de P_i et de Q est soit vide, soit P_i , soit Q. Si c'est Q c'est que Q est lui-même un P_i et la propriété est établie. Ecartons ce cas ; en appelant P_j les classes P_i qui sont incluses dans Q nous aurons : $Q = \bigcup_{j \in J} P_j$. Toute réunion A de classes de C est une réunion dénombrable de classes Q donc de P_i ou de $\bigcup_{j \in J} P_j$. Nous aurons : $A = \bigcup_{k \in K} P_k$.

Je prendrai comme valeur de la fonctionnelle pour Q :

$$t(Q) = \sum_j \lambda_{P_j} m(P_j)$$

Cette série est absolument convergente puisque $|\lambda_{P_j}|$ est borné et $\sum m(P_j)$ absolument convergente. $t(Q)$ est indépendant de la répartition car si je remplace une classe P_j par une somme de classes $P_{j'}$, j'aurai

$$\lambda_{P_j} \sum_{j'} m(P_{j'}) = \sum_{j'} \lambda_{P_{j'}} m(P_{j'})$$

Je prendrai comme valeur de la fonctionnelle pour A

$$t(A) = \sum_k \lambda_{P_k} m(P_k)$$

tout $|t(A)|$ et tout $|t(UA)|$ est majoré par :

$$\text{Sup. } [\sum_i (\lambda_{P_i}) k(m(P_i))] = h(t)$$

Nous avons fait correspondre à tout f étagée et à tout m une fonctionnelle $t = f \times m$. L'opération ainsi définie est distributive par rapport aux additions et séries commutativement convergentes.

$$\text{En effet soit } m = \sum_n m_n \quad t = \sum_n f \times m_n$$

$$\text{J'aurais } t(A) = \sum_n (\sum_k \lambda_{P_k} m_n(P_k))$$

Les 2 sommations étant commutativement convergentes on peut les intervertir et

$$t(Z) = \sum_k [\sum_n \lambda_{P_k} m_n(P_k)]$$

$$\begin{aligned}
 t(Z) &= \sum_k \left[\lambda_{p_k} \sum_{k \in I} m_n(P_k) \right] \\
 &= \sum_k \lambda_{p_k} m(P_k) \\
 &= f \times m(A)
 \end{aligned}$$

Les 2 fonctionnelles $f \times m$ sont identiques ou

$$\sum_n (f \cdot m_n) = f \cdot \left(\sum_n m_n \right) \quad \text{C. Q. F. D.}$$

(Si l'on suppose en outre la distributivité de l'opération $y \cdot z$ par rapport à l'addition dans F on démontre de la même façon la distributivité de l'opération $f \cdot m$ par rapport à l'addition dans $\mathcal{F}$).

Après avoir démontré la continuité de l'opération, par rapport à f nous pourrions l'étendre à la fermeture de l'ensemble des fonctions pour lequel elle est définie, c'est-à-dire à l'ensemble des fonctions semi-étagées.

Nous allons montrer que si $|f| < K h(f \cdot m) < K h(m)$ en effet nous avons vu que :

$$\begin{aligned}
 h(f \cdot m) &\leq \sup_{\pi} \left[\sum_{\substack{P \\ i \in \pi}} |\lambda_{p_i}| |m(P_i)| \right] \\
 &\leq \sup_{\pi} \left[\sum_{\substack{P \\ i \in \pi}} \sup |\lambda_{p_i}| |m(P_i)| \right] \\
 &\leq \sup_{\pi} K \sum |m(P_i)| = K h(m)
 \end{aligned}$$

Donc étant donné $g_n \rightarrow g$ et $\text{Max } g_n < K$ la suite t_n de fonctionnelles $t_n = g_n \cdot m$ est telle que $h(t_n) < K h(m)$

$h(t_p - t_q) \leq \sup |g_p - g_q| h(m)$ étant donné ε il est toujours possible de trouver $\eta = \frac{\varepsilon}{h(m)}$ tel que $|g_p - g_q| < \eta$

entraîne $h(t_p - t_q) < \varepsilon$ donc la suite t_n converge vers une valeur t que l'on posera égale à $g \cdot m$ et l'opération est continue par rapport à $\mathcal{F}$. Si $h(t_n) < H$ pour n suffisamment grand $H(t) < H$ or $h(t) < K h(m)$ pour $\sup |g| < K$.

L'opération $f \cdot m$ est continue par rapport à m en effet soit une suite convergente m_n de mesures donc

$$h(m_n) < H \text{ quel que soit } n$$

$$h(m_p - m_q) < \eta \text{ pour } p \text{ et } q$$

suffisamment grands et une fonction f bornée donc $|f| < K$ j'aurais $t_n = f. f. m_n$ donc

$$h(t_n) < K h(m_n) < K H$$

$$h(t_p - t_q) < K h(m_p - m_q) < K$$

On assurera $h(t_p - t_q) < \eta$ en prenant $\eta < \frac{\varepsilon}{K}$ donc la suite des t_n converge vers une valeur t et l'on aura $t = f. m$ en effet

$$f. m - t = \text{Lim}(f. m - f. m_n) = \text{Lim } \tau_n$$

en posant : $\tau_n = f. m - f. m_n = f(m - m_n)$

$$\tau_n \rightarrow 0 \text{ car } h(\tau_n) < K h(m - m_n) \text{ et } h(m - m_n) \rightarrow 0.$$

Il faudra faire attention si nous écrivons la fonctionnelle sous la forme $\int dm$ que le résultat du produit par f devra s'écrire $f. \int dm$ en utilisant l'opération sur $\varphi \times \mu$ ou $\int f dm$ en utilisant l'opération sur $F \times G$.

QUOTIENT de DEUX FONCTIONNELLES - ADDITIVES -

Soit deux fonctionnelles $m_1(A)$ et $m_2(A)$ prenant leurs valeurs sur un même corps G . Considérons la fonctionnelle $q(A)$ -

$$q(A) = m_1(A) m_2^{-1}(A)$$

Cette fonctionnelle est bornée pour tout ensemble pour lequel $m_2(A) \neq 0$. Considérons les ensembles A m_2 - négligeables, m_1 - non négligeables. Ils sont m_1 équivalents et m_2 équivalents aux sous-ensembles d'un ensemble H m_2 - négligeable m_1 - non négligeable ou m_1 irrégulier.

Considérons les classes P_i appartenant à une répartition régulière de $E \cap C H$. On a soit : $m_2(P_i) \neq 0$ soit P_i m_2 - négligeable.

Si les P_i sont m_2 - négligeables, ou ils sont m_1 - négligeables ou : $m_1(P_i) \neq 0$ mais dans ce dernier cas ils sont m_1 - équivalents à un sous-ensemble de H donc leur intersection avec $C H$ est m_1 - négligeable. Donc de toute façon ils sont m_1 - négligeables.

La valeur de la fonctionnelle m_2 pour les classes d'une répartition régulière de $E \cap C H$ est ou nulle ou différente de zéro. Si elle est nulle les P_i sont m_1 - négligeables et m_2 - négligeables. $q(P_i)$ est donc ou bien fini (si $q(P_i) \neq 0$) ou bien indéterminé (sur un ensemble H m_1 - négligeable, m_2 - négligeable).

On posera : $H' = E \cap C(H \cup H'')$. On pourra écrire pour tout ensemble N , réunion d'une partie classifiée de $H \cup H''$ et de classes de H' :

$$m_1(N) = m_1((H \cup H'') \cap N) + \sum_i m_1(P_i)$$

pour $P_i \subset N \subset H'$.

Le second terme du second membre peut s'écrire également :

$$\sum q(P_i) m_2(P_i).$$

Etant donnée une répartition en classes π de H' , nous pouvons définir une fonction $f_\pi(x)$ sur H' par :

$$f_\pi(x) = q(P_j) \text{ pour } x \in P_j \in \pi$$

Si nous considérons une suite de répartition régulière qui converge vers la classification qui sépare H' nous avons une suite de fonctions $f_1(x) \dots f_n(x) \dots$.

Je me propose de chercher si cette suite de fonctions converge vers une fonction $f(x)$, pour certaines suites de répartitions et si $g(x)$ est indépendant de ces suites. Examinons d'abord ce qui se passe en un point x_0 . Soit $L(x_0)$ la plus grande des limites des $f_\pi(x_0)$ et $l(x_0)$ la plus petite des limites des $f_\pi(x_0)$ quand on prend π arbitraire.

$L(x_0)$ et $l(x_0)$ existent nécessairement pour tout x_0 ce qui détermine 2 fonctions $L(x)$ et $l(x)$ à valeurs finies ou infinies. Il existe une infinité de valeurs de $q(P_j)$ dans l'intervalle $(L - \eta)$, $(L + \eta)$ (ou supérieures à $\frac{1}{\eta}$ si $L(x_0) = +\infty$) ; donc une infinité de classes P_j contenant toutes le point x_0 et convergent vers ce point x_0 . De même il existe une infinité de classes P'_j convergeant vers x_0 pour lesquelles les $q(P'_j)$ sont situés dans $l - \eta$, $l + \eta$ (ou inférieures à $-\frac{1}{\eta}$ si $l(x_0) = -\infty$).

Soit une partie classifiée M nous allons construire deux classifications π et π' ne contenant respectivement que des P_j et des P'_j .

Pour construire choisissons un point x_0 . Il existe une classe $P(x_0)$ dans M qui est un P_j .

Il existe un point $x_1 \in M \cap C P(x_0)$, donc une classe $P(x_1) \subset M \subset P(x_0)$.

En continuant indéfiniment nous arrivons nécessairement à définir une répartition en classes de M car sinon il y aurait des points x

de M qui n'appartiendraient pas à la réunion d'une famille dénombrable de classes et qui par conséquent appartiendraient au complémentaire d'une partie classifiée.

On pourrait donc trouver une nouvelle classe $P(x)$ ce qui permettrait de continuer la classification. De la même façon nous pouvons construire une répartition en classes π' de M dont les classes soient toutes des π'_j . Nous avons donc déterminé deux fonctions étagées $f_{\pi}(x)$ et $f_{\pi'}(x)$.

Quand η tend vers zéro $f_{\pi}(x)$ et $f_{\pi'}(x)$ tendent respectivement vers $L(x)$ et $l(x)$ qui sont donc semi-étagées. Il en est de même de $L(x) - l(x)$ et par suite l'ensemble des points x tels que :

$L(x) - l(x) > \varepsilon > 0$ est une partie classifiée K_ε . Pour tout A inclus dans K_ε on a :

$$\begin{aligned} \int_A (L(x) - l(x)) \, dm_2 &= \lim \int_A f_{\pi}(x) \, dm_2 - \lim \int_A f_{\pi'}(x) \, dm_2 \\ &= m_1(A) - m_2(A) = 0 \end{aligned}$$

Or on ne peut donc avoir

$$\begin{aligned} \int_A (L(x) - l(x)) \, dm &= 0 \\ L(x) - l(x) &> \varepsilon \quad A \subset K_\varepsilon \end{aligned}$$

que si K_ε est négligeable.

Comme sur tout A on a :

$$\begin{aligned} \int_A L(x) \, dm_2 &= \lim \int_A f_{\pi}(x) \, dm_2 = \lim \int_A dm_1 = \int_A dm_1 \\ \int_A l(x) \, dm_2 &= \lim \int_A f_{\pi'}(x) \, dm_2 = \lim \int_A dm_1 = \int_A dm_1 \end{aligned}$$

Les fonctions $L(x)$ et $l(x)$ sont finies (sauf peut être sur un ensemble m_2 - négligeable).

En tout point on a : $L = l$ (sauf peut-être sur un ensemble m_2 - négligeable) et $f_n(x)$ converge vers une fonction unique $f(x)$ (sauf peut-être sur un ensemble m_2 - négligeable).

Remarque :

1/ - Si l'on trouve deux suites de fonctions convergeant l'une vers g_1 l'autre vers g_2 on a $g_1 = g_2$ sauf sur un ensemble négligeable inclus dans K .

2/ - Si l'on complète $g(x)$ sur K par des valeurs bornées on a

$$\int_K g(x) dm_2 = 0 = \int_K dm_1 \text{ on pourra donc écrire}$$

$$\int_A dm_1 = \int_A g dm_2$$

sans s'occuper de K et par conséquent :

$$m_1(A) = \int_{CH' \cap A} dm_1 + \int_{A \cap H'} g dm_2$$

g est la dérivée de la fonctionnelle m_1 par rapport à la fonctionnelle m .
On peut convenir d'écrire symboliquement

$$dm_1 = g dm_2 \text{ sur } H$$

mais alors g n'est plus une fonction semi-étagée, c'est une distribution.

m_1 sera dite continue par rapport à m_2 si CH' est vide. Ce sera le cas si quel que soit A on a :

$$m_1(A) m_2^{-1}(A) \leq \Lambda$$

Toute fonctionnelle continue par rapport à une autre fonctionnelle possède par rapport à cette dernière une dérivée qui est une fonction semi-étagée, bornée.

Exemple :

dm_2 est la fonctionnelle définie par Maurice Girault (1)

dm_1 la fonctionnelle définie par :

$$m_1(A) = \text{Pr. } \{X \in A\}$$

X étant une variable aléatoire définie par :

$$X = X_1 + \dots + X_n$$

avec :

$$X_1 = \begin{cases} \frac{2}{3} & \text{probabilité } \frac{3}{4} \\ -\frac{2}{3} & \text{probabilité } \frac{1}{4} \end{cases}$$

(1) Page 245 -

$$X_n = \begin{cases} \frac{2}{3^n} \text{probabilité } \frac{1}{2} (1 + \frac{1}{2^n}) \\ \frac{2}{3^n} \text{probabilité } \frac{1}{2} (1 - \frac{1}{2^n}) \end{cases}$$

Au point : $x = \sum \frac{2 \epsilon_n}{3^n}$ (avec $\epsilon_n = \pm 1$)

$m_1 m_2^{-1}$ vaudra $\prod (1 + \frac{\epsilon_n}{2^n})$ qui est convergent quel que soit x sur l'ensemble triadique de Cantor sur $-1, +1$.

PRODUIT de FONCTIONNELLES σ -ADDITIVES -

Supposons que nous ayons une classification $C(F_A \cup F_B)$ qui résulte des classifications $C(F_A)$ et $C(F_B)$ sur F_A et F_B . Supposons que nous connaissions deux fonctionnelles σ additives m_A et m_B soient définis sur un même corps. Pour simplifier l'exposé et fixer les idées nous supposons également que m_A et m_B sont des mesures prenant leur valeur sur $0 \leq x \leq 1$ mais cette dernière hypothèse n'est pas indispensable.

Je dis qu'il existe une mesure σ sur les parties classifiées de $F_A \times F_B$ définies sur les produits $P_A \times P_B$ de parties classifiées de F_A et de F_B par : $\sigma(P_A \times P_B) = m_A(P_A) m_B(P_B)$.

Nous avons bien une fonctionnelle σ -additive. En effet si $P_B = P'_B \cup P''_B$ où P'_B et P''_B sont deux ensembles disjoints ; $P_A \times P'_B$ et $P_A \times P''_B$ sont aussi des ensembles disjoints et nous avons :

$$\begin{aligned} (P_A) \times (P'_B \cup P''_B) &= m_A(P_A) m_B(P'_B \cup P''_B) \\ &= m_A(P_A) (m_B(P'_B) + m_B(P''_B)) \\ &= m_A(P_A) m_B(P'_B) + m_A(P_A) m_B(P''_B) \\ &= \sigma(P_A \times P'_B) + \sigma(P_A \times P''_B) \end{aligned}$$

Nous avons la même démonstration en échangeant les rôles de P_A et de P_B .

La fonctionnelle est σ -additive car :

$$1/ - \text{elle est bornée ; en effet on a : } \sigma(N) + \sigma(CN) = 1$$

Or : $\sigma(CN) \geq 0$ puisque toute partie classifiée peut être obtenue à partir des limites de $Q_i = \bigcup P_j$.

Or : $\sigma(P_j) \geq 0$.

$$\sigma\left(\bigcup_j P_j\right) \geq 0$$

$$\sigma(Q_1) \geq 0$$

$$\lim \sigma(Q_1) \leq 0.$$

Donc $\sigma(N) \leq 1$.

2/ - Elle est continue.

En effet $P_A \times P_B \rightarrow \emptyset$ implique que l'un au moins de P_A et de P_B tend vers $\emptyset$ donc que sa mesure tend vers 0 ; donc le produit tend vers zéro.

Si $\bigcup(P_A \times P_B) \rightarrow \emptyset$, on a :

$$(\bigcup P_A) \times (\bigcup P_B) \rightarrow \emptyset$$

Or :

$$\sigma(\bigcup P_A \times P_B) \leq \sigma((\bigcup P_A) \times (\bigcup P_B)) \rightarrow 0$$

Pour une union de parties classifiées on peut grouper dans les $\bigcup P_A \times P_B$ tous les $P_{B_{j1}}$ qui se projettent sur le même P_{A_1} et écrire :

$$\sigma(P_{A_1} \times P_{B_{j1}}) = \sum (m(P_{A_1}) (\sum (P_{B_{j1}})))$$

En passant à la limite pour une partie classifiée M qui se projette en M_A sur F_A et dont on appelle M_x les intersections avec $\{x\} \subset F_B$, on pourra écrire

$$\sigma(M) = \int_{M_A} dm_A \int_{M_x} dm_B$$

Probabilisation d'un produit d'ensembles classifiés.

Nous dirons qu'un ensemble E est probabilisé s'il est classifié et si l'on a défini sur l'ensemble des parties classifiées une fonctionnelle σ -additive, m , satisfaisant aux conditions.

$$1/ - 0 \leq m(A).$$

$$2/ - m(E) = 1.$$

De ces deux conditions on déduit :

$$m(A) \leq 1.$$

la fonctionnelle est dite une distribution de probabilité sur E .

Nous dirons qu'un ensemble $E = E_1 \times E_2$ est probabilisé par rapport à E_1 et à E_2 s'il est classifié et si l'on a défini une fonctionnelle σ -additive m sur les parties classifiées de la classification satisfaisant à

$$1/ - 0 \leq m(A) \quad A \subset E_1 \times E_2$$

$$2/ - m(A_1 \times A_2) = m(A_1 \times E_2) m(E_1 \times A_2) \quad A_1 \subset E_1, \quad A_2 \subset E_2$$

Si m n'est pas la fonctionnelle nulle (ce qui serait un cas sans intérêt) 2' entraîne :

$$3/ - m(E) = 1.$$

En effet en faisant $A_1 = E_1$ $A_2 = E_2$ dans 2' il vient :

$$m(E_1 \times E_2) = m(E_1 \times E_2) m(E_1 \times E_2)$$

$m(E) \neq 0$ entraîne $m(E) = 1$ et par conséquent $m(A) \leq 1$.

Un produit de deux distributions de probabilité satisfait manifestement aux conditions 1/ - et 2/ -

Théorème III - 6 :

Nous allons montrer que s'il existe sur $E_1 \times E_2$ une fonctionnelle σ -additive m satisfaisant 1/ - et 2/ - il est possible de faire une application de $E_1 \times E_2$ sur $E'_1 \times E'_2$ telle que la valeur de la fonctionnelle m pour les origines des parties classifiées de $E'_1 \times E'_2$ soient identiques à la valeur d'un produit de deux distributions de probabilité E'_1 et E'_2 .

Considérons la fonctionnelle φ définie sur $E_1 \times E_2$ par :

$$\varphi(A_1 \times A_2) = m(A_1 \times A_2) m^{-1}(A_1 \times E_2)$$

Nous avons :

$$0 < \varphi < 1$$

En effet :

$$0 \leq m(A_1 \times A_2) \leq m(A_1 \times E_2)$$

nous pouvons considérer $\varphi(A_1 \times A_2)$ pour une valeur donnée de A_2 comme le quotient de deux fonctionnelles σ -additives sur E_1 :

$m(A_1 \times A_2)$ et $\mu = m(A_1 \times E_2)$ $m(A_1 \times A_2)$ étant continue par rapport à μ nous avons quand A_1 tend vers $\{x\}$:

$$\varphi(A_1 \times A_2) \longrightarrow N(x, A_2)$$

N étant une fonction bornée de x et une fonctionnelle par rapport à A_2 . Nous avons entre m , N et μ la relation :

$$m(A_1 \times A_2) = \int_{A_1} N(x, A_2) d\mu$$

$N(x, A_2)$ est une fonctionnelle additive de A_2 ;

en effet si $A_2 = A'_2 \cup A''_2$ avec A'_2 et A''_2 disjoints on aura :

$$\begin{aligned} \varphi(A_1 \times (A'_2 \cup A''_2)) &= m(A_1 \times (A'_2 \cup A''_2)) m^{-1}(A_1 \times E_2) \\ &= (m(A_1 \times A'_2) + m(A_1 \times A''_2)) m^{-1}(A_1 \times E_2) \\ &= m(A_1 \times A'_2) m^{-1}(A_1 \times E_2) + m(A_1 \times A''_2) m^{-1}(A_1 \times E_2) \\ \varphi(A_1 \times (A'_2 \cup A''_2)) &= \varphi(A_1 \times A'_2) + \varphi(A_1 \times A''_2) \end{aligned}$$

Cette égalité est vraie quel que soit A_1 donc à la limite on aura :

$$N(x, A'_2 \cup A''_2) = N(x, A'_2) + N(x, A''_2)$$

N est même une fonctionnelle σ -additive ; en effet :

$$0 \leq N(x, A_2) \leq 1. \text{ Et si } A_2 \longrightarrow \emptyset \quad N(A_1 \times A_2) \longrightarrow \emptyset \text{ donc : } \varphi(x, A_2) \longrightarrow 0.$$

Nous noterons : n_x la mesure sur E_2 définie au moyen de $N(x, A_2)$. Soit y un élément de E_2 et z l'élément générateur de E_2 nous pouvons poser :

$$S_x(y) = \int_{z < y} dn_x$$

Nous appellerons E' l'ensemble dont l'élément générateur t est défini au moyen de 2 nombres u et v compris entre 0 et 1. La relation d'ordre sur E' est définie par :

$$\begin{array}{lll} t_1 < t_2 & \text{si} & u_1 < u_2 \\ t_1 < t_2 & \text{si} & u_1 = u_2 \quad v_1 < v_2 \end{array}$$

La mesure de la classe $T(u_1, v_1) < t \leq (u_2, v_2)$ est définie par $m(T) = u_2 - u_1$; nous définissons la relation suivante entre t et y .

$$u = \theta ; \quad v = (y - y_1)(y_2 - y_1)^{-1} \text{ si } y_2 \neq y_1$$

$$v = 1 \quad \text{si} \quad y_2 = y$$

étant donné que l'ensemble des éléments y de E , tels que $s_x(y) = \theta$, peut se mettre sous la forme :

$$A(\theta, x, y) = \{y_1 \leq y \leq y_2\}$$

Nous désignerons par $t_x(y)$ la valeur de t ainsi définie. A deux valeurs différentes de y correspondent nécessairement 2 valeurs différentes de $t_x(y)$.

A tout $y \in E_2$ nous pouvons faire correspondre les ensembles $L_x(y)$, $M_x(y)$, $N_x(y)$ définis comme suit :

$L_x(y)$ est l'ensemble $0 \leq t \leq t_x(y)$

$M_x(y)$ est l'ensemble $\bigcup_{z < y} L_x(z)$

$N_x(y)$ est l'ensemble $L_x(y) \cap C M_x(y)$

$N_x(y)$ existe toujours ; en effet :

$$L_x(z) \subset L_x(y) \text{ pour } z < y$$

$$\bigcup_{z < y} L_x(z) \subset L_x(y)$$

Et : $t_x(y) \notin L_x(z)$ pour $z < y$

$$t_x(y) \notin \bigcup_{z < y} L_x(z)$$

Tout t_1 appartient à un $N_x(y)$ et à un seul. En effet nous pouvons répartir les y en deux classes non vides, la classe I dans laquelle nous mettrons tous les y tels que $t_x(y) < t_1$, la classe II dans laquelle nous mettrons tous les y tels que $t_x(y) \geq t_1$.

Supposons que la 1ère classe possède un élément y_0 plus grand que tous les autres. Nous aurons :

$$t_x(y_0) = t_0$$

Donc pour tout y de la classe II nous aurons :

$$t_x(y) \geq t_1$$

et par suite

$$\int_{+y_0}^y dn_x \geq t_1 - t_0$$

qui ne pourrait pas tendre vers zéro quand l'ensemble $y_0 < z \leq y$ tend vers $\emptyset$. La classe II contient donc un élément y_0 plus petit que tous les autres. Nous avons : $t_x(y_0) = t_0$
et par suite :

$$L_x(y_0) = \{0 \leq t \leq t_0\}$$

$$L_x(y) = \{0 \leq t \leq s_x(y)\} \quad \text{pour } y < y_0$$

avec $t_x(y) < t_1$. Donc : $M_x(y_0) = \bigcup_{y < y_0} L_x(y) \subset \{0 \leq t < t_1\}$

puisque tout $t_x(y) < t_1$ donc : $t_1 \notin M_x(y_0)$

et par suite $t_1 \in L_x(y_0)$.

Si à tout t $0 \leq t \leq 1$ nous faisons correspondre l'élément y tel que $t \in L_x(y)$ nous définissons une application de E_2^1 sur E_2 qui dépend de $x \in E_1$ et par conséquent une application de $E_1 \times E_2^1$ sur $E_1 \times E_2$.

Nous savons d'après la théorie des ensembles classifiés qu'il existe une classification sur $E_1 \times E_2^1$ dont les images des parties classifiées dans l'application sur $E_1 \times E_2$ sont les parties classifiées de $E_1 \times E_2$.

Soit une classe $A_1 \times A_2$ de $E_1 \times E_2$, image de la classe P de $E_1 \times E_2^1$. Nous posons :

$$\mu(P) = m(A_1 \times A_2) = \varphi(A_1 \times A_2) m(A_1 \times E_2)$$

Si $A_1 = \bigcup A_i$ les A_i étant des sous-classes de A_1 :

$$\mu(P) = \sum m(A_i \times A_2) = \sum \varphi(A_i \times A_2) m(A_1 \times E_2)$$

En prenant des répartitions de plus en plus fines de A_1 et en passant à la limite :

$$\mu(P) = \int_{A_1} N(x, A_2) dm$$

$$\text{Or : } N(x, A_2) = \int_{A_2} ds(y)$$

Or dans l'application de $E_1 \times E_2^1$ sur $E_1 \times E_2$ pour une valeur donnée de x , A_2 est l'image d'un ensemble $A_2'(x)$ et par suite nous pouvons écrire :

$$N(x, A_2) = \int_{A_2'(x)} dl$$

dl étant la mesure de Lebesgue de $A_2'(x)$.

En portant dans $\mu(P)$ il vient :

$$\mu(P) = \int_{A_1} dm \int_{A_2'(x)} dl$$

De la même façon que nous avons défini une application de E'_2 sur E_2 nous pouvons définir une application de E'_1 sur E_1 en prenant pour E'_1 l'ensemble $u = V, W$ et en faisant correspondre à tout x l'ensemble $N(x)$ défini au moyen des ensembles $L(x)$ $M(x)$ par :

$L(x)$ est l'ensemble $0 \leq u \leq v(x)$

$M(x)$ est l'ensemble $\bigcup_{z \leq x} L(z)$

$N(x)$ est l'ensemble $L(x) \cap^C M(x)$ avec :

$$v(x) = \int_{z \leq x} dm(A_1 \times E_2) \quad W(x) = (x - x_1)(x_2 - x_1)^{-1} \quad x_2 \neq x_1$$

$$W(x) = 1 \quad x_2 \neq x_1$$

On démontre comme précédemment que tout u appartient à un $N(x)$ et l'application ainsi définie nous donne :

$$\int_{A_1} dm = \int_{A'_1} du.$$

u étant la mesure de Lebesgues sur E'_1 et par conséquent :

$$\mu(P) = \int_{A'} du \int_{A'_2(x)} dt$$

de quel on peut écrire également $\mu(P) = \int_P d\sigma$; P étant l'image de P' . La propriété est donc bien démontrée. La généralisation des produits de fonctionnelles σ - additives et du théorème précédent à un produit de n ensembles classifiés se traite par récurrence et ne présente aucune difficulté partielle. Nous étendrons les résultats précédents à un espace fonctionnel après avoir traité les fonctionnelles π - multiplicatives.

FONCTIONNELLES SOUS-ADDITIVES -

Ce sont des fonctionnelles φ pour lesquelles on a $\varphi(A \cup B) \leq \varphi(A) + \varphi(B)$ quels que soient A et B . Les fonctionnelles σ - additives sont sous-additives. (La théorie des fonctionnelles sous-additives permet d'ailleurs d'étendre les fonctionnelles σ - additives à des sous-ensembles de E qui ne sont pas des parties classifiées, c'est la théorie classique des mesures extérieures). Mais il y a en outre des fonctionnelles sous-additives qui ne sont pas additives par exemple :

$$1/ - k(\varphi(A)), \quad \varphi(A) \text{ étant une fonctionnelle } \sigma - \text{ additives}$$

$$2/ - \left| \int_A f^p dm \right|^{1/q}$$

$$3/ - \sup_t m^t(A) = \varphi(A)$$

Je ne développerai pas la théorie des fonctionnelles sous-addi-

tives et me contenterai de dire que l'on peut définir des ensembles équivalents et des répartitions régulières par rapport à une fonctionnelle sous-additive. Si l'on cherche à définir le quotient d'une fonctionnelle sous-additive $g(A)$ par une fonctionnelle σ -additive $m(A)$ on trouve sous certaines conditions une fonction $f(x)$ semi-étagée mais l'on n'a pas $g(A) = \int_A f \, dm$.



IMP. LOUIS-JEAN — GAP (H.-A.)

Dépôt légal n° 122 — 1958