



HAL
open science

Contribution à l'étude des partages, probabilités et statistiques sur le simplexe unité

Michel Maurin

► **To cite this version:**

Michel Maurin. Contribution à l'étude des partages, probabilités et statistiques sur le simplexe unité. Modélisation et simulation. Université Joseph-Fourier - Grenoble I, 1983. Français. NNT: . tel-00307403

HAL Id: tel-00307403

<https://theses.hal.science/tel-00307403>

Submitted on 29 Jul 2008

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.

THESE

présentée à

l' Université Scientifique et Médicale de Grenoble

pour obtenir le grade de

**DOCTEUR INGENIEUR
Mathématiques appliquées
«option Statistique»**

par

MAURIN Michel



CONTRIBUTION A L'ETUDE DES PARTAGES, PROBABILITES ET

STATISTIQUES SUR LE SIMPLEXE UNITE.



Thèse soutenue le 25 octobre 1983 devant la commission d'examen.

B. VAN CUTSEM	Président
A. LE BRETON	} Examineurs
V. COHEN	

UNIVERSITE SCIENTIFIQUE ET MEDICALE DE GRENOBLE

Année universitaire 1982-1983

Président de l'Université : M. TANCHE

MEMBRES DU CORPS ENSEIGNANT DE L'U.S.M.G.

(RANG A)

SAUF ENSEIGNANTS EN MEDECINE ET PHARMACIE

PROFESSEURS DE 1ère CLASSE

ARNAUD Paul	Chimie organique
ARVIEU Robert	Physique nucléaire I.S.N.
AUBERT Guy	Physique C.N.R.S.
AYANT Yves	Physique approfondie
BARBIER Marie-Jeanne	Electrochimie
BARBIER Jean-Claude	Physique expérimentale C.N.R.S. (labo de magnétisme)
BARJON Robert	Physique nucléaire I.S.N.
BARNOUD Fernand	Biosynthèse de la cellulose-Biologie
BARRA Jean-René	Statistiques - Mathématiques appliquées
BELORISKY Elie	Physique
BENZAKEN Claude (M.)	Mathématiques pures
BERNARD Alain	Mathématiques pures
BERTRANDIAS Françoise	Mathématiques pures
BERTRANDIAS Jean-Paul	Mathématiques pures
BILLET Jean	Géographie
BONNIER Jean-Marie	Chimie générale
BOUCHEZ Robert	Physique nucléaire I.S.N.
BRAVARD Yves	Géographie
CARLIER Georges	Biologie végétale
CAUQUIS Georges	Chimie organique
CHIBON Pierre	Biologie animale
COLIN DE VERDIÈRE Yves	Mathématiques pures
CRABBE Pierre (détaché)	C.E.R.M.O.
CYROT Michel	Physique du solide
DAUMAS Max	Géographie
DEBELMAS Jacques	Géologie générale
DEGRANGE Charles	Zoologie
DELOBEL Claude (M.)	M.I.A.G. Mathématiques appliquées
DEPORTES Charles	Chimie minérale
DESRE Pierre	Electrochimie
DOLIQUE Jean-Michel	Physique des plasmas
DUCROS Pierre	Cristallographie
FONTAINE Jean-Marc	Mathématiques pures
GAGNAIRE Didier	Chimie physique

.../...

GASTINEL Noël	Analyse numérique - Mathématiques appliquées
GERBER Robert	Mathématiques pures
GERMAIN Jean-Pierre	Mécanique
GIRAUD Pierre	Géologie
IDELMAN Simon	Physiologie animale
JANIN Bernard	Géographie
JOLY Jean-René	Mathématiques pures
JULLIEN Pierre	Mathématiques appliquées
KAHANE André (détaché DAFCO)	Physique
KAHANE Josette	Physique
KOSZUL Jean-Louis	Mathématiques pures
KRAKOWIAK Sacha	Mathématiques appliquées
KUPTA Yvon	Mathématiques pures
LACAZE Albert	Thermodynamique
LAJZEROWICZ Jeannine	Physique
LAJZEROWICZ Joseph	Physique
LAURENT Pierre	Mathématiques appliquées
DE LEIRIS Joël	Biologie
LLIBOUTRY Louis	Géophysique
LOISEAUX Jean-Marie	Sciences nucléaires I.S.N.
LOUP Jean	Géographie
MACHE Régis	Physiologie végétale
MAYNARD Roger	Physique du solide
MICHEL Robert	Minéralogie et pétrographie (géologie)
MOZIERES Philippe	Spectrométrie - Physique
OMONT Alain	Astrophysique
OZENDA Paul	Botanique (biologie végétale)
PAYAN Jean-Jacques (détaché)	Mathématiques pures
PEBAY PEYROULA Jean-Claude	Physique
PERRIAUX Jacques	Géologie
PERRIER Guy	Géophysique
PIERRARD Jean-Marie	Mécanique
RASSAT André	Chimie systématique
RENARD Michel	Thermodynamique
RICHARD Lucien	Biologie végétale
RINAUDO Marguerite	Chimie CERMAV
SENGEL Philippe	Biologie animale
SERGERAERT Francis	Mathématiques pures
SOUTIF Michel	Physique
VAILLANT François	Zoologie
VALENTIN Jacques	Physique nucléaire I.S.N.
VAN CUTSEN Bernard	Mathématiques appliquées
VAUQUOIS Bernard	Mathématiques appliquées
VIALON Pierre	Géologie

PROFESSEURS DE 2ème CLASSE

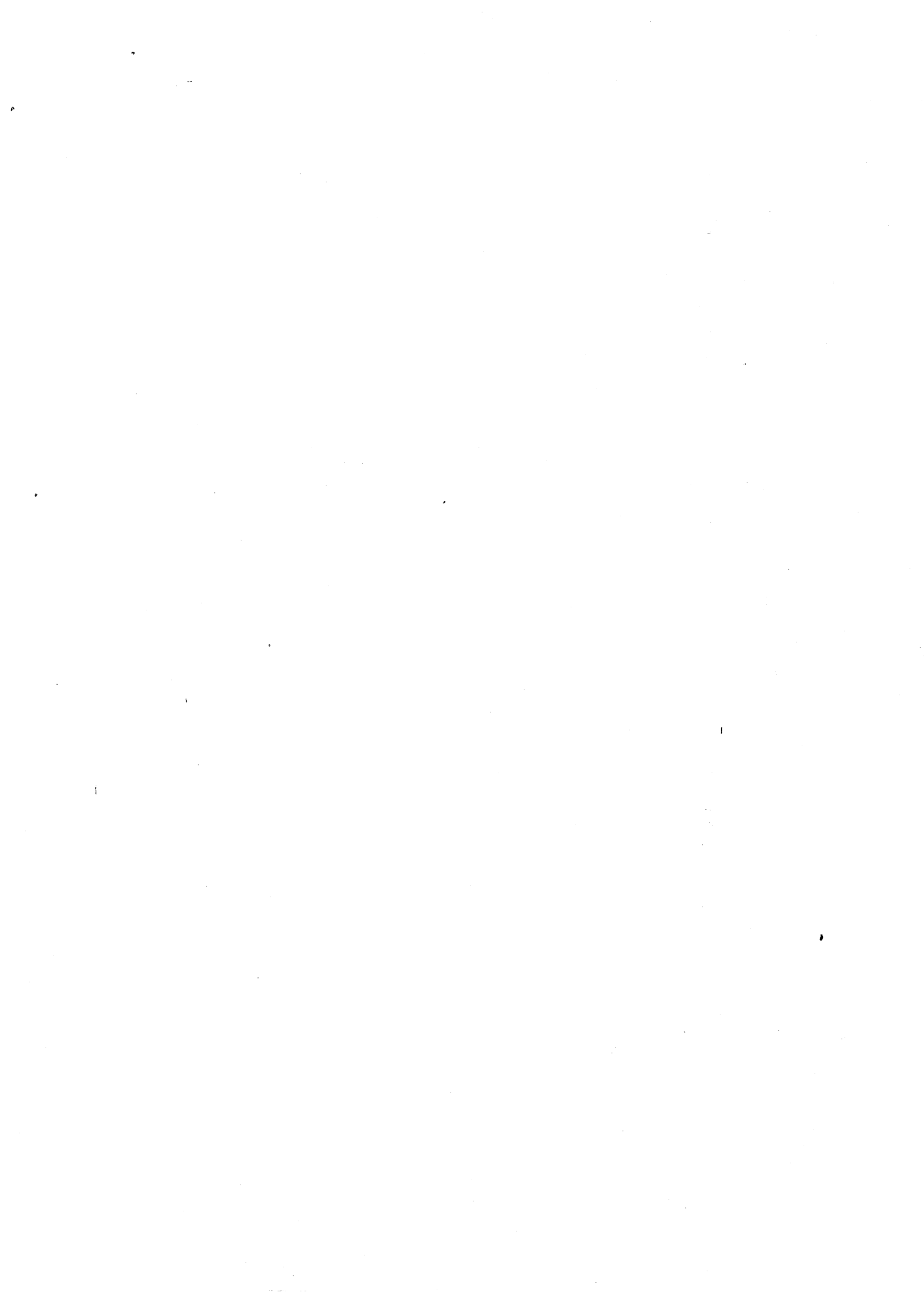
ADIBA Michel	Mathématiques pures
ARMAND Gilbert	Géographie

.../...

AURIAULT Jean-Louis	Mécanique
BEGUIN Claude (M.)	Chimie organique
BOEHLER Jean-Paul	Mécanique
BOITET Christian	Mathématiques appliquées
BORNAREL Jean	Physique
BRUN Gilbert	Biologie
CASTAING Bernard	Physique
CHARDON Michel	Géographie
COHENADDAD Jean-Pierre	Physique
DENEUVILLE Alain	Physique
DEPASSEL Roger	Mécanique des fluides
DOUCE Roland	Physiologie végétale
DUFRESNOY Alain	Mathématiques pures
GASPARD François	Physique
GAUTRON René	Chimie
GIDON Maurice	Géologie
GIGNOUX Claude (M.)	Sciences nucléaires I.S.N.
GUITTON Jacques	Chimie
HACQUES Gérard	Mathématiques appliquées
HERBIN Jacky	Géographie
HICTER Pierre	Chimie
JOSELEAU Jean-Paul	Biochimie
KERCKOVE Claude (M.)	Géologie
LE BRETON Alain	Mathématiques appliquées
LONGEQUEUE Nicole	Sciences nucléaires I.S.N.
LUCAS Robert	Physiques
LUNA Domingo	Mathématiques pures
MASCLE Georges	Géologie
NEMOZ Alain	Thermodynamique (CNRS - CRTBT)
OUDET Bruno	Mathématiques appliquées
PELMONT Jean	Biochimie
PERRIN Claude (M.)	Sciences nucléaires I.S.N.
PFISTER Jean-Claude (détaché)	Physique du solide
PIBOULE Michel	Géologie
PIERRE Jean-Louis	Chimie organique
RAYNAUD Hervé	Mathématiques appliquées
ROBERT Gilles	Mathématiques pures
ROBERT Jean-Bernard	Chimie physique
ROSSI André	Physiologie végétale
SAKAROVITCH Michel	Mathématiques appliquées
SARROT REYNAUD Jean	Géologie
SAXOD Raymond	Biologie animale
SOUTIF Jeanne	Physique
SCHOOL Pierre-Claude	Mathématiques appliquées
STUTZ Pierre	Mécanique
SUBRA Robert	Chimie
VIDAL Michel	Chimie organique
VIVIAN Robert	Géographie

*A SABINE, PATRICE, ANNE-FLORENCE et BENEDICTE
pour toutes les heures de vie familiale que
cette thèse leur a dérobées;*

*A TAMISE et ses semblables qui n'allument pas
des buchers mais qui protègent la flamme.*



Je suis très sincèrement reconnaissant envers Monsieur le Professeur B. VAN CUTSEM qui m'a accueilli dans son équipe, qui a dirigé cette thèse et qui m'a fait l'honneur de présider le jury.

Je remercie Messieurs les Professeurs A. LE BRETON de Grenoble, et V. COHEN de l'Ecole Nationale des Ponts et Chaussées qui ont accepté d'être membres du jury.

Je remercie Monsieur l'Ingénieur Général des Ponts et Chaussées J. BONITZER qui a suivi le déroulement de ce travail.

Mes remerciements s'adressent également à Monsieur le Professeur J. FONLUPT qui m'a donné l'occasion de nouer des contacts avec l'IRMA;

à Monsieur l'Ingénieur en Chef des Ponts et Chaussées C. LAMURE chargé du Centre d'Evaluation et de Recherche des Nuisances et de l'Energie de l'Institut de Recherche des Transports;

à Monsieur l'Ingénieur des Ponts et Chaussées M. PRUVOST qui m'a prodigué ses conseils et encouragements aux moments utiles;

à Monsieur le Professeur L. TARTAR pour ses renseignements amicaux;

à Madame C. ARPINO qui mit beaucoup de dévouement et de qualité pour la dactylographie;

au Centre Informatique Recherche de l'I.R.T. sur le matériel duquel les applications numériques ont été faites;

à l'U.E.R. de Mathématiques de l'Université de Lyon 1, où je pus enrichir ma documentation;

au service de reprographie qui s'est chargé de l'exécution du tirage de cette thèse.

TABLE DES MATIERES

Introduction	1
PREMIERE PARTIE : <u>Distributions absolument continues</u> <u>sur le simplexe unit�</u>	
I - <u>Etude g�om�trique sur le simplexe</u>	7
1 - G�n�ralit�s	7
2 - Etude de λ_1	7
3 - Etude de λ_M	11
4 - Etude conjointe de λ_1 et λ_M	12
II - <u>Lois de λ_k, $k = 1, \dots, M$ avec la mesure uniforme sur le simplexe unit�</u>	16
1 - G�n�ralit�s	16
2 - Mesure uniforme sur S_1^M	17
3 - Historique et commentaires	24
4 - Lois limites quand M tend vers plus l'infini	30
5 - Mesures uniformes sur \bar{E}_T , $T > 1/M$.	34
III - <u>Lois conjointes des couples $\lambda_1 \lambda_2$ et $\lambda_k \lambda_M$ avec la mesure uniforme</u>	36
1 - Distribution du couple $\lambda_1 \lambda_M$	36
2 - Distribution du couple $\lambda_k \lambda_M$, $k = 2, \dots, M-1$	37
3 - Le simplexe ordonn� et ses densit�s	39
4 - Distribution du couple $\lambda_1 \lambda_2$	42
IV - <u>Autres mesures sur simplexe unit�</u>	49
1 - G�n�ralit�s	49
2 - Calculs avec les $\mu(E_1^{i_1} \cap \dots \cap E_k^{i_k})$, $k = 2, \dots, M-1$	50

IV A - Lois fermées pour la convolution sur \mathbb{R}_+^* , et mesures sur le simplexe	52
3 - La normalisation par la somme	52
4 - Distributions gamma et de DIRICHLET	57
5 - Distribution de λ_1 avec une distribution de DIRICHLET	63
6 - Distribution de λ_M et $\lambda_1 \lambda_2$ avec une distribution de DIRICHLET	68
IV B - Changements de variables sur le simplexe unité	71
7 - Etude de l'application $\bar{\varphi}$	72
8 - Calcul du Jacobien	75
9 - Application aux densités de probabilité	79
IV C - Autres questions	81
10- Densités associées	81
11- Inégalité en fonction de l'effectif M	85
12- Densité de DIRICHLET inverses et/ou décentrées	87
13- Autres normalisations par la somme	92
14- Autres densités sur S_1^M	96

DEUXIEME PARTIE :

Distributions discontinues

Introduction	99
V - <u>Distributions discontinues sur le simplexe unité</u>	100
1 - Discussion arithmétique	100
2 - La distribution π_{MN}^1	105
3 - Exemples	123
4 - Autres distributions discontinues	144

TROISIEME PARTIE :
Structures statistiques sur le simplexe unit 

Introduction	153
<u>VI - La structure invariante de DIRICHLET</u>	154
1 - G�n�ralit�s	154
2 - Etude de $\sigma_M(y)$ sur S_1^M	155
3 - Etude de la densit� invariante par Σ_M	156
4 - La structure statistique invariante de DIRICHLET	167
5 - Exemples, simulations de partages al�atoires	172
<u>VII - La structure g�n�rale de DIRICHLET</u>	175
1 - Etude de la densit�	175
2 - La structure statistique de DIRICHLET	181
3 - Le partage du corps �lectoral fran�ais pour les pr�sidentielles 1981	206
<u>VIII - Distributions multinomiales et de DIRICHLET</u>	210
1 - Distributions conjugu�es	210
2 - Identit�s num�riques	211
3 - Convergence entre les distributions	214
4 - Autre convergence vers une distribution de DIRICHLET	218
<u>VIII bis - Autres structures, Structure de FISHER VON-MISES, de DIRICHLET "sph�rique"</u>	223

QUATRIEME PARTIE :
Compléments et applications

Introduction	225
<u>IX - Applications des parts ordonnées et extrémales, fonctions définies sur simplexe unité</u>	228
1 - Etude de ϕ	228
2 - Bornes de ϕ fonction de λ_M	230
3 - Bornes de ϕ fonction de λ_1	233
4 - Applications aux probabilités	240
<u>X - Applications de la distribution de DIRICHLET</u>	251
1 - Distributions gamma et DIRICHLET invariante	251
2 - Application à l'analyse de la variance	253
3 - Structures dérivées	263
<u>ANNEXES</u>	
1) Une identité entre les coefficients du triangle de PASCAL	271
2) Tables de probabilité pour $\phi_{up}(n_i)$, pour $n \in \bar{S}_N^M$ avec la distribution équiprobable	275
3) Convergence presque - sûre de la distribution multinomiale	291
Références bibliographiques	295

Les problèmes de partage sont en effet fort anciens, et ils n'ont jamais cessé d'enrichir l'actualité mathématique. La notion si banale de partage pourrait d'ailleurs servir de fil conducteur à une explication méthodique de l'histoire des mathématiques.

G. Th. GUILBAUD - Les problèmes de partage.

Les partages font partie de la vie quotidienne de chacun, et certains d'entre eux sont devenus célèbres par le soin qui les entourèrent - par exemple le Traité de Verdun qui découpe l'empire Carolingien en trois parts égales - ou par les disputes qui s'ensuivirent - celles d'Achille et d'Agamemnon, de Clovis et du Sicambre, de Pizarre et d'Almagro..... (*). Dans son analyse des partages liée à la théorie des jeux, G. Th. Guilbaud a reconnu cette présence et confère au "partage" un rôle de ferment dans la recherche mathématique. La question que posa le chevalier de Méré à Pascal n'est-elle pas un bel exemple pour illustrer ce point de vue ?

La présente étude est consacrée aux partages aléatoires. Le partage d'un "objet" entre M acteurs ($M \geq 2$) consiste à découper ce dernier en autant de parts qu'il y a de participants, et à affecter l'une d'elles à chacun d'eux. Nous supposons que cet objet et que chacune des parts susceptibles d'intervenir lors d'un partage sont mesurables ; de la sorte si A est la mesure de l'objet partagé et a_i ; $i=1, \dots, M$ les mesures des parts, les rapports a_i/A sont des nombres réels sans dimension, positifs ou nuls de somme égale à 1. Il est naturel de faire correspondre à tout partage le point y dont pour tout $i=1, \dots, M$ la coordonnée y_i est égale au rapport a_i/A . Par conséquent à

(*) Les héritiers de Louis le pieux ne manquèrent pas de reprendre leurs disputes peu de temps après leur accord de Verdun.

tout partage entre M participants il correspond un point image y du simplexe unité S_1^M de l'espace euclidien réel \mathbb{R}^M ; ci-après nous nous limitons à cette représentation géométrique et nous considérons que le simplexe de \mathbb{R}^M est l'ensemble des partages entre M acteurs qui sont susceptibles de se produire. En ce qui concerne "l'aléa" qui s'applique aux partages, nous envisageons le simplexe muni de la tribu de ses boréliens $\mathcal{B}(S_1^M)$, et la restriction \mathcal{P} au simplexe unité de la mesure uniforme sur l'hyperplan qui le supporte. L'espace de probabilité $(S_1^M, \mathcal{B}(S_1^M), \mathcal{P})$ est donc au centre de nos considérations.

Cette étude a été menée autour de trois directions. Dans un partage on peut s'intéresser en particulier à la plus grande part, c'est à dire selon notre représentation à la coordonnée maximale $\sup y_i$; c'est une variable aléatoire au même titre que le partage. L'étude de la loi de probabilité de $\sup y_i$, et des lois des autres parts ordonnées constituent la première direction. Les considérations des partages aléatoires conduisent naturellement à étudier les lois de probabilité sur le simplexe unité, ce qui constitue la deuxième direction de la thèse. Ces deux directions relèvent du Calcul des Probabilités; en revanche la troisième est Statistique, elle correspond à l'étude des structures statistiques sur simplexe unité. En ce qui concerne les deux axes probabilistes il est cependant plus commode de regrouper ce qui touche aux distributions absolument continues d'un côté, aux distributions discontinues d'un autre. Pour cette raison ce travail est rédigé en quatre parties. La première reprend tout ce qui concerne les deux premières directions à propos des distributions absolument continues, la seconde en fait de même pour les distributions discontinues, la troisième correspond à la direction statistique et la dernière reprend divers compléments.

Dans la première partie, la distribution de la variable $\sup y_i$ peut être facilement étudiée à l'aide de considérations géométriques sur le simplexe unité. Sur celui-ci la distribution la plus simple est sans doute la distribution uniforme, la loi de $\sup y_i$ est alors presque immédiate. Déjà Witworth en 1887 la connaissait et peut être date-t-elle d'une époque plus ancienne; elle a été redécouverte à plusieurs reprises dont ici même. Dans cette partie on s'appli-

que également à l'étude des lois des autres coordonnées du point y quand elles sont ordonnées par ordre de valeur décroissante, (avec la distribution uniforme sur le simplexe). Ces résultats sont dus à la géométrie et à la combinatoire et sont pour la plupart établis de longue date; lorsque l'on fait tendre l'effectif des parts M vers plus l'infini on a divers résultats que l'on rappelle également.

Naturellement la distribution de $\sup y_i$ dépend de la distribution sur le simplexe, et ce qui a été fait avec la distribution uniforme pourrait être recommencé avec d'autres. De façon générale on note la place particulière des distributions invariantes par le groupe \sum_M des permutations de M objets. Les considérations géométriques précédentes restent valables, mais conduisent à des calculs plus complexes. Dans ces conditions l'on s'est limité à étudier quelques méthodes générales pour acquérir des distributions absolument continues sur le simplexe unité sans chercher cependant à épuiser cet ensemble. A cette occasion on rencontre les distributions de Dirichlet (associées aux densités de même nom), et on introduit les distributions décentrées de Dirichlet.

Dans la deuxième partie on fait de même pour certaines distributions discontinues sur le simplexe unité, définies sur un sous-ensemble fini du simplexe, invariant par le groupe \sum_M et indicé par l'entier N . L'étude de la plus grande coordonnée nécessite cette fois-ci une discussion arithmétique, et l'on est à même d'établir la loi de la plus grande coordonnée pour plusieurs distributions, la distribution multinomiale invariante par \sum_M par exemple. Le cas le plus commode est ici celui où les points du sous-ensemble sont équiprobables, on peut alors manuellement faire une tabulation de la loi suivie par $\sup y_i$ pour différentes valeurs de M et N ; quelques unes de ces tables sont données en annexe.

On indique également d'autres distributions discontinues sur le simplexe, (conformément à la deuxième direction de la thèse), et quelques résultats asymptotiques quand on fait tendre le paramètre N vers plus l'infini.

Dans les deux cas continu et discontinu, les lois de variables réelles positives qui sont fermées pour la convolution conduisent à des lois sur le simplexe qui sont relativement commodes à manipuler.

La troisième partie et la troisième direction sont Statistiques. Une distribution paramétrée conduit à la structure statistique correspondante sur le simplexe. Nous abordons ici l'analyse de la structure de Dirichlet associée aux distributions et aux densités de même nom. On envisage deux cas, celui des distributions invariantes par le groupe \sum_M et celui des distributions

générales, qui dépendent respectivement de 1 et M paramètres. Le résultat essentiel est relatif à la nature de ces structures, ce sont des structures exponentielles, avec toutes les conséquences statistiques attachées à cette propriété. On indique les statistiques exhaustives, des estimateurs sans biais de variance minimale, les estimateurs du maximum de vraisemblance et quelques tests sur les valeurs des paramètres. Les propriétés statistiques reposent en fait sur des lemmes mathématiques relatifs à des propriétés des fonctions eulériennes Γ et $\psi = \Gamma'/\Gamma$; dans le cas invariant c'est la formule de duplication de Gauss qui intervient, dans le cas général on utilise sur un résultat d'Hermite. Les calculs numériques auxquels on est conduit quand on dispose d'un échantillon de partages sont faciles à programmer. On étudie également la structure statistique multinomiale, elle aussi exponentielle et en outre conjuguée avec la structure de Dirichlet. La littérature nous donne quelques exemples d'échantillons comme celui de l'analyse du sérum de canetons blancs de Pékin, ou de dépôts de pollen sur des fossiles.

Dans la quatrième partie, nous abordons divers compléments. On doit rappeler que la distribution uniforme sur le simplexe unité est une distribution que l'on rencontre très couramment en Statistique quand on s'intéresse aux statistiques d'ordre. Ceci renforce donc son intérêt en sus de la relative facilité des calculs auxquels elle conduit. La littérature s'est intéressée à des

variables du type $S = \sum_{i=1}^M \varphi(y_i)$ où y_i $i=1, \dots, M$

sont les coordonnées d'un point du simplexe unité, et φ une application réelle de $[0, 1]$ dans \mathbb{R} . Lorsque φ est convexe (ou concave), on indique des bornes B_i et B_A de S qui ne dépendent que de $\sup y_i$, (ou de $\inf y_i$).

On en déduit les lois de B_i et B_A quand on connaît celles de $\sup y_i$:

(ou de $\inf y_i$; on démontre ici que la variable $\sup y_i$ possède davantage d'information de Shannon que $\inf y_i$, et qu'il vaut donc mieux travailler avec $\sup y_i$). Dans l'analyse de la variance sur des plans d'expérience, on rencontre un vecteur de données \mathcal{X} dont toutes les coordonnées suivent une loi gamma décentrée à l'exception d'une seule qui suit une loi gamma. A partir de ces coordonnées on peut construire un point sur simplexe unité qui suit une loi de Dirichlet décentrée. Les tests d'hypothèse reviennent

alors à tester une loi de Dirichlet décentrée contre une loi de Dirichlet. En dernier on peut modifier le paramétrage de la loi générale de Dirichlet de façon à mettre en évidence une direction dans le cône positif de l'espace des paramètres et un scalaire de concentration (comme c'est le cas par exemple dans la loi de Fisher von-Mises) ; ce nouveau paramétrage conduit à de nouvelles structures statistiques sur le simplexe qui sont elles aussi exponentielles.

PREMIERE PARTIE :Distributions absolument continues
sur le simplexe unit I - ETUDE GEOMETRIQUE SUR LE SIMPLEXE UNITE.I.1 - GENERALITES.

Les distributions des parts d'un partage peuvent  tre  tablies   l'aide de consid rations g om triques pr alables sur le simplexe unit  de tout espace \mathbb{R}^n .

Soit l'espace affine d'origine O associ    l'espace euclidien \mathbb{R}^n , $n \geq 2$, rapport    une base orthonorm e \vec{e}_i , $i = 1, \dots, n$, et le rep re correspondant $O\vec{e}_i$, $i = 1, \dots, n$ de l'espace affine ; le point y de cet espace a pour coordonn es y_i , $i = 1, \dots, n$. Soit S_1^n le simplexe unit 

$$S_1^n = \left\{ y \in \mathbb{R}^n : y_i \geq 0 \quad i = 1, \dots, n, \quad \sum_{i=1}^n y_i = 1 \right\}.$$

On note e^i le point du simplexe   l'extr mit  du vecteur $O\vec{e}_i$ qui a toutes ses coordonn es nulles   l'exception de la coordonn e num ro i  gale   1, et

$\omega = \sum_{i=1}^n e^i / n$ le point du simplexe qui a toutes ses coordonn es  gales   $1/n$.

A tout point $y \in S_1^n$ on associe les coordonn es ordonn es $\lambda_1 \geq \dots \geq \lambda_n \geq 0$.
Pour simplifier, nous poserons  ventuellement λ la plus grande coordonn e sup y_i et δ la plus petite inf y_i .

I.2 - Etude de λ_1 .

2.1. Pour tout i appartenant   l'ensemble $\{1, 2, \dots, n\}$ et tout t r el compris entre z ro et un, on d finit les ensembles

$$E_t^i = \left\{ y \in S_1^n : y_i \geq t \right\},$$

$$E_t = \bigcup_{i=1}^n E_t^i.$$

(*) La notation habituelle pour le classement des grandeurs ordonn es se fait par ordre croissant, contrairement aux notations ci-contre.

Notations habituelles: $\lambda_{(1)} \leq \dots \leq \lambda_{(n)}$, la r gle de correspondance avec les notations ci-contre est:

$$(\nu) + k = n + 1.$$

On a les propriétés évidentes suivantes : Pour tout $i = 1, \dots, M$ et $t \in [0, 1]$, E_t^i est inclus dans S_1^M ; c'est l'homothétique de S_1^M par l'homothétie de centre e^i et de rapport $1-t$; E_t^i est donc un polyèdre convexe. On a aussi pour tout $i = 1, \dots, M$ et tous $t, t' \in [0, 1]$, $t' \geq t$,

$$E_{t'}^i \subset E_t^i .$$

2.2. Proposition 1 L'ensemble E_t vérifie :

$$E_t = \left\{ y \in S_1^M : \lambda_1 \geq t \right\} .$$

Démonstration : Tout point de E_t appartient à un $E_t^{\bar{i}}$ au-moins, on a donc $y_{\bar{i}} \geq t$, c'est à dire $\lambda_1 \geq t$.

Réciproquement, soit un point y de S_1^M qui vérifie $\lambda_1 \geq t$. Le maximum $\sup y_i$ est atteint par un indice \bar{i} au-moins, y appartient donc à $E_t^{\bar{i}}$ donc à E_t ■

Proposition 2 Pour tous $t, t' \in [0, 1]$ qui vérifient $t' \geq t$ on a la relation d'inclusion suivante: $E_{t'} \subset E_t$, pour tout $t \leq \frac{1}{M}$ on a $E_t = S_1^M$.

Cela résulte des propriétés de E_t^i $i = 1, \dots, M$ et du fait que l'on a nécessairement $\lambda_1 \geq 1/M$. ■

2.3. Soit μ une mesure définie sur le simplexe unité, absolument continue par rapport à la mesure uniforme de LEBESGUE P sur S_1^M . Dans la suite nous utiliserons la mesure $\mu(E_t)$. On rappelle la formule générale suivante (ref. 24,68)

$$\mu(E_t) = \sum_{k=1}^M (-1)^{k-1} \sum_{i_1, \dots, i_k} \mu(E_t^{i_1} \cap \dots \cap E_t^{i_k})$$

dans laquelle \sum_{i_1, \dots, i_k} est la sommation étendue à toutes les combinaisons de

M indices pris k à k sans répétition. Ceci nous conduit à étudier les intersections $E_t^{i_1} \cap \dots \cap E_t^{i_k}$ de M ensembles E_t^i distincts pris k à k , pour k variant de 1 à M .

Proposition 3 Pour tout entier k appartenant à $\{2, \dots, M\}$, pour tout t inférieur ou égal à $1/k$, et tout sous-ensemble d'indices i_1, \dots, i_k différents l'intersection $E_t^{i_1} \cap \dots \cap E_t^{i_k}$ est non vide.

En effet, le point $M^{i_1 \dots i_k} = (e^{i_1} + \dots + e^{i_k})/k$

est un point évident de l'intersection pour tout $t \leq 1/k$ ■

Remarque: Par extension la propriété est vraie pour $k=1$, et pour tout E_t^i , $i=1, \dots, M$.

Théorème 1 Pour tout entier k appartenant à $\{2, \dots, M\}$, et tout sous-ensemble d'indices différents i_1, \dots, i_k , l'intersection

$E_t^{i_1} \cap \dots \cap E_t^{i_k}$ est :

- vide si $1 - kt < 0$;
- réduite à $M^{i_1 \dots i_k}$ si $1 - kt = 0$;
- l'homothétique de S_1^M dans l'homothétie de centre $M^{i_1 \dots i_k}$ et de rapport $1 - kt$ si $1 - kt$ est strictement positif.

Démonstration: Les points 1 (proposition 3) et 2 sont immédiats. Pour le point 3, soient y un point de S_1^M , et \bar{y} son image par l'homothétie ci-dessus. On a

$$\bar{y} - M^{i_1 \dots i_k} = (1 - kt) \cdot (y - M^{i_1 \dots i_k})$$

$$\bar{y} = (1 - kt) y + kt M^{i_1 \dots i_k}$$

On désigne par I l'ensemble des indices $\{i_1, \dots, i_k\}$, et J son complémentaire dans $\{1, \dots, M\}$, on a

$$\bar{y}_i = (1 - kt) \sum_J y_j e^{i_j} + \sum_I ((1 - kt) y_j + t) e^{i_j}$$

Les coordonnées de \bar{y} sont toutes positives ou nulles de somme égale à 1, en outre les coordonnées \bar{y}_i sont supérieures ou égales à t .

Réciproquement, soit \bar{y} un point de $E_t^{i_1} \cap \dots \cap E_t^{i_k}$, et soit y son image par l'homothétie inverse :

$$y = M^{i_1 \dots i_k} + (\bar{y} - M^{i_1 \dots i_k}) / (1 - kt)$$

$$y = \sum_J \frac{y_j}{1-kt} e^j + \sum_I \frac{y_i - t}{1-kt} e^i ,$$

y a toutes ses coordonnées positives ou nulles, de somme égale à 1. ■

Conséquence $E_t^i \cap \dots \cap E_t^k$ est un polyèdre convexe.

2.4. Soit pour tout t compris entre zéro et un : $\bar{E}_t = \{y \in S_1^M : \lambda_i \leq t\}$,
c'est un sous-ensemble de S_1^M non vide pour tout $t \geq \frac{1}{M}$. Pour tous $t, t' \in [0, 1]$
qui vérifient $t' \geq t$ on a l'inclusion $\bar{E}_t \subset \bar{E}_{t'}$.

Les deux ensembles E_t et \bar{E}_t ont une intersection non vide, incluse dans
un nombre fini de variétés de dimension $M-2$ dans S_1^M , donc de mesures P et μ
nulles, pour tout $t \in [0, 1]$,

$$P(E_t) + P(\bar{E}_t) = 1.$$

Théorème 2 Pour $t \in]\frac{1}{M}, \frac{1}{M-1}]$, \bar{E}_t est l'homothétique de S_1^M par l'homothétie
de centre ω et de rapport $1-tM$; \bar{E}_t est réduit à ω pour $t = 1/M$.

Démonstration: Le dernier point est immédiat;

soit t tel que $\frac{1}{M} < t < \frac{1}{M-1}$, et soient $y \in S_1^M$ et \bar{y} son image par l'homothétie

ci-dessus:
$$\bar{y} = M t \omega + (1 - M t) y ,$$

$$\bar{y} = \sum_{i=1}^M (t + (1 - M t) y_i) e^i .$$

Les coordonnées \bar{y}_i ont une somme égale à 1, et vérifient les inégalités

$$0 \leq 1 + t(1 - M) \leq \bar{y}_i \leq t , \quad i=1, \dots, M .$$

Réciproquement, soit \bar{y} un point de \bar{E}_t et soit y son image par l'homothétie
inverse:

$$y = \omega + (\bar{y} - \omega) / (1 - M t) ,$$

$$y = \frac{1}{1 - M t} \sum_{i=1}^M (y_i - t) e^i .$$

Sachant que $\frac{1}{n} < t \leq \frac{1}{n-1}$ et que $\bar{y}_i \leq t$, les coordonnées y_i sont positives ou nulles, de somme égale à 1. ■

1.3 - Etude de λ_M .

3.1. Soit pour tout u appartenant à l'ensemble $[0, 1]$,

$$G_u = \left\{ y \in S_1^M : \lambda_M \geq u \right\}$$

c'est un sous-ensemble de S_1^M , non vide pour $0 \leq u \leq \frac{1}{M}$, identique à S_1^M pour $u=0$, réduit à ω pour $u=1/M$.

Pour tous $u, u' \in [0, \frac{1}{M}]$ qui vérifient $u \geq u'$ on a l'inclusion

$$G_u \subset G_{u'}.$$

3.2.

Théorème 3 Pour tout $u \in [0, 1/M]$, G_u est l'homothétique de S_1^M par l'homothétie de centre ω et de rapport $1 - Mu$; pour $u=1/M$, G_u est réduit à ω .

Démonstration: Le dernier point est immédiat; pour $0 \leq u < \frac{1}{M}$, soient y un point de S_1^M , et \bar{y} son image par l'homothétie ci-dessus

$$\bar{y} = Mu \omega + (1 - Mu) y,$$

$$\bar{y} = \sum_{i=1}^M (u + (1 - Mu) y_i) e^i.$$

Les coordonnées \bar{y}_i sont supérieures ou égales à u , de somme égale à 1.

Réciproquement, soient \bar{y} un point de G_u , et y son image par l'homothétie inverse:

$$y = \omega + (\bar{y} - \omega) / (1 - Mu),$$

$$y = \sum_{i=1}^M \frac{\bar{y}_i - u}{1 - Mu} e^i.$$

Les coordonnées \bar{y}_i sont positives ou nulles, de somme égale à 1. ■

Conséquence: G_u est un polyèdre convexe.

3.3. Soit pour tout u compris entre zéro et un : $\bar{G}_u = \{y \in S_1^M, \lambda_M \leq u\}$

Les deux ensembles G_u et \bar{G}_u ont une intersection non vide, incluse dans un nombre fini de variétés de dimension $M-2$ dans S_1^M , donc de mesures $\bar{\mu}$ et μ nulles, pour tout $u \in [0, 1]$

$$P(G_u) + P(\bar{G}_u) = 1.$$

I.4. - Etude conjointe de λ_1 et λ_M .

4.1. Les points y du simplexe qui vérifient simultanément $\lambda_1 \geq t$ et $\lambda_M \geq u$ sont ceux du sous-ensemble $E_t \cap G_u$ qui est égal à

$$G_u \cap \left(\bigcup_i E_t^i \right) = \bigcup_i (G_u \cap E_t^i),$$

la mesure μ de cet ensemble s'exprime comme suit:

$$\mu(G_u \cap E_t) = \sum_{k=1}^M (-1)^{k-1} \sum_{i_1 \dots i_k} \mu(G_u \cap E_t^{i_1} \cap \dots \cap E_t^{i_k}).$$

4.2. Nous étudions les intersections $G_u \cap E_t^{i_1} \cap \dots \cap E_t^{i_k}$.

Les points extrémaux de G_u sont: $\Delta_u^i = \Pi_u \omega + (1 - \Pi_u) e^i$, $i=1, \dots, M$;

ceux de $E_t^{i_1} \cap \dots \cap E_t^{i_k}$ sont de deux sortes:

α) ceux qui appartiennent à l'hyperplan engendré par e^{i_1}, \dots, e^{i_k}

$$\bar{e}^i = (1 - kt) e^i + kt M^{i_1 \dots i_k} \quad i \in I = \{i_1 \dots i_k\};$$

β) les autres, qui constituent un ensemble vide si et seulement si $k=M$

$$\bar{e}^j = (1 - kt) e^j + kt M^{i_1 \dots i_k} \quad j \in J = \bar{I}.$$

Proposition 4 L'intersection $G_u \cap E_t^{i_1} \cap \dots \cap E_t^{i_k}$ est non vide pour

$$1 - (M-k)u - kt \geq 0, \quad k=1, \dots, M.$$

Démonstration: Soit pour tout $k=1, \dots, M-1$

$$\begin{aligned} \Delta_u^{i_1 \dots i_k} &= \sum_I \Delta_u^i \\ &= \Pi_u \omega + (1 - \Pi_u) M^{i_1 \dots i_k} \end{aligned}$$

un point de G_u , et $\bar{M}^{i_1 \dots i_k} = \sum_J e^j = M \frac{1-kt}{M-k} \omega + k \frac{M-1}{M-k} M^{i_1 \dots i_k}$

un point de $E_t^{i_1} \cap \dots \cap E_t^{i_k}$.

Les 4 points ω , $M^{i_1 \dots i_k}$, $\Delta_u^{i_1 \dots i_k}$ et $\bar{M}^{i_1 \dots i_k}$ sont alignés et l'on a

$$\bar{M}^{i_1 \dots i_k} \in G_u \quad \text{si} \quad M u \leq M \frac{1 - k t}{M - k} .$$

Pour $k=M$ l'intersection est non vide si et seulement si ni G_u ni $E_t^{i_1} \cap \dots \cap E_t^{i_M}$ ne sont vides, c'est à dire si et seulement si $u \leq 1/M$ et $t \leq 1/M$;

la condition $1 - (M-k)u - kt \geq 0$ s'applique encore. ■

Théorème 4 Pour k entier appartenant à $\{1, \dots, M\}$ et $t \geq 1/M \geq u$, et tout sous-ensemble d'indices différents i_1, \dots, i_k l'intersection

$G_u \cap E_t^{i_1} \cap \dots \cap E_t^{i_k}$ est:

- vide si $1 - (M-k)u - kt < 0$,
- réduite à $\Delta_u^{i_1 \dots i_k}$ si $1 - (M-k)u - kt = 0$,
- homothétique de G_u par l'homothétie de centre $\Delta_u^{i_1 \dots i_k}$ et de rapport $\frac{1 - (M-k)u - kt}{1 - Mu}$ si $1 - (M-k)u - kt > 0$.

Démonstration: 1°) on étudie le cas $u < 1/M$.

α) le premier point résulte de la proposition 4;

β) les deux points $\Delta_u^{i_1 \dots i_k}$ et $\bar{M}^{i_1 \dots i_k}$ sont confondus quand $1 - (M-k)u - kt = 0$;

- avec $k < M$, c'est le seul point commun à G_u et $E_t^{i_1} \cap \dots \cap E_t^{i_k}$ puisque $\Delta_u^{i_1 \dots i_k}$ et $\bar{M}^{i_1 \dots i_k}$ sont respectivement situés sur des faces de ces ensembles convexes dont les faces sont portées par des hyperplans de dimensions k et $M-k$ non parallèles,

- avec $k=M$, ce cas ne se présente pas si $1 - Mu > 0$.

γ) soient y un point de G_u ($y_i \geq u$), et \bar{y} son image par l'homothétie ci-dessus

$$\bar{y} - \Delta_u^{i_1 \dots i_k} = \frac{1 - (M-k)u - kt}{1 - Mu} (y - \Delta_u^{i_1 \dots i_k}) ,$$

$$\bar{y} = \frac{1 - (M-k)u - kt}{1 - Mu} y + \frac{k(t-u)}{1 - Mu} \Delta_u^{i_1 \dots i_k} ,$$

les coordonnées \bar{y}_i ont pour somme 1, on a également:

$$\bar{y} = \sum_J \frac{[1-(n-k)u - kt] y_j + ku(t-u)}{1-\pi u} e^j + \sum_I \frac{[1-(n-k)u - kt] y_i + ku(t-u) + (1-\pi u)(t-u)}{1-\pi u} e^i,$$

on vérifie que

$$(1-(n-k)u - kt) y_j + ku(t-u) \geq u(1-\pi u),$$

puisque l'on a $(1-(n-k)u - kt)(y_j - u) \geq 0$, et que

$$(1-(n-k)u - kt) y_j + ku(t-u) + (1-\pi u)(t-u) \geq t(1-\pi u) \geq u(1-\pi u),$$

puisque la première inégalité se déduit de la précédente en remarquant que

$$(u + t - u)(1-\pi u) \geq t(1-\pi u) \quad ;$$

donc le point \bar{y} appartient à $G_u \cap E_t^{i_1} \cap \dots \cap E_t^{i_k}$.

Réciproquement, soient \bar{y} un point de $G_u \cap E_t^{i_1} \cap \dots \cap E_t^{i_k}$, et y son image par l'homothétie inverse

$$\frac{1-(n-k)u - kt}{1-\pi u} y = \bar{y} - \frac{k(t-u)}{1-\pi u} \delta_u^{i_1 \dots i_k},$$

les coordonnées y_n ont une somme égale à 1, on a encore

$$\frac{1-(n-k)u - kt}{1-\pi u} y = \sum_J \left(\bar{y}_j - \frac{ku(t-u)}{1-\pi u} \right) e^j + \sum_I \left(\bar{y}_i - \frac{ku(t-u)}{1-\pi u} - (t-u) \right) e^i ;$$

on a bien $y_j \geq u$ puisque l'inégalité $\bar{y}_j - \frac{ku(t-u)}{1-\pi u} \geq u \frac{1-(n-k)u - kt}{1-\pi u}$

est vérifiée lorsque $\bar{y}_j \geq u$;

et on a bien $y_i \geq u$ puisque l'inégalité $\bar{y}_i - \frac{ku(t-u)}{1-\pi u} - (t-u) \geq u \frac{1-(n-k)u - kt}{1-\pi u}$

est vérifiée lorsque $\bar{y}_i \geq t$.

Donc le point y appartient à G_u .

2°) cas $u = 1/M$.

la quantité $1-(n-k)u - kt$ est égale à $k(1/n - t)$;

- lorsque $t > 1/M$ l'intersection $G_u \cap E_t^{i_1} \cap \dots \cap E_t^{i_k}$ est vide,

- lorsque $t = 1/M$ l'intersection se réduit à $G_{1/n} = \omega$. ■

4.3. Application.

On en déduit notamment

$$P(G_u \cap \bar{E}_t) = P(G_u) - P(G_u \cap E_t)$$

relation également valable pour toute mesure μ absolument continue par rapport à P .

II - LES LOIS DE λ_k $k=1, \dots, M$ AVEC LA MESURE UNIFORME SUR LE SIMPLEXE UNITE.

II.1. - Généralités.

i) Pour toute mesure μ absolument continue par rapport à P , et pour tout sous-ensemble \mathcal{D} P -mesurable, on définit la probabilité suivante sur le simplexe unité de \mathbb{R}^M

$$P_{\text{prob}}(\mathcal{D}) = \mu(\mathcal{D}) / \mu(S^M_1)$$

ii) Nous envisageons pour commencer la mesure uniforme P elle-même; c'est le cas le plus simple et celui qui a été le plus étudié. Nous verrons au chapitre IV comment on peut l'obtenir en pratique, et aussi combien cette distribution se rencontre couramment en Statistique. Le cas de la mesure uniforme sur \overline{E}_t pour tout $t > 1/M$ est un autre cas qui conduit à quelques résultats simples.

iii) La considération des coordonnées y_k et des coordonnées ordonnées λ_k d'un point y du simplexe unité permet d'introduire les sous-ensembles \mathcal{D} suivants:
- Si λ_k est la k^{e} coordonnée de y ordonnée par ordre décroissant, l'ensemble $\{y : \lambda_k \geq t\}$ est réalisé lorsque k coordonnées au-moins sont supérieures ou égales à t , c'est à dire lorsque k sous-ensembles $E_t^{i_k}$ au-moins sont simultanément réalisés. La mesure d'un tel ensemble s'exprime à l'aide d'une formule combinatoire classique. En effet si l'on pose

$$S_{p,t} = \sum_{i_1, \dots, i_p} \mu(E_t^{i_1} \cap \dots \cap E_t^{i_p})$$

on a le résultat (ref. 24 tome 1, chap IV.5) :

$$\mu\{y : \lambda_k \geq t\} = \sum_{d=k}^M (-1)^{d-k} C_{d-1}^{k-1} S_{d,t}$$

- De façon analogue on définit l'ensemble $E_{k,t}$ comme étant l'ensemble qui correspond au fait que k et exactement k coordonnées sont supérieures ou égales à t , c'est à dire lorsque k et exactement k sous-ensembles $E_t^{i_k}$ sont réalisés. On sait que sa mesure s'exprime à l'aide d'une autre formule combinatoire (ref. 24 tome 1, chap IV.3) :

$$\mu\{E_{k,t}\} = \sum_{d=k}^M (-1)^{d-k} C_d^k S_{d,t}$$

II.2. - Mesure uniforme sur S_1^M .

Pour tout t appartenant à $[0, 1/p]$, les considérations géométriques et les homothéties du chapitre 1 nous donnent:

$$P(E_t^{i_1} \cap \dots \cap E_t^{i_p}) = (1-pt)^{n-1} \cdot P(S_1^n),$$

$$P(S_{pt}) = (1-pt)^{M-1} \cdot C_n^p \times P(S_1^n),$$

les intersections $E_t^{i_1} \cap \dots \cap E_t^{i_p}$ étant évidemment des ensembles P-mesurables.

2.2.1. Loi de λ_k , $k=1, \dots, M$.

Théorème 5 Avec la mesure uniforme sur le simplexe unité, et pour tout t compris entre 0 et 1, la fonction de répartition de la variable λ_k a la valeur suivante:

$$P_{\text{rob}}(\lambda_k \leq t) = 1 - k C_n^k \sum_{d=k}^* (-1)^{d-k} C_{n-k}^{d-k} \frac{(1-st)^{n-1}}{d},$$

où \sum^* est la sommation étendue à l'ensemble des entiers d qui vérifient $k \leq d \leq \inf(M, [1/t])$.

En effet, pour $0 \leq t \leq 1/k$ on a:

$$P_{\text{rob}}(\lambda_k \geq t) = \frac{P\{y: \lambda_k \geq t\}}{P\{S_1^n\}} = \sum_{d=k}^* (-1)^{d-k} C_{d-1}^{k-1} C_n^d (1-st)^{n-1},$$

la sommation étant limitée aux termes tels que $1-st \geq 0$, c'est à dire $d \leq [1/t]$ sachant que l'on a aussi $d \leq M$. On a encore:

$$\begin{aligned} P_{\text{rob}}(\lambda_k \geq t) &= \frac{M!}{(k-1)!} \sum_{d=k}^* (-1)^{d-k} \frac{(1-st)^{M-1}}{d (M-d)! (d-k)!} \\ &= k C_n^k \sum_{d=k}^* (-1)^{d-k} C_{n-k}^{d-k} \frac{(1-st)^{M-1}}{d} \end{aligned}$$

Dans cette expression le monôme $(1-st)^{M-1}$ n'apparaît que lorsque t appartient à l'intervalle $[0, 1/d]$, on observe donc que la variable λ_k ne prend ses valeurs que sur $[0, 1/k]$ et par conséquent la probabilité $P_{\text{rob}}(\lambda_k \leq t)$ est égale à 1 sur $[1/k, 1]$.

Pour tout t compris entre 0 et 1, la densité de probabilité f_k de la variable λ_k a la valeur suivante:

$$f_k(t) = k(M-1) C_M^k \sum_{d=k}^x (-1)^{d-k} C_{n-k}^{d-k} (1-t)^{M-2},$$

d'où les moments :

$$\begin{aligned} E_{\lambda_k}(t^q) &= k(M-1) C_M^k \sum_{d=k}^M (-1)^{d-k} C_{n-k}^{d-k} \int_0^{1/d} t^q (1-t)^{n-2} dt, \\ &= k C_M^k / C_{n+q-1}^q \times \sum_{d=k}^M (-1)^{d-k} \frac{C_{n-k}^{d-k}}{d^{q+1}}. \end{aligned}$$

En particulier on a l'espérance mathématique :

$$\bar{\lambda}_k = C_{n-1}^{k-1} \sum_{d=k}^M (-1)^{d-k} C_{M-k}^{d-k} / d^2.$$

2.2.2.

On en déduit notamment les distributions particulières des parts extrémales λ_1 et λ_M . On note $L_M^{\lambda_1}(\cdot)$ et $L_M^{\lambda_M}(\cdot)$ les fonctions de répartition des variables λ_1 et λ_M .

Corollaires 1) Pour tout t compris entre 0 et 1, la fonction de répartition de la variable λ_1 a la valeur suivante:

$$L_M^{\lambda_1}(t) = \sum_{p=1}^* (-1)^p C_M^p (1-t)^{M-1},$$

Les variations de $L_M^{\lambda_1}$ sont dessinées et tabulées dans les figures et tableaux ci-joints.

On vérifie que la fonction $L_M^{\lambda_1}$ est identiquement nulle sur $[0, 1/M]$.

En effet $L_M^\lambda(t)$ est le polynôme $\sum_{p=1}^M (-1)^p C_M^p (1-t)^{M-1}$ dont les coefficients $(-1)^q C_{M-1}^q \sum_{p=1}^M (-1)^p C_M^p t^q$ sont nuls (réf. 24,54,69).

2) Les moments de λ_1 sont :

$$E_{\lambda_1}(t^q) = \frac{C_M^1}{C_{M+q-1}^q} \sum_{d=1}^M (-1)^{d-1} \frac{C_{M-1}^{d-1}}{\Delta^{q+1}}$$

$$= \frac{1}{C_{M+q-1}^1} \sum_{d=1}^M (-1)^{d-1} \frac{C_M^d}{\Delta^q}$$

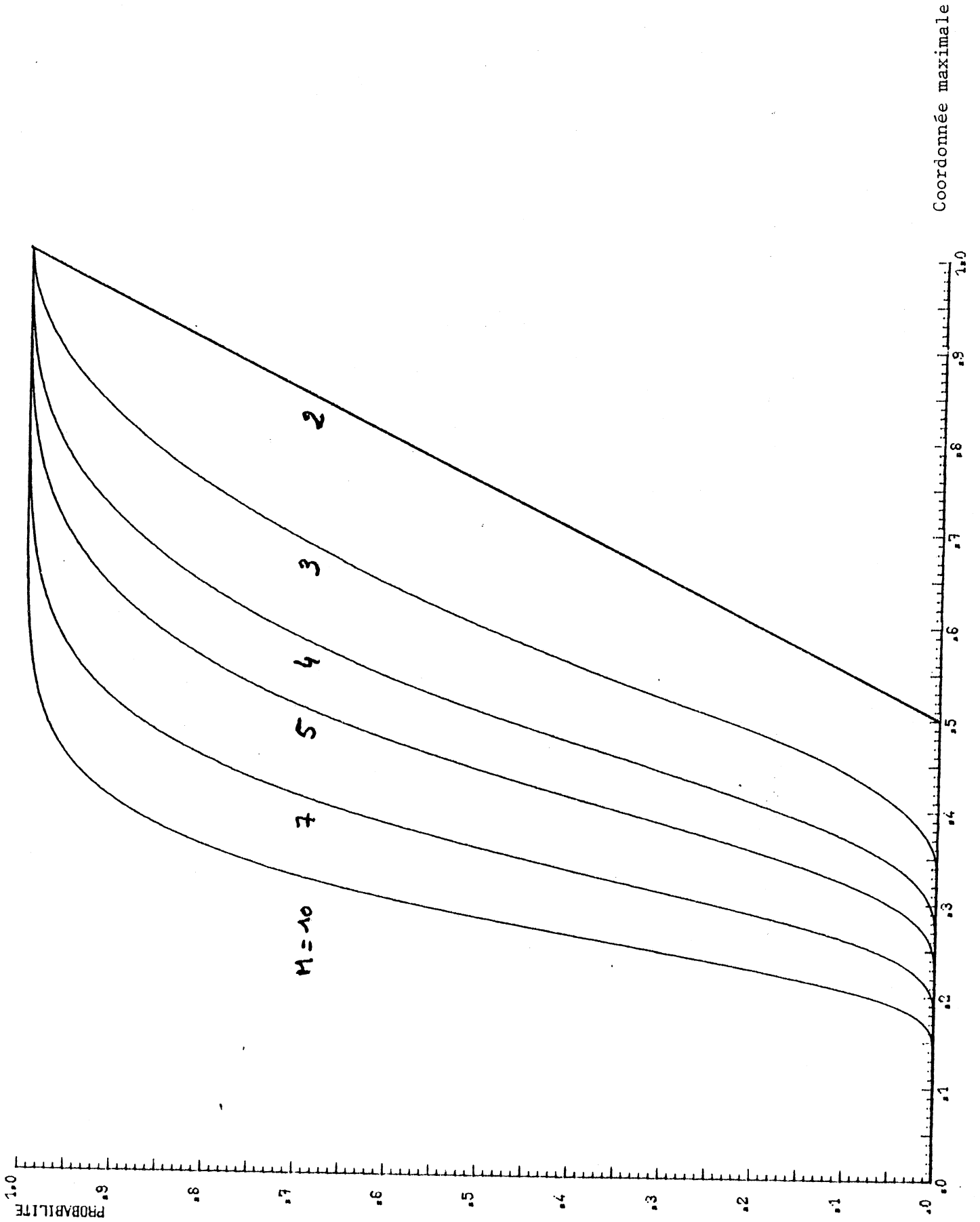
et l'espérance

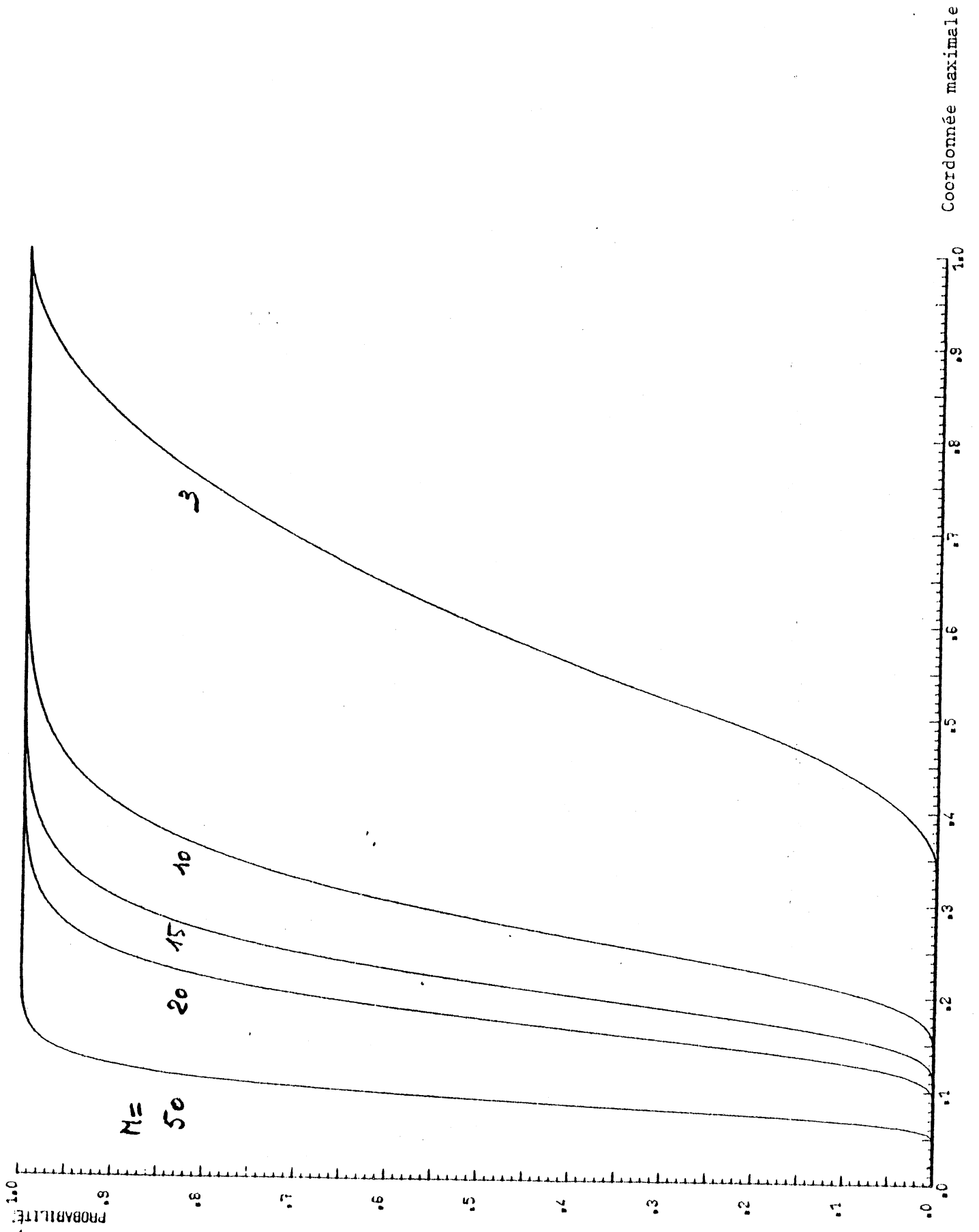
$$\bar{\lambda}_1 = \frac{1}{M} \sum_{d=1}^M (-1)^{d-1} \frac{C_M^d}{\Delta}$$

3) La fonction de répartition de la variable aléatoire λ_M a la valeur suivante:

$$\left\{ \begin{array}{l} L_M^\lambda(t) = 1 - (1 - Mt)^{M-1} \quad t \in [0, 1/M] \\ L_M^\lambda(t) = 1 \quad t \in [1/M, 1] \end{array} \right.$$

Remarque : ce résultat peut être établi directement à l'aide du théorème 3.





FONCTION DE REPARTITION DE LA PLUS GRANDE COORDONNEE
D'UN POINT DU SIMPLEXE UNITE DE RM - MESURE UNIFORME.

X	M= 2	3	4	5	6
.01	.0000000	.0000000	.0000000	.0000000	.0000000
.03	.0000000	.0000000	.0000000	.0000000	.0000000
.05	.0000000	.0000000	.0000000	.0000000	.0000000
.07	.0000000	.0000000	.0000000	.0000000	.0000000
.09	.0000000	.0000000	.0000000	.0000000	.0000000
.11	.0000000	.0000000	.0000000	.0000000	.0000000
.13	.0000000	.0000000	.0000000	.0000000	.0000000
.15	.0000000	.0000000	.0000000	.0000000	.0000000
.17	.0000000	.0000000	.0000000	.0000000	.0000000
.19	.0000000	.0000000	.0000000	.0000000	.0000021
.21	.0000000	.0000000	.0000000	.0000000	.0000515
.23	.0000000	.0000000	.0000000	.0000065	.0011837
.25	.0000000	.0000000	.0000000	.0005074	.0074691
.27	.0000000	.0000000	.0005117	.0039063	.0253915
.29	.0000000	.0000000	.0040965	.0148019	.0601490
.31	.0000000	.0000000	.0138223	.0377304	.1127563
.33	.0000000	.0000000	.0327679	.0749168	.1803983
.35	.0000000	.0025003	.0634999	.1260766	.2580768
.37	.0000000	.0121004	.1052679	.1884676	.3402741
.39	.0000000	.0289001	.1559624	.2580479	.4223590
.41	.0000000	.0528987	.2134746	.3311322	.5009711
.43	.0000000	.0840986	.2756908	.4046282	.5738787
.45	.0000000	.1224985	.3404978	.4760401	.6397902
.47	.0000000	.1680993	.4057868	.5434676	.6981781
.49	.0000000	.2208986	.4694427	.6056044	.7490932
.51	.0199986	.2796980	.5294020	.6617404	.7929841
.53	.0599988	.3372986	.5847068	.7117582	.8305136
.55	.0999987	.3924988	.6354990	.7560151	.8623924
.57	.1399986	.4452988	.6819706	.7949677	.8892825
.59	.1799986	.4956979	.7243147	.8290591	.9117943
.61	.2199993	.5436987	.7627229	.8587109	.9304857
.63	.2599993	.5892989	.7973871	.8843274	.9458652
.65	.2999985	.6324986	.8284989	.9062914	.9583933
.67	.3399980	.6732984	.8562511	.9249682	.9684865
.69	.3799983	.7116988	.8808351	.9407035	.9765185
.71	.4199986	.7476985	.9024433	.9538235	.9828224
.73	.4599987	.7812989	.9212673	.9646356	.9876932
.75	.4999985	.8124989	.9374994	.9734277	.9913906
.77	.5399984	.8412990	.9513316	.9804685	.9941406
.79	.5799985	.8676991	.9629557	.9860078	.9961382
.81	.6199987	.8916991	.9725637	.9902759	.9975495
.83	.6599983	.9132993	.9803478	.9934838	.9985144
.85	.6999984	.9324993	.9864998	.9958239	.9991481
.87	.7399985	.9492994	.9912119	.9974687	.9995444
.89	.7799983	.9636995	.9946759	.9985719	.9997773
.91	.8199984	.9756996	.9970840	.9992679	.9999034
.93	.8599985	.9852997	.9986280	.9996720	.9999646
.95	.8999984	.9924998	.9995000	.9998800	.9999899
.97	.9399983	.9972998	.9998920	.9999688	.9999982
.99	.9799982	.9997000	.9999960	.9999960	.9999999
				1.0000000	1.0000000

FONCTION DE REPARTITION DE LA PLUS GRANDE COORDONNEE
D'UN POINT DU SIMPLEXE UNITE DE RM - MESURE UNIFORME.

X	M= 7	8	9	10	11
.01	.0000000	.0000000	.0000000	.0000000	.0000000
.03	.0000000	.0000000	.0000000	.0000000	.0000000
.05	.0000000	.0000000	.0000000	.0000000	.0000000
.07	.0000000	.0000000	.0000000	.0000000	.0000000
.09	.0000000	.0000000	.0000000	.0000000	.0000000
.11	.0000000	.0000000	.0000000	-.0000022	.0000020
.13	.0000000	.0000038	.0000030	.0000237	.0001550
.15	.0000022	.0000150	.0002056	.0011893	.0042250
.17	.0000468	.0007140	.0037483	.0118888	.0278772
.19	.0012391	.0069001	.0215140	.0484668	.0888215
.21	.0086170	.0289485	.0661777	.1201463	.1875012
.23	.0306828	.0756111	.1409608	.2210042	.3086701
.25	.0737312	.1474612	.2383268	.3365855	.4341037
.27	.1379734	.2377353	.3462128	.4527423	.5505807
.29	.2183982	.3369055	.4536717	.5598128	.6513159
.31	.3077961	.4363264	.5532318	.6529124	.7343705
.33	.3992308	.5298480	.6409664	.7306655	.8006222
.35	.4873753	.6139442	.7155805	.7937655	.8522338
.37	.5688217	.6871291	.7774106	.8439054	.8917347
.39	.6417370	.7492779	.8276601	.8831123	.9215465
.41	.7054492	.8010785	.8678916	.9133785	.9437787
.43	.7600819	.8436369	.8997185	.9364859	.9601762
.45	.8062555	.8782072	.9246394	.9539461	.9721373
.47	.8448497	.9060227	.9439660	.9670022	.9807622
.49	.8768254	.9282067	.9588086	.9766582	.9869053
.51	.9031101	.9457416	.9700900	.9837157	.9912227
.53	.9245450	.9594699	.9785697	.9888086	.9942141
.55	.9418733	.9701062	.9848663	.9924331	.9962544
.57	.9557501	.9782543	.9894806	.9949740	.9976227
.59	.9667490	.9844195	.9928135	.9967262	.9985235
.61	.9753686	.9890214	.9951832	.9979127	.9991046
.63	.9820397	.9924054	.9968387	.9987004	.9994711
.65	.9871320	.9948528	.9979733	.9992118	.9996966
.67	.9909596	.9965905	.9987342	.9995359	.9998316
.69	.9937874	.9977990	.9992324	.9997357	.9999099
.71	.9958362	.9986200	.9995498	.9998550	.9999537
.73	.9972880	.9991632	.9997458	.9999238	.9999774
.75	.9982910	.9995117	.9998627	.9999619	.9999895
.77	.9989638	.9997277	.9999295	.9999820	.9999955
.79	.9993997	.9998559	.9999660	.9999921	.9999982
.81	.9996707	.9999285	.9999847	.9999968	.9999993
.83	.9998311	.9999672	.9999937	.9999989	.9999998
.85	.9999203	.9999864	.9999977	.9999996	.9999999
.87	.9999663	.9999950	.9999993	.9999999	1.0000000
.89	.9999876	.9999985	.9999998	1.0000000	1.0000000
.91	.9999963	.9999996	1.0000000	1.0000000	1.0000000
.93	.9999992	.9999999	1.0000000	1.0000000	1.0000000
.95	.9999999	1.0000000	1.0000000	1.0000000	1.0000000
.97	1.0000000	1.0000000	1.0000000	1.0000000	1.0000000
.99	1.0000000	1.0000000	1.0000000	1.0000000	1.0000000

4) Les moments de λ_M sont

$$E_{\lambda_M}(t^q) = \frac{1}{M^q C_{M+q-1}^q}$$

et

$$\overline{\lambda_M} = \frac{1}{M^2}$$

2.2.3.

Théorème 6 La probabilité pour que k et exactement k coordonnées de y soient supérieures à t est égale à

$$C_M^k \sum_{d=k}^* (-1)^{d-k} C_{M-k}^{d-k} (1-st)^{M-1}, \quad t \in [0, 1/k].$$

En effet on a
$$\text{Prob}(\xi_{k,t}) = \frac{P(\xi_{k,t})}{P(S_M)}$$

$$= \sum_{d=k}^* (-1)^{d-k} C_d^k C_M^d (1-st)^{M-1},$$

$$= C_M^k \sum_{d=k}^* (-1)^{d-k} C_{M-k}^{d-k} (1-st)^{M-1},$$

chaque monôme $(1-st)$ n'apparaît que sur $[0, 1/d]$. ■

II.3 - Historique et commentaires.

i) La distribution de la plus grande part λ_1 d'un partage, avec distribution uniforme sur le simplexe, a été établie à plusieurs reprises (dont cette fois-ci). La littérature fait remonter la première découverte en 1887, par W.A. WHITWORTH dans son ouvrage "choice and chance" (réf. 83). Quant à la loi de la plus petite part λ_M , elle peut être plus ancienne encore (réf. 19,31,66).

ii) Il revient semble-t-il à R.A. FISHER de l'avoir établi ensuite de façon indépendante en 1929. (Réf. 24 tome 2, 84).

Le contexte est différent en apparence. Il s'agit du signal aléatoire

$$A(t) = \sum_{i=1}^M \chi_i \cos \omega_i t + \gamma_i \sin \omega_i t$$

où les ω_i sont donnés, χ_i et γ_i des variables aléatoires normales centrées réduites indépendantes ; FISHER étudie la plus grande des quantités

$$\frac{\chi_i^2 + \gamma_i^2}{\sum \chi_j^2 + \gamma_j^2} \quad . \text{ L'on verra ci-après (chap.IV.4) la rai-}$$

son pour laquelle on a bien un point du simplexe unité avec distribution uniforme. La procédure suivie par FISHER consiste à identifier les coefficients d'un polynôme de degré $M-1$. (Réf. 26).

iii) On peut remarquer par ailleurs combien la fonction de répartition de λ_1 présente une certaine ressemblance avec la densité de probabilité de la variable somme de M variables aléatoires indépendantes de distribution uniforme sur $[0, 1]$, qui a pour valeur:

$$f_M(x) = \frac{1}{(M-1)!} \sum_{k=0}^{[x]} (-1)^k C_M^k (x-k)^{M-1}$$

densité établie par LAPLACE en 1814 (Réf. 84 Chap.8), et que RENYI attribue à LOBATCHEWSKI pour les variables uniformes sur $[-1, 1]$, (Réf. 68 Chap.4).

D'après FELLER cette ressemblance résulte d'une "bonne chance innattendue", elle fait l'objet du théorème de recouvrement de L.W. STEVENS (1939) que l'on rappelle ci-dessous (Réf. 24 tome 2, 76):

Théorème 7 Si l'intervalle $[0, \beta]$ est découpé en M intervalles successifs à l'aide de $M-1$ réalisations indépendantes de la loi uniforme sur $[0, \beta]$, ordonnées par ordre croissant a posteriori, la probabilité $\varphi_M(\beta)$ pour que la longueur d'aucune de ces intervalles ne soit supérieure à t est égale à

$$\varphi_M(\beta) = \frac{t^{M-1}}{\beta^{M-1}} (M-1)! \int_0^{\beta/t} (\beta/t)$$

Pour β fixé, c'est aussi la probabilité pour que le plus grand des intervalles soit inférieur ou égal à t , soit avec les notations ci-contre :

$$\varphi_M(\beta) = L_M^\lambda(t/\beta),$$

et avec $\beta=1$, $L_M^\lambda(t) = t^{M-1} (M-1)! \int_0^1 (1/t)$.

FISHER lui-même fit le rapprochement entre son résultat et celui de STEVENS en 1940 (Réf. 28).

i°) On indique ci-dessous la raison qui fait qu'en découpant $[0, 1]$ par $M-1$ réalisations de la loi uniforme sur $[0, 1]$ ordonnées a posteriori on a un point du simplexe unité qui suit la distribution uniforme.

Proposition 5 Soient $0 \leq x_1 \leq \dots \leq x_k \leq \dots \leq x_{M-1} \leq 1$ les valeurs ordonnées de $M-1$ variables aléatoires indépendantes uniformes sur $[0, 1]$;

On pose :

$$y_1 = x_1$$

$$y_2 = x_2 - x_1$$

:

$$y_k = x_k - x_{k-1}$$

:

$$y_M = 1 - x_{M-1}$$

Le point de coordonnées y_i est un point du simplexe unité S_1^M qui suit la distribution définie par la mesure uniforme.

Démonstration : le point de coordonnées x_1, \dots, x_{n-1} est un point aléatoire de

l'ensemble $\{y \in \mathbb{R}^{M-1} ; y_1 \leq y_2 \leq \dots \leq y_{n-1}\}$ inclus dans l'hypercube $[0, 1]^{M-1}$ avec la densité constante $(n-1)!$.

Le point de coordonnées y_i est à l'évidence un point du simplexe unité S_1^M , et l'on vérifie que le jacobien

$$\frac{D(x_1, x_2, \dots, x_{n-1})}{D(y_1, y_2, \dots, y_{n-1})} \quad \text{est égal à 1.} \quad \blacksquare$$

σ) B. DE FINETTI étudie également cette partition de $[0, 1]$ en segments consécutifs, et se place "sul simplesso" (Réf. 25). Il envisage uniquement le sous-ensemble :

$$\tilde{S}_1^M = \left\{ y \in \mathbb{R}^M ; y_n \leq y_{n-1} \leq \dots \leq y_1 ; \sum_{i=1}^M y_i = 1 \right\},$$

qui est un polyèdre convexe avec les points sommets extrémaux

$$\left\{ \begin{array}{l} c_1 = e^1 \\ \vdots \\ c_k = \frac{1}{k} \sum_{i=1}^k e^i \\ \vdots \\ c_M = \frac{1}{M} \sum_{i=1}^M e^i \end{array} \right. ,$$

et le barycentre

$$c = \frac{1}{M} \sum_{i=1}^M c_i = \frac{1}{M} \sum_{b=1}^M \left(\frac{1}{b} + \dots + \frac{1}{M} \right) e^b.$$

Pour la distribution uniforme P sur le simplexe, les coordonnées de c sont les espérances mathématiques de chaque coordonnée λ_k , $k=1, \dots, M$

$$\bar{\lambda}_k = \frac{1}{M} \sum_{q=k}^M \frac{1}{q} . \quad \text{Pour la plus grande coordonnée } (\lambda=M),$$

cette relation est établie par FISHER avec des moyens analytiques (réf. 27)

v i) La distribution de λ_1 a été re-établie et utilisée par divers auteurs depuis, dans diverses circonstances, par exemple GARWOOD 1940, GREENWOOD 1946, MORAN 1947, IRWIN 1955, KENDALL-MORAN 1963 (réf. 31,35,40,45).

Dans de nombreux cas le "préalable" géométrique est considéré comme allant de soi ; il y a des démonstrations qui ne font pas intervenir les homothéties ; celle de FISHER à l'aide de polynômes, ou la première que donne GARWOOD (passage à la limite de rapports de coefficients du binôme).

v ii) Les probabilités de $\{ \lambda_k \geq t \}$ et de $F_{k,t}$ figurent toutes deux dans IRWIN (réf. 40) ; celle de $F_{k,t}$ est dans KENDALL et MORAN (réf. 45), celle de $\{ \lambda_k \geq t \}$ est dans FISHER, WILKS, RAO et SOBEL (réf. 27, 67, 84).

WILKS obtient ce résultat par l'intermédiaire de calculs analytiques et d'identification de moments, RAO et SOBEL par des intégrations récurrentes de la densité de DIRICHLET (chap. IV) dans le simplexe.

v iii) Ultérieurement des résultats plus complexes ont été établis.

La loi de la somme des q plus grandes coordonnées est établie par MAULDON en 1950 (réf. 54), celle de la somme d'un ensemble de q coordonnées ordonnées consécutives par BARTON et DAVID en 1955, (réf. 6), celle de la somme de coordonnées consécutives pondérées par leur rang avant classement et autres.. par BARTON et DAVID en 1956 (réf. 8). Le résultat de MAULDON est aussi établi par RAO et SOBEL (ref. 67).

i x) Proposition 6.

Pour tout entier $k = 1, \dots, M$ on a les identités suivantes :

$$\frac{1}{M} \sum_{p=k}^M \frac{1}{p} = C_{M-1}^{k-1} \sum_{q=k}^M (-1)^{q-k} \frac{C_{M-k}^{q-k}}{q^2} \quad ;$$

Pour $k=1$, elle s'écrit également

$$\frac{1}{M} \sum_{p=1}^M \frac{1}{p} = \frac{1}{M} \sum_{q=1}^M (-1)^{q-1} \frac{C_M^q}{q} .$$

En effet, il suffit de considérer les espérances mathématiques de $\lambda_k, k=1, \dots, M$ que l'on a calculées avec la densité de probabilité d'une part,

et avec les coordonnées du barycentre $c = \frac{1}{M} \sum_{j=1}^M c_j$ de l'autre.

L'égalité est triviale pour $k=M$.

x) Exemple .

Soit le partage en cinq parts du papyrus Rhind rappelé par G.Th. GUILBAUD (réf. 37), et le partage moyen sur S_4^5 pour la distribution uniforme :

$$\begin{array}{ll} \lambda_1 = 46/120 = 0,38333 & ; \quad \bar{\lambda}_1 = 0,45666 \\ \lambda_2 = 35/120 = 0,291667 & ; \quad \bar{\lambda}_2 = 0,25666 \\ \lambda_3 = 24/120 = 0,2 & ; \quad \bar{\lambda}_3 = 0,15666 \\ \lambda_4 = 13/120 = 0,10833 & ; \quad \bar{\lambda}_4 = 0,09 \\ \lambda_5 = 2/120 = 0,0166 & ; \quad \bar{\lambda}_5 = 0,04 \end{array}$$

Si l'on suppose que c'est un point aléatoire du simplexe S_4^5 qui suit la distribution uniforme on obtient par exemple

$$\text{Prob} (\lambda_1 \geq 46/120) = 0,6934 .$$

L'intervalle $[0,383 \quad 0,520]$ est centré sur la moyenne, sa probabilité est égale à 0,46 environ.

Si la plus grande part est le seul résumé que l'on conserve d'un partage, le partage du papyrus n'est pas un partage "rare" dans le cas de la distribution uniforme sur simplexe.

II.4. Lois limites quand M tend vers plus l'infini.

Comme l'a remarqué P. LEVY "La loi de λ_1 n'est donc pas simple" (réf.52), et son emploi peut devenir tout à fait malaisé lorsque le nombre M de participants est élevé. C'est la même situation que dans les problèmes d'occupation ou de rencontres, où précisément "le problème présente un intérêt particulier quand les effectifs sont grands" (réf. 24 tome 1 chap. IV). On s'efforce habituellement de lever cette difficulté pratique en étudiant le comportement asymptotique des lois lorsque $M \rightarrow +\infty$. On l'envisage ci-dessous sous deux points de vue.

4.1. Lois limites de λ_k .

Proposition 7 Avec la mesure uniforme sur le simplexe unité, la variable λ_k converge en probabilité vers zéro.

Démonstration. La densité de probabilité de λ_1 est égale à

$$\sum_{k=1}^* (-1)^{k+1} k (n-1) C_n^k (1-k t)^{M-2},$$

avec \sum^* sommation étendue à $1 \leq k \leq \inf(M, [1/\epsilon])$.

Pour tout ϵ positif, pour tout $M \geq M_\epsilon = [1/\epsilon]$, la densité de λ_1 est la somme de M_ϵ termes au plus, chacun d'eux étant majoré en valeur absolue par $k M C_n^k (1-k t)^{M-2} < \frac{M^{k+1} \beta^{n-2}}{(k-1)!}$ avec $\beta < 1$.

Ce majorant tend vers zéro quand $M \rightarrow \infty$, on a donc

$$\lim_{M \rightarrow \infty} \text{Prob} \{ |\lambda_1| \geq \epsilon \} = 0.$$

Comme $0 \leq \lambda_2 \leq \lambda_1$, $k=1, \dots, M$, il en est de même pour tout λ_k . ■

On a d'autres résultats à propos des espérances $\bar{\lambda}_k$ quand on fait tendre M vers $+\infty$, (ref. 25). On considère pour cela les valeurs de $\bar{\lambda}_k$ qui résultent des coordonnées du barycentre $c = \frac{1}{M} \sum c_j$.

On a: $\lim_{M \rightarrow \infty} \bar{\lambda}_1 = \frac{1}{M} (\gamma + \log M)$, avec $\gamma = 0,577\dots$ la constante d'EULER-MASCHERONI,

$$\lim_{M \rightarrow \infty} \bar{\lambda}_k = \frac{1}{M} \log \frac{M}{k-1} \quad k=2, \dots, M-1,$$

et la valeur exacte $\bar{\lambda}_M = 1/M^2$.

4.2. Loix limites de $\varphi_k(M), \lambda_k$.

L'analyse précédente est simple mais de peu d'intérêt pratique. On rencontre dans la littérature des résultats plus précis que nous rappelons ci-dessous.

i) Le cas le plus simple est celui de λ_M .

Théorème 8 Quand on fait tendre M vers $+\infty$, la variable aléatoire λ_M converge en loi vers la variable aléatoire qui suit la distribution exponentielle.

En effet on a $\text{Prob} (M^2 \lambda_M \leq M^2 u) = 1 - (1 - \pi u)^{M-1}$, $0 \leq u \leq 1/M$

$$\text{Prob} (M^2 \lambda_M \leq \sigma) = 1 - \left(1 - \frac{\sigma}{M}\right)^{M-1}, \quad 0 \leq \sigma \leq M$$

c'est une expression qui tend vers $1 - e^{-\sigma}$, $\forall \sigma$ quand $M \rightarrow +\infty$. ■

ii) L'observation des limites des espérances $\bar{\lambda}_k$ suggère que les différentes coordonnées ordonnées ont des comportements différents lorsque l'on fait tendre M vers $+\infty$; que les petites parts sont de l'ordre de $1/M^2$, que les parts intermédiaires sont de l'ordre de $1/M$, et que les plus grandes sont de l'ordre de $\log M / M$.

Le théorème suivant est dû à D.A. DARLING (ref. 19);

Théorème 9 Soient a et b deux nombres réels positifs qui vérifient $0 < a < b$.
Quand on fait tendre M vers $+\infty$, on a :

α) Cas des parts intermédiaires :

Le nombre de parts qui sont comprises entre a/M et b/M suit à la limite la distribution normale de moyenne $M(e^{-a} - e^{-b})$ et de variance $M(e^{-a} - e^{-b}) - M(ae^{-a} - be^{-b})^2$;

β) Cas des petites parts :

Le nombre de parts comprises entre a/M^2 et b/M^2 suit à la limite la distribution de POISSON de paramètre $b - a$;

γ) Cas des grandes parts :

Le nombre de parts comprises entre $(a + \log M)/M$ et $(b + \log M)/M$ suit à la limite la distribution de POISSON de paramètre $e^{-a} - e^{-b}$.

Ces résultats sont obtenus à partir de transformées de LAPLACE et de méthodes analytiques. ■

iii) P. LEVY a étudié de façon plus précise le cas de λ_1 (Réf. 52).

Théorème 10 Quand on fait tendre M vers $+\infty$, la variable $M \lambda_1 - \log M$ converge en loi vers la variable qui suit la distribution des extrêmes de FRECHET-FISHER-TIPPETT de type 1.

$$\text{C'est-à-dire } \lim_{M \rightarrow +\infty} \text{Prob} (M \lambda_1 - \log M \leq t) = e^{-e^{-t}},$$

$$\text{ou } \lim_{M \rightarrow +\infty} \text{Prob} \left(\lambda_1 \leq \frac{\log M / u}{M} \right) = e^{-u} \quad \blacksquare$$

Remarque - On sait que l'espérance mathématique de cette distribution est γ , on retrouve la valeur limite de $\bar{\lambda}_1$ de DE FINETTI.

- Ce résultat se trouve en substance chez COCHRAN (Réf. 15).

(c) D.E. BARTON et F.N. DAVID ont établi des résultats du même ordre (Réf. 7).

Théorème 11 Quand on fait tendre M vers $+\infty$, on a :

α) Cas des grandes parts, k constant,

$$\bullet \quad \bar{\lambda}_k \sim \log M/M, \quad \text{var}(\lambda_k) \sim \frac{1}{M^2} \left(\frac{1}{6} - \sum_{j=1}^{k-1} \frac{1}{j^2} \right)$$

•• On pose $X = M \lambda_k - \log M$,
 $\mathcal{L} e^{-X}$ suit à la limite la distribution de χ^2 avec $2k+4$ degrés de liberté ;

β) Cas des petites parts, $k = M+1 - R$ constant,

$$\bullet \quad \bar{\lambda}_k \sim R/M^2, \quad \text{var}(\lambda_k) \sim R/M^4 ;$$

•• On pose $X = M^2 \lambda_k$, $\mathcal{L} X$ suit à la limite la distribution de χ^2 avec $2R$ degrés de liberté ;

γ) Cas des parts intermédiaires $k = 1 + \theta M$, $0 < \theta < 1$,

$$\bullet \quad \bar{\lambda}_k \sim \frac{1}{M} \log \frac{1}{\theta}, \quad \text{var}(\lambda_k) \sim \frac{1}{M^3} \left[\frac{1}{\theta} - 1 - \left(\log \frac{1}{\theta} \right)^2 \right],$$

•• On pose $X = \sqrt{M} (M \lambda_k + \log \theta)$, X suit à la limite la distribution normale centrée de variance $\frac{1}{\theta} - 1 - \left(\log \frac{1}{\theta} \right)^2$. ■

4.3. Exemple .

Dans l'étude du découpage du territoire rural français par les grands axes routiers, il a été dénombré 1378 mailles routières connexes de superficie totale 480.000 km², la plus grande d'elles ayant une superficie de 4 380 km² environ (réf. 13,56).

On suppose que ce partage suit la distribution uniforme sur le simplexe S_1^n avec $M = 1378$ et $\lambda_1 = 0,0081$.

L'expression exacte de $\text{Prob} \{ \lambda_1 \leq 0,0081 \}$ nécessite le calcul et la somme de 124 termes ! Avec l'expression simplifiée de P. LEVY on a immédiatement

$$1378 \times 0,081 - \text{Log } 1378 = 3,9334$$

d'où

$$\text{Prob} \{ 1378 \lambda_1 - \text{Log } 1378 \leq 3,9334 \} \approx 0,98.$$

Dans le cadre de nos hypothèses (partage aléatoire qui suit la loi uniforme sur le simplexe), un tel partage présente une première part λ_1 qui a une valeur assez remarquable.

II.5. Mesures uniformes sur $\bar{E}_T, T > 1/M$.

La mesure uniforme sur \bar{E}_T est un autre exemple de mesure sur le simplexe unité avec laquelle on peut établir simplement les lois de λ_1 et λ_M .

Cette probabilité est définie par

$$\forall \omega \subset S_1^M \quad P_T(\omega) = \frac{P(\omega \cap \bar{E}_T)}{P(\bar{E}_T)},$$

elle est absolument continue par rapport à P , invariante par le groupe des permutations des vecteurs de base de \mathbb{R}^M .

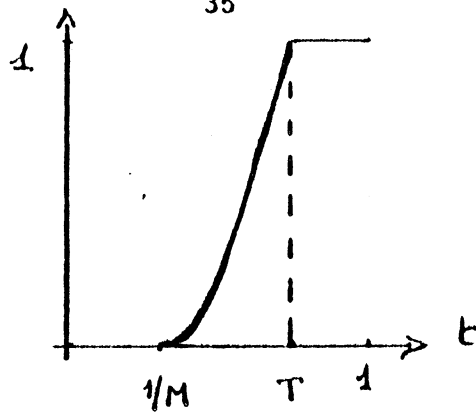
5.1. Loi de λ_1 .

Nous avons noté que pour tout $t, t' \in]\frac{1}{M}, 1]$ tels que $t \geq t'$,
 $\bar{E}_{t'} \subset \bar{E}_t$,

par conséquent :

Proposition 8 Pour tout T supérieur à $1/M$ et la mesure uniforme sur \bar{E}_T , la fonction de répartition de la variable λ_1 a la valeur suivante:

$$\begin{cases} \text{Prob} (\lambda_1 \leq t) = \frac{L_M^*(t)}{L_M^*(T)} & 0 \leq t \leq T \\ \text{Prob} (\lambda_1 \leq t) = 1 & T \leq t \leq 1 \end{cases}$$



La démonstration résulte de ce qui précède. ■

5.2. Loi de λ_M .

On doit considérer l'ensemble $\bar{G}_u \cap \bar{E}_T$, non vide pour $u \leq \frac{1}{M} \leq T$, réduit à ω pour $T = \frac{1}{M}$. On a

$$\begin{aligned} \mu(\bar{G}_u \cap \bar{E}_T) &= \mu(\bar{E}_T) - \mu(G_u \cap \bar{E}_T), \\ &= \mu(\bar{E}_T) - \mu(G_u) + \mu(G_u \cap E_T), \\ &= \mu(\bar{E}_T) + \sum_{k=0}^M (-1)^{k-1} \sum_{i_1, \dots, i_k} \mu(G_u \cap E_T^{i_1} \cap \dots \cap E_T^{i_k}). \end{aligned}$$

On a vu que $\mu(\bar{E}_T) = \sum_{p=0}^{[1/T]} (-1)^p C_M^p (1-pT)^{M-1} \mu(S_1^M)$,

D'après le théorème 4 on a également

$$\mu(G_u \cap E_T^{i_1} \cap \dots \cap E_T^{i_k}) = (1 - (M-k)u - kT)^{M-1} \mu(S_1^M),$$

par conséquent on a :

Proposition 9. La fonction de répartition de la variable λ_M a la valeur suivante:

$$\frac{P(\bar{G}_u \cap \bar{E}_T)}{P(\bar{E}_T)} = 1 + \frac{\sum_{k=0}^{[1/T-u]} (-1)^{k-1} C_M^k (1 - (M-k)u - kT)^{M-1}}{\sum_{p=0}^{[1/T]} (-1)^p C_M^p (1-pT)^{M-1}},$$

la sommation du numérateur étant étendue à tous les termes k tels que $1 - (M-k)u - kT \geq 0$, $k \leq M$.

La démonstration résulte de ce qui précède. ■

III - LOIS CONJOINTES DE COUPLES λ_1, λ_2 ET λ_k, λ_n AVEC LA MESURE UNIFORME

Les résultats géométriques et les formules combinatoires précédentes permettent d'établir la distribution conjointe de λ_k, λ_n , $k=1, \dots, n-1$, avec la mesure uniforme sur le simplexe unité.

La distribution conjointe du couple λ_1, λ_2 est établie avec un procédé plus proche des manipulations probabilistes avec des densités marginales et conditionnelles. Quand on se limite au sous-ensemble

$$\tilde{S}_1^M = \left\{ y \in S_1^M : y_1 \geq y_2 \geq \dots \geq y_n \right\},$$

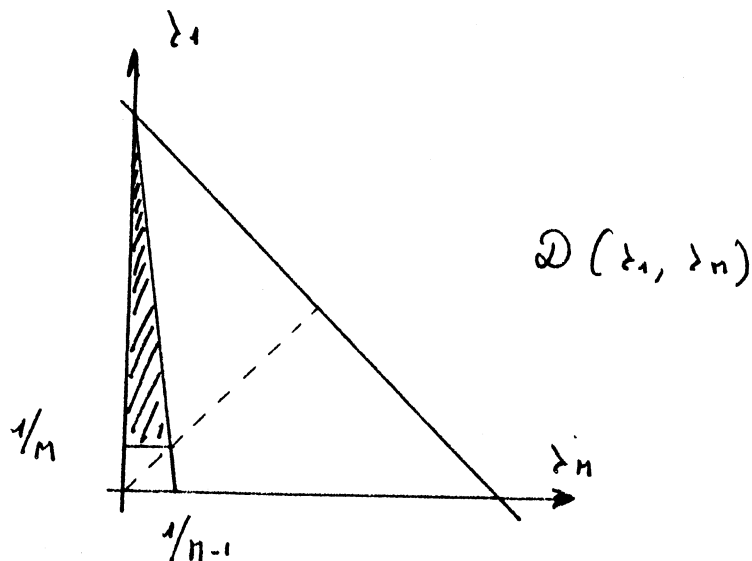
les densités marginales de chaque λ_i sont immédiates ; avec la mesure uniforme sur S_1^M on a rapidement la densité conjointe de λ_1, λ_2 (\tilde{S}_1^M est appelé ici le "simplexe ordonné").

La littérature contient également de nombreux résultats sur des distributions conjointes de plusieurs coordonnées ordonnées, et sur des distributions conjointes limites ($M \rightarrow \infty$).

III.1 - Distribution du couple λ_1, λ_n .

Dans le plan λ_1, λ_n le domaine $\mathcal{D}(\lambda_1, \lambda_n)$ parcouru par ce couple de coordonnées est défini par :

$$\begin{cases} 1/M \leq \lambda_1 \leq 1 \\ \lambda_1 + (n-1)\lambda_n \leq 1 \end{cases}$$



L'ensemble des points y tels que $\lambda_1 \leq t$ et $\lambda_n \geq u$ est l'ensemble $\bar{E}_t \cap G_u$ dont la probabilité est :

$$\sum_{k=0}^n (-1)^k \sum_{i_1 \dots i_k} \mu(G_u \cap E_t^{i_1} \cap \dots \cap E_t^{i_k}) / \mu(S_1^n),$$

avec la sommation étendue à tous les termes tels que $1 - (n-k)u - kt \geq 0$.

On a vu qu'avec la distribution uniforme sur le simplexe on a la probabilité

$$P(G_u \cap E_t^{i_1} \cap \dots \cap E_t^{i_k}) / P(S_1^n) = (1 - (n-k)u - kt)^{n-1}.$$

Proposition 10 On a la probabilité suivante:

$$P\{\lambda_1 \leq t, \lambda_n \geq u\} = \sum_{k=0}^{\lfloor \frac{1-nu}{t-u} \rfloor} (-1)^k C_n^k (1 - (n-k)u - kt)^{n-1}.$$

La démonstration est immédiate. ■

Remarque : Cette expression est donnée par DARLING (Réf 19) à l'issue d'un calcul analytique où interviennent des transformées de LAPLACE et des résidus. DARLING montre également que lorsque le nombre de parts M tend vers plus l'infini, la distribution du couple λ_1, λ_M converge en loi vers la distribution produit des deux distributions suivies par λ_1 et λ_M vues au chapitre précédent.

III.2. Distribution du couple λ_k, λ_{n-k} $k=2, \dots, n-1$.

On peut calculer comme en II.2 diverses probabilités à partir des probabilités élémentaires $\mu(G_u \cap E_t^{i_1} \cap \dots \cap E_t^{i_k}) / \mu(S_1^n)$. Pour tout entier $k=2, \dots, n-1$ et tout ensemble d'indices différents i_1, i_2, \dots, i_k on pose:

$$S_{k,t,u} = \sum_{i_1 \dots i_k} \mu(G_u \cap E_t^{i_1} \cap \dots \cap E_t^{i_k}),$$

égal à $C_n^k (1 - (n-k)u - kt)^{n-1} \cdot P(S_1^n)$ dans le cas de la dis-

tribution uniforme sur le simplexe (pour ensemble non vide, c'est-à-dire

$1 - (n-k)u - kt \geq 0$).

Proposition 11 Pour tout entier $k=1, \dots, M$, on a les probabilités suivantes:

$$\begin{aligned}
 \text{i) Prob} \{ y : \lambda_k \geq t \quad \lambda_M \leq u \} \\
 &= \sum_{\delta \geq k}^* (-1)^{\delta-k} C_{\delta-1}^{k-1} C_M^\delta (1 - (M-\delta)u - \delta t)^{M-1}, \\
 &= k C_M^k \sum_{\delta \geq k}^* (-1)^{\delta-k} C_{M-k}^{\delta-k} \frac{(1 - (M-\delta)u - \delta t)^{M-1}}{\delta}.
 \end{aligned}$$

$$\begin{aligned}
 \text{ii) Prob} \{ y : k \text{ et exactement } k \text{ coordonnées } \geq t \text{ et } \lambda_n \geq u \} \\
 &= \sum_{\delta \geq k}^* (-1)^{\delta-k} C_\delta^k C_M^\delta (1 - (M-\delta)u - \delta t)^{M-1}, \\
 &= C_M^k \sum_{\delta \geq k}^* (-1)^{\delta-k} C_{M-k}^{\delta-k} (1 - (M-\delta)u - \delta t)^{M-1};
 \end{aligned}$$

la sommation $\overset{*}{\sum}$ est étendue à tous les termes tels que $1 - (M-\delta)u - \delta t \geq 0$
c'est-à-dire

$$\delta \leq \inf \left(M, \left[\frac{1 - Mu}{t - u} \right] \right).$$

La démonstration est contenue dans les formules combinatoires vues au chapitre II. ■

III.3 - Le simplexe ordonné et ses densités.

3.1. Définitions. Soit \tilde{y} un point de \tilde{S}_1^M dont les coordonnées sont ordonnées, $\tilde{y}_i = y_{i^*}$, $i=1, \dots, M$. On appelle "simplexe ordonné" l'ensemble suivant :

$$\tilde{S}_1^M = \{ \tilde{y} \in S_1^M : \tilde{y}_1 \geq \tilde{y}_2 \geq \dots \geq \tilde{y}_M \} ,$$

c'est un sous-ensemble non vide, c'est celui auquel se limite DE FINETTI, que l'on trouve également chez WILKS ou FELLER. Pour $M = 3$ et 4 on peut en donner une représentation graphique (figures jointes extraites de Réf. 25). Il existe autant de simplexes ordonnés qu'il y a de numérotation des axes de R^M , à savoir $M!$, qui se déduisent les uns des autres par permutation des axes ; ils ont donc la même mesure de LEBESGUE $\Gamma(M)/M! = 1/M$.

Soit ct_n l'application multiforme qui à tout \tilde{y} de \tilde{S}_1^M , fait correspondre les y de S_1^M déduits par les $M!$ permutations de Σ_M . L'ensemble $ct_n(\tilde{y})$ possède $M!$ points distincts si et seulement si les \tilde{y}_i sont tous différents, dans le cas contraire on considère qu'il en est de même en comptant chaque point de $ct_n(\tilde{y})$ avec son ordre de multiplicité, à savoir $k_1! \times k_2! \times \dots \times k_n!$ si \tilde{y} possède n lots de coordonnées égales d'effectifs k_1, k_2, \dots, k_n avec $\sum k_i \leq M$.

Si g est une densité de probabilité sur S_1^M , on définit

$$\tilde{g}(\tilde{y}) = \sum_{M!} g(ct_n(y))$$

une fonction positive sur S_1^M somme des $M!$ valeurs prises par g aux points de $ct_n(\tilde{y})$.

On se limite aux densités g A.C.[†] sur S_1^M par rapport à la mesure P de LEBESGUE, pour lesquelles la frontière de \tilde{S}_1^M est un ensemble de mesure nulle. La fonction \tilde{g} est une densité de probabilité sur \tilde{S}_1^M , A.C par

† l'abréviation A.C. signifie absolument continue.

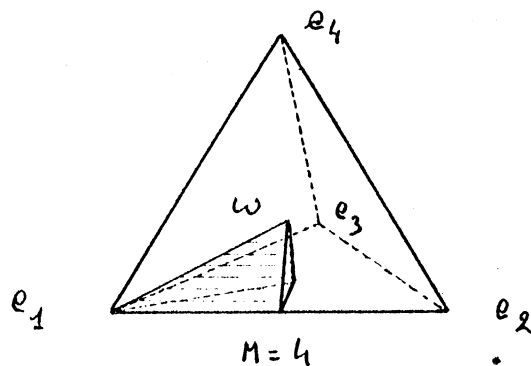
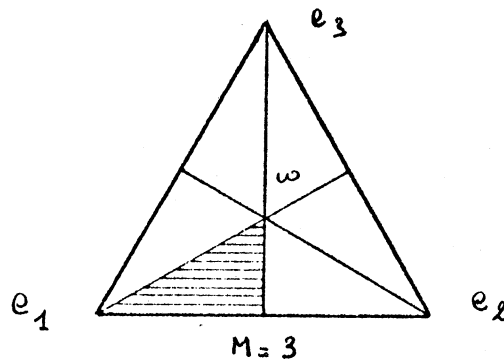
rapport à la restriction \tilde{P} de la mesure de LEBESGUE. Réciproquement à toute densité \tilde{g} A.C. sur \tilde{S}_1^M par rapport à \tilde{P} on fait correspondre la densité g qui pour tout y de $\mathcal{A}_M(\tilde{y})$ a la valeur:

$$g(y) = \frac{1}{M!} \tilde{g}(\tilde{y}),$$

c'est une densité unique et invariante par Σ_M sur S_1^M , à un ensemble de mesure P nulle près, inclus dans $\{y: y_{i_1} = y_{i_2} = \dots = y_{i_k}, k \geq 2\}$.

Remarques : 1) L'intérêt du simplexe ordonné réside dans le fait que les distributions marginales de \tilde{g} sont les distributions des coordonnées ordonnées λ_i (avec g et \tilde{g} A.C, et g invariante par Σ_M). C'est la remarque que DE FINETTI applique pour calculer l'espérance mathématique de chaque λ_i avec la densité constante sur le simplexe (Chap. II.3).

2) L'intérêt des densités g A.C et invariantes par Σ_M réside dans le fait que l'on peut appliquer la remarque ci-dessus, avec \tilde{g} .



Simplexes ordonnés \tilde{S}_1^M .

3.2. Densités marginales et conditionnelles.

Soit \tilde{g} une densité A.C. sur \tilde{S}_1^M , et \tilde{g}_1 la densité marginale en λ_1 ; on en déduit que $\tilde{g}(\lambda_1, \dots, \cdot) / \tilde{g}_1(\lambda_1)$ est la densité conditionnelle de $\tilde{g}_2, \dots, \tilde{g}_M$ par rapport à λ_1 .

3.2.1. Dans \mathbb{R}^{M-1} on effectue le changement de variables $\eta_p = \frac{\lambda_p}{1-\lambda_1}$, $p=2, \dots, M$ dont le jacobien est égal à $(1-\lambda_1)^{M-2}$; le point $\tilde{\eta}$ de coordonnées η_p , $p=2, \dots, M$ est un point de l'ensemble suivant

$$\tilde{S}_1^{M-1}(\lambda_1) = \left\{ \tilde{\eta} \in \tilde{S}_1^{M-1} ; 1_2 \leq \frac{\lambda_1}{1-\lambda_1} \right\},$$

inclus dans \tilde{S}_1^{M-1} , strictement inclus pour $\lambda_1 < 1/2$. La réunion des $(M-1)!$ sous-ensembles déduits de $\tilde{S}_1^{M-1}(\lambda_1)$ par les permutations du groupe Σ_{M-1} n'est autre que le sous-ensemble $\bar{E}_{\lambda_1/1-\lambda_1}$ de S_λ^{M-1} (chap. I), de point courant

x de coordonnées x_2, \dots, x_M .

Théorème 12 Sur $\tilde{S}_1^{M-1}(\lambda_1)$ la densité de probabilité conditionnelle de la variable $\tilde{\eta}$ par rapport à λ_1 est donnée par l'expression :

$$\tilde{f}(\tilde{\eta}) = \frac{M! (1-\lambda_1)^{M-2}}{\tilde{g}_1(\lambda_1)} \cdot g(\lambda_1, (1-\lambda_1)\eta_2, \dots, (1-\lambda_1)\eta_M).$$

La démonstration est immédiate. ■

On en déduit que la densité de probabilité du point aléatoire x a la valeur suivante :

$$f(x) = \frac{M (1-\lambda_1)^{M-2}}{\tilde{g}_1(\lambda_1)} \cdot g(\lambda_1, (1-\lambda_1)x_2, \dots, (1-\lambda_1)x_M),$$

elle est définie sur $\bar{E}_{\lambda_1/1-\lambda_1}$ et invariante par Σ_{M-1} .

On en conclut que l'on peut connaître la densité conditionnelle de λ_2 par rapport à λ_1 si l'on connaît la densité de $\sup(x_i)$ pour la densité f dans le simplexe unité de \mathbb{R}^{M-1} .

III.4 - Distribution du couple λ_1, λ_2 .

4.1. Domaine décrit par λ_1 et λ_2 .

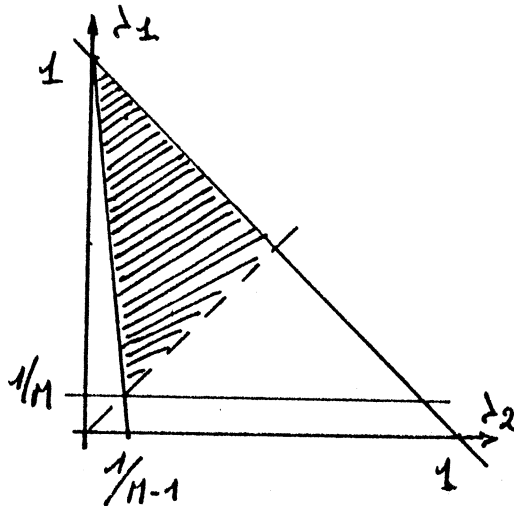
Les deux premiers coefficients λ_1 et λ_2 vérifient les conditions nécessaires

$$\begin{cases} \lambda_1 + \lambda_2 \leq 1 \\ \lambda_2 \leq \lambda_1 \\ \lambda_1 + (\eta-1)\lambda_2 \geq 1 \end{cases}$$

qui délimitent le domaine $\mathcal{D}(\lambda_1, \lambda_2)$ dans le plan λ_1, λ_2 .

Pour tout $\lambda \in \left[\frac{1}{\eta}, 1 \right]$ on a également

$$\frac{1-\lambda}{\eta-1} \leq \lambda_2 \leq \inf(\lambda, 1-\lambda).$$



$\mathcal{D}(\lambda_1, \lambda_2)$

On applique ce qui précède au cas de la mesure uniforme sur S_1^η , on a vu que les valeurs des densités sont les suivantes (chap. II):

$$g(y) = \Gamma(M) \text{ et } L_M^{\lambda'}(t) = (M-1) \sum_{k=1}^{[1/t]} (-1)^{k-1} C_n^k (1-t)^{M-2-k}$$

pour tout $\tilde{\eta}$, la densité \tilde{f} a donc la valeur $\tilde{f}(\tilde{\eta}) = \frac{M! (1-\lambda)^{M-2} \Gamma(M)}{L_M^{\lambda'}(\lambda)}$ et est constante.

4.2. Densité de $x \in S_1^{n-1}$.

Proposition 12 Pour tout λ_1 supérieur ou égal à $1/2$ la densité conditionnelle de λ_2 par rapport à λ_1 a pour valeur

$$L_{n-1}^{\lambda'}\left(\frac{\lambda_2}{1-\lambda_1}\right) \times \frac{1}{1-\lambda_1};$$

pour tout λ_1 appartenant à l'intervalle $]1/n, 1/2]$ la densité conditionnelle de λ_2 par rapport à λ_1 a pour valeur

$$L_{n-1}^{\lambda'}\left(\frac{\lambda_2}{1-\lambda_1}\right) \times \frac{1}{L_{n-1}^{\lambda'}(\lambda_1/(1-\lambda_1))} \times \frac{1}{1-\lambda_1}.$$

Démonstration: i) Pour $\lambda_1 \geq 1/2$ la densité \tilde{f} est une constante égale à $M! \Gamma(M)/M(n-1)$ sur le simplexe ordonné, ce qui correspond à la densité f constante sur S_1^{n-1} égale à $\Gamma(n-1)$. C'est là le cas de la distribution uniforme sur simplexe unité que l'on sait traiter, (chap. II), on en déduit le premier résultat.

ii) Pour $\frac{1}{n} < \lambda_1 < \frac{1}{2}$ la densité \tilde{f} est une constante égale à $\frac{M! \Gamma(M)(1-\lambda_1)^{M-2}}{L_M^{\lambda'}(\lambda_1)}$

sur le simplexe ordonné, ce qui correspond à la densité f constante sur $\bar{E}(\lambda_1/(1-\lambda_1))$ égale à $M! \Gamma(M) (1-\lambda_1)^{M-2} / L_M^{\lambda'}(\lambda_1)$.

On reconnaît un autre cas que l'on sait traiter, (chap. IV), on en déduit le deuxième résultat. ■

4.3. Identité vérifiée par $L_M^{\lambda'}(t)$.

Proposition 13 Pour tout entier M supérieur ou égal à 3, et pour tout $t \in [1/M, 1]$ on a l'identité suivante:

$$M(M-1)(1-t)^{M-2} L_{n-1}^{\lambda'}\left(\frac{t}{1-t}\right) = L_n^{\lambda'}(t).$$

Démonstration: Pour tout $t \in \left[\frac{1}{M}, \frac{1}{2} \right]$ on a :

$$\mu\left(\bar{E} \frac{t}{1-t}\right) = \frac{L^{\lambda'}(t)}{M \Gamma(M) (1-t)^{M-2}}$$

mais on a aussi

$$\mu\left(\bar{E} \frac{t}{1-t}\right) = \frac{L^{\lambda}_{n-1}\left(\frac{t}{1-t}\right)}{\mu(S_1^n)} = \frac{L^{\lambda}_{n-1}\left(\frac{t}{1-t}\right)}{\Gamma(n-1)} ;$$

pour tout $t \geq \frac{1}{2}$ le terme $L^{\lambda}_{n-1}\left(\frac{t}{1-t}\right)$ se réduit à 1, la relation ci-dessus est encore vérifiée. ■

Remarque, cette relation peut être directement vérifiée à partir de l'expression analytique de $L^{\lambda}_M(t)$.

4.4. Densité sur $\mathcal{D}(\lambda_1, \lambda_2)$.

Proposition 14 La densité ρ_M^{λ} de la distribution conjointe de λ_1 et λ_2 sur $\mathcal{D}(\lambda_1, \lambda_2)$ a pour valeur :

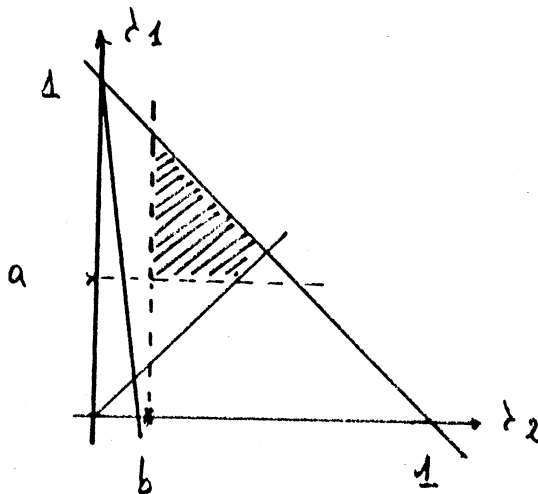
$$\rho_M^{\lambda}(\lambda_1, \lambda_2) = M(M-1)(1-\lambda_1)^{M-2} \cdot L^{\lambda'}\left(\frac{\lambda_2}{1-\lambda_1}\right) \cdot \frac{1}{1-\lambda_1}$$

en tout point de $\mathcal{D}(\lambda_1, \lambda_2)$.

La démonstration résulte des propositions qui précèdent. ■

4.5. Fonction de répartition en λ_1 et λ_2 .

4.5.1. Soit $\mathcal{D}_{a,b}$ le sous-ensemble $\left\{ \lambda_1, \lambda_2 \in \mathcal{D}(\lambda_1, \lambda_2), a, b \in \mathcal{D}(\lambda_1, \lambda_2), \lambda_1 \geq a, \lambda_2 \geq b \right\}$.
L'expression de ρ_M^{λ} est assez commode pour calculer les probabilités qui correspondent à $\mathcal{D}_{a,b}$.



$\mathcal{D}_{a,b}$

Proposition 15 Soient les nombres entiers suivants $A_1 = \lfloor 1-a/b \rfloor$, $A_2 = \lfloor 1-a/a \rfloor$.
La probabilité \mathbb{I}_{ab} qui correspond au sous-ensemble \mathcal{D}_{ab} est égale à

i) $a \geq 1/2$

$$\mathbb{I}_{ab} = M \sum_{k=1}^{A_1} (-1)^{k-1} C_{n-1}^k (1-a-kb)^{M-1},$$

ii) $0 < a < 1/2$

$$\begin{aligned} \mathbb{I}_{ab} &= M \sum_{k=1}^{A_1} (-1)^{k-1} C_{n-1}^k (1-a-kb)^{M-1}, \\ &+ M \sum_{q=1}^{A_2} (-1)^q \frac{C_{n-1}^q}{q+1} (1-(q+1)a)^{M-1}. \end{aligned}$$

Démonstration: Soit

$$\begin{aligned} \mathbb{I}_{ab} &= \iint_{\mathcal{D}_{ab}} L_n^\lambda(\lambda_1, \lambda_2) d\lambda_1 d\lambda_2 = \int_a^{1-b} M(n-1) (1-\lambda_1)^{M-2} d\lambda_1 \int_b^{\mathcal{F}(\lambda_1)} L_{n-1}^{\lambda_2} \left(\frac{\lambda_2}{1-\lambda_1} \right) \frac{d\lambda_2}{1-\lambda_1}, \\ &= M(n-1) \int_a^{1-b} (1-\lambda_1)^{M-2} \left[L_{n-1}^{\lambda_2} \left(\frac{\mathcal{F}(\lambda_1)}{1-\lambda_1} \right) - L_{n-1}^{\lambda_2} \left(\frac{b}{1-\lambda_1} \right) \right] d\lambda_1 \end{aligned}$$

avec $\mathcal{F}(\lambda_1) = 1-\lambda_1$ si $\lambda_1 \geq 1/2$, $\mathcal{F}(\lambda_1) = \lambda_1$ si $\lambda_1 < 1/2$.

i) Soit $a \geq 1/2$

$$\mathbb{I}_{ab} = M(n-1) \int_a^{1-b} (1-\lambda_1)^{M-2} \left[1 - L_{n-1}^{\lambda_2} \left(\frac{b}{1-\lambda_1} \right) \right] d\lambda_1,$$

On effectue le changement de variables $u = \frac{b}{1-\lambda_1}$, $du = \frac{b d\lambda_1}{(1-\lambda_1)^2}$ d'où

$$\mathbb{I}_{ab} = M(n-1) \int_{\frac{b}{1-a}}^1 b^{\frac{n-1}{M}} \frac{1}{u^M} \left[1 - L_{n-1}^{\lambda_2}(u) \right] du.$$

On pose $A = \left[\frac{1-a}{b} \right] \geq 1$, on sait que dans l'expression de L_{n-1}^{λ} le monôme $(1-ku)^{n-2}$, $0 \leq k \leq n-1$ n'intervient que sur l'intervalle $0 \leq u \leq 1/k$ d'où

$$I_{ab} = M(n-1) b^{n-1} \sum_{k=1}^A \int_{\frac{b}{1-a}}^{1/k} (-1)^{k-1} C_{n-1}^k \frac{(1-ku)^{n-2}}{u^n} du.$$

On effectue le nouveau changement de variable $t = b/u$, $dt = -\frac{b du}{u^2}$ d'où :

$$I_{ab} = n(n-1) \sum_{k=1}^A (-1)^{k-1} C_{n-1}^k \int_{kb}^{1-a} (t-kb)^{n-2} dt$$

$$I_{ab} = M \sum_{k=1}^A (-1)^{k-1} C_{n-1}^k (1-a-kb)^{n-1}.$$

ii) Soit $a < 1/2$, on a

$$I_{ab} = \int_a^{1/2} n(n-1) (1-\lambda_1)^{n-2} \left[L_{n-1}^{\lambda_1} \left(\frac{\lambda_1}{1-\lambda_1} \right) - L_{n-1}^{\lambda_1} \left(\frac{b}{1-\lambda_1} \right) \right] d\lambda_1 \\ + \int_{1/2}^{1-b} n(n-1) (1-\lambda_1)^{n-2} \left[1 - L_{n-1}^{\lambda_1} \left(\frac{b}{1-\lambda_1} \right) \right] d\lambda_1,$$

$$I_{ab} = n(n-1) \int_a^{1-b} (1-\lambda_1)^{n-2} \left[1 - L_{n-1}^{\lambda_1} \left(\frac{b}{1-\lambda_1} \right) \right] d\lambda_1 \\ + n(n-1) \int_a^{1/2} (1-\lambda_1)^{n-2} \left[L_{n-1}^{\lambda_1} \left(\frac{\lambda_1}{1-\lambda_1} \right) - 1 \right] d\lambda_1.$$

Le premier terme est l'analogue du terme calculé en i), pour calculer le second on effectue le changement de variable suivant, $u = \lambda_1 / (1-\lambda_1)$, ce terme devient :

$$M(n-1) \int_{\frac{a}{1-a}}^1 \frac{1}{(1+u)^M} \left[L_{n-1}^{\lambda_1}(u) - 1 \right] du.$$

On pose $A_2 = \left[\frac{1-a}{a} \right]$, $A \geq A_2 \geq 1$.

On applique la même règle d'apparition du monôme $(1-ku)^{n-2}$ d'où l'expression suivante du second terme :

$$M(n-1) \sum_{k=1}^{A_2} \int_{\frac{a}{1-a}}^{1/k} (-1)^k C_{n-1}^k (1-ku)^{n-2} \frac{du}{(1+u)^M},$$

à laquelle on applique le changement de variable suivant $t = 1/(1+u)$

d'où

$$n(n-1) \sum_{k=1}^{A_2} \int_{\frac{k}{k+1}}^{1-\alpha} C_{n-1}^k [t(1+k) - k]^{n-2} dt$$

$$= M \sum_{k=1}^{A_2} (-1)^k \frac{C_{n-1}^k}{k+1} (1 - (k+1)\alpha)^{n-1},$$

soit pour tout $\alpha < 1/2$

$$I_{a,b} = M \sum_{k=1}^A (-1)^{k-1} C_{n-1}^k (1 - \alpha - kb)^{n-1}$$

$$+ M \sum_{q=1}^{A_2} (-1)^q \frac{C_{n-1}^q}{q+1} (1 - (q+1)\alpha)^{n-1}.$$

4.6. Compléments .

La littérature contient divers autres résultats relatifs à des lois conjointes relatives à plusieurs coordonnées ordonnées λ_i .

BARTON et DAVID donnent la loi conjointe de deux coordonnées quelconques λ_i ; λ_j ou davantage ; les moments et les cumulants pour deux coordonnées ou davantage ; les corrélations et les distributions du rapport ou de la différence de deux coordonnées (Réf. 7). Ils donnent également la loi conjointe de la somme de deux sous-ensembles de coordonnées consécutives

$$S_{ij} = \sum_i^j \lambda_n, \quad S_{ke} = \sum_k^e \lambda_n, \quad i < j < k < e, \quad (\text{Ref. 6}).$$

RAO et SOBEL donnent la densité de la loi conjointe des k premières ou des k dernières coordonnées ordonnées (Réf. 67).

Ces expressions sont assez complexes, les formes limites quand $n \rightarrow +\infty$ sont plus simples. On rappelle que lorsque $n \rightarrow +\infty$ la loi conjointe de λ_1 et λ_n tend vers le produit des deux lois marginales limites en λ_1 et λ_n (DARLING, Réf. 19) ; BARTON et DAVID établissent également la loi conjointe limite de S_{ij} et S_{ke} lorsque $n \rightarrow +\infty$, et que $\frac{i}{n}, \frac{j}{n}, \frac{k}{n}, \frac{e}{n}$ tendent vers des limites d vérifiant $0 \leq d_i < d_j < d_k < d_e$. La loi limite est une loi normale sur \mathbb{R}^2 (Ref. 6).

4.7. Exemple. On reprend le partage du papyrus Rhind.

$$i) \quad M = 5, \quad \lambda_1 = \frac{46}{120}, \quad \lambda_5 = \frac{2}{120},$$

$$1 - M \lambda_5 = \frac{11}{12}, \quad \lambda_1 - \lambda_5 = \frac{11}{30}, \quad \left[\frac{1 - M \lambda_5}{\lambda_1 - \lambda_5} \right] = 2,$$

d'où

$$\text{Prob} \left(\lambda_1 \leq \frac{46}{120}, \lambda_5 \geq \frac{2}{120} \right) = \left(\frac{11}{12} \right)^4 - 5 \left(\frac{11}{30} \right)^4 + 10 \left(\frac{11}{60} \right)^4 = 0,2598.$$

L'expression asymptotique produit des lois de λ_1 et de λ_5 donne 0,2165.

On a encore, par exemple

$$\text{Prob} \left(\lambda_1 \geq \frac{46}{120}, \lambda_5 \geq \frac{2}{120} \right) = 0,4463.$$

$$ii) \text{ On a également } A = \left[\frac{120-46}{35} \right] = 2, \quad A_2 = \left[\frac{120-46}{46} \right] = 1$$

d'où

$$I_{\frac{46}{120}, \frac{2}{120}} = 0,1934.$$

IV - AUTRES MESURES SUR LE SIMPLEXE UNITE .

IV.1 - Généralités .

L'on considère à présent d'autres mesures μ sur le simplexe unité que la mesure uniforme \mathbb{P} précédente. Dans la deuxième partie (Chap. V) il s'agira de certaines distributions discontinues réparties sur des ensembles finis de points ; ci-dessous l'on envisage des mesures absolument continues par rapport à la mesure de LEBESGUE. Pour obtenir de telles mesures on procède de trois façons :

- a) on construit des densités de points aléatoires qui appartiennent nécessairement au simplexe ;
- b) on envisage des changements de variables sur simplexe unité ;
- c) on rappelle d'autres densités de la littérature.

Au cours de cet examen on rencontre les densités de probabilité dites de DIRICHLET. Ces densités ont des propriétés analytiques intéressantes, elles conduisent à des "structures statistiques" sur simplexe unité qui sont reprises dans la troisième partie.

Deux points généraux doivent être mentionnés :

i) Dans l'ensemble des mesures de probabilité sur simplexe unité on doit noter la place particulière des mesures qui sont invariantes par le groupe \sum_M des permutations des indices de 1 à M ; elles s'introduisent naturellement par les simplifications qu'elles entraînent (III.3).

ii) La question des parties ordonnées et extrémales est assez facilement résolue avec la mesure uniforme grâce aux homothéties vues en I. Cette simplification ne se reproduit évidemment plus avec les autres mesures μ ; aussi est-il nécessaire d'envisager d'autres techniques d'intégration. Lorsque la mesure μ et le domaine $\mathcal{D} \subset S_1^M$ sont connus nous supposerons dans la suite que $\mu(\mathcal{D})$ peut être calculé au besoin par des méthodes d'intégration numérique approchées.

La question des parts extrémales λ_1 et λ_M est plus spécialement évoquée en IV.5, IV.6, et IV.11 ; en V.5 on retrouve ce que l'on désigne par le "critère de COCHRAN" dans la littérature, dont on rappelle la distribution asymptotique quand M tend vers plus l'infini, (ref. 15).

IV.2 - Calculs avec les $\mu(E_t^{i_1} \cap \dots \cap E_t^{i_k})$.

2.1. On a vu que les probabilités liées aux parts ordonnées d'un partage s'expriment à l'aide des

$$S_{kt} = \sum_{i_1, \dots, i_k} \mu(E_t^{i_1} \cap \dots \cap E_t^{i_k}),$$

où la sommation est étendue à l'ensemble des lots de k indices différents pris dans $\{1, 2, \dots, M\}$, $k \leq M$.

On a notamment (rappel)

$$\mu(E_t) = \sum_{k=1}^{[1/t]} (-1)^{k-1} S_{kt},$$

$$\mu(y: \lambda_q \geq t) = \sum_{k=q}^* (-1)^{k-q} C_{k-1}^{q-1} S_{kt},$$

$$\begin{aligned} \mu(y: q \text{ et exactement } q \text{ parts } \geq t) \\ = \sum_{k=q}^* (-1)^{k-q} C_k^q S_{kt}, \end{aligned}$$

avec \sum^* la sommation telle que $\sup k = \inf(M, [1/t])$.

Dans le cas où la mesure μ est invariante par \sum_M on a le résultat supplémentaire

$$S_{kt} = C_M^k \mu(E_t^1 \cap \dots \cap E_t^k).$$

2.2. Nous supposons que μ est une probabilité absolument continue (A.C.) par rapport à la mesure uniforme \mathbb{P} sur le simplexe unité ; on note

$h(y_1, \dots, y_M)$ la valeur de la densité de μ au point y , $y_i \geq 0$, $i = 1, \dots, M$, $\sum_{i=1}^M y_i = 1$. Dans la suite il sera commode d'exprimer cette densité h en fonction de $M-1$ variables indépendantes y_j ; on note

$\bar{h}^q(y_1, \dots, y_{q-1}, y_{q+1}, \dots, y_M)$ la valeur de la densité de μ exprimée en fonction des variables $y_1, \dots, y_{q-1}, y_{q+1}, \dots, y_M$.

Remarque : Pour tout entier $k=1, \dots, M-1$ la probabilité $\mu(E_t^{i_1} \cap \dots \cap E_t^{i_k})$ peut s'exprimer à l'aide des densités marginales de \bar{h}^q , l'indice q étant différent des indices i_1, \dots, i_k :

Soient i_1, \dots, i_k des indices différents et soit q un autre indice ; pour simplifier et sans restreindre la généralité on suppose que q est l'indice M , on introduit donc la densité \bar{h}^M . Pour évaluer

$\mu(E_t^{i_1} \cap \dots \cap E_t^{i_k})$ on factorise \bar{h}^M en un produit d'une densité marginale et d'une densité conditionnelle :

- Soit par exemple, avec des notations évidentes

$$\bar{h}^M(y_1, \dots, y_{M-1}) = \bar{h}_m^M(y_1) \times \bar{h}_c^M(y_2, \dots, y_{M-1} / y_1),$$

d'où

$$\begin{aligned} \mu(E_t^1) &= \int_{E_t^1} \bar{h}^M dy_1 \dots dy_{M-1} \\ &= \int_t^1 \bar{h}_m^M(y_1) dy_1 \int_{E_t^1/y_1} \bar{h}_c^M(y_2, \dots, y_{M-1}/y_1) dy_2 \dots dy_{M-1} \\ &= \int_t^1 \bar{h}_m^M(y_1) dy_1 ; \end{aligned}$$

- et de manière générale, si on pose

$$\bar{h}^M(y_1, \dots, y_{n-1}) = \bar{h}_m^M(y_1, \dots, y_k) \times \bar{h}_c^M(y_{k+1}, \dots, y_{n-1} / y_1, \dots, y_k)$$

on trouve encore :

$$\mu(E_t^1 \cap \dots \cap E_t^k) = \int_{\substack{y_1 \geq t \\ \vdots \\ y_k \geq t}} \bar{h}_m^M(y_1, \dots, y_k) dy_1 \dots dy_k$$

pour tout $k = 1, \dots, M-1$.

La probabilité de $E_t^1 \cap \dots \cap E_t^k$ s'exprime donc simplement à l'aide des densités marginales de \bar{h}^M . Lorsque μ et h sont invariantes par le groupe des permutations Σ_M , les densités marginales de tout \bar{h}^M sont nécessairement identiques entre elles quand elles ont le même nombre de variables. Dans ce dernier cas pour tout $k = 1, \dots, M-1$ on peut donc écrire :

$$S_{kt} = C_M^k \int_{\substack{y_1 \geq t \\ \vdots \\ y_k \geq t}} \bar{h}_m^M(y_1, \dots, y_k) dy_1 \dots dy_k.$$

A LOIS FERMEES POUR LA CONVOLUTION SUR \mathbb{R}_+ ET MESURES SUR LE SIMPLEXE.

IV.3 - La Normalisation par la somme.

En pratique une méthode simple pour obtenir un point de S_n^M consiste à diviser M scalaires réels positifs (ou nuls, non tous nuls) par leur somme. Nous appliquons cette méthode ci-dessous.

3.1. Soit H une densité de probabilité sur \mathbb{R}_+^{*M} , par rapport à la mesure de LEBESGUE sur \mathbb{R}_+^{*M} . On considère le changement de variables suivant, qui traduit la méthode énoncée :

$$\begin{cases} \Delta = \sum_1^M x_i & , \Delta \neq 0 \\ y_j = x_j / \Delta & j=1, \dots, M \end{cases}$$

Δ est un scalaire réel positif et y est un point de S_1^M , ($y_i \geq 0$, $\sum_{i=1}^M y_i = 1$).

Propriété 16 La loi de probabilité sur le simplexe, image de la loi sur \mathbb{R}_+^{*M}

de densité H est définie par la densité h définie sur S_1^M donnée

par la relation $h(y) = \int_0^\infty H(\Delta y) \Delta^{M-1} d\Delta$.

En effet dans ce changement de jacobien $\frac{D(x_1, \dots, x_M)}{D(y_1, \dots, y_{M-1}, \Delta)} = \Delta^{M-1}$

le point y ($y_i \geq 0$, $\sum y_i = 1$) appartient au simplexe unité de \mathbb{R}^M , la variable Δ varie sur $]0, +\infty[$.

On a $\int_{\mathbb{R}_+^{*M}} H(x) dx_1 \dots dx_M = \int_{S_1^M} \int_0^\infty H(\Delta y) \Delta^{M-1} d\Delta dy_1 \dots dy_{M-1}$

il en résulte que $h(y) = \int_0^\infty H(\Delta y) \Delta^{M-1} d\Delta$ ■

Remarque : La densité h est naturellement invariante par Σ_M quand H est invariante par Σ_M .

Cas particulier : Soit x une v.a. positive ou nulle de densité f définie sur \mathbb{R}_+^* , et un échantillon de M tirages indépendants x_1, \dots, x_M de cette loi. Ceci donne la densité suivante dans \mathbb{R}_+^{*M} :

$$H(x) = f(x_1) \dots f(x_M) \quad \text{qui est invariante par } \Sigma_M,$$

on en déduit :

$$h(y_1, \dots, y_n) = \bar{h}^M(y_1, \dots, y_{n-1}) = \int_0^\infty f(sy_1) \dots f(sy_n) s^{M-1} ds$$

avec $\sum y_i = 1$.

Remarques: a) Cette situation pour y_1, \dots, y_{M-1} est désignée par MOSIMANN comme "partiellement indépendante excepté pour la contrainte" (ref. 57).

b) Le procédé que l'on appelle ici la "normalisation par la somme" est en fait la construction de la distribution marginale en y de la distribution sur \mathbb{R}_+^{*M} de densité H , quand on considère que cette distribution est définie sur le produit cartésien

$$\mathbb{R}_+^* \times S_1^M, \quad s \in \mathbb{R}_+^*, \quad y \in S_1^M.$$

On peut obtenir une distribution sur le simplexe S_Δ^M en considérant la distribution suivie par le vecteur de coordonnées $sy_i, i=1, \dots, M$ conditionné par rapport à Δ . Cette distribution sera étudiée sur un cas particulier.

3.2. Pour tout n entier supérieur ou égal à 1, soit la densité f^{*n} de la n -ième itérée de f par convolution :

$$f^{*1}(x) = f(x)$$

$$f^{*n}(x) = \int_0^x f(t) f^{*(n-1)}(x-t) dt \quad n \geq 2.$$

Théorème 13 Dans le cas où H est une densité sur \mathbb{R}_+^{*M} produit de M densités f sur \mathbb{R}_+^* , les densités marginales de toute densité \bar{h}^q peuvent s'exprimer à l'aide des densités f^{*2} .

Démonstration - Sans restreindre la généralité nous prendrons l'indice q égal à l'indice M , et les indices i_1, \dots, i_k égaux aux indices $1, \dots, k$. Cette situation à laquelle on peut toujours se ramener à l'aide d'une permutation sur les vecteurs de base de \mathbb{R}^M à l'avantage de simplifier les notations puisque l'on se limite aux densités marginales de \bar{h}^M qui pour tout $k=1, \dots, M-1$ ont pour valeurs respectives $\bar{h}_m^M(y_1, \dots, y_k)$. La démonstration est faite par récurrence sur l'entier k .

i) Soient $\gamma^{(2)}$ le vecteur de coordonnées y_2, \dots, y_{n-1} , $\bar{h}_m^{M(2)}$ la densité marginale correspondante et $y_m^{(2)} = 1 - \sum_{j=2}^{n-1} y_j$. On a par définition:

$$\begin{aligned} \bar{h}_m^{M(2)}(\gamma^{(2)}) &= \int_0^{y_m^{(2)}} \bar{h}^M(y_1, \dots, y_{n-1}) dy_1 \\ &= \int_0^\infty f(s y_2) \dots f(s y_{n-1}) s^{M-2} ds \int_0^{s y_m^{(2)}} f(s y_1) f(s y_m^{(2)} - s y_1) s dy_1 \end{aligned}$$

d'où

$$\bar{h}_m^{M(2)}(\gamma^{(2)}) = \int_0^\infty f(s y_2) \dots f(s y_{n-1}) f^{*2}(s y_m^{(2)}) s^{M-2} ds.$$

On obtient également la densité $\bar{h}_m^{M(\Sigma)}$ à l'aide d'une normalisation par la somme de $M-2$ variables qui suivent la loi de densité f et une variable qui suit la loi de f^{*2} .

ii) Soient $Y^{(P)}$ le vecteur y_{p+1}, \dots, y_{n-1} , $\bar{h}_m^{M(P)}$ la densité marginale correspondante et $y_M^{(P)} = 1 - \sum_{j=p}^{n-1} y_j$. Pour la récurrence on suppose que l'on a

$$\bar{h}_m^{M(P)}(Y^{(P)}) = \int_0^\infty f(dy_p) \dots f(dy_{n-1}) f^{*P}(s y_n^{(P)}) s^{M-P} ds.$$

Comme la densité marginale en y_{p+1}, \dots, y_{n-1} est donnée par

$$\int_0^{y_n^{(P+1)}} \bar{h}_m^{M(P)}(Y^{(P)}) dy_p = \int_0^\infty f(dy_{p+1}) \dots f(dy_{n-1}) s^{n-p-1} ds \times \left[\int_0^{y_n^{(P+1)}} f(dy_p) f^{*P}(s y_n^{(P+1)} - s y_p) s dy_p \right],$$

on obtient :

$$\bar{h}_m^{M(P+1)}(Y^{(P+1)}) = \int_0^\infty f(dy_{p+1}) \dots f(dy_{n-1}) f^{*(P+1)}(s y_n^{(P+1)}) s^{n-p-1} ds,$$

en posant $y_M^{(P+1)} = 1 - \sum_{j=p+1}^{M-1} y_j$, $Y^{(P+1)}$ le vecteur de coordonnées

$$y_{p+1}, \dots, y_{M-1}.$$

On a ainsi la forme générale des densités marginales de \bar{h}_m^M ainsi que de toute autre densité \bar{h}_m^q ; elles peuvent être exprimées à l'aide des densités f^{*2} sur \mathcal{R}_+^* ; elles résultent également d'une normalisation par la somme dans $\mathcal{R}_+^{*(M-P)}$.

Conséquences : a) en particulier la densité marginale de chaque coordonnée y_k est la suivante :

$$\bar{h}_m^{q(n-1)}(y_k) = \int_0^\infty f(dy_k) f^{*(n-1)}(s - s y_k) s ds,$$

$k \neq q$

et l'on vérifie bien que l'on a :

$$\int_0^1 \frac{1}{h_m} q^{(n-1)}(u) du = \int_0^{\infty} f^{*M}(s) ds = +1.$$

b) Pour tout $k=1, \dots, M-1$, les probabilités S_{kt} s'expriment de même à l'aide des densités f^{*k} .

IV.4 - Distributions Gamma et de DIRICHLET.

Ces développements montrent le rôle particulier et commode des distributions sur \mathcal{R}_+^* qui sont "fermées" pour la convolution. C'est le cas notamment des distributions stables avec un exposant caractéristique $0 < \alpha < 1$; et de certaines distributions indéfiniment divisibles (réf. 24 tome 2). On examine ici le cas des distributions gamma, d'autres exemples sont présentés en IV.13.

4.1. Distributions gamma identiques.

Ce sont des lois indéfiniment divisibles notées $\Gamma(\alpha, \lambda)$, ($\alpha > 0, \lambda > 0$) de densités

$$f_{\alpha, \lambda}(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} \mathbb{1}_{\mathcal{R}_+}(x).$$

On sait que la r -ième itérée par convolution de $\Gamma(\alpha, \lambda)$ est $\Gamma(r\alpha, \lambda)$.

Théorème 14 La "normalisation par la somme" d'un produit de distributions indépendantes $\Gamma(\alpha_i, \lambda)$ donne la distribution de DIRICHLET sur le simplexe unité, invariante par $\sum \eta$.

Démonstration: de façon directe on a l'expression suivante pour la densité produit définie sur \mathcal{R}_+^{*M}

$$H(x_1, \dots, x_n) = \frac{\lambda^{\alpha M}}{\Gamma(\alpha)^M} (x_1 \dots x_n)^{\alpha-1} e^{-\lambda \sum x_i},$$

d'où l'on déduit que la distribution que suit le point y de coordonnées $y_i = \frac{x_i}{\sum_{j=1}^M x_j}$ $i=1, \dots, M$ a une densité h qui a pour valeur:

$$\begin{aligned} h(y_1, \dots, y_n) &= \frac{\lambda^{\alpha M}}{\Gamma(\alpha)^M} (y_1 \dots y_n)^{\alpha-1} \int_0^\infty \lambda^{\alpha n-1} e^{-\lambda s} ds \\ &= \frac{\Gamma(n\alpha)}{\Gamma(\alpha)^M} (y_1 \dots y_n)^{\alpha-1}. \end{aligned}$$

On reconnaît une densité de DIRICHLET invariante par $\sum M$ (Ref. 84)

que l'on désigne ci-après par $D(\cdot, \dots, j; a_1, \dots, a_n)$ quand on explicite les variables y_1, \dots, y_n et les paramètres a_1, \dots, a_n , ou par

$D(\cdot; a)$ avec uniquement le point-variable y et le vecteur a de coordonnées a_1, \dots, a_n .

On sait que les densités marginales de la densité de DIRICHLET sont de nouvelles densités de DIRICHLET, (Ref. 84)

$$\begin{aligned} \bar{h}_m^M \textcircled{(n-p)}(y_1, \dots, y_p) &= D(y_1, \dots, y_p; \alpha, \alpha, \dots, \alpha, (n-p)\alpha) \\ &= \frac{\Gamma(n\alpha)}{\Gamma(\alpha)^p \Gamma((n-p)\alpha)} (y_1 \dots y_p)^{\alpha-1} \left(1 - \sum_{j=1}^p y_j\right)^{(n-p)\alpha-1}, \end{aligned}$$

$$\bar{h}_m^q(y_1) = D(y_1, 1-y_1; \alpha, (n-1)\alpha) = \frac{\Gamma(n\alpha)}{\Gamma(\alpha) \Gamma((n-1)\alpha)} y_1^{\alpha-1} (1-y_1)^{(n-1)\alpha-1};$$

expressions que l'on retrouve avec les calculs du paragraphe IV.3.

Le cas de la distribution exponentielle $\alpha = 1$ donne la densité de DIRICHLET constante égale à $\Gamma(M)$; on retrouve de la sorte la mesure uniforme du chapitre II avec les quantités $y_i = x_i / \sum x_j, i=1, \dots, M; x_i$ étant M variables aléatoires exponentielles indépendantes.

Commentaires :

Cela explique le résultat de FISHER (Chap. II.3) ; en effet les X_k et Y_k de son signal aléatoire suivent la même loi normale centrée, d'où il s'ensuit que $X_k^2 + Y_k^2$ suit la loi exponentielle $\Gamma(1, 1)$. FISHER lui-même fit le rapprochement avec le théorème de recouvrement de STEVENS (Réf.28).

On voit donc comment tous deux ont examiné une question qui conduit à la densité constante $\Gamma(M)$ sur le simplexe unité S_1^M .

4.2. Le cas général.

On vérifie immédiatement qu'avec le produit de M distributions $\Gamma(a_i, \lambda)$ indépendantes, la normalisation par la somme donne la densité de DIRICHLET générale qui a pour expression:

$$D(y_1, \dots, y_n; a_1, \dots, a_n) = \frac{\Gamma(\sum a_i)}{\prod \Gamma(a_i)} \prod y_i^{a_i - 1},$$

à condition que ce soit le même paramètre λ . Ce résultat a été retrouvé à diverses reprises (JOHNSON ref.41, MOSIMANN 57, WILKS 84, ci-contre).

4.3. Compléments

On peut également obtenir une distribution sur simplexe quand on conditionne par rapport à la somme $\Delta = \sum_1^M x_i$ une distribution dans \mathbb{R}_+^M . Avec les notations précédentes on a ainsi $cl(\Delta) = \int_{S_1^M} H(\Delta y) \Delta^{n-1} dy_1 \dots dy_{n-1}$ la valeur de la densité marginale de la somme Δ et $\Delta^{n-1} H(\Delta y) / cl(\Delta)$

la valeur de la densité de y conditionnée par rapport à Δ , $y \in S_1^M$.

Si l'on reprend le cas des distributions gamma $\Gamma(a_i, \lambda)$, on sait que la somme Δ suit la loi $\Gamma(\sum a_i, \lambda)$, on a la factorisation suivante :

$$\begin{aligned} \prod_{G_i, \lambda} (\Delta y_i) \Delta^{n-1} &= \frac{e^{-\lambda \Delta} \Delta^{\sum (a_i - 1)}}{\lambda^{\sum a_i} \prod \Gamma(a_j)} \Delta^{n-1} \prod_1^M y_k^{a_k - 1} \\ &= \frac{e^{-\lambda \Delta} \Delta^{\sum a_i - 1}}{\lambda^{\sum a_i} \Gamma(\sum a_j)} \times \frac{\Gamma(\sum a_k)}{\prod \Gamma(a_j)} \times \\ &\quad \times \prod y_n^{a_n - 1} \end{aligned}$$

On retrouve la densité marginale de Δ , et on constate que la distribution de y conditionnée par rapport à Δ est la distribution de DIRICHLET précédente qui ne dépend pas de Δ .

4.4. Commentaires sur la densité de DIRICHLET.

La densité de DIRICHLET - et la distribution qui lui correspond - sont dues à une relation établie par LEJEUNE-DIRICHLET en 1839 sur le simplexe unité de \mathbb{R}^3 et de \mathbb{R}^M (Réf. 80, 82, 84). Il semble que ce soit WILKS qui la baptise ainsi le premier (Réf. 84), mais elle apparaît dans des travaux antérieurs. D'après JOHNSON, K. PEARSON signale lui-même que L.N.G. FILON l'a mentionnée en 1901 (Réf. 41). GOOD l'utilise et en donne un historique rapide jusque vers 1960, (l'époque de son baptême) de la part d'auteurs Bayésiens, et ne mentionne pas FILON, (Réf. 33). Vers cette époque JOHNSON, MOSIMANN, MAULDON l'utilisent et la retrouvent, sans la nommer (Réf. 41, 55, 57). DARLING indique la fonction caractéristique de la distribution qui correspond à la mesure uniforme, (la densité constante $\Gamma(M)$), (Réf. 19). C'est sur \mathbb{R}^{n-1} la généralisation de la distribution classique Béta de 1^o espèce; cependant elle ne doit pas être confondue avec de nombreuses autres distributions que l'on désigne par "distributions Béta généralisées et distributions Béta généralisées multivariées", qui sont essentiellement des généralisations matricielles de la distribution Béta scalaire, de la même façon que les matrices de WISHART sont des généralisations matricielles du χ^2 scalaire (Réf. 43, 61, 77).

La distribution de DIRICHLET est caractérisée de différentes manières; elle intervient aussi le plus souvent avec seulement $M-1$ variables $y_i \geq 0, i=1, \dots, M-1$ indépendantes, $\sum y_i \leq 1$.

i) MOSIMANN s'appuie sur un résultat de LUKACS; y_1, \dots, y_M suivent une distribution de DIRICHLET si et seulement si \bar{h}^M est la densité qui résulte de la normalisation par la somme du produit de densités

$$\prod_{i=1}^M f_i(y_i), \quad y_i \in \mathbb{R}_+^* \quad \text{avec} \quad y_j / \sum_1^M y_k \quad j=1, \dots, M-1$$

indépendants de la somme $\sum_1^M y_i$ (condition de LUKACS); alors les f_i sont égales entre elles et égales à la densité d'une loi gamma, (ref. 57).

i i) DARROCH et RATCLIFF montrent que sous réserve de la continuité des distributions en jeu, le vecteur y de coordonnées y_1, \dots, y_M suit une distribution de DIRICHLET si et seulement si pour tout $j = 1, \dots, M-1$,

$$y_j / (y_j + y_M) \quad \text{et} \quad y_k$$

sont indépendants, quel que soit $k = 1, \dots, M-1$, $k \neq j$ (réf. 20).

ii) Pour CONNOR et MOSIMANN un vecteur y_1, \dots, y_M est "complètement neutre" si $y_1, \frac{y_2}{1-y_1}, \dots, \frac{y_M}{1-\sum_{j=1}^{M-1} y_j}$ sont indépendantes. Ils signalent que moyennant des conditions de régularité, KRUSKALL a montré que y_1, \dots, y_M suit une distribution de DIRICHLET si et seulement si pour tout $j = 1, \dots, M-1, \frac{y_j}{y_j + y_M}$ et y_k sont indépendants, quel que soit $k = 1, \dots, M-1, k \neq j$ (ref. 16).

La densité de DIRICHLET a donné lieu à une étude et une tabulation très détaillées, sous la forme particulière de l'intégrale de DIRICHLET incomplète de type I par SOBEL, UPPULURI et FRANKOWSKI (Réf. 73). Dans cette analyse l'intégrale $I_p^{M-1}(\alpha, T-1)$ est obtenue en intégrant

sur $y_j \leq \rho \quad j = 1, \dots, M-1$ la densité où $a_j = a \quad j = 1, \dots, M-1, a_M = T - (M-1)a$;

voir aussi RAO et SOBEL (Réf. 67).

Parmi toutes les densités de DIRICHLET, la densité constante est très répandue en statistique. On a vu comment on l'obtient à partir de la distribution uniforme sur $[0, 1]$, Chap. II.3. Plus généralement si X est une v.a. qui possède une f.r. F continue, on sait que $F(X)$ est la v.a. uniforme sur $[0, 1]$. WILKS reprend cette propriété pour découper le segment $[0, 1]$ de manière uniforme à l'aide de "coverages" (Réf. 84). PYKE indique d'autres méthodes qui conduisent à la distribution uniforme sur $[0, 1]$, (Réf. 66) ; tous ces cas, on l'a vu, aboutissent à la distribution constante sur le simplexe unité de \mathbb{R}^M , $M \geq 2$.

IV.5 - Distribution de λ_1 avec une distribution de DIRICHLET.

Dans certains cas, il est possible d'établir plus ou moins complètement la f.r. de λ_1 lorsque le point y du simplexe unité suit une distribution de DIRICHLET.

5.1. Etude de l'intervalle $[1/M, 1/(M-1)]$.

5.1.1. Le théorème 2 du chapitre II signifie que pour tout t appartenant à l'intervalle $[1/M, 1/(M-1)]$, le changement de variables

$$y_i = (1-tM)\xi_i + t, \quad i=1, \dots, M \quad \text{de jacobien} \quad \frac{D(y_1, \dots, y_{n-1})}{D(\xi_1, \dots, \xi_{n-1})} = (1-tM)^{M-1}$$

établit une correspondance biunivoque entre le point $\xi = (\xi_1, \dots, \xi_n) \in S_1^M$ et le point y du sous-ensemble \bar{E}_t .

On a donc :

$$\begin{aligned} \mu(E_t) &= \int_{S_1^M} R((1-tM)\xi + t) \cdot |1-tM|^{n-1} d\xi_1 \dots d\xi_{n-1} \\ &= \frac{\Gamma(n\alpha)}{\Gamma(t)^M} |1-tM|^{n-1} \int_{S_1^M} \prod_{j=1}^M [(1-tM)\xi_j + t]^{\alpha-1} d\xi_1 \dots d\xi_{n-1} \end{aligned}$$

lorsque y suit une distribution de DIRICHLET invariante.

5.1.2. On a l'identité :

$$\prod_{j=1}^M [(1-tM)\xi_j + t] = \sum_{q=0}^M (1-tM)^q t^{M-q} \sigma_q,$$

où les σ_q sont les fonctions symétriques fondamentales des ξ_i ($i=1, \dots, n$; $\sigma_0=1, \sigma_M=1$); le produit est un polynôme en t de degré M , avec des coefficients monômes en ξ_i . Lorsque le paramètre α de la densité invariante est un entier; la fonction sous l'intégrale est un polynôme en t de degré $M(\alpha-1)$ avec des coefficients que l'on sait intégrer sur le simplexe unité avec l'intégrale de DIRICHLET.

Par exemple :

$$\int_{S_1^M} \sigma_q d\xi_1 \dots d\xi_{n-1} = C_M^q \frac{\Gamma(2)^q \Gamma(1)^{n-q}}{\Gamma(n+q)} = \frac{C_M^q}{\Gamma(n+q)}.$$

5.1.3. La cas $\alpha = 2$ est immédiat. On a en effet:

$$P_{\text{rob}}^{(n)}(\lambda_1 \leq t) = \Gamma(2M) |1-tM|^{M-1} \sum_{q=0}^M \frac{C_M^q}{\Gamma(M+q)} (1-tM)^q t^{M-q};$$

soit pour $M = 3$:

$$P_{\text{rob}}^{(3)}(\lambda_1 \leq t) = \Gamma(6) |1-3t|^2 \left[\frac{t^3}{\Gamma(3)} + \frac{3}{\Gamma(4)} t^2(1-3t) + \frac{3}{\Gamma(5)} t(1-3t)^2 + \frac{1}{\Gamma(6)} (1-3t)^3 \right]$$

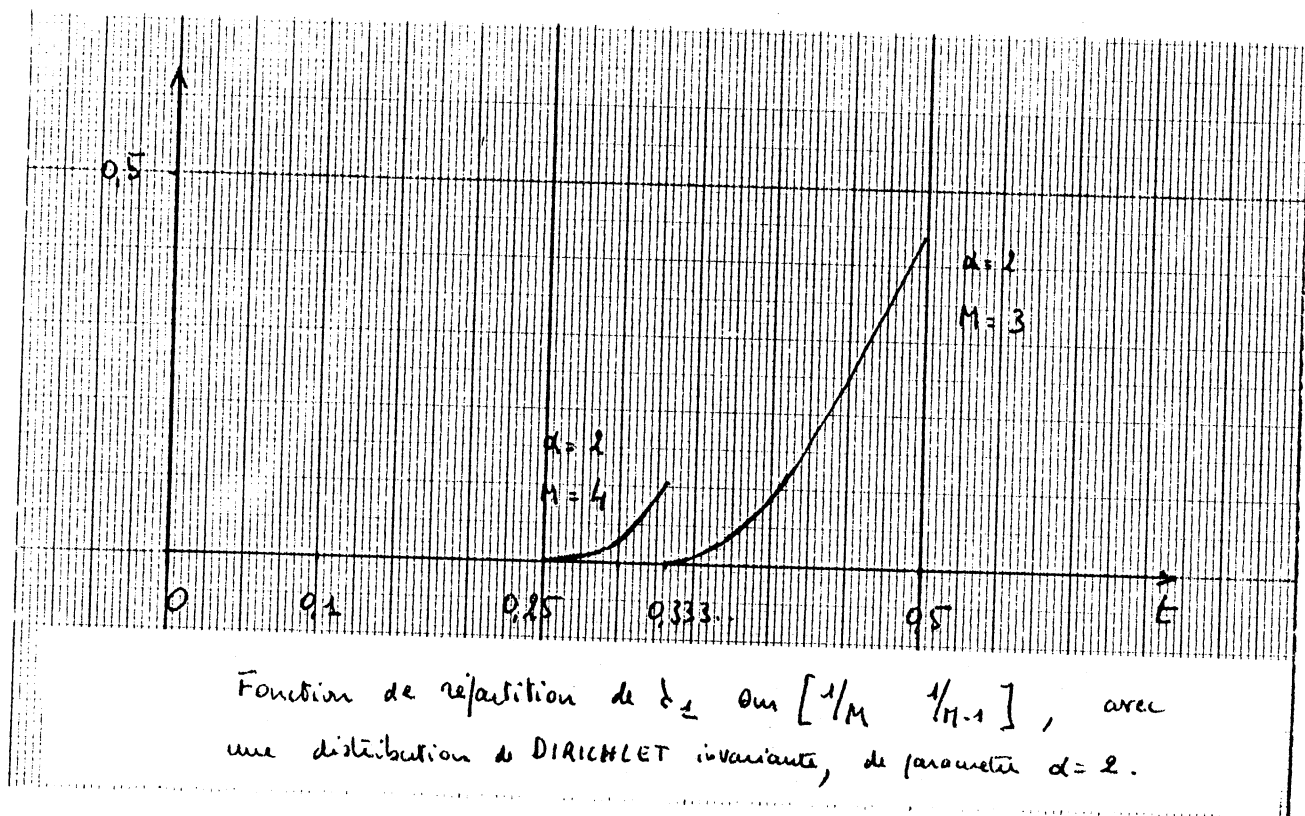
sur $[1/3 \quad 1/2]$;

et pour $M = 4$:

$$P_{\text{rob}}^{(4)}(\lambda_1 \leq t) = \Gamma(8) |1-4t|^3 \left[\frac{t^4}{\Gamma(4)} + \frac{4}{\Gamma(5)} t^3(1-4t) + \frac{6}{\Gamma(6)} t^2(1-4t)^2 + \frac{4}{\Gamma(7)} t(1-4t)^3 + \frac{1}{\Gamma(8)} (1-4t)^4 \right]$$

sur $[1/4 \quad 1/3]$.

Les variations graphiques de ces deux probabilités sont indiquées dans la figure ci-contre:



5.2. Etude des autres intervalles $[1/k+1, 1/k]$, $k \geq 1$.

La f.r de λ_1 s'obtient très facilement sur $[1/2, 1]$ à l'aide de la fonction Beta incomplète ; c'est un polynôme quand α est un nombre entier positif.

On a :

$$\mu(\bar{E}_t) = 1 - \mu(E_t)$$

$$P_{rob}(\lambda_1 \leq t) = 1 - \frac{1}{B(\alpha, (n-1)\alpha)} \int_t^1 u^{\alpha-1} (1-u)^{(n-1)\alpha-1} du$$

On en déduit donc la f.r complète de λ_1 dans le cas $\alpha = 2$, $M = 3$ (cf 5.1.3.).

Sur $[1/k+1, 1/k]$ $k = 2, \dots, M-1$ on peut utiliser le

changement de variables classique qui transforme un simplexe unité en un hypercube $[0, 1]^{M-1}$ (Réf. 34, 82, 84) :

$$\begin{cases} y_1 = \theta_1 \\ y_2 = \theta_2(1-\theta_1) \\ \vdots \\ y_{n-1} = \theta_{n-1}(1-\theta_{n-2}) \dots (1-\theta_2)(1-\theta_1) \end{cases}$$

Tout ensemble $E_t^1 \cap \dots \cap E_t^k$, $k = 2, \dots, n-1$ se transforme comme suit :

$$\begin{aligned} t &\leq \theta_1 \leq 1 - t(k-1) \\ \frac{t}{1-\theta_1} &\leq \theta_2 \leq 1 - t \frac{k-2}{1-\theta_1} \\ &\vdots \\ \frac{1}{(1-\theta_1) \dots (1-\theta_{k-1})} &\leq \theta_k \leq 1 \end{aligned}$$

La probabilité de cet ensemble s'exprime ainsi comme une succession d'intégrales Béta incomplètes en θ_i , comprises entre les bornes ci-dessus. Les calculs sont algébriques quand le paramètre α est un nombre entier.

5.3. Le critère de COCHRAN et extensions .

5.3.1. COCHRAN étudie la distribution de $\frac{\text{sup } x_i}{\sum x_j}$ où les x_i sont M variables indépendantes qui suivent la loi du χ^2 avec ν degrés de liberté (Réf. 15). D'après ce qui précède on a une distribution de DIRICHLET avec le paramètre $\alpha = \nu/2$. On retrouve pour le $\text{sup } y_i$ les mêmes expressions Béta incomplètes. COCHRAN insiste sur le cas algébrique des ν pairs, malgré cela dit-elle, "the expressions rapidly become complicated".

Elle précise deux points :

a) La table des valeurs de $\frac{\text{sup } x_i}{\sum x_j}$ dépassées avec une probabilité de 5% pour $3 \leq M \leq 10$ et $2\alpha = 1, 2, 3, 4, 5, 6, 8, 10$. Cette table est également donnée par EISENHART, HASTRAY, WALLIS (Réf. 22, 44).

b) Pour M élevé, elle généralise le résultat de P. LEVY; la loi de la variable $\frac{M \nu \ln 2}{2} - \log M$ tend vers la loi des extrêmes de FRECHET - FISHER - TIPPETT quand M tend vers plus l'infini, (contesté par DARLING).

5.3.2. DARLING étudie la distribution de $\frac{\sum x_j}{\text{sup } x_i}$ où les x_i indépendantes suivent une même loi, pour la recherche de valeurs aberrantes. (outliers). Dans le cas des x_i qui suivent une loi du χ^2 il indique la fonction de répartition de sa variable qui somme toute est une expression assez peu complexe, (Réf. 18). Quand M tend vers plus l'infini il indique certaines formes limites ; par exemple quand les x_i suivent la distribution $\Gamma(\alpha, 1)$, la loi de la variable $d_M - \frac{1}{\lambda_1 \times c_M}$ tend vers la loi des

extrêmes de FRECHET - FISHER - TIPPETT ; avec

$$c_M = \alpha M / \log^2 M$$

$$d_M = \log M - (\alpha - 1) \log \log M + \log \Gamma(\alpha) .$$

Il indique d'autres lois limites pour d'autres lois suivies par les x_i .

5.4. Remarque.

Avec une loi conjointe de densité :

$$h(y_1, \dots, y_n) = \int_0^\infty f(y_1) \dots f(y_{n-1}) f(y_n) s^{n-1} ds, \quad \sum_{i=1}^n y_i = 1,$$

chaque coordonnée y_i est une v.a sur $[0, 1]$ dont la densité a pour valeur:

$$\bar{h}_m(u) = \int_0^\infty f(su) \int_0^{1-(n-1)s} (s-su) s ds.$$

On a donc un ensemble de M v.a qui suivent la même loi. Le problème de la distribution suivie par $\sup y_i$ n'est cependant pas celui des valeurs extrêmes comme l'ont étudié FRECHET, FISHER et TIPPETT (Réf. 38), puisqu'ici les y_i sont liées par la relation $\sum_{i=1}^n y_i = 1$.

Exemple. On prend la densité de la distribution gamma qui a pour valeur

$$f(x) = \frac{1}{\Gamma(\alpha)} x^{\alpha-1} e^{-x} \mathbb{1}_{\mathbb{R}_+^+}(x),$$

d'où

$$h(y_1, \dots, y_n) = \frac{\Gamma(n\alpha)}{\Gamma(\alpha)^n} (y_1 \dots y_n)^{\alpha-1},$$

et

$$\bar{h}_m(u) = \frac{\Gamma(n\alpha)}{\Gamma(\alpha) \Gamma((n-1)\alpha)} u^{\alpha-1} (1-u)^{(n-1)\alpha-1}.$$

Si l'on a M variables y_i qui suivent cette distribution bêta, et qui sont indépendantes, on sait alors que la densité de la distribution de $\sup y_i$

a pour valeur

$$g(x) = M \left[\frac{\Gamma(n\alpha)}{\Gamma(\alpha) \Gamma((n-1)\alpha)} \right]^M \left[\int_0^x u^{\alpha-1} (1-u)^{(n-1)\alpha-1} du \right]^{n-1} x^{\alpha-1} (1-x)^{(n-1)\alpha-1}$$

avec $\alpha = 1$: $g(x) = M(n-1) \left[1 - (1-x)^{n-1} \right]^{n-1} (1-x)^{n-2}$, on vérifie

qu'elle est différente de la densité de la distribution de WHITWORTH, cf II.2.

IV.6 - Distribution de λ_M et de λ_1, λ_2 avec une distribution de DIRICHLET.

6.1. Le cas de la variable $\lambda_M = \inf\{y_i\}$ est beaucoup plus simple. Nous savons en effet que cette variable ne prend ses valeurs que sur l'intervalle

$$[0 \quad 1/M]$$

Le théorème 3 du chapitre II entraîne que le changement de variables

$$y_i = (1-uM)\xi_i + u; \quad i=1, \dots, M \text{ de jacobien } \frac{D(y_1, \dots, y_{n-1})}{D(\xi_1, \dots, \xi_{n-1})} = (1-uM)^{M-1} \text{ établit une}$$

correspondance biunivoque entre le point $\xi = (\xi_1, \dots, \xi_n)$ de S_1^M et le point y de son sous-ensemble G_u pour tout u appartenant à l'intervalle

$$[0 \quad 1/M]$$

On a donc

$$\mu(G_u) = \int_{S_1^M} h((1-uM)\xi + u) \cdot |1-uM|^{M-1} d\xi_1 \dots d\xi_{n-1};$$

avec une densité invariante de DIRICHLET définie par le paramètre α , nous sommes ramenés à des calculs identiques à ceux du paragraphe III.5.1.

Par exemple, en prenant $\alpha = 2$:

$$\text{Prob}_{(M)}(\lambda_M \leq u) = 1 - \frac{\Gamma(2M)}{\Gamma(M)^2} |1-uM|^{M-1} \sum_{q=0}^{M-1} \frac{C_M^q}{\Gamma(M+q)} (1-uM)^q u^{M-q}.$$

6.2. Le cas des densités invariantes de DIRICHLET est un de ceux que l'on sait traiter en partie pour le couple λ_1, λ_2 , cf III.3.

6.2.1. Proposition 17 Pour tout λ_1 supérieur ou égal à $1/2$, le point x de coordonnées $x_p = y_p / (1 - \lambda_1)$ pour tout indice p tel que $y_p < \lambda_1$, suit une distribution de DIRICHLET invariante par le groupe Σ_{M-1} sur le simplexe S_1^{M-1} .

Démonstration : Le point y de S_1^M suit la distribution de DIRICHLET invariante par Σ_M , définie par le paramètre α . On a établi que la densité marginale de λ_1 à la valeur suivante :

$$\tilde{g}_1(\lambda_1) = \frac{M \Gamma(M\alpha)}{\Gamma(\alpha)^M \Gamma((M-1)\alpha)} \lambda_1^{\alpha-1} (1-\lambda_1)^{(M-1)\alpha-1}$$

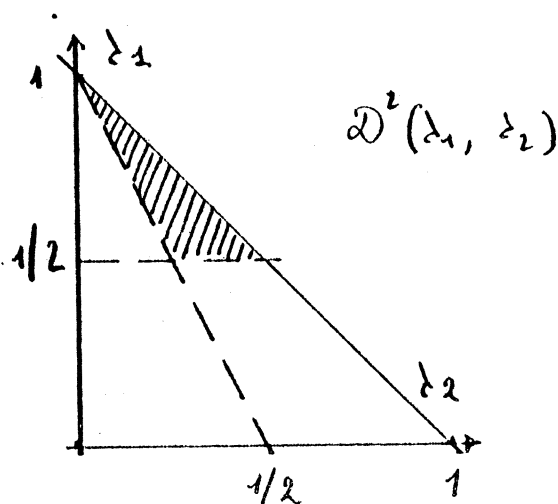
L'événement $\{\lambda_1 = \lambda_2\}$ est de probabilité nulle. On définit le point \tilde{y} de coordonnées $y_p = \lambda_p / (1 - \lambda_1)$, $p \geq 2$ sur le simplexe ordonné S_1^{M-1} , la densité \tilde{g} de la distribution de \tilde{y} a donc pour valeur

$$\tilde{g}(\tilde{y}) = (M-1)! \frac{\Gamma((M-1)\alpha)}{\Gamma(\alpha)^{M-1}} (1 - \lambda_1)^{\alpha-1},$$

ce qui correspond à la densité de DIRICHLET invariante par Σ_{M-1} sur S_1^{M-1} définie par le paramètre α . ■

6.2.2. Soit $\mathcal{D}^2(\lambda_1, \lambda_2)$ le sous-ensemble de $\mathcal{D}(\lambda_1, \lambda_2)$ défini par

$$\begin{cases} 1/2 \leq \lambda_1 \leq 1 \\ \frac{1-\lambda_1}{2} \leq \lambda_2 \leq \lambda_1 \end{cases}$$



On en déduit la densité d_M^λ définie sur $\mathcal{D}^2(\lambda_1, \lambda_2)$ qui a pour valeur

$$d_M^\lambda(\lambda_1, \lambda_2) = D_M^\lambda(\lambda_1) \cdot D_{M-1}^\lambda\left(\frac{\lambda_2}{1-\lambda_1}\right) \cdot \frac{1}{1-\lambda_1},$$

avec

$$D_M^\lambda(\lambda_1) = \frac{M \Gamma(M\alpha)}{\Gamma(\alpha) \Gamma((M-1)\alpha)} \lambda_1^{\alpha-1} (1-\lambda_1)^{(M-1)\alpha-1}.$$

Remarque - La fonction de répartition de la loi du couple (λ_1, λ_2) s'exprime avec la fonction Bêta incomplète sur $\mathcal{D}^2(\lambda_1, \lambda_2)$.

En effet si le point $(a, b) \in \mathcal{D}^2(\lambda_1, \lambda_2)$,

$$I_{ab} = \int_a^{1-b} D_M^\lambda(\lambda_1) d\lambda_1 \int_b^{1-\lambda_1} D_{M-1}^\lambda\left(\frac{\lambda_2}{1-\lambda_1}\right) \frac{d\lambda_2}{1-\lambda_1}$$

$$= \int_a^{1-b} \left[1 - D_{M-1}^\lambda\left(\frac{b}{1-\lambda_1}\right) \right] D_M^\lambda(\lambda_1) d\lambda_1$$

$$= D_M^\lambda(1-b) - D_M^\lambda(a) - J_{ab},$$

en posant

$$J_{ab} = \int_a^{1-b} D_{M-1}^\lambda\left(\frac{b}{1-\lambda_1}\right) D_M^\lambda(\lambda_1) d\lambda_1.$$

On a

$$J_{ab} = \frac{M(M-1) \Gamma(M\alpha)}{\Gamma(\alpha)^2 \Gamma((M-2)\alpha)} b^{\alpha-1} \int_a^{1-b} [\alpha-1-\lambda_1(M\alpha-2)] \lambda_1^{\alpha-2} (1-\lambda_1-b)^{(M-2)\alpha-1} d\lambda_1$$

On effectue le changement de variables $t = \frac{x-a}{1-b}$ $dt = \frac{dx}{1-b}$;

$$\int_a^b \frac{\eta(\eta-1) \Gamma(\eta \alpha)}{\Gamma(\alpha)^2 \Gamma((\eta-1)\alpha)} b^{\alpha-1} \int_{\frac{a}{1-b}}^1 (1-b)^{(\eta-1)\alpha-3} t^{\alpha-2} (1-t)^{(\eta-2)\alpha-1} x$$

$$\left[(1-b)(\alpha-1) - t(\eta\alpha-2) \right] dt ;$$

$$\int_a^b = \frac{\eta(\eta-1) \Gamma(\eta \alpha)}{\Gamma(\alpha)^2 \Gamma((\eta-2)\alpha)} b^{\alpha-1} (1-b)^{(\eta-1)\alpha-3} x$$

$$\lambda \left[(1-b)(\alpha-1) \int_{\frac{a}{1-b}}^1 t^{\alpha-2} (1-t)^{(\eta-2)\alpha-1} dt - (\eta\alpha-2) \int_{\frac{a}{1-b}}^1 t^{\alpha-1} (1-t)^{(\eta-2)\alpha-1} dt \right]$$

B - Changement de variables sur le simple unité.

Quel que soit le point y d'un simplexe unité, et quelle que soit l'application φ définie sur $[0, 1]^M$ à valeur dans \mathbb{R}_+ , on définit

$$z_i = \varphi(y_i) / \sum_{j=1}^M \varphi(y_j) \quad i=1, \dots, M \quad . \text{ Les } z_i \text{ sont les coordonnées d'un}$$

point z qui appartient également au même simplexe unité. On note $\bar{\varphi}$ cette application du simplexe unité dans lui-même ; on étudie dans cette partie les applications $\bar{\varphi}$ et leurs conséquences sur les densités de probabilité \mathcal{R} définies sur le simplexe unité.

IV-7 - Etude de l'application $\bar{\varphi}$.

Nous examinons des conditions pour que $\bar{\varphi}$ soit une application bijective (automorphisme algébrique) ; on se limite ici au cas des applications φ continues dérivables à dérivée continue. Sachant que φ et $\lambda\varphi$, λ réel $\neq 0$ conduisent à la même application $\bar{\varphi}$, on étudie uniquement les φ à valeur dans \mathbb{R}_+ . On étudie des conditions pour que $\bar{\varphi}$ soit surjective, puis injective.

7.1. Par définition l'application $\bar{\varphi}$ est surjective si pour tout z de S_1^M il existe un y de S_1^M tel que $z = \bar{\varphi}(y)$. Pour tout z ,

on considère les quantités $t z_i$ avec le scalaire réel positif t ,

$$x_i(t) = \varphi^{-1}(t z_i) \geq 0 \quad i=1, \dots, M \quad \text{et} \quad \mathcal{A}(t) = \sum_1^M x_i(t).$$

Par conséquent si pour tout z de S_1^M il existe $t^*(z)$ tel que $\mathcal{A}[t^*(z)] = 1$ alors $\bar{\varphi}$ est une surjection.

Théorème 15 Si φ est une application définie continue dérivable strictement croissante de $[0, 1[$ dans \mathbb{R}_+ , avec $\varphi(0) = 0$,

l'application $\bar{\varphi}$ est une surjection du simplexe sur lui-même; lorsque $\varphi(1)$ est fini on prend comme intervalle de définition de φ l'intervalle $[0, 1]$.

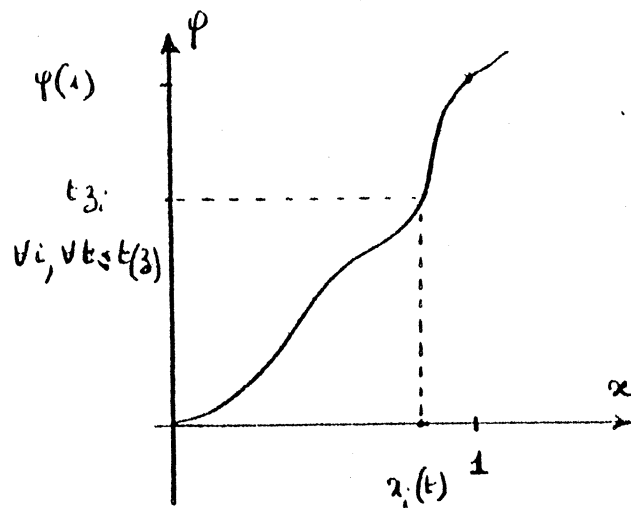
Démonstration :

i) Supposons $\varphi(1)$ fini, on pose $t(z) = \varphi(1) / \sup(z_i)$, c'est une quantité définie pour tout z du simplexe, supérieure strictement à $\varphi(1)$

excepté pour les points-extrémaux $z = e^k$, $k=1, \dots, M$. Pour

$t \in [0, t(z)]$ tous les $t z_i$ appartiennent à l'espace d'arrivée $[0, \varphi(1)]$. Il en résulte que les applications x_i sont définis dans $[0, 1]$, et que l'application $\mathcal{A} = \sum_1^M x_i$ est continue dérivable. On a la valeur de la dérivée $\mathcal{A}'(t) = \sum_1^M z_i / \varphi'(x_i(t))$ positive, $\mathcal{A}(0) = 0$ et

$\mathcal{A}(t(z))$ supérieur ou égal à 1, (l'égalité n'est obtenue que pour les $z = e^k$) ; on en déduit l'existence d'un $t^*(z)$ unique tel que $\mathcal{A}(t^*(z)) = 1$.



Le point de coordonnées $\varphi^{-1}(t^*(z), z_i)$ $i=1, \dots, M$ est un point du simplexe dont son image par $\bar{\varphi}$ est le point z , unique par construction. Dans le cas des $z = e^k$ on a la solution évidente $y = e^k$ $k=1, \dots, M$.

ii) On a $\varphi(x) = +\infty$. Pour tout z différent des e^k on généralise immédiatement ce qui précède en prenant $t(z) = +\infty$; on a

$$\lim_{t \rightarrow \infty} \lambda(t) > 1 \quad \text{d'où un } t^*(z) \text{ unique.}$$

Pour $z = e^k$ on vérifie par continuité que $y = e^k$ est solution unique de $\bar{\varphi}(y) = e^k$.

Remarque : Les points e^k sont invariants par $\bar{\varphi}$, de même que $\omega = \frac{1}{M} \sum_k e^k$.

Théorème 16 - Sous les mêmes conditions $\bar{\varphi}$ est injective, donc bijective.

Soient y^1 et y^2 deux points distincts, qui diffèrent nécessairement par l'intermédiaire de 2 coordonnées au moins, d'indices k et l , pour lesquelles on a les inégalités strictes :

$$y_k^1 > y_k^2, \quad y_l^1 < y_l^2.$$

Supposons que l'on ait $z_k^1 \leq z_k^2$ et $z_e^1 \geq z_e^2$, ceci entraîne

que $\frac{\varphi(y_k^1)}{\phi^1} \leq \frac{\varphi(y_k^2)}{\phi^2}$ et $\frac{\varphi(y_e^2)}{\phi^2} \leq \frac{\varphi(y_e^1)}{\phi^1}$ d'où

$$\varphi(y_k^1) \cdot \varphi(y_e^2) \leq \varphi(y_k^2) \cdot \varphi(y_e^1) \quad \text{ce qui est en contra-}$$

diction avec l'hypothèse, on a donc $z_k^1 > z_k^2$ ou $z_e^1 > z_e^2$ ■

Note . On a posé $\phi^1 = \sum_i \varphi(y_i^1)$ et $\phi^2 = \sum_i \varphi(y_i^2)$.

7.2. Les applications $\bar{\varphi}$ ont les propriétés suivantes :

i) $\bar{\varphi}$ ne modifie pas l'ordre des coordonnées classées par valeurs croissantes (ou décroissantes).

ii) Les points extrémaux e^k sont des points fixes de $\bar{\varphi}$; plus généralement à tout point frontière $y_{k_1} = y_{k_2} = \dots = y_{k_e} = 0$

correspond un point de la même frontière $z_{k_1} = z_{k_2} = \dots = z_{k_e} = 0$.

iii) L'application qui correspond à $\varphi(x) = x$ est la bijection identité.

iv) Chaque bijection $\bar{\varphi}$ possède une bijection réciproque $\bar{\varphi}^{-1}$; la composée $\bar{\varphi} \circ \bar{\varphi}^{-1}$ de deux bijections est naturellement une bijection. On n'étudie pas ici la question de savoir si pour $\bar{\varphi}^{-1}$ et $\bar{\varphi} \circ \bar{\varphi}$ il existe des applications \bar{S}_1 et \bar{S}_2 de $[0, 1]$ dans \mathcal{R}_+ , nulles à l'origine, telles que $\bar{\varphi}^{-1} = \bar{S}_1$ et $\bar{\varphi} \circ \bar{\varphi} = \bar{S}_2$; c'est-à-dire l'étude du sous-groupe éventuel des automorphismes $\bar{\varphi}$ construits à partir des applications monotones croissantes positives φ de $[0, 1]$ dans \mathcal{R}_+ , nulles à l'origine. On verra cependant un cas particulier simple de sous-groupe en 8.2.

IV.8 - Calcul du Jacobien.

Les applications bijectives $\bar{\varphi}$ précédentes définissent un changement de variables entre sous-ensembles de \mathbb{R}^{n-1} avec un Jacobien $\frac{D(\beta)}{D(\gamma)}$ noté $J_{\varphi}(\gamma)$.

8.1. Théorème 17 - Le Jacobien J_{φ} de l'application $\bar{\varphi}$ est une fonction de $\gamma_1, \dots, \gamma_M$ qui est invariante par le groupe des permutations Σ_M , sa valeur est égale à :

$$\frac{1}{\left[\sum_i \varphi(\gamma_i) \right]^M} \sum_{k=1}^M \varphi'(\gamma_1) \cdots \varphi'(\gamma_{k-1}) \varphi(\gamma_k) \varphi'(\gamma_{k+1}) \cdots \varphi'(\gamma_M).$$

Démonstration :

i) Sur le simplexe unité S_1^M on prend les $M-1$ variables indépendantes $\gamma_1, \dots, \gamma_{M-1}$ pour γ et $\beta_1, \dots, \beta_{M-1}$ pour β . On pose $\phi = \sum_1^M \varphi(\gamma_i)$, $\varphi_i = \varphi(\gamma_i)$, $\varphi'_i = \varphi'(\gamma_i)$.

$$\text{On a } dz_j = \frac{\varphi'_j}{\phi} dy_j - \frac{\varphi_j}{\phi^2} \left(\sum_1^{M-1} \varphi'_k dy_k - \varphi'_M \sum_1^{n-1} dy_q \right)$$

$$= \frac{\varphi'_j}{\phi} dy_j - \frac{\varphi_j}{\phi^2} \sum_{k=1}^M (\varphi'_k - \varphi'_M) dy_k$$

pour tout $j=1, \dots, M-1$.

D'où le Jacobien

$$\frac{D(\beta_1, \dots, \beta_{M-1})}{D(\gamma_1, \dots, \gamma_{M-1})} = \frac{1}{\phi^{M-1}} D(1, 2, \dots, M-1; M; \phi),$$

en posant le déterminant suivant :

$$D = \begin{vmatrix} \varphi'_1 - \frac{\varphi_1}{\phi} (\varphi'_1 - \varphi'_n) & -\frac{\varphi_1}{\phi} (\varphi'_2 - \varphi'_n) & \dots & -\frac{\varphi_1}{\phi} (\varphi'_{n-1} - \varphi'_n) \\ -\frac{\varphi_2}{\phi} (\varphi'_1 - \varphi'_n) & \varphi'_2 - \frac{\varphi_2}{\phi} (\varphi'_2 - \varphi'_n) & \dots & -\frac{\varphi_2}{\phi} (\varphi'_{n-1} - \varphi'_n) \\ & & & \vdots \\ -\frac{\varphi_{n-1}}{\phi} (\varphi'_1 - \varphi'_n) & \dots & \dots & \varphi'_{n-1} - \frac{\varphi_{n-1}}{\phi} (\varphi'_{n-1} - \varphi'_n) \end{vmatrix}$$

de terme général $\varphi'_i \delta_{ij} - \frac{\varphi_i}{\phi} (\varphi'_i - \varphi'_n)$, $i=1, \dots, n-1$, $j=1, \dots, n-1$.

i) Si l'on utilise la linéarité de D en fonction de la première ligne, on peut écrire que D est la somme de deux déterminants. Le premier est égal au produit de φ'_1 par le déterminant $D(2, 3, \dots, n-1; n; \phi)$ de dimensions $(n-2, n-2)$; le second est le produit de $-\frac{\varphi_1}{\phi} (\varphi'_1 - \varphi'_n)$ par le déterminant suivant :

$$\begin{vmatrix} 1 & \varphi'_2 - \varphi'_n & \dots & \varphi'_{n-1} - \varphi'_n \\ -\frac{\varphi_2}{\phi} & & & \\ \vdots & & & \\ -\frac{\varphi_{n-1}}{\phi} & & & \end{vmatrix} \begin{vmatrix} & & & \\ & & & \\ & & & \\ & & & \end{vmatrix} D(2, 3, \dots, n-1; n; \phi)$$

Ce dernier déterminant est égal au produit $\varphi'_2 \varphi'_3 \dots \varphi'_{n-1}$ (ajouter à la ligne $k \geq 2$ la 1^o ligne multipliée par φ_k / ϕ).

On en déduit la relation :

$$D(1, 2, \dots, n-1; n; \phi) = \varphi'_1 D(2, 3, \dots, n-1; n; \phi) - \frac{\varphi'_1 - \varphi'_n}{\phi} \varphi_1 \varphi'_2 \dots \varphi'_n ;$$

et les relations analogues par récurrence, jusqu'à

$$D(n-1; n; \phi) = \varphi'_{n-1} - \frac{\varphi'_{n-1} - \varphi'_n}{\phi} \varphi_{n-1} ,$$

d'où le résultat annoncé :

$$D(1, 2, \dots, n-1; n; \phi) = \frac{1}{\phi} \sum_{k=1}^M \varphi'_1 \dots \varphi'_{k-1} \varphi_k \varphi'_{k+1} \dots \varphi'_n \quad \blacksquare$$

Remarque : En tout point intérieur $y_j \neq 0, j=1, \dots, M$ on a $\varphi_j \neq 0, j=1, \dots, M$.

on peut mettre le Jacobien sous la forme suivante :

$$\frac{D(z_1, \dots, z_{n-1})}{D(y_1, \dots, y_{n-1})} = z_1 \dots z_M \sum_{k=1}^M \frac{\varphi'_1}{\varphi_1} \dots \frac{\varphi'_{k-1}}{\varphi_{k-1}} \frac{\varphi'_{k+1}}{\varphi_{k+1}} \dots \frac{\varphi'_n}{\varphi_n} .$$

8.2. Exemple .

Pour tout u positif on considère l'application φ_u qui pour tout $x \in [0, 1]$ a la valeur $\varphi_u(x) = x^u$, elle vérifie les conditions des théorèmes précédents .

$$i) \text{ On a } z_i = \frac{y_i^u}{\sum_{k=1}^M y_k^u} \quad i=1, \dots, M ; \quad \varphi(x) = \frac{x}{u} \varphi'(x) ,$$

Le Jacobien se met sous la forme simplifiée suivante :

$$\begin{aligned} J_u(y) &= \frac{u^{M-1}}{(\sum_{k=1}^M y_k^u)^M} \sum_{k=1}^M y_1^{u-1} \dots y_k^{u-1} \dots y_M^{u-1} , \\ &= \frac{u^{M-1} (y_1 \dots y_M)^{u-1}}{(\sum_{k=1}^M y_k^u)^M} , \end{aligned}$$

et à l'intérieur du simplexe on a $J_u(y) = u^{M-1} \frac{z_1 \dots z_n}{y_1 \dots y_n}$.

ii) Proposition 18 - Les bijections $\bar{\varphi}_u$ définies à partir des φ_u ont une structure de groupe isomorphe au groupe multiplicatif des réels positifs.

En effet si l'on considère $\bar{\varphi}_u$ et $\bar{\varphi}_v$, la composée $\bar{\varphi}_u \circ \bar{\varphi}_v$ est une nouvelle bijection sur le simplexe identique à $\bar{\varphi}_{uv}$ construite à partir de φ_{uv} .

On a donc la loi de composition interne $\bar{\varphi}_u \circ \bar{\varphi}_v = \bar{\varphi}_{uv}$ dans l'ensemble des $\bar{\varphi}_u, u > 0$, avec une bijection inverse $\bar{\varphi}_u^{-1} = \bar{\varphi}_{1/u}$ et l'élément neutre $\bar{\varphi}_1$ qui est la bijection identité.

Les automorphismes $\bar{\varphi}_u$ du simplexe unité constituent donc un sous-groupe du groupe des automorphismes, doté de la loi de composition interne ci-dessus, isomorphe par construction au groupe multiplicatif des réels positifs. ■

Remarque : On observe simplement ce résultat sur l'expression des jacobiens en tout point intérieur des simplexes S_1^M , en effet, $y \in S_1^M$ on a :

$$z = \bar{\varphi}_u(y)$$

$$J_u(y) = u^{M-1} \frac{z_1 \dots z_n}{y_1 \dots y_n}$$

$$w = \bar{\varphi}_v(z)$$

$$J_v(z) = v^{M-1} \frac{w_1 \dots w_n}{z_1 \dots z_n}$$

d'où $J_u(y) J_v(z) = J_{uv}(y)$.

IV.9 - Application aux densités de probabilité.

9.1. Soit h la densité d'une distribution A.C. sur le simplexe unité S_1^n , et $\bar{\varphi}$ une bijection définie à partir d'une application φ positive continue dérivable croissante, nulle à l'origine. Avec ce changement de variable $z = \bar{\varphi}(y)$ l'image g de la densité h est une nouvelle densité sur simplexe S_1^n dont la valeur en tout point z est donnée par les relations suivantes :

$$h(y) = g(\bar{\varphi}(y)) \cdot \left| \frac{D(z)}{D(y)} \right|,$$

$$g(z) = h(\bar{\varphi}^{-1}(z)) \cdot \left| \frac{D(y)}{D(z)} \right|.$$

A tout domaine \mathcal{D}_y correspond un domaine $\mathcal{D}_z = \bar{\varphi}(\mathcal{D}_y)$:

$$\mathcal{D}_z = \{ z : z = \bar{\varphi}(y), y \in \mathcal{D}_y \}$$

et réciproquement à tout \mathcal{D}_z correspond le domaine $\bar{\varphi}^{-1}(\mathcal{D}_z) = \{ y \in S_1^n : \bar{\varphi}(y) \in \mathcal{D}_z \}$.

9.2. Ce changement de variables permet deux points de vue :

a) h et φ sont connus, on en déduit une nouvelle densité g sur S_1^n pour le point aléatoire $z = \bar{\varphi}(y)$; c'est la densité de la distribution image de h par $\bar{\varphi}$. La probabilité de tout événement \mathcal{D}_z est la probabilité de l'événement $\bar{\varphi}^{-1}(\mathcal{D}_z)$ pour la densité h . Dans ce cas, le calcul du jacobien est inutile, en revanche on doit noter que des ensembles \mathcal{D}_z du type ${}_z E_t^i$ conduisent sur le simplexe $S_1^n(y)$ à des $\bar{\varphi}^{-1}({}_z E_t^i)$ qui ne rentrent pas dans le cadre des événements examinés au chapitre II.

b) g est connu, on peut en déduire que toute fonction de y qui a une expression de la forme suivante : $h_p(y) = g[\bar{\varphi}(y)] \cdot \left| \frac{D\bar{\varphi}(y)}{D(y)} \right|$ est une densité de probabilité sur $S_1^n(y)$ dont l'image par $\bar{\varphi}$ est précisément g . La probabilité de tout événement ω_y pour la densité h_p est égale à celle de $\omega_z = \bar{\varphi}(\omega_y)$ pour la densité g .

Dans aucun des cas, on a simultanément la densité constante $\Gamma(M)$ et les domaines E'_t déduits de S_1^n par simple homothétie. Cela signifie que l'on ne peut appliquer les résultats simples du chapitre II, et qu'il faut faire éventuellement appel à des techniques numériques d'intégration comme on l'a signalé en IV.1.

On remarque aussi que le jacobien $\frac{D(z)}{D(y)}$ est une fonction en y invariante par le groupe Σ_M , on en déduit que h et g sont simultanément invariantes par Σ_M , de même que ω_y et $\bar{\varphi}(\omega_y)$.

9.3. Exemple .

Soit les automorphismes $\bar{\varphi}_u$ $u > 0$ précédents pour lesquels les jacobiens $\frac{D(z)}{D(y)}$ ont pour expression $u^{n-1} \frac{z_1 \dots z_n}{y_1 \dots y_n}$ en tout point intérieur.

A partir de toute densité g on en déduit les densités h_u dont les valeurs sont égales à

$$h_u(y) = g(\bar{\varphi}_u(y)) \frac{u^{n-1}}{(\sum y_i^u)^n} (y_1 \dots y_n)^{u-1},$$

et dont les images par $\bar{\varphi}_u$ sont g . Compte tenu de leur structure de groupe on établit une relation d'équivalence \mathcal{R} entre densités h_u comme suit :

$h_u \mathcal{R} h_v$ si et seulement si il existe une application $\bar{\varphi}_w$ telle que l'image de h_u par $\bar{\varphi}_w$ soit h_v ; chaque classe peut être représentée par son élément h_1 .

i) Si on prend pour h_{α} une densité de DIRICHLET invariante par Σ_M de valeur

$$h_{\alpha}(y) = \frac{\Gamma(n\alpha)}{\Gamma(\alpha)^n} (y_1 \cdots y_n)^{\alpha-1},$$

les éléments de la classe sont les densités $h_{u,\alpha}$ dont les valeurs sont égales à

$$h_{u,\alpha}(y) = u^{n-1} \frac{\Gamma(n\alpha)}{\Gamma(\alpha)^n} \frac{(y_1 \cdots y_n)^{\alpha u-1}}{(\sum y_k^u)^{n\alpha}}$$

ii) les éléments de la classe qui correspond à $\Gamma(M)$ sont les densités $h_{u,1}$ dont les valeurs sont égales à

$$h_{u,1}(y) = u^{n-1} \frac{\Gamma(M)}{\Gamma(1)^n} \frac{(y_1 \cdots y_n)^{u-1}}{(\sum y_k^u)^M}$$

C - Autres questions.

IV.10 - Densités associées.

10.1 - Soit G une densité continue sur le simplexe unité S_1^{n+1} . Ce simplexe est aussi une partie de \mathbb{R}_+^M comme suit :

$$\left\{ x \in \mathbb{R}_+^M \quad \sum_{i=1, i \neq \bar{k}}^{n+1} x_i \leq 1 \right\}$$

Il y a autant de façons de procéder ainsi qu'il y a d'indices $\bar{k} = 1, \dots, n+1$. Ci-après, pour des raisons de commodité qui ne restreignent pas la généralité, on retient le cas $\bar{k} = n+1$. Sur cette partie de \mathbb{R}_+^M , on pose :

$$\tilde{G}_{M,1}(x_1, \dots, x_n) = G(x_1, \dots, x_n, 1 - x_1 - x_2 - \dots - x_n),$$

et $\tilde{G}_{M,1}(x_1, \dots, x_n) = 0$ ailleurs.

Alors \tilde{G}_{n+1} est une densité sur S_1^n , on en déduit que l'application g qui à tout y de S_1^M fait correspondre

$$g_{n+1}(y) = \int_0^1 \tilde{G}_{n+1}(sy) s^{n-1} ds$$

est une densité sur S_1^M (cf IV.3). On l'appelle densité associée à G pour l'indice \bar{k} . On a donc $M+1$ densités sur S_1^M associées à toute densité sur S_1^{n+1} .

Pour chaque densité associée, on a l'interprétation géométrique suivante : pour tout $k=1, \dots, M+1$ soit le point de $e^{\bar{k}}$ extrémité du vecteur de base $\overrightarrow{0e^{\bar{k}}}$ de

l'espace affine ; et soit le sous-ensemble $\{z \in S_1^{n+1} : z_{\bar{k}} = 0\}$ que l'on peut identifier au simplexe unité S_1^M avec $y_p = z_p$ $p \neq \bar{k}$ (fig en IV.11)

La formule générale qui donne la valeur de la densité associée à G pour l'indice \bar{k} peut s'écrire :

$$g_{\bar{k}}(y) = \int_0^1 G[sy, (1-s)e^{\bar{k}}] s^{n-1} ds$$

définie en tout point $y \in S_1^{n+1} \cap \{z : z_{\bar{k}} = 0\}$.

Lorsque la densité G est invariante par Σ_{M+1} , on observe que toutes les densités associées $g_{\bar{k}}$ sont identiques, et elles sont aussi invariantes par Σ_M .

Exemple : Soit la densité de DIRICHLET qui pour tout $z \in S_1^{M+1}$ a pour valeur :

$$G(z) = \frac{\Gamma\left(\sum_1^{n+1} a_e\right)}{\prod_1^{n+1} \Gamma(a_e)} \prod_1^{n+1} z_e^{a_e - 1} ;$$

Proposition 19 Si G est sur S_1^{n+1} la densité déduite de la densité H définie dans \mathcal{R}_+^{*n+1} par la normalisation par la somme, alors la densité $g_{\bar{k}}$ associée à G pour tout indice $\bar{k} = 1, \dots, n+1$ est une densité sur S_1^n , obtenue par la normalisation par la somme de la densité marginale de H en $z_1, \dots, z_{\bar{k}-1}, z_{\bar{k}+1}, \dots, z_{n+1}$. (On note Z le point courant de \mathcal{R}_+^{*n+1} de coordonnées z_1, \dots, z_{n+1}).

Démonstration : Pour simplifier les notations on prend le cas $\bar{k} = n+1$ qui ne restreint pas la généralité. On considère le changement de variables $st = u$, $t = v$ dont le jacobien $D(u, v) / D(s, t)$ a pour valeur v , avec lequel la domaine $\{0 \leq t < +\infty, 0 \leq s \leq 1\}$ est transformé dans le domaine $\{0 \leq u < +\infty, u \leq v < +\infty\}$.

On obtient donc, avec des notations évidentes :

$$\begin{aligned} g_{n+1}(y_1, \dots, y_n) &= \int_0^\infty u^{n-1} du \int_u^\infty H(uy_1, \dots, uy_n, v-u) dv \\ &= \int_0^\infty u^{n-1} du \int_0^\infty H(uy_1, \dots, uy_n, v) dv \\ &= \int_0^\infty \bar{H}_n^{n+1}(uy_1, \dots, uy_n) u^{n-1} du \quad \blacksquare \end{aligned}$$

Conséquences : Si la densité H est un produit de densités f_j sur \mathcal{R}_+^* , $j=1, \dots, n+1$, la densité marginale en $z_1, \dots, z_{\bar{k}-1}, z_{\bar{k}+1}, \dots, z_{n+1}$ est le produit $\prod_{j \neq \bar{k}} f_j$, la densité associée à tout indice \bar{k} est celle qui résulte de la normalisation par la somme de ce produit $\prod_{j \neq \bar{k}} f_j$.

Pour tout $\bar{k} = 1, \dots, M$, elle admet pour densités associées les densités définies sur S_1^n qui ont pour valeur :

$$g_{\bar{k}}(y) = \frac{\Gamma(\sum_{e \neq \bar{k}} a_e)}{\prod_{e \neq \bar{k}} \Gamma(a_e)} \prod_{j \neq \bar{k}} y_j^{a_j - 1},$$

avec y le point de coordonnées y_i , $i = 1, \dots, n+1$, $i \neq \bar{k}$, $y \in S_1^n$.

Conséquence : En particulier la densité constante sur S_1^{n+1} admet la densité constante sur S_1^n pour toute densité associée.

10.2. Naturellement le procédé d'association de densités peut être itéré plusieurs fois ; et on peut le combiner avec la "normalisation par la somme" du chapitre IV.3.

Soit G une densité sur S_1^{n+1} obtenue par la normalisation par la somme d'une densité définie dans \mathbb{R}_+^{n+1} . Comme indiqué ci-dessus, on en déduit les densités $\tilde{G}_{\bar{k}}$ et $g_{\bar{k}}$ relatives à l'indice \bar{k} , égal ici à $n+1$.

On a les relations suivantes ; (notations de IV.3 et IV.10) :

$$G(z_1, \dots, z_{n+1}) = \int_0^\infty H(tz_1, \dots, tz_{n+1}) t^M dt,$$

$$\tilde{G}_{n+1}(x_1, \dots, x_n) = \int_0^\infty H(tx_1, \dots, tx_n, t - t(x_1 + \dots + x_n)) t^n dt,$$

$$g_{n+1}(y_1, \dots, y_n) = \int_0^1 s^{n-1} ds \int_0^\infty H(stx_1, \dots, stx_n, t-st) t^n dt.$$

IV.11 - Inégalités en fonction de l'effectif M .

Soit G une densité définie sur le simplexe unité S_1^{n+1} et $g_{\bar{k}}$ la densité associée pour l'indice \bar{k} , définie sur le simplexe unité S_1^n .

On pose les probabilités suivantes :

$$\theta_{n+1}^\lambda(t) = \int G(z) dz_1 \dots dz_n$$

où l'intégrale est étendue à l'ensemble :

$$\bar{E}_{n+1,t} = \left\{ z \in S_1^{n+1}, \lambda_1(z) \leq t \right\},$$

et

$$\theta_{n,\bar{k}}^\lambda(u) = \int g_{\bar{k}}(y) dy_1 \dots dy_{n-1}$$

où l'intégrale est étendue à l'ensemble :

$$\bar{E}_{n,u} = \left\{ y \in S_1^n, \lambda_1(y) \leq u \right\}.$$

Cette dernière intégrale s'écrit encore :

$$\begin{aligned} \theta_{n,\bar{k}}^\lambda(u) &= \int_0^1 ds \int G(sy, (1-s)e^{\bar{k}}) dy_1 \dots dy_{n-1} \\ &= \int G(z^{\bar{k}}) dz^{\bar{k}} \end{aligned}$$

où l'intégrale est étendue à l'ensemble

$$\left\{ z^{\bar{k}} \in S_1^{n+1}, z^{\bar{k}} = sy + (1-s)e^{\bar{k}}, y \in S_1^n \cap \{z: z_{\bar{k}} = 0\}, \lambda_1(y) \leq u, s \in [0,1] \right\}.$$

Soit la partition suivante de ce dernier ensemble :

$$\{z^{\bar{k}}\} = \{z_1^{\bar{k}}\} \cup \{z_2^{\bar{k}}\} \quad \text{avec } s \geq u \text{ pour } \{z_1^{\bar{k}}\} \text{ et } s < u \text{ pour } \{z_2^{\bar{k}}\}.$$

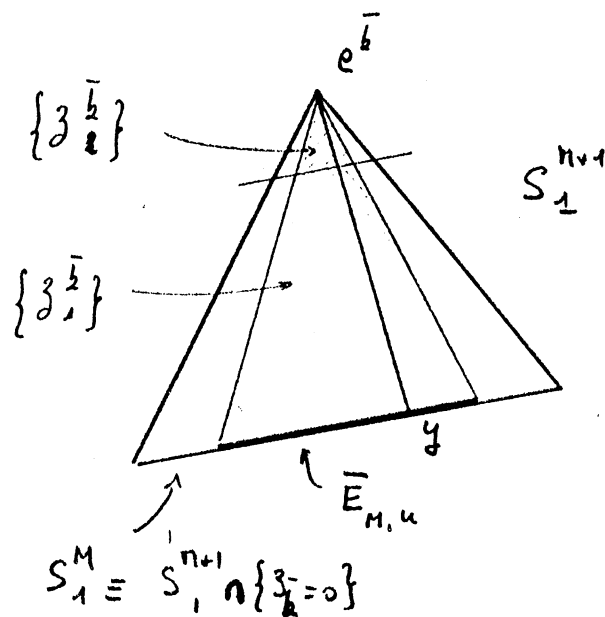
On a évidemment $\{z_1^{\bar{k}}\} \subset \bar{E}_{n+1,u} = \{z \in S_1^{n+1} : \lambda_1(z) \leq u\}$, d'où la

Proposition 20 - Pour tout u appartenant à $[0, 1]$, on a l'inégalité valable pour toute densité G sur S_1^{n+1} :

$$\theta_{n, \bar{k}}^+(u) - \int_{\{z_{\bar{k}}^{\bar{k}}\}} G(z) dz \leq \theta_{n+1}^+(u) ;$$

pour tout u appartenant à $[0, 1/n]$, on a l'inégalité plus forte évidente $\theta_{n, \bar{k}}^+(u) \leq \theta_{n+1}^+(u)$.

Démonstration : Le premier résultat résulte de l'inclusion $\{z_{\bar{k}}^{\bar{k}}\} \subset \bar{E}_{n+1, u}$; le deuxième résultat est immédiat puisque l'intégrale $\theta_{n, \bar{k}}^+(u)$ est nécessairement nulle pour $0 \leq u \leq 1/n$.



Proposition 21 - Avec la densité G qui correspond à la mesure uniforme sur S_1^n , on a les inégalités plus précises :

$$\text{pour tout } u \in [0, 1] : [1 - (1-u)^n] \cdot L_n^\lambda(u) \leq L_{n+1}^\lambda(u),$$

$$\text{pour tout } u \in [0, \frac{1}{n}] : L_n^\lambda(u) \leq L_{n+1}^\lambda(u),$$

où L_n^λ est la fonction de répartition de λ_1 lorsque y suit la distribution uniforme sur S_1^n (cf chap. II).

En effet on a immédiatement :

$$\int_{\{z^{\bar{k}}\}} G(z) dz = (1-t)^M \int_{\{z^{\bar{k}}\}} G(z) dz$$

puisque $\{z^{\bar{k}}\}$ et $\{z^{\bar{k}}\}$ sont homothétiques dans S_1^{n+1} , dans une homothétie de centre $e^{\bar{k}}$, de rapport $(1-t)$ ■

IV.12 - Densités de DIRICHLET inverses et/ou décentrées.

On rappelle que si x_1, \dots, x_M sont des variables aléatoires indépendantes qui suivent des distributions gamma $\Gamma(a_i, \lambda)$, le vecteur de coordonnées

$x_i / \sum_1^M x_i$, $i=1, \dots, M$ suit une distribution de DIRICHLET, sur le simplexe unité S_1^M (chap. IV.4).

12.1. i) Soit x_0 une autre variable qui suit la distribution gamma $\Gamma(a_0, \lambda)$ indépendante des précédentes, et les nouvelles variables :

$$z_0 = x_0, \quad z_j = \frac{x_j}{x_0} \quad j=1, \dots, M \quad \text{avec le Jacobien} \quad \frac{D(x_0, \dots, x_M)}{D(z_0, \dots, z_M)}$$

qui est égal à z_0^M .

La distribution image du produit des distributions des x_j $j=0, \dots, M$ par ce changement de variables a une densité K qui a pour valeur :

$$K(z_0, z_1, \dots, z_M) = \frac{1}{\prod_0^M \Gamma(a_j)} \prod_0^M (z_0 z_j)^{a_j-1} e^{-z_0 z_j}, \quad z_0^{a_0-1} e^{-z_0} z_1^{a_1-1} e^{-z_0 z_1} \dots z_M^{a_M-1} e^{-z_0 z_M}$$

et la densité de la distribution marginale en z_1, \dots, z_M a pour valeur :

$$H(z_1, \dots, z_M) = \int_0^\infty K(z_0, \dots, z_M) dz_0 \\ = \frac{\Gamma(a_0 + a_1 + \dots + a_M)}{\Gamma(a_0) \Gamma(a_1) \dots \Gamma(a_M)} \frac{z_1^{a_1-1} \dots z_M^{a_M-1}}{(1 + z_1 + \dots + z_M)^{a_0 + a_1 + \dots + a_M}}$$

c'est la densité de "DIRICHLET inverse" introduite par TIAO et GUTTMAN (réf.43,78), définie sur \mathcal{R}_+^M .

Cette distribution apparaît également en substance dans OLKIN et SOBEL, et KHATRI et MITRA (réf.47,62).

ii) On peut donc appliquer à cette densité la procédure de normalisation par la somme (cf IV.3) pour obtenir une densité sur le simplexe unité de \mathcal{R}^M .

On a

$$\begin{aligned}
 h(y_1, \dots, y_n) &= \int_0^\infty H(sy_1, \dots, sy_n) s^{n-1} ds \\
 &= \frac{\Gamma(a_0 + \dots + a_n)}{\Gamma(a_0) \dots \Gamma(a_n)} y_1^{a_0-1} \dots y_n^{a_n-1} \int_0^\infty \frac{s^{a_0 + \dots + a_n - 1}}{(1+s)^{a_0 + \dots + a_n}} ds \\
 &= \frac{\Gamma(a_0 + \dots + a_n)}{\Gamma(a_0) \dots \Gamma(a_n)} y_1^{a_0-1} \dots y_n^{a_n-1} .
 \end{aligned}$$

On retrouve dans ce cas la densité de DIRICHLET $D(\cdot \dots \cdot; a_1, \dots, a_n)$ également notée $D(\cdot; a)$, avec a le vecteur de coordonnées $a_i, i=1, \dots, n$.

12.2. i) Les $x_j, j=1, \dots, M$ suivent à présent des distributions gamma décentrées de paramètres γ_j (réf. 5), qui ont pour densités les expressions suivantes :

$$f_{0j; \gamma_j}(t) = e^{-\gamma_j} \sum_{m \geq 0} \frac{\gamma_j^m}{m!} \frac{t^{a_j + m - 1} e^{-t}}{\Gamma(a_j + m)} \mathbb{1}_{\mathcal{R}_+}(t) .$$

On définit z_0 et $z_j, j=1, \dots, M$ comme ci-dessus, les z_j suivant à présent des distributions bêta de 2^e espèce décentrées. Avec les mêmes notations la densité de la distribution image du produit des distributions des x_i a pour valeur :

$$\begin{aligned}
 h(z_0, \dots, z_n) &= \frac{z_0^{a_0-1} e^{-z_0} z_0^M}{\Gamma(a_0)} \sum_{\substack{m_1 \geq 0 \\ \vdots \\ m_n \geq 0}} \prod_{j=1}^M \frac{e^{-\gamma_j} \gamma_j^{m_j}}{m_j!} \frac{(z_0 z_j)^{a_j + m_j - 1} e^{-z_0 z_j}}{\Gamma(a_j + m_j)} , \\
 &= \frac{z_0^{a_0-1} e^{-z_0(1+z_1+\dots+z_n)}}{\Gamma(a_0)} e^{-\sum_{j=1}^M \gamma_j} \sum_{\varphi \geq 0} z_0^{a_0 + \dots + a_n + \varphi} . d\varphi ,
 \end{aligned}$$

en posant

90

$$d\mu_{\varphi} = \sum_{\sum m_j = \varphi} \prod_{j=1}^M \frac{\gamma_j^{m_j}}{m_j!} \frac{z_j^{a_j + m_j - 1}}{\Gamma(a_j + m_j)} \quad ;$$

et la densité de la distribution marginale en z_1, \dots, z_n a pour valeur :

$$H(z_1, \dots, z_n) = \frac{e^{-\sum \gamma_j}}{\Gamma(a_0)} \sum_{\varphi \geq 0} \frac{\Gamma(a_0 + a_1 + \dots + a_n)}{(1 + z_1 + \dots + z_n)^{a_0 + a_1 + \dots + a_n + \varphi}} d\mu_{\varphi} .$$

C'est la densité d'une nouvelle distribution sur \mathbb{R}_+^n que l'on peut appeler distribution de "DIRICHLET inverse décentrée".

ii) On peut lui appliquer également la procédure de normalisation par la somme, d'où les valeurs des densités :

$$h(y_1, \dots, y_n) = \frac{e^{-\sum \gamma_j}}{\Gamma(a_0)} \sum_{\varphi \geq 0} \int \Gamma(a_0 + \dots + a_n + \varphi) \sum_{j=1}^* \prod_{j=1}^M \frac{\gamma_j^{m_j}}{m_j!} \frac{y_j^{a_j + m_j - 1}}{\Gamma(a_j + m_j)}$$

en posant

$$J = \int_0^{\infty} \frac{s^{a_1 + \dots + a_n + \varphi - 1}}{(1+s)^{a_0 + \dots + a_n + \varphi}} ds ,$$

d'où

$$h(y_1, \dots, y_n) = e^{-\sum \gamma_j} \sum_{\varphi \geq 0} \Gamma(a_1 + \dots + a_n + \varphi) \left[\sum_{j=1}^* \prod_{j=1}^M \frac{\gamma_j^{m_j}}{m_j!} \frac{y_j^{a_j + m_j - 1}}{\Gamma(a_j + m_j)} \right] ,$$

$$= e^{-\sum \gamma_j} \sum_{\varphi \geq 0} \frac{1}{\varphi!} \sum_{m_1! \dots m_n!}^* \frac{\varphi!}{m_1! \dots m_n!} \gamma_1^{m_1} \dots \gamma_n^{m_n} D(y_j, \dots, a_j + m_j, \dots) ,$$

avec $\sum_{j=1}^*$ la sommation étendue à l'ensemble des $m_j \geq 0$ tels que

$$\sum_{j=1}^n m_j = \varphi$$

. C'est la densité d'une nouvelle distribution sur le simplexe unité que l'on peut appeler distribution de DIRICHLET "décentrée". C'est une distribution mélange d'une infinité de distributions de DIRICHLET, pondérées par le produit de M distributions de POISSON .

On note $DD(\dots; a_1, \dots, a_n; \gamma_1, \dots, \gamma_n)$ ou $DD(\cdot; a; \gamma)$ cette densité, et

$DD(a; \gamma)$ la distribution correspondante sur le simplexe, avec γ le vecteur de coordonnées γ_i $i=1, \dots, M$. Elles sont invariantes par \sum_M si et seulement si tous les paramètres γ_i d'une part et tous les exposants a_i de l'autre sont respectivement égaux entre eux.

L'introduction ci-dessus de cette distribution a été faite en s'inspirant de la littérature, (distributions inverses); on peut l'obtenir directement:

Théorème 18 Pour $i=1, \dots, M$, soient α_i des variables positives indépendantes qui suivent respectivement les distributions $\Gamma(a_i; \lambda; \delta)$. Dans \mathbb{R}_+^{M+1} la normalisation par la somme du produit de ces distributions conduit à la distribution de DIRICHLET décentrée $DD(a; \gamma)$.

Démonstration: Avec les notations de IV.4 et ci-dessus on a les valeurs de densités suivantes:

$$H(x_1, \dots, x_n) = e^{-\sum \gamma_i - \lambda \sum x_i} \sum_{\varphi \geq 0} \sum^* \frac{\prod (\lambda x_j)^{a_j + m_j - 1} \lambda \gamma_j^{m_j}}{m_j! \Gamma(a_j + m_j)},$$

d'où

$$\begin{aligned} h(y_1, \dots, y_n) &= e^{-\sum \gamma_i} \sum_{\varphi \geq 0} \sum^* \prod \frac{y_j^{a_j + m_j - 1} \gamma_j^{m_j}}{m_j! \Gamma(a_j + m_j)} \int_0^\infty (\lambda \delta)^{\sum a_j + \varphi - 1} \lambda e^{-\lambda \delta} d\delta \\ &= e^{-\sum \gamma_i} \sum_{\varphi \geq 0} \frac{1}{\varphi!} \sum^* \frac{\varphi!}{m_1! \dots m_n!} \gamma_1^{m_1} \dots \gamma_n^{m_n} \times DD(y; a; \gamma). \end{aligned}$$

12.3. Cas de $M = 2$.

Dans ce cas on a une distribution définie sur le simplexe S_1^2 , qui est l'équivalent d'une distribution Bêta de 1^o espèce décentrée et qui dépend de 4 paramètres $a_1, a_2, \gamma_1, \gamma_2$ et dont la densité a pour valeur :

$$h(y_1, y_2) = e^{-(\gamma_1 + \gamma_2)} \sum_{\varphi \geq 0} \frac{1}{\varphi!} \sum^* \frac{\varphi!}{m_1! m_2!} \gamma_1^{m_1} \gamma_2^{m_2} B(y; a_1 + m_1, a_2 + m_2)$$

avec $y_1 + y_2 = 1$,

avec $B(\cdot; \alpha, \beta)$ la densité bêta de première espèce définie sur S_1^2

et \sum^* la sommation étendue aux couples d'entiers $\geq 0, m_1, m_2$ de somme φ .

Remarque : On introduit le plus souvent la distribution Bêta décentrée de 2° espèce définie sur \mathcal{R}_+ ; c'est une distribution mélange de distributions Bêta de 2° espèce, avec pour pondérations les probabilités d'une distribution de POISSON (réf. 5).

On rencontre aussi dans la littérature des mélanges de distributions Bêta de 1° espèce définies sur $[0, 1]$; les pondérations sont les probabilités d'une distribution de POISSON (réf 39,51,60), ou d'une distribution de PASCAL (réf.11).

Les distributions de DIRICHLET décentrées introduites ici ont un système de pondération qui correspond à un produit d'autant de distributions de POISSON qu'il y a d'exposants α_i ; pour $M = 2$ la distribution Bêta décentrée est pondérée par un produit de deux distributions de POISSON, elle est plus générale que les mélanges rappelés ci-dessus.

IV.13 - Autres normalisations par la somme.

On poursuit cette procédure sur des produits de distributions $\prod_{i=1}^M F_i$ où les F_i sont fermées pour la convolution sur \mathcal{R}_+ (cf. IV.4).

13.1 La distribution de DIRICHLET décentrée.

Les distributions $\Gamma(a; \gamma; \delta)$ vérifient la relation de convolution (ref. 5)

$$\Gamma(a_1; \gamma_1; \delta) * \Gamma(a_2; \gamma_2; \delta) = \Gamma(a_1+a_2; \gamma_1+\gamma_2; \delta) ,$$

elles conduisent on l'a vu, à la distribution $DD(a; \delta)$; ses propriétés sont étudiées en X.2 .

13.2 La distribution de DIRICHLET - BESSEL - FELLER .

On prend la densité suivante qui pour tout $\rho > 0$ a pour valeur :

$$f_\rho(x) = e^{-x} \frac{\rho}{x} I_\rho(x) * \mathbb{1}_{\mathcal{R}_+}(x)$$

dans laquelle intervient la fonction de BESSEL d'expression

$$I_\rho(x) = \left(\frac{x}{2}\right)^\rho \sum_{k=0}^{\infty} \frac{(x/2)^{2k}}{k! \Gamma(k+\rho+1)} .$$

On sait que c'est la densité d'une distribution indéfiniment divisible, qui vérifie

$$f_\rho * f_\eta = f_{\rho+\eta} \quad (\text{réf.24 Tome 2}).$$

On en déduit sur S_1^M la densité suivante qui a pour valeur :

$$h(y_1, \dots, y_n) = \bar{h}^M(y_1, \dots, y_{n-1}) = \frac{\rho_1 \dots \rho_n}{y_1 \dots y_n} \int_0^\infty \prod_{i=1}^n I_{\rho_i}(\rho_i y_i) \frac{e^{-s}}{s} ds$$

dont on connaît les valeurs des densités marginales :

$$\bar{h}_m(y_1, \dots, y_k) = \frac{\rho_1 \dots \rho_k \sum_{i=1}^M \rho_i}{y_1 \dots y_k (1 - \sum_{i=1}^k y_i)} \int_0^\infty \prod_{i=1}^k I_{\rho_i}(\rho_i y_i) \frac{I_{\sum \rho_i}(s - \sum_{i=1}^k y_i) e^{-s}}{s} ds$$

$$\bar{h}_m(y_j) = \frac{\rho_j \sum_{i=1}^M \rho_i}{y_j (1 - y_j)} \int_0^\infty I_{\rho_j}(\rho_j y_j) \frac{I_{\sum \rho_i}(s - y_j) e^{-s}}{s} ds$$

Quand on développe les fonctions de BESSEL I_ρ on obtient :

$$h(y_1, \dots, y_n) = \rho_1 \dots \rho_n \sum_{k_i \geq 0} \prod_{i=1}^n \left(\frac{y_i}{2}\right)^{\rho_i + 2k_i - 1} \int_0^\infty \frac{s^{\sum(\rho_i + 2k_i) - 1} e^{-s}}{\prod k_i! \Gamma(k_i + \rho_i + 1)} ds$$

avec $\sum_{i=1}^n y_i = +1$,

$$h(y_1, \dots, y_n) = \frac{\rho_1 \dots \rho_n}{2^{\sum \rho_i}} \sum_{\varphi \geq 0} \frac{\Gamma(2\varphi + \sum \rho_i)}{2^{2\varphi}} \sum_{i=1}^n \frac{y_i^{\rho_i + 2k_i - 1}}{k_i! \Gamma(k_i + \rho_i + 1)}$$

ou encore :

$$h(y_1, \dots, y_n) = \frac{\rho_1 \dots \rho_n}{2^{\sum \rho_i}} \sum_{\varphi \geq 0} \frac{1}{2^{2\varphi}} \sum_{i=1}^n D(y; \dots, \rho_i + 2k_i \dots) \prod_{j=1}^n \frac{\Gamma(\rho_j + 2k_j)}{k_j! \Gamma(\rho_j + 1)}$$

C'est un autre exemple de mélange de distribution de DIRICHLET sur le simplexe, que l'on peut appeler la distribution de DIRICHLET - BESSEL - FELLER.

(La sommation \sum^* est étendue aux M-uples entiers k_i de somme Q).

Le système de pondération du mélange est le produit des M distributions définies sur \mathcal{N} avec pour tout n :

$$t_i(n) = \frac{\rho_i}{2^{\rho_i}} \frac{\Gamma(\rho_i + 2n)}{2^{2n} n! \Gamma(\rho_i + n + 1)}$$

Proposition 22 - On a les identités suivantes :

$$i) \quad \sum_{n \geq 0} \frac{\Gamma(\rho + 2n)}{2^{2n} n! \Gamma(\rho + n + 1)} = \frac{2^\rho}{\rho}$$

$$ii) \quad \sum_{Q \geq 0} \frac{1}{2^{2Q}} \sum_{j=1}^* \prod_{j=1}^M \frac{\Gamma(\rho_j + 2k_j)}{k_j! \Gamma(\rho_j + k_j + 1)} = \frac{2^{\sum \rho_i}}{\prod \rho_i}$$

En effet ces identités résultent de la sommation des probabilités des distributions discrètes introduites ci-dessus. ■

Remarque i) La distribution de DIRICHLET - BESSEL - FELLER est invariante par \sum_M si et seulement si tous les ρ_i sont égaux à une même valeur.

ii) Pour $M = 2$, la densité de la distribution de DIRICHLET - BESSEL - FELLER a la valeur suivante pour tout $y \in [0, 1]$:

$$DBF(y) = \frac{\rho_1 \rho_2}{2^{\rho_1 + \rho_2}} \sum_{Q \geq 0} \frac{\Gamma(\rho_1 + \rho_2 + 2Q)}{2^{2Q}} \sum_{\substack{k_1 + k_2 \\ = Q}} \frac{y^{\rho_1 + 2k_1 - 1} (1-y)^{\rho_2 + 2k_2 - 1}}{k_1! k_2! \Gamma(\rho_1 + k_1 + 1) \Gamma(\rho_2 + k_2 + 1)}$$

13.3. La distribution de GAUSS-LAPLACE inverse.

On désigne ainsi la distribution dont la densité a l'expression suivante, $\eta > 0$

$$f_{\eta}(x) = \frac{\eta}{\sqrt{2\eta}} \frac{e^{-\eta/2x}}{x^{3/2}} \mathbb{1}_{\mathbb{R}_+}(x),$$

on sait que c'est une Loi stable sur \mathbb{R}_+^* avec l'exposant caractéristique égal à $1/2$; on a donc (réf. 4,24,58)

$$f_{\eta} * f_{\xi} = f_{\eta+\xi}.$$

On en déduit sur S_1^n la densité dont les valeurs sont :

$$h(y_1, \dots, y_n) = \bar{h}^n(y_1, \dots, y_n) = \frac{\eta \dots \eta_n}{(\eta)^{n/2}} \frac{1}{(y_1 \dots y_n)^{3/2}} \int_0^{\infty} \frac{e^{-\frac{1}{2s} \sum y_i^2} s^{n-1}}{s^{3n/2}} ds$$

avec $\sum y_i = 1$.

On effectue le changement de variable $u = 1/s$, d'où l'expression :

$$h(y_1, \dots, y_n) = \Gamma\left(\frac{n}{2}\right) \frac{\eta \dots \eta_n}{\eta^{n/2}} \frac{1}{(y_1 \dots y_n)^{3/2}} \left(\sum \frac{\eta_i^2}{y_i}\right)^{-n/2},$$

c'est la densité d'une distribution qui a des moments d'ordre entier quelconque, contrairement à la distribution de GAUSS-LAPLACE inverse qui a servi pour la construire. Cette distribution est invariante par $\sum_{\mathcal{M}}$ si et seulement si tous les η_i sont égaux à une même valeur.

Les densités marginales ont les valeurs suivantes:

$$\bar{h}_m(y_1, \dots, y_k) = \Gamma\left(\frac{k+1}{2}\right) \frac{\eta \dots \eta_k \left(\sum_{k+1}^n \eta_i\right)}{\eta^{(k+1)/2}} \frac{1}{[y_1 \dots y_k (1 - \sum_{i=1}^k y_i)]^{3/2}} \left[\sum_{i=1}^k \frac{\eta_i^2}{y_i} + \left(\sum_{i=1}^k \frac{\eta_i^2}{1 - \sum_{j=1}^k y_j}\right)^2 \right]^{-\frac{k+1}{2}}$$

$$\bar{h}_m(y_1) = \frac{\eta \left(\sum_{i=2}^n \eta_i\right)}{\eta} \frac{1}{(y_1 (1 - y_1))^{3/2}} \left[\frac{\eta_1^2}{y_1} + \frac{\sum_{i=2}^n \eta_i^2}{1 - y_1} \right]^{-1}.$$

IV.14 - Autres distributions sur S_1^M .

Nous indiquons quelques autres exemples de densités classiques, soit des densités dans \mathbb{R}_+^M auxquelles on applique la normalisation par la somme, soit directement des densités sur le simplexe unité ; toutes ces densités sont issues de JOHNSON et KOTZ tome 4 (Réf. 43).

14.1. La distribution F multiple .

On a dans \mathbb{R}_+^M

$$H(F_1, \dots, F_M) = \frac{\Gamma(a_0 + \dots + a_M)}{\Gamma(a_0) \dots \Gamma(a_M)} \prod_{i=0}^M (\lambda a_i)^{a_i} \frac{F_1^{a_1-1} \dots F_M^{a_M-1}}{(\lambda a_0 + \lambda \sum_{j=1}^M a_j F_j)^{a_0 + \dots + a_M}}$$

d'où, en appliquant la normalisation par la somme:

$$\begin{aligned} h(j_1, \dots, j_M) &= \frac{\Gamma(a_0 + \dots + a_M)}{\Gamma(a_0) \dots \Gamma(a_M)} \int_0^{a_1-1} \dots \int_0^{a_M-1} \prod_{i=0}^M (\lambda a_i)^{a_i} \int_0^\infty \frac{s^{a_0 + \dots + a_M - 1}}{(\lambda a_0 + \lambda \sum_{k=1}^M a_k j_k)^{a_0 + \dots + a_M}} ds \\ &= \frac{\Gamma(a_0 + \dots + a_M)}{\Gamma(a_0) \dots \Gamma(a_M)} \frac{1}{j_1 \dots j_M} \prod_{j=1}^M \left(\frac{a_j j_j}{\sum_{k=1}^M a_k j_k} \right)^{a_j}, \end{aligned}$$

la valeur de la densité en tout point de coordonnées $j_i: i=1, \dots, M$ de S_1^M .

La densité h est invariante par \sum_M si et seulement si tous les paramètres $a_i: i=1, \dots, M$ sont égaux à une même valeur, auquel cas on retrouve la densité de la distribution invariante de DIRICHLET.

14.2. Autres exemples .

La normalisation par la somme d'une densité dans \mathbb{R}_+^M peut s'appliquer également aux distributions de PARETO, BURR, logistiques et log-normales. Les expressions respectives des H sont en référence 43, les h correspondants résultent des formules générales de IV.3.

14.3. Distributions sur le simplexe .

Parmi les distributions directement définies sur simplexe unité, on peut citer :

a) La distribution de OBENCHAIN du point aléatoire de coordonnées $i=1, \dots, M$

$$e^{y_i} / \sum_1^M e^{y_k}$$

où Y_i $i=1, \dots, M$ est un vecteur aléatoire qui suit une loi normale (réf.43) ;

b) La distribution de DIRICHLET généralisée de CONNOR et MOSIMANN dont la densité a pour valeur (réf.16,43) :

$$R(y_1, \dots, y_n) = \frac{1}{\prod_1^{n-1} B(a_k, b_k)} \prod_1^n y_k^{a_k-1} \left[\prod_{k=1}^{n-1} \left(1 - \sum_{p=1}^{k-1} y_p \right)^{b_k-1} \right]^{-1}$$

Note : une erreur s'est glissée dans cette expression en référence 43.

c) La distribution des valeurs propres d'une matrice aléatoire de WISHART offre un exemple de distribution conjointe que l'on peut ramener sur le simplexe unité, et pour laquelle on dispose de renseignements sur les plus grandes coordonnées :

c1 - Si A et B sont deux matrices de WISHART dans \mathbb{R}^P avec des paramètres m_A et $m_B \geq f$, on connaît la distribution conjointe des zéros du déterminant de $A - \lambda B$ (réf. 2,43). On a également établi les distributions des zéros extrémaux λ_{np} et λ_{ij} (réf.43).

c2 - En outre on sait que la somme des valeurs propres d'une matrice de WISHART (la trace de $A - \lambda I$) suit la distribution scalaire du χ^2 avec m_A, P degrés de liberté.

c3 - Les résultats classiques sont présentés en termes de $\sup \lambda_i$ et $\inf \lambda_i$, mais connaissant la loi de $\sum \lambda_j$ on peut se ramener sur le simplexe unité. On rappelle simplement ici l'intérêt que présentent les taux $\sup \lambda_i / \sum \lambda_j$ pour l'interprétation des résultats d'une analyse factorielle des correspondances, (réf. 50).

d) KENDALL et STUART ont étudié la même variable $\sum x_j / \sup x_i$ que DARLING (cf IV.5.3), dans le cas où les x_j $j=1, \dots, M$ sont des variables aléatoires indépendantes qui suivent la loi uniforme sur $[0, \theta]$. Les calculs sont analogues, le résultat est ici que $\sum x_j / \sup x_i$ est distribué comme la somme de $M-1$ variables indépendantes uniformes sur $[0, 1]$, augmentée de 1. (réf.46).

DEUXIEME PARTIE :Distributions discontinuesIntroduction.

Le partage d'une quantité peut également être envisagé en considérant que le point y du simplexe unité S_1^M associé au partage varie sur un sous-ensemble dénombrable ou fini du simplexe. La coordonnée maximale d'un point d'un tel sous-ensemble est une variable qui ne prend ses valeurs que sur un sous-ensemble dénombrable ou fini de $[0, 1]$.

Dans ce nouveau cas, bien entendu, les techniques de calcul que l'on doit employer sont différentes de celles que l'on utilise avec les distributions continues. Ci-dessous nous introduisons une famille de sous-ensembles finis de S_1^M qui dépendent de l'entier N supérieur ou égal à 1, sous-ensembles invariants par le groupe Σ_M des permutations de M objets.

La définition de ces sous-ensembles s'accompagne d'une discussion arithmétique préalable en M et N (chap. V.1).

Nous envisageons les distributions de probabilités discrètes définies sur ces sous-ensembles ; lorsqu'elles sont elles-mêmes invariantes par le groupe

Σ_M des permutations, il est assez facile verrons-nous d'établir la distribution que suit la plus grande coordonnée. La plus simple d'entre elles est celle où tous les points du simplexe (en nombre fini) sont équiprobables, il est alors possible de tabuler à la main les probabilités de la distribution de la plus grande coordonnée pour des petites valeurs de M et N (de 1 à 10 environ) (Annexe 2). Dans le cas général nous établissons une formule elle aussi générale (chap V.2). A cette occasion nous introduirons quelques distributions de probabilités pour lesquelles les expressions analytiques des probabilités prennent une tournure assez simple. Là encore les distributions fermées pour la convolution jouent un rôle particulier (chap V.3)(il s'agit ici de distributions sur \mathcal{N}). Comme pour les distributions absolument continues, nous examinons ensuite divers autres procédés pour obtenir des lois de probabilités discrètes sur le simplexe unité, sans plus nous soucier de la plus grande coordonnée, (chap V.4). La distribution multinomiale intervient dans un procédé assez général, pour lequel nous envisageons des propriétés asymptotiques quand on fait tendre le paramètre N vers plus l'infini

V - DISTRIBUTIONS DISCONTINUES SUR LE SIMPLEXE UNITE

V.1 - Discussion arithmétique.

1.1. Sous-ensemble discrets du simplexe.

Pour tout entier N positif ou nul, et tout entier M supérieur ou égal à 1, on définit dans \mathbb{R}^M le simplexe :

$$S_N^M = \left\{ y \in \mathbb{R}^M ; y_i \geq 0 \quad i=1, \dots, M, \sum_{i=1}^M y_i = N \right\},$$

et le sous-ensemble \bar{S}_N^M des points n dont toutes les coordonnées n_i sont des nombres entiers compris entre 0 et N :

$$\bar{S}_N^M = \left\{ n \in \mathbb{N}^M ; \sum_{i=1}^M n_i = N \right\}.$$

On vérifie immédiatement que \bar{S}_N^M est invariant par le groupe des permutations Σ_M . Ci-après pour simplifier nous considérons indifféremment \bar{S}_N^M ou le sous-ensemble de S_1^M des points de coordonnées rationnelles n_i/N $i=1, \dots, M$, $\sum_{i=1}^M n_i = N$. On peut encore remarquer que \bar{S}_N^M peut être identifié à l'ensemble isomorphe des M -uples (n_1, \dots, n_M) de M nombres entiers positifs ou nuls, rangés par l'indice $i=1, \dots, M$ et de somme N . Nous dirons que M est la dimension, N la somme et n_i les coordonnées du M -uplet ; l'ensemble des points de coordonnées n_i/N , $i=1, \dots, M$ est encore appelé le sous-ensemble N -discret de S_1^M .

On définit l'application λ_1 qui à tout point n de coordonnées n_i de \bar{S}_N^M fait correspondre $\lambda_1(n) = \sup n_i$. La discussion arithmétique qui suit est commode pour l'étude de l'application λ_1 ; elle est le pendant de l'étude géométrique dans le cas des distributions continues (chapitre I).

1.2. Une partition de $\{1, 2, \dots, N\}$.

1.2.1. Pour tout N entier positif ou nul, on définit les suites numériques b_i et b_a telles que pour tout $q \in \mathcal{N}^*$ on a :
($[x]$ désigne la partie entière de x) :

$$b_i(q) = \left[\frac{N}{q+1} + 1 \right] ,$$

$$b_a(q) = \left[\frac{N}{q} \right] .$$

Proposition 1 - Les deux suites b_i et b_a sont décroissantes, et vérifient les conditions suivantes :

- Pour tout $q \geq N+1$ $b_a(q) = 0$
- pour tout $q \geq N$ $b_i(q) = 1$
- pour tout $q \geq 1$ et pour tout $k \geq 1$,
 $b_i(q+k) \leq b_i(q) - 1$ avec égalité pour $k=1$.

Les démonstrations résultent immédiatement des expressions de $b_i(q)$ et $b_a(q)$. ■

1.2.2.

Proposition 2 - Pour tout $N \geq 1$ et pour tout q appartenant à $\{1, \dots, N\}$ et vérifiant $b_i(q) \leq b_a(q)$ les ensembles

$$I_q = \left\{ k \in \mathcal{N} : b_i(q) \leq k \leq b_a(q) \right\}$$

constituent une partition de $\{1, \dots, N\}$.

Démonstration :

- i) De tels ensembles I_q existent, I_1 et I_N par exemple ;
- ii) Soient I_{q_1} et I_{q_2} deux tels ensembles ; ils sont nécessairement disjoints, puisque l'on a :

$$b_a(\sup(q_1, q_2)) \leq b_i(\inf(q_1, q_2)) - 1 ;$$

iii) Quel que soit $f \in \{1, \dots, N\}$, on lui associe le quotient $q = \lfloor N/p \rfloor$ et le reste $z = N - fq$, $0 \leq z < f$ de la division euclidienne de N par f , (cette division est unique).

On a immédiatement :

$$\frac{N}{q} = f + \frac{z}{q}, \quad f \leq \frac{N}{q}, \quad f \leq b_q(q) = \left\lfloor \frac{N}{q} \right\rfloor,$$

et

$$N < f(q+1), \quad f > \frac{N}{q+1}, \quad f \geq b_{q+1}(q) = \left\lfloor \frac{N}{q+1} + 1 \right\rfloor,$$

donc tout f appartient à l'ensemble $\mathcal{I}_{\lfloor N/p \rfloor}$ et à lui-seul. ■

Cas particulier : Pour tout diviseur q_d de N , l'ensemble \mathcal{I}_{q_d} est défini ; en effet $f_d = \frac{N}{q_d}$ appartient à \mathcal{I}_{q_d} .

Conséquence : La partition par les ensembles \mathcal{I}_q classe les entiers f de 1 à N en fonction de leur quotient dans la division euclidienne de N par eux-mêmes, on observe ainsi que les seuls quotients q possibles sont ceux qui vérifient la condition $\left\lfloor \frac{N}{q+1} + 1 \right\rfloor \leq \left\lfloor \frac{N}{q} \right\rfloor$.

Notamment le quotient $q = 2$ est atteint si et seulement si $\left\lfloor \frac{N}{3} + 1 \right\rfloor \leq \left\lfloor \frac{N}{2} \right\rfloor$, relation vérifiée pour $N = 2, 4, 5$ et $N \geq 6$ (puisque $\frac{x}{2} \geq \frac{x}{3} + 1$ ssi $x \geq 6$) ;

le quotient $q = 3$ est atteint si et seulement si $\left\lfloor \frac{N}{4} + 1 \right\rfloor \leq \left\lfloor \frac{N}{3} \right\rfloor$

relation vérifiée pour $N = 3, 6, 7, 9, 10, 11$ et $N \geq 12$, ($\frac{x}{3} \geq \frac{x}{4} + 1$ ssi $x \geq 12$) etc...

1.2.3. Exemples

Pour $N = 5$, on a $\mathcal{I}_1 = \{3, 4, 5\}$; $\mathcal{I}_2 = \{2\}$; $\mathcal{I}_5 = \{1\}$

Pour $N = 9$, on a $\mathcal{I}_1 = \{5, 6, 7, 8, 9\}$; $\mathcal{I}_2 = \{4\}$; $\mathcal{I}_3 = \{3\}$; $\mathcal{I}_4 = \{2\}$; $\mathcal{I}_9 = \{1\}$;

On trouvera d'autres exemples de ces partitions dans le tableau ci-contre.

N = 1	N = 2	N = 3	N = 4	N = 5	
q	I _q	q	I _q	q	I _q
1	1	1	{ 2 3 }	1	{ 3 5 }
		2	1	2	2
		3		4	1
				5	
N = 6	N = 7	N = 8	N = 9	N = 10	
q	I _q	q	I _q	q	I _q
1	{ 4 6 }	1	{ 5 8 }	1	{ 6 10 }
2	3	2	{ 3 4 }	2	{ 4 5 }
3	2	4	2	3	3
6	1	8	1	4	2
				9	1
				10	1
N = 11	N = 12	N = 13	N = 14	N = 15	
q	I _q	q	I _q	q	I _q
1	{ 6 11 }	1	{ 7 13 }	1	{ 8 15 }
2	{ 4 5 }	2	{ 5 6 }	2	{ 6 7 }
3	3	3	4	3	{ 4 5 }
5	2	4	3	4	3
11	1	6	2	7	2
		12	1	14	1
				15	

Partitions de $\{1, \dots, N\}$ en fonction de $q = \lfloor N/r \rfloor$,

intervalles I_q.

1.3. Application à λ_1 .

Dans la suite on s'intéresse aux entiers f qui sont atteints par l'application λ_1 , ainsi que au quotient q et au reste r qui leur sont associés dans la division euclidienne de N par f . On définit l'ensemble :

$$E_{M,N} = \left\{ f \in N : f = d_1(n), n \in \bar{S}_N^M \right\}.$$

A l'évidence on a les résultats suivants :

- Quand M divise N : $E_{M,N} = \left\{ \frac{N}{M}, \frac{N}{M} + 1, \dots, N \right\}$,
- Quand M ne divise pas N : $E_{M,N} = \left\{ \left[\frac{N}{M} + 1 \right], \dots, N \right\}$,

(le raisonnement par l'absurde est immédiat).

Proposition 3 - Soit l'ensemble $F_{M,N} = \left\{ q = \left[\frac{N}{f} \right], f \in E_{M,N} \right\}$,

- Quand M divise N on a $F_{M,N} \subset \{1, \dots, M\}$
- Quand M ne divise pas N on a $F_{M,N} \subset \{1, \dots, M-1\}$, la valeur $M-1$ n'étant atteinte par un quotient q que si et seulement si la condition $b_1(M-1) \leq b_0(M-1)$ est vérifiée.

Démonstration :

- i) Quand M divise N , M est évidemment le quotient de N par N/M , et N/M est le seul point f de $E_{M,N}$ pour lequel on a $\left[\frac{N}{f} \right] = M$. Le sous-ensemble $\left\{ \frac{N}{M} + 1, \dots, N \right\}$ est recouvert par une partition constituée d'ensembles I_q avec nécessairement $q \leq M-1$.

ii) Quand M ne divise pas N , le plus petit élément de $E_{M,N}$ est égal à $b_i(M-1)$.

- Si l'on a $b_i(M-1) \leq b_s(M-1)$, l'ensemble I_{M-1} existe et il fait partie de la partition des I_q qui recouvre $E_{M,N}$ avec nécessairement $q \leq M-1$, la valeur $M-1$ étant atteinte pour tous les f de I_{M-1} .

- Si au contraire $b_s(M-1) < b_i(M-1)$, aucun quotient $[N/p]$ n'est égal à $M-1$. La valeur $b_i(M-1)$ appartient à un ensemble $I_{\bar{q}}$ tel que l'on a $b_s(M-1) < b_s(\bar{q})$ d'où l'on déduit nécessairement $\bar{q} < M-1$. L'ensemble $E_{M,N}$ est recouvert par une partition constituée par $I_{\bar{q}} \cap E_{M,N}$ et d'autres ensembles I_q avec $q < \bar{q}$. ■

V.2 - La distribution $\pi_{M,N}^1$.

2.1. Préliminaires.

2.1.1. Soit $\pi_{M,N}$ une probabilité définie sur \bar{S}_N^M , invariante par le groupe des permutations Σ_M , avec $\omega_{M,N}(n)$ la valeur de la probabilité de cette distribution en tout point n appartenant à \bar{S}_N^M .

Soit $\pi_{M,N}^1$ la probabilité image de $\pi_{M,N}$ par l'application λ_1 qui à tout point n fait correspondre $\lambda_1(n) = \sum p n_i$. Par construction $\pi_{M,N}^1$ est définie sur $E_{M,N}$. Soit l'ensemble $\Omega_p = \{n \in \bar{S}_N^M : \lambda_1(n) = p\}$, en tout point f de $E_{M,N}$ la distribution $\pi_{M,N}^1$ a une probabilité

qui est égale à $\omega_{M,N}^1(p) = \sum_{n \in \Omega_p} \omega_{M,N}(n)$.

2.1.2. Pour calculer les probabilités $\omega_{M,N}^1(p)$, il est commode de considérer l'ordre de multiplicité Q avec lequel λ_1 atteint $f_1(Q \geq 1)$.

En effet, par exemple avec les valeurs $N = 9$ et $M = 4$ les 3 points suivants de \bar{S}_9^4 $(3, 3, 3, 0)$, $(3, 3, 2, 1)$ et $(3, 2, 2, 2)$ conduisent à la même valeur pour $\lambda_1(n)$ mais avec des ordres de multiplicité différents (à savoir $Q = 3, 2, 1$).

Lorsque une valeur f est atteinte avec l'ordre Q , le point n possède Q coordonnées égales à f et $M - Q$ coordonnées qui sont inférieures à f , de somme $N' = N - fQ$. A cela correspond la décomposition arithmétique suivante de N par f : $N = fQ + N'$. Connaissant la division euclidienne de N par f : $N = fq + r$, on en déduit que les décompositions arithmétiques de N par f sont de la forme $N = f(q - k) + (r + fk)$ pour $k = 0, \dots, q-1$.

On pose les ensembles suivants :

$$\Omega_{f,Q} = \left\{ n \in \bar{S}_N^M : \lambda_1(n) = f, f \text{ atteint exactement sur } Q \text{ coordonnées} \right\}$$

et la partition

$$\Omega_f = \bigcup_{Q=1}^q \Omega_{f,Q}, \quad q = \lfloor N/f \rfloor.$$

Les ensembles $\Omega_{f,Q}$ et Ω_f sont évidemment invariants par le groupe des permutations Σ_M .

2.1.3. Pour tout f de $E_{M,N}$ on a le quotient $q = \lfloor N/f \rfloor$ de la division de N par f . Pour tout $k = 0, \dots, q-1$ on pose les quantités $Q = q - k$, $M' = M - Q = M - q + k$, $N' = N - fQ = r + fk$.

On a les propriétés immédiates suivantes :

- a) Le nombre M' de coordonnées inférieures à f est compris entre $M - q$ et $M - 1$:

- Quand M divise N , la borne inférieure des valeurs de M' , est égale à zéro, elle est atteinte uniquement pour le point $f = N/M$.
 - Quand M ne divise pas N , la borne inférieure des valeurs de M' est égale à 1, elle est atteinte si et seulement si la condition $b_i(M-1) \leq b_i(N-1)$ est vérifiée, elle est alors atteinte uniquement pour les points de l'ensemble I_{n-1} .
- b) La somme N' des coordonnées de n inférieures à f est comprise entre $N-fq = \underline{q}$ et $N-f$. Quand M divise N , la borne inférieure de N' est égale à zéro, elle est atteinte uniquement pour le point N/M ; simultanément avec le zéro pour M' . Par ailleurs $N' = \underline{q} + fk$ est strictement inférieur à f si et seulement si l'ordre de multiplicité φ de f est égal au quotient $q = [N/p]$.

2.1.4. Lorsqu'un point n appartient à $\Omega_{f\varphi}$ tous les points transformés de n par les permutations de Σ_n appartiennent à $\Omega_{f\varphi}$. Tout point qui possède φ coordonnées égales à f et $M' = n - \varphi$ coordonnées n_j quelconques vérifiant les conditions $\sum n_j = N'$, $n_j < f$ appartient à $\Omega_{f\varphi}$. D'où l'idée de "factoriser" tout point n de $\Omega_{f\varphi}$ en un élément du produit cartésien $A_{f\varphi} \times \bar{S}_N^{M'}$, où $A_{f\varphi}$ est l'ensemble des φ -uples qui ont toutes leurs coordonnées égales à f . Soit l'application qui à tout $n \in \bar{S}_N^M$ fait correspondre le point n' de $\bar{S}_N^{M'}$ dont les coordonnées n'_j sont égales aux coordonnées n_j de n qui sont inférieures à f , n' est la "projection" de n sur $\bar{S}_N^{M'}$.

Inversement à tout point n' de $\bar{S}_N^{M'}$ on fait correspondre la classe des points \tilde{n} de \bar{S}_N^M qui ont φ coordonnées égales à f , et les M' autres coordonnées égales aux coordonnées de n' . On remarque que la classe des \tilde{n} contient $\binom{\varphi}{M}$ points n distincts.

Trois cas peuvent se produire :

- a) le point n' a toutes ses coordonnées inférieures à f . Dans ces conditions tous les points \tilde{n} ont Q et exactement Q coordonnées égales à f , ils appartiennent tous à $\Omega_f \varphi$; nous dirons que le point n' est $fQM N$ -acceptables;
- b) le point n' a au-moins une coordonnée supérieure à f . Dans ces conditions tous les points \tilde{n} ont leur plus grande coordonnée supérieure à f et ils n'appartiennent plus à $\Omega_f \varphi$; nous dirons qu'ils sont $fQM N$ -inacceptables;
- c) le point n' a une coordonnée au plus égale à f . Dans ces conditions, les points \tilde{n} appartiennent à Ω_f mais l'ordre de multiplicité de f est supérieur à Q ; nous dirons encore qu'ils sont $fQM N$ -inacceptables.

Commentaire : La définition ci-contre de la $fQM N$ -acceptabilité revient à dire que n' est $fQM N$ -acceptable si et seulement si il existe un point n de $\Omega_f \varphi$ tel que n' est la projection de n sur $\bar{S}_{N'}^{M'}$.

Conséquence : Tous les points de $\bar{S}_{N'}^{M'}$ sont $fQM N$ -acceptables si et seulement si N' est inférieur strictement à f , c'est-à-dire si et seulement si $Q = q$ ou $k = 0$.

2.1.5. Le balayage de Ω_p .

Pour tout $f \in E_{MN}$ et $q = [N/p]$ nous utiliserons la formule de sommation :

$$\omega_{MN}^u(p) = \sum_{Q=1}^q \sum_{\Omega_f \varphi} \omega_{MN}(n)$$

qui nous invite à faire les sommations partielles sur $\Omega_f \varphi$.

Pour faire ces dernières, on peut dans un premier temps recenser les points de $n \in \Omega_f \varphi$ qui interviennent dans la sommation (avant les pondérations par $\omega_{MN}(n)$). On utilise pour cela la "factorisation" de n dans l'espace produit $A_{f\varphi} \times \bar{S}_{N'}^{M'}$.

A l'évidence un point n appartient à $\Omega_f \mathcal{Q}$ si et seulement si sa projection n' sur $\bar{\Sigma}_{N'}^{n'}$ est $f \mathcal{Q} \cap N$ -acceptable, auquel cas les $C_n^{\mathcal{Q}}$ points de la classe des \tilde{n} appartiennent à $\Omega_f \mathcal{Q}$.

Pour tout $f' \geq f$ on définit l'ensemble $\Omega_{f'} = \{ n' \in \bar{\Sigma}_{N'}^{n'} : \sup(n_j) = f' \}$

Les points n' qui sont $f \mathcal{Q} \cap N$ -inacceptables sont alors les points de $\bigcup_{f'=f}^{N'} \Omega_{f'}$. On observe ainsi que le problème initial du recensement de Ω_f renvoie éventuellement à $r+1+f(k-1)$ problèmes analogues dans $\bar{\Sigma}_{N'}^{n'}$, à savoir les recensements des $\Omega_{f'}$, $f' = f, \dots, r+fk$.

Remarques : 1) Ces nouveaux problèmes ne se posent pas quand on a $\mathcal{Q} = q$,

2) Ils sont a priori plus simples que le problème initial puisque la dimension M' et la somme N' de $\bar{\Sigma}_{N'}^{n'}$ sont respectivement inférieures aux paramètres M et N initiaux. On a également la propriété suivante qui va dans le sens de la simplification.

Proposition 4 - Pour tout $M \geq 1$ et $N \geq 0$, pour tout $f \in E_{MN}$ et tout $f' \geq f$, l'ordre de multiplicité \mathcal{Q}' avec lequel f' est atteint est inférieur à q .

En effet, en utilisant les notations précédentes et des extensions évidentes sur $\bar{\Sigma}_{N'}^{n'}$ on a

$$\mathcal{Q}' \leq q' = \left[\frac{N'}{f'} \right] = \left[\frac{r+fk}{f'} \right] \leq \left[\frac{r+fk}{f} \right],$$

c'est à dire $\mathcal{Q}' \leq q' \leq k \leq q-1$ ■

Il en résulte que pour tout $\Omega_{f'}$ la partition en $\Omega_{f'} \mathcal{Q}'$ contient moins de sous-ensembles (donc de sommes partielles) que la partition de Ω_f en $\Omega_f \mathcal{Q}$.

2.2. Le calcul des probabilités $\omega_{MN}^1(t)$.

2.2.1. Distributions conditionnelles et marginales.

Pour le calcul des $\omega_{MN}^1(p)$ il sera commode d'introduire les distributions marginales et conditionnelles de Π_{MN} . (Pour simplifier on supprime les indices M et N quand ils ne sont pas nécessaires).

La probabilité Π_{MN} étant invariante par Σ_M , on sait que toutes les distributions marginales qui dépendent du même nombre u , $u=1, \dots, n-1$, de variables sont égales entre elles, de même que toutes les distributions conditionnelles par rapport à un même nombre de variables. Avec des notations évidentes on pose $\omega_m(n_1, \dots, n_u)$ les valeurs des probabilités marginales de la loi Π qui dépendent de u variables, et $\omega_{\Sigma/n_1, \dots, n_u}(n_{u+1}, \dots, n_n)$ les valeurs des probabilités conditionnelles de Π par rapport à u variables. On rappelle la relation élémentaire

$$\omega(n_1, \dots, n_n) = \omega_m(n_1, \dots, n_u) \cdot \omega_{\Sigma/n_1, \dots, n_u}(n_{u+1}, \dots, n_n)$$

avec
$$\sum_{i=1}^M n_i = N - \sum_{i=1}^u n_i.$$

Proposition 5 - Pour tout entier $u=1, \dots, n-2$ la distribution conditionnelle de Π par rapport à u variables est invariante par le groupe des permutations Σ_{n-u} .

En effet Σ_{n-u} est un sous-groupe de Σ_M , l'invariance de la distribution conditionnelle résulte de l'expression de $\omega_{\Sigma/n_1, \dots, n_u}(n_{u+1}, \dots, n_n)$ et de l'invariance de $\omega(n_1, \dots, n_u, \dots, n_n)$ par Σ_{n-u} lorsque n_1, \dots, n_u sont fixées. ■

Complément. Pour $u = M-1$ la distribution conditionnelle par rapport à n_1, \dots, n_{M-1} est égale à la distribution de DIRAC

$$\omega_{c/n_1, \dots, n_{M-1}}(n_M) = \int_{n_M, N - \sum_{i=1}^{M-1} n_i}^{\dots},$$

(son invariance par \sum_1 est triviale).

Remarque - On n'a évidemment pas la même propriété pour les distributions marginales.

2.2.2. Applications.

Pour tout $u \leq M-1$ et $\bar{n}_i, i=1, \dots, u, \sum \bar{n}_i = \bar{N} < N$, la sommation des $\omega(n)$ sur l'ensemble des points n^* de \bar{S}_N^M qui ont leurs coordonnées $n_i = \bar{n}_i, i=1, \dots, u$, est par définition égale à la probabilité marginale $\omega_{N-M}(\bar{n}_1, \dots, \bar{n}_u)$ de la distribution π_{N-M} .

De la même façon la probabilité du point n' de $\bar{S}_{N-\bar{N}}^{M-u}$ qui a pour coordonnées n_{u+1}, \dots, n_M est la distribution conditionnelle de π_{N-M} par rapport à n_1, \dots, n_u .

On reprend la factorisation de ω sur l'espace produit $A_{1\varphi} \times \bar{S}_N^M$, on en déduit que la sommation des $\omega_{\pi_{N-M}}(n)$ sur l'ensemble des n^* tels que $n_1 = \dots = n_\varphi = 1$ est égale à la probabilité $\omega_{\pi_{N-M}}(n_1 = \dots = n_\varphi = 1)$.

Soit π'_{N-M} la probabilité image de π_{N-M} par l'application qui à tout $n \in \bar{S}_N^M$ fait correspondre le point $n' \in \bar{S}_N^{M'}$. Par définition c'est la distribution conditionnelle de π_{N-M} par rapport aux φ coordonnées qui sont égales à 1. On en déduit donc la probabilité suivante pour tout point $n' \in \bar{S}_N^{M'}$:

a) Pour tout $Q < M-1$

$$\begin{aligned} \omega'_{\pi'N'}(n') &= \omega_{\pi'N' / n_1 \dots n_Q} (n_{Q+1}, \dots, n_m) \\ &= \frac{\omega_{\pi'N'}(t, t, \dots, t, n_{Q+1}, \dots, n_m)}{\omega_{\pi'N'}(n_1 = \dots = n_Q = t)} \end{aligned}$$

On a vu que $\pi'_{\pi'N'}$ est invariante par le groupe $\sum n - Q$;

b) Pour $Q = M-1$

$$\omega'_{\pi'N'}(n_m) = \int_{n_m} N - \sum_{i=1}^{m-1} n_i$$

2.2.3. Calcul de $\omega'_{\pi'N'}(p)$.

Le calcul est fait pas à pas. Le résultat est repris sous la forme d'un théorème à la fin du paragraphe.

Pour tout $f \in E_{MN}$ et $q = [N/p]$ on applique la formule :

$$\begin{aligned} \omega'_{\pi'N'}(p) &= \sum_{Q=1}^q \sum_{\Omega \neq Q} \omega_{\pi'N'}(n) \\ &= \sum_{\Omega \neq q} \omega_{\pi'N'}(n) + \sum_{Q=1}^{q-1} \sum_{\Omega \neq Q} \omega_{\pi'N'}(n) \end{aligned}$$

i) La sommation sur $\Omega_{f,q}$ est immédiate. En effet on a vu que tous les points n' de $\bar{\Sigma}_{N'}^{f,q}$ sont $fQMN$ -acceptables (en effet $N' = z < f$). Sachant qu'il y a C_M^q possibilités qu'un point de $\bar{\Sigma}_N^M$ ait q coordonnées égales à f , on a immédiatement

$$\sum_{\Omega_{f,q}} \omega_{\pi_N}(n) = C_M^q \omega_{\pi_{Nm}}(n_1 = \dots = n_q = f)$$

dans le cas général, et

$$\sum_{\Omega_{f,q}} \omega_{\pi_N}(n) = \omega_{\pi_N}\left(\frac{N}{M}, \dots, \frac{N}{M}\right)$$

quand M divise N et que $f = \frac{N}{M}$, $q = M$.

ii) Pour tout $k = 1, \dots, q-1$ on a de la même façon :

$$\sum_{n \in A_{f, q-k} \times \bar{\Sigma}_{z+f, k}^{n-q+k}} \omega_{\pi_N}(n) = C_M^{q-k} \omega_{\pi_{Nm}}(n_1 = \dots = n_{q-k} = f) \cdot$$

On sait cependant que tous les points de l'ensemble produit $A_{f, q-k} \times \bar{\Sigma}_{z+f, k}^{n-q+k}$

ne sont pas $fQMN$ -acceptables. La sommation cherchée doit être faite sur $\Omega_{f, q-k}$ uniquement. Pour y parvenir on a vu qu'il faut retirer les points de $\bar{\Sigma}_{z+f, k}^{n-q+k}$ qui appartiennent à tout $\Omega_{f'}$ pour $f' \geq f$; c'est à dire :

$$\sum_{n \in \Omega_{f, q-k}} \omega_{\pi_N}(n) = C_M^{q-k} \omega_{\pi_{Nm}}(n_1 = \dots = n_{q-k} = f) \times \left[1 - \sum_{f'=f}^{z+k} \sum_{\Omega_{f'}} \omega'_{\pi_{qk, z+f, k}}(n') \right]$$

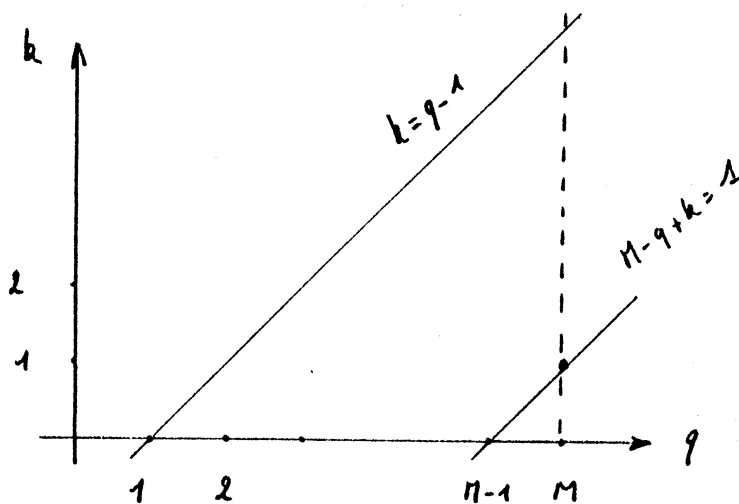
avec pour tout $n' \in \overline{r+fk}^{\eta-q+k}$ $\eta-q+k \geq 2$:

$$\omega'_{\eta-q+k, r+fk}(n') = \frac{\omega_{MN}(t, \dots, t, \eta-q+k, \dots, \eta)}{\omega_{MN}(\eta_1 = \dots = \eta_{q-k} = t)}$$

et pour $\eta-q+k = 1$:

$$\omega'_{\eta-q+k, r+fk}(\eta_n) = \int_{\eta_n, N - \sum_{i=1}^{\eta-1} \eta_i} \left(\sum_{i=1}^{\eta-1} \eta_i = r+fk \right).$$

Ce dernier cas ne se présente que si $M \geq 2$, $q = M$ (c'est à dire M divise N et $f = N/M$) et si $k = 1$



iii) On récapitule les résultats dans le théorème suivant :

Théorème 1 - Pour tout $M \geq 1$, $N \geq 0$, $f \in E_{MN}$, $q = [N/f]$, $r = N - fq$ respectivement le quotient et le reste de la division euclidienne de N par f , la probabilité π'_{MN} image de π_{MN} par l'application λ_1 a les probabilités suivantes :

$$\omega_{\pi N}^{-1}(t) = C_{\pi}^q \omega_{\pi N m} (n_1 = \dots = n_q = t) \\ + \sum_{k=1}^{q-1} C_{\pi}^{q-k} \omega_{\pi N m} (n_1 = \dots = n_{q-k} = t) \times \\ \left[1 - \sum_{p=1}^{q-k} \sum_{r=1}^{r_{p+k}} \omega_{\pi-q+k, r_{p+k}}'(n') \right],$$

où l'on a

$$\omega_{\pi-q+k, r_{p+k}}'(n') = \frac{\omega_{\pi N} (n_1 = \dots = n_{q-k} = t, n_{q-k+1}, \dots, n_{\pi})}{\omega_{\pi N m} (n_1 = \dots = n_{q-k} = t)},$$

et les cas particuliers :

$$\bullet \quad C_{\pi}^q \omega_{\pi} (n_1 = \dots = n_q = t) = \omega\left(\frac{N}{M}, \dots, \frac{N}{M}\right) \text{ quand } M \text{ divise } N$$

et que $f = M \quad q = N/M;$

$$\bullet \quad \omega_{\pi N'}'(n_{\pi}) = \int_{n_{\pi}, N - \sum_{i=1}^{\pi-1} n_i} \text{ quand } \pi' = \pi - q + k = 1, \sum_{i=1}^{\pi-1} n_i = N'.$$

La démonstration résulte des constructions précédentes

Commentaires. Le théorème précédent résout entièrement le calcul des $\omega_{\pi N}^{-1}(p)$, on peut ajouter les commentaires suivants :

1) Pour tout f de l'ensemble $\Gamma_f = \left\{ \left[\frac{N}{2} + 1 \right], \dots, N \right\}$ le quotient q est égal à 1, la formule générale prend la forme simplifiée

$$\omega_{\pi N}^{-1}(p) = M \omega_{\pi N m} (n_1 = t)$$

qui permet un calcul direct.

2) Pour les autres quotients q supérieurs ou égaux à 2, la formule générale fait intervenir des termes de la forme

$$\sum_{\Omega_{t'}^{n'}} \omega_{\pi'N'}^{n'}(n')$$

On remarque que ce sont des termes analogues à ceux que l'on calcule, sur un espace $\bar{\Sigma}_{N'}^{M'}$ avec $M' < M$ et $N' < N$.

Pour tout M et N donnés, la formule générale renvoie donc à des problèmes a priori plus simples, le nombre de ces nouveaux problèmes posés à cette occasion étant nécessairement fini.

3) Si l'on emploie la notation naturelle suivante

$$\sum_{\Omega_{t'}^{n'}} \omega_{\pi'N'}^{n'}(n') = \omega_{\pi'N'}^{t'}(t'),$$

le théorème précédent donne l'expression plus concise :

$$\omega_{\pi N}^{p} = \sum_{k=0}^{q-1} C_M^{q-k} \omega_M^{(n_1 = \dots = n_{q-k} = p)} \left[1 - \sum_{p'=p}^{r+k} \omega_{\pi-q+k, r+k}^{t'} \right]$$

qui met bien en évidence la "cascade" de renvois possibles, sur des couples (M', N') , $M' < M$, $N' < N$.

2.3. Sur les distributions $\pi_{M'N'}$.

Pour tout $M \geq 1$ et $N \geq 0$ et une probabilité π_{MN} sur $\bar{\Sigma}_N^M$ invariante par Σ_M , le calcul de la probabilité π_{MN}^1 image de π_{MN} par λ_1 renvoie à des calculs analogues avec des probabilités $\pi_{M'N'}$ sur $\bar{\Sigma}_{N'}^{M'}$ avec $M' < M$ et $N' < N$,

lesquels peuvent à leur tour renvoyer à d'autres probabilités $\pi_{n''N''}''$ sur $\sum_{N''} \pi_{N''}''$, avec $n'' < n'$ et $N'' < N'$, etc... Le nombre (fini) de ces nouvelles questions "en cascade" dépend de M , N et ρ . A priori toute distribution rencontrée $\pi_{n''N''}''$ peut dépendre du "chemin suivi", ce qui limite l'intérêt d'une tabulation préalable de toute $\omega_{n''N''}''$.

Dans ce chapitre on démontre qu'il n'en est rien et donc par conséquent que l'on peut tabuler à l'avance les $\omega_{n''N''}''$ avant d'aborder le problème de ω_{nN} . La formule générale est donc une formule constructive de proche en proche avec M et N croissants.

On étudie ensuite une famille particulière de probabilités π_{nN} , qui conduit aux exemples de V.3.

2.3.1.

Lemme 1 - Toute distribution conditionnelle d'une distribution conditionnelle de π_{nN} est une distribution conditionnelle de π_{nN} .

Démonstration : Compte tenu de l'invariance par \sum_{n_1} de π_{nN} et de l'invariance par \sum_{n_1, \dots, n_u} des distributions conditionnelles $\pi_{nNc/n_1, \dots, n_u}$ par rapport à u variables, pour simplifier le déroulement des calculs, on n'envisage les distributions conditionnelles de π_{nN} et de $\pi_{nNc/n_1, \dots, n_u}$ que par rapport à des variables bien numérotées ; ceci ne restreint pas la généralité.

i) Avec des notations évidentes la distribution $\pi_{nNc/n_1, \dots, n_u}$ a pour probabilités

$$\omega_{nNc/n_1, \dots, n_u}(n_{u+1}, \dots, n_n) = \frac{\omega_{nN}(n_1, \dots, n_u, \dots, n_n)}{\sum_{n_{u+1}, \dots, n_n} \omega_{nN}(n_1, \dots, n_u, \dots, n_n)}$$

où la sommation \sum^1 est étendue à l'ensemble des points n dont les coordonnées n_1, \dots, n_u sont fixées, les autres étant variables positives ou nulles de somme $N' = N - \sum_j^u n_j$.

On a ainsi une probabilité $\pi'_{N'N'}$ définie sur $\bar{S}_{N'}^{M-u}$.

ii) On envisage la distribution conditionnelle $\pi'_{N'c} / n_{u+1}, \dots, n_v$ qui a pour probabilités :

$$\frac{\omega'_{N'c} / n_{u+1}, \dots, n_v (n_{u+1}, \dots, n_n)}{\pi'_{N'c} / n_{u+1}, \dots, n_v} = \frac{\omega'_{N'N'} (n_{u+1}, \dots, n_v, \dots, n_n)}{\sum^2 \omega'_{N'} (n_{u+1}, \dots, n_v, \dots, n_n)}$$

où la sommation \sum^2 est étendue à l'ensemble des points n' de $\bar{S}_{N'}^{M-u}$ dont les coordonnées n_{u+1}, \dots, n_v sont fixées, les autres étant variables positives ou nulles de somme $N'' = N' - \sum_{u+1}^v n_j$.

On a encore $N'' = N - \sum_1^v n_k$; \sum^3 est donc aussi la sommation \sum^3 étendue sur l'ensemble des points n de \bar{S}_N^M dont les coordonnées n_1, \dots, n_v sont fixées, les autres étant variables positives ou nulles de somme $N - \sum_1^v n_k$.

iii) On a donc

$$\frac{\omega'_{N'c} / n_{u+1}, \dots, n_v (n_{u+1}, \dots, n_n)}{\pi'_{N'c} / n_{u+1}, \dots, n_v} = \frac{\omega_{NN} (n_1, \dots, n_u, \dots, n_v, \dots, n_n)}{\sum^3 \omega_{NN} (n_1, \dots, n_u, \dots, n_v, \dots, n_n)}$$

c'est à dire

$$\frac{\omega'_{N'c} / n_{u+1}, \dots, n_v (n_{u+1}, \dots, n_n)}{\pi'_{N'c} / n_{u+1}, \dots, n_v} = \frac{\omega_{Nc} / n_1, \dots, n_v (n_{u+1}, \dots, n_n)}{\pi_{Nc} / n_1, \dots, n_v}$$

Conséquence :

a) Toutes les probabilités $\pi_{n'N'}^I$, $\pi_{n''N''}''$, $\pi_{n^*N^*}^*$ etc.. que l'on rencontre dans la formule générale du théorème 1 appartiennent à la même famille des distributions conditionnelles de π_{nN} , chacune étant uniquement définie par les paramètres $M^* N^*$ et ne dépend pas de l'enchaînement des renvois $M' N'$, $M'' N''$, ..., $M^* N^*$ tels qu'ils se présentent effectivement dans les calculs ;

b) Les tabulations préalables des distributions images $\pi_{n'N'}^I$ peuvent donc être faites à l'avance, (quand on est dans le cadre de la distribution π_{nN}) ;

c) Cette conséquence prend toute son importance lorsque toute probabilité π_{nN} sur laquelle on travaille est elle-même par construction la distribution conditionnelle d'une distribution $\pi_{\tilde{n}\tilde{N}}$ avec $n < \tilde{n}$, $N < \tilde{N}$.

2.3.2. Les calculs généraux du théorème et du lemme précédents prennent une tournure facile dans le cas particulier suivant :

Soit une distribution T définie sur \mathcal{N} , avec les probabilités $t \geq 0$, $t(e)$; et soient T^{*2} la λ -ième itérée de T par convolution, avec les probabilités $t^{*2}(e)$.

Lemme 2 - Pour tout $M \geq 1$ et $N \geq 0$, on envisage dans \mathcal{N}^M la distribution produit $\prod_1^M T_i$ de M distributions indépendantes de loi T , et la probabilité $T \pi_{nN}$ égale à la restriction du produit $\prod_{i=1}^M T_i$ sur \bar{S}_N^M . Alors la distribution $T \pi_{nN}$ est invariante par Σ_M , et la probabilité de $T \pi_{nN}$ en tout point n de \bar{S}_N^M est égale à

$$\omega_{nN} = \frac{t(n_1) \dots t(n_M)}{t^{*M}(e)}$$

Démonstration :

i) A priori l'expression

$$\omega_{n_N}(n) = \frac{t(n_1) \times \dots \times t(n_M)}{\sum_{\bar{n} \in \mathcal{J}_N^M} t(\bar{n}_1) \dots t(\bar{n}_M)}$$

est invariante par \sum_M .

On établit les distributions marginales successives de $T \Pi_{n_N}$ par récurrence.

Pour simplifier on pose ω le dénominateur de $\omega_{n_N}(n)$.

ii) Distribution marginale en $M-1$ variables.

Soit $N^{(2)} = N - \sum_1^{M-2} n_i$, par construction la distribution marginale en $M-1$ variables a pour probabilités

$$\omega_{\Pi_{N^{(2)}}}(n_1, \dots, n_{M-2}, N^{(2)}) = \frac{1}{\omega} t(n_1) \dots t(n_{M-2}) \sum_{n_{M-1}=0}^{N^{(2)}} t(n_{M-1}) t(N^{(2)} - n_{M-1})$$

d'où
$$\omega_{\Pi_{N^{(2)}}}(n_1, \dots, n_{M-2}, N^{(2)}) = \frac{1}{\omega} t(n_1) \dots t(n_{M-2}) t^{*2}(N^{(2)})$$

avec
$$N^{(2)} + \sum_1^{M-2} n_j = N.$$

iii) Distribution marginale en $M-2$ variables.

On suppose que l'on a la forme suivante pour les probabilités en $M-2+1$ variables

$$\omega_{\Pi_{N^{(2)}}}(n_1, \dots, n_{M-2}, N^{(2)}) = \frac{1}{\omega} t(n_1) \dots t(n_{M-2}) t^{*2}(N^{(2)}),$$

avec
$$N^{(2)} + \sum_1^{M-2} n_j = N.$$

Si l'on pose $N^{(2)} = N - n_{n-2}$, on a :

$$\omega_{\Pi_{N^m}}(n_1, \dots, n_{n-2}, N^{(2)}) = \frac{1}{\mathcal{Q}} t(n_1) \dots t(n_{n-2}) \sum_{n_{n-2}=0}^{N^{(2)}} t(n_{n-2}) t^{*(2)}(N - n_{n-2})$$

c'est à dire :

$$\omega_{\Pi_{N^m}}(n_1, \dots, n_{n-2}, N^{(2)}) = \frac{1}{\mathcal{Q}} t(n_1) \dots t(n_{n-2}) t^{*(2)}(N^{(2)})$$

avec $\sum_1^{n-2} n_i + N^{(2)} = N$.

i) La distribution marginale avec 1 seule variable a pour probabilités

$$\omega_{\Pi_{N^m}}(n_1, N^{(n-1)}) = \frac{1}{\mathcal{Q}} t(n_1) t^{*(n-1)}(N^{(n-1)})$$

avec $n_1 + N^{(n-1)} = N$, soit encore

$$\omega_{\Pi_{N^m}}(n_1) = \frac{1}{\mathcal{Q}} t(n_1) t^{*(n-1)}(N - n_1)$$

ii) On en déduit que $\mathcal{Q} = t^{*M}(N)$. ■

Proposition 6 - Pour tout u entier compris entre 1 et $M-2$, les distributions conditionnelles de $\mathcal{T} \Pi_{N^m}$ par rapport à n_1, \dots, n_u ont pour probabilités

$$\omega_{\Pi_{N^m} / n_1, \dots, n_u}(n_{u+1}, \dots, n_n) = \frac{t(n_{u+1}) \dots t(n_n)}{t^{*(M-u)}(N - \sum_1^u n_i)} ;$$

pour $u = M-1$ on a le cas particulier $\omega_{\pi_N / n_1, \dots, n_{M-1}}(n_n) = \int_{n_n}^{n-1} \dots$

En effet on connaît les distributions marginales ; on en déduit immédiatement l'expression ci-dessus. ■

Remarques :

1) On vérifie sur ces expressions que les distributions conditionnelles sont invariantes par $\sum_{\pi-u}$, et que les marginales ne le sont pas ;

2) Toute probabilité $\tau_{\pi_{M,N}}$ peut être naturellement considérée comme la distribution conditionnelle d'une distribution $\tau_{\tilde{\pi}_{M,N}}$ avec $M < \tilde{M}$ et $N < \tilde{N}$.

Applications . Pour tout M et N positifs et pour tout $f \in E_{M,N}$ le calcul de la distribution $\tau_{M,N}^1$ fait appel aux ensembles $\tilde{S}_{N'}^{M'}$ particuliers suivants pour $k = 0, \dots, q-1$ avec $M' = M - q + k$, $N' = N + k$.

La formule générale du théorème 1 devient ici :

$$\omega_{\pi_N}^{-1}(f) = C_M^q \frac{t(p)^q t^*(M-q)(z)}{t^{*M}(N)}$$

$$+ \sum_{k=1}^{q-1} C_M^{q-k} \frac{t(p)^{q-k} t^*(M-q+k)(z+k)}{t^{*M}(N)} \times$$

$$\left[1 - \sum_{p'=p}^{z+k} \sum_{n \in a_{p'}} \frac{t^{(M-q+k)} \dots t(n)}{t^{*(M-q+k)}(z+k)} \right]$$

dans le cas général, avec la convention

$$t^{*(0)}(0) = 1 \quad \text{quand } M \text{ divise } N, \quad f = \frac{N}{M}, \quad q = M;$$

que l'on peut écrire encore sous la forme plus concise :

$$\omega'_{n,N}(p) = C_n^q \frac{t(p)^q t^{*(M-q)}(z)}{t^{*M}(N)}$$

$$+ \sum_{k=1}^{q-1} C_M^{q-k} \frac{t(p)^{q-k} t^{*(n-q+k)}(z+p^k)}{t^{*M}(N)} \left[1 - \sum_{p=p'}^{z+p^k} \omega'_{n-q+k, z+p^k}(z') \right].$$

V.3 - Exemples.

3.1. La distribution de BOSE-EINSTEIN.

La distribution $\prod_{n \in N}$ la plus simple sur \bar{S}_N^M est sans doute le cas où tous les points n sont équiprobables ; on peut alors se ramener à des problèmes de dénombrement des points de \bar{S}_N^M et de $\Omega_{\neq \emptyset}$.

3.1.1. Si l'on considère les M -uplets (n_1, \dots, n_M) on reconnaît le problème statistique de BOSE-EINSTEIN qui cherche à répartir N objets indiscernables dans M classes. Le nombre total des M -uplets (le cardinal de \bar{S}_N^M) est donc donné par le résultat classique C_{M-1+N}^{M-1} que l'on note ici par commodité BE_N^M (on a $M \geq 1$ et $N \geq 0$).

La probabilité $\omega(n)$ est donc $1/BE_N^M$ pour tout $n \in \bar{S}_N^M$; pour tout \neq on note $BE_N^M(\neq)$ l'effectif des points n de Ω_{\neq} .

Remarques : Avec l'image géométrique des points de \bar{S}_N^M on peut simplement établir les relations

$$BE_N^M = \sum_{k=0}^M BE_N^{M-k}$$

$$BE_N^M = BE_{N-1}^M + BE_N^{M-1}$$

$$BE_{N-M}^M = \sum_{k=1}^{M-1} (-1)^k C_N^k BE_N^{M-k} \quad (N \geq M)$$

qui sont des relations classiques (à des degrés divers) entre les coefficients du triangle de PASCAL (Annexe 1), (réf. 69).

3.1.2. De manière générale le dénombrement des $n \in \Omega_{f, \varphi}$ se fait par l'intermédiaire du dénombrement des points $f \in MN$ -acceptables de tout $\bar{S}_N^{n'}$ concerné.

i) Quand on a $\varphi = q$, tous les points n' de \bar{S}_2^{n-q} sont $f \in MN$ -acceptables d'où

$$\text{card}(\Omega_{f, q}) = C_n^q B E_2^{n-q};$$

ii) Quand M divise N et $f = N/M$ on a un point n unique on pose la convention complémentaire $B E_0^0 = 1$;

iii) Pour tout $k = 1, \dots, q-1$, avec $\varphi = q-k$, les points n' de \bar{S}_{2+k}^{n-q+k} sont au nombre de $B E_{2+k}^{n-q+k}$, mais ils ne sont pas tous $f \in MN$ -acceptables. Le nombre de points n' qui ne sont pas acceptables est précisément égal à :

$$\sum_{p'=p}^{2+k} B E_{2+p'}^{n-q+k} (f')$$

avec quand $M - q + k = 1$ la forme particulière $B E_{N'}^1 (f') = \delta_{1, N'}$.

iv) Pour tout $M \geq 1$, $N \geq 0$ et $f \in E_{MN}$, on en déduit la formule :

$$B E_N^M (f) = C_n^q B E_2^{n-q} + \sum_{k=1}^{q-1} C_n^{q-k} \left[B E_{2+k}^{n-q+k} - \sum_{p'=p}^{2+k} B E_{2+p'}^{n-q+k} (f') \right].$$

Le calcul est direct pour tout $f \in \mathbb{I}_1$, $BE_N^M(f) = C_n^q BE_2^{n-q}$;

pour les autres f de E_{MN} , cette formule renvoie à des dénombrements de points de Ω_f' dans $\mathbb{S}_N^{n'}$. Ce cas particulier des points équiprobables conduit à des calculs relativement faciles qui peuvent être faits à la main pour des petites valeurs M et N .

3.1.3. Exemples.

$$3.1.3.1. \quad M=3, N=4, BE_4^3 = C_6^2 = 15, E_{34} = \{2, 3, 4\}.$$

Les valeurs $p = 4$ et 3 correspondent au quotient $q = 1$, d'où les effectifs respectifs de solutions acceptables 3 et 6 , par application de la formule explicite pour $q = 1$.

La valeur $p = 2$ correspond au quotient $q = 2$:

$$i) \quad f=2, q=2 \quad \text{on a} \quad C_n^q C_{n-q-1}^{n-q-1} = C_3^2 C_0^0 = 3$$

ii) $f=2, q=1$ on est ramené au cas $M'=2, N'=2$, avec $f' < 2$; la table correspondante (2,2) (Annexe 2) indique qu'il y a 1 solution acceptable, soit ici $1 \times C_3^1 = 3$.

On en déduit la table suivante :

N	q	$BE_4^3(f)$
4	1	3
3	1	6
2	2	6

$$3.1.3.2. \quad M=4, N=6, BE_6^4 = C_9^3 = 84, E_{46} = \{2, 3, 4, 5, 6\}.$$

Les valeurs $f = 6, 5$ et 4 correspondent au quotient $q = 1$, d'où les effectifs respectifs de solutions acceptables $4, 12$ et 24 , (calcul direct).

La valeur $f = 3$ correspond au quotient $q = 2$:

i) $f = 3, Q = 2$ on a $C_4^2 C_1^1 = 6$

ii) $f = 3, Q = 1$ on est ramené au cas $M' = 3, N' = 3$ avec $f' < 3$; la table correspondante (3 3) (Annexe 2) indique qu'il y a 7 solutions acceptables soit ici $7 \times C_4^1 = 28$.

La valeur $f = 2$ correspond au quotient $q = 3$:

i) $f = 2, Q = 3$ on a $C_4^3 C_0^0 = 4$

ii) $f = 2, Q = 2$ on est ramené au cas $M' = 2, N' = 2$, avec $f' < 2$; la table correspondante (2 2) indique qu'il y a 1 solution acceptable soit ici $1 \times C_4^2 = 6$.

iii) $f = 2, Q = 1$ on est ramené au cas $M' = 3, N' = 4$, avec $f' < 2$; la table correspondante (3 4) ne donne aucune solution acceptable.

On en déduit la table suivante :

f	q	$BE_c^h(f)$
6	1	4
5	1	12
4	1	24
3	2	34
2	3	10

3.1.3.3. On trouvera en annexe 2 d'autres exemples de tables de $BE_N^M(f)$ pour M compris entre 1 et 6 et N compris entre 1 et 8 à 10.

3.2. Les distributions $T \cap \eta_N$.

La forme analytique particulière de ces probabilités se prête relativement bien à la poursuite des calculs dans certains cas.

3.2.1. La forme analytique des probabilités de π_{MN}^1 est établie à l'aide des multiples itérées de T par convolution. Dans le cas général pour tout $|s| \leq 1$ la fonction génératrice de la distribution T a pour valeur :

$$g_T(s) = \sum_{l \geq 0} s^l t(l)$$

la fonction génératrice de T^{*2} a pour valeur $g_T^2(s)$ et l'on en déduit les probabilités

$$t^{*2}(l) = \frac{1}{l!} \frac{d^l (g_T^2)}{ds^l} (s=0),$$

c'est la dérivée à l'origine d'une application composée, et au besoin on peut la calculer à l'aide de la formule de FAA di BRUNO par exemple, connaissant les dérivées multiples de g à l'origine.

3.2.2. L'utilisation de la formule générale est cependant grandement facilitée lorsque la distribution T est fermée pour la convolution, auquel cas les itérées successives T^{*2} sont plus simples à calculer.

Les paragraphes suivants sont consacrés à l'étude des quatre distributions suivantes :

- distribution de POISSON
- distribution binomiale négative ou de PASCAL
- distribution binomiale
- distribution dite de FELLER-BESSEL

qui sont fermées pour la convolution.

3.3. La distribution multinomiale .

3.3.1. Pour tout μ positif, on considère la distribution de POISSON qui a tout l entier positif ou nul a la probabilité :

$$t(\mu, l) = \frac{e^{-\mu} \mu^l}{l!} .$$

On sait que c'est sur \mathcal{N} une distribution fermée pour la convolution qui vérifie la relation $P_\mu * P_\nu = P_{\mu+\nu}$; avec les probabilités

$$t^{*2}(\mu, l) = t(2\mu, l) .$$

Pour tout $M \geq 1$ et $N \geq 0$ nous prenons M distributions de POISSON identiques indépendantes pour définir une distribution $\tau_{\mathcal{N}, N}$ sur \bar{S}_N^M . Dans ces conditions on a la probabilité suivante pour tout $n \in \bar{S}_N^M$:

$$\begin{aligned} \bar{\omega}_{\mathcal{N}, N}(n) &= e^{-M\mu} \prod_{i=1}^M \frac{\mu^{n_i}}{n_i!} \times \frac{e^{M\mu} N!}{(M\mu)^N} \\ &= \frac{N!}{n_1! \dots n_M!} \times \frac{1}{M^N} . \end{aligned}$$

On reconnaît la distribution multinomiale $m(\cdot; N; \theta)$ définie sur \bar{S}_N^M avec le paramètre vectoriel $\theta_i = 1/M$, $i=1, \dots, M$ (on note ω le point de S_1^M dont toutes les coordonnées ω_i , $i=1, \dots, M$ sont égales à $1/M$).

3.3.2. Pour tout $\theta \in S_1^M$ et pour le cas $\theta = \omega$ invariant par \sum_M on rappelle les valeurs des moments d'ordre 1 et 2 de la distribution multinomiale :

	θ	ω
$\bar{E}(n_i)$	$N \theta_i$	N/M
$\sigma^2(n_i)$	$N \theta_i (1 - \theta_i)$	$N (M-1) / M^2$
$\text{cov}(n_i, n_j)$	$- N \theta_i \theta_j$	$- N / M^2$
$\text{cor}(n_i, n_j)$	$\frac{- \theta_i \theta_j}{\sqrt{\theta_i \theta_j (1 - \theta_i) (1 - \theta_j)}}$	$- 1 / (M-1)$

Les distributions marginales en n_1, \dots, n_u sont des distributions marginales sur \bar{S}_N^{u+1} avec le paramètre $\theta'_i = \theta_i \quad i=1, \dots, u, \quad \theta'_{u+1} = 1 - \sum_{i=1}^u \theta_i$.

Les distributions conditionnelles par rapport à n_1, \dots, n_u sont des distributions multinomiales sur $\bar{S}_{N - \sum_{i=1}^u n_i}^{M-u}$ avec le paramètre $\theta'_j = \frac{\theta_j}{1 - \sum_{i=1}^u \theta_i}$ $j = u+1, \dots, M$.

Pour l'application de la formule générale on a

$$\frac{t^q (P) \quad t^{*(M-q)}(n)}{t^{*M}(N)} = \frac{N!}{M^N} \frac{(M-q)^2}{(1!)^q 2!},$$

avec la convention $0^0 = 1$, $0! = 1$ pour $q = M$, $z = 0$;

$$\frac{t(p)^{q-k} t^{*(n-q+k)}_{(z+tk)}}{t^{*M}(N)} = \frac{N!}{M^N} \frac{(M-q+k)^{z+tk}}{(t!)^{q-k} (z+tk)!}$$

On en déduit la formule générale :

$$\begin{aligned} \omega_{nN}^1(p) &= \frac{N!}{M^N} C_M^q \frac{(n-1)^2}{(t!)^q z!} \\ &+ \frac{N!}{M^N} \sum_{k=1}^{q-1} C_M^{q-k} \frac{(M-q+k)^{z+tk}}{(t!)^{q-k} (z+tk)!} \left[1 - \sum_{p=p}^{z+tk} \omega_{n-1, z+tk}^1(t) \right]. \end{aligned}$$

et pour tout $t \in \left\{ \left\lfloor \frac{N}{2} + 1 \right\rfloor, \dots, N \right\}$ la formule explicite :

$$\omega_{nN}^1(t) = \frac{N!}{M^N} M \frac{(n-1)^2}{t! z!} = M C_N^t \frac{(n-1)^2}{n^{t+2}}$$

3.4. La distribution hypergéométrique .

3.4.1. Pour tout L entier positif et ξ appartenant à $]0, 1[$, on considère la distribution binomiale $B(L, \xi)$ qui a tout $\ell = 0, \dots, L$ a la probabilité

$$t(\xi, L, \ell) = C_L^\ell \xi^\ell (1-\xi)^{L-\ell} .$$

On sait que c'est une distribution fermée pour la convolution qui vérifie la relation : $B(L, \xi) * B(L', \xi) = B(L+L', \xi)$,

avec les probabilités $t^{*L}(\xi, L, e) = t(\xi, 2L, e)$.

Pour tout $M \geq 1$ et $N \geq 0$ nous prenons M distributions binomiales identiques indépendantes pour définir une distribution $\tau_{\pi_{n,N}}$ sur \bar{S}_N^M . Dans ces conditions on a la probabilité suivante pour tout $n \in \bar{S}_N^M$:

$$\omega_{\pi_N}(n) = \frac{1}{C_{ML}^N} \prod_{j=1}^M C_L^{n_j}.$$

On reconnaît une distribution hypergéométrique multiple, invariante par Σ_M , paramétrée par L .

(on utilisera la convention supplémentaire $C_a^b = 0$ si $b > a$).

3.4.2. On rappelle les principales propriétés de la distribution hypergéométrique multiple sur \bar{S}_N^M , invariante par Σ_M :

$$E(n_i) = N/M$$

$$\sigma^2(n_i) = \frac{N(M-1)}{M^2} \frac{ML-N}{ML-1}$$

$$\text{cov}(n_i, n_j) = -\frac{N}{M^2} \frac{ML-N}{ML-1}$$

$$\text{cor}(n_i, n_j) = -\frac{1}{M-1}.$$

Les distributions marginales et conditionnelles sont des hypergéométriques multiples.

Pour l'application de la formule générale on a :

$$\frac{t(t)^q t^{*(M-q)}(r)}{t^{*M}(N)} = \frac{1}{C_{ML}^N} (C_L^P)^q C_{(n-q)L}^2$$

$$\frac{t(t)^{q-k} t^{*(M-q+k)}(r+k)}{t^{*M}(N)} = \frac{1}{C_{ML}^N} (C_L^P)^{q-k} C_{(n-q+k)L}^{2+k}$$

d'où la formule

$$\begin{aligned} \omega_{nN}^1(t) = & \frac{1}{C_{ML}^N} C_n^q (C_L^P)^q C_{(n-q)L}^2 \\ & + \frac{1}{C_{ML}^N} \sum_{k=1}^{q-1} C_M^{q-k} (C_L^P)^{q-k} C_{(n-q+k)L}^{2+k} \left[1 - \sum_{p=p}^{2+k} \omega_{n-q+k, r+k}^1(t') \right] \end{aligned}$$

et pour tout $t \in \left\{ \left[\frac{N}{2} + 1 \right], \dots, N \right\}$, on a la formule explicite

$$\omega_{nN}^1(t) = \frac{1}{C_{ML}^N} M C_L^P C_{(n-1)L}^2$$

3.5. La distribution multinomiale - DIRICHLET composée (réf. 57).

3.5.1. Pour tout α positif et ξ appartenant à $]0, 1[$, on considère la distribution binomiale négative (de PASCAL) $\mathcal{P}_{\alpha, \xi}$ qui a tout ℓ entier positif ou nul a la probabilité

$$t(\xi, \alpha, \ell) = \frac{\Gamma(\alpha + \ell)}{\ell! \Gamma(\alpha)} \xi^\alpha (1-\xi)^\ell$$

On sait que c'est sur \mathcal{N} une distribution fermée pour la convolution qui vérifie la relation : $\mathcal{P}_{\alpha, \xi} * \mathcal{P}_{\beta, \xi} = \mathcal{P}_{\alpha+\beta, \xi}$ (réf. 24 ; 84), avec donc les probabilités $t^{**}(\xi, \alpha, \ell) = t(\xi, \alpha, \ell)$.

Pour tout $M \geq 1$ et $N \geq 0$, nous prenons M distributions de PASCAL indépendantes identiques pour définir une distribution $\Gamma \Pi_{n, N}$ sur \bar{S}_N^M . Dans ces conditions on a la probabilité suivante pour tout $n \in \bar{S}_N^M$:

$$w_{n, N}^{(n)} = \frac{\Gamma(n\alpha) N!}{\Gamma(n\alpha + N)} \prod_{j=1}^M \frac{\Gamma(\alpha + n_j)}{\Gamma(\alpha) n_j!}$$

paramétrée par α .

Remarque. Avec $\alpha = 1$ on a la distribution géométrique sur \mathcal{N} ; laquelle conduit à la distribution de BOSE-EINSTEIN équipondérée sur \bar{S}_N^M .

3.5.2. Cette probabilité peut être obtenue différemment.

Soit $m(\cdot; N / \theta)$ la distribution multinomiale définie sur \bar{S}_N^n conditionnelle par rapport au paramètre vectoriel $\theta \in S_1^n$, et soit H une distribution définie sur le simplexe unité S_1^n , absolument continue par rapport à la mesure de LEBESGUE, de densité h . Après le déconditionnement on obtient les probabilités suivantes pour tout n de \bar{S}_N^n :

$$w_{n, N}^h(n) = \int_{S_1^n} m(n; N / \theta) h(\theta) d\theta.$$

C'est une manière très générale de définir des distributions $\Pi_{n, N}^h$ sur \bar{S}_N^M , nous y reviendrons au chapitre suivant.

Si l'on considère que la densité φ est la densité de la distribution de DIRICHLET $D(a)$ avec a le vecteur $a_i > 0 \quad i=1, \dots, M, \sum_1^M a_i = A$, on obtient en tout $n \in \bar{S}_N^M$ la probabilité :

$$\omega_{n,N}^D = \frac{\Gamma(A) N!}{\Gamma(A+N)} \prod_{i=1}^M \frac{\Gamma(a_i + n_i)}{\Gamma(a_i) n_i!},$$

(utiliser l'intégrale de DIRICHLET).

Cette distribution est invariante par Σ_M si et seulement si tous les a_i sont égaux à un même nombre α ; on reconnaît la probabilité précédente.

3.5.3. Cette distribution sur \bar{S}_N^M a été introduite en 1962 par J.E. MOSIMANN à l'aide du déconditionnement par DIRICHLET de la distribution multinomiale ("compound multinomial distribution", réf. 57), et avant lui en 1923 par EGGENBERGER et POLYA dans le cas binomial $M = 2$. Ci-dessous on indique quelques unes de ses propriétés; cette distribution composée figure également dans le tome 4 de JOHNSON et KOTZ (Ref. 43).

Proposition 7 - Pour tout $M \geq 1$ et $N \geq 0$, et tous $a_i \quad i=1, \dots, M$ réels positifs ou nuls non tous nuls, on a l'identité numérique suivante :

$$\sum_{\sum_{j=1}^M n_j = N} \prod_{i=1}^M \frac{\Gamma(a_i + n_i)}{\Gamma(a_i) n_i!} = \frac{\Gamma(\sum a_k + N)}{\Gamma(\sum a_k) N!}$$

Cette identité résulte immédiatement de la valeur des probabilités $\omega_{n,N}^D$ ■

On rappelle les principaux moments d'ordre 1 et 2 de cette distribution, dans le cas général et le cas invariant par Σ_M :

a	$\alpha \mathbb{1}$
$E(n_i) = a_i \frac{N}{A}$	$\frac{N}{M}$
$\sigma^2(n_i) = a_i \frac{N}{A} \left(1 - \frac{a_i}{A}\right) \frac{A+N}{A+1}$	$\frac{N(\pi-1)}{M^2} \frac{\pi\alpha + N}{\pi\alpha + 1}$
$\text{cov}(n_i, n_j) = -N \frac{a_i a_j}{A^2} \frac{A+N}{A+1}$	$-\frac{N}{M^2} \frac{\pi\alpha + N}{\pi\alpha + 1}$
$\text{cor}(n_i, n_j) = -\sqrt{\frac{a_i a_j}{(A-a_i)(A-a_j)}}$	$-\frac{1}{M-1}$

($\mathbb{1}$ vecteur de \mathbb{R}^M dont toutes les coordonnées sont égales à 1).

Les distributions marginales en n_1, \dots, n_u sont des distributions du même type sur \bar{S}_N^{u+1} avec le paramètre vectoriel $a'_i = a_i \quad i=1, \dots, u$

$$a'_{u+1} = A - \sum_{j=1}^u a_j \quad ;$$

les distributions conditionnelles par rapport à n_1, \dots, n_u sont du même type sur $\bar{S}_{N-\sum_{j=1}^u n_j}^{M-u}$ avec le paramètre vectoriel $a'_j = a_j \quad j=u+1, \dots, M$.

Pour l'application de la formule générale on a :

$$\frac{t(f)^q t^{*(\pi-q)}(z)}{t^{*M}(N)} = \frac{\Gamma(\pi\alpha) N!}{\Gamma(\pi\alpha + N)} \left[\frac{\Gamma(\alpha + f)}{\Gamma(\alpha) f!} \right]^q \frac{\Gamma((\pi-q)\alpha + z)}{\Gamma((\pi-q)\alpha) z!}$$

avec la convention $\frac{\Gamma(0)}{\Gamma(0)} = 1$ lorsque M divise N et $f = \frac{N}{M}$.

$$\frac{t(1)^{q-k} t^{*(n-q+k)} (z+f k)}{t^{*M}(N)} = \frac{\Gamma(n\alpha) N!}{\Gamma(n\alpha+N)} \left[\frac{\Gamma(\alpha+f)}{\Gamma(\alpha) f!} \right]^{q-k} \frac{\Gamma((n-q+k)\alpha + z+f k)}{\Gamma((n-q+k)\alpha) (z+f k)!}.$$

On en déduit la formule générale :

$$\begin{aligned} \omega_{\pi, N}^{D^1}(t) &= \frac{\Gamma(n\alpha) N!}{\Gamma(n\alpha+N)} C_M^q \left[\frac{\Gamma(\alpha+f)}{\Gamma(\alpha) f!} \right]^q \frac{\Gamma((n-q)\alpha + z)}{\Gamma((n-q)\alpha) z!} \\ &+ \frac{\Gamma(n\alpha) N!}{\Gamma(n\alpha+N)} \sum_{k=1}^{q-1} C_M^{q-k} \left[\frac{\Gamma(\alpha+f)}{\Gamma(\alpha) f!} \right]^{q-k} \frac{\Gamma((n-q+k)\alpha + z+f k)}{\Gamma((n-q+k)\alpha) (z+f k)!} \\ &\times \left[1 - \sum_{f'=1}^{z+f k} \omega_{\pi-q+k, z+f k}^{D^1}(f') \right], \end{aligned}$$

et pour tout

$$f \in \left\{ \left[\frac{N}{2} + 1 \right], \dots, N \right\}$$

la formule explicite :

$$\omega_{\pi, N}^{D^1}(t) = M \frac{\Gamma(n\alpha) N!}{\Gamma(n\alpha+N)} \frac{\Gamma(\alpha+f)}{\Gamma(\alpha) f!} \frac{\Gamma((n-1)\alpha + z)}{\Gamma((n-1)\alpha) z!}.$$

3.5.4. On a vu que le cas des points équipondérés est un cas particulier de la distribution multinomiale DIRICHLET composée.

Ceci permet en particulier d'établir simplement les distributions marginales et conditionnelles de la distribution équipondérée :

On a notamment

$$\omega_{\pi, N, M}^{\circ}(n_1, \dots, n_u) = \frac{N! \Gamma(M)}{\Gamma(M+N)} \cdot \frac{\Gamma(M-u + N - \sum_{i=1}^u n_i)}{\Gamma(\pi-u) (N - \sum_{i=1}^u n_i)!}$$

$$= C_{\pi-u+N-\sum_{i=1}^u n_i}^{\pi-u-1} / C_{\pi, N-1}^{M-1}$$

$$= C_{\pi-u+N-\sum_{i=1}^u n_i}^{\pi-u-1} / BE_N^M$$

$$\omega_{\pi, N, M}^{\circ}(n_1) = C_{\pi+N-2-n_1}^{M-2} / C_{\pi, N-1}^{\pi-1}$$

$$\omega_{\pi, N, M}^{\circ} / n_1 \dots n_u (n_{u+1}, \dots, n_n) = \frac{(N - \sum_{i=1}^u n_i)! \Gamma(\pi-u)}{\Gamma(M-u+N - \sum_{i=1}^u n_i)} = \frac{1}{BE_{N-\sum_{i=1}^u n_i}^{M-u}}$$

les distributions conditionnelles sont elles-mêmes équipondérées.

3.6. Une distribution due à FELLER-BESSEL .

3.6.1. Le quatrième exemple est sans doute moins classique.

Au cours du chapitre IV nous avons utilisé la densité de probabilité de FELLER-BESSEL qui pour tout $\rho > 0$ a pour valeur

$$f_{\rho}(x) = e^{-x} \frac{\rho}{x} I_{\rho}(2x) \times \mathbb{1}_{\mathbb{R}_+}(x) ,$$

avec
$$I_p(x) = \left(\frac{x}{2}\right)^p \sum_{k=0}^{\infty} \frac{(x/2)^{2k}}{k! \Gamma(k+p+1)} \mathbb{1}_{\mathbb{R}_+}(x)$$

C'est la densité d'une distribution fermée pour la convolution, qui vérifie

$$f_p * f_\eta = f_{p+\eta} \quad (\text{réf. 24, tome 2}).$$

Cette densité est aussi le mélange suivant de densités gamma

$$g_{a,\lambda}(x) = \frac{\lambda^a}{\Gamma(a)} x^{a-1} e^{-\lambda x} \mathbb{1}_{\mathbb{R}_+}(x),$$

puisque l'on a

$$f_p(x) = \sum_{k=0}^{\infty} u(p,k) \cdot g_{2k+p,1}(x)$$

Pour tout $p > 0$ et $k \geq 0$, les coefficients $u(p,k)$ constituent une série positive de somme 1.

Définition : Pour tout p positif on appelle distribution discrète de FELLER-BESSEL la distribution u_p qui pour tout $k \in \mathbb{N}$ a la probabilité :

$$u(p,k) = \frac{p}{2^p} \frac{\Gamma(2k+p)}{k! 2^{2k} \Gamma(k+p+1)}$$

Proposition 8 - La famille des distributions discrètes de FELLER-BESSEL est fermée pour la convolution.

En effet pour tout p et q positifs, on a les densités suivantes :

$$f_p(x) = \sum_{k \geq 0} u(p, k) \cdot g_{2k+p, 1}(x)$$

$$f_q(x) = \sum_{l \geq 0} u(q, l) \cdot g_{2l+q, 1}(x)$$

et leur produit de convolution :

$$f_p * f_q(x) = f_{p+q}(x) = \sum_{m \geq 0} u(p+q, m) \cdot g_{2m+p+q, 1}(x);$$

mais on a aussi :

$$\begin{aligned} f_p * f_q(x) &= \sum_{k \geq 0} \sum_{l \geq 0} u(p, k) \cdot u(q, l) \cdot g_{2k+p, 1} * g_{2l+q, 1}(x) \\ &= \sum_{m \geq 0} \left(\sum_{n=0}^m u(p, n) \cdot u(q, m-n) \right) \cdot g_{2m+p+q, 1}(x). \end{aligned}$$

Les densités de probabilité gamma ci-contre ont pour valeur $\frac{e^{-x} x^{p+q-1}}{\Gamma(p+q) x^{2m}}$ sont égales au produit d'une application positive sur \mathbb{R}_+^* par des applications linéairement indépendantes ; par identification sur les densités gamma linéairement indépendantes on en déduit :

$$u(p+q, m) = \sum_{n=0}^m u(p, n) \cdot u(q, m-n),$$

or le deuxième membre est la probabilité en m de la distribution $u_p * u_q$; on a par conséquent $u_p * u_q = u_{p+q}$.

3.6.2. Pour tout $M \geq 1$ et $N \geq 0$, nous prenons M distributions discrètes de FELLER-BESSEL indépendantes pour définir une distribution $T\pi_{n,N}$ sur \bar{S}_N^M , la distribution est invariante par Σ_M si et seulement si les paramètres ρ_i sont égaux entre eux. Dans ces conditions on a la probabilité suivante pour tout $n \in \bar{S}_N^M$

$${}_g w_{n,N}(n) = \frac{\rho_1 \dots \rho_M}{\Sigma \rho_i} \frac{N! \Gamma(N + \Sigma \rho_i + 1)}{\Gamma(2N + \Sigma \rho_i)} \prod_{i=1}^M \frac{\Gamma(2n_i + \rho_i)}{n_i! \Gamma(n_i + \rho_i + 1)}$$

dans le cas général, et

$$w_{n,N}(n) = \frac{\rho^M}{\pi \rho} \frac{N! \Gamma(N + M\rho + 1)}{\Gamma(2N + M\rho)} \prod_{i=1}^M \frac{\Gamma(2n_i + \rho)}{n_i! \Gamma(n_i + \rho + 1)}$$

dans le cas invariant par Σ_M .

Proposition 9 - Pour tout $M \geq 1$ et $N \geq 0$, et tous ρ_i ($i=1, \dots, M$) réels positifs, on a l'identité numérique :

$$\sum_{\Sigma y_i = N} \prod_{i=1}^M \frac{\Gamma(2n_i + \rho_i)}{n_i! \Gamma(n_i + \rho_i + 1)} = \frac{\Sigma \rho_i}{\rho_1 \dots \rho_M} \frac{\Gamma(2N + \Sigma \rho_i)}{\Gamma(N + \Sigma \rho_i + 1) N!}$$

Cette identité résulte immédiatement de la valeur des probabilités ${}_g w_{n,N}(n)$ ■

Cette identité est à rapprocher des propositions au chapitre IV.13.

Les distributions marginales en n_1, \dots, n_u sont des distributions du

même type sur $\bar{S} \frac{u+1}{N}$ avec les paramètres $\rho'_i = \rho_i \quad i=1, \dots, u$

$$\text{et } \rho'_{u+1} = \sum_{j=1}^u \rho_j$$

Les distributions conditionnelles par rapport à n_1, \dots, n_u sont du

même type sur $\bar{S} \frac{M-u}{N - \sum_{j=1}^u n_j}$ avec les paramètres $\rho'_j = \rho_j \quad j=u+1, \dots, M$.

3.6.3.

Pour l'application de la formule générale on a :

$$\frac{t(t)^q t^{*(n-q)}(z)}{t^{*n}(N)} = \left[\frac{\rho \Gamma(z+\rho)}{z! \Gamma(z+\rho+1)} \right]^q \frac{(n-q)\rho \Gamma(2z+(n-q)\rho)}{z! \Gamma(z+(n-q)\rho+1)} \frac{n! \Gamma(N+n\rho+1)}{n\rho \Gamma(2N+n\rho)}$$

avec la convention $\frac{0 \cdot \Gamma(0)}{0!} = 1$ lorsque M divise N et $t = \frac{N}{n}$;

$$\frac{t(t)^{q-k} t^{*(n-q+k)}(z+k)}{t^{*n}(N)} = \left[\frac{\rho \Gamma(z+\rho)}{z! \Gamma(z+\rho+1)} \right]^{q-k} \frac{(n-q+k)\rho \Gamma(2z+2zk+(n-q+k)\rho)}{(z+k)! \Gamma(z+k+(n-q+k)\rho+1)}$$

$$\times \frac{n! \Gamma(N+n\rho+1)}{n\rho \Gamma(2N+n\rho)}$$

On en déduit la formule générale suivante :

$$\begin{aligned} \omega'_{\pi, N}(t) &= \frac{N! \Gamma(N + \pi\rho + 1)}{\pi\rho \Gamma(2N + \pi\rho)} C_{\pi}^q \left[\frac{\rho \Gamma(2t + \rho)}{t! \Gamma(t + \rho + 1)} \right]^q \frac{(\pi - q)\rho \Gamma(2z + (\pi - q)\rho)}{z! \Gamma(z + (\pi - q)\rho + 1)} \\ &+ \frac{N! \Gamma(N + \pi\rho + 1)}{\pi\rho \Gamma(2N + \pi\rho)} \sum_{k=1}^{q-1} C_{\pi}^{q-k} \left[\frac{\rho \Gamma(2t + \rho)}{t! \Gamma(t + \rho + 1)} \right]^{q-k} \frac{(\pi - q + k)\rho \Gamma(2z + 2tk + (\pi - q + k)\rho)}{(z + tk)! \Gamma(z + tk + (\pi - q + k)\rho + 1)} \\ &\left[1 - \sum_{t'=t}^{z+tk} \omega'_{\pi - q + k, z + tk}(t') \right] \end{aligned}$$

et pour tout $t \in \left\{ \left\lfloor \frac{N}{2} + 1 \right\rfloor, \dots, N \right\}$ la formule explicite :

$$\omega'_{\pi, N}(t) = \frac{N! \Gamma(N + \pi\rho + 1)}{\rho \Gamma(2N + \pi\rho)} \frac{\rho \Gamma(2t + \rho)}{t! \Gamma(t + \rho + 1)} \frac{(\pi - 1)\rho \Gamma(2z + (\pi - 1)\rho)}{z! \Gamma(z + (\pi - 1)\rho + 1)}$$

3.7. Compléments.

3.7.1. Les trois distributions invariantes sur \bar{S}_N^M : multinomiale, multinomiale-DIRICHLET composée, hypergéométrique multiple ont les mêmes espérances $E(n_i)$ et mêmes coefficients de corrélation $\text{cor}(n_i, n_j)$. En revanche les variances et covariances sont différentes :

	variances	covariances
multinomial :	$\frac{N(N-1)}{M^2}$	$-\frac{N}{M^2}$
multinomial de DIRICHLET :	$\frac{N(N-1)}{M^2} \quad \frac{\pi\alpha + N}{\pi\alpha + 1}$	$-\frac{N}{M^2} \quad \frac{\pi\alpha + N}{\pi\alpha + 1}$
hypergéométrique :	$\frac{N(N-1)}{M^2} \quad \frac{\pi L - N}{\pi L - 1}$	$-\frac{N}{M^2} \quad \frac{\pi L - N}{\pi L - 1}$

les paramètres α et L étant positifs on en déduit que $\frac{\eta_{\alpha+N}}{\eta_{\alpha+1}} \geq 1$
 et $\frac{\eta_{L-N}}{\eta_{L-1}} \leq 1$.

A ce titre la distribution multinomiale apparaît comme une distribution intermédiaire entre les deux autres familles de distributions envisagées.

3.7.2. La distribution multinomiale $m(\cdot; N; \theta)$ est aussi celle des trois qui a donné lieu à des études sur les valeurs extrêmes $\sup(n_i)$, $\inf(n_i)$, $\sup(n_i) - \inf(n_i)$, $\sup(n_i) / \inf(n_i)$, pour les points de \bar{S}_N^M avec des θ_i égaux (invariance par Σ_M) ou non.

Ces questions remontent également à WHITWORTH ; à partir de sa proposition n° XXIII, GOOD en déduit la fonction génératrice de la probabilité

Prob $\left\{ \sup(n_i) \leq t, i=1, \dots, M, \sum n_i = N \right\}$. On a par exemple :

$$\sum_{N=0}^{\infty} \frac{M^N x^N}{N!} \text{Prob} \left\{ \sup(n_i) \leq t, n \in \bar{S}_N^M \right\} = \left(1 + x + \frac{x^2}{2} + \dots + \frac{x^p}{p!} \right)^M$$

pour le cas invariant (réf.32). Il en déduit aussi le cas général et la fonction génératrice des probabilités Prob $\left\{ \inf(n_i) \geq t, n \in \bar{S}_N^M \right\}$; il indique également une formule constructive exacte pour Prob $\left\{ \sup(n_i) \leq t, n \in \bar{S}_N^M \right\}$, et encore des méthodes d'approximation pour calculer les coefficients d'une fonction génératrice (méthodes du point-selle).

GREENWOOD et GLASGOW (réf. 36) ont étudié l'espérance mathématique et la variance de $\sup(n_i)$ et $\inf(n_i)$ dans le cas binomial ($M = 2$), et de façon moins complète pour $M = 3$, invariants par Σ_M . Ils donnent des approximations de ces quantités.

Comme on le verra en troisième partie, une distribution multinomiale peut être "approchée de près" par une distribution de DIRICHLET, ou encore par une distribution normale multivariée. Ces "ressemblances" permettent de calculer des approximations des distributions de $\text{sup}(n_i)$ connaissant les mêmes distributions dans les distributions approchées. KOZELKA (réf. 48) indique une approximation normale pour la distribution des variables

$$t_i = \left(\frac{n_i}{N} - \frac{1}{M} \right) / \sqrt{\frac{M-1}{M^2 N}}, \quad i=1, \dots, M, \quad (*)$$

avec

$$P\{ \text{sup}(t_i) \geq t \} \approx \frac{M}{2\pi} \int_t^{+\infty} e^{-u^2/2} du$$

lorsque $\theta_i = 1/M$, $i=1, \dots, M$.

JOHNSON et YOUNG (réf. 44) ont étudié ces approximations, notamment le critère de COCHRAN qui conduit à un $\text{sup}(y_i)$ pour une distribution de DIRICHLET. Ils donnent des tables de ces diverses approximations pour $\text{sup}(n_i)$ pour les valeurs $2 \leq M \leq 8$, $N = 25, 50$ et 100 . MALLOWS, JOHNSON et KOTZ signalent également des tables non publiées de RAPPEPORT pour $\text{sup}(n_i)$ et $\text{inf}(n_i)$, dans le cas invariant et des valeurs comprises dans $5 \leq M \leq 50$, $10 \leq N \leq 200$ (réf. 44,53).

V.4 - Autres distributions discontinues.

Comme dans le cas des distributions continues (1^o partie chapitre IV) nous examinons ici d'autres distributions discrètes sur le simplexe unité, invariantes ou non par le groupe Σ_M . Comme au chapitre IV l'objectif est de dresser une certaine liste de ces distributions, sans plus se soucier des distributions des parts extrémales.

Remarque - A toute probabilité $\pi_{n,N}$ on peut appliquer la transformation du chapitre III 7 à 9, et partant de là obtenir autant de nouvelles probabilités sur le simplexe.

* Cette expression est erronée en référence 48.

4.1. Le procédé du chapitre V.2.3.

On peut le généraliser au produit de M distributions indépendantes définies sur \mathcal{N} , distinctes ou non, et en déduire une distribution $g \Pi_{n, N}$ qui pour tout $n \in \bar{S}_N^M$ a la probabilité

$$g \Pi_{n, N}(n) = \frac{\prod_{i=1}^M \bar{p}_i^{t_i(n_i)}}{t_1! \times t_2! \times \dots \times t_M! (\sum n_i)!}$$

avec $\sum_{i=1}^M n_i = N$.

Par exemple si l'on applique ce procédé à M distributions de POISSON \mathcal{P}_{μ_i} on obtient la distribution multinomiale générale $m(\cdot; N; \theta)$ avec pour tout $i=1, \dots, M$, les paramètres $\theta_i = \mu_i / \sum_{j=1}^M \mu_j$.

On obtient la distribution multinomiale-DIRICHLET composée avec M distributions de PASCAL $\mathcal{P}_{a_i; \xi}$, et la distribution hypergéométrique multiple avec M distributions binomiales $\mathcal{B}(L_i; \xi)$. La même construction avec des distributions de PASCAL $\mathcal{P}_{a_i; \xi_i}$ ou binomiales $\mathcal{B}(L_i; \xi_i)$ donnent d'autres probabilités $g \Pi_{n, N}$ où subsistent les différents paramètres $\xi_i, i=1, \dots, M$.

4.2. La normalisation par la somme.

4.2.1.

Proposition 10 - Pour tout entier $M \geq 1$, soit \mathcal{D}_M une distribution définie dans \mathcal{N}^M . La distribution image de \mathcal{D}_M par l'opération de "normalisation par la somme" est définie dans l'ensemble des points de coordonnées rationnelles du simplexe unité S_1^M .

Démonstration : Tout point z de coordonnées $z_i = n_i / \sum_{j=1}^M n_j$, $i=1, \dots, M$ a ses coordonnées rationnelles.

Inversement pour tout $i=1, \dots, M$ soient f_i et q_i des entiers vérifiant

$f_i \leq q_i$, $z_i = f_i / q_i$, $\sum_{i=1}^M z_i = 1$, et z le point de coordonnées z_i du simplexe S_1^M .

Soit q le PPCM de q_1, \dots, q_M avec $q = q_i \beta_i$ $i=1, \dots, M$;

d'où $q = \sum \beta_i f_i$. Le point n de coordonnées entières

$n_i = \beta_i f_i$ $i=1, \dots, M$ est "normalisé" au point z de coordonnées z_i . ■

Proposition 11 - Soit $\bar{\Theta}_n$ la distribution image de Θ_n par normalisation par la somme. Si $\Theta_n(n)$ est la probabilité de Θ_n en tout point de \mathbb{N}^M , alors en tout point z du simplexe unité de coordonnées rationnelles égales à f_i / q_i $i=1, \dots, M$ la distribution $\bar{\Theta}_n$ a pour probabilité :

$$\bar{\Theta}_n(z) = \frac{1}{1 - \Theta_n(\vec{0})} \sum_{k \geq 1} \Theta_n(kqz_1, \dots, kqz_j, \dots, kqz_M),$$

où q est le PPCM des dénominateurs q_i $i=1, \dots, M$ et $\vec{0}$ le point de coordonnées nulles.

Démonstration

i) tout point de coordonnées entières kqz_i se projette sur le point du simplexe de coordonnées z_i $i=1, \dots, M$, et seuls les points de coordonnées kqz_i , $k \geq 1$ font ainsi (démonstration par l'absurde).

ii) Seul le point $\vec{0}$ ne se projette pas ainsi sur S_1^M . ■

4.2.2. Exemples.

4.2.2.1. On peut prendre par exemple une distribution Θ_n produit de M distributions indépendantes T_i , d'où les probabilités suivantes :

$$\theta_n(n_1, \dots, n_n) = \prod_{i=1}^M t_i(n_i),$$

et

$$\bar{\theta}_n(z) = \frac{1}{1 - \prod_{i=1}^M t_i(0)} \sum_{k=1}^M \prod_{j=1}^M t_j(kqz_j).$$

4.2.2.2. Soit g_{MN} une distribution sur \bar{S}_N^M et T une distribution sur \mathcal{N} ; on peut prendre le produit $\Theta_n = g_{MN} * T$, d'où les probabilités :

$$\theta_n(n_1, \dots, n_n) = g_{MN}^{\omega}(n_1, \dots, n_n) * t(N),$$

et

$$\bar{\theta}_n(z) = \frac{1}{1-t_0} \sum_{k=1}^{\infty} t(kq) * g_{MN, kq}^{\omega}(kqz_1, \dots, kqz_j, \dots, kqz_M).$$

Lorsque $T = T_1 * T_2 * \dots * T_M$ et g_{MN} est une des distributions de 4.1. on retrouve les mêmes distributions qu'en 4.2.2.1.; dans certains autres cas on rencontre des distributions de la littérature:

i) g_{MN} est la distribution multinomiale $m(\cdot; N; \theta)$ et pour tout $N' > N$, T est la distribution binomiale $B(N', \xi)$ avec $t(0) = (1-\xi)^{N'}$;

Θ_n est la distribution multinomiale $m(n_1, \dots, n_n, N'-N; N'; \xi\theta_1, \dots, \xi\theta_n, 1-\xi)$;

ii) \mathcal{P}_{MN}^{π} est la distribution multinomiale $m(\cdot; N; \theta)$ et \mathcal{T} la distribution de PASCAL $\mathcal{P}_2 \xi$ avec $t(\theta) = (1-\xi)^{\theta}$;

Θ_M est la distribution dite multinomiale négative (réf. 43, 47, 58) ;

iii) \mathcal{P}_{MN}^{π} est la distribution multinomiale et \mathcal{T} la distribution logarithmique avec $t(\theta) = \theta$; Θ_M est la distribution dite logarithmique multiple (réf. 4, 43).

Cependant, quelle que soit la distribution Θ_M , la distribution $\bar{\Theta}_M$ que l'on en déduit en normalisant par la somme paraît être d'usage mal commode.

4.3. La composition par la distribution multinomiale, résultats asymptotiques.

4.3.1. Soit h une distribution de probabilité sur le simplexe unité $S_1^M(\theta)$, $\theta_i \geq 0$ $i=1, \dots, M$, $\sum_1^M \theta_i = 1$; et soit la distribution multinomiale $m(\cdot; N; \theta)$ définie sur les points de coordonnées n_i/N $i=1, \dots, M$, $\sum_1^M n_i = N$ du simplexe unité $S_1^M(y)$.

On suppose que m est conditionnée par θ , à la suite du déconditionnement par la distribution de densité h on obtient une distribution $\pi_{n,N}^h$ qui est définie en tout point n de \bar{S}_N^M ou n/N de S_1^M avec en ce point une probabilité égale à :

$$\omega_{M,N}^h = \int_{S_1^M} m(n; N/\theta) h(\theta) d\theta$$

C'est là une autre façon très générale d'avoir des distributions sur \bar{S}_N^M ou S_1^M . Le chapitre V.3.5. est un cas particulier dans lequel h est une densité de DIRICHLET $D(a)$, cas particulier important puisque l'on verra en troisième partie que les distributions multinomiale et de DIRICHLET sont "conjuguées" au sens Bayésien. De façon générale, pour toute densité h on a les résultats asymptotiques suivants :

Théorème 2 Lorsque l'on fait tendre N vers plus l'infini, les moments de la distribution $\prod_{M,N}^k$ tendent vers les moments correspondants de la distribution de densité h , la convergence est en $1/N$ (MOSIMANN ref. 57).

Démonstration a) On calcule aisément les moments factoriels de la distribution $\prod_{M,N}^k$. En effet pour tout $i=1, \dots, M$ et k_i entier positif ou nul, on prend $N > \sum_{i=1}^M k_i$, on a :

$$\begin{aligned} E \left[n_1^{(k_1)} \dots n_M^{(k_M)} \right] &= E \left[\int_{S_1^M(\theta)} n_1^{(k_1)} \dots n_M^{(k_M)} m(n; N/\theta) h(\theta) d\theta \right] \\ &= E \left[\int_{S_1^M(\theta)} \frac{N!}{(n_1 - k_1)! \dots (n_M - k_M)!} \prod \theta_i^{n_i - k_i} \prod \theta_j^{k_j} h(\theta) d\theta \right] \\ &= \frac{N!}{(N - \sum k_i)!} \int_{S_1^M(\theta)} \prod \theta_j^{k_j} h(\theta) d\theta \\ &= \int_{S_1^M(\theta)} \prod \theta_j^{k_j} h(\theta) d\theta \times N^{\sum k_i} \left[1 + O\left(\frac{1}{N}\right) \right]. \end{aligned}$$

b) D'autre part on a également

$$E \left[\left(\frac{n_1}{N} \right)^{k_1} \dots \left(\frac{n_M}{N} \right)^{k_M} \right] = E \left[\left(\frac{n_1}{N} \right)^{k_1} \dots \left(\frac{n_M}{N} \right)^{k_M} \right] \times \left[1 + O\left(\frac{1}{N}\right) \right],$$

d'où la relation

$$E \left[\left(\frac{n_1}{N} \right)^{k_1} \dots \left(\frac{n_M}{N} \right)^{k_M} \right] = \int_{S_1^M(\theta)} \prod \theta_j^{k_j} h(\theta) d\theta \times \left[1 + O\left(\frac{1}{N}\right) \right].$$

4.3.2. Soit $\{\Omega, \mathcal{F}, \mu\}$ un espace de probabilité sur lequel sont définis les vecteurs aléatoires à valeur dans S_1^m dont les lois images sont la loi définie par $\pi_{n,N}^h$ et par la densité h .

Théorème 3 - Quel que soit le point θ du simplexe unité on pose $Y_{h,N}$ le vecteur dont la loi image est $\pi_{n,N}^h$ et Y_h le vecteur dont la loi est définie par la densité h . Alors $Y_{h,N}$ converge presque sûrement vers Y_h quand on fait tendre N vers $+\infty$.

Démonstration

i) le résultat est trivial si θ est situé en un sommet du simplexe unité.

ii) Quand il n'en est pas ainsi, on se ramène au simplexe S_1^{m-k} tel que θ soit un point intérieur.

On définit le vecteur aléatoire $Y_N^{m,\theta}$ dont la distribution image est la distribution multinomiale $m(\cdot; N; \theta)$. D'après la loi forte des grands nombres généralisée aux variables multinomiales, on a la convergence presque sûre de $Y_N^{m,\theta}$ vers le point θ ; c'est à dire :

Pour tous ε et η positifs il existe un $N(\theta)$ tel que pour tout $N > N(\theta)$ on a

$$\text{Prob} \left\{ \| Y_N^{m,\theta} - Y_h \| \geq \eta \right\} < \varepsilon,$$

sachant que $\| \cdot \|$ est une norme dans \mathbb{R}^m (on sait qu'elles sont équivalentes (réf 14)).

On sait aussi que pour tout θ , $N(\theta)$ est uniformément borné en θ , (Annexe 3).

iii) On prend $\bar{N} = \sup_{\theta} (N(\theta))$

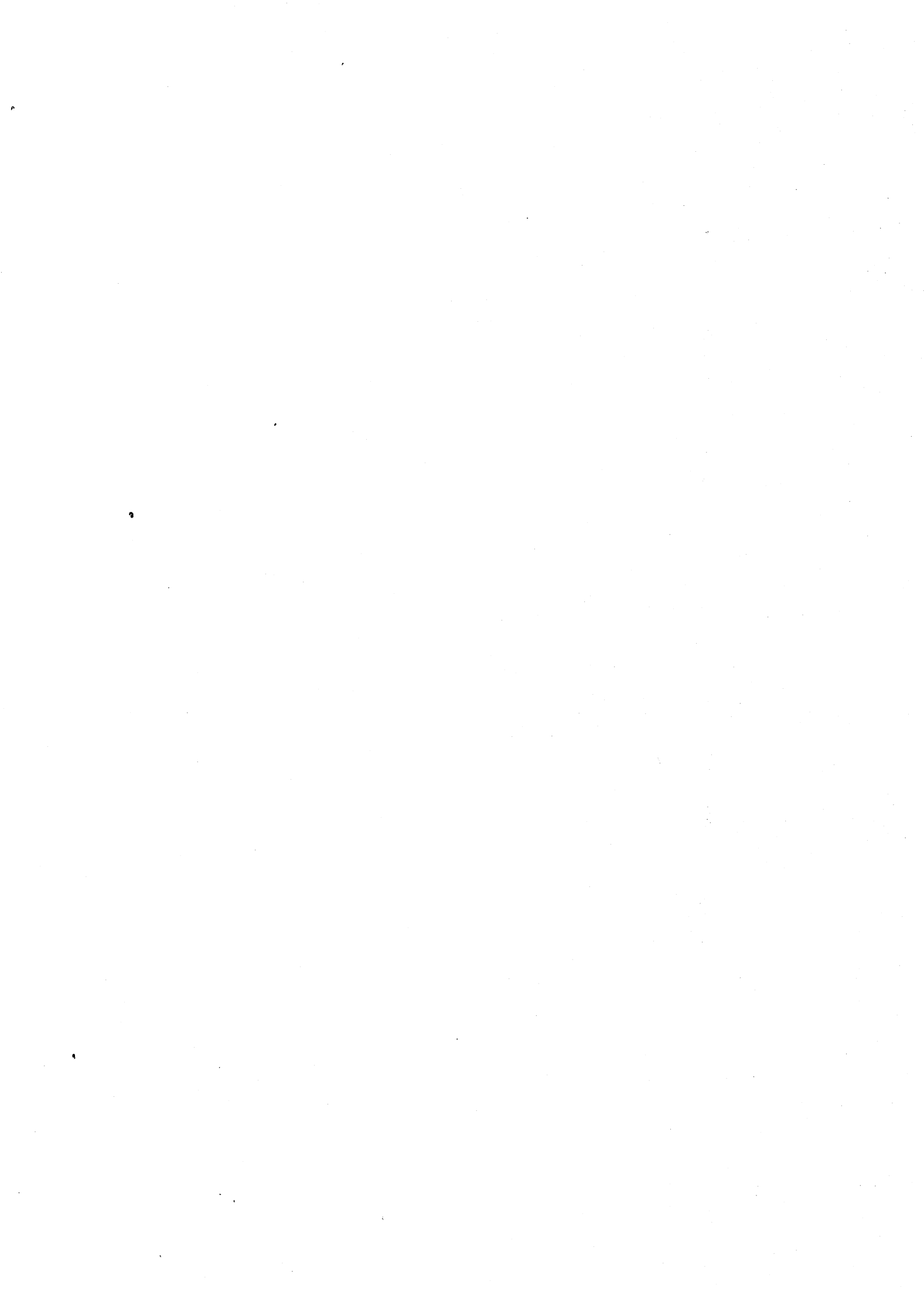
Il en résulte que pour tous ε et η positifs, pour tout θ et tout $N > \bar{N}$, la probabilité de l'événement conditionné par rapport à θ

$$\left\{ \| \gamma_{e,N} - \gamma_e \| \geq \eta / \theta \right\}$$

est inférieure à $\varepsilon \int l(\theta) d\theta$; on en déduit donc que

$$\text{Prob} \left\{ \| \gamma_{e,N} - \gamma_e \| \geq \eta \right\} < \varepsilon \int l(\theta) d\theta = \varepsilon .$$

Il en résulte la convergence presque sûre $\gamma_{e,N} \xrightarrow{p.s.} \gamma_e$
 quand on fait tendre N vers $+\infty$. ■



TROISIEME PARTIE :

Structures statistiques sur le simplexe unité

INTRODUCTION

Jusqu'ici on s'est intéressé aux probabilités du type $\mu(\mathcal{D})$ où μ est une probabilité sur le simplexe unité, et \mathcal{D} un sous-ensemble mesurable défini notamment à l'aide des coordonnées extrémales de $y \in S_1^n$. Les mesures envisagées dépendent la plupart du temps de paramètres.

A présent on transite du point de vue Probabiliste au point de vue Statistique. L'on suppose que l'on dispose d'un échantillon y^1, \dots, y^p de points-partages indépendants qui obéissent tous à une Loi commune. Un problème usuel en statistique consiste à poser a priori une famille de distributions dépendant d'un nombre fini de paramètres, et à deviner ou inférer les paramètres de la distribution suivie, au vu de l'échantillon. On débouche sur la notion de structure statistique telles qu'elles sont définies dans BARRA (Réf. 5).

Dans le cas présent, le support des structures est le simplexe unité S_1^n doté de la tribu de ses boréliens $\mathcal{B}(S_1^n)$ et de la mesure uniforme P. Ayant introduit les densités $D(\cdot; a)$ de DIRICHLET on étudie la structure statistique associée dite de DIRICHLET :

structure invariante^(*) : $\left(S_1^n, \mathcal{B}(S_1^n), \{ D(\cdot; a), a_i = \alpha > 0, i=1..n \} \right)^P$,

structure générale : $\left(S_1^n, \mathcal{B}(S_1^n), \{ D(\cdot; a) \mid a \in \mathbb{R}_+^{*n} \} \right)^P$.

Cette structure se révèle être une structure exponentielle (dite de FISHER-DARMOIS - KOOPMAN - PITTMAN, réf. 4) avec toutes les conséquences classiques au niveau de l'estimation des paramètres et des tests statistiques

(*) La même lettre-symbole P est employée pour désigner la mesure uniforme sur simplexe et la taille de l'échantillon, le contexte suffit pour lever toute confusion.

sur ceux-ci. Pour une étude directe, il est nécessaire d'étudier la densité de DIRICHLET de manière analytique et d'établir ou de rappeler certaines propriétés des fonctions eulériennes Γ , $\Psi = \Gamma'/\Gamma$, Ψ' . La structure invariante par le groupe Σ_M ne dépend que d'un paramètre tandis que la structure générale dépend de M paramètres, (chapitres VI et VII).

Une autre structure est la structure multinomiale définie sur les points de coordonnées n_i/N du simplexe N -discret. C'est une autre structure exponentielle. Entre celle-ci et la structure de DIRICHLET il existe de nombreuses relations qui font l'objet du chapitre VIII. Il est dit deux mots de la structure de FISHER VON-MISES sur l'hypersphère unité, elle même exponentielle, et qui peut également être appliquée sur les points du simplexe unité après une transformation de variable.

Sur le plan du vocabulaire on a indifféremment utilisé les termes exhaustif ou suffisant, en revanche les termes ancillaire et ancillarité (absent de LITRE présent dans ROBERT) ont été utilisés pour désigner les termes libre et liberté (Statistique).

VI - LA STRUCTURE INVARIANTE DE DIRICHLET.

VI.1 - Généralités.

Soit sur le simplexe unité de \mathbb{R}^M l'expression générale des densités de DIRICHLET :

$$D(y_1, \dots, y_n; a_1, \dots, a_n) = \frac{\Gamma(a_1 + \dots + a_n)}{\Gamma(a_1) \dots \Gamma(a_n)} y_1^{a_1-1} \dots y_n^{a_n-1},$$

avec $a_i > 0$, $y_i \geq 0$ $i=1..M$, $\sum_1^M y_i = 1$; pour simplifier on la note $D(y; a)$ quand il n'y a pas d'ambiguïté, (y et a vecteurs de \mathbb{R}^M).

Lorsque tous les exposants a_i sont égaux, on obtient la densité invariante par le groupe des permutations Σ_M qui a pour valeur :

$$D(y_1, \dots, y_n; \alpha) = \frac{\Gamma(M\alpha)}{\Gamma(\alpha)^M} (y_1 \dots y_n)^{\alpha-1},$$

où α réel positif est la valeur commune des exposants. On note $D(\alpha)$ la distribution, $D(\cdot; \alpha)$ sa densité et pour tout point y du simplexe S_1^M $D(y; \alpha)$ sa valeur.

Dans la densité invariante on introduit la forme symétrique fondamentale de poids M : $\sigma_M(y)$ égale au produit $\prod_1^M y_i$ invariant par Σ_M .

Le cas $a_i = n_i$ entier donne un monôme algébrique, le cas $a_i = 1$ est celui de la densité constante $\Gamma(M)$ qui correspond à la mesure uniforme (LEBESGUE) sur le simplexe.

VI.2 - Etude de σ_M sur S_1^M .

2.1. La fonction σ_M est définie continue positive ou nulle sur S_1^M ; nulle sur la frontière du simplexe où le minimum est atteint.

A l'intérieur de S_1^M les extrêma de σ_M sont aussi ceux de $\log \sigma_M$, ils sont solution du système de LAGRANGE :

$$\frac{\partial}{\partial y_i} \left[\log \sigma_M + \mu (\sum y_j - 1) \right] = 0 \quad i=1, \dots, M$$

$$\frac{1}{y_i} + \mu = 0$$

Il y a donc un extrémum unique atteint au point ω , $\omega_i = 1/M$ $i=1 \dots M$; σ_M admet le maximum M^{-M} .

2.2. Soit $\sigma \in]0, M^{-M}[$ et $\sigma(\sigma)$ la variété des $y \in S_1^M$ tels que $\sigma_M(y) = \sigma$, et soit $V(\sigma) = \{ y \in S_1^M : \sigma_M(y) \geq \sigma \}$;

$V(\sigma)$ est un ensemble de niveau et on vérifie que $V(\sigma)$ est un ensemble convexe.

Il en résulte que

pour $\sigma_1 \geq \sigma_2 \iff V(\sigma_1) \subset V(\sigma_2)$

pour $\sigma = M^{-M}$, $\sigma(\sigma)$ et $V(\sigma)$ se réduisent à ω

pour $\sigma = 0$, $V(\sigma) \equiv S_1^M$, $\sigma(\sigma)$ est la frontière ∂S_1^M du simplexe.

VI.3 - Etude de la densité invariante par Σ_M , $D(\cdot; \alpha)$.

La densité $D(\cdot; \alpha)$ est définie sur l'intérieur $\overset{\circ}{S}_1^M$ pour

tout $\alpha > 0$, et définie sur S_1^M pour tout $\alpha \geq 1$.

3.1. Variations en fonction de y .

La densité $D(\cdot; \alpha) = \frac{\Gamma(M\alpha)}{\Gamma(\alpha)^M} \sigma_M^{\alpha-1}$ varie comme $\sigma_M^{\alpha-1}$, (immédiat).

Il en résulte que

si $\alpha = 1$, $D(\cdot; \alpha)$ est la fonction constante $\Gamma(M)$;

si $\alpha > 1$, $D(\cdot; \alpha)$ varie comme σ_M , est nulle sur la

frontière du simplexe, et maximum en ω , de valeur $\Gamma(M\alpha) M^{-n(\alpha-1)} / \Gamma(\alpha)^M$;

si $0 < \alpha < 1$, $D(\cdot; \alpha)$ varie comme $1/\sigma_M$, est minimale

en ω , et tend vers $+\infty$ quand

y tend vers la frontière ∂S_1^M .

3.2. Variations en fonction de α ($\alpha > 0$).

On a
$$\frac{1}{D(y; \alpha)} \frac{\partial}{\partial \alpha} D(y; \alpha) = M [\psi(M\alpha) - \psi(\alpha)] + \log \sigma_M(y)$$

avec
$$\psi(u) = \Gamma'(u) / \Gamma(u)$$

3.2.1. On pose $J(\cdot, n)$ la fonction réelle qui à tout α réel positif fait correspondre $\psi(n\alpha) - \psi(\alpha)$.

Lemme 1 Pour tout $M \geq 2$, la fonction $J(\cdot, M)$ est définie continue dérivable strictement décroissante pour tout $\alpha > 0$.

En effet :

i) On utilise la formule de multiplication de GAUSS (Réf. 82).

$$\Gamma(M\alpha) = (2\pi)^{\frac{1-M}{2}} M^{M\alpha - \frac{1}{2}} \Gamma(\alpha) \Gamma(\alpha + \frac{1}{M}) \dots \Gamma(\alpha + \frac{M-1}{M})$$

d'où par dérivation

$$M \psi(M\alpha) = M \log M + \sum_{k=0}^{M-1} \psi(\alpha + k/M),$$

et

$$J(\alpha, M) = \log M + \frac{1}{M} \sum_{k=0}^{M-1} (\psi(\alpha + k/M) - \psi(\alpha)).$$

ii) Soit $g_k(\alpha) = \psi(\alpha + k/M) - \psi(\alpha)$ fonction continue positive

dérivable, et

$$g_k'(\alpha) = \int_0^{\infty} \frac{t e^{-\alpha t} (e^{-kt/M} - 1)}{1 - e^{-t}} dt ;$$

cette dérivée est négative ; donc la fonction $g_k(\cdot)$ est décroissante sur \mathbb{R}_+^* . Par ailleurs ψ' est décroissante, de limite nulle à $+\infty$ (*), on en déduit que $g_k(\cdot)$ tend vers zéro quand α tend vers $+\infty$, (théorème des accroissements finis).

(*) Par exemple on sait que $\psi'(n+1)$ est égal au reste de rang n de la série convergente $\sum_{k=1}^{\infty} 1/k^2$ (Réf. 1,72,82).

iii) Par conséquent $J(\cdot, M)$ est une fonction définie sur \mathbb{R}_+^* positive continue décroissante, qui tend vers $\log M$ quand α tend vers $+\infty$ ■

Valeurs particulières :

- Au voisinage de l'origine on sait que l'on a $\psi(u) = -\frac{1}{u} - \gamma + O(u)$, donc $J(\cdot, M)$ est un infiniment grand équivalent à $(M-1)/\alpha M$ quand $\alpha \rightarrow 0^+$;

- au voisinage de $+\infty$, $J(\cdot, M)$ décroît vers sa limite $\log M$;

- pour toute valeur entière n de α on a $\psi(1) = -\gamma$

$$\text{et } \psi(n) = -\gamma + 1 + \frac{1}{2} + \dots + \frac{1}{n-1}, \quad n \geq 2$$

où $\gamma = 0,577\dots$ la constante d'EULER-MASCHERONI (Réf. 72,80,82).

On a donc

$$J(1, M) = 1 + \frac{1}{2} + \dots + \frac{1}{M-1}$$

$$J(n, M) = \frac{1}{n} + \dots + \frac{1}{nM-1}, \quad n \geq 2.$$

• Comme
$$\int_n^{Mn} \frac{dx}{x} \leq J(n, M) \leq \int_{n-1}^{Mn-1} \frac{dx}{x},$$

on a
$$\log M \leq J(n, M) \leq \log \frac{Mn-1}{n-1}.$$

• On a encore (Ref. 72) :

$$\psi(n) = \log n - \frac{1}{2n} - \frac{1}{12n^2} + \frac{1}{n^2} O\left(\frac{1}{n}\right),$$

d'où

$$J(n, M) = \log M + \frac{1}{2n} \left(1 - \frac{1}{M}\right) + \frac{1}{12n^2} \left(1 - \frac{1}{M^2}\right) + \frac{1}{n^2} O\left(\frac{1}{n}\right).$$

3.2.2. On en déduit le

Théorème 1 Pour tout $u > \log M$, l'équation en α : $J(\alpha, n) = u$ a une racine réelle unique.

La démonstration est immédiate. ■

Soit $\alpha = \eta_n(u)$ la solution en α de l'équation ci-dessus, la fonction η_n ainsi définie est continue positive dérivable pour $u > \log M$; sa dérivée a pour valeur :

$$\eta_n'(u) = \frac{1}{M \psi'(\alpha M) - \psi'(\alpha)}$$

et par continuité on a les valeurs limites

$$\lim_{u \rightarrow +\infty} \eta_n(u) = 0, \quad \lim_{u \rightarrow \log M} \eta_n(u) = +\infty.$$

3.2.3. Conséquences, variations de $D(\cdot; j \cdot)$ en α .

i) On appelle $d(\alpha) = D(\omega; \alpha) = \frac{\Gamma(M\alpha)}{\Gamma(\alpha)^M} M^{-n(\alpha-1)}$ la valeur maximum de $D(\cdot; j \alpha)$ en y .

La fonction d est croissante sur \mathbb{R}_+^* .

• En effet $\frac{d'(\alpha)}{d(\alpha)} = M \zeta(\alpha, M) - M \log M$

donc $d'(\alpha)$ est positive sur $]0 + \infty[$;

• Les valeurs aux bornes sont :

- quand $\alpha \rightarrow 0^+$, on a $\Gamma(\alpha) = \frac{1}{\alpha} (1 + o(\alpha))$,

d'où $d(\alpha) = \frac{\alpha^{M-1}}{M} M^{-n\alpha} M^M (1 + o(\alpha))$ qui tend vers zéro ;

- quand $\alpha \rightarrow +\infty$, la formule de STIRLING :

$$\Gamma(\alpha) = \alpha^\alpha e^{-\alpha} \left(\frac{2\pi}{\alpha} \right)^{1/2} (1 + o(1/\alpha))$$

entraîne $\frac{\Gamma(M\alpha)}{\Gamma(\alpha)^M} = M^{M\alpha} \alpha^{\frac{M-1}{2}} \times H(M) (1 + o(1/\alpha))$

et $d(\alpha) = \alpha^{\frac{M-1}{2}} M^M \times H(M) (1 + o(1/\alpha))$ qui tend vers $+\infty$;

(On a posé $H(M) = \left(\frac{2\pi}{M} \right)^{\frac{1-M}{2}} M^{-1/2}$) ;

- de plus $d(1) = \Gamma(M)$.

ii) Soit y un point de la frontière ∂S_1^n , pour lequel $\sigma_M = 0$;

$$\begin{aligned} \text{alors } D(y; \alpha) &= 0 \text{ si } \alpha > 1, \\ &= \Gamma(M), \text{ si } \alpha = 1, \\ &= +\infty, \text{ si } \alpha < 1. \end{aligned}$$

iii)

• Pour tout $y \neq \omega$ et $y \notin \partial S_1^M$, $D(y; \cdot)$ est une fonction croissante puis décroissante en α .

$$\text{En effet } \frac{1}{D(y; \alpha)} \frac{\partial}{\partial \alpha} D(y; \alpha) = M \int(\alpha, M) + \log \sigma_H(y)$$

et d'après les propriétés précédentes de \int et σ_H , il y a une racine unique $\alpha(y) = \int_M \left(-\frac{1}{M} \log \sigma_H(y) \right)$ qui annule la dérivée

$D'_\alpha(y; \alpha)$; cette dérivée est positive sur $]0, \alpha(y)[$, négative au-delà.

• Les valeurs aux bornes de $D(y; \cdot)$ sont :

$$\text{- Quand } \alpha \rightarrow 0^+ \quad D(y; \alpha) = \frac{\alpha^{M-1}}{M} \sigma_H^{\alpha-1} [1 + o(\alpha)],$$

et $D(y; \alpha)$ tend vers zéro.

$$\text{- Quand } \alpha \rightarrow +\infty \quad D(y; \alpha) = \alpha^{\frac{M-1}{2}} \left(M^M \sigma_H(y) \right)^\alpha \frac{H(M)}{\sigma_H(y)} \left(1 + o\left(\frac{1}{\alpha}\right) \right),$$

avec $M^M \sigma_H < 1$, et $D(y; \alpha)$ tend aussi vers zéro.

- On peut donc dresser le tableau de variation ci-joint pour tout point y intérieur au simplexe et différent de ω :

α	0	$\int_M \left(-\frac{1}{M} \log \sigma_H(y) \right)$	$+\infty$
$D(y; \alpha)$	0	$\nearrow \quad 0(\alpha(y)) \quad \searrow$	0

iv) Pour tout $\sigma \in]0, M^{-M}[$, on pose $\lambda(\sigma) = \rho_M(-\frac{1}{M} \log \sigma)$;
 et pour tout $\alpha \in \mathbb{R}_+^*$, $\mu(\alpha) = e^{-M J(\alpha, M)}$.

• Les propriétés précédentes entraînent immédiatement que les
 2 fonctions λ et μ sont définies continues dérivables croissantes
 sur leurs intervalles de définitions respectifs ; λ et μ sont deux appli-
 cations réciproques.

•• Les valeurs aux bornes sont :

- Quand $\alpha \rightarrow 0^+$, on a $\lim_{\alpha \rightarrow 0} J(\alpha, M) = +\infty$, $\lim_{\alpha \rightarrow 0} \mu(\alpha) = 0$;
- Quand $\alpha \rightarrow +\infty$, $\lim_{\alpha \rightarrow +\infty} J(\alpha, M) = \log M$, $\lim_{\alpha \rightarrow +\infty} \mu(\alpha) = M^{-M}$;
- On a $\mu'(\alpha) = -M J'(\alpha, M) \cdot \mu(\alpha)$ positif ;

sachant qu'au voisinage de zéro ψ' est équivalent à $1/u^2$,
 $\mu'(\alpha)$ est donc équivalent à $(M-1) e^{\frac{1-M}{\alpha}} / \alpha^2$ au voisinage de
 zéro, et par conséquent $\lim_{\alpha \rightarrow 0^+} \mu'(\alpha) = 0$.

v) Tous les points $y : y \neq \omega, y \notin S_1^M$ d'une même variété $\sigma(\sigma)$
 ont par définition la même valeur de la forme symétrique σ_M ;
 les variations en α de $D(\cdot ; \cdot)$ sont donc les mêmes pour tout y
 de $\sigma(\sigma)$.

A toute variété $\sigma(\tilde{\sigma})$ on fait correspondre l'exposant $\tilde{\alpha} = \lambda(\tilde{\sigma})$
 et la densité $D(\cdot ; \lambda(\tilde{\sigma}))$; on a vu que pour tous les \tilde{y} de $\sigma(\tilde{\sigma})$
 la fonction en α $D(\tilde{y} ; \cdot)$ atteint son maximum pour $\tilde{\alpha} = \lambda(\tilde{\sigma})$.

Réciproquement, à tout exposant $\tilde{\alpha}$, il correspond une variété
 unique $\sigma(\mu(\tilde{\alpha}))$.

- On définit
$$\theta(\tilde{\alpha}) = \frac{\Gamma(M \tilde{\alpha})}{\Gamma(\tilde{\alpha})^M} \mu(\tilde{\alpha})^{\tilde{\alpha}-1},$$

la valeur du maximum de $D(\cdot; \tilde{\alpha})$ en tout point de $\sigma(\mu(\tilde{\alpha}))$.

• Proposition 1

La fonction θ est définie continue positive dérivable sur \mathbb{R}_+^* , elle admet en $\alpha = 1$ son minimum égal à $\Gamma(M)$.

En effet
$$\frac{\theta'(\alpha)}{\theta(\alpha)} = M \zeta(\alpha, M) + \log \mu(\alpha) + (\alpha-1) \frac{\mu'(\alpha)}{\mu(\alpha)}$$

c'est à dire
$$\frac{\theta'(\alpha)}{\theta(\alpha)} = (\alpha-1) \frac{\mu'(\alpha)}{\mu(\alpha)} = M(1-\alpha) \zeta'(\alpha, M);$$

la dérivée θ' est continue, $\theta'(\alpha)$ est du signe de $(\alpha-1)$, θ est donc décroissante sur $]0, 1[$ et croissante au-delà; pour $\alpha = 1$ on a la densité de DIRICHLET constante et $\theta(1) = \Gamma(M)$. Ce minimum est atteint pour les points y de la variété $\sigma(\mu(1))$ avec $\sigma = \mu(1) = e^{-M(1 + \frac{1}{n-1})}$ ■

.. D'après une formule précédente (Réf. 80), on a pour les grandes valeurs de M

$$1 + \frac{1}{2} + \dots + \frac{1}{n-1} = \gamma + \log M - \frac{1}{2M} - \frac{1}{12M^2} + \frac{1}{n^2} O\left(\frac{1}{n}\right),$$

d'où

$$\mu(1) = \frac{e^{-M\gamma}}{M^M} e^{\left(\frac{1}{2} + \frac{1}{12M} + \frac{1}{n} O\left(\frac{1}{n}\right)\right)},$$

$$= M^{-n} e^{-n\gamma} e^{1/2} \left[1 + \frac{1}{12M} + \frac{1}{n} O\left(\frac{1}{n}\right) \right].$$

... Les valeurs aux bornes sont :

$$\text{- Quand } \alpha \rightarrow 0^+ , \quad \frac{\Gamma(n\alpha)}{\Gamma(\alpha)^M} = \frac{\alpha^{n-1}}{M} (1 + o(\alpha)) ,$$

$$\text{et } \mu(\alpha)^{\alpha-1} = e^{-n(\alpha-1) J(\alpha, n)} = e^{-\frac{n-1}{\alpha}} (1 + o(\alpha)) ,$$

$$\text{donc } \lim_{\alpha \rightarrow 0^+} \theta(\alpha) = +\infty .$$

$$\text{- quand } \alpha \rightarrow +\infty , \quad \frac{\Gamma(n\alpha)}{\Gamma(\alpha)^M} = M^{n\alpha} \alpha^{\frac{n-1}{2}} H(M) (1 + o(\frac{1}{\alpha})) ,$$

$$\text{donc } \lim_{\alpha \rightarrow +\infty} \theta(\alpha) = +\infty .$$

.... On a les variations suivantes, (figures ci-après) :

α	0	1	$+\infty$
$\mu(\alpha)$	0	$\mu(1)$	M^{-M}
$\theta(\alpha)$	$+\infty$	$\Gamma(M)$	$+\infty$
$d(\alpha)$	0	$\Gamma(M)$	$+\infty$

3.3. Exemple : La densité Béta .

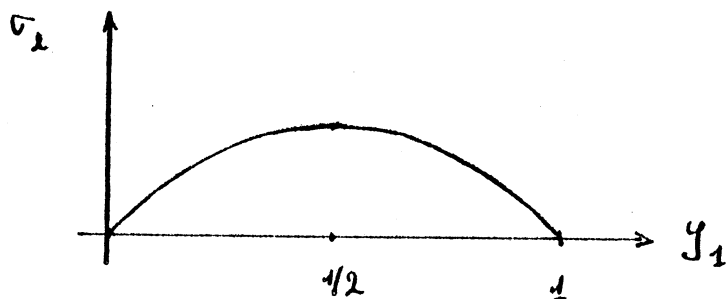
Le cas $M = 2$ correspond au cas particulier de la densité des lois Béta de 1^o espèce, et permet une illustration graphique simple.

Pour tout y_1 appartenant à $[0, 1]$, la valeur de la densité est égale à

$$D(y_1; \alpha) = \frac{\Gamma(\alpha)^2}{\Gamma(2\alpha)} (y_1(1-y_1))^{\alpha-1},$$

avec $\sigma_n(y) = y_1(1-y_1)$ compris entre 0 et $1/4$.

Toute variété $\sigma(\sigma)$ se réduit aux deux points d'abscisse $y_1 = \frac{1}{2} \pm \sqrt{\frac{1}{4} - \sigma}$ racines de l'équation $y_1(1-y_1) = \sigma$, $0 \leq \sigma \leq 1/4$.



i) Variation en α ; la formule de GAUSS se réduit ici à la formule de duplication de LEGENDRE, on a

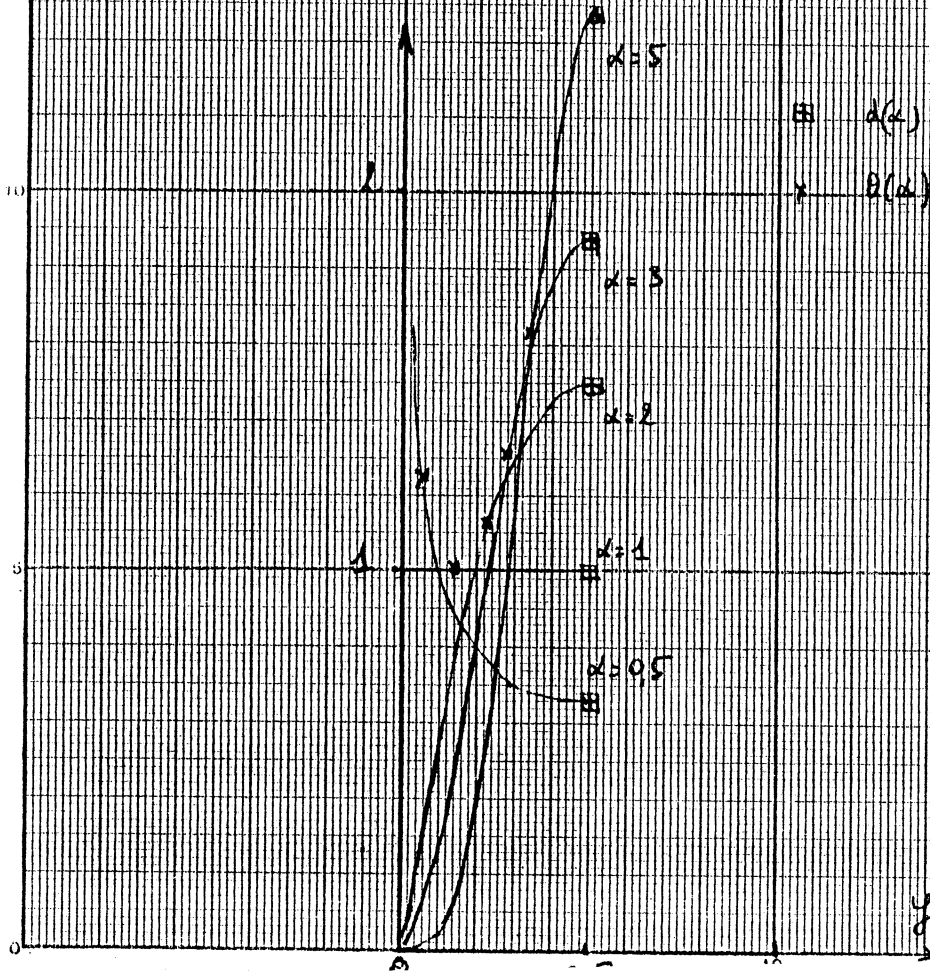
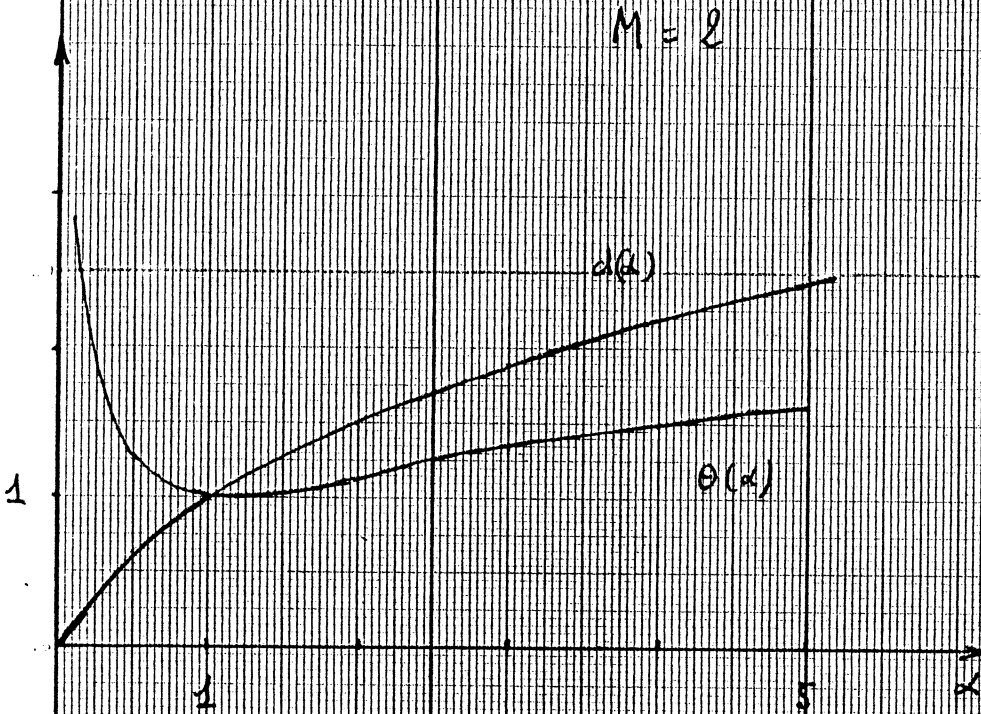
$$\begin{aligned} S(\alpha, 2) &= \Psi(2\alpha) - \Psi(\alpha) \\ &= \log 2 + \frac{1}{2} [\Psi(\alpha + \frac{1}{2}) - \Psi(\alpha)]. \end{aligned}$$

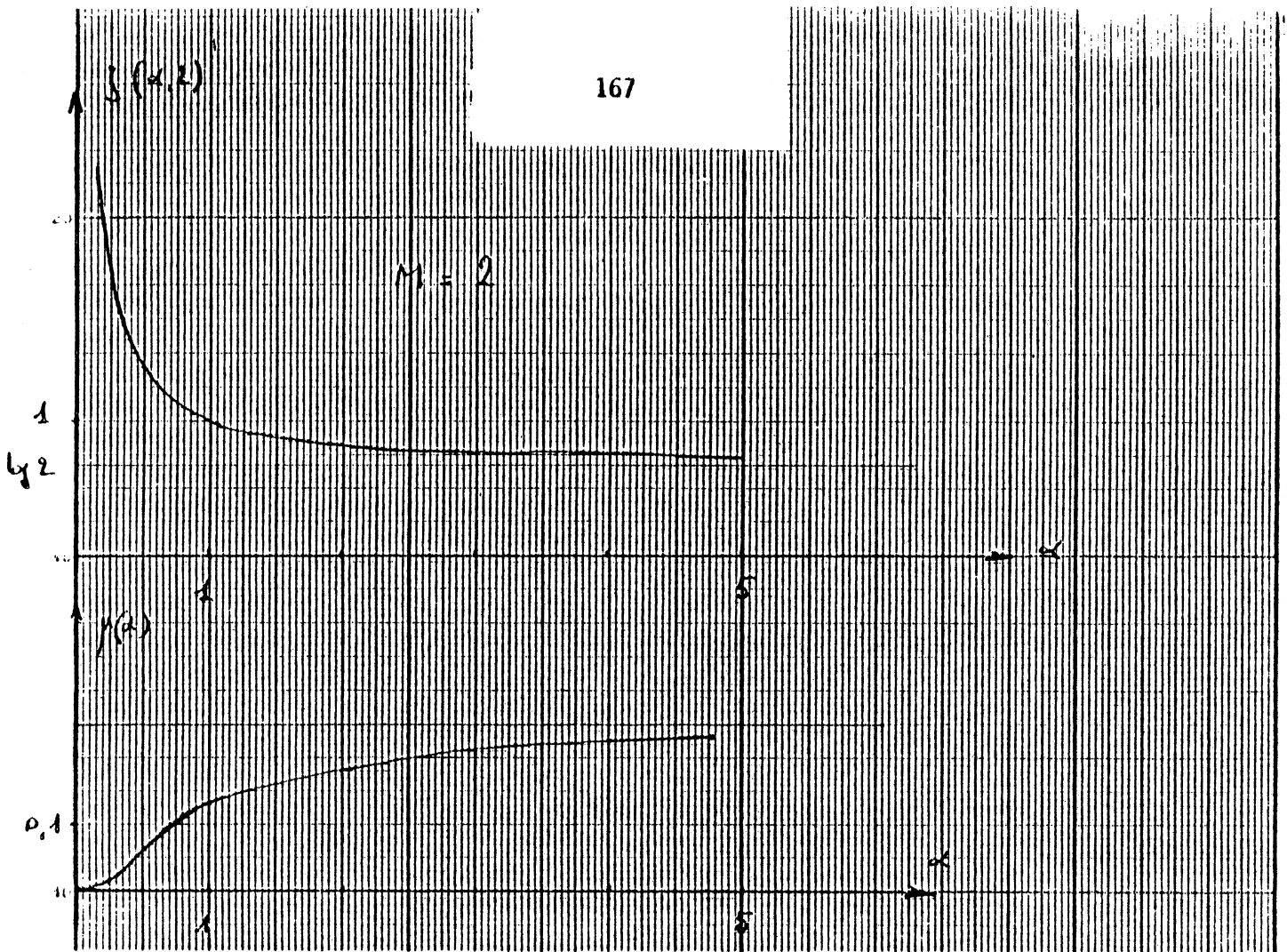
ii) On étudie comme ci-dessus μ , d , θ et

aussi $y = \frac{1}{2} - \sqrt{\frac{1}{4} - \mu}$ l'abscisse de l'un des deux points de $[0, 1]$ auxquels se réduisent les variétés $\sigma[\mu(\alpha)]$. On a le tableau suivant:

α	$S(\alpha, 2)$	$\mu(\alpha)$	$d(\alpha)$	$\theta(\alpha)$	$y(\alpha)$
0,5	1,386	0,062	0,636	1,270	0,067
1	1,000	0,135	1,000	1,000	0,161
2	0,833	0,188	1,500	1,128	0,250
3	0,783	0,209	1,875	1,307	0,297
5	0,745	0,225	2,460	1,468	0,340
\vdots					
$+\infty$	$\log 2 = 0,693$	0,25	$+\infty$	$+\infty$	0,5

(variations graphiques dans figures ci-contre)





VI.4 - La structure statistique invariante de DIRICHLET.

4.1. Définition .

On appelle structure statistique invariante (par le groupe $\bar{\Sigma}_M$) de DIRICHLET la structure échantillon dominée suivante :

$$\left(S_1^M, \mathcal{B}(S_1^M), \{ D(\cdot; \alpha), \alpha \in \mathbb{R}_+^* \} \right)^P$$

associée à la densité de DIRICHLET qui a pour valeur:

$$\forall y \in S_1^M, \quad D(y; \alpha) = \frac{\Gamma(M\alpha)}{\Gamma(\alpha)^M} \left(\prod_{i=1}^M y_i \right)^{\alpha-1}, \quad \alpha > 0.$$

Toutes les densités sont définies sur le simplexe unité pour $\alpha > 0$, dominées par exemple par la densité constante $D(\cdot; 1)$ qui correspond à la mesure uniforme de LEBESGUE sur le simplexe.

Ci-après on suppose que l'on a un échantillon de P points-partages y^p $p=1 \dots P$ indépendants qui suivent la distribution $D(\alpha)$.

4.2. Estimation de α .

i) Avec une probabilité égale à 1, la part y_i^p est positive pour tout $i=1 \dots M$ et pour tout $p=1 \dots P$. A toute observation de partage y^p on associe le produit $\sigma_n^p = \sigma_n(y^p) = \prod_{i=1}^M y_i^p$, et la moyenne géométrique de l'échantillon $\bar{\sigma}_n = (\prod_p \sigma_n^p)^{1/P}$.

Théorème 2 - La structure invariante de DIRICHLET est une structure exponentielle complète, la quantité $\text{Log } \bar{\sigma}_n$ est un estimateur sans biais de $-M \cdot \zeta(\alpha, n)$, de variance $M [\psi'(\alpha) - M \psi'(M\alpha)]/P$; de plus cet estimateur est de variance minimum parmi les estimateurs sans biais.

En effet α varie dans \mathbb{R}_+^* variété de dimension 1, on a immédiatement

$$\text{Log } D(y; \alpha) = (\alpha - 1) \text{Log } \sigma_n(y) + \text{Log} \left(\frac{\Gamma(n\alpha)}{\Gamma(\alpha)^M} \right);$$

On reconnaît la forme des distributions exponentielles (de FISHER - DARMOIS - KOOPMAN - PITMAN, (réf. 4, 5, 17, 29). On sait que la statistique

$T = \sum_p \text{Log } \sigma_n^p$ est une statistique exhaustive minimale, avec le paramètre naturel α ; on sait encore que $\text{Log } \bar{\sigma}_n = T/P$ est un estimateur efficace sans biais de variance minimum de

$$-\frac{d}{d\alpha} \text{Log} \frac{\Gamma(n\alpha)}{\Gamma(\alpha)^M} = -M \zeta(\alpha, n) = M [\psi(\alpha) - \psi(M\alpha)]$$

avec la variance

$$\begin{aligned} -\frac{1}{P} \frac{d^2}{d\alpha^2} \text{Log} \frac{\Gamma(n\alpha)}{\Gamma(\alpha)^M} &= -\frac{M}{P} \zeta'(\alpha, n) \\ &= \frac{M}{P} [\psi'(\alpha) - M \psi'(M\alpha)]. \end{aligned}$$

On en déduit aussitôt que $-\frac{1}{M} \log \bar{\sigma}_n$ est un estimateur efficace sans biais de $J(\alpha, M)$ avec la variance $-\frac{1}{MP} J'(\alpha, M)$.

ii) La vraisemblance de l'échantillon des y^P s'écrit

$$L(y^P; \alpha) = (\sigma_n^1 \dots \sigma_n^P)^{\alpha-1} \left(\frac{\Gamma(n\alpha)}{\Gamma(\alpha)^M} \right)^P$$

d'où

$$\log L = (\alpha-1) \sum_P \log \sigma_n^P + P \log \frac{\Gamma(n\alpha)}{\Gamma(\alpha)^M}$$

$$\frac{\partial \log L}{\partial \alpha} = P \left[\log \bar{\sigma}_n + M \cdot J(\alpha, M) \right]$$

L'équation du maximum de vraisemblance possède une racine unique

$$\hat{\alpha} = \mathcal{J}_M \left(-\frac{1}{M} \log \bar{\sigma}_n \right)$$

D'après ce qui précède, on vérifie que $L(y^P; \cdot)$ est maximale en $\hat{\alpha}$. L'application \mathcal{J}_M étant strictement décroissante, $\hat{\alpha}$ est lui-même un estimateur exhaustif, fonction monotone de la statistique exhaustive $\log \bar{\sigma}_n$.

On sait encore que $\hat{\alpha}$ est asymptotiquement efficace et normal avec la variance

$$-\frac{1}{P} E \left(\frac{\partial^2 \log L}{\partial \alpha^2} \right)$$

c'est à dire $-\frac{1}{MP} J'(\alpha, M)$.

Remarque : la distribution $D(\alpha)$ invariante par Σ_M se rencontre dans un certain contexte Bayésien. A cet effet on trouve dans GOOD (Réf. 33) et COX et HINCKLEY (Réf. 17) par exemple divers procédés chargés d'évaluer le paramètre α ; procédés qui naturellement sont différents de l'estimation ci-contre.

4.3. Applications numériques .

La résolution de l'équation de vraisemblance passe par la résolution en de

$$\Psi(\alpha n) - \Psi(\alpha) = -\frac{1}{MP} \sum_{i,p} \log y_i^p .$$

On peut le faire numériquement à l'aide de la méthode classique de la corde sécante et des tangentes de NEWTON (Réf. 3).

On a besoin pour cela de calculer pour tout α positif les expressions $\Psi(\alpha)$, $\Psi(n\alpha)$, $\Psi'(\alpha)$, $\Psi'(n\alpha)$. Nous avons retenu les expressions suivantes : (Réf. 1),

$$\left\{ \begin{array}{l} \Psi(3) = \Psi(3+n) - \sum_{k=0}^{n-1} \frac{1}{3+k} \\ \Psi(u) = \log u - \frac{1}{2u} - \frac{1}{12u^2} + \frac{1}{120u^4} - \frac{1}{252u^6} + \frac{1}{240u^8} - \frac{1}{132u^{10}} \end{array} \right.$$

$$\left\{ \begin{array}{l} \Psi'(3) = \Psi'(3+n) + \sum_{k=0}^{n-1} \frac{1}{(3+k)^2} \\ \Psi'(u) = \frac{1}{u} + \frac{1}{2u^2} + \frac{1}{6u^3} - \frac{1}{30u^5} + \frac{1}{42u^7} - \frac{1}{30u^9} + \frac{5}{66u^{11}} \end{array} \right.$$

On peut prendre par exemple n de l'ordre de la dizaine ; un programme écrit en FORTRAN extrait la racine $\hat{\alpha}$ du système de vraisemblance avec une précision en 10^{-6} en 5 ou 6 itérations.

4.4. Test sur α .

4.4.1. Par construction la structure statistique invariante de DIRICHLET est une structure exponentielle ; il en résulte les propriétés habituelles relatives aux tests UMP, UMPB sur le paramètre naturel α , (Réf. 5, 29, 51).

On ramène la recherche des régions critiques et des régions d'acceptation sur l'espace décrit par une statistique exhaustive. Il est commode ici de prendre la statistique $\bar{\sigma}_M$ liée à l'estimateur exhaustif $\log \bar{\sigma}_n$, ou à $\hat{\alpha} = \rho_M \left(-\frac{1}{M} \log \bar{\sigma}_n \right)$.

On en déduit les tests suivants (Ref. 5) :

i) Quels que soient $\eta_1 \in]0, 1[$ et $\alpha_0 > 0$, le test défini par la région critique $\bar{\sigma}_n \geq c_1$ où c_1 vérifie :

$$\eta_1 = \int_{V(c_1)} D(y; \alpha_0) dy$$

est U.M.P. de niveau η_1 comme test de l'hypothèse $\alpha \leq \alpha_0$ contre l'hypothèse $\alpha > \alpha_0$.

ii) Quels que soient $\eta_2 \in]0, 1[$ et $\alpha_0 > 0$, le test défini par la région critique $\bar{\sigma}_n \leq c_2$ où c_2 vérifie :

$$1 - \eta_2 = \int_{V(c_2)} D(y; \alpha_0) dy$$

est U.M.P. de niveau η_2 comme test de l'hypothèse $\alpha \geq \alpha_0$ contre l'hypothèse $\alpha < \alpha_0$.

iii) Quels que soient η_1 et η_2 positifs de somme inférieure à 1, et $\alpha_0 > 0$, le test défini par la région critique $(\bar{\sigma}_n \geq c_1) \cup (\bar{\sigma}_n \leq c_2)$ où c_1 et c_2 vérifient les relations ci-dessus est U.M.P.B. de niveau $\eta_1 + \eta_2$ comme test de l'hypothèse $\alpha = \alpha_0$ contre l'hypothèse $\alpha \neq \alpha_0$.

4.4.2. L'application de ces tests repose sur le calcul des intégrales :

$$I(\alpha, M, \sigma) = \int_{V(\sigma)} D(y; \alpha) dy$$

définies sur S_1^M pour $\alpha > 0$, $M \geq 2$ et $0 \leq \sigma \leq M^{-M}$.

On a en particulier $I(\alpha, M, 0) = 1$ et $I(\alpha, M, M^{-M}) = 0$. La puissance de ces tests s'exprime également avec les intégrales ci-dessus. Dans le cas particulier $\alpha_0 = 1$, la densité $D(\cdot; \alpha_0)$ se réduit à la constante $\Gamma(M)$, le calcul des $I(1, M, \sigma)$ est ici le calcul de la mesure de LEBESGUE des domaines $V(\sigma)$.

La fonction de répartition de la statistique $\sigma_M = \prod_{i=1}^M y_i$ est
 $1 - I(\alpha, M, \sigma)$; on pose $f(\alpha, n, \sigma)$ la densité $-\frac{\partial I}{\partial \alpha}$, (cf. chap X.1)

4.5. Estimation ensembliste de α .

Un résultat classique de la statistique relie l'estimation ensembliste d'un paramètre aux tests relatifs à ce paramètre. Toute valeur $\tilde{\alpha}$ qui n'est pas rejetée dans un test contre α fait partie de l'intervalle de confiance de α (pour le même seuil que le test). Dans le cas présent l'intervalle $[\lambda(c_1), \lambda(c_2)]$ est une estimation ensembliste de α avec le seuil de confiance $1 - \pi = I(\alpha, n, c_2) - I(\alpha, n, c_1)$.

Sous sa forme générale ce résultat est par exemple érigé en théorème dans BARRA (Réf. 5, VI.3 théorème 2) ; pour leur part COX et HINCKLEY mentionnent que c'est une "bonne" règle d'estimation ensembliste (Réf. 17, chapitre 7), la nuance entre les deux "is partly a matter of taste".

VI.5 - Exemple, simulations de partages aléatoires.

On simule sur le simplexe unité des partages qui suivent une distribution invariante de DIRICHLET. On applique les calculs d'estimation précédents sur un échantillon de tels partages.

5.1. La première simulation que l'on envisage consiste à prendre $M-1$ tirages indépendants de la v.a. uniforme sur $[0, 1]$. Une fois ordonnées ces $M-1$ valeurs sont les bornes de M segments dont la longueur est égale à 1. On a vu que l'on obtient un partage aléatoire qui suit la distribution uniforme de densité $D(\cdot; 1)$, (1^o partie chap. II.3).

La simulation de la variable uniforme sur $[0, 1]$ est faite avec le sous-programme ALEAT de la bibliothèque CII. On a fait varier M et P dans les divers essais. Pour chacun de ceux-ci on indique l'estimateur $\hat{\alpha}$ du maximum de vraisemblance, l'estimation $\hat{\sigma}_\alpha^2 = -1/M P f(\alpha, n)$ de sa variance et l'intervalle de confiance $[\alpha_1, \alpha_2]$ au seuil de 95% en supposant que $\hat{\alpha}$ suit une loi normale.

Les résultats sont :

simulation		$\hat{\alpha}$	$\hat{\sigma}_{\alpha}^2$	α_1	α_2
M = 5	P = 20	1,056	0,0209	0,772	1,340
	P = 100	0,977	0,0035	0,861	1,094
M = 10	P = 20	1,066	0,0097	0,873	1,259
	P = 100	1,008	0,0017	0,927	1,090

Sur ces résultats la valeur réelle (et connue, $\alpha = 1$) du paramètre appartient toujours à l'intervalle $[\alpha_1, \alpha_2]$, l'estimation de $\hat{\sigma}_{\alpha}^2$ décroît avec la taille de l'échantillon P et avec le nombre de parts M .

5.2. La deuxième simulation utilise la normalisation par la somme de M v.a indépendantes qui suivent toute la même loi gamma $\Gamma(\alpha, \theta)$. Les partages aléatoires suivent une distribution invariante (1° partie, chap. IV.3 et IV.4). A titre d'exemple les simulations sont ici faites pour $M = 5$ et 10 , $P = 20$ et 100 , et les paramètres $\alpha = 1$ (loi uniforme), et $\alpha = 3$.

5.2.1. Simulations avec lois exponentielles ($\alpha = 1$) :

simulation		$\hat{\alpha}$	$\hat{\sigma}_{\alpha}^2$	α_1	α_2
M = 5	P = 20	1,117	0,0236	0,815	1,419
	P = 100	1,003	0,0037	0,883	1,123
M = 10	P = 20	1,023	0,0089	0,839	1,208
	P = 100	1,010	0,0017	0,928	1,091

Les estimations $\hat{\sigma}_{\alpha}^2$ sont du même ordre de grandeur qu'avec le découpage uniforme en 5.1.

5.2.2. Simulations avec lois $\Gamma(3, \theta)$, (θ réel positif):

simulation		$\hat{\alpha}$	$\hat{\sigma}_{\alpha}^2$	α_1	α_2
M = 5	P = 20	2,490	0,1341	1,772	3,207
	P = 100	3,106	0,0428	2,700	3,512
M = 10	P = 20	2,616	0,0669	2,0184	3,1221
	P = 100	3,044	0,0184	2,7781	3,3101

On retrouve les mêmes conclusions pour la décroissance de $\hat{\sigma}_{\alpha}^2$ en fonction de M et P ; en outre cette estimation augmente lorsque α prend les valeurs 1 puis 3.

VII - LA STRUCTURE GÉNÉRALE DE DIRICHLET

La distribution (générale) de DIRICHLET est également définie sur le simplexe unité de \mathbb{R}^M . Pour tous points α de \mathbb{R}_+^M et y du simplexe, de coordonnées α_i , resp. y_i , $i=1, \dots, M$ elle est définie par la densité de probabilité de DIRICHLET qui a pour valeur, (ref. 84)

$$D(y_1, \dots, y_M; \alpha_1, \dots, \alpha_M) = \frac{\Gamma(\sum_{i=1}^M \alpha_i)}{\prod_{i=1}^M \Gamma(\alpha_i)} \prod_{i=1}^M y_i^{\alpha_i - 1}$$

On note $D(\alpha)$ la distribution, $D(\cdot; \alpha)$ la densité et $D(y; \alpha)$ sa valeur.

VII.1 - Etude de la densité.

Le produit de facteurs $g(y) = \prod_{i=1}^M y_i^{\alpha_i - 1}$ est défini continu dérivable en tout point intérieur au simplexe unité S_1^M . Il s'annule

(et est défini) pour $y_i = 0$ si et seulement si $\alpha_i \geq 1$. Il prend la valeur infinie (par continuité) en $y_i = 0$ pour $\alpha_i < 1$. C'est une expression algébrique pour α_i entier.

1.1. Variation en fonction de y .

On a
$$\frac{dg}{g} = \sum_{i=1}^M \frac{\alpha_i - 1}{y_i} dy_i,$$

Soit
$$\frac{\partial g}{\partial y_j} = g(y) \cdot \left(\frac{\alpha_j - 1}{y_j} - \frac{\alpha_n - 1}{y_n} \right) \quad j=1, \dots, n-1.$$

On a les conditions nécessaires d'un extrémum si $\frac{\alpha_i - 1}{y_i}$ est le même pour tout

i , d'où la solution $\bar{y}_i = \frac{\alpha_i - 1}{\sum \alpha_k - M}$, solution acceptable si et seule-

ment si pour tout i , $\bar{y}_i \in [0, 1]$. Ceci est vérifié si pour tout k , α_k

est supérieur à 1, ou si pour tout k , α_k est compris entre 0 et 1 (au sens strict).

Si pour tout $k=1, \dots, M$, $\alpha_k > 1$, on a évidemment un maximum; comme le souligne

BARDNORFF-NIELSEN g est fortement unimodale (réf. 4) c'est-à-dire que

$\log g$ est convexe; si pour tout k , $0 < \alpha_k < 1$, $1/g$ est fortement unimodale.

1.2. Variation en fonction de α .

Nous nous limitons au cas $\alpha > \theta$ avec λ réel positif, et θ un point

i) Proposition 2 Pour tout point θ de l'intérieur du simplexe unité, l'application qui à tout λ réel positif fait correspondre $D(y; \lambda \theta)$ est: -croissante puis décroissante pour tout y de \bar{S}_1^M différent de θ ; -croissante pour $y = \theta$.

En effet on a
$$D(y; \lambda \theta) = \frac{\Gamma(\lambda)}{\prod \Gamma(\theta_i \lambda)} \cdot \left(\prod y_j^{\theta_j \lambda} \right) \cdot \left(\prod y_j^{\theta_j} \right)^{\lambda-1}$$

d'où
$$D'_\lambda(y; \lambda \theta) = D(y; \lambda \theta) \cdot \left[\Psi(\lambda) - \sum \theta_i \Psi(\theta_i \lambda) + \sum \theta_j \log y_j \right]$$

qui a le même signe que
$$\sum \theta_i [\Psi(\lambda) - \Psi(\theta_i \lambda)] + \sum \theta_j \log y_j.$$

On démontrera en VII.2.2 que l'application qui à tout λ positif fait correspondre $\Psi(\lambda) - \Psi(\theta_i \lambda)$ est positive décroissante avec les valeurs limites:

$$\lim_{\lambda \rightarrow 0^+} \Psi(\lambda) - \Psi(\theta_i \lambda) = +\infty, \quad \lim_{\lambda \rightarrow +\infty} \Psi(\lambda) - \Psi(\theta_i \lambda) = -\log \theta_i.$$

On rappelle par ailleurs l'inégalité classique de la théorie de l'information $\sum \theta_i \log y_i < \sum \theta_j \log \theta_j$ (ref. 49,68), avec égalité si et seulement si $\theta_i = y_i$ pour tout $i = 1 \dots M$. Par conséquent pour tout y différent de θ , il existe une racine réelle unique finie $\bar{\lambda}_y$ de l'équation $D'_\lambda(y; \lambda \theta) = 0$, donc D'_λ est positive sur $]0, \bar{\lambda}_y[$ et négative sur $]\bar{\lambda}_y, +\infty[$; au contraire pour $y = \theta$, D'_λ est uniquement positive sur \mathbb{R}_+^* . ■

ii) Les valeurs aux bornes sont:

- Quand $\lambda \rightarrow 0^+$ on a
$$\Gamma(\theta_i \lambda) = \frac{1}{\lambda^{\theta_i}} (1 + o(\lambda)),$$

d'où
$$\frac{\Gamma(\lambda)}{\prod \Gamma(\theta_i \lambda)} = \lambda^{\lambda-1} \prod \theta_i \cdot (1 + o(\lambda)),$$

donc
$$\lim_{\lambda \rightarrow 0^+} D(y; \lambda \theta) = 0,$$

pour tout point y à l'intérieur du simplexe unité.

- Quand $\lambda \rightarrow +\infty$, on applique la formule de STIRLING

$$\Gamma(\lambda \theta_i) = (\lambda \theta_i)^{\lambda \theta_i} e^{-\lambda \theta_i} \left(\frac{2\pi}{\lambda \theta_i} \right)^{1/2} \cdot \left(1 + o\left(\frac{1}{\lambda}\right) \right),$$

$$\frac{\Gamma(\lambda)}{\prod \Gamma(\lambda \theta_i)} = \left(\frac{\lambda}{2\pi} \right)^{\frac{\lambda-1}{2}} \cdot \left(\prod \theta_i \right)^{\lambda-1} \cdot \frac{1}{\left(\prod \theta_j^{\theta_j} \right)^\lambda} \left(1 + o\left(\frac{1}{\lambda}\right) \right),$$

$$\text{et } D(y; \lambda \theta) = \left(\frac{\lambda}{\lambda n}\right)^{\frac{n-1}{2}} \left(\frac{\prod_{j=1}^n y_j^{\theta_j}}{\prod_{j=1}^n \theta_j^{\theta_j}}\right)^{\lambda} \cdot \frac{(\prod_{j=1}^n \theta_j)}{\prod_{j=1}^n y_j} \cdot \left(1 + o\left(\frac{1}{\lambda}\right)\right).$$

$$\text{- au point } y = \theta \text{ on a } D(\theta; \lambda \theta) = \left(\frac{\lambda}{\lambda n}\right)^{\frac{n-1}{2}} \frac{1}{(\prod_{j=1}^n \theta_j)^{1/2}} \left(1 + o\left(\frac{1}{\lambda}\right)\right)$$

et par conséquent

$$\lim_{\lambda \rightarrow +\infty} D(\theta; \lambda \theta) = +\infty$$

$$\text{- en tout point } y \text{ différent de } \theta \text{ on a } \frac{\prod_{j=1}^n y_j^{\theta_j}}{\prod_{j=1}^n \theta_j^{\theta_j}} < 1$$

et par conséquent

$$\lim_{\lambda \rightarrow +\infty} D(y; \lambda \theta) = 0.$$

iii) Remarque: la valeur de λ qui annule la dérivée $\frac{\partial D(y; \lambda \theta)}{\partial \lambda}$ est la même pour tous les points de la variété $\{y \in S_1^n; \prod_{j=1}^n y_j^{\theta_j} = \eta\}$. En tous ces points la fonction $D(\cdot; \cdot)$ est simultanément croissante ou décroissante en λ .

1.3. Courbes de niveau.

- Pour tout i on suppose que a_i est supérieur ou égal à 1.
Le maximum de $D(\cdot; a)$ est égal dans ce cas à :

$$\frac{\Gamma(A)}{\prod_{j=1}^n \Gamma(a_j)} \cdot \frac{1}{(A-n)^{A-n}} \cdot \prod_{j=1}^n (a_j - 1)^{a_j - 1}$$

(On désigne par A la somme des a_i).

On définit les ensembles :

$$\sigma_{a-\mathbb{1}, \eta} = \left\{ y \in S_1^n : \prod_{j=1}^n y_j^{a_j-1} = \eta \right\},$$

$$V_{a-\mathbb{1}, \eta} = \left\{ y \in S_1^n : \prod_{j=1}^n y_j^{a_j-1} \geq \eta \right\},$$

(On désigne par $\mathbb{1}$ le vecteur de coordonnées égales à 1).

Ce sont deux ensembles non vides pour

$$0 \leq \eta \leq \eta_A = \frac{\prod_{i=1}^M (a_i - 1)^{a_i - 1}}{(A - M)^{A - M}}$$

et on vérifie que $V_{a-1, \eta}$ est convexe, en particulier $V_{0, \eta}$ est convexe.

- Si pour tout i $a_i \leq 1$, $D(\cdot; a)$ présente un minimum qui a la même expression que précédemment. On définit de même les "courbes de niveau" et l'ensemble de niveau :

$$\bar{V}_{a-1, \eta} = \left\{ y \in S_1^M : \prod_{i=1}^M y_i^{a_i - 1} \leq \eta \right\},$$

qui est encore convexe.

- Si pour tout i $a_i > 0$, on définit également v et V comme ci-dessus, V n'a plus de propriétés de convexité.

1.4. Rapport de deux densités.

On étudie le rapport $D(\cdot; a) / D(\cdot; b)$ qui varie comme le produit $\prod_{i=1}^M y_i^{a_i - b_i}$; on se limite au cas où $a_i \geq b_i$ pour tout i .

Soient les sous-ensembles :

$$w_{a,b,\zeta} = \left\{ y \in S_1^M : \frac{D(y; a)}{D(y; b)} = \zeta \right\},$$

$$W_{a,b,\zeta} = \left\{ y \in S_1^M : \frac{D(y; a)}{D(y; b)} \geq \zeta \right\}.$$

Proposition 3 Les ensembles $w_{a,b,\zeta}$ et $W_{a,b,\zeta}$ sont non vides pour

$$0 \leq \zeta \leq \zeta_A = \frac{\Gamma(A)}{(A-B)^{A-B} \Gamma(B)} \times \prod_{i=1}^M \frac{(a_i - b_i)^{a_i - b_i} \Gamma(b_i)}{\Gamma(a_i)},$$

en outre $W_{a,b,\zeta}$ est convexe.

En effet on a

$$\frac{D(y; a)}{D(y; b)} = \frac{\Gamma(A)}{\Gamma(A+B) \Gamma(B)} \prod_{i=1}^M \frac{\Gamma(a_i+1-b_i) \Gamma(b_i)}{\Gamma(a_i)} \cdot D(y; a+1-b)$$

(avec $A = \sum_{i=1}^M a_i$, $B = \sum_{i=1}^M b_i$, $\mathbb{1}$ le vecteur de coordonnées 1).

Par hypothèse les paramètres a_i+1-b_i sont supérieurs ou égaux à 1, on en déduit la valeur de ζ_0 d'après VII 1.3, et aussi que

$\omega_{a,b,\zeta}$ (resp $W_{a,b,\zeta}$) est égal à $\nu_{a-b,\gamma}$ (resp $V_{a-b,\gamma}$)

avec

$$\gamma = \int \frac{\Gamma(B)}{\Gamma(A)} \prod_{i=1}^M \frac{\Gamma(a_i)}{\Gamma(b_i)} \cdot \quad \blacksquare$$

1.5. Dispositions mutuelles des modes.

On suppose ci-dessous que $a_i \geq b_i \geq 1$ pour tout $i=1 \dots M$.

Soient μ_a , μ_b , μ_{a+b} les points modaux des densités

$D(\cdot; a)$, $D(\cdot; b)$ et $D(\cdot; a+1-b)$ de coordonnées respectives

$$\frac{a_i-1}{A-M}, \quad \frac{b_i-1}{B-M}, \quad \frac{a_i-b_i}{A-B}, \quad i=1 \dots M.$$

Définition. Soient J, K, N trois sous-ensembles qui constituent une partition de l'ensemble $\{1 \dots M\}$, J et K simultanément vides ou non. On définit la relation binaire \mathcal{R}_{JKN} par

$$y \mathcal{R}_{JKN} z \Leftrightarrow \begin{cases} y_j > z_j & j \in J \\ y_k < z_k & k \in K \\ y_n = z_n & n \in N \end{cases}$$

Proposition 4 : J et K étant non vides \mathcal{R}_{JKN} est une relation d'ordre partiel strict (Réf. 70) ; $\mathcal{R}_{\emptyset\emptyset N}$ est la relation identité.

L'identité est triviale quand J et K sont vides, dans le cas contraire on vérifie simplement la transitivité et l'anti-symétrie. ■

Théorème 3 : On définit les ensembles

$$J = \left\{ j : \frac{b_j - 1}{B - M} > \frac{a_j - 1}{A - M} \right\}$$

$$K = \left\{ k : \frac{b_k - 1}{B - M} < \frac{a_k - 1}{A - M} \right\}$$

$$N = \left\{ n : \frac{b_n - 1}{B - M} = \frac{a_n - 1}{A - M} \right\}$$

tels que $\mathcal{M}_a \mathcal{R}_{JKN} \mathcal{M}_b$ par construction.

Alors \mathcal{M}_a est intermédiaire entre \mathcal{M}_{ab} et \mathcal{M}_b pour \mathcal{R}_{JKN} , (à la limite les 3 modes sont confondus).

En effet, J et K sont simultanément vides (condition nécessaire entre points sur simplexe unité). Il suffit d'appliquer la relation classique pour tous b et d positifs, a et c positifs ou nuls :

$$\frac{a}{b} \leq \frac{c}{d} \Leftrightarrow \frac{a}{b} \leq \frac{a+c}{b+d} \leq \frac{c}{d} \quad (\text{égalités simultanées}).$$

qui s'écrit ici

$$\frac{a_j - b_j}{A - B} < \frac{a_j - 1}{A - M} < \frac{b_j - 1}{B - M}$$

$$\frac{b_k - 1}{B - M} < \frac{a_k - 1}{A - M} < \frac{a_k - b_k}{A - B}$$

d'où

$$\mathcal{M}_{ab} \mathcal{R}_{JKN} \mathcal{M}_a \mathcal{R}_{JKN} \mathcal{M}_b \quad \blacksquare$$

VII.2 - La structure statistique de DIRICHLET.

Définition: On appelle structure statistique de DIRICHLET la structure suivante, associée à la distribution $D(a)$

$$\left(S_1^M, \mathcal{B}(S_1^M), \{ D(a), a \in \mathbb{R}_+^{*M} \} \right)^P.$$

C'est une structure dominée par la densité constante sur le simplexe unité.

2.1. Premiers résultats.

2.1.1. Etude du vecteur aléatoire $\partial \log D / \partial a$.

Soient $\frac{\partial \log D}{\partial a_i} = \log y_i + \Psi(\sum a_k) - \Psi(a_i)$ pour tout $i=1, \dots, M$;

les coordonnées d'un point aléatoire tel que les y_i suivent la distribution de DIRICHLET $D(a)$. On rappelle que $\partial \log D / \partial a$ est centré.

Par ailleurs les coordonnées $\partial \log D / \partial a_i$ sont linéairement indépendantes, d'où il résulte que la matrice de variance-covariance est définie positive. Sachant

$$\text{que: } \frac{\partial^2}{\partial a_i^2} \int D \, dy = \int \frac{\partial^2}{\partial a_i^2} \log D \cdot D \, dy + \int \left(\frac{\partial}{\partial a_i} \log D \right)^2 \cdot D \, dy = 0,$$

$$\frac{\partial^2}{\partial a_i \partial a_j} \int D \, dy = \int \frac{\partial^2}{\partial a_i \partial a_j} \log D \cdot D \, dy + \int \left(\frac{\partial}{\partial a_i} \log D \right) \left(\frac{\partial}{\partial a_j} \log D \right) \cdot D \, dy = 0,$$

la matrice de variance-covariance de $\frac{\partial \log D}{\partial a}$ est la suivante:

$$\Gamma_{ij} = - \int \frac{\partial^2 \log D}{\partial a_i \partial a_j} \cdot D \, dy = - \Psi'(\sum a_k),$$

$$\Gamma_{ii} = - \int \frac{\partial^2 \log D}{\partial a_i \partial a_i} \cdot D \, dy = \Psi'(a_i) - \Psi'(\sum a_k),$$

soit encore: $\Gamma_{ij} = \Psi'(a_i) \delta_{ij} - \Psi'(\sum a_k)$.

C'est la matrice d'information de FISHER, son déterminant est positif pour tout point $a \in \mathbb{R}_+^{*M}$ (ref. 5, 17, 29).

Conséquences analytiques:

- Appliqué au cas $M = 2$, on en déduit l'inégalité suivante :

$$\frac{1}{\Psi'(\alpha)} + \frac{1}{\Psi'(\beta)} < \frac{1}{\Psi'(\alpha + \beta)}$$

pour tous α, β réels positifs, (cf. le signe de $\det(\Gamma_\beta)$).

- La relation $\int \frac{\partial \log D}{\partial a_i} D dy = 0$ s'écrit

$$-\frac{\Gamma(\sum a_i)}{\prod \Gamma(a_i)} \int \log y_i \cdot \prod y_j^{a_j-1} dy_1 \dots dy_{n-1} = \psi(\sum a_i) - \psi(a_i),$$

et dans le cas $M = 2$:

$$\frac{-1}{\Gamma(\alpha, \beta)} \int_0^1 \log y \cdot y^{\alpha-1} (1-y)^{\beta-1} dy = \psi(\alpha+\beta) - \psi(\alpha).$$

2.1.2. Propriétés de la densité de DIRICHLET.

Nous rappelons ici d'autres propriétés de cette densité qui se traduisent facilement en termes de statistique.

- Soient $Q = \{1, \dots, q\}$ et $R = \{s+1, \dots, M\}$ une partition de l'ensemble $\{1, \dots, M\}$, pour tout $s = 1, \dots, M-2$.

On pose $z_q = y_q \quad q \in Q, \quad z_{s+1} = \sum_{s+1}^M y_r, \quad z \in S_1^{s+1},$

$$u_r = y_r / z_{s+1} \quad r \in R, \quad u \in S_1^{M-s-1},$$

$$A_n = \sum_{i=1}^M a_i$$

Proposition 5 : La densité marginale de $D(\cdot; a)$ en $y_1 \dots y_s$ est la densité de DIRICHLET $D(\cdot; a_q, A_{s+1})$.

Démonstration dans WILKS, (référence 84). ■

Par conséquent la distribution conditionnelle de u par rapport à z est la distribution de DIRICHLET $D(\cdot; a_r)$, comme le montre la factorisation suivante:

$$\frac{\Gamma(A_1)}{\prod_{i=1}^M \Gamma(a_i)} \prod_{i=1}^M y_i^{a_i-1} = \frac{\Gamma(A_1)}{\prod_{i=1}^q \Gamma(a_i) \cdot \Gamma(A_{s+1})} \prod_{i=1}^q y_i^{a_i-1} \times \left(\sum_{s+1}^M y_r \right)^{A_{s+1}-1} \\ \times \frac{\Gamma(A_{s+1})}{\prod_{r=s+1}^M \Gamma(a_r)} \prod_{r=s+1}^M \left(\frac{y_r}{\sum y_r} \right)^{a_r-1} \cdot \left(1 / \sum y_r \right)^{M-s-1}$$

c'est à dire :

$$D(y_1, y_2; a_q, a_r) = D(y_1; a_q, A_{s+1}) \times D(u_r / y_q; a_r).$$

Ci-dessous (2.1.4.) on remarque que la structure de DIRICHLET est une structure exponentielle (ref. 4, 5, 17). Sur le plan statistique on retrouve ici directement le fait général que la distribution des y_2 ou u_2 conditionnés par les y_q ou z_q est elle-même une loi de type exponentiel, (ref. 4, 74). que les y_2 ou u_2 conditionnées par y_q ou z_q sont ancillaires (libres) pour les a_q , et que les y_q ou z_q ne sont pas ancillaires pour les a_r .

Proposition 6: Soient \mathcal{Q}_ℓ $\ell=1, \dots, L$ une partition de $\{1, \dots, M\}$, avec i_ℓ l'indice courant de \mathcal{Q}_ℓ , soit \mathcal{Z} le point de coordonnées $\mathcal{Z}_\ell = \sum y_{i_\ell}$ du simplexe unité S_1^L , et \mathcal{b} le point de coordonnées $\mathcal{b}_\ell = \sum a_{i_\ell}$. L'image de la loi de DIRICHLET $D(\mathcal{a})$ par l'application qui à tout \mathcal{y} fait correspondre \mathcal{Z} est la loi de DIRICHLET $D(\mathcal{b})$

Démonstration dans WILKS, (référence 84). ■

2.1.3. Statistiques d'échantillon.

On pose $A = \sum_1^M a_k$; on rappelle que l'espérance mathématique et la matrice de variance-covariance de la distribution $D(\mathcal{a})$ sont respectivement le vecteur de coordonnées a_i/A et la matrice d'éléments $\frac{a_i}{A} (\delta_{ij} - \frac{a_j}{A}) / A+1$ (ref. 84).

Lorsque l'on a un échantillon de P observations de partages indépendantes \mathcal{y}^p qui suivent la loi $D(\mathcal{a})$, on rappelle que les statistiques d'échantillon $\bar{\mathcal{y}} = \frac{1}{P} \sum_p \mathcal{y}^p$ et $\frac{1}{P-1} \sum_p (\mathcal{y}^p - \bar{\mathcal{y}})(\mathcal{y}^p - \bar{\mathcal{y}})'$ sont des estimateurs sans biais de l'espérance mathématique et de la variance-covariance de la distribution $D(\mathcal{a})$.

2.1.4. Structure exponentielle.

Théorème 4: La structure statistique de DIRICHLET est une structure exponentielle complète. Elle possède la statistique exhaustive (vectorielle)

$$\mathcal{Q}_i = \sum_{i=1}^M \log y_i^p$$

avec les paramètres naturels a_i .

En effet le paramètre \mathcal{a} varie dans le cône positif de \mathcal{R}^M qui est une variété de dimension M , la densité de DIRICHLET a pour expression

$$D(\mathcal{y}; \mathcal{a}) = \frac{\Gamma(\sum a_k)}{\prod \Gamma(a_i)} \prod_1^M y_i^{a_i-1}$$

C'est la forme générale d'une densité de probabilité d'une structure exponentielle, avec les paramètres naturels a_i et les statistiques exhaustives

$\mathcal{Q}_i = \sum_p \log y_i^p$. Avec les notations de FOURGEAUD-FUCHS (ref. 29 chap.14),

les \mathcal{Q}_i/P sont des estimateurs sans biais de variance minimum parmi les estimateurs sans biais de $\Psi(a_i) - \Psi(\sum a_k)$ avec la matrice de variance-covariance $\Gamma(\mathcal{a})/P$.

(ci-dessous on note $\overline{\log y_i} = \frac{1}{P} \sum_p \log y_i^p$).

2.2. Le système du maximum de vraisemblance .

Les estimateurs \hat{a}_i du maximum de vraisemblance sont solutions du système

$$\begin{cases} \psi(\sum a_k) - \psi(a_i) = -\varphi_i / p \\ i=1, \dots, M \end{cases}$$

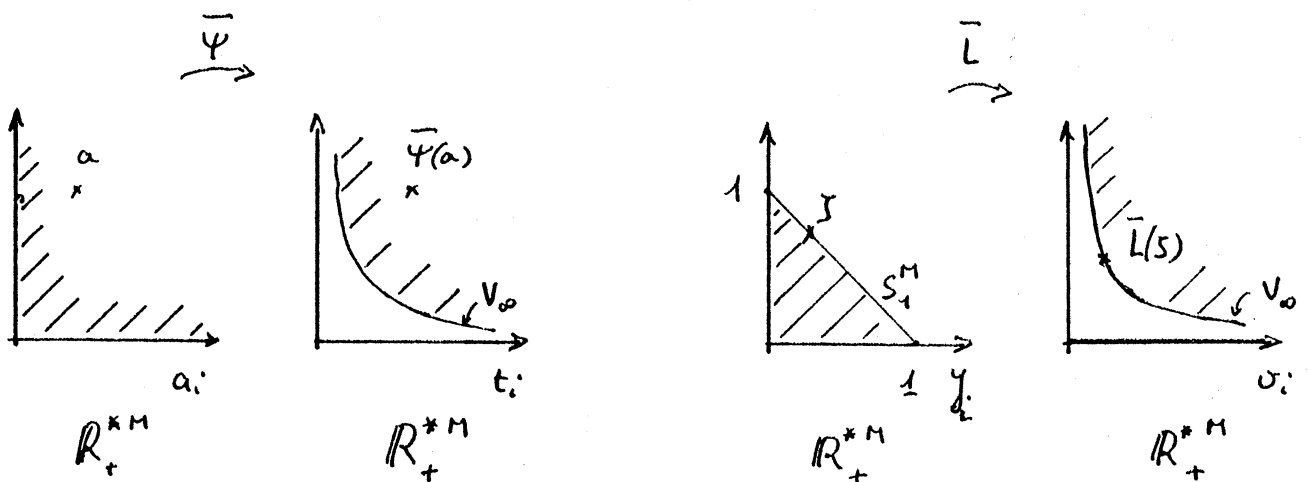
Sous des conditions très générales on sait que ce système possède une solution unique (réf. 4). Ci-dessous on en fait une étude directe.

2.2.1. Soit $\bar{\psi}$ l'application qui à tout point a de \mathbb{R}_+^{*M} de coordonnées a_i positives fait correspondre le point t défini par ses coordonnées :

$$t_i = \psi(\sum a_k) - \psi(a_i) \quad \text{pour tout } i=1, \dots, M .$$

C'est une application définie continue différentiable en tout point a ; la matrice $\partial t / \partial a$ est l'opposée de la matrice de variance-covariance précédente, le jacobien de cette application garde donc un signe constant.

L'image de \mathbb{R}_+^{*M} par $\bar{\psi}$ est incluse de façon triviale dans le cône positif de \mathbb{R}^M , puisque l'application ψ est croissante sur \mathbb{R}_+^* , (les coordonnées t_i sont positives). Lorsque des coordonnées a_i tendent vers zéro, les coordonnées t_i tendent vers plus l'infini (et uniquement dans ce cas).



Les applications $\bar{\psi}$ et \bar{L} .

2.2.2. Soit \bar{L} l'application qui à tout point y de \mathbb{R}_+^{*M} de coordonnées y_i positives fait correspondre le point σ défini par ses coordonnées $\sigma_i = -\log y_i$ pour tout $i = 1, \dots, M$.

C'est une application définie continue différentiable de jacobien non nul de \mathbb{R}_+^{*M} dans \mathbb{R}^M . Soit V_∞ la variété image par \bar{L} du simplexe unité, c'est une variété de dimension $M-1$ dont le point courant σ a pour coordonnées $\sigma_i = -\log \xi_i$, $\xi_i > 0$, $\sum \xi_i = 1$.

Proposition 7 - Toute droite passant par σ qui a des coefficients directeurs positifs ou nuls n'a aucun autre point commun avec V_∞ que σ .

En effet la fonction logarithme étant croissante sur \mathbb{R}_+^* , il n'est pas possible qu'il y ait un autre point ξ' du simplexe unité de coordonnées $\xi'_i > 0$ tel que l'on ait $-\log \xi'_i - \log \xi_i$ positif pour tout $i = 1, \dots, M$. ■

On note encore qu'en tout point σ , les composantes du vecteur normal à V_∞ sont de même signe.

On définit la relation d'ordre partiel $\stackrel{\text{im}}{\leq}$ dans l'espace \mathbb{R}^M , produit des ordres totaux \leq sur chaque axe de coordonnée ; on définit ainsi la relation stricte $\stackrel{\text{im}}{<}$ produit des relations $<$ sur \mathbb{R} .

Soit C_∞ l'ensemble réunion des cônes de sommet σ parallèles au cône positif de \mathbb{R}^M ; c'est l'ensemble des points u de \mathbb{R}^M tels qu'il existe au point σ de V_∞ vérifiant $\sigma \stackrel{\text{im}}{\leq} u$; on définit de même C_∞° l'ensemble des u tels que l'on ait $\sigma \stackrel{\text{im}}{<} u$.

Proposition 8 . L'ensemble C_∞ est convexe.

En effet soient deux points u^0 et u^1 de C_∞ et le compromis $u^\mu = \mu u^1 + (1-\mu) u^0$ pour $0 \leq \mu \leq 1$. Par hypothèse il existe s^0 et s^1 du simplexe unité tels que $\bar{L}(s^0) \stackrel{\text{im}}{\leq} u^0$ et $\bar{L}(s^1) \stackrel{\text{im}}{\leq} u^1$.

La fonction logarithme étant concave on a :

$$-\log(\mu s^1 + (1-\mu) s^0) \leq -\mu \log s^1 - (1-\mu) \log s^0 ,$$

$$\bar{L}(\mu s^1 + (1-\mu) s^0) \stackrel{\text{im}}{\leq} \mu \bar{L}(s^1) + (1-\mu) \bar{L}(s^0) ,$$

$$\bar{L}(\mu s^1 + (1-\mu) s^0) \stackrel{\text{im}}{\leq} u^\mu ,$$

donc u^μ appartient également à C_∞ . \blacksquare

2.2.3.

Lemme 2 - La fonction qui à tout x réel positif fait correspondre la valeur $x \psi'(x)$ est positive décroissante sur \mathbb{R}_+^* .

En effet on a la formule suivante dans WHITTAKER et WATSON (réf. B2, chap. 12.32)

$$\psi'(z) = \frac{1}{2z^2} + \frac{1}{z} + \int_0^\infty \frac{4tz}{(t^2+z^2)^2} \frac{dt}{e^{2\pi t} - 1}$$

$$= \frac{1}{2z^2} + \frac{1}{z} + \frac{4}{z} \int_0^\infty \frac{u du}{(1+u^2)^2 (e^{2\pi u} - 1)}$$

Ce qui démontre le lemme. \blacksquare

Remarque : On le démontre également en remarquant que $\Psi'(x) = \zeta(2, x)$ et en utilisant la formule d'HERMITE pour $\zeta(\Delta, x)$ (réf. 82 chap. 13.2), (ζ fonction de RIEMANN).

Conséquences 1) Pour tout x positif, et tout θ compris entre zéro et un, l'application qui à tout x fait correspondre $\Psi(x) - \Psi(\theta x)$ est une application décroissante.

En effet, l'application est dérivable, sa dérivée est du signe de $x \Psi'(x) - \theta x \Psi'(\theta x)$ donc négative.

2) Les valeurs aux bornes de cette application sont (réf. 1)

• quand $x \rightarrow +\infty$
$$\Psi(x) = \log x - \frac{1}{2x} - \frac{1}{12x^2} + \frac{1}{24} o\left(\frac{1}{x}\right)$$

d'où
$$\Psi(x) - \Psi(\theta x) = -\log \theta + \frac{1}{2x} \left(\frac{1}{\theta} - 1\right) + \frac{1}{24} o\left(\frac{1}{x}\right)$$

et
$$\lim_{x \rightarrow +\infty} \Psi(x) - \Psi(\theta x) = -\log \theta.$$

• quand $x \rightarrow 0^+$
$$\Psi(x) = -\frac{1}{2} - \gamma + x \frac{\pi^2}{6} + x o(x)$$

$$\Psi(x) - \Psi(\theta x) = \frac{1}{2x} \left(\frac{1}{\theta} - 1\right) + x(1-\theta) \frac{\pi^2}{6} + x o(x)$$

et
$$\lim_{x \rightarrow 0^+} \Psi(x) - \Psi(\theta x) = +\infty.$$

3) Ceci complète la démonstration de la proposition 2 en VII.1.2.

Remarques - Lorsque θ est de la forme $1/n$ pour tout n entier ≥ 2 , la formule de multiplication de GAUSS conduit aux mêmes propriétés de $\Psi(x) - \Psi(x/n)$ (cf chap VI.3) ;

- On a encore $\frac{1}{\Psi'(0^+)} < \frac{\theta}{\Psi'(x)}$, cette relation

entraîne l'inégalité $\frac{1}{\Psi'(a)} + \frac{1}{\Psi'(b)} < \frac{1}{\Psi'(a+b)}$ démontrée en VII.2.1.

Proposition 9 - L'image $\bar{\Psi}(a)$ de tout point de coordonnées non nulles de \mathbb{R}_+^{*M} appartient à C_∞^0 .

En effet les coordonnées de $\bar{\Psi}(a)$ sont égales à

$$t_i = \Psi(\sum a_k) - \Psi\left(\frac{a_i}{\sum a_k} \times \sum a_k\right) \quad \text{pour tout } i=1, \dots, M.$$

D'après les conséquences du lemme précédent on a $t_i > -\log\left(\frac{a_i}{\sum a_k}\right)$,

d'où $\bar{L}\left(+\frac{a_i}{\sum a_k}\right) < \bar{\Psi}(a)$ sachant que le point

de coordonnées $a_i / \sum a_k$ appartient au simplexe unité. ■

2.2.4.

Lemme 3 - L'application Ψ^{-1} est définie croissante sur \mathbb{R} . Quand x tend vers plus l'infini $\Psi^{-1}(x)$ est un infiniment grand équivalent à e^x , quand x tend vers moins l'infini, $\Psi^{-1}(x)$ est un infiniment petit équivalent à $-1/x$.

En effet l'application Ψ est définie continue croissante sur \mathbb{R}_+^* à valeur dans \mathbb{R} , son application réciproque Ψ^{-1} est donc définie croissante sur \mathbb{R} . Si l'on pose $x = \Psi(z)$, on cherche l'expression de $z(x)$ aux voisinages de $+\infty$ et $-\infty$.

• Quand x tend vers moins l'infini z tend vers zéro, on a

$$x = -\frac{1}{z} - \gamma + o(z)$$

$$x + \gamma = -\frac{1}{z} (1 - o(z))$$

d'où
$$z = \frac{-1}{x + \gamma} (1 - o(z))$$

$$z = -\frac{1}{x} \left(1 + o\left(\frac{1}{x}\right)\right) \cdot \left(1 - o(z)\right)$$

• Quand x tend vers plus l'infini z tend vers plus l'infini, on a

$$x = \log z - \frac{1}{2z} + o\left(\frac{1}{z}\right)$$

On prend $\bar{x} = \log z$ et $\bar{z} = e^x$, qui vérifient

$$\bar{x} = x + \frac{1}{2z} + o\left(\frac{1}{z}\right)$$

$$\bar{x} - x = \log \frac{\bar{z}}{z} = \frac{1}{2z} + o\left(\frac{1}{z}\right)$$

$$\frac{\bar{z}}{z} = 1 + \frac{1}{2z} + o\left(\frac{1}{z}\right)$$

d'où
$$z = \left(\bar{z} + \frac{1}{2}\right) \cdot \left(1 + o\left(\frac{1}{\bar{z}}\right)\right) = \left(e^x + \frac{1}{2}\right) \cdot \left(1 + o\left(\frac{1}{e^x}\right)\right)$$
 ■

Théorème 5 - L'application $\bar{\Psi}$ est une application bijective entre \mathbb{R}_+^{*M} et \mathcal{C}_∞° .

Démonstration : Nous démontrons ci-après que l'application $\bar{\Psi}$ est injective

(cf proposition 10). Pour démontrer qu'elle est surjective on démontre ici que le système des M équations :

$$\psi\left(\sum a_k\right) - \psi(a_i) = t_i \quad i=1, \dots, M$$

possède une racine en a pour tout point t de C_∞^0 .

i) On pose l'inconnue intermédiaire $x = \psi\left(\sum a_k\right)$, elle vérifie les relations

$$a_i = \psi^{-1}(x - t_i) \quad , \quad i=1, \dots, M$$

et

$$\sum_j a_j = \psi^{-1}(x) \quad .$$

ii) On considère la fonction réelle g qui à tout x réel fait correspondre

$$g(x) = \psi^{-1}(x) - \sum_i \psi^{-1}(x - t_i) \quad ;$$

on remarque que si x^* est un zéro réel de g alors $a_i^* = \psi^{-1}(x^* - t_i)$ est solution du système initial.

iii) La fonction g est définie continue sur \mathbb{R} .

• Quand x tend vers plus l'infini, $g(x)$ est la différence d'infiniments grands équivalent à $e^x (1 - \sum e^{-t_i})$.

Sachant que par hypothèse il existe un point de coordonnées δ_i du simplexe unité tel que $t_i > -\log \delta_i$ on a $1 - \sum e^{-t_i} > 0$, on en déduit que $g(x)$ tend vers plus l'infini quand x tend vers plus l'infini.

• Quand x tend vers moins l'infini, $g(x)$ est un infiniment petit de

la forme $\frac{1}{x} \left(\sum_k \frac{1}{1 - t_k/x} - 1 \right)$.

Pour $|t_k/x| < 1$ on a $\frac{1}{1 - t_k/x} > 1 + \frac{t_k}{x}$,

d'où $\sum_k \frac{1}{1 - t_k/x} > M + \frac{1}{x} \sum t_k > M - \frac{1}{x} \sum \log \delta_k$,

soit encore, la fonction logarithme étant concave

$$\sum_k \frac{1}{1-t_k/\alpha} > M + \frac{1}{\alpha} M \log M,$$

pour α vérifiant $\alpha < \inf(-t_k)$. Pour tout M

supérieur ou égal à deux, le coefficient multiplicateur de $1/\alpha$ dans $g(\alpha)$ est positif quand α tend vers moins l'infini, donc $g(\alpha)$ tend vers zéro par valeurs négatives quand α tend vers moins l'infini.

iv) La fonction g étant continue, il existe donc une racine réelle à distance finie de l'équation $g(\alpha) = 0$, donc une racine unique α^* du système envisagé. ■

2.2.5. Résolution.

Le système du maximum de vraisemblance est un système du type ci-dessus, avec au deuxième membre les quantités $t_i = -\varphi_i / p$.

Théorème 6 - Pour tout ensemble de P partages - observations indépendants, P supérieur ou égal à 2, le système du maximum de vraisemblance en a_i possède une racine unique avec une probabilité 1.

En effet pour tout f le point de coordonnées $-\log y_i^p$ $i=1, \dots, M$ est un point de la variété V_∞ . L'ensemble C_∞ étant convexe le point de coordonnées $-\frac{1}{p} \sum \log y_i^p = -\overline{\log y_i}$ $i=1, \dots, M$ appartient à C_∞ , il appartient à l'ouvert C_∞° dès que deux points y_{t_1} et y_{t_2} sont différents (événement de probabilité égale à 1). D'après le théorème précédent le système a donc une solution \hat{a}_i unique. ■

D'autre part la matrice formée par les dérivés d'ordre deux du logarithme de la vraisemblance par rapport aux paramètres a_i a_j est proportionnelle à la matrice de variance - covariance changée de signe $-\Gamma(a)$; elle est donc définie négative en tout point a . ■

2.2.6. Inégalités particulières.

Lemme 4 Soit g une application réelle définie dérivable sur \mathbb{R}_+^* , qui vérifie les propriétés suivantes :

- g' est positive strictement convexe ,
- l'application qui à tout x réel fait correspondre $x \cdot g'(x)$ est décroissante au sens large.

Alors pour tout entier $n \geq 2$ et tous les réels positifs a_i et b_i $i=1, \dots, n$ de sommes respectives $A = \sum a_i$, $B = \sum b_i$ on a l'inégalité

$$(B - A) [g(B) - g(A)] \leq \sum_{i=1}^n (b_i - a_i) [g(b_i) - g(a_i)] ,$$

avec égalité si et seulement si pour tout $i=1, \dots, n$ on a $a_i = b_i$.

Démonstration: L'égalité est triviale si pour tout $i=1, \dots, n$ on a $a_i = b_i$. On remarque qu'il suffit de démontrer le résultat pour $n=2$.

On pose: $\delta_1 = b_1 - a_1$, $\delta_2 = b_2 - a_2$ et $\delta = B - A = \delta_1 + \delta_2$.

a) $\delta_1 = \delta_2 = 0$, on a l'égalité;

b) $\delta_1 = 0$, $\delta_2 \neq 0$, l'inégalité se réduit à

$$(b_2 - a_2) [g(a_1 + b_2) - g(a_1 + a_2) - g(b_2) + g(a_2)] \leq 0 ;$$

or on a $x g'(x + a_1) < (x + a_1) g'(x + a_1) \leq x g'(b_2)$,

donc l'application définie par $g(x + a_1) - g(x)$ est strictement décroissante sur \mathbb{R}_+^* , on a donc bien l'inégalité stricte dans ce cas;

c) $\delta_1 \neq 0$, $\delta_2 \neq 0$,

On remarque que la quantité $(y-x)[g(y) - g(x)]$ peut se mettre sous la forme intégrale

$$(y-x) [g(y) - g(x)] = \int_x^y \int_x^y g'(u) du dx .$$

Le changement de variables :

$$u = \frac{\int u}{\int_1} + \frac{a_2 \int_1 - a_1 \int_2}{\int_1}, \quad v = \frac{\int v}{\int_1} + \frac{a_2 \int_1 - a_1 \int_2}{\int_1}$$

transforme le domaine d'intégration de $(b_1 - a_1) [g(b_1) - g(a_1)]$ dans celui de $(B - A) (g(B) - g(A))$;

le changement de variables :

$$u = \frac{\int u}{\int_2} + \frac{a_1 \int_2 - a_2 \int_1}{\int_2}, \quad v = \frac{\int v}{\int_2} + \frac{a_1 \int_2 - a_2 \int_1}{\int_2}$$

en fait de même pour $(b_2 - a_2) [g(b_2) - g(a_2)]$. Par conséquent l'inégalité que l'on cherche à démontrer se met sous la forme suivante:

$$\int_A^B \int_A^B \left[g'(u) - \frac{\int_1^2}{\int^2} g'\left(\frac{t_1}{\int}\right) - \frac{\int_2^1}{\int^1} g'\left(\frac{t_2}{\int}\right) \right] du dv \leq 0,$$

en posant

$$t_1 = (u - a_2) \int_1 + a_1 \int_2$$

$$t_2 = (u - a_1) \int_2 + a_2 \int_1$$

$$\text{et } u = \frac{t_1 + t_2}{\int}$$

L'application $x g'$ étant positive décroissante, on en déduit que pour tout $\theta \in]0, 1[$ on a $x g'(x) \leq \theta x g'(\theta x)$, d'où :

$$\frac{\int_1}{\int} g'\left(\frac{t_1}{\int}\right) + \frac{\int_2}{\int} g'\left(\frac{t_2}{\int}\right) \leq \frac{\int_1^2}{\int^2} g'\left(\frac{t_1}{\int}\right) + \frac{\int_2^1}{\int^1} g'\left(\frac{t_2}{\int}\right)$$

g' étant convexe, on a encore :

$$g'(u) = g'\left(\frac{t_1 + t_2}{\int}\right) < \frac{\int_1}{\int} g'\left(\frac{t_1}{\int}\right) + \frac{\int_2}{\int} g'\left(\frac{t_2}{\int}\right),$$

(inégalité stricte puisque $\int_1 \neq 0$ et $\int_2 \neq 0$).

La fonction que l'on intègre sur $[A, B] \times [A, B]$ est continue négative, son intégrale est donc strictement négative, l'inégalité stricte en résulte. ■

Remarque: L'application \bar{f} qui à tout point a de \mathbb{R}^M de coordonnées $a_i, i=1, \dots, M$ fait correspondre le point de coordonnées $g(a_i) = g(\sum a_k)$, $i=1, \dots, M$ est une fonction strictement monotone, (Ref. 52').

Applications: Cette inégalité s'applique dans de nombreux cas classiques:

1) $g(x) = 1/x$, on a

$$(B-A) \log B/A \leq \sum (b_i - a_i) \log b_i/a_i ;$$

Ceci généralise l'inégalité relative à la divergence de JEFFREYS (ref. 49)

$$\sum (p_i - q_i) \log p_i/q_i \geq 0, \quad \sum p_i = \sum q_i = 1.$$

2) $g(x) = 1/x^p$, $p > 1$, on a

$$(B-A) \left(\frac{1}{B^{p-1}} - \frac{1}{A^{p-1}} \right) \leq \sum (b_i - a_i) \left(\frac{1}{b_i^{p-1}} - \frac{1}{a_i^{p-1}} \right).$$

3)

Proposition 10 Pour tout entier $n \geq 2$ et pour tous les réels positifs a_i et $b_i, i=1, \dots, n$ de sommes respectives $A = \sum a_i, B = \sum b_i$, on a l'inégalité suivante:

$$(B-A) [\psi(B) - \psi(A)] \leq \sum_i (b_i - a_i) [\psi(b_i) - \psi(a_i)].$$

Démonstration On sait que l'application ψ' est convexe au sens strict, le lemme 2 montre que ψ' vérifie les autres hypothèses du lemme ci-dessus. ■

Conséquence: Soient a et b deux points de \mathbb{R}_+^M , et $\bar{\psi}(a), \bar{\psi}(b)$ leur image par $\bar{\psi}$. L'inégalité précédente signifie que dans \mathbb{R}^M le produit scalaire

$(b-a)' [\bar{\psi}(b) - \bar{\psi}(a)]$ est négatif pour $a \neq b$. Ceci démontre que $\bar{\psi}$ est une application injective (cf. théorème 5), en outre les hyperplans d'équation

$(b-a)' \tau = K$ séparent $\bar{\psi}(a)$ et $\bar{\psi}(b)$ pour tout K compris entre

$(b-a)' \bar{\psi}(b)$ et $(b-a)' \bar{\psi}(a)$.

- On note également que l'application $-\bar{\psi}$ est strictement monotone.

2.2.7. Compléments.

Proposition 11- L'application \bar{L} est une bijection entre l'ensemble C_∞ et l'ensemble des points y de \mathbb{R}_+^M de coordonnées positives de somme inférieure ou égale à 1.

En effet soit le point J de coordonnées J_i du simplexe unité, et soit le point y de coordonnées $y_i = J J_i$, $0 < J \leq 1$. Son image par \bar{L} a pour coordonnées $-\log J_i = -\log J$ avec $-\log J \geq 0$.

Inversement soit t un point de C_∞ à distance finie, le point de coordonnées $y_i = e^{-t_i}$ a ses coordonnées positives de somme inférieure à 1, son image par \bar{L} est le point t . ■

Remarque. L'ensemble des points y de coordonnées positives ou nulles de somme inférieure ou égale à 1 est dans \mathbb{R}^M isomorphe au simplexe unité S_1^{M+1} . L'application \bar{L} réalise une bijection entre C_∞ et S_1^{M+1} .

2.3. Estimation ensembliste.

On sait que pour de grands échantillons (P élevé) l'estimateur du maximum de vraisemblance \hat{a}_i , $i=1, \dots, M$ est tel que la distribution de $\sqrt{P} (\hat{a}_i - a_i)$ converge en loi vers la loi normale sur \mathbb{R}^M d'espérance nulle et de matrice de variance-covariance $\Gamma(a)^{-1}$.

Dans \mathbb{R}^M l'hyper-ellipsoïde d'équation

$$P (\hat{a} - a)' \Gamma(\hat{a}) (\hat{a} - a) = \lambda_1,$$

où λ_1 est défini par la relation $\Gamma\left(\frac{M}{2}, \lambda\right) ([0 \ \lambda_1]) = \eta$ constitue une estimation ensembliste asymptotique de a avec le seuil de confiance η (réf. 2).

2.4. Méthode numérique .

Le système à résoudre s'écrit de façon globale

$$\bar{\Psi}(a) + \mathcal{Q}/P = 0 ;$$

soit $\frac{\partial \bar{\Psi}}{\partial a}$ la matrice $-\Gamma(a)$ inversible comme on l'a vu, qui est aussi la matrice $\frac{\partial}{\partial a} (\bar{\Psi} + \mathcal{Q}/P)$.

Pour un tel système on peut utiliser la méthode itérative de NEWTON dans \mathbb{R}^M qui consiste à partir d'un point initial a^0 et d'appliquer les approximations successives suivantes :

$$\begin{aligned} a^{n+1} &= a^n - [-\Gamma(a^n)]^{-1} (\bar{\Psi}(a^n) + \mathcal{Q}/P) \\ &= a^n - X \end{aligned}$$

avec à chaque étape la résolution du système linéaire

$$[-\Gamma(a^n)] X = \bar{\Psi}(a^n) + \mathcal{Q}/P .$$

(réf. 3)

Un point de départ possible est le point a^0 qui correspond à la densité invariante par Σ_M , $a_i^0 = a_i^*$ solution de l'équation

$$\Psi(M a_i^0) - \Psi(a_i^0) = - \frac{1}{MP} \sum_{i:P} \log y_i^P \quad (\text{chap. VI}) .$$

Le calcul numérique de Ψ, Ψ' nécessaire pour le calcul du 2° membre \mathcal{Q}/P et de la matrice Γ sont les mêmes que précédemment.

Naturellement ces manipulations numériques sont menées en utilisant des moyens de calcul informatique. Dans les exemples qui suivent, il a fallu en pratique une dizaine d'itérations pour avoir $|a_i^{n+1} - a_i^n| \leq 10^{-3}$.

Remarque : Pour éviter un "dérapiage" dans les premières itérations, on peut prévoir une instruction qui remplace tout a_i^{n+1} négatif intempestif par la valeur initiale a_i^0 .

2.5. Le cas des données nulles .

On a supposé que pour tout $i = 1, \dots, M$ et pour tout $p = 1, \dots, P$ aucune des parts y_i^p des partages de l'échantillon n'est nulle.

Quand on se place comme on le fait ici dans l'hypothèse où les y suivent une distribution de DIRICHLET $D(a)$, une telle éventualité est de probabilité nulle. Il peut cependant se trouver des parts nulles dans un échantillon de partages observés. Si cela se présente on peut adopter plusieurs attitudes:

i) éliminer les partages de l'échantillon qui ont des parts nulles ; ceci peut être raisonnable si leur effectif est faible ;

ii) remplacer les y_i^p nuls par une valeur positive ε du même ordre de grandeur que les erreurs de mesure qui s'appliquent sur les valeurs y_i^p (modifier les autres en conséquence) ;

iii) appliquer certaines propriétés de la distribution de DIRICHLET.

On a vu en effet que si y suit la loi $D(a)$, pour toute partition

φ_ℓ $\ell = 1, \dots, L$ de $\{1, \dots, M\}$, le point z de coordonnées $z_\ell = \sum_{\varphi_\ell} y_i$ suit la loi $D(b)$ avec b le point de coordonnées $b_\ell = \sum_{\varphi_\ell} a_i$.

Par conséquent, on peut s'arranger pour regrouper certaines variables y_i (plus précisément celles qui ont des parts nulles et certaines autres au choix) de façon à obtenir un échantillon de partages z^p tels qu'il n'y ait aucune part nulle z_ℓ^p $\ell = 1, \dots, L$. L'estimation précédente conduit aux estimateurs du maximum de vraisemblance $\hat{b}_\ell = \sum_{\varphi_\ell} \hat{a}_i$.

Cette procédure permet de conserver la totalité de l'échantillon, elle n'est évidemment pas unique (au même titre que la partition des φ_ℓ $\ell = 1, \dots, L$), et peut au besoin être combinée avec les précédentes.

2.6. Lois des statistiques exhaustives .

2.6.1. Soit $\mathcal{F}_M(a, 1)$ la loi image par l'application \bar{L} de la loi $D(a)$ sur le simplexe S_1^M ; c'est une distribution sur la variété V_ω .

Soit un échantillon de taille P de $D(a)$; on note $\mathcal{F}_M(a, P)$ la distribution suivie par le vecteur τ^P de \mathbb{R}_+^{*M} qui a pour coordonnées $\tau_i^P = -\frac{1}{P} \sum_r \log y_i^r$ pour tout $i=1, \dots, M$, (statistiques exhaustives de l'échantillon). On a vu qu'avec une probabilité égale à un, le point τ^P appartient à C_∞ .

Proposition 12 Pour tout P supérieur ou égal à 1, l'espérance mathématique et la matrice de variance-covariance de la distribution $\mathcal{F}_M(a, P)$ sont respectivement le vecteur de coordonnées $\Psi(\sum a_k) - \Psi(a_i)$ $i=1, \dots, M$, et la matrice $\Gamma(a) / P$.

En effet pour $P = 1$, on a $\tau_i^1 = \Psi(\sum a_k) - \Psi(a_i) - \frac{\partial \log D}{\partial a_i}$

pour tout $i=1, \dots, M$. On a rappelé que le point de coordonnées $\frac{\partial \log D}{\partial a_i}$

est centré avec la variance-covariance $\Gamma(a)$ (cf VII.2.1) ; la propriété en résulte. ■

- A l'aide du théorème de la limite centrale multivariée (réf. 2 chap. IV)

on sait encore que la distribution de $\sqrt{P} [\tau^P - E_{\mathcal{F}} \tau^P]$ converge en loi vers la distribution normale d'espérance mathématique le vecteur nul de \mathbb{R}^M , et de variance-covariance la matrice $\Gamma(a)$, quand on fait tendre P vers plus l'infini.

2.6.2. Transformées de LAPLACE.

Propriété 13 - La transformée de LAPLACE de la distribution $\mathcal{F}_M(a, 1)$ image de la loi $D(a)$ par \bar{L} est égale à :

$$R_c(z_j) > - a_j \quad j=1, \dots, M$$

$$\int_{\mathcal{F}} (z_1, \dots, z_M) = \frac{\Gamma(\sum a_j)}{\Gamma(\sum a_j + \sum z_j)} \prod_{j=1}^M \frac{\Gamma(a_j + z_j)}{\Gamma(a_j)}$$

En effet, dans l'espace \mathbb{R}_+^M des statistiques exhaustives

$T_i = -\log y_i$, $i=1, \dots, M$, la mesure image par \bar{L} de $D(a)$

est sur la variété \mathcal{V}_∞ , elle a donc pour expression analytique en tout point T de \mathbb{R}_+^{M-1} l'expression :

$$(-1)^{M-1} \frac{\Gamma(\sum a_k)}{\prod_j \Gamma(a_j)} \prod_{i=1}^{M-1} e^{-a_i T_i} \cdot e^{-(a_n-1) T_n} \int_{\mathcal{T}_n}$$

où $\int_{\mathcal{T}_n}$ est la distribution de DIRAC au point $T_n = -\log [1 - (e^{-T_1} + \dots + e^{-T_{n-1}})]$.

La transformée de LAPLACE est donc égale à :

$$(-1)^{M-1} \frac{\Gamma(\sum a_k)}{\prod_j \Gamma(a_j)} \int \prod_{i=1}^{M-1} e^{-(a_i + \beta_i) T_i} \cdot [1 - (e^{-T_1} + \dots + e^{-T_{n-1}})]^{(a_n + \beta_n - 1)} dT_1 \dots dT_{n-1}$$

soit, après changement de variables $y_i = e^{-T_i}$, $i=1, \dots, M-1$

$$\int_{\mathcal{F}} (\beta_1, \dots, \beta_n) = \frac{\Gamma(\sum a_k)}{\Gamma(\sum a_k + \sum \beta_k)} \prod_{i=1}^M \frac{\Gamma(a_i + \beta_i)}{\Gamma(a_i)}$$

définie pour $\operatorname{Re}(\beta_j) + a_j > 0$ (notamment $\operatorname{Re}(\beta_j) > 0$).

(on utilise la généralisation naturelle de la formule de DIRICHLET aux exposants $a_i + \beta_i$ complexes de partie réelle positive).

Conséquences. 1) La transformée de LAPLACE de la distribution $\mathcal{F}_M(a, P)$

$$\text{est égale à } \left[\int_{\mathcal{F}} \left(\frac{\beta_1}{P}, \dots, \frac{\beta_n}{P} \right) \right]^P ;$$

2) Pour tout $i=1, \dots, M$, la transformée de LAPLACE de la distribution marginale de τ_i est égale à :

$$\frac{\Gamma(\sum a_k)}{\Gamma(\sum a_k + 3)} = \frac{\Gamma(a_i + 3)}{\Gamma(a_i)}$$

et celle de τ_i^p est égale à :

$$\left[\frac{\Gamma(\sum a_k)}{\Gamma(\sum a_k + 3/p)} , \frac{\Gamma(a_i + 3/p)}{\Gamma(a_i)} \right]^p$$

2.6.3. Fonction caractéristique.

La transformée de LAPLACE de l'image de $D(a)$ par \bar{L} est définie pour tous les β_i de parties réelles positives ou nulles, la fonction caractéristique de la distribution des statistiques exhaustives a donc pour valeur, pour tout t_n réel, $n=1, \dots, M$:

$$\frac{\Gamma(\sum a_k)}{\Gamma(\sum a_k - i \sum \beta_k)} = \prod_{n=1}^M \frac{\Gamma(a_n - i \beta_n)}{\Gamma(a_n)}$$

Quand on prend $a_n = 1$ pour tout $n=1, \dots, M$, on peut comparer cette expression à la fonction caractéristique de la distribution invariante $D(1)$ que donne DARLING (Ref. 19).

2.6.4. Densité sur V_∞ .

L'image de S_1^M par \bar{L} est V_∞ . Dans ce changement de variables on a

$$\begin{aligned} \prod_1^M y_i^{a_i-1} \prod_1^{M-1} dy_j &= \prod_1^M y_i^{a_i} \cdot \frac{1}{y_M} \prod_1^{M-1} \frac{dy_j}{y_j} \\ &= \prod_1^M e^{-a_i \tau_i} \cdot (-1)^{M-1} \frac{1}{y_M} \prod_1^{M-1} d\tau_j \end{aligned}$$

La variété V_ω a pour équation $\sum_1^M e^{-\tau_i} = +1$ et pour vecteur normal

le vecteur de coordonnées $e^{-\tau_i}$ et de cosinus directeurs, $k=1, \dots, M$

$$\cos \theta_k = e^{-\tau_k} / \sqrt{\sum_1^M e^{-2\tau_i}}$$

Si σ_ω est la mesure de surface (uniforme) de V_ω ; on a naturellement

$$\begin{aligned} d\sigma_\omega &= d\tau_1 \dots d\tau_{M-1} / \cos \theta_M \\ &= \left(\sum_1^M e^{-2\tau_i} \right)^{-1/2} d\tau_1 \dots d\tau_{M-1} / y_M \end{aligned}$$

On en déduit que l'image $\mathcal{F}_n(a, 1)$ de $D(a)$ par \bar{L} est une distribution de structure exponentielle canonique (ref. 5), sa densité par rapport à la mesure σ_ω a pour expression

$$\frac{d\mathcal{F}_n(a, 1)}{d\sigma_\omega} = \frac{\Gamma(\sum_1^n a_k)}{\prod \Gamma(a_i)} e^{-\sum_1^M a_j \tau_j} \left(\sum_1^M e^{-2\tau_i} \right)^{-1/2}$$

Il en résulte aussi que $\frac{\prod \Gamma(a_i)}{\Gamma(\sum a_k)}$ est la transformée de LAPLACE de la mesure

m définie sur V_ω par la relation $d_m = \left(\sum_1^M e^{-2\tau_i} \right)^{-1/2} d\sigma_\omega$.

2.6.5. Distributions associées.

- Pour tout $P \geq 2$, la distribution image de $\mathcal{F}_n(a, P)$ par l'application \bar{L}^{-1} est une distribution sur le simplexe unité S_{1}^{M+1} .

- Pour tout $P \geq 2$, la distribution image de $\mathcal{F}_n(a, P)$ par l'application $\bar{\Psi}^{-1}$ est une distribution dans le cône positif \mathbb{R}_+^{*M} . Par construction c'est la distribution des estimateurs du maximum de vraisemblance \hat{a}_i . On a vu en VII.2.3. que la distribution du vecteur de coordonnées $\bar{P}(\hat{a}_i - a_i)$ converge en loi vers la distribution normale $\mathcal{N}(0, \Gamma^{-1}(a))$ d'espérance nulle et de variance-covariance $\Gamma^{-1}(a)$.

Proposition 14 - Soit la matrice $\Gamma(\hat{\alpha})$ dont on note l'élément général

Γ_{ij}^{-1} $i=1, \dots, M$, $j=1, \dots, M$. Pour tout k de 1 à M , quand on fait tendre P vers $+\infty$ la distribution de $\sqrt{P}(\hat{a}_k - a_k)$ converge en loi vers la distribution normale d'espérance nulle et de variance Γ_{kk}^{-1} . Quand on fait tendre P vers $+\infty$ la distribution de $\sqrt{P}\left(\sum_1^M (\hat{a}_k - a_k)\right)$ converge en loi vers la distribution normale d'espérance nulle et de variance $\sum_{ij} \Gamma_{ij}^{-1}$.

Cette proposition résulte des propriétés classiques de la loi normale multivariée $\mathcal{N}(\mu, \Sigma)$ de $X \in \mathbb{R}^M$, relative à la distribution marginale normale de toute coordonnée du vecteur aléatoire, et relative à la distribution normale de tout $A X$ (d'espérance $A\mu$ et de variance-covariance $A\Sigma A'$). Pour la somme des coordonnées de X il suffit de prendre pour A la matrice ligne $(1, \dots, 1)$. (réf. 2). ■

2.7. Tests sur α .

2.7.1. Test d'une hypothèse simple contre une hypothèse simple.

On se place dans l'espace C_∞ de la statistique suffisante T^P , d'un échantillon de taille $P \geq 2$; qui a pour coordonnées $i=1, \dots, M$

$$T_i^P = -\frac{1}{P} \sum_P \log y_i^P.$$

Soient a et b deux points de \mathbb{R}_+^{*M} . Si l'on désire tester l'hypothèse $\{a\}$ contre l'hypothèse $\{b\}$ on sait que l'on doit calculer le rapport des densités des lois $\mathcal{F}_M(a; P)$ et $\mathcal{F}_M(b; P)$. Dans le cas présent la région critique (rejet de a) est définie par une relation de

la forme $\sum_i (b_i - a_i) T_i \leq K$. Pour tous a et b différents, il est évidemment qu'il existe toujours des hyperplans H_K perpendiculaires à $b-a$ qui découpent C_∞ en deux sous-ensembles $C_{\infty, K}^+$ et $C_{\infty, K}^-$ tels que

l'on ait $\int_{C_{\infty, K}^+} dF_H(a; P) \geq 0$, $\int_{C_{\infty, K}^-} dF_H(a; P) \geq 0$ et que leur

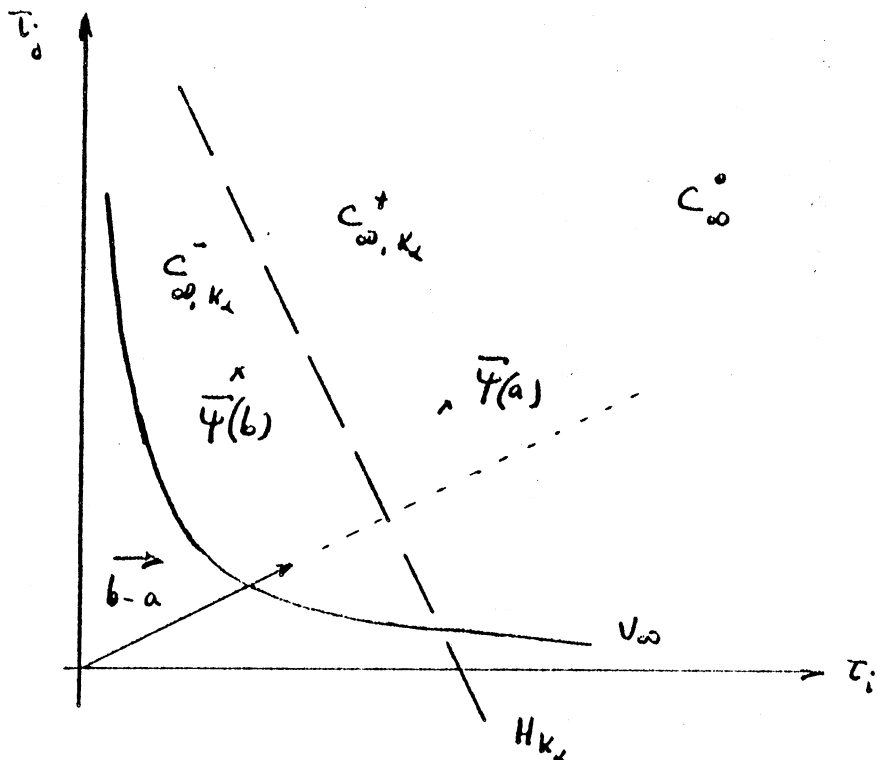
somme soit égale à un ; et de même pour tout α de $]0, 1[$ il existe

un hyperplan H_α d'équation $\sum (b_i - a_i) T_i = K_\alpha$ tel que

$$\int_{C_{\infty, K_\alpha}^-} dF_H(a; P) = \alpha.$$

Pour tout $\alpha \in]0, 1[$ il existe donc un test de puissance maximum défini par la région d'acceptation de l'hypothèse $\{a\}$ C_{∞, K_α}^+ de seuil α .

Sachant que $(b-a)'(\bar{\Psi}(b) - \bar{\Psi}(a)) < 0$, on vérifie que si H_K sépare $\bar{\Psi}(a)$ et $\bar{\Psi}(b)$, $\bar{\Psi}(b)$ est dans la région critique $C_{\infty, K}^-$.



On remarque de façon classique que la région critique ne dépend pas du point b , et qu'elle est la même pour tout point b_γ de la demi-droite

\overrightarrow{ab} décrite par le point $b_\gamma = a + \gamma(b-a)$, $\gamma > 0$. De façon plus générale deux points a et b engendrent une droite D_{ab} décrite par le point $b_\chi = a + \chi(b-a)$, $\chi \in \mathbb{R}$; quand on se limite aux points b_χ qui appartiennent au cône positif ouvert de \mathbb{R}_+^{*n} , on a la structure statistique exponentielle suivante :

$$\left(S_1^M, \mathcal{B}(S_1^M), \left\{ U(\chi) e^{-\chi \sum_{i=1}^M (b_i - a_i) \tau_i} g(\tau), \chi \in \mathbb{R}, a_i + \chi(b_i - a_i) > 0 \quad (i=1, \dots, n) \right\} \right)^P$$

qui ne dépend que d'un paramètre χ et possède la statistique suffisante $\left[\sum (b_i - a_i) \tau_i, \sum \tau_i^p \right]$.

On en déduit les tests habituels UMP, UNPB sur χ (réf.5).

Par exemple, pour le test UMP de $\chi \leq \chi_0$ contre $\chi > \chi_0$ et pour le seuil α la région critique est de la forme $\sum (b_i - a_i) \tau_i \leq k_\alpha$

où l'hyperplan H_{k_α} d'équation $\sum_{i=1}^M (b_i - a_i) \tau_i = k_\alpha$ définit la région critique C_{ω, k_α}^- telle que $\int_{C_{\omega, k_\alpha}^-} dF_M(a; P) = \alpha$;

la puissance de ce test est égale à $1 - \int_{C_{\omega, k_\alpha}^+} dF_n(a; P)$.

2.7.2. Test sur un paramètre d'intérêt unique a_τ .

On peut désirer tester un seul paramètre a_τ ; c'est le cas classique des structures exponentielles à M paramètres qui contient donc $M-1$ paramètres fantômes. On sait que l'on doit conditionner par les statistiques suffisantes pour les paramètres fantômes, lorsque le paramètre principal a_τ est fixé (réf. 17,51).

Dans le cas présent la morphologie de la densité de DIRICHLET permet facilement de ramener le nombre de paramètres fantômes à 1.

En effet, sans réduire la généralité, on suppose que le paramètre principal est a_1 ; dans ce cas, la distribution des statistiques y_1 et $\sum_2^M y_j$ sur S_1^2 est la distribution Béta dont la densité a pour valeur $y_1 \in [0, 1]$

$$\frac{\Gamma(\sum a_i)}{\Gamma(a_1) \Gamma(A_2)} y_1^{a_1-1} (1-y_1)^{A_2-1} \quad \text{avec le nouveau paramètre } A_2 = \sum_2^M a_j$$

C'est une distribution qui a la même distribution marginale en y_1 , qui a le même paramètre d'intérêt a_1 et qui n'a qu'un seul paramètre fantôme A_2 .

Le problème qui reste posé est celui des distributions conditionnelles des distributions $\mathcal{F}_2(a, P)$, par rapport à la statistique

$-\frac{1}{P} \sum_{p=1}^P \left(\log \sum_{j=2}^M y_j^p \right)$, on s'est ainsi ramené à un nouveau problème, plus simple, qui est propre à la distribution Béta.

2.7.3. Test du rapport de vraisemblance.

On a calculé les estimateurs du maximum de vraisemblance pour les paramètres a_i $i=1, \dots, M$ dans le cas de la distribution invariante par Σ_M et de la distribution générale.

On peut donc appliquer le test du rapport des maximum de vraisemblance tel qu'il se pratique dans une structure statistique paramétrée, quand le paramètre θ varie dans un espace Ω de dimension n et que l'on désire tester l'hypothèse nulle H_0 que θ est dans un sous-espace Ω_0 de dimension k . Dans notre exemple Ω est le cône positif de \mathbb{R}^M , variété de dimension M et Ω_0 la demi-droite bissectrice du cône, variété de dimension 1. Soit λ le rapport des maximum de la vraisemblance dans Ω et Ω_0 ; quand l'effectif de l'échantillon tend vers plus l'infini, on sait que la distribution de $-2 \log \lambda$ converge en loi vers la distribution du χ^2 avec ici $M-1$ degrés de liberté. (réf. 17).

On a donc un moyen classique de tester l'hypothèse que l'échantillon est constitué de points qui suivent une distribution de DIRICHLET invariante par

$$\Sigma_M$$

2.8. Remarque .

Les procédures statistiques ci-contre se situent dans le cadre d'une distributions dont on sait a priori que c'est une distribution de DIRICHLET ; l'on a ci-dessus développé une technique d'estimation et d'ajustement compte tenu des propriétés analytiques propres à DIRICHLET et aux structures exponentielles.

On n'a pas abordé le problème de la validité de cet ajustement, ou son bien fondé, à savoir s'il s'agit vraiment d'une distribution de DIRICHLET ou non.

Une distribution de DIRICHLET peut être caractérisée de différentes manières (chapitre IV) ; elle dépend de M paramètres, ce qui signifie notamment que les M moments du 1^o ordre et les $M(M+1)/2$ moments du 2^o ordre ne dépendent que de M paramètres. C'est pour s'affranchir de telles contraintes que CONNOR et MOSIMANN par exemple ont introduit des distributions sur le simplexe unité moins sévères sur le plan des contraintes. (Réf. 16).

VII.3 - Le partage du corps électoral français pour les présidentielles 81

3.1. Introduction .

Un corps électoral est constitué d'un nombre entier d'électeurs inscrits équipondérés, amenés à se prononcer sur un certain nombre d'options politiques. Après scrutin la part respective de chaque option et le découpage du corps sont donnés en comptabilisant le nombre de voix recueillies par chaque option. Sur le plan national les élections présidentielles et les référendums proposent le même jeu d'options à tous les électeurs, ce qui n'est pas forcément le cas des autres élections (législatives,...). Nous avons retenu les présidentielles françaises de 1981 comme exemples de partage. Elles se sont déroulées en deux tours, avec 10 candidats au premier et 2 au second. Pour être complet nous avons regroupé l'ensemble des voix non exprimées (pour raisons diverses) dans la part d'un acteur du partage complémentaire, le 1^o ou le 3^o selon le cas. Nous examinons ces partages en supposant qu'ils obéissent à une distribution de DIRICHLET. Il va de soi que ces élections ne sont ici qu'un prétexte pour disposer de points-partages dans un simplexe unité, et qu'il n'est nullement question de faire de la politologie à cette occasion. Les applications numériques qui suivent ne sont qu'une application formelle des résultats précédents.

Au 1er tour il y a eu 36 324 913 électeurs qui se sont répartis comme suit (réf.59) :

Mmes, Mrs	Giscard d'Estaing	8 181 681	-	22,52%
	Mitterand	7 490 130	-	20,62%
	Non exprimés	7 376 772	-	20,31%
	Chirac	5 026 425	-	14,33%
	Marchais	4 451 960	-	12,26%
	Lalonde	1 123 304	-	3,01%
	Laguiller	668 164	-	1,84%
	Crépeau	641 689	-	1,77%
	Debré	480 805	-	1,32%
	Garaud	384 215	-	1,06%
	Bouchardeau	319 768	-	0,88%

Au 2° tour les 36 260 407 inscrits ont voté comme suit :

Mitterand	15 683 355	-	43,25%
Giscard	14 575 623	-	40,20%
Non expr.	6 001 429	-	16,55%

3.2. Echantillon des partages électoraux .

Les chiffres ci-contre sont les résultats nationaux. Si nous supposons que ces partages obéissent à des distributions de DIRICHLET il est nécessaire de disposer d'un échantillon (aléatoire) de partages pour estimer les coefficients α (distribution invariante) ou α_i (distribution générale). Pour acquérir cet échantillon nous avons pris les dépouillements de voix au niveau de chaque région en supposant que tous ces partages obéissent à la même distribution, soit un échantillon de 24 points-partages (*). Ces suppositions sont ici faites à titre de commodité pratique pour pouvoir acquérir en une seule fois un échantillon, alors que ces suppositions peuvent ne pas refléter la réalité. Nous excluons, avons nous dit, toute discussion relative à la validité du modèle statistique dans cet exemple, ceci toutefois ne doit pas faire perdre de vue que c'est à chaque spécialiste des données traitées qu'il appartient de définir un échantillon correct de partages.

3.3. Distribution invariante .

L'on a supposé dans un premier temps que les partages suivent une distribution invariante par ΣM . Ci-dessous nous indiquons les valeurs des quantités $-\log \bar{v}_n / M$ et $\hat{\alpha}$, et l'intervalle de confiance $[d_1, d_2]$ de $\hat{\alpha}$, intervalle au seuil de 95% dans le cadre d'une variation normale de $\hat{\alpha}$.

(*) On a pris les 22 régions de programme, augmentées de la ville de Paris (extraite de la région Ile de France) et l'ensemble des DOM-TOM.

M	$-\frac{1}{M} \log \frac{1}{\sigma_n}$	$\hat{\alpha}$	α_1	α_2
1° tour M = 11	3,146	0,740	0,628	0,852
2° tour M = 3	1,194	3,714	2,308	5,119

Puisque nous n'en sommes pas à un abus près sur ces données qui servent ici de prétexte à des calculs formels, nous avons modifié les données du 1° tour en regroupant certains candidats qui représentent des options politiques assez proches :

R1 : regroupement des 3 candidats gaullistes,

R2 : R1 plus regroupement de Mesdames Laguiller et Bouchardeau,

R3 : R2 plus regroupement de Messieurs Mitterand et Crépeau.

M	$-\frac{1}{M} \log \frac{1}{\sigma_n}$	$\hat{\alpha}$	α_1	α_2
R1 : 9	2,824	0,849	0,703	0,994
R2 : 8	2,513	1,163	0,944	1,382
R3 : 7	2,269	1,488	1,179	1,797

Ces regroupements ont pour effet de réduire le nombre de petites parts, et l'on voit comment cela se traduit sur l'estimation $\hat{\alpha}$. Dans le cas du regroupement R2 on peut accepter l'hypothèse que l'on a la distribution de DIRICHLET constante sur S_1^p .

3.4. Distribution générale.

L'on reprend les mêmes données en supposant cette fois-ci que les partages suivent une distribution de DIRICHLET sur un simplexe unité.

L'estimation des paramètres α_i par la méthode du maximum de la vraisemblance donne les résultats suivants:

1° tour :

i	Mit.	Gis.	Mar.	Ch.	Lal.	Crep.	Deb.	Gar.
$\hat{\alpha}_i$	29,44	33,44	16,63	21,79	4,45	2,86	2,29	1,90

i	Lag.	Bouch.	N.exp.
$\hat{\alpha}_i$	3,01	1,58	30,87

2° tour :

i	Mit.	Gis.	N. exp.
$\hat{\alpha}_i$	26,64	25,95	11,01

VIII - DISTRIBUTIONS MULTINOMIALES ET DE DIRICHLET

Ce sont deux familles de distributions exponentielles sur le simplexe unité. Dans ce chapitre on examine certaines des propriétés et des relations (exactes ou approchées) qui les réunissent.

VIII.1 - Distributions conjuguées.

1.1. La distribution de DIRICHLET est la conjuguée de la distribution multinomiale au sens où l'entend BARDNORFF-NIELSEN (Ref.4): si

$a(\theta) b(x) e^{\theta t(x)}$ est la représentation canonique minimale d'une structure statistique exponentielle pour la statistique exhaustive et le paramètre naturel θ , la structure conjuguée a pour représentation

$d(\gamma, x) a(\theta)^\gamma e^{x\theta}$, autre structure exponentielle.

Dans le cas présent en effet, on considère la distribution multinomiale $m(N; \theta)$. Cette distribution dépend d'un nombre entier N et d'un point θ du simplexe unité (non situé en un sommet), la probabilité de toute observation n_i , $\sum_i n_i = N$ est égale à $m(n; N; \theta) = \frac{N! \prod \theta_i^{n_i}}{\prod n_i!}$. La structure associée est exponentielle, et on constate que la structure conjuguée est celle de DIRICHLET de densité:

$$D(\theta; a) = \frac{\Gamma(\sum a_i)}{\prod \Gamma(a_i)} \prod_1^M \theta_i^{a_i - 1}$$

1.2. De façon plus générale, en termes bayesiens, cette conjugaison signifie que les distributions a priori et a posteriori $g_1(\omega)$, $g_2(\omega/\gamma)$ du paramètre ω appartiennent à une même famille, compte tenu de la "composition" classique de g_1 par une distribution $f(\gamma/\omega)$ de l'échantillon γ :

$$g_2(\omega/\gamma) = \frac{g_1(\omega) f(\gamma/\omega)}{\int g_1(\omega') f(\gamma/\omega') d\omega'}$$

lorsque par définition la famille des g est conjuguée pour f . Au sens algébrique le passage de g_1 à g_2 est une loi externe dans la famille des g . Ici Y est le vecteur des n_i (échantillon), ω est le vecteur des paramètres θ_i , f une distribution multinomiale $m(N; \theta)$, g des distributions conjuguées de DIRICHLET $g_1 = D(\cdot; a)$, $g_2 = D(\cdot; a + n)$, (Ref. 17).

1.3. Remarque : Cette notion de conjugaison entre structures exponentielles différentes ne doit pas être confondue avec la conjugaison entre densités exponentielles au sens de KULLBACK (Ref. 49). Cette fois-ci les densités conjuguées appartiennent à la même structure.

VIII.2 - Identités numériques.

2.1.

Il existe de nombreuses identités entre les distributions $m(N; \theta)$ et $D(a)$ sommées sur des sous-ensembles particuliers des simplexes respectifs $S_1^{k+1}(\frac{a}{n})$ et $S_1^{k+1}(y)$.

Soit la densité de DIRICHLET :

$$D(y; A + 1, N + 1 - A - k) = \frac{\Gamma(N+1)}{\Gamma(N+1-A-k) \prod_1^k \Gamma(A_i+1)} \prod_1^k y_i^{A_i} (1-y_i)^{N-A-k},$$

avec $A_i = \sum_1^k A_i$, $y_i = \sum_1^k y_i \leq 1$; A (resp y) les vecteurs de \mathbb{R}^k (resp. \mathbb{R}^{k+1}) de coordonnées A_i (resp. y_i et $1-y_i$);

et la distribution multinomiale :

$$m(n; N; \theta) = \frac{N!}{\prod n_i! (N-n)!} \prod \theta_i^{n_i} (1-\theta_i)^{N-n},$$

avec $n_i = \sum_1^k n_i \leq N$, $\theta_i = \sum_1^k \theta_i \leq 1$, n (resp. θ) les vecteurs de \mathbb{N}^{k+1} (resp. S_1^{k+1}) de coordonnées n_i , $N-n$. (resp. θ_i , $1-\theta_i$).

On rappelle ci-dessous les identités du type :

$$\sum_{n \in \mathcal{D}} m(n; N; \theta) = \int_{y \in \mathcal{D}^*} D(y; \Delta+1, N+1-d.-k) dy$$

établies particulièrement par I. OLKIN, M. SOBEL et C. KHATRI, S. MITRA.

On a par exemple :

$$i) \quad \mathcal{D} = \left\{ n : 0 \leq n_i \leq \Delta_i \leq N \quad i=1, \dots, k \right\}$$

$$\mathcal{D}^* = \left\{ \theta : \theta_i \leq y_i \leq 1 \quad i=1, \dots, k \right\}$$

$$ii) \quad \mathcal{D} = \left\{ n : \Delta_{i+1} \leq n_i \leq N \quad i=1, \dots, k \right\}$$

$$\mathcal{D}^* = \left\{ \theta : 0 \leq y_i \leq \theta_i \quad i=1, \dots, k \right\}$$

$$iii) \quad \mathcal{D} = \left\{ n : \begin{array}{l} 0 \leq n_i \leq \Delta_i \quad i=1 \dots j \\ \Delta_{i+1} \leq n_i \leq N \quad i=j+1 \dots k \end{array} \right\}$$

$$\mathcal{D}^* = \left\{ y : \begin{array}{l} \theta_i \leq y_i \leq 1 \quad i=1 \dots j \\ 0 \leq y_i \leq \theta_i \quad i=j+1 \dots k \end{array} \right\}$$

$$\text{iv) } \omega = \left\{ \begin{array}{l} n: \\ \quad \lambda_{1+1} \leq n_1 \\ \quad \lambda_{1+2} + 2 \leq n_1 + n_2 \\ \quad \vdots \\ \quad \lambda + k \leq n. \end{array} \right\}$$

$$\omega^* = \left\{ \begin{array}{l} y: \\ \quad 0 \leq y_1 \leq \theta_1 \\ \quad 0 \leq y_2 \leq \theta_1 + \theta_2 - y_1 \\ \quad \vdots \\ \quad 0 \leq y_k \leq \theta_k - (y_1 + \dots + y_{k-1}) \end{array} \right\}$$

(Réf.47,62.)

2.2. Remarques .

Dans ces identités les densités de DIRICHLET sont des monômes algébriques en y_i avec des exposants entiers. Les paramètres θ_i et λ_i tiennent lieu simultanément de paramètre dans un membre et de borne d'intégration dans l'autre.

Ces identités généralisent l'identité plus classique

$$\sum_{j=0}^{\lambda} \binom{\lambda}{j} \theta^j (1-\theta)^{\lambda-j} = \frac{\Gamma(n+1)}{\Gamma(\theta+1) \Gamma(n-\theta)} \int_0^1 x^{\theta} (1-x)^{n-\theta-1} dx ,$$

entre distributions binomiales et bêta, réf.68.

Dans les mêmes articles, on peut noter de nombreuses relations qui ont pour but de calculer de manière approchée mais précise les probabilités cumulées de la distribution multinomiale. Les identités rappelées ci-dessus permettent d'appliquer ces approximations aux densités de DIRICHLET à exposants entiers.

Dans un ordre d'idée comparable, ces deux mêmes articles contiennent également des relations analogues entre les distributions multinomiales négatives, et les densités de DIRICHLET inverses de TIAO et GUTTMAN qui y apparaissent en substance sans être désignées comme telles, (les communications de TIAO-GUTTMANN et de OLKIN-SOBEL sont contemporaines). (réf.62,78).

VIII.3 - Convergence entre les distributions.

Les relations précédentes concernent des exposants entiers et des domaines d'intégration très particuliers. On examine ici la convergence mutuelle de ces deux distributions, ce qui permet d'obtenir de nouvelles relations, approchées mais plus générales.

3.1.

Soient pour tout $i = 1, \dots, M$, $a_i > 0$, $\theta_i \geq 0$, $\sum_{i=1}^M \theta_i = 1$ les paramètres respectifs de distributions multinomiale et de DIRICHLET.

Proposition 15 i) Quel que soit λ réel positif, les moments du 1^o ordre de la distribution de DIRICHLET de paramètres $a_i = \lambda \theta_i$ sont identiques à ceux de la distribution multinomiale;

ii) Si $\lambda = N - 1$ les moments du 2^o ordre sont également identiques;

iii) Si l'on fait tendre N vers plus l'infini, et si $\lambda = N [1 + O(\frac{1}{N})]$ les moments des deux distributions ont la même limite finie, la convergence vers cette limite est en $1/N$.

En effet

i) On a trivialement

$$E_D(y_i) = a_i / \sum a_k, \quad E_m \left(\frac{n_i}{N} \right) = \theta_i / \sum \theta_k ;$$

$$ii) \quad \bar{E}_D (y_i y_j) = \frac{a_i a_j}{(\sum a_k) (1 + \sum a_k)}$$

$$E_D (y_i^2) = \frac{a_i (1 + a_i)}{(\sum a_k) (1 + \sum a_k)}$$

$$\bar{E}_m \left(\frac{y_i y_j}{N^2} \right) = \frac{N-1}{N} \theta_i \theta_j$$

$$E_m \left(\frac{y_i^2}{N^2} \right) = \frac{\theta_i}{N} (1 + (N-1) \theta_i)$$

La vérification du deuxième point est immédiate.

- On a encore

$$\bar{E}_D (y_1 y_2 y_3) = \frac{a_1 a_2 a_3}{(\sum a_k) (1 + \sum a_k) (2 + \sum a_k)}$$

et

$$\bar{E}_m \left(\frac{y_1 y_2 y_3}{N^3} \right) = \theta_1 \theta_2 \theta_3 \frac{(N-1)(N-2)}{N^2}$$

ces moments sont différents même avec $\lambda = N-1$, l'identité entre moments n'est donc pas vérifiée pour les moments d'ordre supérieur ou égal à 3.

iii) On peut écrire cependant

$$\begin{aligned} E_D (y_1^{k_1} \dots y_n^{k_n}) &= \frac{\Gamma(\sum a_i)}{\Gamma(\sum_{j=1}^n a_j + k_j)} \prod_{r=1}^n \frac{\Gamma(a_r + b_r)}{\Gamma(a_r)} \\ &= \frac{\Gamma(\lambda)}{\Gamma(\lambda + K)} \prod_{r=1}^n \lambda \theta_r (\lambda \theta_r + 1) \dots \\ &\quad \dots (\lambda \theta_r + k_r - 1) \end{aligned}$$

$$= \frac{\lambda^K \Gamma(\lambda)}{\Gamma(\lambda+K)} \prod_2 \theta_2^{k_2} \cdot \left(1 + O\left(\frac{1}{\lambda}\right)\right)$$

$$= \prod_2 \theta_2^{k_2} \cdot \left(1 + O\left(\frac{1}{\lambda}\right)\right)$$

en posant $K = \sum_j^M k_j$, $a_j = \lambda \theta_j$.

D'autre part, pour $N \geq K$ les moments factoriels de la distribution multinomiale sont:

$$E_m \left(n_1^{(k_1)} \dots n_M^{(k_M)} \right) = N^{(K)} \cdot \prod_j \theta_j^{k_j}$$

d'où

$$E_m \left(\left(\frac{n_1}{N}\right)^{k_1} \dots \left(\frac{n_M}{N}\right)^{k_M} \right) = \frac{N^{(K)}}{N^K} \cdot \prod_j \theta_j^{k_j} \cdot \left(1 + O\left(\frac{1}{N}\right)\right)$$

$$= \prod_j \theta_j^{k_j} \left(1 + O\left(\frac{1}{N}\right)\right).$$

Le cas $\lambda = N-1$ assure on l'a vu l'identité des moments jusqu'à l'ordre 2. ■

Remarque: Les deux premiers points se trouvent en substance dans JOHNSON (Ref. 41), qui précise en outre les coefficients classiques γ_1 et γ_2 de dissymétrie et d'aplatissement, (différents puisqu'ils dépendent des moments d'ordre 3 et 4).

3.2.

Soit $\{\Omega, \mathcal{F}, \mu\}$ un espace de probabilité sur lequel sont définis les vecteurs aléatoires à valeur dans S_1^M qui suivent les distributions multinomiales $m(N; \theta)$ et de DIRICHLET $D(a)$.

Théorème 7 Quel que soit θ un point du simplexe unité S_1^M non situé en un sommet, et λ un scalaire réel positif, on pose $Y_{0,\lambda}$ le vecteur aléatoire qui suit la loi $D(\lambda\theta)$. Alors $Y_{0,\lambda}$ converge en probabilité vers θ , $Y_{0,\lambda} \xrightarrow{P} \theta$, lorsque l'on fait tendre λ vers plus l'infini. En outre, quels que soient η et la norme $\|\cdot\|$ dans \mathbb{R}^M , il existe λ^* tel que pour tout $\lambda \geq \lambda^*$ la probabilité $P_2(\|Y_{0,\lambda} - \theta\| \geq \eta)$ est inférieure à une fonction positive décroissante en λ qui tend vers 0 quand λ tend vers plus l'infini.

Démonstration: i) on suppose que θ appartient à l'intérieur du simplexe unité, dans le cas contraire en effet on peut se ramener à un point intérieur d'un simplexe S_1^{n-1} .

ii) On sait que dans \mathbb{R}^m toutes les normes sont équivalentes. Pour l'une d'entre elles notée $\| \cdot \|$, on considère la boule ouverte $B(\theta, \eta)$ de centre θ et de rayon η de telle façon qu'elle soit à l'intérieur du simplexe, et ${}^c B(\theta, \eta)$ son complémentaire dans S_1^n (fermé). On a vu que pour tout $y \neq \theta$ la quantité $D(y; \lambda \theta)$ est une fonction positive décroissante en λ , pour λ supérieur à la racine $\bar{\lambda}_y$ de l'équation $\psi(\lambda) - \sum_k \theta_k \psi(\theta_k \lambda) = - \sum_j \theta_j \log y_j$.

iii) La fonction qui à tout y fait correspondre $-\sum_j \theta_j \log y_j$ est minorée par $-\sum_k \theta_k \log \theta_k$, elle atteint son minimum sur le compact ${}^c B(\theta, \eta)$ en y^* auquel correspond la racine λ^* de l'équation :

$$\psi(\lambda) - \sum_k \theta_k \psi(\theta_k \lambda) = \inf_{y \in {}^c B(\theta, \eta)} \left(- \sum \theta_j \log y_j \right).$$

Par construction, pour tout $y \in {}^c B(\theta, \eta)$ la fonction $D(y; \cdot)$ est décroissante en λ pour $\lambda \geq \lambda^*$, donc la probabilité $P_2(\|Y_{\theta, N} - \theta\| \geq 1)$ qui est inférieure à $\sup_{y \in {}^c B(\theta, \eta)} (D(y; \lambda \theta))$ tend vers 0 quand λ tend vers plus l'infini, † (convergence en probabilité).

En outre, par construction $\sup_{y \in {}^c B(\theta, \eta)} (D(y; \lambda \theta))$ est une fonction décroissante en λ pour $\lambda \geq \lambda^*$, qui tend vers 0 quand λ tend vers plus l'infini. ■

Théorème 8 Avec le même θ que dans le théorème précédent, on pose $Y_{n, N}$ le vecteur aléatoire qui suit la loi multinomiale $m(N; \theta)$. Alors $Y_{n, N}$ converge presque sûrement vers θ lorsque l'on fait tendre N vers plus l'infini, $Y_{n, N} \xrightarrow{P_2} \theta$.

Il s'agit de la loi forte des grands nombres appliquée à la loi multinomiale. En particulier, on sait qu'il existe N^* tel que pour tout $N \geq N^*$ on a $P_2(\|Y_{n, N} - \theta\| \geq 1)$ inférieur à une fonction décroissante positive en N qui tend vers 0 quand $N \rightarrow +\infty$.

Corollaire : Avec le même θ que dans les théorèmes précédents, on a

$$Y_{\theta, \lambda} - Y_{n, N} \xrightarrow{P_2} 0 \text{ quand on fait tendre } \lambda \text{ et } N \text{ vers plus l'infini.}$$

† On rappelle que l'application qui à tout λ réel positif fait correspondre $\psi(\lambda) - \sum \theta_j \psi(\theta_j \lambda)$ est positive décroissante sur \mathbb{R}_+^* (conséquence du lemme 2), donc $\lambda^* \geq \bar{\lambda}_y$ $\forall y \in {}^c B(\theta, \eta)$.

En effet on a établi que $Y_{D,\lambda} \xrightarrow{P_2} \theta$ resp. $Y_{m,N} \xrightarrow{P_2} \theta$ quand λ , resp. N tendent vers plus l'infini.

Pour les variables aléatoires, on sait que si l'on a $X_n \xrightarrow{P_2} 0$ et $Y_n \xrightarrow{P_2} 0$ alors $X_n - Y_n \xrightarrow{P_2} 0$ quand n et n' tendent vers plus l'infini; ceci résulte par exemple de l'inégalité triangulaire de la distance entre variables aléatoires (X, X') , distance qui sert à définir la convergence en probabilité, (Ref. 79).

Etendu aux vecteurs aléatoires de \mathbb{R}^M (†) , on en déduit donc que l'on a

$Y_{D,\lambda} - Y_{m,N} \xrightarrow{P_2} 0$ quand on fait tendre λ et N vers plus l'infini.

On rappelle en outre que les deux distances $(Y_{D,\lambda} - \theta, 0)$ et $(Y_{m,N} - \theta, 0)$ dont la somme majore $(Y_{D,\lambda} - Y_{m,N}, 0)$ sont respectivement majorées par des fonctions positives en λ , resp. N , décroissantes pour tout λ , resp. N , supérieurs à λ^* , resp. N^* , et de limite zéro quand on fait tendre λ , resp. N , vers plus l'infini.

(†) Pour cette extension, voir par exemple les théorèmes du chapitre 4 de la référence 29.

Remarque : Cette convergence permet donc de remplacer une distribution par une autre sur des domaines du simplexe très ressemblants, on est ainsi conduit à des relations numériques d'une toute autre nature que celles du paragraphe VIII.2. On trouve dans JOHNSON des tables comparatives de sommation pour les distributions binomiales et bêta avec $\lambda = N-1$, $N = 25, 50$ ($N = 100$ chez BENEDETTI) Réf. 9,41.

VIII.4 - Autre convergence vers une distribution de DIRICHLET.

Pour remplacer une distribution multinomiale par une distribution de DIRICHLET, et inversement, on a vu ci-dessus qu'il est préférable que les exposants α_i soient de somme élevée ($\sum \alpha_i = N-1$).

4.1. La distribution multinomiale - DIRICHLET composée.

On a vu au chapitre V une autre convergence dans laquelle interviennent ces deux distributions. Soit $m(N/\theta)$ une distribution multinomiale conditionnée par θ et $h(\cdot)$ une densité de probabilité pour θ ; on a la distribution définie par la probabilité suivante:

$$\bar{\omega}_h(\cdot) = \int_{S_1^n} m(\cdot; N/\theta) h(\theta) d\theta$$

en tout n_i/N . On a noté la convergence de $\bar{\omega}_h$ vers h quand on fait tendre N vers plus l'infini, donc la possibilité de remplacer $\bar{\omega}_h$ par h pour N élevé et a_i quelconque. Avec la densité de DIRICHLET $h(\cdot) = D(\cdot; a)$, la distribution $\bar{\omega}_h$ a pour expression :

$$\bar{\omega}_h\left(\frac{n}{N}\right) = \frac{N! \Gamma(\sum a_k)}{\Gamma(N + \sum a_i)} \times \prod_{j=1}^M \frac{\Gamma(a_j + n_j)}{n_j! \Gamma(a_i)}$$

la distribution multinomiale-DIRICHLET composée.

A présent la convergence vers une distribution de DIRICHLET ne dépend pas de la valeur élevée des exposants a_i .

- On signale l'approximation numérique de KHATRI et MITRA :

Si
$$\bar{\omega}_H(\cdot) = \int_{S_1^n} m(\cdot; N/\theta) dH(\theta)$$

on a
$$\sum_{n \in \mathcal{D}} \bar{\omega}_H(n) = H\left(\frac{c_1+1}{N+1}, \dots, \frac{c_{n-1}+1}{N+1}\right) + \eta$$

avec

$$\mathcal{D} = \left\{ \begin{array}{l} n_i \leq c_i \quad i=1 \dots n-1 \\ \sum c_i \leq N \end{array} \right\}$$

et
$$\eta \leq \frac{A}{8(N+2)} \quad \text{ou} \quad \eta \leq \frac{A^*}{4\sqrt{N+2}}$$

selon les conditions de régularité vérifiées par H , (Réf. 47).

4.2. Application.

La distribution multinomiale DIRICHLET composée n'est pas une distribution exponentielle, puisque les statistiques n_i et les paramètres a_i ne sont pas factorisés, dans les termes $\Gamma(a_i + n_i)$ par exemple. Les pratiques statistiques habituelles ne sont donc pas très faciles avec cette distribution.

J.E. MOSIMANN étudie des échantillons de partages qu'il suppose suivre une distribution multinomiale - DIRICHLET composée, et il estime les coefficients de corrélation ρ_{ij} entre les données n_i/N et n_j/N . Compte tenu de la convergence de cette distribution vers une distribution de DIRICHLET, on peut substituer l'une à l'autre et utiliser les méthodes d'estimation sur densité de DIRICHLET vues ci-dessus. Connaissant les \hat{a}_i on en déduit les $\rho_{ij}(\hat{a})$ à l'aide des formules classiques sur ces densités (Réf. 84). Ci-dessous on applique les méthodes d'estimation des α_i sur les données de MOSIMANN et l'on compare les valeurs des corrélations qui en résultent. (Réf.57).

4.2.1. Les données DEMERS-CARBONNEAU-MOSIMANN.

Il s'agit de mesures faites sur le sérum de canetons blancs de Pékin âgés de trois semaines. On s'intéresse aux teneurs en préalbumine (y_1), albumine (y_2) et globuline (y_3) sur un échantillon d'analyses de 23 sérums (tableau 3).

Tableau 3. *Relative frequencies of serum proteins in white Pekin ducklings as determined by electrophoresis*

p_1 , pre-albumin; p_2 , albumin; p_3 , globulins

p_1	p_2	p_3	p_1	p_2	p_3
0.178	0.346	0.476	0.060	0.435	0.505
.162	.307	.531	.089	.418	.493
.083	.448	.469	.050	.485	.465
.087	.474	.439	.073	.378	.549
.078	.503	.419	.064	.562	.374
0.010	0.456	0.504	0.085	0.465	0.450
.049	.363	.588	.094	.388	.518
.100	.317	.583	.014	.449	.537
.075	.394	.531	.060	.544	.396
.084	.445	.471	.031	.569	.400
			0.025	0.491	0.484
			.045	.613	.342
			.0195	.526	.4545

Si l'on fait l'hypothèse que ces données représentent des partages aléatoires qui suivent une distribution de DIRICHLET on a les résultats suivants :

$$\begin{aligned} \text{i) en invariant} \quad \hat{\alpha} &= 1,116 \\ \text{intervalle} & [0,724 \quad 1,509] \quad (95\%) ; \end{aligned}$$

$$\begin{aligned} \text{ii) en général} \quad \hat{a}_1 &= 3,215 \\ \hat{a}_2 &= 20,382 \\ \hat{a}_3 &= 21,685 \end{aligned}$$

Compte tenu de ces valeurs le coefficient de corrélation ρ_{12} est égal à - 0,250 ; celui que trouve MOSIMANN est égal à - 0,252.

4.2.2. Les données CLISBY - SEARS - MOSIMANN .

Ces données portent sur l'analyse de pollens déposés, et sur les teneurs respectives en provenance de 4 origines, les genres Pinus (y_1) Abies (y_2) Quercus (y_3) et Alnus (y_4).

On dispose d'un échantillon de 73 partages en 4 parts, dont certaines sont nulles (notamment y_2 et y_4). Pour éviter cette complication on a remplacé les valeurs nulles par 10^{-5} . Si l'on applique les techniques d'estimation adaptées à la structure de DIRICHLET générale on obtient :

$$\begin{aligned} \hat{a}_1 &= 12,156 \\ \hat{a}_2 &= 0,202 \\ \hat{a}_3 &= 1,310 \\ \hat{a}_4 &= 0,389 \end{aligned}$$

Ci-dessous on indique les coefficients de corrélation déduits de ces valeurs pour une distribution de DIRICHLET, et ceux qu'obtient MOSIMANN :

	1	2	3	4		1	2	3	4
1	1	-0,30	-0,79	-0,46	1	-0,305	-0,811	-0,427	
2		1	-0,04	-0,06		1	-0,039	-0,020	
3			1	-0,02				1	-0,054
4				1					1

Corrélations de MOSIMANN

Corrélations par la distribution
de DIRICHLET

Forest pollen counts from the Bellas Artes core, Clisby & Sears (1955)

Only counts for which 100 total grains were obtained are included here. Several counts are adjusted slightly in the case where grains of spruce (*Picea*) were included in the 100 total. Seventy-three levels are arranged in order of occurrence.

<i>Pinus</i>	<i>Abies</i>	<i>Quercus</i>	<i>Alnus</i>	<i>Pinus</i>	<i>Abies</i>	<i>Quercus</i>	<i>Alnus</i>
	(top of core)						
94	0	5	1	81	1	15	3
75	2	14	9	81	1	16	2
81	2	13	4	76	2	18	4
95	2	3	0	87	3	7	3
89	3	1	7	91	1	5	3
S4	5	7	4	94	0	5	1
81	3	10	6	88	1	11	0
97	0	2	1	93	4	2	1
86	1	8	5	84	0	3	8
86	2	11	1	87	1	12	0
S2	2	10	6	89	1	6	4
72	1	16	11	73	0	13	14
89	0	9	2	87	3	8	2
93	4	2	1	94	1	3	2
87	1	11	1	81	2	9	8
S5	0	12	3	83	0	9	3
91	0	7	2	94	0	4	2
95	1	3	1	69	7	18	6
94	3	3	0	90	0	8	2
85	1	12	2	86	1	8	5
91	1	4	4	90	0	7	3
99	1	0	0	74	5	16	5
90	2	8	0	82	2	11	5
91	0	8	1	87	3	9	1
79	1	19	1	68	3	26	3
S9	0	7	4	77	3	11	9
95	2	1	2	86	2	7	5
90	3	5	2	79	1	11	9
93	1	6	0	87	0	11	2
90	2	7	1	79	1	17	3
S9	2	9	0	74	0	19	7
88	1	9	2	80	0	14	6
99	0	1	0	85	3	9	3
86	1	10	3				
88	0	7	5				
	(bottom of core)						
91	0	7	2				
84	0	14	2				
84	1	12	3				
97	0	3	0				
83	0	13	4				

La convergence de la distribution multinomiale-DIRICHLET composée vers la distribution de DIRICHLET permet ici de "récupérer" avec avantage les propriétés statistiques de cette dernière.

VIII Bis - AUTRES STRUCTURES, STRUCTURES DE FISHER VON-MISES, DE DIRICHLET
" SPHERIQUE "

Le balayage des distributions de probabilité avec densités absolument continues (chap. IV) ou discrètes (chap. V) conduit à autant de structures associées paramétrées.

1 - Nous mentionnons une autre structure possible, liée à une structure très étudiée. Si y est un point du simplexe unité de \mathbb{R}^M de coordonnées y_i le point ζ de coordonnées $\zeta_i = +\sqrt{y_i}$ est un point de l'hypersphère unité de \mathbb{R}^M situé dans le cône positif. Ordinairement on utilise la distribution de FISHER VON-MISES et la structure associée pour des données que l'on observe sur l'hypersphère unité, (Ref. 4, 17). Cette distribution - très étudiée - est paramétrée par un vecteur modal de \mathbb{R}^M et un paramètre de forme ou de concentration réel positif ou nul. En toute rigueur cette distribution est définie sur l'hypersphère entière et non pas uniquement sur la partie comprise dans le cône positif. Cependant pour un vecteur modal qui appartient à ce cône, et avec un paramètre de forme élevé on peut faire l'approximation que quasiment toute la probabilité est distribuée dans le cône positif, et partant de là on peut l'appliquer sur les racines carrées positives des coordonnées d'un point du simplexe unité. On a par exemple appliqué cette distribution sur des y_i qui décrivent les différentes parties temporelles d'un budget-temps (Ref.75)

2 - Avec le changement de variable ci-dessus, on a :

$$\prod_1^M y_i^{a_i-1} = \prod_1^M \zeta_i^{2a_i-1} / \prod_1^M \zeta_k$$

$$\prod_1^{n-1} dy_j = 2^{n-1} \prod_1^n \zeta_k \cdot \prod_1^{n-1} d\zeta_k / \zeta_n$$

d'où

$$D(y; a) = 2^{n-1} \frac{\Gamma(\sum a_i)}{\prod \Gamma(a_j)} \cdot \prod \zeta_k^{2a_j-1} \cdot \frac{d\zeta_1 \dots d\zeta_{n-1}}{\zeta_n}$$

En tout point ζ de l'hypersphère unité le vecteur normal est $\vec{o\zeta}$ lui-même, avec les cosinus directeurs $\cos \theta_k = \zeta_k$, $k=1 \dots M$. Si σ est la mesure de surface (uniforme) de l'hypersphère de \mathbb{R}^M , on a :

$$d\sigma = \frac{1}{|\cos \theta_k|} \prod_{j \neq k} d\zeta_j$$

et en particulier, dans le cas présent $d\sigma = d\zeta_1 \dots d\zeta_{n-1} / \zeta_n$.

On en déduit donc une nouvelle distribution sur le quadrant positif de l'hyper-

sphère unité qui a pour densité $2^{n-1} \frac{\Gamma(\sum a_k)}{\prod \Gamma(a_i)} \cdot \prod s_j^{2a_j-1}$,

par rapport à la mesure uniforme sur le quadrant positif de la sphère unité, (peut-on l'appeler distribution sphérique de DIRICHLET ?). Pour les points du quadrant positif c'est une nouvelle distribution de probabilité, de la même façon que la distribution de FISHER VON-MISES peut être adaptée sur le simplexe unité.

Remarque: Par ailleurs on sait que (Ref. 72)

$$2^{n-1} \int_{\substack{\sum s_i = 1 \\ s_i > 0}} \prod_{i=1}^n s_i^{2a_i-1} d\sigma = \frac{\prod \Gamma(a_i)}{\Gamma(\sum a_i)}$$

On vérifie donc (s'il en est besoin !) que l'expression que l'on a établie est bien une densité de probabilité.

Inversement à partir de cette formule on peut directement poser la densité de probabilité en question, les calculs qui précèdent montrent qu'elle correspond à l'image de la distribution de DIRICHLET sur le simplexe unité par le changement de variable $s_i = \sqrt{y_i}$.

QUATRIEME PARTIE :

Compléments et applications

INTRODUCTION .

Les considérations précédentes peuvent être poursuivies dans plusieurs directions.

1) En premier le concept de partage se trouve lié à un grand nombre de domaines. Il suffit de rappeler que l'on a traité des exemples de partage électoral, de découpage de territoire, d'analyses de sérum ou de dépôts de pollen. Les références qui traitent de partages envisagent les découpages de budget-temps, de zones urbaines, d'activités de production dans l'entreprise, de régimes alimentaires, de minéraux, de composition chimique d'ossements de rats, du recouvrement de carapaces de tortues... (réf.57). La liste peut être longue, elle s'applique également à tout ce qui peut être considéré comme un "compromis" de points dans un espace affine (réf.63). Par conséquent toute formalisation d'un partage trouve un vaste champ d'applications pratiques.

2) Un autre champ, probabiliste et statistique est celui qui est lié au découpage du segment $[0, 1]$ par la distribution uniforme. Dans la première partie, on a vu qu'un tel mode de découpage conduit à la distribution uniforme sur tout simplexe unité de \mathbb{R}^n , et que l'on peut en déduire un grand nombre de statistiques dont on connaît les distributions. Par conséquent on dispose d'un grand nombre de tests possibles basés sur le "spacing" ou découpage aléatoire de $[0, 1]$ par la variable qui suit la loi uniforme. Ces tests sont rappelés dans PYKE (ref. 66). Outre les tests fondés sur les parts extrémales, d'autres sont basés sur les distributions de sommes

$$\phi(j) = \sum_{i=1}^n \psi(j_i) \quad \text{avec } \psi \text{ application de } [0, 1] \text{ dans } \mathbb{R}^1,$$

à la suite de la communication de GREENWOOD (réf. 35). Les principales applications ψ que l'on rencontre dans la littérature ont les expressions suivantes, (ref. 19 , 66) :

$\varphi(x) = \log x$; $1/x$ étudiés par DARLING

$\varphi(x) = x^2$ étudié par MORAN

$\varphi(x) = x^\alpha$, $\alpha > 0$ étudié par KIMBALL, WEISS

$\varphi(x) = (x - 1/M)^2$ étudié par KIMBALL

$\varphi(x) = |x - 1/M|$ étudié par SHERMAN

D'autres tests sont dus aux travaux de WEISS et de BLUMENTHAL (réf. 10). Pour situer leur importance pratique, il suffit de rappeler que les distributions uniformes sur $[0, 1]$ ou sur tout simplexe unité se rencontrent fréquemment en statistique (réf. 66,84).

3) Les parts extrémales d'un partage sont utilisées dans des tests ci-dessus relatifs au découpage uniforme de $[0, 1]$. On les retrouve avec le point de vue de la théorie de l'information. En effet, par rapport au partage égalitaire $\frac{1}{M}, \frac{1}{M}, \dots, \frac{1}{M}$, tout partage y_1, \dots, y_M présente le gain d'information de SHANNON égal à $\log M + \sum_{j=1}^M y_j \log y_j$ (réf. 49). Si l'on ne connaît qu'une seule part, mettons y_k , le nouveau partage $y_k, (1-y_k)/(M-1), \dots, (1-y_k)/(M-1)$ présente le gain d'information

$$\log M + y_k \log y_k + (1-y_k) \log (1-y_k)/(M-1).$$

Cette quantité est positive, décroissante pour $y_k \in [0, 1/M[$, croissante sur $]1/M, 1]$. Par conséquent, pour tout partage, la part qui apporte à elle seule le plus d'information est soit la plus grande $\sup y_i$ soit la plus petite $\inf y_i$ (en IX.3 nous verrons que c'est la plus grande).

On retrouve également de manière géométrique et analytique le rôle particulier des parts extrémales d'un partage quand on étudie la somme $\sum \varphi(y_i)$ avec φ concave ou convexe. On aborde ci-dessous le cas $\varphi(x) = -x \log x$.

4) Les structures statistiques sur simplexe sont une partie du formalisme qui peut s'appliquer aux partages. On a déjà mis à profit la structure de DIRICHLET pour simplifier la structure multinomiale - DIRICHLET composée de MOSIMANN, ou pour se substituer au besoin à la structure multinomiale ; on mentionne quelques applications possibles sur la distribution gamma et dans le cadre de l'Analyse de la Variance.

On remarque que le paramètre vectoriel α de la distribution de DIRICHLET peut être décomposé sous la forme θA d'un point du produit cartésien $S_1^n \times \mathbb{R}_+^k$. Ceci entraîne l'existence de deux structures statistiques dérivées obtenues en fixant respectivement θ ou A .

IX - APPLICATION DES PARTS ORDONNEES ET EXTREMALES, FONCTIONS DEFINIES
SUR SIMPLEXE UNITE

IX.1 - Etude de ϕ .

Définition : Pour toute application φ de $[0, 1]$ dans \mathcal{R} , on définit l'application ϕ du simplexe unité S_1^n dans \mathcal{R} qui à tout point y de S_1^n fait correspondre la valeur $\phi(y) = \sum_{j=1}^n \varphi(y_j)$.

C'est une application continue et différentielle avec φ , qui est invariante par le groupe des permutations Σ_n .

1.1. Extrémums à l'intérieur du simplexe .

A l'intérieur du simplexe, les extrémums de ϕ sont solutions du système de LAGRANGE :

$$\frac{\partial}{\partial y_i} \left[\phi(y) - \mu \left(\sum_{j=1}^n y_j - 1 \right) \right] = 0 \quad , \quad i=1, \dots, n \quad ,$$

qui prend la forme simplifiée :

$$\varphi'(y_i) - \mu = 0 \quad .$$

Nous nous limitons au cas des fonctions φ concaves ou convexes sur $[0, 1]$, avec une dérivée φ' continue. Dans ces conditions le système de LAGRANGE possède la solution unique $y_j = 1/n$, qui définit le point ω du simplexe unité. Avec φ convexe, on sait que la valeur de $\phi(\omega) = n \varphi(1/n)$ est un minimum global ; à l'inverse $\phi(\omega)$ est un maximum quand φ est concave. Nous nous limitons ci-dessous au cas φ convexe.

Remarque : Toutes les applications φ rappelées dans l'introduction sont effectivement convexes ou concaves, et à dérivée continue (option de SHERMAN excepté).

1.2. Rappels sur les fonctions convexes .

On pose $\mathcal{D}_\alpha^\varphi$ l'ensemble de niveau des points y de S_1^n qui vérifient la condition $\phi(y) \leq \alpha$; on sait que c'est un sous-ensemble convexe, non vide pour tout $\alpha \geq \phi(\omega)$; on rappelle que ϕ est une application convexe sur S_1^n .

1.3. Maximum de ϕ .

Soit e^i le point du simplexe unité qui a toutes ses coordonnées nulles sauf celle de rang i , égale à 1. On rappelle que S_1^n est l'enveloppe convexe des points e^i , $i: 1, \dots, M$ (réf. 63).

Proposition 1 : Le maximum de ϕ sur le simplexe unité est atteint aux points e^i , $i: 1, \dots, M$.

En effet tout point y du simplexe unité est de la forme $y = \sum_{i=1}^M \alpha_i e^i$ avec $\alpha_i \geq 0$, $\sum_1^M \alpha_i = 1$, il en résulte que $\phi(y) \leq \sum \alpha_i \phi(e^i)$.

D'autre part les $\phi(e^i)$ sont tous égaux à $\varphi(1) + (n-1)\varphi(0)$, $i: 1, \dots, M$.

On a donc $\phi(y) \leq \varphi(1) + (n-1)\varphi(0)$, la quantité $\varphi(1) + (n-1)\varphi(0)$

est un maximum global de ϕ atteint en tous les points e^i ■

1.4. Compléments.

Ces propriétés sont également valables pour tout simplexe

$$S_\gamma^n = \left\{ y \in \mathbb{R}^M, y_i \geq 0, \sum_1^M y_i = \gamma \right\}$$

défini pour tout γ positif inférieur ou égal à 1, et pour tout sous-ensemble homothétique de S_1^n par l'homothétie $H(\rho)$ de rapport ρ compris entre zéro et 1, et de centre le point de coordonnées $\gamma/M, \dots, \gamma/M$.

En effet les M points $e^i(\gamma)$ qui ont toutes leurs coordonnées nulles sauf une égale à γ constituent les points extrémaux de S_γ^n ; les M points

$$s^i(\gamma, \rho) = \left(\rho + \frac{1-\rho}{M} \right) e^i(\gamma) + \sum_{j \neq i} \frac{1-\rho}{M} e^j(\gamma)$$

sont les points extrémaux de l'homothétique $H_\rho(S_\gamma^n)$.

Le point de coordonnées r/M appartient à tous ces sous-ensembles, on démontre de la même façon que :

$$\min \phi(y) = M \varphi(r/M)$$

$$\max \phi(y) = (n-1) \varphi\left(r \frac{1-p}{M}\right) + \varphi\left(r \left(p + \frac{1-p}{M}\right)\right)$$

pour tout y appartenant à $H_p(S_r^M)$, quelque soit r positif et p compris entre zéro et un au sens large.

Ces propriétés sont évidemment valables quelle que soit la dimension M supérieure ou égale à 2, et triviales pour $M = 1$.

IX.2 - Bornes de ϕ fonction de λ_M .

On rappelle que les λ_i $i=1, \dots, M$ sont les coordonnées y_i du point y , rangées par ordre décroissant ; on a évidemment

$$\phi(y) = \sum \varphi(\lambda_i)$$

Pour simplifier on note $\lambda(y)$ ou λ la borne supérieure $\sup(y_i)$, et $\varrho(y)$ ou ϱ la borne inférieure $\inf(y_i)$.

Théorème 1 : Soit φ une application définie sur $[0, 1]$ à valeur dans \mathbb{R} , convexe et à dérivée continue. Alors les deux applications réelles b_i et b_s qui ont les valeurs suivantes :

$$b_i(x) = \varphi(x) + (n-1) \varphi\left(\frac{1-x}{n-1}\right) \quad x \in [0, 1/n]$$

$$b_s(x) = (n-1) \varphi(x) + \varphi(1-(n-1)x) \quad x \in [0, 1/n]$$

sont définies continues dérivables décroissantes sur $[0, 1/n]$; pour tout y appartenant au simplexe unité, $b_i(\varrho)$ et $b_s(\varrho)$ sont respectivement des bornes inférieures et supérieures de $\phi(y)$.

En effet soit $t \in \mathbb{R}^{n-1}$ le point de coordonnées λ_i pour tout $i=1, \dots, n-1$,
 et ϕ^0 l'application qui à t fait correspondre $\phi^0(t) = \sum_1^{n-1} \varphi(\lambda_i)$.

• Borne inférieure

Le point t appartient au simplexe $S_{1,0}^{n-1}$, qui contient le point
 $1-0/n-1, \dots, 1-0/n-1$; φ étant convexe le minimum de ϕ^0 égal
 à $(n-1) \varphi(1-0/n-1)$ est atteint en ce point.

On a bien $\phi(y) = \phi^0(t) + \varphi(0) \geq b_1(0)$.

• Borne supérieure

Plus précisément le point t appartient au sous-ensemble de $S_{1,0}^{n-1}$
 des points qui vérifient $\inf_i \phi(t_i) \geq 0$.

On vérifie que le sous-ensemble est l'homothétique du simplexe $S_{1,0}^{n-1}$
 dans l'homothétie de centre $\frac{1-0}{n-1}, \dots, \frac{1-0}{n-1}$ et de rapport
 $\rho = \frac{1-n0}{1-0}$ (cf chapitre I.3).

Le maximum de ϕ^0 est atteint aux points-extrémaux suivants:

$$\begin{aligned} \phi^0\left(1-0, \frac{1-n0}{1-0}\right) &= \left[\frac{1-n0}{1-0} + \frac{1}{n-1}\left(1 - \frac{1-n0}{1-0}\right)\right] e^{i(1-0)} + \sum_{j \neq i} \frac{1}{n-1} \left(1 - \frac{1-n0}{1-0}\right) e^{j(1-0)}, \\ &= \frac{1-(n-1)0}{1-0} e^{i(1-0)} + \sum_{j \neq i} \frac{0}{1-0} e^{j(1-0)} \end{aligned}$$

de \mathbb{R}^{n-1} . On a bien $\phi(y) = \phi^0(t) + \varphi(0) \leq b_n(0)$.

• Propriétés de b_i :

b_i est continue dérivable avec φ ,

$$b_i'(\vartheta) = \varphi'(\vartheta) - \varphi'\left(\frac{1-\vartheta}{n-1}\right)$$

négatif pour $\vartheta < \frac{1-\vartheta}{n-1}$ ou $\vartheta < \frac{1}{n}$; nul pour $\vartheta = \frac{1}{n}$.

Les valeurs aux bornes sont :

$$b_i(0) = \varphi(0) + (n-1) \varphi\left(\frac{1}{n-1}\right) , \quad b_i\left(\frac{1}{n}\right) = n \varphi\left(\frac{1}{n}\right) ,$$

$$b_i'(0) = \varphi'(0) - \varphi'\left(\frac{1}{n-1}\right) , \quad b_i'\left(\frac{1}{n}\right) = 0 .$$

• Propriétés de b_d :

b_d est continue dérivable avec φ ,

$$b_d'(\vartheta) = (n-1) \left[\varphi'(\vartheta) - \varphi'(1-(n-1)\vartheta) \right]$$

négatif pour $\vartheta < 1-(n-1)\vartheta$ ou $\vartheta < \frac{1}{n}$; nul pour $\vartheta = \frac{1}{n}$.

Les valeurs aux bornes sont :

$$b_d(0) = (n-1) \varphi(0) + \varphi(1) , \quad b_d\left(\frac{1}{n}\right) = n \varphi\left(\frac{1}{n}\right) ,$$

$$b_d'(0) = (n-1) (\varphi'(0) - \varphi'(1)) , \quad b_d'\left(\frac{1}{n}\right) = 0 \quad \blacksquare$$

Corollaire : Si φ est concave, les deux bornes b_i et b_d sont des fonctions croissantes de ϑ , (prendre $\psi = -\varphi$ convexe).

Conséquence. On en déduit la double inclusion suivante dans l'ensemble des parties de $S^{\eta-1}$:

$$\left\{ y : b_i(y) \leq u \right\} \subset \mathcal{D}_u^p \subset \left\{ y : b_i(y) \leq u \right\}$$

Ces sous-ensembles étant non vides pour tout u supérieur à $M \varphi(1/M)$.

IX.3 - Bornes de ϕ fonction de λ_i .

3.1. On a un résultat analogue avec $\lambda_i = \Delta_{i,p}(y_i)$.

Théorème 2 Avec la même application φ du théorème précédent, les deux applications réelles B_i et B_Δ qui ont les valeurs suivantes :

$$B_i(x) = \varphi(x) + (\eta-1) \varphi((1-x)/(\eta-1)) \quad , \quad x \in [1/\eta \quad 1] \quad ,$$

$$B_2(x) = \varphi(x) + \varphi(1-x) + (\eta-2) \varphi(0) \quad , \quad x \in [1/2 \quad 1] \quad ,$$

$$B_\Delta(x) = (q+1) \varphi(x) + \varphi(1-(q+1)x) + (\eta-q-2) \varphi(0) \quad ,$$

$$x \in \left] \frac{1}{q+2} \quad \frac{1}{q+1} \right] \quad , \quad q = 1, \dots, \eta-2 \quad ,$$

$$B_0\left(\frac{1}{M}\right) = \eta \varphi\left(\frac{1}{M}\right) \quad ,$$

sont définies continues croissantes sur $[1/M \quad 1]$; pour tout y appartenant au simplexe unité, $B_i(\lambda)$ et $B_\Delta(\lambda)$ sont respectivement des bornes inférieures et supérieures de $\phi(y)$.

En effet, soit $t \in \mathbb{R}^{\eta-1}$ le point de coordonnées λ_i pour tout $i = 1, \dots, \eta$, et $\bar{\phi}$ l'application qui à t fait correspondre $\bar{\phi}(t) = \sum_{i=1}^{\eta} \varphi(\lambda_i)$. Le point t appartient au sous-ensemble

$$\bar{S}_{1-\lambda}^{n-1}(\lambda) = \left\{ t \in \mathbb{R}^{n-1}, 0 \leq t_i \leq \lambda, \sum_{i=1}^n t_i = 1-\lambda \right\}$$

sous-ensemble convexe de $S_{1-\lambda}^{n-1}$ qui contient le point de coordonnées $(1-\lambda)/(n-1), \dots, (1-\lambda)/(n-1)$.

• Borne inférieure

φ étant convexe, $\bar{\varphi}$ atteint son minimum au point de coordonnées $(1-\lambda)/(n-1), \dots, (1-\lambda)/(n-1)$, minimum égal à $(n-1) \varphi((1-\lambda)/(n-1))$. On a bien

$$\phi(y) = \bar{\varphi}(t) + \varphi(\lambda) \geq \beta_1(\lambda) \quad \text{pour tout } y \in S_1^n.$$

• Borne supérieure

i) Pour $\lambda \geq 1/2$, les deux ensembles $\bar{S}_{1-\lambda}^{n-1}(\lambda)$ et $S_{1-\lambda}^{n-1}$ sont identiques, on a donc $\phi(y) \leq \beta_2(\lambda)$ pour tous les y tels que $\lambda \geq 1/2$.

ii) Pour λ appartenant à $]1/n, 1/2]$ on a l'inclusion

stricte $\bar{S}_{1-\lambda}^{n-1}(\lambda) \subset S_{1-\lambda}^{n-1}$. L'ensemble $\bar{S}_{1-\lambda}^{n-1}(\lambda)$ est

un convexe défini par des liaisons et contraintes linéaires. Les points extrémaux possèdent donc $n-2$ contraintes saturées, c'est à dire $n-2$ coordonnées égales à zéro ou λ (réf. 63); ils sont permutablement par le groupe de permutation Σ_{n-1} et ont tous les même nombre de coordonnées égales à λ . Par conséquent, ce nombre est égal au plus grand entier $k(\lambda)$ tel que le produit $\lambda \cdot k(\lambda)$ soit strictement inférieur à $1-\lambda$. Tout point extrémal de $\bar{S}_{1-\lambda}^{n-1}(\lambda)$ possède ainsi $k(\lambda)$ coordonnées égales à λ , $n-2-k(\lambda)$ coordonnées nulles et la dernière égale à $1-\lambda(1+k(\lambda))$, positive éventuellement nulle.

Par construction $k(\lambda) \in [1/n, 1/2]$, avec nécessairement $1 \leq k \leq n-2$,

donc $k(\lambda)$ est constant égal à q sur l'intervalle $I_q =]\frac{1}{q+2}, \frac{1}{q+1}]$

$q = 1, \dots, n-2$.

Le maximum de $\bar{\phi}$ sur le convexe $\bar{S}_{1-\lambda}^{n-1}(\lambda)$ est atteint en tout point extrême, il est égal à $k \varphi(\lambda) + \varphi(1 - (k+1)\lambda) + (n-k-2)\varphi(0)$ pour $\lambda \in \mathbb{I}_k$. On a donc $\phi(y) \leq B_0(\lambda)$ pour tous les y tels que $\lambda \in \mathbb{I}_k$.

iii) Pour $\lambda = 1/M$ on a évidemment $\phi(y) = M \varphi(1/M)$.

• Propriété de B_1 :

B_1 possède la même expression analytique que B_0 ; elle est continue dérivable avec φ ,

$$B_1'(\lambda) = \varphi'(\lambda) - \varphi'\left(\frac{1-\lambda}{n-1}\right) \quad ; \text{ la dérivée est positive pour}$$

$$\lambda > \frac{1-\lambda}{n-1} \text{ ou } \lambda > \frac{1}{n}, \text{ nulle pour } \lambda = \frac{1}{n} ;$$

B_1 est croissante.

Les valeurs aux bornes sont :

$$B_1\left(\frac{1}{M}\right) = M \varphi\left(\frac{1}{M}\right), \quad B_1(1) = \varphi(1) + (n-1)\varphi(0),$$

$$B_1'\left(\frac{1}{n}\right) = 0, \quad B_1'(1) = \varphi'(1) - \varphi'(0).$$

• Propriété de B_q :

On vérifie que B_q est continue en tout $\lambda = 1/n$, $n = 2, \dots, M$.

Sur tout intervalle \mathbb{I}_q , $q = 0, \dots, M-2$, B_q est dérivable ; on a

$$B_q'(\lambda) = (q+1) \left[\varphi'(\lambda) - \varphi'(1 - (q+1)\lambda) \right] \text{ positive sur } \mathbb{I}_q$$

avec $\lim_{\lambda \rightarrow \frac{1}{q+2}} B_q'(\lambda) = 0$; B_q est croissante.

$$\lambda \rightarrow \frac{1}{q+2}$$

Les valeurs en $1/M$ et 1 de B_Δ et B'_Δ sont les mêmes que celles de B_i , B'_i . ■

Corollaire. Si φ est concave, les deux bornes B_i et B_Δ sont des fonctions décroissantes de λ .

Conséquence : On en déduit la double inclusion suivante dans l'ensemble des parties de S_λ^n

$$\{y : B_\Delta(\lambda) \leq u\} \subset \mathcal{D}_u^\varphi \subset \{y : B_i(\lambda) \leq u\}.$$

3.2. Comparaison des bornes.

Les bornes introduites ci-dessus vérifient le

Théorème 3 Pour tout partage y on a :

$$b_i(\varnothing) \leq B_i(\lambda) \leq \phi(y) \leq B_\Delta(\lambda) \leq b_\Delta(\varnothing).$$

• Démonstration de $b_i(\varnothing) \leq B_i(\lambda)$:

i) Les deux fonctions b_i et B_i ont la même expression analytique ;

ii) Pour tout y on a nécessairement $(n-1)\lambda + \varnothing \geq 1$

$$\text{d'où } \lambda \geq \bar{\lambda} = \frac{1-\varnothing}{n-1}, \text{ avec } \varnothing \leq \frac{1}{M} \text{ et } \bar{\lambda} \geq \frac{1}{M}.$$

Il suffit donc de démontrer $b_i(\varnothing) \leq B_i(\bar{\lambda})$, c'est-à-dire

$$\varphi(\varnothing) + (n-1)\varphi\left(\frac{1-\varnothing}{n-1}\right) \leq \varphi\left(\frac{1-\varnothing}{n-1}\right) + (n-1)\varphi\left(1 - \frac{1-\varnothing}{n-1}\right).$$

$$\text{iii) or } \varnothing + (n-2)\frac{1-\varnothing}{n-1} = (n-1)\left(1 - \frac{1-\varnothing}{n-1}\right),$$

l'application φ étant convexe l'inégalité cherchée en découle.

• Démonstration de $B_\lambda(\lambda) \leq b_\lambda(0)$:

i) Pour tout $\lambda \geq 1/M$ on a évidemment

$$B_\lambda(\lambda) \leq (n-1) \varphi(\lambda) + \varphi(1 - (n-1)\lambda),$$

le deuxième membre étant la valeur prise par l'application b_λ dont l'intervalle de définition est étendu à $[0, 1]$.

ii) Pour tout y on a nécessairement $\lambda + (n-1)\vartheta \leq 1$,

d'où $\lambda \leq \lambda^0 = 1 - (n-1)\vartheta$, avec $\vartheta \leq \frac{1}{M}$ et $\lambda^0 \geq \frac{1}{n}$.

Il suffit de démontrer $b_\lambda(\lambda^0) \leq b_\lambda(\vartheta)$, c'est-à-dire

$$(n-1) \varphi(1 - (n-1)\vartheta) + \varphi(1 + (n-1)^2\vartheta - (n-1)) \leq (n-1) \varphi(\vartheta) + \varphi(1 - (n-1)\vartheta)$$

iii) Or

$$(n-2) [1 - (n-1)\vartheta] + [1 + (n-1)^2\vartheta - (n-1)] = (n-1)\vartheta$$

φ étant convexe l'inégalité cherchée en découle. ■

Conséquence :

On en déduit la succession d'inclusions suivantes dans l'ensemble des parties de S_1^n :

$$\{y: b_\lambda(\vartheta) \leq u\} \subset \{y: B_\lambda(\lambda) \leq u\} \subset \mathcal{O}_u^r \subset \{y: B_\lambda(\lambda) \leq u\} \subset \{y: b_\lambda(\vartheta) \leq u\}.$$

Remarque : On peut poursuivre, et montrer qu'il existe des bornes inférieures et supérieures de ϕ qui dépendent des parts ordonnées λ_1 et λ_2 , croissantes en λ_1 et λ_2 telles que

$$B_\lambda(\lambda) \leq B_\lambda(\lambda_1, \lambda_2) \leq \phi(y) \leq B_\lambda(\lambda_1, \lambda_2) \leq B_\lambda(\lambda).$$

3.3. Compléments.

3.3.1. B_1 possède une expression analytique unique sur le intervalle $[\frac{1}{M}, 1]$, tandis que B_Δ change d'expression à chaque intervalle $\frac{1}{q}$ $q = 0, \dots, M-2$.

On pose $B_\Delta^*(\lambda) = M \varphi(0) + \frac{\varphi(\lambda) - \varphi(0)}{\lambda}$ définie sur l'intervalle $[\frac{1}{M}, 1]$.

On vérifie que B_Δ^* prend les mêmes valeurs que B_Δ en tout point $\lambda = 1/q$:

$$\begin{aligned} B_\Delta^*\left(\frac{1}{q}\right) &= M \varphi(0) + q \left(\varphi\left(\frac{1}{q}\right) - \varphi(0) \right) , \\ &= B_\Delta\left(\frac{1}{q}\right) \quad q = 1, \dots, M. \end{aligned}$$

3.3.2. Propriétés de B_Δ^* . B_Δ^* est continue dérivable avec φ ,

$$\begin{aligned} B_\Delta^{*'}(\lambda) &= \frac{\varphi'(\lambda)}{\lambda} - \frac{\varphi(\lambda) - \varphi(0)}{\lambda^2} \\ &= \frac{1}{\lambda} \left[\varphi'(\lambda) - \frac{\varphi(\lambda) - \varphi(0)}{\lambda} \right] . \end{aligned}$$

Avec φ convexe, on en déduit que $B_\Delta^{*'}$ est positive ou nulle sur $[\frac{1}{M}, 1]$, (réf. 14) ; B_Δ^* est donc croissante sur $[\frac{1}{M}, 1]$ avec les valeurs aux bornes :

$$B_\Delta^*\left(\frac{1}{M}\right) = B_\Delta\left(\frac{1}{M}\right) = B_1\left(\frac{1}{M}\right) , \quad B_\Delta^*(1) = B_\Delta(1) = B_1(1) .$$

Proposition 2 On a $B_\Delta^*(\lambda) \geq B_\Delta(\lambda)$ pour tout λ de l'intervalle $[\frac{1}{M}, 1]$.

En effet :

$$\begin{aligned}
 B_0^*(\lambda) - B_0(\lambda) &= \frac{\varphi(\lambda) - \varphi(0)}{\lambda} - q(\varphi(\lambda) - \varphi(0)) - [\varphi(1-q\lambda) - \varphi(0)] \\
 &= [\varphi(\lambda) - \varphi(0)] \times \frac{1-q}{\lambda} - [\varphi(1-q\lambda) - \varphi(0)] \\
 &= \frac{1-q}{\lambda} \varphi(\lambda) + \frac{\lambda-1+q\lambda}{\lambda} \varphi(0) \\
 &\quad - \varphi(1-q\lambda) .
 \end{aligned}$$

Soient $t = (1-q\lambda)/\lambda$ et $1-t = (\lambda-1+q\lambda)/\lambda$

positifs ou nuls sur I_{q-1} , φ étant convexe on a

$$t \varphi(\lambda) + (1-t) \varphi(0) \geq \varphi(\lambda t) = \varphi(1-q\lambda)$$

d'où la conclusion. ■

On en déduit donc (au besoin) une fonction majorante B_0^* croissante, qui possède une expression analytique unique sur $[\frac{1}{n}, 1]$.

3.4. Exemples .

3.4.1. Soit $\varphi(x) = x^\alpha$, $\alpha > 1$, le problème de KIMBALL, (réf. 19, 63).

On en déduit

$$B_0(\lambda) = \lambda^\alpha + (n-1) \left(\frac{1-\lambda}{n-1} \right)^\alpha ,$$

$$B_q(\lambda) = q \lambda^\alpha + (1-q\lambda)^\alpha , \quad \lambda \in I_{q-1} , \quad q = 1, \dots, n-1 ,$$

$$B_0^*(\lambda) = \lambda^{\alpha-1} .$$

Dans les figures ci-après, on a représenté ces applications pour $\alpha = 2$ (le cas étudié par P.A.P. MORAN), et pour diverses valeurs de M .

3.4.2. Soit $\psi(x) = -x \log x$, concave positive sur $]0, 1[$.

Cette application de $]0, 1[$ dans \mathcal{R} se rencontre dans la théorie de l'information de C.E. SHANNON (réf. 49,68). La quantité correspondant à $\psi(y)$ est l'entropie de la distribution des y_i , et possède également des propriétés intéressantes pour décrire la dispersion de la distribution (réf. 23,30). L'application ψ étant concave, les bornes B_i^{ψ} et B_{Δ}^{ψ} sont décroissantes, depuis $\log M$ jusqu'à 0. La fonction complémentaire est ici $B_i^{\psi*}$ décroissante entre les mêmes valeurs. On a

$$B_{\Delta}^{\psi}(\lambda) = -\lambda \log \lambda - (1-\lambda) \log (1-\lambda) / (M-1),$$

$$B_i^{\psi}(\lambda) = -q \lambda \log \lambda - (1-q\lambda) \log (1-q\lambda) \quad \lambda \in]\frac{1}{M}, 1[,$$

$$B_i^{\psi*}(\lambda) = -\log \lambda.$$

Ces applications sont représentées dans les figures ci-contre, pour diverses valeurs de M .

IX.4 - Application aux probabilités.

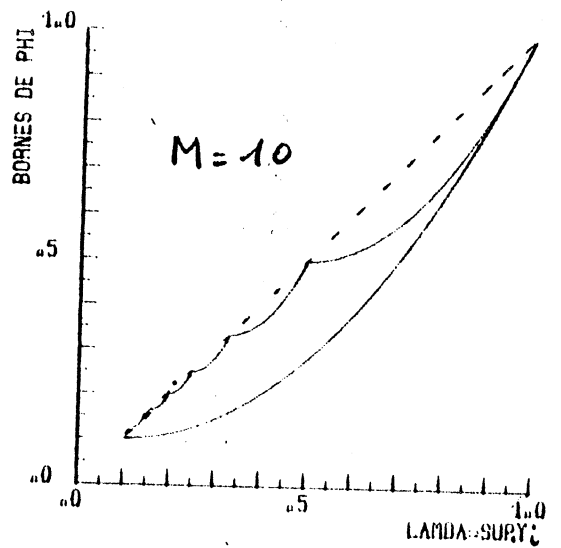
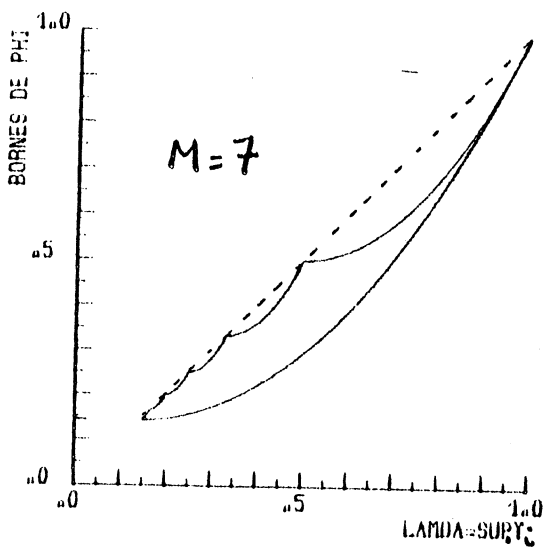
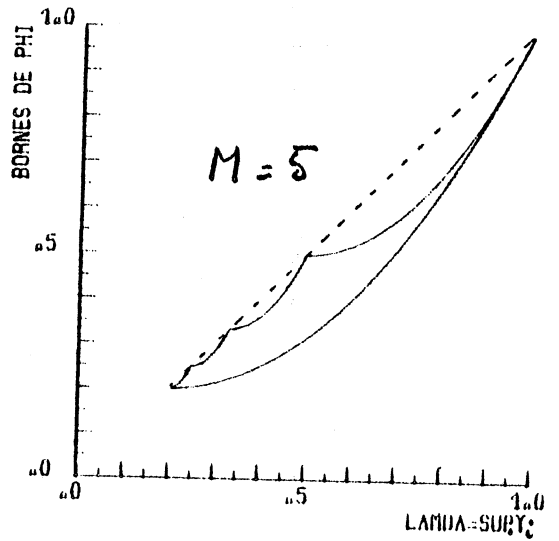
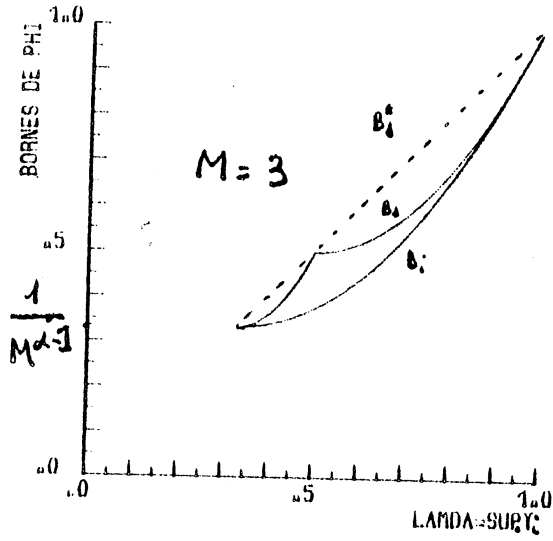
On suppose que y est un point aléatoire du simplexe unité ; il en résulte que $\lambda(y)$, $\rho(y)$ et $\phi(y)$ sont des variables aléatoires. Les relations d'inclusion précédentes entraînent que l'on a

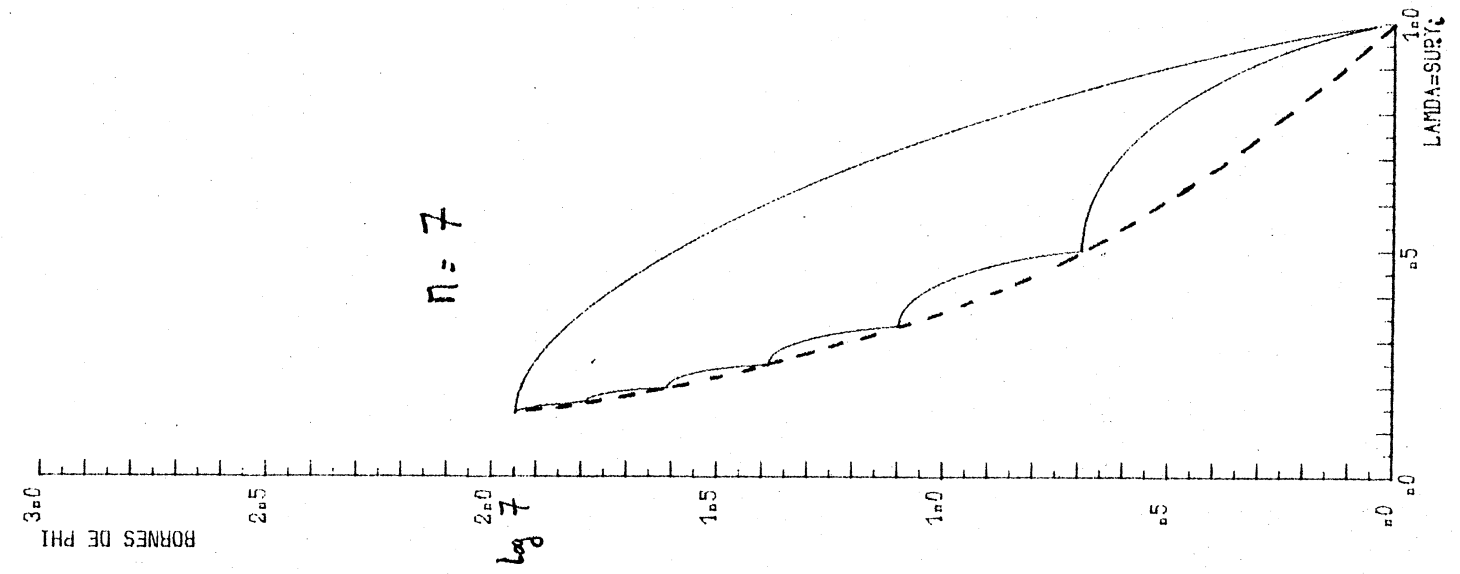
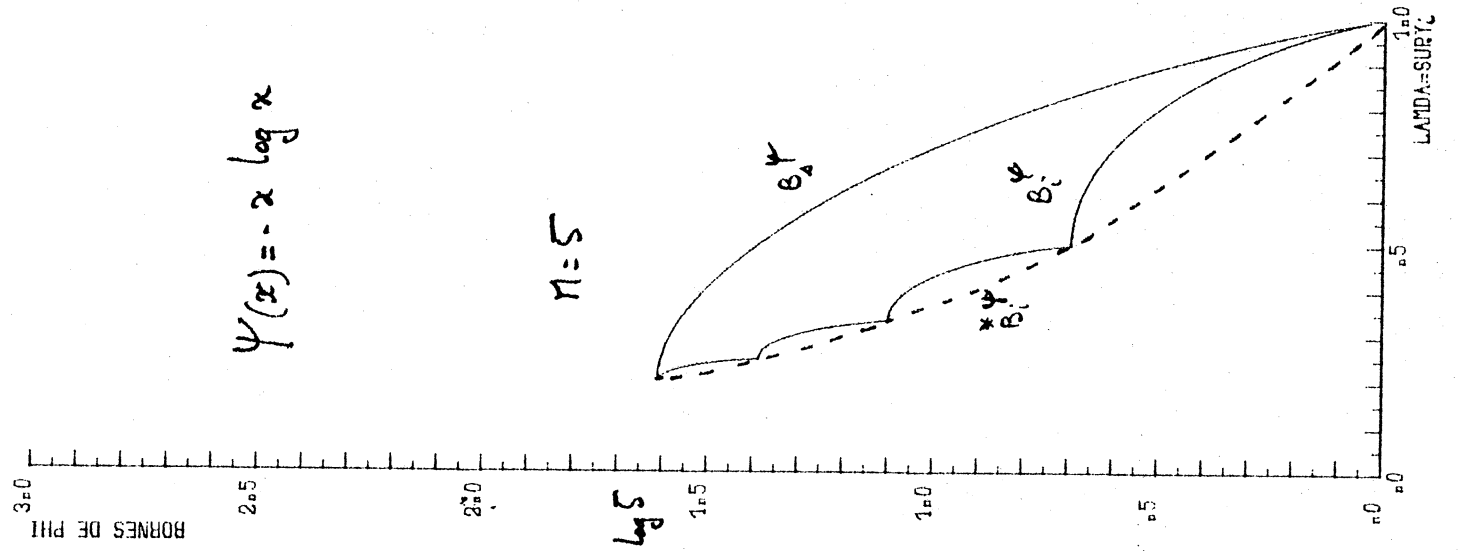
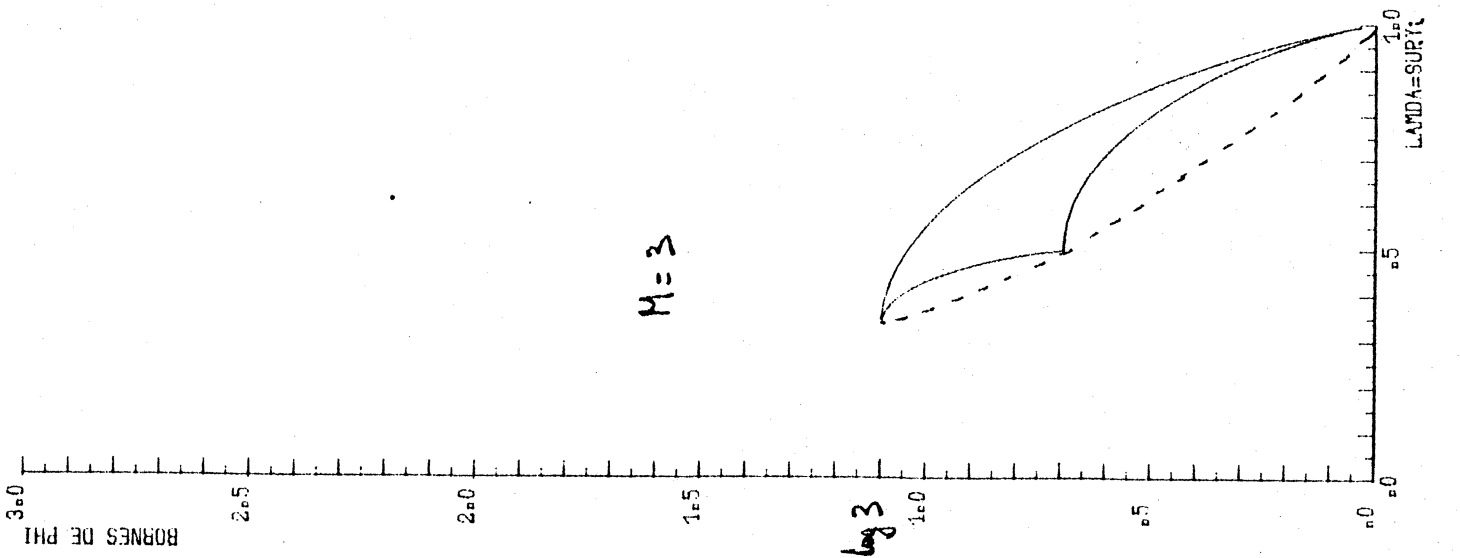
$$\text{Prob} \{ y : B_{\Delta}(\lambda) \leq u \} \leq \text{Prob} \{ \omega_u^{\psi} \} \leq \text{Prob} \{ y : B_i(\lambda) \leq u \},$$

ce qui donne des bornes (atteintes) pour $\text{Prob}(\omega_u^{\psi})$.

$$\varphi(x) = x^\alpha, \quad \alpha = 2 \quad (\text{MORAN}).$$

Applications $B_i(\lambda)$, $B_0(\lambda)$, $B_i^*(\lambda)$.





$$\psi(x) = -\alpha \log x$$

4.1. Cas de la mesure uniforme .

C'est le cas étudié dans les chapitres II et III. On rappelle que :

$$\text{Prob} \{ y: \lambda \leq x \} = L_H^\lambda(x) = \sum_{k=0}^{[1/x]} (-1)^k C_M^k (1-kx)^{M-1} ,$$

d'où les inégalités suivantes :

$$L_H^\lambda(B_d^{-1}(u)) \leq \text{Prob}(\mathcal{D}_u^\Psi) \leq L_H^\lambda(B_i^{-1}(u)) .$$

On a ainsi deux probabilités qui sont respectivement bornes inférieures et supérieures de $\text{Prob}(\mathcal{D}_u^\Psi)$, elles ont des variations paramétrées évidentes en fonction du paramètre t variant sur $[\frac{1}{M}, 1]$:

$$\text{i) } \quad \pi(u) = L_H^\lambda(B_d^{-1}(u)) \quad \Leftrightarrow \quad \begin{cases} u = B_d(t) \\ \eta = L_H^\lambda(t) \end{cases}$$

$$\text{ii) } \quad \pi(u) = L_H^\lambda(B_i^{-1}(u)) \quad \Leftrightarrow \quad \begin{cases} u = B_i(t) \\ \eta = L_H^\lambda(t) \end{cases}$$

Les figures ci-après illustrent ces variations, pour les applications

$$\Psi(x) = x^2 \quad \text{et} \quad \Psi(x) = -x \log x .$$

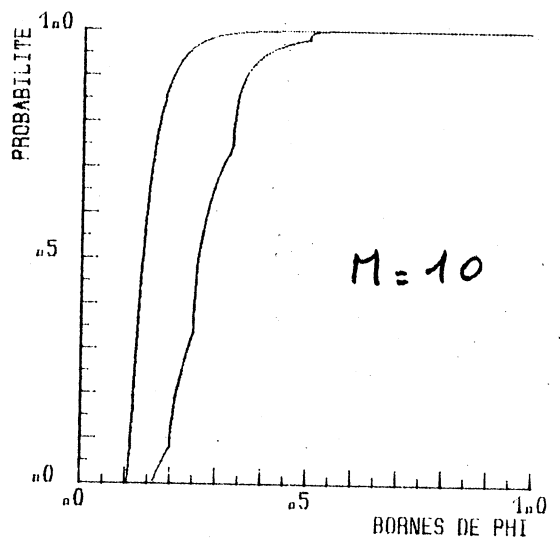
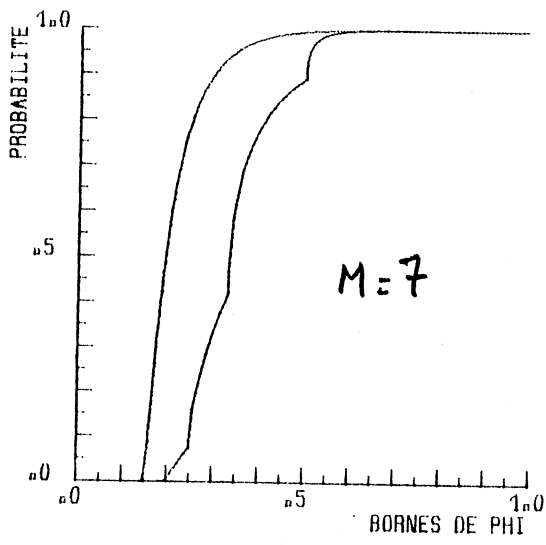
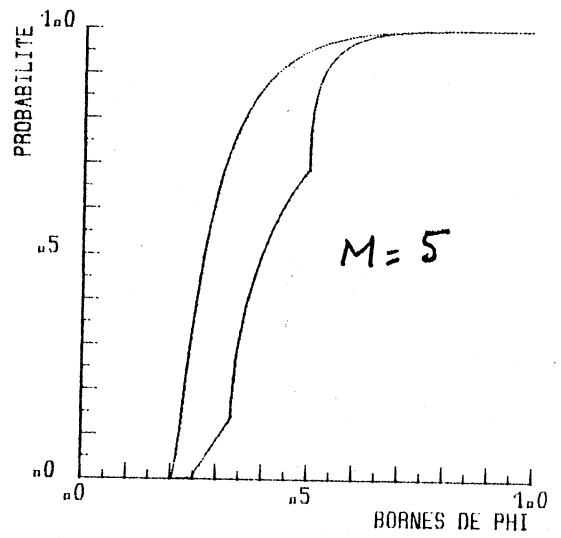
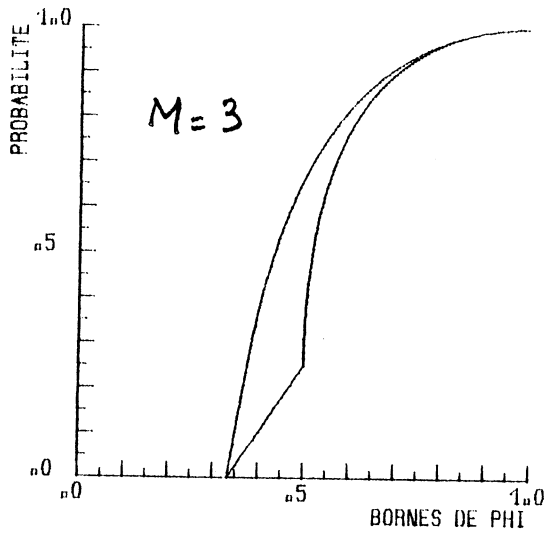
4.2. Autres mesures sur S_A^H .

La relation

$$\text{Prob} \{ y: B_d(\lambda) \leq u \} \leq \text{Prob} \{ \mathcal{D}_u^\Psi \} \leq \text{Prob} \{ y: B_i(\lambda) \leq u \}$$

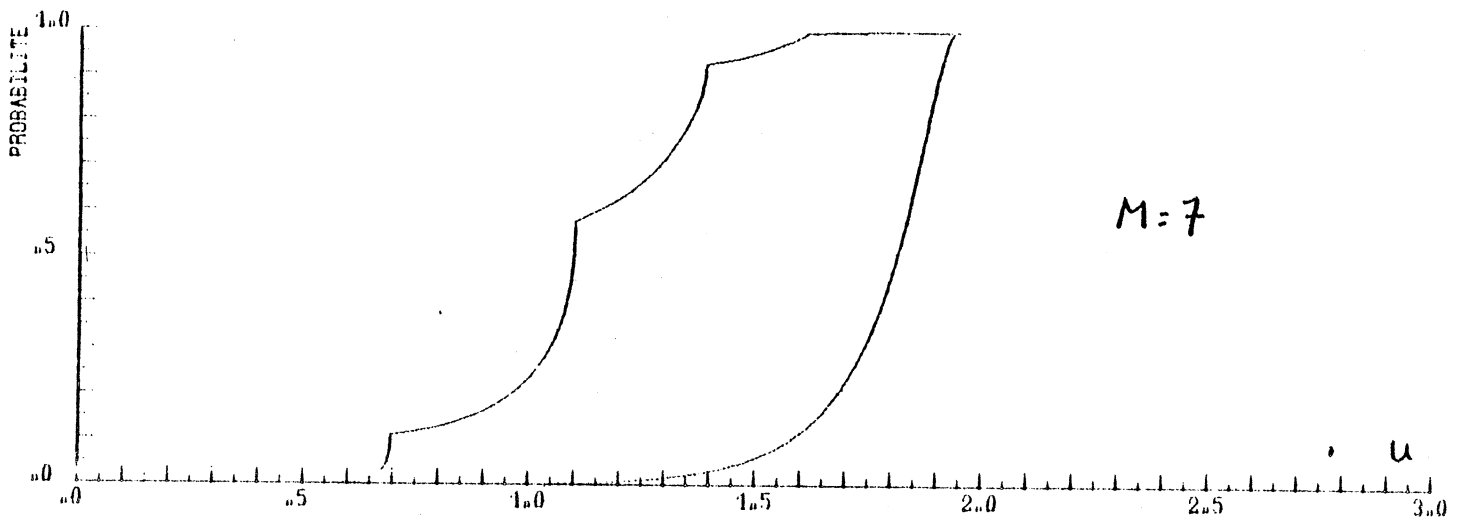
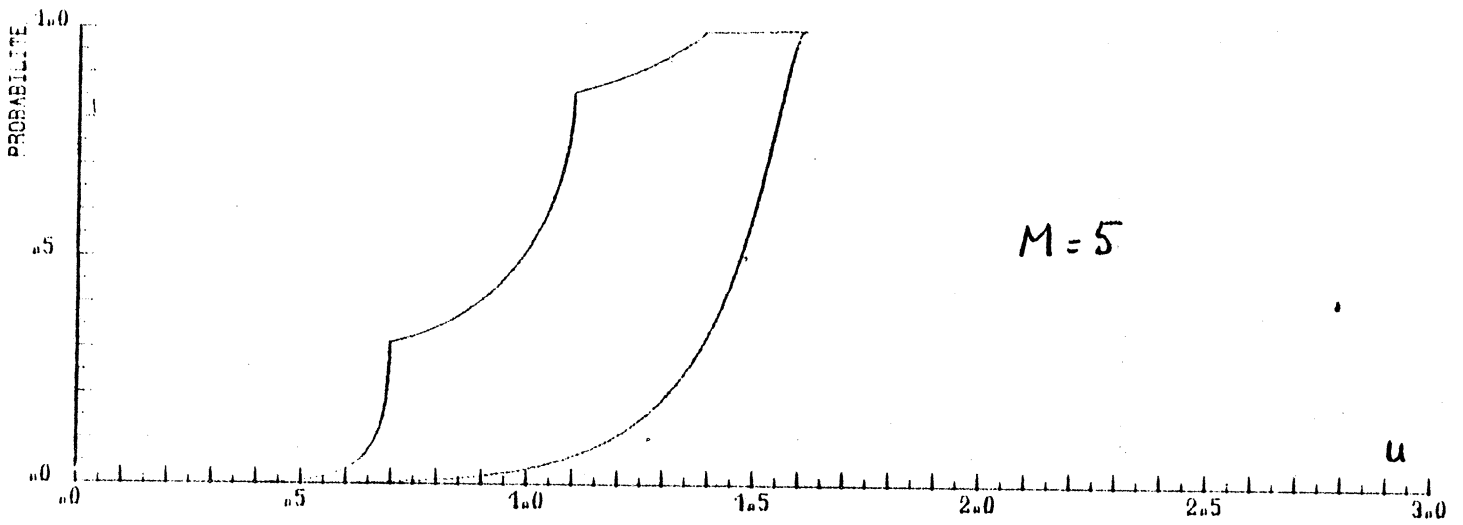
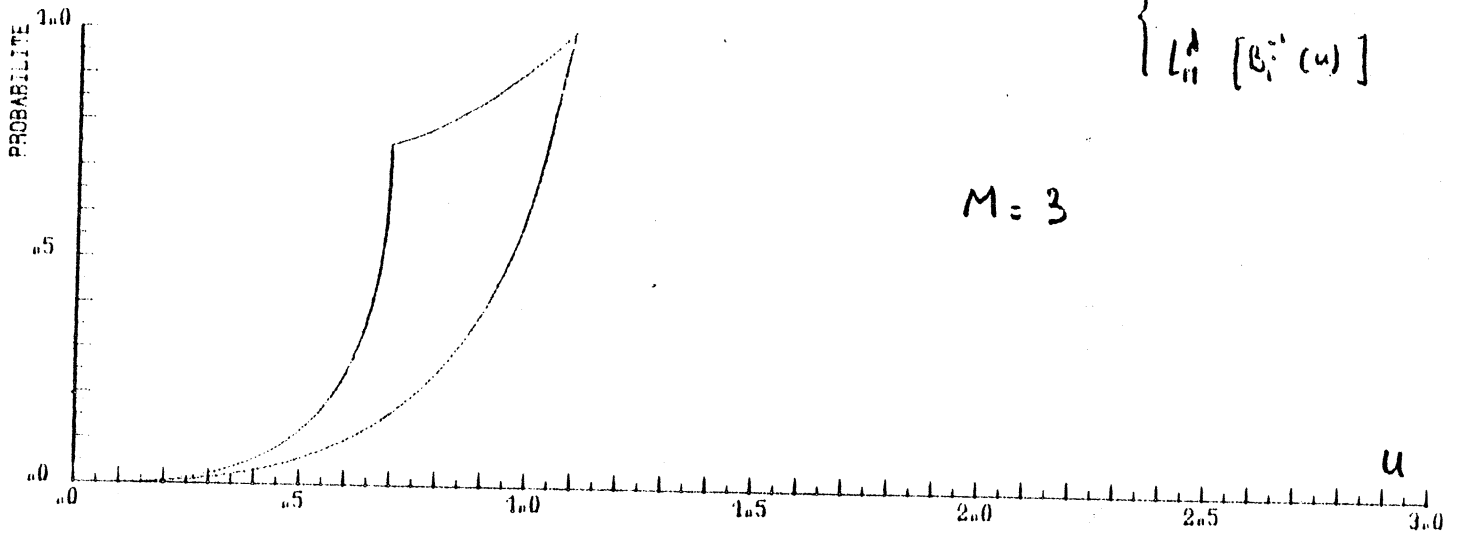
$\varphi(x) = x^2$ (cas de MORAN)

$L_n^\lambda [B_\delta^{-1}(u)]$ et $L_n^\lambda [B_i^{-1}(u)]$



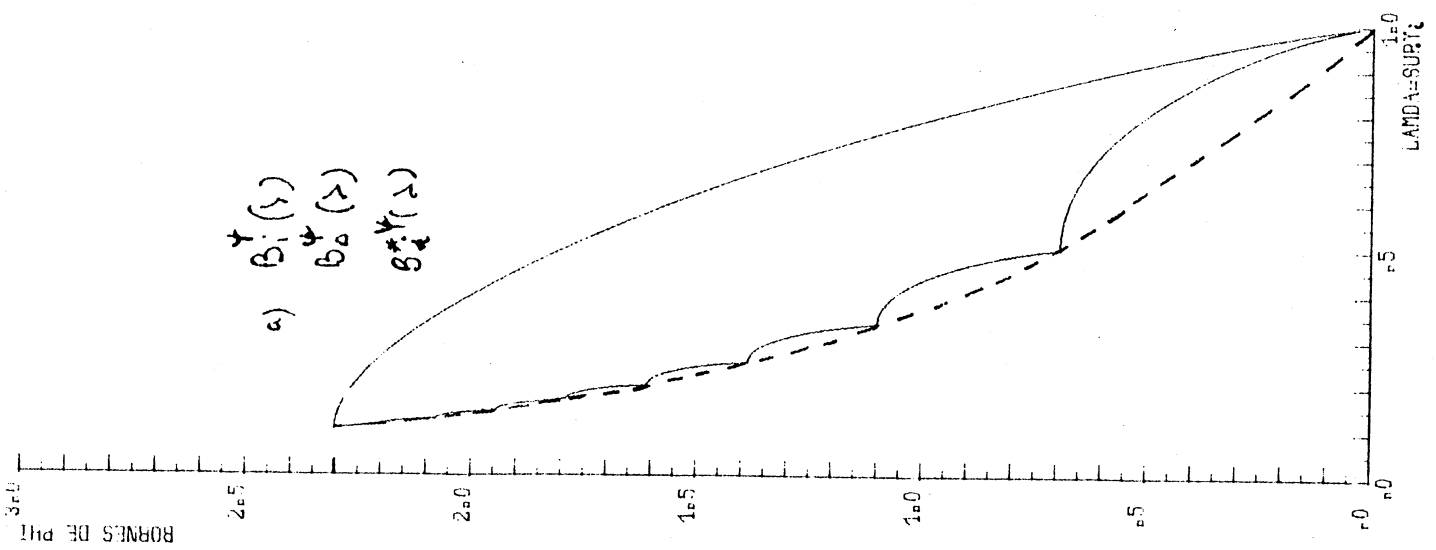
$$\psi(x) = -x \log x$$

$$\begin{cases} L_n^{\Delta} [B_{\delta}^{-1}(u)] \\ L_n^{\Delta} [B_{\delta}^{-1}(u)] \end{cases}$$

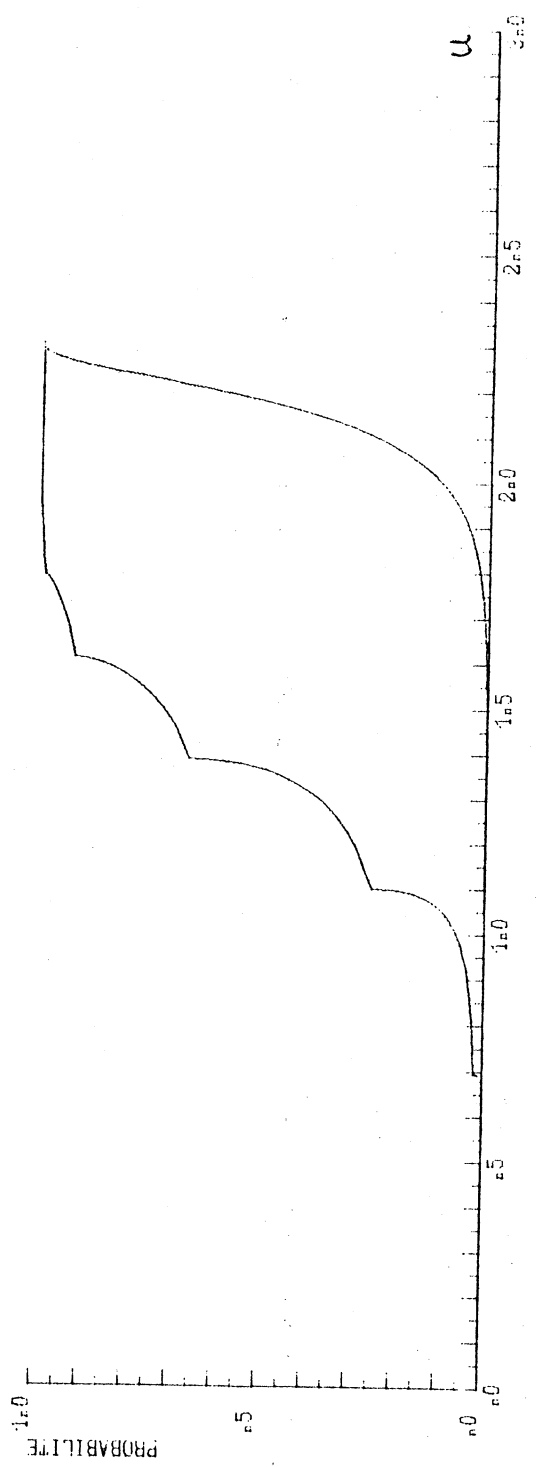


$$\psi(x) = -x \log x, \quad M = 10$$

a) $B_1(\lambda)$
 $B_0(\lambda)$
 $B_2^*(\lambda)$



b) $L_M^A [B_0^*(u)]$
 $L_M^A [B_1^*(u)]$



est valable pour toute probabilité sur le simplexe unité. En pratique les distributions des deux bornes de Prob $\{ \mathcal{D}P_u \}$ sont connues pour autant que l'on connaisse la distribution de $\sup y_i$.

4.3. Exemple .

Nous reprenons le partage du papyrus Rhind, les 2 applications $\varphi(x) = x^2$,

$\Psi(x) = -x \log x$ et la distribution uniforme sur le simplexe unité S_1^5 .

Le partage est le suivant $\frac{46}{120}, \frac{35}{120}, \frac{24}{120}, \frac{13}{120}, \frac{2}{120}$ avec $\lambda = 0,3833$

et $M = 5$.

i) Avec $\varphi(x) = x^2$ on a $\phi = 0,2840$, on en déduit :

$$0,060 \leq \text{Prob} \{ \phi \leq 0,2840 \} \leq 0,574 ;$$

ii) Avec $\Psi(x) = -x \log x$ on a $\Psi = 1,3578$, on en déduit :

$$0,265 \leq \text{Prob} \{ \Psi \leq 1,3578 \} \leq 0,971.$$

4.4. Etude du cas $-x \log x$.

4.4.1. Généralités .

D.A. DARLING a étudié de façon générale la distribution de $\phi(y)$ lorsque le point y est un point aléatoire du simplexe unité avec la distribution uniforme ; il a établi la fonction caractéristique de cette distribution et en déduit notamment les deux premiers moments (réf. 19).

$$m_1 = M(M-1) \int_0^1 (1-r)^{M-2} \varphi(r) dr$$

$$m_2 = M(M-1) \int_0^1 (1-r)^{M-2} \varphi^2(r) dr$$

$$+ M(M-1)^2 \iint_{\substack{r_1 \geq 0 \\ r_2 \geq 0}} (1-r_1-r_2)^{M-3} \varphi(r_1) \varphi(r_2) dr_1 dr_2 .$$

Dans le cas présent, le calcul de ces expressions est une conséquence immédiate des résultats établis pour la densité générale de DIRICHLET $D(a)$ du chapitre VII.2. Nous utilisons le fait que $\partial \log D / \partial a_i$ est un point aléatoire centré pour $D(a)$.

4.4.2. Proposition 3 Avec l'application $\varphi(z) = -z \log z$, les deux premiers moments de la distribution de ϕ sont :

$$\text{moyenne } \phi = \psi(n+1) - \psi(2)$$

$$\text{variance } \phi = \frac{2 \psi'(3)}{n+1} - \psi'(n+2) + \frac{n-1}{2(n+1)^2} \quad (*)$$

• La moyenne :

On rappelle que $\frac{\partial \log D(y; a)}{\partial a_i} = \psi(A) - \psi(a_i) + \log y_i$, $i=1, \dots, M$

avec $A = \sum_{k=1}^M a_k$, d'où

$$\psi(A) - \psi(a_i) = - \frac{\Gamma(A)}{\Gamma(a_i)} \int \log y_i \prod_{j=1}^M y_j^{a_j-1} dy_1 \dots dy_{n-1}$$

On prend le cas particulier $M=2$, $a_1=a$, $a_2=b$ (cas particulier de la distribution Béta (a, b)) :

$$\psi(a+b) - \psi(a) = - \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \int_0^1 \log y y^{a-1} (1-y)^{b-1} dy$$

soit avec $a=2$, $b=n-1$:

$$- n(n-1) \int_0^1 y \log y (1-y)^{n-2} dy = \psi(n+1) - \psi(2).$$

(*) Dans ce calcul les notations ψ, ψ' désignent les fonctions polygamma.

. Le moment d'ordre 2 :

$$i) \text{ On a encore } \int \left(\frac{\partial \log D}{\partial a_i} \right)^2 D(y; a) dy = - \int \frac{\partial^2 \log D}{\partial a_i^2} D(y; a) dy ,$$

soit avec la distribution Béta (a, b) :

$$\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \int_0^1 y^{a-1} (1-y)^{b-1} \log^2 y dy = \psi'(a) - \psi'(a+b) + (\psi(a+b) - \psi(a))^2 ,$$

soit en prenant $a = 3$, $b = n-1$:

$$n(n-1) \int_0^1 y^2 \log^2 y (1-y)^{n-2} dy = \frac{2}{n+1} \left[\psi'(3) - \psi'(n+2) + (\psi(n+2) - \psi(3))^2 \right].$$

$$ii) \text{ On a aussi } \int \frac{\partial \log D}{\partial a_i} \frac{\partial \log D}{\partial a_j} D(y; a) dy = - \int \frac{\partial^2 \log D}{\partial a_i \partial a_j} D(y; a) dy$$

d'où

$$\int \log y_i \log y_j D(y; a) dy = [\psi(a_i) - \psi(a_j)] [\psi(a_i) - \psi(a_j)] - \psi'(a_i) .$$

Si l'on prend $M = 3$, $a_1 = 2$, $a_2 = 2$, $a_3 = n-2$:

$$(n+1)n(n-1) \int y \log y \log z (1-y-z)^{n-3} dy dz = (\psi(n+2) - \psi(2))^2 - \psi'(n+2) ,$$

c'est à dire

$$n(n-1)(n-2) \int y \log y \log z (1-y-z)^{n-3} dy dz = \frac{n-1}{n+1} \left[(\psi(n+2) - \psi(2))^2 - \psi'(n+2) \right].$$

• La variance :

$$\begin{aligned} \text{On en déduit } \sigma^2_{\hat{\beta}} &= \frac{2}{n+1} \psi'(3) - \psi'(n+2) \\ &+ \frac{2}{n+1} (\psi(n+2) - \psi(3))^2 + \frac{n-1}{n+1} (\psi(n+2) - \psi(2))^2 \\ &- (\psi(n+1) - \psi(2))^2. \end{aligned}$$

$$\begin{aligned} \sigma^2_{\hat{\beta}} &= \frac{2}{n+1} \psi'(3) - \psi'(n+2) \\ &+ \psi^2(n+2) - \psi^2(n+1) + \frac{2}{n+1} (\psi^2(3) - \psi^2(2)) \\ &- 2\psi(2)[\psi(n+2) - \psi(n+1)] - \frac{4\psi(n+2)}{n+1} (\psi(3) - \psi(2)). \end{aligned}$$

Sachant que $\psi(n+1) = -\gamma + \sum_{l=1}^n 1/l^2$,

la dernière ligne

s'écrit : $-2[\psi(2) + \psi(n+2)] / (n+1)$,

La ligne intermédiaire s'écrit :

$$\frac{\psi(2) + \psi(3) + \psi(n+1) + \psi(n+2)}{n+1}$$

d'où

$$\sigma^2_{\hat{\beta}} = \frac{2\psi'(3)}{n+1} - \psi'(n+2) + \frac{n-1}{2(n+1)^2}$$

X - APPLICATION DE LA DISTRIBUTION DE DIRICHLET

X.1 - Distributions gamma et DIRICHLET invariante.

La distribution de DIRICHLET invariante intervient à propos de la distribution gamma définie sur \mathbb{R}_+ . On la désigne par $\Gamma(\alpha, \lambda)$, sa densité a pour

$$\text{expression } \alpha > 0, \lambda > 0, \quad f_{\alpha, \lambda}(x) = \lambda^\alpha x^{\alpha-1} e^{-\lambda x} / \Gamma(\alpha) \times \mathbb{1}_{\mathbb{R}_+}(x).$$

Le paramètre de forme α et le paramètre de taille λ sont reliés aux premiers moments de la loi par les relations $\alpha = m_1^2 / \sigma^2$, $\lambda = m_1 / \sigma^2$.

1.1. Estimation sur α .

1.1.1. Rappels. Soit $L(x_1, \dots, x_N) = \prod_{i=1}^N f_{\alpha, \lambda}(x_i)$ la vraisemblance d'un échantillon de N observations indépendantes ; on a

$$\log L = (\alpha - 1) \sum \log x_i - \lambda \sum x_i + N \alpha \log \lambda - N \log \Gamma(\alpha).$$

On pose les moyennes $\bar{m} = \sum x_i / N$, $\bar{j} = (\prod x_i)^{1/N}$, m

et $\log \bar{j}$ sont des statistiques exhaustives pour les paramètres naturels λ et $\alpha - 1$. On a le système du maximum de vraisemblance :

$$\begin{cases} \frac{\partial \log L}{\partial \alpha} = N (\log \bar{j} - \Psi(\alpha) + \log \lambda) \\ \frac{\partial \log L}{\partial \lambda} = -N (\bar{m} - \alpha / \lambda) \end{cases}$$

Les estimateurs $\hat{\alpha}$ et $\hat{\lambda}$ du maximum de vraisemblance sont solutions de $\hat{\alpha} = \hat{\lambda} \bar{m}$ et $\Psi(\hat{\alpha}) - \log \hat{\lambda} = \log \bar{j}$; en particulier l'estimateur $\hat{\alpha}$ est solution de l'équation :

$$\Psi(\alpha) - \log \alpha = \log \bar{j} / \bar{m}.$$

Remarques 1) Avec les nouvelles variables α et $\mu = \alpha / \lambda$ on a $\hat{\mu} = \bar{m}$, c'est-à-dire $(\alpha / \hat{\lambda}) = \hat{\alpha} / \hat{\lambda}$;

2) On sait que \bar{m} et $\log \bar{g}/\bar{m}$ sont des statistiques indépendantes (COX et LEWIS, rappelé dans réf. 4),

3) On a
$$\psi(x) - \log x = -\frac{1}{2x} - 2 \int_0^{\infty} \frac{t dt}{(x^2+t^2)(e^{2xt}-1)}, \quad (\text{ref. 1}),$$

l'application $\psi(\cdot) - \log(\cdot)$ est négative décroissante sur \mathcal{R}_+^* avec les valeurs limites suivantes: $\lim_{x \rightarrow +\infty} \psi x - \log x = 0^-$, $\lim_{x \rightarrow 0^+} \psi x - \log x = -\infty$.

On a donc une solution réelle positive finie $\hat{\alpha}$ unique, avec une probabilité égale à 1 (c'est-à-dire $\bar{g} < \bar{m}$).

1.1.2. Pour estimer α on peut utiliser concurremment les résultats dus à la distribution de DIRICHLET invariante.

En effet soient x_1, \dots, x_M des observations indépendantes de la loi $\Gamma(\alpha, \lambda)$ de somme $A = \sum x_i$. On a vu que le point y de coordonnées x_i/A suit la distribution de DIRICHLET $D(1)$ sur le simplexe unité S_1^M .

Pour estimer α on peut donc par exemple diviser l'échantillon de N observations x_i en P blocs de M observations ($N = P \times M$), calculer sur chaque bloc les

$$y_i^p = x_i^p / \sum_1^M x_k^p, \quad i = 1, \dots, M, \quad p = 1, \dots, P,$$

et résoudre l'équation

$$\psi(M\alpha) - \psi(\alpha) = -\frac{1}{M \cdot P} \sum_p \sum_i \log y_i^p$$

du chapitre VI.

1.2. Loi du produit des observations indépendantes.

On sait que la somme $A = \sum_1^N x_i$ suit la loi $\Gamma(N\alpha, \lambda)$.

On note w le produit $\prod_1^N x_i$.

Sachant que le point de coordonnées $y_i = \frac{z_i}{\Delta}$ $i=1, \dots, M$ suit sur S_1^M la distribution $D(\alpha)$, le produit $\sigma_N = \prod_1^N y_i = \omega / \Delta^N$ suit la loi dont la densité a pour valeur, pour tout σ réel positif

$$\varphi(\alpha, N, \sigma) = - \frac{\partial}{\partial \sigma} I(\alpha, N, \sigma)$$

avec

$$I(\alpha, N, \sigma) = \int_{V(\sigma)} D(y; \alpha) dy.$$

(notations du chapitre VI).

Proposition 4 La densité de la distribution du produit ω de N observations indépendantes de $\Gamma(\alpha, \lambda)$ a pour valeur

$$f(\omega) = \frac{\lambda^{N\alpha}}{\Gamma(N\alpha)} \int_0^\infty \varphi(\alpha, N, \frac{\omega}{\Delta^N}) \cdot \Delta^{N(\alpha-1)-1} e^{-\lambda\Delta} d\Delta.$$

En effet $\varphi(\alpha, N, \frac{\omega}{\Delta^N})$ est la valeur de la densité de probabilité de

$\frac{\omega}{\Delta^N} = \sigma_N$ conditionné par Δ , et $\frac{1}{\Delta^N} \varphi(\alpha, N, \frac{\omega}{\Delta^N})$ celle de la densité de probabilité de ω conditionné par Δ .

Sachant que la densité de Δ a pour valeur $\lambda^{N\alpha} \Delta^{N\alpha-1} e^{-\lambda\Delta} / \Gamma(N\alpha)$ on a bien

$$f(\omega) = \frac{\lambda^{N\alpha}}{\Gamma(N\alpha)} \int_0^\infty \varphi(\alpha, N, \frac{\omega}{\Delta^N}) \cdot \Delta^{N\alpha-N-1} e^{-\lambda\Delta} d\Delta$$

X.2 - Application à l'analyse de la variance.

2.1. Introduction.

2.1.1. L'analyse de la variance sur des plans d'expériences (avec erreurs normales équivariées) est une technique statistique très répandue, on trouve par exemple des exposés classiques et pratiques en langue française dans DUGUE et GIRAULT (ref. 21), et CHAPOUILLE (ref. 12).

Ci-dessous nous employons la présentation et les notations de BARRA (réf. 5) pour le modèle I (I chiffre romain). On suppose que l'on a fait N expériences d'où un vecteur résultant α de \mathbb{R}^N d'espérance m et de variance-covariance $\sigma^2 I_{N,N}$. A priori on se place dans le cas où m est un vecteur d'un sous-espace V de \mathbb{R}^N , avec V_\perp l'orthogonal de V dans \mathbb{R}^N .

On considère ensuite des sous-espaces V_i de V tels que les sous-espaces orthogonaux W_i des V_i dans V soient deux à deux orthogonaux et engendrent V .

On désigne par α_{V_\perp} , α_{W_i} les projections d'un vecteur α sur les sous-espaces V_\perp et W_i qui engendrent \mathbb{R}^N . On sait que les normes carrées $\|\alpha_{V_\perp}\|^2$ et $\|\alpha_{W_i}\|^2$ suivent des lois connues dans le modèle :

$$\|\alpha_{V_\perp}\|^2 \text{ suit la loi gamma } \Gamma\left(\frac{N - \dim V}{2}, \frac{1}{2\sigma^2}\right),$$

$$\|\alpha_{W_i}\|^2 \text{ suit la loi gamma décentrée } \Gamma\left(\frac{\dim W_i}{2}, \frac{\|m_{W_i}\|^2}{2\sigma^2}, \frac{1}{2\sigma^2}\right),$$

avec le paramètre de décentrage $\delta_i = \|m_{W_i}\|^2 / 2\sigma^2$.

L'hypothèse $H_0 : \{m \in V\}$ étant prise comme cadre a priori, les hypothèses habituellement considérées sont du type $H_i : \{m_{W_i} = 0\}$ auquel cas les lois gamma correspondantes sont centrées.

De tels espaces V , V_i et W_i se présentent dans l'Analyse de la variance des plans d'expériences où l'on examine l'effet de facteurs F_α et de leurs interactions, quand on décompose les fonctions-résultats $\eta(e)$, (η réel, e est un point de l'ensemble produit des modalités prises par les facteurs). Un m_{W_i} nul s'interprète en termes statistiques, cela correspond à l'absence de l'influence d'un facteur ou d'une interaction entre certains d'entre eux.

2.1.2. Les $\|x v_i\|^2$ et $\|x w_i\|^2$ constituent les éléments nécessaires pour construire sur le simplexe unité le point de coordonnées

$$y_i = \frac{\|x w_i\|^2}{\sum \|x w_j\|^2}, \quad i=1, \dots, M,$$

qui par construction suit la distribution de DIRICHLET décentrée

$DD(a; \delta)$, avec les vecteurs a et δ de coordonnées

$$a_i = \frac{\dim W_i}{2} \quad \text{et} \quad \delta_i, \quad (\text{chap. IV.12}).$$

Un test à hypothèses linéaires multiples consiste à choisir parmi les 2^I hypothèses

$$\delta_i = 0 \quad i \in \mathcal{P}_1$$

$$\delta_{i'} > 0 \quad i' \in \mathcal{P}_2$$

pour toute partition $\mathcal{P}_1, \mathcal{P}_2$ de l'ensemble des parties de $[1, i, I]$.

On sait que l'on peut se ramener à tester les I hypothèses $\delta_i = 0$ (c'est-à-dire $m \in V_i$) (réf. 65).

En l'occurrence il s'agit de comparer deux à deux des densités de DIRICHLET décentrées $DD(y; a; \delta)$ et $DD(y; a; \delta')$ avec

$\delta_j = \delta'_j \quad j \neq i$ et $\delta'_i = 0$, au vu d'un seul point échantillon (ici $R = 1$).

2.2. Approximations pour distributions de DIRICHLET décentrées.

Comme le souligne SCHEFFE il est commode de remplacer des lois du χ^2 ou de F décentrées par des lois centrées approchées du même type (réf. 71-Annexe). Dans le cas d'une loi gamma décentrée $\Gamma(a_i, \delta_i, \lambda)$ de densité

$$e^{-\gamma_i} \sum_{n \geq 0} \frac{\gamma_i^n}{n!} \frac{\lambda^{a_i+n} x^{a_i+n-1} e^{-\lambda x}}{\Gamma(a_i+n)} * \mathbb{1}_{\mathbb{R}_+}(x)$$

de moyenne $\frac{a_i + \gamma_i}{\lambda}$ et de variance $\frac{a_i + 2\gamma_i}{\lambda^2}$, SCHEFFE

propose en substance de prendre la loi centrée $\Gamma(b_i, \mu_i)$ avec

$$b_i = \frac{(a_i + \gamma_i)^2}{a_i + 2\gamma_i} \quad \text{et} \quad \mu_i = \lambda \frac{a_i + \gamma_i}{a_i + 2\gamma_i}. \quad \text{Cette loi}$$

approchée possède les mêmes moments d'ordre 1 et 2 que ceux de la loi initiale. Malheureusement cette méthode appliquées aux $\|x w_i\|^2$ donne des μ_i différents, et par conséquent on ne peut en déduire une distribution de DIRICHLET comme cela est fait en IV.4.

On peut raisonner autrement en cherchant à simplifier le mélange de lois gamma $\Gamma(a_i+n, \lambda)$ par la loi unique qui correspond au coefficient multiplicateur $\gamma_i^n / n!$ le plus important. On sait que ce maximum est atteint pour $n = \lceil \gamma_i \rceil$. En pratique on peut retenir la distribution $\Gamma(a_i + \gamma_i, \lambda)$ qui possède la même moyenne que la loi initiale, et une variance égale à $(a_i + \gamma_i) / \lambda^2$; la différence entre les deux approximations $\Gamma(b_i, \mu_i)$ et $\Gamma(a_i + \gamma_i, \lambda)$ tient à ce que l'on remplace $\frac{a_i + \gamma_i}{a_i + 2\gamma_i}$ par 1.

Cette approximation par la technique du "coefficient dominant" est donc peu différente de la technique basée sur les deux premiers moments; dans le cas présent elle offre l'avantage de donner un paramètre de taille λ qui ne varie pas avec i (contrairement à μ_i), donc de conduire à la distribution de DIRICHLET $D(a + \gamma)$.

Remarque : On peut appliquer la technique précédente à une distribution de DIRICHLET décentrée :

$$D D (a ; \gamma) = e^{-\sum \gamma_i} \sum_{K \geq 0} \frac{(\sum \gamma_i)^K}{K!} \sum_{\substack{k_1 + \dots + k_n = K}} \frac{K!}{k_1! \dots k_n!} \theta_1^{k_1} \dots \theta_n^{k_n} D(a, k')$$

en posant $\theta_i = \gamma_i / \sum \gamma_i$, et k^v le vecteur de coordonnées k_i , $i=1, \dots, n$.

α) Sur l'ensemble des k^v tels que $\sum k_i = K$ le coefficient dominant est atteint pour un k^v/K "proche" de θ . En pratique on peut retenir la distribution unique $D(a + K\theta)$, d'où l'approximation

$$e^{-\sum \gamma_i} \sum_{K \geq 0} \frac{(\sum \gamma_i)^K}{K!} D(a + K\theta).$$

β) Sur l'ensemble des $K \geq 0$, le coefficient dominant dans l'approximation ci-dessus est atteint pour $K = [\sum \gamma_i]$; en pratique on peut retenir la seule distribution $D(a + \gamma)$.

γ) On retrouve ainsi $D(a + \gamma)$ lorsqu'on applique directement la technique du coefficient dominant sur $D D (a ; \gamma)$. Il y a donc une certaine cohérence dans cette approximation, bien que l'on n'a pas ici l'égalité entre les moments d'ordre 1 et 2 des distributions $D D (a ; \gamma)$ et $D(a + \gamma)$.

2.3. Distributions marginales, et distributions des parts regroupées.

La distribution de DIRICHLET décentrée vérifie deux propriétés de la distribution de DIRICHLET vues au chapitre VII.2.1., qui sont directement applicables en termes statistiques.

2.3.1. Distributions marginales.

Proposition 5 Soit $DD(a, \gamma)$ une distribution de DIRICHLET décentrée sur le simplexe S_1^M . Pour tout d compris entre 1 et M on pose Q l'ensemble $1 \leq q \leq d$ et R l'ensemble $d+1 \leq r \leq M$, $A_d = \sum_0^d a_i$, $\Gamma_0 = \sum_0^M \gamma_i$, $\theta_i = \gamma_i / \Gamma_0$, $\theta^d = \sum_0^d \theta_i$, a^d et γ^d les vecteurs de \mathbb{R}^d de coordonnées a_q, γ_q $q \in Q$, $\beta_q = \gamma_q$ $q \in Q$ et $\beta_{d+1} = \sum_R \gamma_r$, ($\beta \in S_1^{d+1}$).

Alors la distribution marginale de $DD(a, \gamma)$ en $y_1 \dots y_d$ est la distribution de DIRICHLET décentrée $DD(a^d, A_{d+1}; \gamma^d, \Gamma_{d+1})$ sur le simplexe S_1^{d+1} .

En effet la série qui définit la densité décentrée étant absolument convergente, toute densité marginale de la série est égale à la somme des densités marginales correspondantes. La densité marginale en β est donc :

$$e^{-\Gamma_0} \sum_{K \geq 0} \frac{\Gamma_0^K}{K!} \sum_{b_1, \dots, b_n}^* \frac{K!}{b_1! \dots b_n!} \theta_1^{b_1} \dots \theta_n^{b_n} D(\beta; \dots a_q + b_q \dots, A_{d+1} + \sum_R b_r)$$

La sommation \sum^* étant étendue aux b_i tels que $\sum b_i = K$.

Soit $K' = \sum_{d+1} b_r$, \sum^* s'écrit encore

$$\sum_{b_1, \dots, b_d}^* \frac{K!}{b_1! \dots b_d! K!} \theta_1^{b_1} \dots \theta_d^{b_d} D(\beta; \dots a_q + b_q \dots, A_{d+1} + K') \cdot \frac{K!}{\prod_R b_r!} \prod_R \theta_r^{b_r}$$

c'est-à-dire

$$\sum_{k_1, \dots, k_r} \frac{K!}{k_1! \dots k_r! k'!} \theta_1^{k_1} \dots \theta_r^{k_r} (1 - \sum_R \theta_r)^{k'} D(\mathcal{Z}; \dots a_1 + k_1 \dots, A_{0..} + k')$$

avec $\sum_{i=1}^r k_i = K - k'$ étendue à $k' + \sum_1^r k_i = K$. ■

2.3.2. Distribution des coordonnées regroupées.

Proposition 6: Soit P_1, \dots, P_ℓ une partition de $[1, \dots, M]$, \mathcal{Z} le point de coordonnées $\mathcal{Z}_\ell = \sum y_{i\ell}$, $i_\ell \in P_\ell$ du simplexe S_1^L , et soient b et η les points de coordonnées $b_\ell = \sum a_{i\ell}$, $\eta_\ell = \sum \gamma_{i\ell}$, $i_\ell \in P_\ell$.

Alors la distribution de \mathcal{Z} est la distribution de DIRICHLET décentrée

$$D D(b; \eta).$$

Démonstration: Il suffit de démontrer la proposition pour un regroupement élémentaire tel que $\mathcal{Z}_1 = \mathcal{Z}$, $q=1, \dots, \ell$, $\mathcal{Z}_{0..1} = \sum_{i=1}^M \mathcal{Z}_i$, $\mathcal{Z} \in S_1^{1..1}$.

(chaque terme de $D D(a; \gamma)$ correspond à une densité des coordonnées regroupées de la forme $D(\mathcal{Z}; \dots a_1 + k_1 \dots, A_{0..1} + \sum_R k_R)$

La sommation de celles-ci conduit aux mêmes calculs combinatoires que ceux de la proposition précédente. ■

Conséquences :

i) Les parts regroupées $\mathcal{Z}_\ell = \sum y_{i\ell}$ sont des statistiques exhaustives pour les paramètres regroupés b_ℓ et η_ℓ selon la même partition en P_ℓ ;

ii) Les deux propriétés précédentes de $D D(a; \gamma)$ sont cohérentes avec l'approximation précédente qui revient à remplacer $D D(a; \gamma)$ par $D(a; \gamma)$.

2.4. Analyse de la variance .

Les tests rappelés ci-contre reviennent à comparer les distributions

$DD(a; \delta)$ et $DD(a; \delta')$ sur S_1^M , où l'on a $\delta'_i = \delta_i$ pour tout indice i sauf pour l'un d'entre eux \bar{i} , où $\delta'_{\bar{i}} = 0$, ($\bar{i} = 1$ par exemple).

On applique la propriété précédente aux deux regroupements suivants

$$z_1 = y_1, \quad z_2 = \sum_2^M y_2, \quad z \in S_1^2,$$

ce qui donne les distributions Béta décentrées de 1^o espèce

$$BD(a_1, A_2; \tilde{\gamma}_1, \Gamma_2), \quad \tilde{\gamma}_1 \text{ égal à } 0 \text{ ou } \gamma_1,$$

et leurs approximations par les distributions Béta de 1^o espèce

$$B(a_1 + \tilde{\gamma}_1, A_2 + \Gamma_2), \quad \text{définies sur } [0, 1].$$

Le rapport des deux densités Béta (densités approchées)

$$\frac{\Gamma(A_2 + \gamma_1 + \Gamma_2)}{\Gamma(A_2 + \Gamma_2)} \frac{\Gamma(a_1)}{\Gamma(a_1 + \gamma_1)} z_1^{\gamma_1}$$

conduit à des régions critiques de la forme $z_1 \geq c$.

Si l'on utilise l'approximation Béta des distributions Beta décentrées, le calcul du seuil et de la puissance de ces tests nécessite uniquement des calculs de la fonction Béta incomplète (réf. 64).

2.5. Remarques sur le modèle II .

Nous terminons cette application par une propriété analytique des densités de DIRICHLET décentrées qui est inspirée par le modèle II de l'Analyse de la variance, sur des plans complets équilibrés.

2.5.1. Le modèle II suppose que les effets et les interactions dus aux modalités des facteurs sont des variables aléatoires qui suivent des lois normales et ne sont plus des quantités fixes. Avec des plans complets équilibrés, les statistiques $\|x_{w_i}\|^2$ suivent des lois gamma (réf. 5,65). On fait les deux remarques suivantes :

α) Si z_i suit la loi gamma $\Gamma(\alpha_i, \lambda/\sigma_i^2)$, z_i/σ_i^2 suit la loi $\Gamma(\alpha_i, \lambda)$, et les $\frac{z_i/\sigma_i^2}{\sum_k z_k/\sigma_k^2}$ suivent la distribution de DIRICHLET $D(\alpha)$;

β) On rappelle que les γ_i suivent des lois gamma

$$\Gamma\left(\frac{\dim w_i}{2}, \mu_i\right) \quad , (\text{réf. 81}) .$$

On est ainsi amené à étudier la distribution de DIRICHLET décentrée conditionnée par les paramètres de décentrage γ_i qui suivent des lois gamma.

2.5.2. Lemme 1 Soit y un point de S_1^n , et z le point de coordonnées $z_i = \frac{\gamma_i \gamma_i}{\sum_k \gamma_k \gamma_k}$, $\gamma_i > 0$ pour tout $i=1, \dots, n$. Le jacobien $\frac{D(z)}{D(y)}$ de ce changement de variable est égal à :

$$\prod_{i=1}^n \frac{\gamma_i}{\sum_k \gamma_k \gamma_k} .$$

En effet

i) Soit le changement de variable $z_i = \frac{\varphi_i(y_i)}{\sum_k \varphi_k(y_k)}$, où les $\varphi_i(y_i)$

sont des fonctions de même signe, monotones de même croissance. Formellement

le calcul de $\frac{D(z)}{D(y)}$ est le même que celui des paragraphes IV 7 et 8,

à la différence qu'à présent φ_i φ'_i désignent ici respectivement

$\varphi_i(y_i)$ et $\varphi'_i(y_i)$. On a donc

$$\frac{D(z)}{D(y)} = z_1 \dots z_n \sum_k \prod_{j \neq k} \frac{\varphi'_j}{\varphi_j} .$$

ii) On prend $\varphi_i(t) = \gamma_i t$, $\gamma_i > 0$. Il y a manifestement bijection entre les simplexes unités de y et z , et l'on a

$$\varphi'_i / \varphi_i = 1 / y_i, \quad z_i / y_i = \gamma_i / \sum_k \gamma_k z_k,$$

d'où

$$\frac{D(z)}{D(y)} = \prod_i \frac{z_i}{y_i}$$

sachant que $\sum_1^m y_i = +1$ ■

Proposition 7 Soit $DD(a; \gamma)$ une distribution de DIRICHLET décentrée. On considère que les γ_i sont des paramètres de conditionnement qui suivent des lois gamma $\Gamma(a_i, \mu_i)$. On pose $\theta_i = \mu_i / (1 + \mu_i)$.

Si l'on effectue le changement de variable $z_i = \theta_i y_i / \sum \theta_k y_k$

sur la distribution marginale de y après le déconditionnement en γ , alors le point z suit la distribution de DIRICHLET centrée $D(a)$.

En effet: la densité marginale $\omega(y; a)$ en y après déconditionnement est la suivante :

$$\begin{aligned} \omega(y; a) &= \sum_k \sum^* D(y; a + k_i) \cdot \prod_j \int_0^\infty \frac{\mu_j^{a_j} \gamma_j^{y_j + k_j - 1} e^{-(1 + \mu_j) \gamma_j}}{k_j! \Gamma(a_j)} d\gamma_j, \\ &= \sum_k \sum^* D(y; a + k_i) \cdot \prod_j \theta_j^{a_j} \left(\frac{1}{1 + \mu_j} \right)^{k_j} \frac{1}{k_j!} \frac{\Gamma(a_j + k_j)}{\Gamma(a_j)}, \\ &= \frac{1}{\prod_k \Gamma(a_k)} \cdot \prod_j \theta_j^{a_j} \gamma_j^{a_j - 1} \sum_k \Gamma(A_1 + k) \sum^* \prod_l \frac{1}{k_l!} \left(\frac{\gamma_l}{1 + \mu_l} \right)^{k_l}, \end{aligned}$$

où \sum^* est étendue sur les k_p vérifiant $\sum k_p = K$. C'est, sous la dernière forme, une sommation multinomiale égale à

$$\frac{1}{K!} \left(\sum \frac{y_i}{1+\mu_i} \right)^K ;$$

La sommation en K est la somme d'une distribution de PASCAL, égale à

$$\frac{\Gamma(A_1)}{\left(1 - \sum \frac{y_i}{1+\mu_i}\right)^{A_1}} = \frac{\Gamma(A_1)}{\left(\sum \theta_i y_i\right)^{A_1}},$$

(sachant que $\sum y_i = 1$).

$$\text{D'où } \mathcal{D}(y; a) = \frac{\Gamma(A_1)}{\prod \Gamma(a_i)} \cdot \prod \left(\frac{\theta_i y_i}{\sum \theta_k y_k} \right)^{a_i - 1} \cdot \prod \left(\frac{\theta_j}{\sum \theta_k y_k} \right)^{a_j}$$

Le changement de variable $z_i = \theta_i y_i / \sum \theta_k y_k$ et le lemme conduisent à la conclusion. \blacksquare

X.3 - Structures dérivées.

3.1. Nouveaux paramètres.

Le paramètre α de la distribution $D(\alpha)$ est un paramètre vectoriel de coordonnées α_i positives pour tout $i=1, \dots, M$. Si on pose A la somme des coordonnées $\sum \alpha_i$, et θ_i le rapport α_i/A , le point Θ de coordonnées θ_i est un point du simplexe unité ouvert S_1^{M-1} . On met ainsi le paramètre α sous la forme ΘA , où ΘA est un point du produit cartésien $S_1^{M-1} \times \mathbb{R}_+^M$. On note $D(\Theta, A)$ la distribution de DIRICHLET quand on explicite Θ et A .

On rappelle que θ est l'espérance de $D(\theta, A)$ et que la matrice d'élément $\theta_i(\delta_{ij} - \theta_j)/(A+1)$ $i, j=1, \dots, M$ est sa matrice de variance-covariance (réf. 84). De façon tout à fait élémentaire, on rappelle ainsi que

le point moyen $\bar{y} = \frac{1}{p} \sum_p y^p$ et la matrice de variance-covariance

$\frac{1}{p-1} \sum_p (y^p - \bar{y})(y^p - \bar{y})'$ de l'échantillon des y^p sont des estimateurs sans biais de l'espérance et de la variance-covariance de $D(\theta, A)$.

De façon imagée on peut donc considérer que θ est un paramètre vectoriel de "direction" dans \mathbb{R}_+^M et A un paramètre scalaire de "concentration" ; comme c'est le cas pour la distribution de FISHER VON-MISES sur l'hypersphère unité (réf.75).

Cette décomposition permet d'introduire une nouvelle présentation de la structure générale, et deux structures dérivées sur le simplexe S_1^M .

3.2. Structure générale de DIRICHLET .

On peut répéter l'analyse de cette structure avec les nouveaux paramètres. On rappelle que la densité de cette distribution a pour valeur, pour tout

$y \in S_1^M$:

$$D(y; \theta, A) = \frac{\Gamma(A)}{\prod_{j=1}^M \Gamma(\theta_j A)} \prod_{j=1}^M e^{(\theta_j A - 1) \cdot \log y_j} .$$

Sous cette forme les paramètres naturels sont $\theta_j A$ pour tout $j=1, \dots, M$.

Pour tout $i=1, \dots, M$ on pose $\overline{\log y_i} = \frac{1}{p} \sum_p \log y_i^p$,

La vraisemblance d'un échantillon de P partages - observations indépendants est la suivante :

$$L = \prod_{j=1}^M e^{P(\theta_j A - 1) \cdot \overline{\log y_j}} \cdot \left[\frac{\Gamma(A)}{\prod_{j=1}^M \Gamma(\theta_j A)} \right]^P .$$

On en déduit :

$$\left\{ \begin{array}{l} \frac{\partial \log L}{\partial A} = P \sum \theta_i \overline{\log y_i} + P [\psi(A) - \sum \theta_i \psi(\theta_i A)] \\ \frac{\partial \log L}{\partial \theta_i} = P A \overline{\log y_i} - P A \psi(\theta_i A) \quad i=1, \dots, M \end{array} \right.$$

d'où le système du maximum de vraisemblance suivant, sous la liaison $\sum \theta_i = +1$:

$$\left\{ \begin{array}{l} \psi(A) = \sum_{i=1}^M \theta_i (\psi(\theta_i A) - \overline{\log y_i}) \\ P A (\overline{\log y_i} - \psi(\theta_i A)) = C \quad , i=1, \dots, M . \end{array} \right.$$

ce qui donne la solution :

$$\left\{ \begin{array}{l} \hat{A} = \sum_i \psi^{-1} [\psi(\hat{A}) + \overline{\log y_i}] \\ \hat{\theta}_i = \frac{1}{\hat{A}} \psi^{-1} [\psi(\hat{A}) + \overline{\log y_i}] \quad i=1, \dots, M , \end{array} \right.$$

et la valeur $C = -P \hat{A} \psi(\hat{A})$ du multiplicateur de LAGRANGE qui correspond à la liaison $\sum \theta_i = +1$.

On a vu que ce système possède une solution unique avec une probabilité 1 pour tout effectif $P \geq 2$ (chap. VII). On remarque que dans la résolution de ce système les variables θ_i et A ne sont pas séparables, mais que \hat{A} peut être calculé indépendamment des $\hat{\theta}_i$, $i=1, \dots, M$.

3.3. Structure de DIRICHLET avec A fixé.

3.3.1. On note $D_A(\theta)$ la distribution de DIRICHLET lorsque le paramètre A est fixé, avec la densité qui a pour valeur, pour tout $y \in S_1^M$:

$$D_A(y; \theta) = \frac{\Gamma(A)}{\prod \Gamma(\theta_i A)} \prod y_i^{\theta_i A - 1}$$

On considère la structure statistique associée :

$$\left(S_1^M, \mathcal{B}(S_1^M), \left\{ D_A(\cdot; \theta), \theta \in \dot{S}_1^M \right\} \right)^P.$$

Théorème 4 - La structure statistique ci-dessus est une structure exponentielle complète. Pour tout indice $\bar{x} = 1, \dots, M$ il existe une statistique (vectorielle à $M-1$ composantes) exhaustive, $j=1, \dots, M, j \neq \bar{x}$

$$T_{j, \bar{x}} = \overline{\log y_j} - \overline{\log y_{\bar{x}}}.$$

En effet le paramètre θ varie dans une variété de dimension $M-1$. On prend $\bar{x} = M$ par exemple, la densité a pour expression :

$$\begin{aligned} D_A(y; \theta) &= \frac{\Gamma(A)}{\prod \Gamma(\theta_i A)} \prod_{j=1}^{M-1} y_j^{\theta_j A - 1} \prod_{j=1}^{M-1} \prod_{i=1}^{M-1} y_j^{\theta_i A - 1} \\ &= \frac{\Gamma(A)}{\prod \Gamma(\theta_i A)} \prod_{j=1}^{M-1} y_j^A \prod_{j=1}^{M-1} \prod_{i=1}^{M-1} y_j^{-1} \prod_{i=1}^{M-1} \left(\frac{y_i}{y_M} \right)^{A \theta_i}. \end{aligned}$$

C'est la forme générale d'une densité de probabilité d'une structure exponentielle, avec les paramètres naturels $\theta_k, k=1, \dots, M, k \neq M$. Les statistiques exhaustives sont

$$T_k^{\bar{x}=M} = A \left(\overline{\log y_k} - \overline{\log y_M} \right).$$

Les résultats sont analogues pour tout indice \bar{x} .

3.3.2. Estimateur du maximum de vraisemblance.

i) La vraisemblance d'un échantillon de P observations indépendantes est égale à

$$L = \prod_{i=1}^M e^{P(\theta_i A - 1) \overline{\log y_i}} \cdot \left[\frac{\Gamma(A)}{\prod \Gamma(\theta_i A)} \right]^P,$$

que l'on maximise sous la liaison $\sum \theta_i = 1$, d'où, $i=1, \dots, M$,

$$\frac{\partial \log L}{\partial \theta_i} = PA \overline{\log y_i} - PA \psi(\theta_i A) - C.$$

ii) Pour résoudre ce système on prend l'inconnue auxiliaire $K = \psi(\theta_i A) - \overline{\log y_i}$

($K = -C/PA$), elle vérifie les relations nécessaires :

$$A\theta_i = \psi^{-1}(K + \overline{\log y_i})$$

et

$$A = \sum_i \psi^{-1}(K + \overline{\log y_i}).$$

Réciproquement, si K^* vérifie la dernière relation, les paramètres

$$\theta_k^* = \frac{1}{A} \psi^{-1}(K^* + \overline{\log y_k})$$

sont solutions du système du maximum de vraisemblance.

Proposition 8 - Pour tout échantillon de partages - observations indépendants

\mathcal{Y}^P de taille $P \geq 1$, le système du maximum de vraisemblance possède une solution unique.

En effet soit g l'application de \mathbb{R} dans \mathbb{R}^+ définie par, $K \in \mathbb{R}$:

$$g(K) = \sum \psi^{-1}(K + \overline{\log y_k})$$

C'est une somme d'applications positives continues croissantes avec les valeurs limites

$$\lim_{k \rightarrow -\infty} g(k) = 0^+,$$

$$\lim_{k \rightarrow +\infty} g(k) = +\infty.$$

Pour tout A réel positif, l'équation $g(k) = A$ possède donc une racine unique k^* , d'où la solution unique du système du maximum de vraisemblance

$$\hat{\theta}_k = \theta_k^*, \quad k=1, \dots, M. \quad \blacksquare$$

Remarque : En pratique le calcul des θ_k^* repose sur la résolution numérique de l'équation $g(k) = A$.

3.3. Structure de DIRICHLET avec θ fixé.

3.3.1. On note $D_\theta(A)$ la distribution de DIRICHLET lorsque θ est fixé, avec la densité qui a pour valeur, pour tout $y \in S_1^M$:

$$D_\theta(y; A) = \frac{\Gamma(A)}{\prod_1^M \Gamma(\theta_j; A)} \prod_1^M y_j^{\theta_j A - 1}.$$

On considère la structure statistique associée :

$$\left(S_1^M, \mathcal{B}(S_1^M), \left\{ D_\theta(\cdot; A), A \in \mathbb{R}_+^* \right\} \right)^P.$$

Théorème 5 - La structure statistique ci-dessus est une structure exponentielle

complète. Elle possède la statistique exhaustive $T = \sum_{k=1}^M \theta_k \overline{\log y_k}$,
(on pose $\overline{\log y_k} = \frac{1}{P} \sum_P \log y_k^p$).

En effet A varie dans \mathcal{R}_+^* variété de dimension 1 ; la densité a pour expression, (on pose $A \theta_i^{-1} = \theta_i^{-1} + \theta_i (A-1)$)

$$D_{\theta} (y; A) = \frac{\Gamma(A)}{\prod_{j=1}^M \Gamma(\theta_j; A)} \prod_{j=1}^M y_j^{\theta_j - 1} \times \left[\prod_{j=1}^M y_j^{\theta_j} \right]^{A-1};$$

c'est la forme générale d'une densité de probabilité d'une structure exponentielle, avec le paramètre naturel A . Pour un échantillon de taille P , la statistique exhaustive est

$$T = \sum_{k=1}^M \theta_k \overline{\log y_k} \quad \blacksquare$$

3.3.2. Estimateur du maximum de vraisemblance .

L'estimateur \hat{A} est solution de l'équation :

$$\psi(A) - \sum_{i=1}^M \theta_i \psi(\theta_i; A) = - \sum_i \theta_i \overline{\log y_i}.$$

Proposition 9 - Pour tout échantillon de partage-observations y^P d'effectif $P \geq 1$, il y a une solution unique finie de l'équation du maximum de vraisemblance.

En effet soit l'application de \mathcal{R}_+^* dans \mathcal{R}_+^* définie par, $A \in \mathcal{R}_+^*$

$$g(A) = \psi(A) - \sum_{i=1}^M \theta_i \psi(\theta_i; A) = \sum_{i=1}^M \theta_i [\psi(A) - \psi(\theta_i; A)],$$

c'est une somme d'applications positives continues décroissantes avec les valeurs limites

$$\lim_{A \rightarrow 0^+} g(A) = +\infty,$$

$$\lim_{A \rightarrow +\infty} g(A) = - \sum_{i=1}^M \theta_i \overline{\log \theta_i}.$$

Les points θ et y^p , $p=1, \dots, P$ sont des points du simplexe unité, on sait que $-\sum_1^M \theta_k \log y_k^p \geq -\sum_1^M \theta_k \log \theta_k$ avec égalité si et seulement si $y_k^p = \theta_k$ pour tout $k=1, \dots, M$. L'événement $y^p = \theta$ étant de probabilité nulle, le deuxième membre qui figure dans l'équation du maximum de vraisemblance est supérieur strictement à la limite de $g(A)$ quand A tend vers plus l'infini, pour tout $P \geq 1$. On a donc avec une probabilité égale à 1 vers une racine unique \hat{A} à distance finie, pour tout $P \geq 1$ ■

Remarques: En pratique le calcul de l'estimateur repose sur la résolution numérique de l'équation $g(A) = -\sum_1^M \theta_i \overline{\log y_i}$.

Complément: Quand l'effectif P tend vers plus l'infini, on sait que la distribution de $\sqrt{P} (\hat{A} - A)$ converge en loi vers la distribution normale d'espérance nulle et de variance $-1/E\left(\frac{\partial^2 \log D_\theta}{\partial A^2}\right)$ noté $\sigma_{\hat{A}}^2$.

on a ci-contre $\frac{\partial^2 \log D_\theta}{\partial A^2} = \psi'(A) - \sum_1^M \theta_k^2 \psi'(\theta_k A)$

d'où $\frac{1}{\sigma_{\hat{A}}^2} = \sum_1^M \theta_k^2 \psi'(\theta_k A) - \psi'(A)$.

Le cas θ fixé généralise le cas de la structure de DIRICHLET invariante par \sum_M , auquel cas θ est égal à ω . Connaissant le maximum de vraisemblance pour un θ fixé, on peut comme en VII.2.7.3. appliquer le test du rapport de vraisemblance pour tester l'hypothèse que a appartient à la demi-droite de vecteur directeur θ contre l'hypothèse que a appartient au cône positif de \mathbb{R}^M .

ANNEXE 1

Une identité entre les coefficients du triangle de PASCAL .

1) Notations

1.1. On pose \bar{S}_N^M l'ensemble des points n de R^M , à coordonnées n_i entières positives ou nulles

$$\bar{S}_N^M = \left\{ n \in \mathbb{N}^M, \sum_{i=1}^M n_i = N \right\}$$

pour tout $M \geq 2$; cet ensemble peut être identifié par isomorphisme

à l'ensemble des M-uples (n_1, n_2, \dots, n_M) , $n_i \in \mathbb{N}$, $\sum n_i = N$.

1.2. En suivant les notations du texte on pose BE_N^M le cardinal

de \bar{S}_N^M , on a vu que $BE_N^M = C_{N+M-1}^N = \binom{N+M-1}{N}$.

1.3. Sur les sous-ensembles de \bar{S}_N^M , on pose μ la mesure de l'effectif, en particulier

$$\mu(\bar{S}_N^M) = BE_N^M.$$

1.4. Soient les sous-ensembles particuliers :

$$\forall j=1, \dots, M$$

$$\bar{S}_N^{M,j} = \left\{ n \in \bar{S}_N^M, \quad n_j = 0 \right\}$$

$$\forall j_1, \dots, j_k \in \{1, \dots, M\}, \quad j_1 \neq j_2 \neq j_3 \neq \dots \neq j_k$$

$$\bar{S}_N^{M, j_1 j_2 \dots j_k} = \left\{ n \in \bar{S}_N^M, \quad n_{j_1} = n_{j_2} = \dots = n_{j_k} = 0 \right\}$$

$$\bar{S}_N^0 = \left\{ n \in \bar{S}_N^M, \quad n_i \geq 1 \quad i=1, \dots, M \right\}$$

2) Propositions.

1 - Pour tout $j=1, \dots, M$, il existe une application bijective entre \bar{S}_N^{M-1} et $\bar{S}_N^{M,j}$. Démonstration immédiate.

2 - Pour tout $k=1, \dots, M-1$, et pour tous j_1, \dots, j_k différents qui appartiennent à l'ensemble $\{1, \dots, M\}$, il existe une application bijective entre \bar{S}_N^{M-k} et $\bar{S}_N^{M, j_1 \dots j_k}$. Démonstration immédiate.

3 - On a l'identité suivante:

$$\bar{S}_N^{M, j_1 \dots j_k} = \bigcap_{i=1}^k \bar{S}_N^{M, j_i}, \quad \text{Démonstration immédiate.}$$

4 - Pour tout N supérieur ou égal à 1, l'ensemble $\bar{S}_N^{M, 1, \dots, M}$ est vide. Démonstration immédiate.

5 - Pour tous $N \geq M \geq 2$, il existe une application bijective entre \bar{S}_N^0 et \bar{S}_{N-M}^M .

En effet soit l'application qui à tout point n de \bar{S}_{N-M}^M fait correspondre le point n' de coordonnées égales à n_i+1 , $i=1, \dots, M$. On obtient un point de \bar{S}_N^M dont toutes les coordonnées sont supérieures ou égales à 1; on vérifie immédiatement que c'est une application injective et surjective entre \bar{S}_{N-M}^M et \dot{S}_N^M .

6 - Le complémentaire de \dot{S}_N^M dans \bar{S}_N^M est égal à l'ensemble $\bigcup_{j=1}^M \bar{S}_N^{M,j}$.
 En effet, tout point du complémentaire de \dot{S}_N^M dans \bar{S}_N^M possède au-moins une coordonnée nulle, donc appartient à $\bar{S}_N^{M,j}$ $j=1, \dots, M$.
 Réciproquement, tout point de l'union des $\bar{S}_N^{M,j}$ $j=1, \dots, M$ appartient au complémentaire de \dot{S}_N^M .

7 - La mesure μ de l'effectif se conserve par les bijections précédentes. (propriété triviale).

3) Théorème Pour tous $N \geq M \geq 1$, on a l'identité suivante:

$$\binom{N-1}{M-1} = \sum_{k=0}^{M-1} (-1)^k \binom{M}{k} \binom{N+M-k-1}{N}$$

En effet, pour tous $N \geq M \geq 2$ on a

$$\begin{aligned} \text{a) } \mu(\dot{S}_N^M) &= \mu(\bar{S}_{N-M}^M) = BE_{N-M}^M \\ &= \binom{N-1}{M-1} \end{aligned}$$

$$\begin{aligned} \text{b) } \mu(\bar{S}_N^M) &= \mu(\bar{S}_N^M) - \mu \bigcup_j \bar{S}_N^{M,j} \\ &= BE_N^M + \sum_{k=1}^{M-1} (-1)^k \sum_{j_1, \dots, j_k} \mu(\bar{S}_N^{M, j_1, \dots, j_k}) \end{aligned}$$

$$\mu \left(\begin{smallmatrix} \circ & M \\ S & N \end{smallmatrix} \right) = \sum_{k=0}^{M-1} (-1)^k C_H^k BE_N^{M-k}$$

d'où

$$BE_{N-M}^M = \sum_{k=0}^{M-1} (-1)^k C_H^k BE_N^{M-k}$$

Pour $M=1$, on vérifie que la relation est encore vraie quel que soit N , les deux membres sont tous deux égaux à 1. ■

4) Remarque:

On a dans RIORDAN la relation suivante, (Ref. 69):

$$\sum (-1)^k \binom{p}{k} \binom{n-k}{n-m} = \binom{n-p}{m}$$

qui conduit au même résultat.

ANNEXE 2

Tables de probabilité de la variable $\sup(n_i)$ lorsque le point n de coordonnées n_i $i=1, \dots, M$ suit la distribution équiprobable sur l'ensemble

$$\bar{S}_N^M = \left\{ n \in \mathcal{N}^M, \sum n_i = N \right\}$$

(distribution dite de BOSE-EINSTEIN dans le texte)

$$M = 2$$

- $N = 1$, $BE_N^M = C_2^1 = 2$

p	q	$BE_N^M(p)$
1	1	2

- $N = 2$, $BE_N^M = C_3^1 = 3$

p	q	$BE_N^M(p)$
2	1	2
1	2	1

- $N = 3$, $BE_N^M = C_4^1 = 4$

p	q	$BE_N^M(p)$
3	1	2
2	1	2

$$M = 2$$

$$\bullet N = 4, \quad BE_N^M = C_5^1 = 5$$

p	q	$BE_N^M(p)$
4	1	2
3	1	2
2	2	1

$$\bullet N = 5, \quad BE_N^M = C_6^1 = 6$$

p	q	$BE_N^M(p)$
5	1	2
4	1	2
3	1	2

$$\bullet N = 6, \quad BE_N^M = C_7^1 = 7$$

p	q	$BE_N^M(p)$
6	1	2
5	1	2
4	1	2
3	2	1

$$M = 2$$

• $N = 7$, $BE_N^M = C_8^1 = 8$

p	q	$BE_N^M(p)$
7	1	2
6	1	2
5	1	2
4	1	2

• $N = 8$, $BE_N^M = C_9^1 = 9$

p	q	$BE_N^M(p)$
8	1	2
7	1	2
6	1	2
5	1	2
4	2	1

• $N = 9$, $BE_N^M = C_{10}^1 = 10$

p	q	$BE_N^M(p)$
9	1	2
8	1	2
7	1	2
6	1	2
5	1	2

$N = 1$, $BE_N^M = C_3^2 = 3$

p	q	$BE_N^M(p)$
1	1	3

$N = 2$, $BE_N^M = C_4^2 = 6$

p	q	$BE_N^M(p)$
2	1	3
1	2	3

$N = 3$, $BE_N^M = C_5^2 = 10$

p	q	$BE_N^M(p)$
3	1	3
2	1	6
1	3	1

$N = 4$, $BE_N^M = C_6^2 = 15$

p	q	$BE_N^M(p)$
4	1	3
3	1	6
2	2	6

$$M = 3$$

• $N = 5$, $BE_N^M = C_7^2 = 21$

p	q	$BE_N^M(p)$
5	1	3
4	1	6
3	1	9
2	2	3

• $N = 6$, $BE_N^M = C_8^2 = 28$

p	q	$BE_N^M(p)$
6	1	3
5	1	6
4	1	9
3	2	9
2	3	1

• $N = 7$, $BE_N^M = C_9^2 = 36$

p	q	$BE_N^M(p)$
7	1	3
6	1	6
5	1	9
4	1	12
3	2	6

$$M = 3$$

• $N = 8$, $BE_N^M = C_{10}^2 = 45$

p	q	$BE_N^M(p)$
8	1	3
7	1	6
6	1	9
5	1	12
4	2	9
3	2	3

• $N = 9$, $BE_N^M = C_{11}^2 = 55$

p	q	$BE_N^M(p)$
9	1	3
8	1	6
7	1	9
6	1	12
5	1	15
4	2	9
3	3	1

• $N = 10$, $BE_N^M = C_{12}^2 = 66$

p	q	$BE_N^M(p)$
10	1	3
9	1	6
8	1	9
7	1	12
6	1	15
5	2	15
4	2	6

$$M = 3$$

- $N = 11$, $BE_N^M = C_{13}^2 = 78$

p	q	$BE_N^M(p)$
11	1	3
10	1	6
9	1	9
8	1	12
7	1	15
6	1	18
5	2	12
4	2	3

- $N = 12$, $BE_N^M = C_{14}^2 = 91$

p	q	$BE_N^M(p)$
12	1	3
11	1	6
10	1	9
9	1	12
8	1	15
7	1	18
6	2	18
5	2	9
4	3	1

- $N = 13$, $BE_N^M = C_{15}^2 = 105$

p	q	$BE_N^M(p)$
13	1	3
12	1	6
11	1	9
10	1	12
9	1	15
8	1	18
7	1	21
6	2	15
5	2	6

$$M = 4$$

$$N = 1, \quad BE_N^M = C_4^3 = 4$$

p	q	$BE_N^M(p)$
1	1	4

$$N = 2, \quad BE_N^M = C_5^3 = 10$$

p	q	$BE_N^M(p)$
2	1	4
1	2	6

$$N = 3, \quad BE_N^M = C_6^3 = 20$$

p	q	$BE_N^M(p)$
3	1	4
2	1	12
1	3	4

$$N = 4, \quad BE_N^M = C_7^3 = 35$$

p	q	$BE_N^M(p)$
4	1	4
3	1	12
2	2	18
1	4	1

$$284$$

$$M = 4$$

$$N = 5, \quad BE_N^M = C_8^3 = 56$$

p	q	$BE_N^M(p)$
5	1	4
4	1	12
3	1	24
2	2	16

$$N = 6, \quad BE_N^M = C_9^3 = 84$$

p	q	$BE_N^M(p)$
6	1	4
5	1	12
4	1	24
3	2	34
2	3	10

$$N = 7, \quad BE_N^M = C_{10}^3 = 120$$

p	q	$BE_N^M(p)$
7	1	4
6	1	12
5	1	24
4	1	40
3	2	36
2	3	4

$$N = 8, \quad BE_N^M = C_{11}^3 = 165$$

p	q	$BE_N^M(p)$
8	1	4
7	1	12
6	1	24
5	1	40
4	2	54
3	2	30
2	4	1

$$M = 4$$

- $N = 9$, $BE_N^M = C_{12}^3 = 220$

p	q	$BE_N^M(p)$
9	1	4
8	1	12
7	1	24
6	1	40
5	1	60
4	2	60
3	3	16

- $N = 10$, $BE_N^M = C_{13}^3 = 286$

p	q	$BE_N^M(p)$
10	1	4
9	1	12
8	1	24
7	1	40
6	1	60
5	2	78
4	2	58
3	3	10

- $N = 11$, $BE_N^M = C_{14}^3 = 364$

p	q	$BE_N^M(p)$
11	1	4
10	1	12
9	1	24
8	1	40
7	1	60
6	1	84
5	2	88
4	2	48
3	3	4

• $N = 1$, $BE_N^M = C_5^4 = 5$

p	q	$BE_N^M(p)$
1	1	5

• $N = 2$, $BE_N^M = C_6^4 = 15$

p	q	$BE_N^M(p)$
2	1	5
1	2	10

• $N = 3$, $BE_N^M = C_7^4 = 35$

p	q	$BE_N^M(p)$
3	1	5
2	1	20
1	3	10

• $N = 4$, $BE_N^M = C_8^4 = 70$

p	q	$BE_N^M(p)$
4	1	5
3	1	20
2	2	40
1	4	5

$$M = 5$$

$$\bullet N = 5, \quad BE_N^M = C_9^4 = 126$$

p	q	$BE_N^M(p)$
5	1	5
4	1	20
3	1	50
2	2	50
1	5	1

$$\bullet N = 6, \quad BE_N^M = C_{10}^4 = 210$$

p	q	$BE_N^M(p)$
6	1	5
5	1	20
4	1	50
3	2	90
2	3	45

$$\bullet N = 7, \quad BE_N^M = C_{11}^4 = 330$$

p	q	$BE_N^M(p)$
7	1	5
6	1	20
5	1	50
4	1	100
3	2	125
2	3	30

$$M = 5$$

• $N = 8$, $BE_N^M = C_{12}^4 = 495$

p	q	$BE_N^M(p)$
8	1	5
7	1	20
6	1	50
5	1	100
4	2	165
3	2	140
2	4	15

• $N = 9$, $BE_N^M = C_{13}^4 = 715$

p	q	$BE_N^M(p)$
9	1	5
8	1	20
7	1	50
6	1	100
5	1	175
4	2	230
3	3	130
2	4	5

• $N = 10$, $BE_N^M = C_{14}^4 = 1001$

p	q	$BE_N^M(p)$
10	1	5
9	1	20
8	1	50
7	1	100
6	1	175
5	2	270
4	2	280
3	3	100
2	5	1

$$M = 6$$

$$N = 1, BE_N^M = C_6^5 = 6$$

p	q	$BE_N^M(p)$
1	1	6

$$N = 2, BE_N^M = C_7^5 = 21$$

p	q	$BE_N^M(p)$
2	1	6
1	2	15

$$N = 3, BE_N^M = C_8^5 = 56$$

p	q	$BE_N^M(p)$
3	1	6
2	1	30
1	3	20

$$N = 4, BE_N^M = C_9^5 = 126$$

p	q	$BE_N^M(p)$
4	1	6
3	1	30
2	2	75
1	4	15

$$290$$

$$M = 6$$

$$N = 5, \quad BE_N^M = C_{10}^5 = 252$$

p	q	$BE_N^M(p)$
5	1	6
4	1	30
3	1	90
2	2	120
1	5	6

$$N = 6, \quad BE_N^M = C_{11}^5 = 462$$

p	q	$BE_N^M(p)$
6	1	6
5	1	30
4	1	90
3	2	195
2	3	140
1	6	1

$$N = 7, \quad BE_N^M = C_{12}^5 = 792$$

p	q	$BE_N^M(p)$
7	1	6
6	1	30
5	1	90
4	1	210
3	2	330
2	3	126

$$N = 8, \quad BE_N^M = C_{13}^5 = 1287$$

p	q	$BE_N^M(p)$
8	1	6
7	1	30
6	1	90
5	1	210
4	2	405
3	2	456
2	4	90

ANNEXE 3

Convergence presque sûre du vecteur aléatoire multinomial

Soit θ un point interne du simplexe unité S_1^M de coordonnées $\theta_i > 0$, $i=1, \dots, M$, $\sum \theta_i = 1$.

Soit $m(\cdot; N; \theta)$ la distribution multinomiale définie par les probabilités $\frac{N!}{n_1! \dots n_n!} \prod \theta_i^{n_i}$ en tout point n de coordonnées $n_i \in \mathcal{N}$, $i=1, \dots, M$.

Théorème - Quand on fait tendre N vers plus l'infini, le vecteur aléatoire n/N de coordonnées n_i/N , $i=1, \dots, M$ qui suit la loi multinomiale $m(\cdot; N; \theta)$ converge presque sûrement vers le point θ .

Plus précisément pour tous ε et γ positifs et tout $\theta \in S_1^M$ il existe un $N(\theta)$ borné tel que pour tout $N > N(\theta)$ on a

$$\text{Prob} \left\{ \left\| \frac{n}{N} - \theta \right\| \geq \gamma \right\} < \varepsilon$$

où $\| \cdot \|$ est une norme dans \mathbb{R}^M .

Démonstration : Le résultat est trivial si θ est en un point extrémal de S_1^M . Dans le cas contraire on peut se ramener à un simplexe S_1^{n-k} tel que θ soit interne au simplexe.

a) pour tout $i=1, \dots, M$ on sait que n_i/N suit une distribution binomiale de paramètre θ_i , et l'on sait que cette variable converge presque sûrement vers θ_i , c'est à dire

$$\forall \varepsilon_1 > 0 \quad \eta_1 > 0 \quad \exists N_1(\theta_i) \quad \text{tel que pour tout } N > N_1(\theta_i)$$

$$\text{Prob} \left\{ \left| \frac{n_i}{N} - \theta_i \right| \geq \eta_1 \right\} < \varepsilon_1.$$

En particulier on peut prendre $N_1(\theta_i) = \left\lceil \frac{\theta_i(1-\theta_i)}{\varepsilon_1 \eta_1^2} \right\rceil \leq \left\lceil \frac{1}{4\varepsilon_1 \eta_1^2} \right\rceil$.

b) Soit $N(\theta) = \sup_i N_1(\theta_i)$,

et l'événement

$$\bigcup_i \left\{ n : \left| \frac{n_i}{N} - \theta_i \right| \geq \eta_1 \right\}$$

On a de manière évidente, (l'ensemble de chacun des deux membres est inclus dans l'ensemble de l'autre membre) :

$$\begin{aligned} \bigcup_i \left\{ n : \left| \frac{n_i}{N} - \theta_i \right| \geq \eta_1 \right\} &= \left\{ n : \sup_i \left| \frac{n_i}{N} - \theta_i \right| \geq \eta_1 \right\} \\ &= \left\{ n : \left\| \frac{n}{N} - \theta \right\|_A \geq \eta_1 \right\} \end{aligned}$$

en posant la norme $\|x\|_A = \sup_i x_i$, $x \in \mathbb{R}^M$.

On en déduit que

$$\text{Prob} \left\{ n : \left\| \frac{n}{N} - \theta \right\|_d \geq \eta_1 \right\} \leq \sum \text{Prob} \left\{ n : \left| \frac{n_i}{N} - \theta_i \right| \geq \eta_1 \right\}$$

et donc pour tout $N > N_1(\theta)$ on a :

$$\text{Prob} \left\{ n : \left\| \frac{n}{N} - \theta \right\|_d \geq \eta_1 \right\} \leq \eta \varepsilon_1.$$

c) Par conséquent pour tout ε et η donnés, il suffit de prendre

$$\eta_1 = \eta \quad , \quad \varepsilon_1 = \varepsilon / M \quad \text{pour avoir le résultat, avec la norme } \|\cdot\|_d.$$

Par ailleurs on sait que dans \mathbb{R}^M toutes les normes $\|\cdot\|$ sont équivalentes (réf. 14).

On a en outre $N(\theta) \leq \left\lceil \frac{M}{4\varepsilon\eta^2} \right\rceil$ pour tout $\theta \in S_1^M$ ■

BIBLIOGRAPHIE

- 1 - ABRAMOWITZ M.
STEGUN I.A. Handbook of Mathematical functions (1970)
Dover
- 2 - ANDERSON T.W. An introduction to multivariate analysis (1958)
John Wiley
- 3 - BAKHVALOV N. Méthodes numériques (1976)
M.I.R.
- 4 - BARNDORFF-NIELSEN O. Information and Exponential families (1977)
John Wiley
- 5 - BARRA J.R. Notions fondamentales de statistique mathématique
(1971)
Dunod
- 6 - BARTON D.E.
DAVID F.N. Sums of ordered intervals and distances (1955)
Mathematika vol 2, pages 150-159
- 7 - " Some notes on ordered random intervals (1956a)
Jour. Roy Stat. Soc. série B vol 18, pages 79-94
- 8 - " Tests for randomness of points on a line (1956b)
Biometrika vol 43, pages 104-112
- 9 - BENEDETTI C. Sulla rappresentabilità di una distribuzione
binomiale mediante una distribuzione B e vice-versa
(1956)
Metron. vol 18 Nos 1-2 pages 121-131
- 10 - BLUMENTHAL S. Test of fit based on partial sums of the ordered
spacings (1970)
Ann. Inst. Statist Math, pages 261-276

- 11 - CHAKRAVARTI I.M.,
LAHA R.G.,
ROY J. Handbook of methods of applied statistics (1967)
John Wiley
- 12 - CHAPOUILLE P. Planification et analyse des expériences (1973)
Masson
- 13 - CHARRE J.G. L'espace dans le territoire français, morcellement
et vacuité (1973)
Centre de Recherche de l'Urbanisme C R U
- 14 - CHOQUET G. Cours d'Analyse, tome 2 Topologie (1964)
Masson
- 15 - COCHRAN W.G. The distribution of the largest of a set of estimated
variances as a fraction of their total (1941)
Ann. of Engenics, vol 11, pages 47-52
- 16 - CONNOR R.J.,
MOSIMANN J.E. Concepts of independance for proposition with a
generalisation of the DIRICHLET distribution (1969)
Journ. of Ann Stat. Ass, vol 64, pages 194-206
- 17 - COX D.R.
HINCKLEY D.V. Theoretical statistics (1974)
Chapman and Hall
- 18 - DARLING D.A. A test for homogeneity and extreme values (1952)
Ann. Math. Stat. vol 23, pages 450-456
- 19 - " On a class of problems relating to the random
division of an interval (1953)
Ann. Math. Stat. vol.24, pages 239-253
- 20 - DARROCH J.N.
RATCLIFF D. A characterization ot the DIRICHLET distribution (1971)
Journ. of Ann. Stat. Ass. vol 66, pages 641-643
- 21 - DUGUE D.
GIRAULT M. Analyse de variance et plans d'expérience (1969)
Dunod

- 22 - EISENHART C.,
HASTRAY M.W.,
WALLIS W.A. Selected techniques of statistical analysis (1947)
Mc graw Hill
- 23 - FAVERGE J.M. Méthodes statistiques en psychologie appliquée,
volume 3 (1973)
PUF
- 24 - FELLER W. An introduction to probability theory and its
applications, 2 volumes (1966,1968)
John Wiley
- 25 - de FINETTI B. Alcune osservazioni in tema di "suddivisione casuale"
(1964)
Gior. Inst. Italiano degli Attuari vol 27, p. 151-173
- 26 - FISHER R.A. Tests of significance in harmonic analysis (1929)
Proc. Roy. Soc. A vol 125, pages 54-59
- 27 - " The sampling distribution of some statistics obtained
from non-linear equations (1939)
Ann. Engen. vol 9, pages 238-249
- 28 - " On the similarity of the distributions found for the
Test of significance in harmonic analysis, and in
STEVENS' problem in geometric probability (1940)
Ann. Engen. vol 10, pages 14-17

réédités dans "Contributions to Mathematical
Statistics", J. Wiley, papers n° 16, 36 et 37
- 29 - FOURGEAUD C. Statistique (1967)
FUCHS A. Dunod
- 30 - GARNER W.R. The relation between information and variance
Mc GILL W.J. analysis (1956)
Psychometrika vol 21 n° 3, pages 219-228

- 31 - GARWOOD F. An application to the theory of probability to the operation of vehicular - controlled traffic signals (1940)
J. Roy. Stat. Soc. Suppl. vol 7, pages 65-77
- 32 - GOOD J.J. Saddlepoint methods for the multinomial distribution (1957)
Ann. Math. Stat. vol 28, pages 861-881
- 33 - " The estimation of probabilities. An essay on modern Bayesian methods (1965)
Research monograph n°30 M I T press
- 34 - GOURSAT E. Cours d'Analyse mathématique (1921)
Gauthier-Villars
- 35 - GREENWOOD M. The statistical study of infectious diseases (1946)
J. Roy Stat. Soc. A vol 109, pages 85-110
- 36 - GREENWOOD R.E.
GLASGOW M.O. Distribution of maximum and minimum frequencies in a sample drawn from a multinomial distribution (1950)
Ann. Math. Stat vol 21, pages 416-424
- 37 - GUILBAUD G. Th. Les problèmes de partages (1952)
Economie appliquée tome V n°1, pages 93-137
- 38 - GUMBEL E.G. Statistics of extremes (1958)
Columbia Un. Press
- 39 - HODGES J.L. On the noncentral Beta-distribution (1955)
Ann. Math. Stat. vol 26, pages 648-652
- 40 - IRWIN J.O. A unified derivation of some well known frequency distribution of interest (1955)
J. Roy Stat. Soc A vol 118, pages 389-404

- 41 - JOHNSON N.L. An approximation to the multinomial distribution :
some properties and applications (1960)
Biometrika vol 47, pages 93-102
- 42 - JOHNSON N.L. vol 1 : Discrete distributions (1969)
 , KOTZ S. J. Wiley
- 43 - " vol 4 : continuous multivariate distribution (1972)
 J. Wiley
- 44 - JOHNSON N.L. Some applications of two approximations to the multi-
 YOUNG D.H. nomial distribution (1960)
 Biometrika vol 47, pages 463-469
- 45 - KENDALL M.G. Geometrical probability (1963)
 MORAN P. Griffin
- 46 - KENDALL M.G. The advanced theory of statistics, volume 2 (1964)
 STUART A. Griffin
- 47 - KHATRI C.G. Some identities and approximations concerning positive
 MITRA S.K. and negative multinomial distribution (1969)
 in Multivariate Analysis, Tome II Ac. Press,
 pages 241-260
- 48 - KOZELKA R.M. Approximate upper percentage points for extreme values
 in multinomial sampling (1956)
 Ann. Math. Stat vol 27, pages 507-512
- 49 - KULLBACK S. Information theory and statistics (1958)
 Dover
- 50 - LEBART L. Techniques de la description statistique (1977)
 MORINEAU A. Dunod
 TABARD N.
- 51 - LEHMAN E.L. Testing statistical hypotheses (1959)
 J. Wiley

- 52 - LEVY P. Sur la division d'un segment par des points choisis au hasard (1939)
C.R. Acad. Sci. Paris vol 208, pages 147-149
Réédité dans Oeuvres complètes Tome 6 pages 312-314, note n° 247
- 52'- LIONS J.L. Quelques méthodes de résolution des problèmes aux limites non linéaires (1969)
Dunod, Gauthier-Villars
- 53 - MALLOWS C.L. An inequality involving multinomial probabilities (1968)
Biometrika vol 55, pages 422-424
- 54 - MAULDON J.G. Random division of an interval (1959)
Proc. Camb. Phil. Soc. vol 47, pages 331-336
- 55 - " A generalisation of the Beta distribution (1959)
Ann. Math. Stat. vol 30, pages 509-520
- 56 - MAURIN M. Effet de coupure et morcellement de l'espace naturel (1979)
IRT-CERNE
- 57 - MOSIMANN J.E. On the compound multinomial distribution, the multivariate Beta distribution, and correlations among proportions (1962)
Biometrika vol 49, pages 65-82
- 58 - " On the compound negative multinomial distribution and correlations among inversely sampled pollen counts (1963)
Biometrika vol 50, pages 47-54
- 59 - "Le Monde" n° du mardi 28 avril 1981
n° du mardi 12 mai 1981
- 60 - NICHOLSON W.L. A computing formula for the power of the analysis of variance test (1954)
Ann. Math. Stat. vol 25, pages 607-610
- 61 - OLKIN I. Multivariate Beta distributions and independence properties of the WISHART distribution (1964)
RUBIN H. Ann. Math. Stat. vol 35, pages 261-269

- 62 - OLKIN I. Integral expressions for tail probabilities of the
 SOBEL M. multinomial and negative multinomial distributions (1965)
 Biometrika vol 52, pages 167-179
- 63 - PALLU de la BARRIERE R. Cours d'automatique théorique (1966)
 Dunod
- 64 - PEARSON K. Tables of the incomplete Beta-Function (1956)
 Cambridge Un. Press
- 65 - PHAM DINH T. Contributions à l'analyse de la variance et aux plans
 d'expérience (1970)
 Thèse Université de Grenoble
- 66 - PYKE R. Spacings (1965)
 J. Roy. Stat. Soc. B vol 27, pages 395-449
- 67 - RAO J.S. Incomplete DIRICHLET integrals with applications to
 SOBEL M. orderd uniform spacings (1980)
 Jour. Mult. Analysis vol 10, pages 603-610
- 68 - RENYI A. Calcul des probabilités (1966)
 Dunod
- 69 - RIORDAN J. An introduction to combinatorial analysis (1958)
 J. Wiley
- 70 - ROBERTS F.S. Measurement theory (1979)
 Addison-Wesley Pg.Cy.
- 71 - SCHEFFE H. The analysis of variance (1959)
 J. Wiley
- 72 - SCHWARTZ L. Méthodes mathématiques pour les sciences physiques (1965)
 Hermann
- 73 - SOBEL M. DIRICHLET distribution type 1 in Selected tables in
 UPPULURI V.R.R. Math. Stat, vol 4 (1977)
 FRANKOWSKI K. Inst. of Math. Stat.

- 74 - SOLER J.L. Notions de liberté en statistiques mathématiques (1970)
Thèse, Université de Grenoble
- 75 - STEPHENS M.A. Use of the Von-MISES distribution to analyse continuous
proportions (1982)
Biometrika vol 69 pages 197-203
- 76 - STEVENS W.L. Solution to a geometrical problem in probability (1939)
Ann. Engen. vol 9, pages 315-320
- 77 - TAN N.Y. Note on the multivariate and the generalised multivariate
Beta distributions (1969)
J. Ann. Stat. Assoc. vol 64, pages 230-241
- 78 - TIAO G.C. The inverted DIRICHLET distribution with applications
GUTTMAN I. (1965)
J. Ann. Stat. Assoc. vol 60 pages 793-805
- 79 - TORTRAT A. Calcul des probabilités (1963)
Masson
- 80 - VALIRON G. Théorie des fonctions, tome I (1955)
Masson
- 81 - VAN CUTSEM B. Note de cours (manuscrit)
- 82 - WHITTAKER E.T. A course of modern analysis (1965)
WATSON G.N. Cambridge Un. Press
- 83 - WHITWORTH W.A. Choice and chance (1887)
Cambridge Un. Press
- 84 - WILKS S.S. Mathematical statistics (1963)
J. Wiley

DERNIERE PAGE D'UNE THESE

3È CYCLE, DOCTEUR INGÉNIEUR OU UNIVERSITÉ

Vu les dispositions de l'arrêté du 16 avril 1974,

Vu les rapports de M. *P. Van Cutsem*

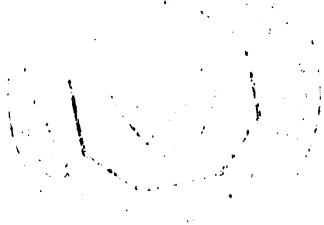
M.

M. *Michel MAURIN* est autorisé
à présenter une thèse en vue de l'obtention du grade de DOCTEUR *Ingénieur*

Grenoble, le *2* *883*

Le Président de l'Université Scientifique
et Médicale

M. TANCHE



Tanche

