

THÈSES DE L'UNIVERSITÉ PARIS-SUD (1971-2012)

NATHALIE CASTELLE

*Principes d'invariance et application à la statistique de modèles
censurés, 1991*

Thèse numérisée dans le cadre du programme de numérisation de la bibliothèque mathématique Jacques
Hadamard - 2016

Mention de copyright :

Les fichiers des textes intégraux sont téléchargeables à titre individuel par l'utilisateur à des fins de
recherche, d'étude ou de formation. Toute utilisation commerciale ou impression systématique est
constitutive d'une infraction pénale.

Toute copie ou impression de ce fichier doit contenir la présente page de garde.



ORSAY
n° d'ordre :

UNIVERSITE PARIS SUD

CENTRE d'ORSAY

T H E S E

présentée pour obtenir

**Le TITRE de DOCTEUR EN SCIENCES
SPECIALITE : MATHEMATIQUES**

par

Nathalie CASTELLE

**Sujet : PRINCIPES D'INVARIANCE ET APPLICATION A LA STATISTIQUE
DE MODELES CENSURES.**

Soutenue le 31 Janvier 1991 devant le Jury composé de :

ASSOUAD

BRETAGNOLLE

DEHEUVELS

HUBER

MASSART

Patrice

Jean

Paul

Catherine

Pascal

Président

NAME: CASTELLE Nathalie

SUBJECT: INVARIANCE PRINCIPLES AND APPLICATION TO THE STATISTIC OF CENSORED MODELS.

ABSTRACT:

Let $X_1, \dots, X_n$ a n -sample with distribution $F(t) = \mathbf{P}(X \leq t)$ continuous over $\mathbf{R}^+$ and let $\hat{F}_k$ the empirical distribution function associated to the k -first observations. We set $\alpha_k = \sqrt{k}(\hat{F}_k - F)$ the empirical bridge. Komlos, Major and Tusnady established in 1975 the existence of a Kiefer K process such that $\sup_{1 \leq k \leq n} \|\alpha_k - K/\sqrt{k}\|_\infty = O((\log^2 n)/\sqrt{n})$ a.s.

We give the complete proof of this strong approximation (called "Kiefer's approximation") via a lemma of normal approximation for a hypergeometric distribution. Otherwise, when the sample is randomly stratified, we establish the strong approximation for the associated empirical sub-functions. This last result enables us to establish the strong approximation for a large class of statistics constructed on randomly censored variables. For instance, we obtain the asymptotic normality for the temporal statistics of the Gehan and Logrank tests on the whole line, and that with optimal convergence rate and under any hypothesis. The same technics enable us to prove that the Cox maximum likelihood estimate converges to a gaussian variable with a rate of order $(\log^2 n)/\sqrt{n}$.

KEY WORDS AND PHRASES: Gaussian Processes, Empirical distribution function, Invariance principle, Strong approximation, Random censored model, Asymptotic properties of tests, Asymptotic properties of estimates.

REMERCIEMENTS

Je tiens en premier lieu à remercier mon directeur de thèse Jean Bretagnolle. Je ne saurais dire combien il m'a appris et cela toujours avec gentillesse et patience. Je lui suis très reconnaissante pour le temps qu'il m'a consacré et les idées qu'il m'a apportées.

Je remercie Patrice Assouad, Paul Deheuvels, Catherine Huber et Pascal Massart pour avoir accepté de faire partie de mon jury. Je tiens à remercier de nouveau Pascal Massart. Il a rapporté cette thèse et ses remarques se sont avérées extrêmement utiles. Mais il a également su me faire profiter de son expérience à divers moments importants pour moi. Je peux maintenant apprécier son exigence et ses conseils. Je remercie Richard Gill pour avoir accepté de rapporter cette thèse et pour les idées qu'il me suggère afin d'étendre les résultats.

Didier Dacunha Castelle a su me motiver et m'orienter vers une thèse, et m'a permis de la réaliser dans les meilleures conditions. Qu'il en soit remercié.

Je voudrais citer particulièrement Françoise Bonvalot. Ensemble, nous avons mené la démonstration du théorème de Kiefer. Ce fut des heures intenses qui resteront pour moi comme un bon souvenir. Françoise m'a apporté rigueur mathématique et sérieux dans le travail. Je l'assure ici de toute mon amitié. Je remercie toute l'équipe du laboratoire de Statistiques au sein de laquelle j'ai pu travailler dans une bonne ambiance et j'ajouterai une mention spéciale à mes amis Anne et Jean Claude.

Je remercie ma famille et Jean Marc qui m'ont soutenue et prodigué leur confiance. Je remercie aussi tous mes amis, même les plus inattendus, qui se sont manifestés à divers moments difficiles.

Je terminerai en remerciant Madame Zielinski pour avoir veillé à l'édition de cet ouvrage.

TABLE DES MATIERES

Introduction —————▶ page 1

Bibliographie —————▶ page 9

1ère partie : Principes d'invariance

Chapitre 1 : Théorème de Kiefer —————▶ page 11

Introduction —————▶ page 11

Quelques lemmes techniques —————▶ page 14

Résultat intermédiaire —————▶ page 15

Preuve du théorème —————▶ page 33

Appendice —————▶ page 38

Chapitre 2 : Approximation forte de sous fonctions de répartition
empirique —————▶ page 54

2nde partie : Application des principes d'invariance à la statistique de modèles censurés —————▶ page 66

Introduction —————▶ page 67

Principe d'invariance et inégalités exponentielles —▶ page 70

*Etude générale de certains processus construits sur des
données censurées* —————▶ page 75

*Application à la comparaison des temps de survie entre
deux populations* —————▶ page 90

*Application à l'estimateur du maximum de vraisemblance
de Cox* —————▶ page 110

INTRODUCTION

Introduction

Nous nous proposons dans ce travail d'étudier les propriétés asymptotiques de certains estimateurs et statistiques construits sur des données censurées à l'aide des principes d'invariance forts pour les fonctions de répartition empirique initiés par Komlos, Major et Tusnady en 1975. La première partie concerne les principes d'invariance forts pour les fonctions de répartition empirique: dans le premier chapitre, nous donnons la preuve complète d'un des théorèmes de Komlos, Major et Tusnady (1975), dans le second chapitre nous établissons un théorème d'invariance applicable à la statistique de modèles censurés. La seconde partie concerne la statistique de modèles censurés: à l'aide du résultat précédent nous établissons la normalité asymptotique avec vitesse de convergence de l'estimateur du maximum de vraisemblance de Cox et de certaines statistiques de test de comparaison des distributions entre deux populations (la normalité asymptotique de ces statistiques n'était jusqu'alors connue que sous l'hypothèse nulle ou des hypothèses contigües). Dans cette introduction, nous présentons d'abord les principes d'invariance forts pour les fonctions de répartition empirique, nous présentons ensuite les modèles censurés puis nous décrivons notre travail.

Présentation des principes d'invariance forts pour les fonctions de répartition empirique

Soient $X_1, \dots, X_n$ des variables indépendantes et de même loi F continue sur $\mathbf{R}^+$: $F(t) = P(X \leq t)$. Pour tout t positif, on définit la fonction de répartition empirique:

$$\hat{F}_n(t) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}(X_i \leq t)$$

Il est naturel de s'intéresser à l'approximation gaussienne du processus $\alpha_n(t) = \sqrt{n}(\hat{F}_n(t) - F(t))$. Est appelé *principe d'invariance fort pour les fonctions de répartition empirique* un résultat qui établit l'existence d'un processus gaussien W_n tel que:

$$\| \alpha_n - W_n \|_{\infty} = O(a_n) \quad \text{p.s.} \quad (1)$$

où a_n décroît vers zéro. Bien entendu, il s'agit de trouver une suite $(W_n, a_n)_n$ telle que a_n décroisse le plus vite possible vers zéro.

Une solution à ce problème fut pour la première fois proposée par Brillinger (1969). Il démontra l'existence d'un pont brownien W_n tel que (1) soit vraie avec:

$$a_n = n^{-1/4}(\log n)^{1/2}(\log \log n)^{1/4}$$

Cette vitesse est loin d'être optimale, comme nous allons le voir. La même année, Kiefer (1969) souleva un nouveau problème en remarquant que le théorème de Brillinger ne donnait aucun renseignement sur la loi jointe des processus W_n . Il proposa la construction d'une suite de ponts browniens indépendants $W_1, \dots, W_n, \dots$ tels que pour tout entier n :

$$\sup_{1 \leq k \leq n} \left\| \alpha_k - \sum_{j=1}^k W_j \right\|_{\infty} = O(a_n) \quad \text{p.s.} \quad (2)$$

Une telle approximation sera dite "de Kiefer". Tous ces résultats vont être repris par Komlos, Major et Tusnady (1975). Ils établiront l'existence d'une suite de ponts browniens tels que (1) soit vraie avec $a_n = (\log n)/\sqrt{n}$. Cette vitesse est optimale. Ils démontreront également un théorème d'approximation forte de Kiefer avec $a_n = (\log^2 n)/\sqrt{n}$. (Notons que cette vitesse est optimale à un facteur $\log n$ près). Enfin, Tusnady proposera en 1977 un principe d'invariance fort pour les fonctions de répartition empiriques de la loi uniforme sur $[0, 1] \times [0, 1]$ (la vitesse restant égale à $(\log^2 n)/\sqrt{n}$). Nous ne savons toujours pas si les résultats de Komlos, Major et Tusnady s'étendent aux dimensions strictement supérieures à deux, c'est à dire si dans ce cas, il existe encore un principe d'invariance fort avec une vitesse $(\log^d n)/\sqrt{n}$ où d dépend (ou ne dépend pas) de la dimension considérée. L'apport des théorèmes de Komlos, Major et Tusnady, qui sont d'ailleurs maintenant communément appelés théorèmes hongrois, a été considérable: les applications statistiques des principes d'invariance sont nombreuses, dès qu'il s'agit d'obtenir des résultats de convergence en loi avec vitesse de convergence. Citons pour mémoire l'estimation de densités par la méthode des noyaux, traitée par de nombreux auteurs: Bickel et Rosenblatt (1973) qui ne disposaient que du théorème de Brillinger, Liero (1982), Revesz (1978), Silverman (1978) et la liste n'est pas exhaustive.

Présentation des modèles censurés

Lors de l'étude d'une variable aléatoire T , les observations du n -échantillon $T_1, \dots, T_n$ peuvent être perturbées par une censure aléatoire C . On observe alors le n -échantillon:

$$(T_1 \wedge C_1, \dots, T_n \wedge C_n)$$

qui ne permet pas l'accès à la fonction de répartition empirique de T . De plus, la loi de $T \wedge C$ peut dépendre d'une variable aléatoire Z (éventuellement vectorielle) qui est appelée covariable. Enonçons tout de suite une hypothèse très souvent retenue: l'indépendance de T et de C conditionnellement à Z :

$$P(T \wedge C \geq x / Z) = P(T \geq x/Z) P(C \geq x/Z) \quad (3)$$

De tels modèles sont largement utilisés en statistique médicale. Les deux problèmes majeurs sont l'estimation de la fonction de survie $P(T \geq x/Z)$ et la construction de tests de non effet de la covariable. Donnons brièvement un historique des résultats. Le premier problème est de trouver un estimateur pour $P(T \geq x/Z)$ et d'étudier ses propriétés.

Sous l'hypothèse de non effet de la covariable, Kaplan et Meier (1958) proposent un tel estimateur pour $P(T \geq x)$ appelé estimateur de Kaplan-Meier ou estimateur produit-limite. Les propriétés de normalité asymptotique de cet estimateur ont été étudiées par Breslow et Crowley (1974). Notons que Burke, Csörgö et Horváth (1981) proposent une généralisation à $P(T \geq x/Z)$. Le second problème consiste à construire des tests de non effet de la covariable. Dans la littérature ce problème est ramené à la construction de tests de comparaison des distributions entre deux populations. Les statistiques habituelles sont celles de Gehan (1965), de Prentice (1978) et la statistique du logrank dûe à Peto et Peto (1972). Dans son livre Gill (1980) propose une démarche générale s'appuyant sur la théorie des martingales pour étudier ces différents estimateurs et statistiques. Il retrouve ainsi la normalité asymptotique de l'estimateur de Kaplan-Meier (à ce sujet, l'article de Gill (1983) améliore ses résultats précédents) et obtient également sous l'hypothèse de non effet de la covariable (ou des hypothèses contigües) la normalité asymptotique des statistiques de Gehan, de Prentice et du logrank. En s'inspirant de ses méthodes, Harrington et Fleming (1982) généralisent ces derniers résultats à une classe assez générale de tests, parmi lesquels ceux de Prentice et du logrank. Ces mêmes outils martingale permettent également l'étude d'un modèle censuré particulier: le modèle de Cox. Un modèle censuré est un modèle de Cox si en plus de l'hypothèse (3) on a:

$$P(T \geq x/Z) = (h(x))^{\exp \langle \beta | Z \rangle}$$

La loi de T est donc résumée par une fonction certaine h et un paramètre d'intérêt β . Ce modèle est un modèle semi-paramétrique. Cox introduit alors une pseudo vraisemblance, appelée vraisemblance de Cox, et tout naturellement un estimateur qui maximise cette pseudo vraisemblance, noté $\hat{\beta}_n$ et appelé estimateur du maximum de vraisemblance de Cox. Nous renvoyons au livre de Cox et Oakes (1984). Andersen et Gill (1982) établissent la consistance et la normalité asymptotique de $\hat{\beta}_n$ et d'un estimateur de $\log h$, ce qui apporte une solution satisfaisante au problème de l'estimation de $P(T \geq x/Z(x))$. Notons ici un fait nouveau: la covariable peut dépendre du temps. Les techniques utilisées jusqu'alors pour obtenir la normalité asymptotique des estimateurs et statistiques de modèles censurés (voir Andersen et Gill (1982) et Gill (1980-1983)) étaient donc les techniques martingales, dont le théorème de Rebolledo (1978-1980). La même étude peut être réalisée avec les théorèmes de principes d'invariance. Cette voie a été ouverte par Burke, Csörgö et Horváth (1981). Une difficulté technique d'importance, inhérente à la censure, est que l'on ne dispose pas de la fonction de répartition empirique. Aussi Burke, Csörgö et Horváth définissent-ils des sous fonctions de répartition empirique et démontrent un principe d'invariance fort pour de tels objets. Ils proposent alors un estimateur pour $P(T \geq x/Z)$ (estimateur de Kaplan-Meier généralisé) et établissent son approximation forte par un processus gaussien. De plus, la vitesse obtenue (que nous ne pensons pas être optimale) est suffisante pour obtenir une loi du logarithme itéré pour cet estimateur. Pour ce dernier point, nous renvoyons à l'article de Csörgö et Horváth (1983).

Description de notre travail

La première partie de cette thèse concerne les principes d'invariance forts pour les fonctions de répartition empirique et est constituée de deux chapitres. Le premier chapitre consiste en la preuve complète du théorème d'approximation forte de Kiefer de l'article de Komlos, Major et Tusnady (1975). Ce théorème repose sur un lemme d'approximation normale d'une loi hypergéométrique dont aucune démonstration n'est proposée dans l'article de Komlos, Major et Tusnady. Ce lemme, que nous appellerons lemme hypergéométrique est doublement important puisqu'il est également indispensable pour démontrer le principe d'invariance fort pour les fonctions de répartition empirique bidimensionnelles proposé par Tusnady (1977). Le second chapitre est une amélioration du théorème de Burke, Csörgö et Horváth (1981) qui établit un principe d'invariance fort pour des sous fonctions de répartition empirique. Ce théorème est indispensable à l'application des principes d'invariance aux modèles censurés. Dans ce cadre, notre amélioration nous a permis l'obtention de vitesses de convergence optimales. Dans la seconde partie de cette thèse nous avons établi une méthode pour l'étude des modèles censurés à l'aide des principes d'invariance forts. Nos applications concernent les statistiques des tests de Gehan et du log-rank, les estimateurs de la variance de ces statistiques, les dérivées première et seconde du logarithme de la vraisemblance de Cox et l'estimateur du maximum de vraisemblance de Cox. Nous établissons l'approximation forte de ces différentes statistiques temporelles par les processus gaussiens associés et obtenons ainsi leur normalité asymptotique avec vitesse de convergence. Cette vitesse est optimale et permet de retrouver pour ces statistiques les propriétés des processus gaussiens qui les approximent: en particulier leurs moments sont égaux et on peut aussi donner une loi du logarithme itéré pour les statistiques considérées. De plus les approximations fortes des statistiques de test sont valables sous n'importe quelle hypothèse (alors que nous ne la connaissions auparavant que sous l'hypothèse nulle ou des hypothèses contigües) ce qui permet théoriquement de calculer la puissance de ces tests. Enfin, nos approximations fortes sont uniformes lorsque le temps décrit $[0, \infty]$ ce qui règle le problème souvent rencontré dans la littérature de la restriction du temps à un intervalle compact et donc de l'arrêt des observations à une date fixe. Nous détaillons maintenant nos résultats.

Première partie - Chapitre un: Théorème de Kiefer

Nous énonçons et démontrons le lemme hypergéométrique, puis nous reprenons en détail la construction proposée par Komlos, Major et Tusnady (1975). Cette construction est bidimensionnelle: la majoration cherchée (2) est uniforme à la fois pour le temps t et pour le nombre d'observations k , il nous faut donc contrôler simultanément ces deux paramètres. Pour cela nous discrétisons le temps à la taille $1/n$ et contrôlons les fluctuations. Nous obtenons un quadrillage de $[0, 1] \times [0, n]$ (l'axe horizontal représentant le temps et l'axe vertical le nombre d'observations) sur lequel nous construisons simultanément le processus gaussien et le processus empirique. Le principe consiste à découper successivement des rectangles en quatre. A chaque étape nous connaissons la loi des processus sur le rectangle

total ainsi que sur sa moitié supérieure et sur sa moitié gauche. Ceci explique pourquoi la structure même des constructions bidimensionnelles est hypergéométrique. Une fois construits ces deux processus, nous contrôlons leur distance en probabilité. Chaque rectangle $[0, m] \times [0, j/n]$ avec $1 \leq m, j \leq n$ peut être considéré comme un vecteur de $\mathbf{R}^n \otimes \mathbf{R}^n$ dont le nombre de composantes non nulles sur la base de Haare est d'ordre $\log^2 n$. L'erreur d'approximation entre les deux processus sur le rectangle $[0, m] \times [0, j/n]$ s'exprime alors comme une somme de $\log^2 n$ "rectangles d'erreur" et le lemme hypergéométrique permet d'établir que la somme des erreurs sur une même bande horizontale est majorée par une variable $Z_n/\sqrt{n}$ où Z_n suit une loi du chi deux à $O(\log n)$ degrés de liberté. Notons que nos contrôles permettent une évaluation des constantes: $O(a_n) \leq (118 \log^2 n)/\sqrt{n}$ convient dans la formule (2).

Première partie - Chapitre deux: Approximation forte de sous fonctions de répartition empirique

Dans le chapitre, nous adaptons les théorèmes d'invariance au problème des modèles censurés. Burke, Csörgö et Horváth (1981) ont suivi une démarche analogue. Rappelons qu'un modèle censuré ne permet pas l'accès à la fonction de répartition empirique. Aussi Burke, Csörgö et Horváth définissent-ils les sous fonctions de répartition $F^j(t) = P(X \leq t, R = j)$ et les sous fonctions de répartition empirique:

$$\hat{F}_n^j(t) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}(X_i \leq t) \mathbf{1}(R_i = j)$$

où X est une variable de loi continue sur $\mathbf{R}^+$ et R une variable discrète à valeurs dans $\{1, \dots, m\}$. Le but est de trouver un principe d'invariance fort pour $(\hat{F}_n^j)_{j=1}^m$. Précisons davantage: il s'agit de construire une suite $W_n^1, \dots, W_n^m$ de processus gaussiens et de trouver des constantes C, K, λ telles que:

$$P\left(\sup_{1 \leq j \leq m} \|\sqrt{n}(\hat{F}_n^j - F^j) - W_n^j\|_\infty \geq \frac{1}{\sqrt{n}}(x + C \log n)\right) \leq K \exp(-\lambda x) \quad (4)$$

Burke, Csörgö et Horváth réalisent la construction d'un vecteur de processus $(W_n^j)_{j=1}^m$ tel que (4) soit vérifiée avec: $K = mK_1$ où K_1 est une constante absolue. Nous avons voulu améliorer ce résultat et une autre construction nous permet de retrouver (4) avec $C = 40$, $K = 7$, $\lambda = 1/20$. Autrement dit, le principe d'invariance obtenu ne dépend plus du nombre de composantes du vecteur $(\hat{F}_n^j)_{j=1}^m$. Notre construction est donc intéressante lorsque m devient très grand. De plus, nous pouvons préciser certaines relations entre les processus gaussiens qui vérifient (4). Pour cela, introduisons quelques notations. Soit $N = (N_j)_{j=1}^m$ la variable de loi multinômiale définie par: $N_j = \sum_{i=1}^n \mathbf{1}(R_i = j)$. Posons $p_j = P(R = j)$ et $u_j(t) = P(X \leq t/R = j)$. Alors notre construction réalisera:

$$W_n^j(t) = \sqrt{p_j} B^j(u_j(t)) - u_j(t) Z_n^j$$

où $Z_n = (Z_n^j)_{j=1}^m$ est un vecteur gaussien centré et de même covariance que $N/\sqrt{n}$, où $(B^j)_{j=1}^m$ est un vecteur de ponts browniens mutuellement indépendants et indépendants

de Z_n et de N , et où les vecteurs Z et N satisfont la relation:

$$\sup_{1 \leq j \leq m} \left| \frac{N_j - np_j}{\sqrt{n}} - Z_n^j \right| \leq \frac{37 \log n}{\sqrt{n}} \quad \text{p.s.}$$

Ce dernier résultat, énoncé en lemme, est en fait un théorème de la limite centrale avec vitesse explicite pour un vecteur multinomial de nombre de paramètres quelconque. Ces propriétés du vecteur $(W_n^j)_{j=1}^m$ ont été cruciales pour obtenir des vitesses optimales dans les approximations fortes des statistiques construites sur des données censurés.

La clé du théorème du chapitre deux est une construction non classique. L'idée consiste à "déconditionner" les variables X et R . On obtient ainsi un problème simple d'approximation forte de vraies fonctions de répartition empirique indépendantes que l'on sait résoudre grâce au principe d'invariance (1). On finit par un "reconditionnement" des processus.

Seconde partie: Statistique de modèles censurés

Nous supposons que la covariable Z est une variable discrète à valeurs dans $\{0, 1\}$ et qui ne dépend pas du temps. L'interprétation statistique est la suivante: les observations du n -échantillon $T_1 \wedge C_1, \dots, T_n \wedge C_n$ sont classifiées en deux groupes disjoints et l'appartenance d'un individu à un groupe ne peut varier au cours du temps. Nous adoptons une formalisation inspirée de Burke, Csörgö et Horváth (1981) qui est bien adaptée aux techniques d'approximation forte. Notons $d = \mathbf{1}_{T \leq C}$. On définit la variable aléatoire R à valeurs dans $\{1, 2, 3, 4\}$ par:

$$\begin{aligned} \{R = 2z + 1\} &= \{Z = z, d = 1\} \\ \{R = 2(z + 1)\} &= \{Z = z, d = 0\} \end{aligned}$$

qui sont respectivement les évènements disjoints:

$$\begin{aligned} &\text{"non censuré dans le } (z+1)\text{-ième groupe"} \\ &\text{"censuré dans le } (z+1)\text{-ième groupe"} \end{aligned}$$

Tout naturellement, l'on définit alors les sous fonctions de répartition $F_j(x) = \mathbf{P}(T \wedge C, R = j)$ et les sous fonctions de répartition empirique:

$$\hat{F}_j(x) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}(T_i \wedge C_i \leq x, R_i = j)$$

Suivant la convention généralement retenue en statistique médicale nous considérons la version continue à gauche des sous queues de répartition empirique:

$$\hat{G}_j(x) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}(T_i \wedge C_i \geq x, R_i = j)$$

Avec le théorème du chapitre deux de la première partie nous savons qu'il existe un processus gaussien $(W_j^n)_{j=1}^4$ tel que l'on ait simultanément:

$$\sup_{1 \leq j \leq 4} \|\sqrt{n}(\hat{F}_j - F_j) - W_j^n\|_\infty \leq 61 \frac{\log n}{\sqrt{n}} \quad \text{p.s.} \quad (1)$$

$$\sup_{1 \leq j \leq 4} \|\sqrt{n}(\hat{G}_j - G_j) - (W_j^n(\infty) - W_j^n)\|_\infty \leq 98 \frac{\log n}{\sqrt{n}} \quad \text{p.s.}$$

Nous pourrions alors obtenir les approximations fortes de statistiques construites sur de tels modèles dès qu'il sera possible de les linéariser en les termes $(\hat{F}_j - F_j)$ et $(\hat{G}_j - G_j)$ et de majorer convenablement les restes. Pour cela, nous avons étudié les intégrales stochastiques du type:

$$S_j^n(t, \beta) = \int_0^t \psi_j(\hat{G}_1(s), \dots, \hat{G}_4(s), \beta) d\hat{F}_j(s)$$

où ψ_j est une fonction réelle bornée sur $[0, 1]^4$ et β un paramètre réel. Ce choix est motivé par le fait que la plupart des processus construits sur des variables censurées s'écrivent comme une somme de telles intégrales. Donnons un exemple. Soit $u = (u_1, \dots, u_4) \in [0, 1]^4$. Alors les processus définis par

$$\int_0^t \psi_1(\hat{G}_1, \dots, \hat{G}_4, \beta) d\hat{F}_1 + \int_0^t \psi_3(\hat{G}_1, \dots, \hat{G}_4, \beta) d\hat{F}_3$$

avec:

$$\psi_1(u, \beta) = \frac{u_1 + u_2}{(e^\beta(u_3 + u_4) + (u_1 + u_2))^\alpha} \quad \text{et} \quad \psi_3(u, \beta) = -\frac{e^\beta(u_3 + u_4)}{(e^\beta(u_3 + u_4) + (u_1 + u_2))^\alpha}$$

correspondent pour des observations arrêtées au temps t à la dérivée au point β du logarithme de la vraisemblance de Cox lorsque $\alpha = 1$, à la statistique du logrank lorsque $\alpha = 1$ et $\beta = 0$ et à la statistique de Gehan lorsque $\alpha = \beta = 0$. Notons $\hat{H}_j(s, \beta) = \psi_j(\hat{G}_1(s), \dots, \hat{G}_4(s), \beta)$ et $H_j(s, \beta) = \psi_j(G_1(s), \dots, G_4(s), \beta)$. La formule suivante est à la base de notre étude:

$$S_j^n(t, \beta) = \int_0^t H_j dF_j + \int_0^t H_j d(\hat{F}_j - F_j) + \int_0^t (\hat{H}_j - H_j) dF_j + \int_0^t (\hat{H}_j - H_j) d(\hat{F}_j - F_j)$$

Le premier terme est un terme de centrage, le dernier terme est un terme d'erreur, les second et troisième termes sont des termes principaux: ils peuvent être approchés par un processus gaussien explicite grâce au principe d'invariance (1). L'approximation gaussienne du troisième terme via (1) n'est évidemment possible qu'après linéarisation de $\hat{H}_j - H_j$. Nous avons donc recherché des conditions de linéarisation des fonctions ψ_j qui permettent d'obtenir une approximation forte de S_j^n avec une vitesse optimale. Pour ne pas restreindre le temps à un intervalle compact nous avons séparé $t \leq t_n$ et $t > t_n$ avec $F_j(t_n) = O((\log n)/n)$ et réglé le cas $t > t_n$ par une inégalité de grande déviation. Le

résultat principal établit l'approximation forte du processus S_j^n par un processus gaussien dès que certaines conditions pour les fonctions ψ_j sont vérifiées. Cette approximation forte est uniforme lorsque t décrit $[0, \infty]$ et lorsque β varie dans un compact et les vitesses obtenues sont optimales. Nous vérifions ensuite ces conditions pour les statistiques temporelles de Gehan et du logrank, pour les estimateurs de la variance de ces statistiques ainsi que pour les dérivées première et seconde du logarithme de la vraisemblance de Cox. Nous obtenons ainsi l'approximation forte de toutes ces statistiques. Notons que si l'on arrête les observations à une date fixe t , la vitesse de convergence des statistiques de Gehan et du logrank renormalisées vers une variable gaussienne est d'ordre $(\log n)/\sqrt{n}$ mais si cette vitesse reste inchangée lorsque $t \rightarrow \infty$ pour la statistique de Gehan, elle se dégrade jusqu'à $(\log^2 n)/\sqrt{n}$ pour la statistique du logrank. Enfin, à partir de l'approximation forte des dérivées première et seconde du logarithme de la vraisemblance de Cox nous obtenons celle de l'estimateur du maximum de vraisemblance de Cox avec une vitesse d'ordre $(\log^2 n)/\sqrt{n}$.

BIBLIOGRAPHIE

References

- [1] ANDERSEN P.K. and GILL R.D., 1982. *Cox's regression model for counting processes: a large sample study*. Annals of Statistics, vol 10, 4, 1100-1120.
- [2] BENNETT G., 1962. *Probability inequalities for the sum of independants random variables*. J. AM. Statis. Assoc., 57, 33-45.
- [3] BERKES I. and PHILIPP W. 1979 *Approximation theorems for independent and weakly dependent random vectors*. Annals of Probability, vol 7, 29-54.
- [4] BICKEL P.J. and ROSENBLATT M., 1973. *On some global measures of the deviations of density function estimates*. Annals of Statistics, vol 1, 1071-1095.
- [5] BRESLOW N. and CROWLEY J., 1974. *A large sample study of the life table and product limit estimates under random censorship*. Annals of Statistics, vol 2, 437-453.
- [6] BRETAGNOLLE J. and MASSART P., 1989. *Hungarian construction from the non asymptotic view point*. Annals of Probability, vol 10, 239-256.
- [7] BRILLINGER D.R., 1969. *The asymptotic representation of the sample distribution function*. Bull. Amer. Math. Soc., vol 75, 545-547.
- [8] BURKE M.D., CSÖRGÖ S. and HORVÁTH L., 1981. *Strong approximations of some biometric estimates under random censorship*. Z. Warschein. Verw. Geb., 56, 87-112
- [9] COX D.R and OAKES D., 1984. *Analysis of survival data*. Chapman and Hall.
- [10] CSÖRGÖ M. and REVESZ P., 1981. *Strong approximations in probability and statistics*. Academic Press, New York, 133-134.
- [11] CSÖRGÖ S. and HORVÁTH L., 1983. *The rate of strong uniform consistency for the product limit estimator*. Z. Wahrschein. Verw. Geb., 62, 411-426.
- [12] DVORETZKY A., KIEFER J.C. and WOLFOWITZ J., 1956. *Asymptotic minimax character of the sample distribution function and of the classical multinomial estimator*. Ann. Math. Stat., 33, 642-669.
- [13] EINMAHL J.H.J. and KONING A.J. 1989. *Limits theorems for a general weighted process under random censoring with applications*. Medical Informatics and Statistics Report, 24. (University of Limburg)
- [14] GEHAN E.A., 1965. *A generalized Wilcoxon test for comparing arbitrarily singly-censored samples*. Biometrika, 52, 203-223.
- [15] GILL R.D., 1980. *Censoring and stochastic integrals*. Mathematical Centre Tracts 124. Amsterdam: Mathematische Centre.

- [16] GILL R.D., 1983. *Large sample behaviour of the product-limit estimator on the whole line.* Annals of Statistics, vol 11, 49-58.
- [17] HARRINGTON D. and FLEMING T., 1982. *A class of rank test procedures for censored survival data.* Biometrika, 69, 3, 553-566.
- [18] KAPLAN E.L. and MEIER P., 1958. *Nonparametric estimation from incomplete observations.* J. AM. Statis. Assoc., 53, 457-481.
- [19] KIEFER J., 1969. *On the deviations in the Skorohod-Strassen approximation scheme.* Z. Wahrschein. Verw. Geb., 13, 321-332.
- [20] KOMLOS J., MAJOR P., and TUSNADY G., 1975. *An approximation of partial sums of independent RV^2 - and the sample DF. I.* Z. Warschein. Verw. Geb., 32, 111-131.
- [21] LIERO H., 1982. *On the maximal deviation of the kernel regression function estimate.* Math. Operationsforsch. Statist., Ser Statistics, vol 13, nx2, 171-182.
- [22] MASSART P., 1990. *The tight constant of the Dvoretzky-Kiefer-Wolfowitz inequality.* A paraitre dans Annals of probability.
- [23] PETO R. and PETO J., 1972. *Asymptotically efficient rank invariant test procedures.* J.R. Statist. Soc., 135, 185-206.
- [24] PRENTICE R.L., 1978. *Linear rank tests with right censored data.* Biometrika, 65, 167-179.
- [25] REBOLLEDO R., 1978. *Sur les applications de la théorie des martingales à l'étude statistique d'une famille de processus ponctuels.* Springer Lect. Notes in Mathematics, 636, 27-70.
- [26] REBOLLEDO R., 1980. *Central limit theorems for local martingales* Z. Wahrsch. verw. Gebiete 51, 269-286.
- [27] REVESZ P., 1978. *A strong law of the empirical density function.* Periodica Math. Hung., Vol 9 (4), 317-324.
- [28] SHORACK G.R. and WELLNER J.A., 1986. *Empirical processes with applications to statistics.* Wiley and Sons.
- [29] SILVERMAN B.W., 1978. *Weak and strong uniform consistency of the kernel estimate of a density and its derivatives.* Annals of Statistics, vol 6, nx1, 177-184.
- [30] SKOROHOD A.V., 1976. *On a representation of random variables.* Th. Proba. Appl., 628-632.
- [31] TUSNADY G., 1977. *A remark on the approximation of the sample DF in the multidimensional case.* Periodica Math. Hung., Vol 8 (1), 53-55.
- [32] WELLNER J., 1978. *Limit theorems for the ratio of the empirical distribution function to the true distribution function.* Z. Warschein. Verw. Geb., 45, 73-88.

PREMIERE PARTIE

PRINCIPES D'INVARIANCE

CHAPITRE 1

THEOREME DE KIEFER

Principe d'invariance pour une fonction de répartition empirique: Théorème de Kiefer

Résumé

Soit $x_1, \dots, x_n$ un n -échantillon de loi uniforme sur $[0, 1]$ défini sur un espace de probabilité Ω suffisamment riche.

On note $\hat{F}_k$, la fonction de répartition empirique associée aux k -premières observations et $\alpha_k(t)$ le pont empirique $\alpha_k(t) = \sqrt{k}(\hat{F}_k(t) - t)$.

L'existence d'un pont brownien B° tel que la déviation uniforme $\|\alpha_n - B^\circ\|_\infty$ soit d'ordre $(\log n)/\sqrt{n}$ en probabilité s'obtient par une construction hongroise, réalisée par Komlos, Major et Tusnady [6], et un lemme d'approximation normale entre une variable binomiale $B(n, 1/2)$ recentrée et la variable gaussienne correspondante (Ce lemme, dû à Tusnady, est énoncé dans le livre de Csörgö et Revesz [3]). Bretagnolle et Massart [2] ont publié une démonstration complète de ce lemme, ce qui leur a permis d'évaluer les constantes lors du contrôle en probabilité de $\|\alpha_n - B^\circ\|_\infty$.

Dans le même article, Komlos, Major et Tusnady [6] ont démontré l'existence d'une suite de ponts browniens indépendants assurant pour $\sup_{1 \leq k \leq n} \|\alpha_k - (\sum_{i=1}^k B_i^\circ / \sqrt{k})\|_\infty$ un contrôle d'ordre $(\log^2 n)/\sqrt{n}$ en probabilité, sans préciser la valeur des constantes. La clé d'un tel résultat repose sur un lemme d'approximation normale pour des lois hypergéométriques de paramètres quelconques. Nous démontrons un tel lemme avec des constantes explicites, formulé en deux inégalités, qui complète et généralise une conjecture de Komlos, Major et Tusnady [6]. L'inégalité de type linéaire diffère de celle du lemme de Tusnady (sur une binomiale symétrique) par un terme correctif qui mesure la dissymétrie de la loi hypergéométrique et qui n'est pas améliorable. A partir de ce lemme et d'une construction de type hongrois, nous effectuons le contrôle en probabilité de $\sup_{1 \leq k \leq n} \|\sqrt{k}\alpha_k - \sum_{i=1}^k B_i^\circ\|_\infty$ en donnant une borne exponentielle pour le terme d'erreur, avec des constantes explicites.

1 Introduction

Soit $X_1, \dots, X_n$ une suite de variables indépendantes de loi uniforme sur $[0, 1]$.

Soit $\hat{F}_m$, $m = 1, \dots, n$, la fonction de répartition empirique définie par:

$$\hat{F}_m(t) = \frac{1}{m} \sum_{i=1}^m \mathbf{1}_{X_i \leq t} \quad \text{où } t \in [0, 1]$$

et soit le processus $\hat{F}_m^\circ$, $m = 1, \dots, n$, défini par:

$$\hat{F}_m^\circ(t) = (\hat{F}_m(t) - t) \quad \text{où } t \in [0, 1]$$

Enfin, soit α_m , $m = 1, \dots, n$, le pont empirique défini par:

$$\alpha_m(t) = \sqrt{m} \hat{F}_m^\circ(t) \quad \text{où } t \in [0, 1]$$

Le théorème 1 est dû à Komlos, Major et Tusnady [6] et à Bretagnolle et Massart [2] pour l'évaluation des constantes.

Theoreme 1 *Pour tout entier $n \geq 2$, il existe un mouvement brownien $B(t)$ tel que pour tout x positif :*

$$\mathbf{P}(\sqrt{n} \|\sqrt{n} \hat{F}_n^\circ(t) - B^\circ(t)\|_\infty > x + 12 \log n) \leq 2 \exp(-x/6)$$

Komlos, Major et Tusnady [6] énoncent le théorème suivant, dit théorème de Kiefer (la première version de ce résultat se trouve en effet dans un article de Kiefer [5]), et donnent les grandes lignes de la démonstration. Afin d'évaluer les constantes, nous démontrons en détail ce théorème dans les sections 3 et 4.

Theoreme 2 *Il existe une suite de mouvements browniens indépendants $B_1(t), \dots, B_n(t), \dots$ tels que pour tout $n \geq 3$ et pour tout x positif :*

$$\mathbf{P}\left(\sup_{1 \leq m \leq n} \sup_{0 \leq t \leq 1} |m \hat{F}_m^\circ(t) - \sum_{i=1}^m B_i^\circ(t)| > (x + C \log n) \log n\right) \leq K \exp(-\lambda x)$$

où $C = 76$, $K = 3.26$, $\lambda = 1/41$.

Notations

- On note $\Phi_{(n, np, np')}$ la fonction de répartition de la variable aléatoire de loi hypergéométrique $H(n, np, np')$ définie par:

$$\mathbf{P}(H(n, np, np') = k) = \frac{\binom{np}{k} \binom{nq}{np' - k}}{\binom{n}{np'}}$$

où np, np', nq, nq' et k sont des entiers, $p + q = 1, p' + q' = 1$, $\max(0, n(p' - q)) \leq k \leq \min(np, np')$.

- On note Ψ_n la fonction de répartition de la loi binomiale symétrique $B(n, 1/2)$ définie par

$$\mathbf{P}(B(n, 1/2) = k) = \binom{n}{k} (1/2)^n$$

où k est un entier de $[0, n]$

- On note Φ la fonction de répartition de la variable aléatoire Y suivant une loi normale centrée réduite .
- On note $B(t)$, ($t \geq 0$) le mouvement brownien et $B^0(t)$, ($0 \leq t \leq 1$) le pont brownien défini par $B^0(t) = B(t) - tB(1)$.

L'argument clé de la preuve du théorème 1 est le lemme suivant démontré par Bretagnolle et Massart [2].

Lemme 1 (Tusnady) *Approximation normale d'une loi binomiale symétrique.*

Soient Φ , Ψ_n définies précédemment, Y une variable aléatoire normale centrée réduite, on pose $B_n = \Psi_n^{-1} \circ \Phi(Y) - n/2$.

On a alors les inégalités suivantes:

1. $|B_n| \leq 1 + (\sqrt{n}/2) |Y|$
2. $|B_n - (\sqrt{n}/2)Y| \leq 1 + Y^2/8$

La démonstration du théorème 2 impose la connaissance d'un lemme d'approximation normale d'une hypergéométrique de paramètres quelconques. Ce lemme n'ayant jamais été énoncé sous cette forme à notre connaissance, nous le démontrons en détail dans l'appendice.

Lemme 2 *Approximation normale d'une loi hypergéométrique*

Soient Φ et $\Phi_{(n,np,np')}$ définies précédemment et Y une loi normale centrée réduite,

$$\text{on pose : } \begin{cases} X_n = \Phi_{(n,np,np')}^{-1} \circ \Phi(Y) - npp' \\ \sigma_n^2 = npp'qq' \\ \delta\delta' = (p-q)(p'-q') \end{cases}$$

Pour tout $\eta > 0$ tel que $|\delta\delta'| \leq 1 - \eta$, il existe des constantes a, b, c, d telles que:

1. $|X_n| \leq a + \sigma_n |Y| + b |\delta\delta'| |Y|^2$
2. $|X_n - \sigma_n Y| \leq c + dY^2$

où b et d ne dépendent que de η .

De plus $b \geq 1/6$ dès que $|\delta\delta'| \geq \rho > 0$, en ce sens l'inégalité 1 n'est pas améliorable.

La détermination des constantes du théorème 2 impose d'améliorer les inégalités 1 et 2 du Lemme 2 dans certains cas (la preuve est également donnée dans l'appendice).

Cas particuliers du Lemme 2 :

1. Si $\delta\delta' = 0$ et $\sigma^2 \geq 4.5$, on obtient l'inégalité: $|X_n| \leq 3 + \sigma|Y|$
2. Si $|\delta\delta'| \leq 1/8$ et $\sigma^2 \geq 4.5$, on obtient :

- (a) $|X_n| \leq 3 + \sigma_n |Y| + (7/5)^2 |\delta\delta'| Y^2$
(b) $|X_n - \sigma_n Y| \leq 3 + 0.41 Y^2$

Dans la suite, si X et Y sont des variables aléatoires, on note $E(X)$ l'espérance de X , $V(X)$ la variance de X et $\mathcal{L}(X/Y)$ la loi de X sachant Y .
On note $\mathcal{M}(n, p_1, \dots, p_k)$ la loi multinomiale définie par :

$$P(\mathcal{M}(n, p_1, \dots, p_k) = (l_1, \dots, l_k)) = \frac{n!}{l_1! \dots l_k!} p_1^{l_1} \dots p_k^{l_k}$$

$$\text{où } (l_1, \dots, l_k) \in \mathbb{N}^k \text{ et } \sum_{i=1}^k l_i = n$$

2 Quelques lemmes techniques

Inégalités exponentielles et maximales

Les inégalités suivantes sont récapitulées dans le livre de Shorack et Wellner [8].
Le lemme 3 est une application de l'inégalité de Cramer-Chernov et est due à Bennett [1] et à Wellner [10].

Lemme 3 *Soit Z une binomiale de moyenne m . Alors pour tout x positif et tout signe ϵ , on a :*

$$P(\epsilon(Z - m) > x) \leq \exp(-m h(\epsilon x/m))$$

où $h(t) = (1+t) \log(1+t) - t$, pour $t > -1$ et $h(t) = +\infty$ pour $t \leq -1$

En utilisant le lemme 3 et un argument martingale, à partir d'inégalités de Jammes et Shorack, Shorack et Wellner ([7], inégalité 11.1.2) ont prouvé le lemme 4.

Lemme 4 *Pour tout $b \in]0, 1[$ et pour tout x positif :*

$$P(n \sup_{t \in [0, b]} |\hat{F}_n^\circ(t)| > x) \leq 2 \exp(-nb(1-b) h(x/(nb)))$$

Dans le cas d'un pont brownien, Bretagnolle et Massart ([2], inégalité 2.4) ont montré :

Lemme 5 *Pour tout $b \in]0, 1/2[$ et pour tout x positif :*

$$P(\sup_{t \in [0, b]} B^\circ(t) > x) \leq \exp(-x^2 / 2b(1-b))$$

On rappelle l'inégalité classique de Kolmogorov :

Lemme 6

$$\mathbf{P} \left(\sup_{t \in [0,1]} B^o(t) > x \right) = \exp(-2x^2)$$

Inégalité de grandes déviations

L'inégalité suivante est due à Massart [7].

Lemme 7 *Pour tout x positif on a :*

$$\mathbf{P}(n \|\hat{F}_n^o\|_\infty > x) \leq 2 \exp(-2x^2/n)$$

3 Résultat intermédiaire: théorème de Kiefer avec une dépendance en n

Afin d'identifier plus simplement les constantes, nous commençons par montrer le lemme intermédiaire suivant:

Lemme 8 *Pour tout entier $n \geq 3$, il existe une suite de mouvements browniens indépendants $B_1(t), \dots, B_n(t)$ tels que pour tout x positif :*

$$\mathbf{P} \left(\sup_{1 \leq m \leq n} \sup_{0 \leq t \leq 1} |m \hat{F}_m^o(t) - \sum_{i=1}^m B_i^o(t)| > (x + C \log n) \log n \right) \leq K \exp(-\lambda x)$$

où $C = 64$, $K = 0.35$, $\lambda = 1/32$.

3.1 construction dyadique

A partir d'une suite de mouvements browniens indépendants $B_1, \dots, B_n$ on construit un processus de même loi que $(1_{X_1 \leq t}, \dots, 1_{X_n \leq t})$ tel que l'inégalité du lemme 8 soit vérifiée. Par abus d'écriture, on notera encore $\hat{F}_m$, $m = 1, \dots, n$ la nouvelle version du processus empirique. L'existence simultanée de ces processus sur un même espace de probabilité Ω est assurée par un lemme de Skorohod [9] dès que Ω est "suffisamment riche". Pour construire le nouveau processus empirique, il suffit de construire ses accroissements sur des intervalles de longueur $1/n$.

Soit $f : [0, 1] \rightarrow \mathbf{R}$ telle que $f(0) = 0$.

On note $AC(f) = (f(\frac{1}{n}), f(\frac{2}{n}) - f(\frac{1}{n}), \dots, f(1) - f(\frac{n-1}{n}))$.

On note Z le vecteur de $\mathbf{R}^n \otimes \mathbf{R}^n$ $Z = (AC(B_1), \dots, AC(B_n))$ et on construit à partir de Z un vecteur X de même loi que le vecteur de $\mathbf{R}^n \otimes \mathbf{R}^n$ $(AC(1_{X_1 \leq t}), \dots, AC(1_{X_n \leq t}))$

Soit N l'entier tel que $2^{N-1} < n \leq 2^N$.

Nous construisons en fait un processus de même loi que $(1_{X_1 \leq t}, \dots, 1_{X_{2^N} \leq t})$.

Base de $\mathbf{R}^n \otimes \mathbf{R}^n$

Pour tous $l, l' \in \mathbf{N}$, on note $\mathbf{1}_{[l, l']}$ le vecteur de $\mathbf{R}^n$ dont les coordonnées de $l+1$ à l' valent 1, et toutes les autres 0.

$$\text{Soit } \begin{cases} e_{j,k} = \mathbf{1}_{[k2^{j-1}, (k+1)2^j]} \\ 0 \leq j \leq N; 0 \leq k \leq 2^{N-j} - 1 \end{cases}$$

On interprétera le vecteur $e_{i,l} \otimes e_{j,k}$ comme l'indicatrice du rectangle $[k2^{j-N}, (k+1)2^{j-N}] \times [l2^i, (l+1)2^i]$ (inclus dans $[0, 1] \times [0, n]$), l'axe horizontal représentant le temps sur $[0, 1]$ et l'axe vertical les d'individus numérotés de 1 à n , le point $(0, 0)$ étant placé dans le coin supérieur gauche.

$$\text{On pose } \begin{cases} \tilde{e}_{j,k} = e_{j-1,2k} - (e_{j,k}/2) \\ 1 \leq j \leq N; 0 \leq k \leq 2^{N-j} - 1 \end{cases}$$

Alors $\tilde{\mathcal{B}} = \{\tilde{e}_{j,k}; 1 \leq j \leq N, 0 \leq k \leq 2^{N-j} - 1\} \cup \{e_{N,0}\}$ est une base orthogonale de $\mathbf{R}^n$ et $\tilde{\tilde{\mathcal{B}}} = \tilde{\mathcal{B}} \otimes \tilde{\mathcal{B}}$ est une base orthogonale de $\mathbf{R}^n \otimes \mathbf{R}^n$.

Innovations gaussiennes

On note $\langle . | . \rangle$ le produit scalaire usuel.

$$\text{On pose } \begin{cases} W_{j,k}^{i,l} = \langle Z | e_{i,l} \otimes e_{j,k} \rangle \\ 0 \leq i \leq N; 0 \leq l \leq 2^{N-i} - 1 \\ 0 \leq j \leq N; 0 \leq k \leq 2^{N-j} - 1 \end{cases}$$

$W_{j,k}^{i,l}$ s'interprète comme l'accroissement sur $[k2^{j-N}, (k+1)2^{j-N}]$ du mouvement brownien $B_{l2^i+1} + B_{l2^i+2} + \dots + B_{(l+1)2^i}$.

$$\text{On pose } \begin{cases} \tilde{W}_{j,k}^{i,l} = \langle Z | e_{i,l} \otimes \tilde{e}_{j,k} \rangle \\ 0 \leq i \leq N; 0 \leq l \leq 2^{N-i} - 1 \\ 1 \leq j \leq N; 0 \leq k \leq 2^{N-j} - 1 \\ \tilde{\tilde{W}}_{j,k}^{i,l} = \langle Z | \tilde{e}_{i,l} \otimes \tilde{e}_{j,k} \rangle \\ 1 \leq i \leq N; 0 \leq l \leq 2^{N-i} - 1 \\ 1 \leq j \leq N; 0 \leq k \leq 2^{N-j} - 1 \end{cases}$$

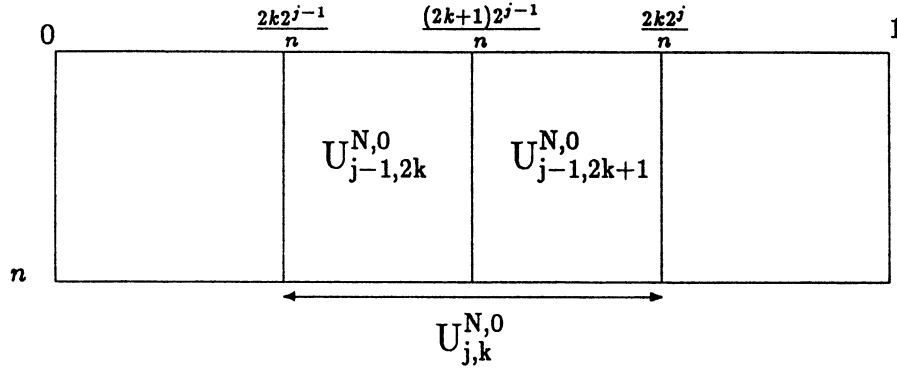
$$\text{On remarque } \begin{cases} \tilde{W}_{j,k}^{i,l} = W_{j-1,2k}^{i,l} - \mathbf{E}(W_{j-1,2k}^{i,l} / W_{j,k}^{i,l}) \\ \tilde{\tilde{W}}_{j,k}^{i,l} = W_{j-1,2k}^{i-1,2l} - \mathbf{E}(W_{j-1,2k}^{i-1,2l} / W_{j,k}^{i,l}, W_{j,k}^{i-1,2l}, W_{j-1,2k}^{i,l}) \\ \mathbf{V}(\tilde{W}_{j,k}^{i,l}) = 2^{i+j-N}/4 \\ \mathbf{V}(\tilde{\tilde{W}}_{j,k}^{i,l}) = 2^{i+j-N}/16 \end{cases}$$

Descente conditionnelle

Soient $\Phi_{n,np,np'}$, Ψ_n , et Φ définies dans les notations.

à l'étape 0

Cette étape correspond à la construction d'une version de la fonction de répartition F_n .
Connaissant le nombre d'observations dans les colonnes de largeur 2^{j-N} , on cherche la loi conditionnelle du nombre d'observations dans les colonnes de largeur 2^{j-1-N} .



On voit que $\mathcal{L}(U_{j-1,2k}^{N,0} / U_{j,k}^{N,0}) = B(U_{j,k}^{N,0}, 1/2)$

On décrit la descente:
$$\begin{cases} U_{N,0}^{N,0} = n \\ U_{j-1,2k}^{N,0} = \Psi_{U_{j,k}^{N,0}}^{-1} \circ \Phi((2^j/4)^{-1/2} \tilde{W}_{j,k}^{N,0}) \\ U_{j-1,2k+1}^{N,0} = U_{j,k}^{N,0} - U_{j-1,2k}^{N,0} \\ \text{pour } j = N, N-1, \dots, 1 \text{ et } 0 \leq k \leq 2^{N-j} - 1 \end{cases}$$

A la fin de cette étape, on a reconstruit la statistique du n -échantillon au pas de subdivision $1/n$.

A l'étape $N - (i - 1)$ pour tout $i = N, N - 1, \dots, 1$

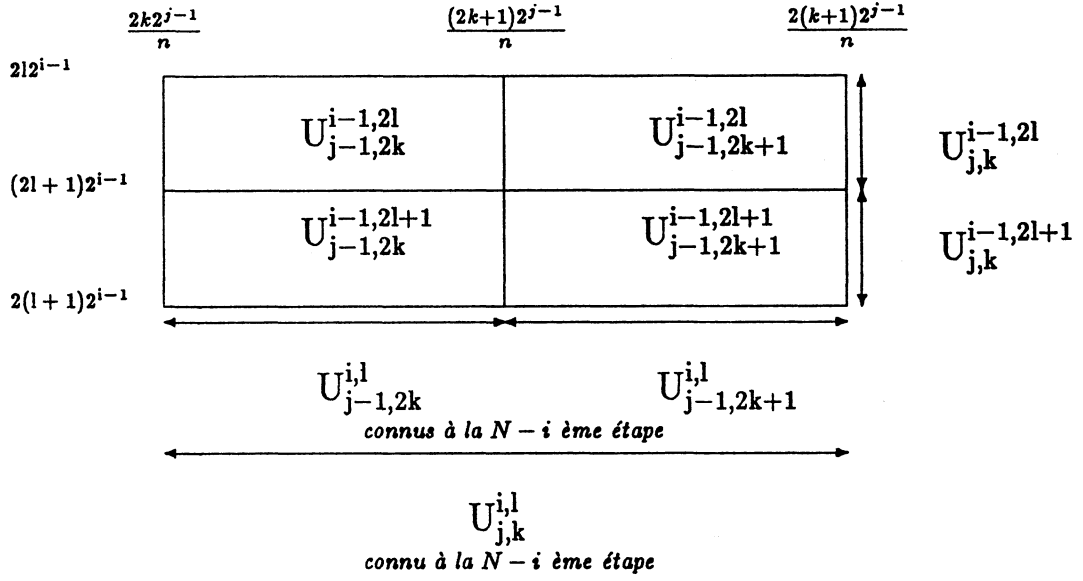
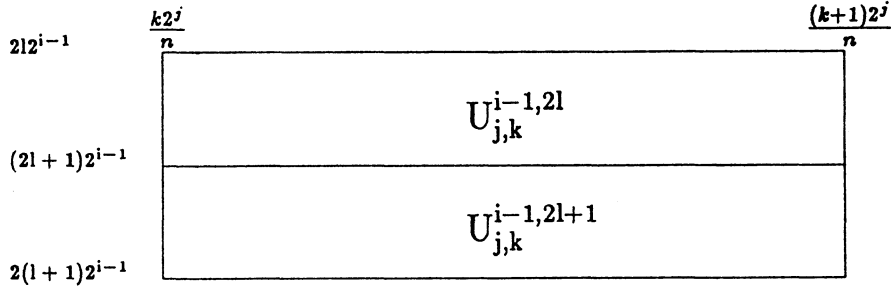
A la fin de la $N - i$ ème étape, on connaît le nombre d'observations noté $U_{j,k}^{N,0}$ sur les rectangles du type $]k2^{j-N}, (k+1)2^{j-N}] \times]l2^i, (l+1)2^i]$, où seul i est fixé.

A la $N - (i - 1)$ ème étape, on va répartir les 2^i observations de la bande $]0, 1] \times]l2^i, (l+1)2^i]$ en deux groupes de taille 2^{i-1} sur les bandes $]0, 1] \times]2l2^{i-1}, (2l+1)2^{i-1}]$ et $]0, 1] \times](2l+1)2^{i-1}, 2(l+1)2^{i-1}]$.

La descente conditionnelle consiste à diviser de façon dyadique l'intervalle de temps $]0, 1]$ sur ces deux bandes.

On suppose connu les nombres d'observations notés $U_{j,k}^{i-1,2l}$ et $U_{j,k}^{i-1,2l+1}$ sur les rectangles $]k2^{j-N}, (k+1)2^{j-N}] \times]2l2^{i-1}, (2l+1)2^{i-1}]$ et $]k2^{j-N}, (k+1)2^{j-N}] \times](2l+1)2^{i-1}, 2(l+1)2^{i-1}]$.

Nous décrivons entièrement la division suivante:



On voit que $\mathcal{L}(U_{j-1,2k}^{i-1,2l} / U_{j,k}^{i,l}, U_{j,k}^{i-1,2l}, U_{j-1,2k}^{i,l}) = H(U_{j,k}^{i,l}, U_{j,k}^{i-1,2l}, U_{j-1,2k}^{i,l})$
et de même pour les trois autres.

On continue la division dyadique jusqu'aux rectangles $]k/n, (k+1)/n[\times]2l2^{i-1}, 2(l+1)2^{i-1}[$
et $]k/n, (k+1)/n[\times]2l2^{i-1}, 2(l+1)2^{i-1}[$ puis on passe à l'étape $N - (i - 2)$

A l'étape $N - (i - 1)$ on décrit la descente:

$$\text{Pour tout } l = 0, \dots, 2^{N-(i-1)} \quad \left\{ \begin{array}{l} U_{N,0}^{i-1,l} = 2^{i-1} \\ U_{j-1,2k}^{i-1,2l} = \Phi_{U_{j,k}^{i,l}, U_{j,k}^{i-1,2l}, U_{j-1,2k}^{i,l}}^{-1} \circ \Phi((2^{i+j-N}/16)^{-1/2} \tilde{W}_{j,k}^{i,l}) \\ U_{j-1,2k}^{i-1,2l+1} = U_{j-1,2k}^{i,l} - U_{j-1,2k}^{i-1,2l} \\ U_{j-1,2k+1}^{i-1,2l} = U_{j,k}^{i-1,2l} - U_{j-1,2k}^{i-1,2l} \\ U_{j-1,2k+1}^{i-1,2l+1} = U_{j-1,2k+1}^{i,l} - U_{j-1,2k+1}^{i-1,2l} \\ \text{pour } j = N, N-1, \dots, 1 \text{ et } 0 \leq k \leq 2^{N-j} - 1 \end{array} \right.$$

On prouve par récurrence que le vecteur de $\mathbf{R}^n \otimes \mathbf{R}^n$ $X = (U_0^{0,0}, U_0^{0,1}, \dots, U_0^{0,n-1})$, avec $U_{0,0}^{0,l} = (U_{0,0}^{0,l}, U_{0,1}^{0,l}, \dots, U_{0,n-1}^{0,l})$ a même loi que le vecteur $(AC(1_{X_1 \leq t}), \dots, AC(1_{X_n \leq t}))$.

On peut vérifier les propriétés suivantes:

- $\mathcal{L}(U_{j-1,2k}^{i,l}/U_{j,k}^{i,l}) = B(U_{j,k}^{i,l}, 1/2)$
- $\mathcal{L}(U_{j,k}^{i-1,2l}/U_{j,k}^{i,l}) = H(2^i, 2^{i-1}, U_{j,k}^{i,l})$
- $\mathcal{L}(U_{j,0}^{i,l}, U_{j,1}^{i,l}, \dots, U_{j,2^{N-j}-1}^{i,l}) = \mathcal{M}_{2^{N-j}}(2^i, \underbrace{\frac{2^j}{n}, \dots, \frac{2^j}{n}}_{2^{N-j}})$

$$\text{Nous posons maintenant } \left\{ \begin{array}{l} \tilde{U}_{j,k}^{i,l} = \langle X \mid e_{i,l} \otimes \tilde{e}_{j,k} \rangle \\ \tilde{\tilde{U}}_{j,k}^{i,l} = \langle X \mid \tilde{e}_{i,l} \otimes \tilde{e}_{j,k} \rangle \end{array} \right.$$

On rappelle que par abus d'écriture, on note encore $\hat{F}_m$, $m = 1, \dots, n$ le nouveau processus empirique reconstitué à partir de X .

Il suffit de contrôler la quantité

$$\mathcal{P} = \mathbf{P} \left(\sup_{1 \leq m \leq n} \sup_{0 \leq t \leq 1} |m \hat{F}_m^o(t) - \sum_{i=1}^m B_i^o(t)| > \log n(x + 64 \log n) \right)$$

3.2 contrôle de $\mathcal{P}$

- Si $x + 64 \log n \geq n/12$

En utilisant les lemmes 6 et 7, on obtient : $\mathcal{P} \leq 4n \exp(-(x + 64 \log n)^2 (\log n)^2 / (2n))$

En utilisant $x + 64 \log n \geq n/12$, on a:

$$\mathcal{P} \leq 4 \exp \left(-\frac{(\log n)^2}{24} x + (\log n) \left(1 - \frac{64 \log^2 n}{24} \right) \right)$$

Avec $n \geq 3$ on obtient:

$$P \leq 0.95 \exp(-x/20) \quad (1)$$

- Si $x + 64 \log n \leq n/12$

On a supposé $2^{N-1} < n \leq 2^N$

Choix de l'échelle L_o en fonction de la taille de l'échantillon

On remarque que si $x + 64 \log n \leq n/12$ alors $n \geq n_o$ où $n_o = 6500$.

Soit L_o , l'entier tel que $2^{L_o-1} \leq 0.57(x + 64 \log n) < 2^{L_o}$.

Pour tout entier $m = 1, \dots, n$, on définit $\Pi_{L_o}(m)$, le point $l_o 2^{L_o}$ de la grille $\{l 2^{L_o}, 1 \leq l \leq 2^{N-L_o}\}$ tel que $(l_o - 1)2^{L_o} < m \leq l_o 2^{L_o}$.

On a

$$P \leq Q + P$$

où

$$\begin{aligned} Q &= \mathbf{P} \left(\max_{1 \leq m \leq n} \|m \hat{F}_m^o - \Pi_{L_o}(m) \hat{F}_{\Pi_{L_o}(m)}^o\|_\infty \geq \left(\frac{1-\lambda}{2}\right)(\log n)(x + 64 \log n) \right) \\ &\quad + \mathbf{P} \left(\max_{1 \leq m \leq n} \left\| \sum_{s=1}^m B_s^o - \sum_{s=1}^{\Pi_{L_o}(m)} B_s^o \right\|_\infty \geq \left(\frac{1-\lambda}{2}\right)(\log n)(x + 64 \log n) \right) \\ P &= \mathbf{P} \left(\max_{1 \leq l_o \leq 2^{N-L_o}} \|l_o 2^{L_o} \hat{F}_{l_o 2^{L_o}}^o - \sum_{s=1}^{l_o 2^{L_o}} B_s^o\|_\infty \geq \lambda(\log n)(x + 64 \log n) \right) \\ &\quad \text{où } \lambda \in [0, 1] \end{aligned}$$

3.2.1 Contrôle de Q

$Q \leq Q_1 + Q_2$ où

$$\begin{aligned} Q_1 &\leq 2^N \max_{1 \leq m < 2^{L_o}} \mathbf{P} \left(\|m \hat{F}_m^o\|_\infty \geq \left(\frac{1-\lambda}{2}\right)(\log n)(x + 64 \log n) \right) \\ Q_2 &\leq 2^N \max_{1 \leq m < 2^{L_o}} \mathbf{P} \left(\left\| \sum_{s=1}^m B_s^o \right\|_\infty \geq \left(\frac{1-\lambda}{2}\right)(\log n)(x + 64 \log n) \right) \end{aligned}$$

En appliquant les lemmes 6 et 7 à Q_1 et Q_2 puis en utilisant la définition de L_o , on obtient

$$Q \leq 8 \exp \left(-\frac{1}{0.57} \left(\frac{1-\lambda}{2} \right)^2 (\log n)^2 x + (\log n) \left(1 - 64 \left(\frac{1-\lambda}{2} \right)^2 \frac{(\log n)^2}{0.57} \right) \right)$$

En utilisant $n \geq 6500$, $x \geq 0$ et en imposant $x(\log n)^2(1-\lambda)^2/(4 \times 0.57) \geq (x/32)$, on obtient $\lambda \leq 0.969$.

En choisissant $\lambda = 0.969$ et en utilisant la définition de L_o , on obtient :

$$Q \leq 1.3 \times 10^{-3} \exp(-x/32) \quad (2)$$

3.2.2 Contrôle de P

$$P \leq 2^{N-L_o} \max_{1 \leq l_o \leq 2^{N-L_o}} P_{l_o} \quad \text{où}$$

$$P_{l_o} = \mathbf{P} \left(\|l_o 2^{L_o} \hat{F}_{l_o 2^{L_o}}^o - \sum_{s=1}^{l_o 2^{L_o}} B_s^o\|_{\infty} \geq 0.969(\log n) (x + 64 \log n) \right)$$

On fixe l_o et l'on contrôle P_{l_o}

1. Soit $l(L)$ l'entier tel que $l_o 2^{L_o} \in]2^L l(L), 2^L(l(L) + 1)]$. On note :

$$\left\{ \begin{array}{ll} \tilde{e}_L = \tilde{e}_{L,l(L)} & \text{et } I_L =]l(L)2^L, (l(L) + 1)2^L] \\ I_{L-1,H} =]2l(L)2^{L-1}, (2l(L) + 1)2^{L-1}] & \text{et } I_{L-1,B} =](2l(L) + 1)2^{L-1}, (2l(L) + 2)2^{L-1}] \\ 2^{L-1} \hat{F}_{L-1,H}^o(t) = \sum_{s \in I_{L-1,H}} (1_{X_s \leq t} - t), & \text{et } 2^{L-1} \hat{F}_{L-1,B}^o(t) = \sum_{s \in I_{L-1,B}} (1_{X_s \leq t} - t) \\ B_{L-1,H}^o = \sum_{s \in I_{L-1,H}} B_s^o, & \text{et } B_{L-1,B}^o = \sum_{s \in I_{L-1,B}} B_s^o. \\ U_{j,k}^{L-1,H} = \langle X | 1_{I_{L-1,H}} \otimes e_{j,k} \rangle & \text{et } U_{j,k}^{L-1,B} = \langle X | 1_{I_{L-1,B}} \otimes e_{j,k} \rangle \end{array} \right.$$

Remarque

I_L est l'intervalle de longueur 2^L contenant $l_o 2^{L_o}$, $I_{L-1,H}$ est l'intervalle égal à la moitié supérieure de I_L et donc de longueur 2^{L-1} . $\sqrt{2^{L-1}} \hat{F}_{L-1,H}^o$ correspond au pont empirique sur la population située dans $I_{L-1,H}$.

2. Le développement de $1_{]0, l_o 2^{L_o}]}$ sur la base orthogonale $\tilde{B}$ donne :

$$1_{]0, l_o 2^{L_o}] = \sum_{L > L_o}^N c_L \tilde{e}_L + \frac{l_o 2^{L_o}}{2^N} e_{N,0} \quad \text{où } 0 \leq c_L \leq 1, L = L_o + 1, \dots, N.$$

La preuve est facile en notant que $1_{]0, l_o 2^{L_o}]}$ est orthogonal à $\tilde{e}_{L,l}$ si $l \neq l(L)$ ou $L \leq L_o$ et que $0 \leq \langle 1_{]0, l_o 2^{L_o}] | \tilde{e}_L \rangle \leq 2^{L-2} = \langle \tilde{e}_L | \tilde{e}_L \rangle$

3. De plus : $\sum_{L > L_o} c_{L,l(L)} \leq (2/3)(N - L_o + 0.5)$. En effet :

$$c_{L,l(L)} = \frac{\langle \tilde{e}_{L,l(L)} | 1_{]0, l_o 2^{L_o}] \rangle}{\|\tilde{e}_{L,l(L)}\|^2} = 2^{1-L} d(l_o 2^{L_o}, 2^L N)$$

où N est l'ensemble des entiers naturels.

$$\text{Or } d(l_o 2^{L_o}, 2^L N) + d(l_o 2^{L_o}, 2^{L-1}(2N + 1)) = 2^{L-1}$$

Donc $\sum_{L>L_o}^N 2^{1-L} (d(l_o 2^{L_o}, 2^L \mathbf{N}) + d(l_o 2^{L_o}, 2^{L-1}(2\mathbf{N} + 1))) = N - L_o$

En utilisant $d(l_o 2^{L_o}, 2^{L-1}(2\mathbf{N} + 1)) \geq d(l_o 2^{L_o}, 2^{L-1}\mathbf{N})$, on obtient :

$$\begin{aligned} N - L_o &\geq \left(\sum_{L>L_o}^N (2^{1-L} + 2^{-L}) d(l_o 2^{L_o}, 2^L \mathbf{N}) \right) - 2^{-N} d(l_o 2^{L_o}, 2^N \mathbf{N}) \\ &\geq (3/2) \sum_{L>L_o} c_{L,l(L)} - (1/2) \end{aligned}$$

d'où le résultat .

En utilisant 2 et 3 on obtient :

$$\begin{aligned} \|l_o 2^{L_o} \hat{F}_{l_o 2^{L_o}}^\circ - \sum_{s=1}^{l_o 2^{L_o}} B_s^\circ\|_\infty &\leq (2/3)(N - L_o + 0.5) (1/2) \max_{L_o < L \leq N} (\|2D_L\|_\infty) \\ &\quad + \|2^N \hat{F}_{2^N}^\circ - \sum_{s=1}^{2^N} B_s^\circ\|_\infty \quad (3) \\ \text{où } 2D_L(t) &= ((2^{L-1} \hat{F}_{L-1,H}^\circ - B_{L-1,H}^\circ) - (2^{L-1} \hat{F}_{L-1,B}^\circ - B_{L-1,B}^\circ))(t) \end{aligned}$$

En utilisant 3, $L_o \geq 2$ et $\log n > (N - 1) \log 2$, on a :

$$\begin{aligned} P_{l_o} &\leq \mathbf{P} \left((2/3) \left(\frac{N-1}{2} \right) \max_{L_o < L \leq N} \|2D_L\|_\infty \geq (N-1) 0.969(1-\alpha)(\log 2)(x + 64 \log n) \right) \\ &\quad + \mathbf{P} \left(\|2^N \hat{F}_{2^N}^\circ - \sum_{s=1}^{2^N} B_s^\circ\|_\infty \geq 0.969\alpha(\log n)(x + 64 \log n) \right) \\ \text{où } \alpha &\in [0, 1] \end{aligned}$$

Les accroissements de $2^N \hat{F}_{2^N}^\circ$ sont fonction des accroissements de $\sum_{s=1}^{2^N} B_s$ (Section 3 - descente conditionnelle - étape 0). Cette construction correspond à la construction hongroise unidimensionnelle utilisée par Komlos, Major et Tusnady [6] et Bretagnolle et Massart [2] pour démontrer le théorème 1.

On a donc pour tout $X \geq 0$: $\mathbf{P} \left(\|2^N \hat{F}_{2^N}^\circ - \sum_{s=1}^{2^N} B_s^\circ\|_\infty \geq X + 12 \log(2^N) \right) \leq 2 \exp(-X/6)$

On prend $X = (0.969 \times \alpha)(\log n)(x + 64 \log n) - 12 \log(2^N)$

En utilisant la définition de L_o , on obtient :

$$\begin{aligned} 2^{N-L_o} \mathbf{P} \left(\|2^N \hat{F}_{2^N}^\circ - \sum_{s=1}^{2^N} B_s^\circ\|_\infty \geq (\log n)(x + 64 \log n)(0.969 \times \alpha) \right) &\leq \\ \frac{2 \times 2}{0.57(x + 64 \log n)(\log n)} \exp \left(-\frac{X}{6} + (\log n) + (\log \log n) \right) \end{aligned}$$

Pour $\alpha = 0.038$, en utilisant $n \geq 6500$, $x \geq 0$, on obtient:

$$64(\log n)^2(0.969 \times \alpha) - 12(\log n) - 12(\log 2) \geq 6(\log n) + 6(\log \log n)$$

Donc

$$2^{N-L_0} \mathbf{P} \left(\left\| 2^N \hat{F}_{2^N}^0 - \sum_{s=1}^{2^N} B_s^0 \right\|_{\infty} \geq (\log n)(x + 64 \log n)(0.969 \times 0.038) \right) \leq 1.5 \times 10^{-3} \exp \left(-\frac{x}{32} \right)$$

On a donc:

$$P \leq 1.5 \times 10^{-3} \exp(-x/32) + (N - L_0) 2^{N-L_0} \max_{1 \leq l_0 \leq 2^{N-L_0}} \max_{L_0 < L \leq N} P_{l_0}^L \quad (4)$$

$$\text{où } P_{l_0}^L = \mathbf{P} \left((2/3)(1/2) \left\| (2^{L-1} \hat{F}_{L-1,H}^0 - B_{L-1,H}^0) - (2^{L-1} \hat{F}_{L-1,B}^0 - B_{L-1,B}^0) \right\| \geq (x + 60 \log n)A \right) \\ \text{avec } A = 0.969 \times 0.962 \times \log 2$$

On fixe $L > L_0$ et l'on contrôle $P_{l_0}^L$

Régularisation à l'échelle $M(L)$ du temps

Soit $M(L)$, l'entier tel que $2^{L+M(L)-N-1} \leq 0.57(x + 60 \log n) < 2^{L+M(L)-N}$.

Pour tout $t \in]0, 1]$, on définit $\Pi_L(t)$ le point $k2^{M(L)-N}$ de la grille $\{k2^{M(L)-N}, 1 \leq k \leq 2^{N-M(L)}\}$ tel que $(k-1)2^{M(L)-N} < t \leq k2^{M(L)-N}$.

Alors $P_{l_0}^L \leq P_1(l_0, L) + P'_1(l_0, L) + P_2(l_0, L) + P((\Theta_L^\epsilon)^c)$ où

$$\Theta_L^\epsilon = \{ (1-\epsilon)2^{L-1+j-N} \leq U_{j,k}^{L-1,H} \leq (1+\epsilon)2^{L-1+j-N} \quad j > M(L) \quad k = 0, \dots, 2^{N-j} - 1 \} \\ \cap \{ (1-\epsilon)2^{L-1+j-N} \leq U_{j,k}^{L-1,B} \leq (1+\epsilon)2^{L-1+j-N} \quad j > M(L) \quad k = 0, \dots, 2^{N-j} - 1 \} \\ \cap \{ (1-\epsilon)2^{L+j-N} \leq U_{j,k}^L \leq (1+\epsilon)2^{L+j-N} \quad j = M(L) \quad k = 0, \dots, 2^{N-j} - 1 \}$$

où $\epsilon \in [0, 1]$

$$P_1(l_0, L) \leq 2^{N-M(L)} \mathbf{P} \left((2/3) \sup_{0 \leq t < 2^{M(L)-N}} |2^{L-1} \hat{F}_{L-1,H}^0(t)| \geq \beta \times A(x + 64 \log n) \right)$$

$$P'_1(l_0, L) \leq 2^{N-M(L)} \mathbf{P} \left((2/3) \sup_{0 \leq t < 2^{M(L)-N}} |B_{L-1,H}^0(t)| \geq \beta' \times A(x + 64 \log n) \right)$$

où $\beta \in]0, 1[$ et $\beta + \beta' \in]0, 1[$

$$P_2(l_o, L) = \mathbf{P} \left(\left\{ \max_{1 \leq k \leq 2^{N-M(L)}} \left| \frac{2}{3} D_L^k \right| \geq A(x + 64 \log n)(1 - \beta - \beta') \right\} \cap \Theta_L^\epsilon \right)$$

$$\text{où } D_L^k = 1/2 \left((2^{L-1} \hat{F}_{L-1,H}^\circ - B_{L-1,H}^\circ) - (2^{L-1} \hat{F}_{L-1,B}^\circ - B_{L-1,B}^\circ) \right) (k 2^{M(L)-N})$$

$P_1(l_o, L)$ et $P'_1(l_o, L)$ mesurent séparément les petites fluctuations de $\hat{F}_{L-1,H}^\circ$ et de $B_{L-1,H}^\circ$. Une évaluation en loi sur des petits intervalles du pont empirique et du pont brownien suffisent pour obtenir le contrôle exponentiel.

La construction des vecteurs X et Z et les lemmes 1 et 2 entraînent que $|(2^{L-1} \hat{F}_{L-1,H}^\circ - B_{L-1,H}^\circ) - (2^{L-1} \hat{F}_{L-1,B}^\circ - B_{L-1,B}^\circ)(k 2^{M(L)-N})|$ est majoré par une variable de loi chi-2 sur l'évènement Θ_L^ϵ qui élimine les trop grandes déviations de U .

Contrôle de $P((\Theta_L^\epsilon)^c)$

En utilisant $U_{j,k}^L = U_{j-1,2k}^L + U_{j-1,2k+1}^L$, et plus précisément en remarquant que tout $U_{j,k}^L, j > M(L)$ est une somme de $U_{M(L)+1,k}^L$, il vient:

$$\begin{aligned} (\Theta_L^\epsilon)^c \subset & \bigcup_{0 \leq k < 2^{N-M(L)-1}} \{U_{M(L)+1,k}^{L-1,H} > (1 + \epsilon) 2^{L+M(L)-N}\} \\ & \bigcup_{0 \leq k < 2^{N-M(L)-1}} \{U_{M(L)+1,k}^{L-1,H} < (1 - \epsilon) 2^{L+M(L)-N}\} \\ & \bigcup_{0 \leq k < 2^{N-M(L)-1}} \{U_{M(L)+1,k}^{L-1,B} > (1 + \epsilon) 2^{L+M(L)-N}\} \\ & \bigcup_{0 \leq k < 2^{N-M(L)-1}} \{U_{M(L)+1,k}^{L-1,B} < (1 - \epsilon) 2^{L+M(L)-N}\} \\ & \bigcup_{0 \leq k < 2^{N-M(L)}} \{U_{M(L),k}^L > (1 + \epsilon) 2^{L+M(L)-N}\} \\ & \bigcup_{0 \leq k < 2^{N-M(L)}} \{U_{M(L),k}^L < (1 - \epsilon) 2^{L+M(L)-N}\} \end{aligned}$$

En appliquant le lemme 1 et en utilisant la définition de $M(L)$ on obtient :

$$\begin{aligned} \mathbf{P}((\Theta_L^\epsilon)^c) \leq & 2^{N-M(L)+1} \exp(-0.57(x + 64 \log n)h(\epsilon)) \\ & + 2^{N-M(L)+1} \exp(-0.57(x + 64 \log n)h(-\epsilon)) \end{aligned}$$

Donc en utilisant la définition de L_o et de $M(L)$, on obtient:

$$\begin{aligned} (N - L_o) \times 2^{N-L_o} \mathbf{P}((\Theta_L^\epsilon)^c) \leq & \frac{8(N - L_o) \exp(-0.57h(\epsilon)x + \log n(-0.57 \times 64h(\epsilon) + 2))}{(0.57)^2(x + 64 \log n)^2} \\ & + \frac{8(N - L_o) \exp(-0.57h(-\epsilon)x + \log n(-0.57 \times 64h(-\epsilon) + 2))}{(0.57)^2(x + 64 \log n)^2} \end{aligned}$$

En imposant la condition $-0.57 \times 64h(\epsilon) + 2 \leq 0$, on obtient $\epsilon \geq 0.35$. On choisit $\epsilon = 0.35$ et on pose désormais $\Theta_L^\epsilon = \Theta^L$. En remarquant $L_o \geq 2$, en utilisant $N - L_o \leq (\log n)/(\log 2)$ et $x \geq 0$, on obtient:

$$(N - L_o)2^{N-L_o} \mathbf{P}(\Theta_L^\epsilon) \leq \frac{16}{(0.57 \times 64)^2 (\log 2)(\log n)} \exp\left(-\frac{x}{32}\right) \quad (5)$$

Contrôle de $P_1(l_o, L)$ et $P'_1(l_o, L)$

On pose $b = 2^{M(L)-N}$. On remarque $b \leq 1/2$.

Les lemmes 4 et 5 entraînent :

$$\begin{aligned} P_1(l_o, L) &\leq 2^{N-M(L)+1} \exp\left(-2^{L-1}b(1-b)h\left(\frac{3A\beta(x+64\log n)}{2^L b}\right)\right) \\ P'_1(l_o, L) &\leq 2^{N-M(L)+1} \exp\left(-\frac{(3A\beta'(x+64\log n))^2}{8b(1-b)2^{L-1}}\right) \end{aligned}$$

Du fait de la définition de $M(L)$ et de L_o , on a

$$\begin{aligned} (N - L_o)2^{N-L_o} (P_1(l_o, L) + P'_1(l_o, L)) &\leq \\ C(n) \left\{ \exp\left(-\frac{0.57}{4}h\left(\frac{3A}{2 \times 0.57}\beta\right)x + (\log n)\left(-64\frac{0.57}{4}h\left(\frac{3A}{1.14}\beta\right) + 2\right)\right) \right\} \\ + C(n) \left\{ \exp\left(-\frac{(3A\beta')^2}{8 \times 0.57}x + (\log n)\left(-64\frac{(3A\beta')^2}{8 \times 0.57} + 2\right)\right) \right\} \\ \text{où } C(n) &= \frac{8(N - L_o)}{(0.57(x + 64\log n))^2} \end{aligned}$$

On impose sur β, β' , les conditions :

$$\begin{aligned} -64 \times \frac{0.57}{4} \times h\left(\frac{3A}{1.14}\beta\right) + 2 &\leq 0 \\ -64 \times \frac{(3A\beta')^2}{8 \times 0.57} + 2 &\leq 0 \end{aligned}$$

soit $\beta \geq 0.435$ et $\beta' \geq 0.195$. En choisissant $\beta = 0.435$ et $\beta' = 0.195$ et en utilisant $N - L_o \leq (\log n)/(\log 2)$, on obtient :

$$\begin{aligned} (N - L_o)2^{N-L_o} \max_{1 \leq l_o \leq 2^{N-L_o}} \max_{L_o < L \leq N} P_1(l_o, L) + P'_1(l_o, L) &\leq \\ \frac{16}{(0.57 \times 64)^2 (\log 2)(\log n)} \exp\left(-\frac{x}{32}\right) &\quad (6) \end{aligned}$$

En résumé, pour $\epsilon = 0.35$ et en utilisant $n \geq 6500$, on obtient en regroupant 5 et 6 :

$$\begin{aligned} (N - L_o)2^{N-L_o} \max_{1 \leq l_o \leq 2^{N-L_o}} \max_{L_o < L \leq N} P_1(l_o, L) + P'_1(l_o, L) + P(\Theta_L^\epsilon) &\leq \\ 2.1 \times 10^{-3} \exp\left(-\frac{x}{32}\right) &\quad (7) \end{aligned}$$

En regroupant 4 et 7 on a:

$$P \leq 7 \times 10^{-3} \exp(-x/32) + (N - L_o) 2^{N-L_o} \max_{1 \leq l_o \leq 2^{N-L_o}} \max_{L_o < L \leq N} P_2(l_o, L) \quad (8)$$

$$\text{où } P_2(l_o, L) = P \left(\left\{ \max_{1 \leq k \leq 2^{N-M(L)}} |D_L^k| \geq r(x + 64 \log n) \right\} \cap \Theta_L \right) \quad (9)$$

$$\text{avec } r = \frac{3}{2} \times 0.969 \times 0.962 \times 0.37 \times (\log 2) \quad (10)$$

Contrôle de $P_2(l_o, L)$

$$P_2(l_o, L) \leq 2^{N-M(L)} \max_{1 \leq k \leq 2^{N-M(L)}} P_2^k(l_o, L) \quad \text{où}$$

$$P_2^k(l_o, L) = P \left(|D_L^k| \geq r(x + 64 \log n) \cap \Theta_L \right)$$

où l'on rappelle : $2D_L^k = \left((2^{L-1} \hat{F}_{L-1,H}^\circ - B_{L-1,H}^\circ) - (2^{L-1} \hat{F}_{L-1,B}^\circ - B_{L-1,B}^\circ) \right) (k2^{M(L)-N})$
et r défini en 10 .

On fixe k et l'on contrôle $P_2^k(l_o, L)$

Etude de $P_2^k(l_o, L)$

Soit $k(j)$ l'entier tel que $k2^{M(L)-N} \in]k(j)2^{j-N}, (k(j)+1)2^{j-N}]$. On note $\tilde{e}_j = \tilde{e}_{j,k(j)}$

En utilisant le développement de $1_{]0, k2^{M(L)-N}]}$ sur $\tilde{B}$, on obtient:

$$1_{]0, k2^{M(L)-N}]} = \sum_{j=M(L)+1}^N c_j \tilde{e}_j + \frac{k2^{M(L)}}{n} e_{N,0} \quad \text{où } 0 \leq c_j \leq 1 \quad (a)$$

De manière évidente on a :

$$2^{L-1} \left(\hat{F}_{L-1,H}^\circ - \hat{F}_{L-1,B}^\circ \right) (k2^{M(L)-N}) = 2 \langle X | \tilde{e}_L \otimes 1_{]0, k2^{M(L)-N}]} \rangle \quad (b)$$

$$\begin{aligned} \left(B_{L-1,H}^\circ - B_{L-1,B}^\circ \right) (k2^{M(L)-N}) &= 2 \left(\langle Z | \tilde{e}_L \otimes 1_{]0, k2^{M(L)-N}]} \rangle \right. \\ &\quad \left. - k2^{M(L)-N} \langle Z | \tilde{e}_L \otimes e_{N,0} \rangle \right) \quad (c) \end{aligned}$$

En utilisant (a,b,c) on obtient la majoration suivante de D_L^k :

$$\begin{aligned} |D_L^k| &= \left| \sum_{j>M(L)} c_j \langle X - Z | \tilde{e}_L \otimes \tilde{e}_j \rangle \right| \\ &\leq \sum_{j>M(L)} \left| \tilde{U}_j^L - \tilde{W}_j^L \right| \end{aligned} \quad (11)$$

Etude de $|D_k^L|$

On pose $\xi_j^L = (2^{L+j-N}/16)^{-1/2} \tilde{W}_j^L$.

Les variables ξ_j^L $j = 1, \dots, N$ sont des variables gaussiennes centrées réduites mutuellement indépendantes (Section 3 : Innovations gaussiennes).

On a :

$$\tilde{U}_j^L = U_{j-1,2k(j)}^{L-1,H} - \frac{1}{2}U_{j,k(j)}^{L-1,H} - \frac{1}{2}U_{j-1,2k(j)}^L + \frac{1}{4}U_{j,k(j)}^L$$

Les lois et les lois conditionnelles des variables qui composent $\tilde{U}_j^L$ sont connues (Section 3 - Descente conditionnelle).

On notera Δ les variables $U_{j,k(j)}^{L-1,H}$, $U_{j,k(j)}^{L-1,B}$, $U_{j-1,2k(j)}^L$ et $U_{j-1,2k(j)+1}^L$ recentrées par leur espérance conditionnelle par rapport à $U_{j,k(j)}^L$:

$$\begin{cases} \Delta_{j,k(j)}^{L-1,H} &= \Delta_{j,k(j)}^{L-1,B} &= U_{j,k(j)}^{L-1,H} - (U_{j,k(j)}^L)/2 \\ \Delta_{j-1,2k(j)}^L &= \Delta_{j-1,2k(j)+1}^L &= U_{j-1,2k(j)}^L - (U_{j,k(j)}^L)/2 \end{cases} \quad (12)$$

On remarque $k(j-1) = 2k(j)$ ou $k(j-1) = 2k(j) + 1$. Par nécessité de calcul, on se ramènera toujours à $\Delta_{j,k(j)}^{L-1,H}$ et $\Delta_{j-1,k(j-1)}^L$. Donc, lorsque aucune confusion ne sera possible, on notera :

$$\Delta_j^{L-1} = \Delta_{j,k(j)}^{L-1,H} \text{ et } \Delta_{j-1}^L = \Delta_{j-1,k(j-1)}^L$$

On décompose $\tilde{U}_j^L$:

$$\tilde{U}_j^L = \gamma_j^L + \frac{\Delta_j^{L-1}\Delta_{j-1}^L}{U_{j,k(j)}^L} \quad \text{avec} \quad \gamma_j^L = \left(U_{j-1,2k(j)}^{L-1,H} - \frac{U_{j,k(j)}^{L-1,H}U_{j-1,2k(j)}^L}{U_{j,k(j)}^L} \right)$$

On décompose $\tilde{W}_j^L$:

$$\tilde{W}_j^L = \sigma_j^L \xi_j^L + \left(\sqrt{\frac{2^{L+j-N}}{16}} - \sigma_j^L \right) \xi_j^L$$

avec $\sigma_j^{L^2} = U_{j,k(j)}^L \left(\frac{U_{j,k(j)}^{L-1,H}}{U_{j,k(j)}^L} \right) \left(\frac{U_{j-1,2k(j)}^L}{U_{j,k(j)}^L} \right) \left(\frac{U_{j,k(j)}^{L-1,B}}{U_{j,k(j)}^L} \right) \left(\frac{U_{j-1,2k(j)+1}^L}{U_{j,k(j)}^L} \right)$

Posons $\omega = |\xi_j^L| \left| \sqrt{(2^{L+j-N}/16)} - \sigma_j^L \right|$. En utilisant $U_{j,k(j)}^L \geq (1-\epsilon)2^{L+j-N}$ sur Θ_L et $|ab| \leq (a^2 + b^2)/2$, on obtient :

$$|\tilde{U}_j^L - \tilde{W}_j^L| \leq |\gamma_j^L - \sigma_j^L \xi_j^L| + \frac{2^{-(L+j-N)}}{2(1-\epsilon)} \left((\Delta_j^{L-1})^2 + (\Delta_{j-1}^L)^2 \right) + \omega \quad (13)$$

• Pour contrôler $|\mathcal{V}_j^L - \sigma_j^L \xi_j^L|$, la construction permet d'utiliser l'inégalité 2 du lemme 2 et plus précisément le second cas particulier de ce lemme pourvu que les contraintes soient vérifiées.

On vérifie la contrainte sur $|\delta\delta'|$: ($j > M(L)$)

$$\delta = \frac{U_{j,k(j)}^{L-1,H} - U_{j,k(j)}^{L-1,B}}{U_{j,k(j)}^L} \quad \text{et} \quad \delta' = \frac{U_{j-1,2k(j)}^L - U_{j-1,2k(j)+1}^L}{U_{j,k(j)}^L}$$

Sur l'évènement Θ_L on a

$$(1-\epsilon)2^{L'+j-N} \leq U_j^{L',H}, U_j^{L',B} \leq (1+\epsilon)2^{L'+j-N} \text{ pour } L' \in \{L, L-1\} \text{ et } j+L' \geq M(L)+L$$

On peut écrire:

$$\begin{aligned} U_j^{L-1,H} &= (1+a)2^{L+j-1-N}, \quad U_j^{L-1,B} = (1+b)2^{L+j-1-N} \\ U_{j-1,2k(j)}^L &= (1+c)2^{L+j-1-N}, \quad U_{j-1,2k(j)+1}^L = (1+d)2^{L+j-1-N} \\ \text{pour } j > M(L) \text{ et } |a|, |b|, |c|, |d| &\leq \epsilon \end{aligned}$$

On calcule $\max_{\Theta_L} |\delta| = \max_{|a|, |b| \leq \epsilon} |a-b|/(2+a+b)$ en prenant $b \leq a$. (le calcul de $\max_{\Theta_L} |\delta'|$ est identique)

$$\begin{aligned} \max_{\Theta_L} |\delta| &= \max_{|a|, |b| \leq \epsilon} \frac{a-b}{2+a+b} \\ &= \max_{|s| \leq 2\epsilon} \frac{1}{s+2} \left[\left(\max_{\mathcal{E}} 2a \right) - s \right] \\ \text{où } \mathcal{E} &= \{a+b=s; a > s/2; |a|, |b| \leq \epsilon\} \\ &= \begin{cases} \max_{|s| \leq 2\epsilon} (2\epsilon - s)/(s+2) & \text{si } s \geq 0 \\ \max_{|s| \leq 2\epsilon} (2\epsilon + s)/(s+2) & \text{si } s \leq 0 \end{cases} = \epsilon \end{aligned}$$

Donc $\max_{\Theta_L} |\delta\delta'| \leq \epsilon^2 \leq 1/8$ pour $\epsilon = 0.35$.

On vérifie la contrainte sur σ^2 : ($j > M(L)$)

$$\sigma^2 = \frac{U_j^L}{16} (1 - \delta^2)(1 - \delta'^2)$$

Sur Θ_L , $U_j^L \geq (1-\epsilon)2^{L+j-N}$, $\max(|\delta|, |\delta'|) = \epsilon$ donc

$$\sigma^2 \geq \frac{(1-\epsilon)^3(1+\epsilon)^2 2^{L+j-N}}{16} \quad (14)$$

Avec $2^{L+j-N} \geq 2^{L+M(L)+1-N} \geq 1.14(x+64 \log n_0)$ on obtient $\sigma^2 \geq 4.5$ sur Θ_L et l'inégalité (b) du second cas particulier du lemme 2 s'applique:

$$|\gamma_j^L - \sigma_j^L \xi_j^L| \leq 3 + 0.41 \left(\xi_j^L \right)^2 \quad \text{pour tout } j > M(L) \text{ et sur } \Theta_L \quad (15)$$

• On multiplie et divise ω par $\sigma_j^L + (2^{L+j-N}/16)^{1/2}$ et on utilise 14 pour minorer σ_j^L :

$$\omega \leq \frac{4}{\tau} 2^{-(L+j-N)/2} |\xi_j^L| |\sigma_j^L|^2 - \left(\frac{2^{L+j-N}}{16} \right) \quad (16)$$

avec $\tau = 1 + (1 + \epsilon)(1 - \epsilon)^{3/2}$. Soit $B = \sigma_j^L{}^2 - (2^{L+j-N}/16)$. On décompose B :

$$B \leq \frac{1}{8} \left\{ |\Delta_{j,k(j)}^{L-1,H}| + |\Delta_{j,k(j)}^{L-1,B}| + |\Delta_{j-1,2k(j)}^L| + |\Delta_{j,2k(j)+1}^L| \right\} + \frac{|U_j^L - 2^{L+j-N}|}{16}$$

On remarque que:

$$\begin{aligned} |U_j^L - 2^{L+j-N}| &= \left| \sum_{j \leq u \leq N-1} 2^{j-u} (U_u^L - \frac{U_{u+1}^L}{2}) \right| \\ &\leq \sum_{j \leq u \leq N-1} 2^{j-u} |\Delta_u^L| \end{aligned}$$

Cette relation et les égalités 12 donnent:

$$B \leq \frac{1}{4} \left\{ |\Delta_{j-1}^L| + |\Delta_j^{L-1}| + \frac{1}{4} \sum_{j \leq u \leq N-1} 2^{j-u} |\Delta_u^L| \right\} \quad (17)$$

Soient des réels $\alpha_1, \alpha_2, \alpha_3$ et $\beta_i = 1/\alpha_i$ pour $i = 1, 2, 3$. Les inégalités 16, 17 et la relation $|ab| \leq (\alpha_i^2 a^2 + \beta_i^2 b^2)/2$ entraînent:

$$\begin{aligned} \omega &\leq \frac{\alpha_1^2 + \alpha_2^2 + \alpha_3^2}{2} (\xi_j^L)^2 \\ &+ \frac{2^{L+j-N}}{2\tau^2} \left[\beta_1^2 (\Delta_j^{L-1})^2 + \beta_2^2 (\Delta_{j-1}^L)^2 + \frac{\beta_3^2}{16} \left(\sum_{j \leq u \leq N-1} 2^{j-u} |\Delta_u^L| \right)^2 \right] \quad (18) \end{aligned}$$

En reportant 15 et 18 dans 13 on a:

$$\begin{aligned} |D_L^k| &\leq \lambda_1 \sum_{j > M(L)} (\xi_j^L)^2 + \lambda_2 \sum_{j > M(L)} 2^{-(L+j-N)} (\Delta_j^{L-1})^2 + \lambda_3' \sum_{j > M(L)} 2^{-(L+j-N)} (\Delta_{j-1}^L)^2 \\ &+ \lambda_4' \sum_{j > M(L)} 2^{-(L+j-N)} \left(\sum_{j \leq u \leq N-1} 2^{(j-u)} |\Delta_u^L| \right)^2 + 3(N - M(L)) \quad (19) \end{aligned}$$

$$\begin{aligned}
\text{avec } \lambda_1 &= 0.41 + (\alpha_1^2 + \alpha_2^2 + \alpha_3^2)/2 \\
\lambda_2 &= 1/(2(1-\epsilon)) + \beta_1^2/(2\tau^2) \\
\lambda'_3 &= 1/(2(1-\epsilon)) + \beta_2^2/(2\tau^2) \\
\lambda'_4 &= \beta_3^2/(32\tau^2)
\end{aligned}$$

On simplifie le terme en λ'_4 :

$$\begin{aligned}
A &= \sum_{j>M(L)} 2^{-(L+j-N)} \left(\sum_{j \leq u \leq N-1} 2^{(j-u)} |\Delta_u^L| \right)^2 \\
&= \sum_{j>M(L)} \left(\sum_{j \leq u \leq N-1} 2^{-(L+u+1-N)/2} 2^{(j-u+1)/2} |\Delta_u^L| \right)^2
\end{aligned}$$

On utilise $|< a | b >| \leq \| a \| \| b \|$:

$$A \leq \sum_{j>M(L)} \left(\sum_{j \leq u \leq N-1} 2^{(j-u+1)/2} \right) \left(\sum_{j \leq u \leq N-1} 2^{-(L+u+1-N)} 2^{(j-u+1)/2} (\Delta_u^L)^2 \right)$$

En majorant $\sum_{j \leq u \leq N-1} 2^{(j-u+1)/2}$ par $\sum_{0 \leq s \leq \infty} 2^{-(s-1)/2}$ et en permutant la somme sur u et la somme sur j , on obtient:

$$A \leq 2 \left(\sum_{0 \leq s \leq \infty} 2^{-(s-1)/2} \right)^2 \sum_{u>M(L)} 2^{-(L+u+1-N)} (\Delta_u^L)^2 \quad (20)$$

Finalement, en réunissant 19 et 20:

$$\begin{aligned}
|D_L^k| &\leq 3(N - M(L)) + \lambda_1 \sum_{j>M(L)} (\xi_j^L)^2 \\
&\quad + \lambda_2 \sum_{j>M(L)} 2^{-(L+j-N)} (\Delta_j^{L-1})^2 + \lambda_3 \sum_{j>M(L)} 2^{-(L+j-N)} (\Delta_{j-1}^L)^2 \quad (21)
\end{aligned}$$

avec $\lambda_3 = 1/(2(1-\epsilon)) + \beta_2^2/(2\tau^2) + \beta_3^2/(8\tau^2(\sqrt{2}-1)^2)$

Contrôle de $P_2^k(l_0, L)$

On avait

$$P_2^k(l_0, L) \leq \mathbf{P} \left(|D_L^k| \geq r(x + 64 \log n) \cap \Theta_L \right)$$

avec r donné en 10

D'après l'inégalité 21, $P_2^k(l_0, L) \leq \pi_1 + \pi'_2 + \pi'_3$ avec:

$$\begin{cases}
\pi_1 = \mathbf{P} \left(N - M(L) + \lambda_1 \sum_{j>M(L)} (\xi_j^L)^2 \geq (r/3)(x + 64 \log n) \right) \\
\pi'_2 = \mathbf{P} \left(N - M(L) + \lambda_2 \sum_{j>M(L)} 2^{-(L+j-N)} (\Delta_j^{L-1})^2 \geq (r/3)(x + 64 \log n) \cap \Theta_L \right) \\
\pi'_3 = \mathbf{P} \left(N - M(L) + \lambda_3 \sum_{j>M(L)} 2^{-(L+j-N)} (\Delta_{j-1}^L)^2 \geq (r/3)(x + 64 \log n) \cap \Theta_L \right)
\end{cases}$$

Lemme 9 *Lemme du chi-2*

Soit $Y(d)$ une variable aléatoire de loi chi-2 à d degrés de libertés . Pour tout réels y on a :

$$P(Y(d) \geq y) \leq \exp\left(-\frac{2y}{5} + \frac{d \log 5}{2}\right)$$

preuve :

$$\begin{aligned} P(Y(d) \geq y) &\leq \inf_{u>0} \exp(-uy) \mathbb{E}(\exp(uY(d))) \\ &\leq \exp\left(-\frac{2y}{5}\right) \left(\mathbb{E} \exp\left(\frac{2Y(1)}{5}\right)\right)^d \\ &\leq \exp\left(-\frac{2y}{5}\right) \left(1 - 2\left(\frac{2}{5}\right)\right)^{-d/2} \end{aligned}$$

Remarque: On utilise ce lemme avec $d = N - M(L)$. Or $M(L) \geq M(N)$ et le choix de l'échelle impose $2^{M(N)} \geq 0.57(x + 64 \log n)$. Donc

$$N - M(L) \leq \frac{\log n}{\log 2} + 1 - \frac{\log(0.57 \times 64 \log n_0)}{\log 2} \leq \frac{\log n}{\log 2}$$

La construction (Section 3 : descente conditionnelle) a imposé les lois conditionnelles :

$$\mathcal{L}(U_{j-1}^L / U_j^L) = B(U_j^L, 1/2) \quad \text{et} \quad \mathcal{L}(U_j^{L-1} / U_j^L) = H(2^L, 2^{L-1}, U_j^L)$$

Par construction dyadique, on peut fabriquer une version de $\{U_{j,k}^L\}_{j>M(L)}^{0 \leq k < 2^{M(L)}}$

notée $\{U_{j,k}^{L,*}\}_{j>M(L)}^{0 \leq k < 2^{M(L)}}$ et une version de $\{U_{j,k}^{L-1,H}, U_{j,k}^{L-1,B}\}_{j>M(L)}^{0 \leq k < 2^{M(L)}}$

notée $\{U_{j,k}^{L-1,H**}, U_{j,k}^{L-1,B**}\}_{j>M(L)}^{0 \leq k < 2^{M(L)}}$ telles que l'on ait les inégalités 1 des lemmes 1 et

2. Finalement, sur Θ_L on a :

1. Sans conditions sur σ^2 :

$$\begin{aligned} \left| U_{j-1}^{L,*} - \frac{U_j^{L,*}}{2} \right| &\leq 1 + \left(\frac{U_j^{L,*}}{4} \right)^{1/2} \left| \xi_j^L \right| \\ &\leq 1 + \sqrt{\frac{1+\epsilon}{4}} 2^{(L+j-N)/2} \left| \xi_j^L \right| \quad \text{pour tout } M(L) < j \leq N \end{aligned}$$

2. Dès que $\sigma^2 \geq 4.5$, on peut utiliser le cas particulier du lemme 2 avec $\delta\delta' = 0$. On vérifie la contrainte sur σ^2 :

$$\begin{aligned}\sigma^2 &= \frac{2^L}{4} \frac{U_j^L}{2^L} \left(\frac{2^L - U_j^L}{2^L} \right) \\ &\geq \frac{(1-\epsilon)2^{L+M(L)+1-N}}{4} \frac{1-\epsilon}{2} \text{ dès que } M(L) < j < N \\ &\geq \frac{(1-\epsilon)^2}{8} \times 1.14 \times 64 \log(n_0) \geq 4.5\end{aligned}$$

On obtient donc:

$$\begin{aligned}\left| U_j^{L-1**} - \frac{U_j^{L**}}{2} \right| &\leq 3 + \left(\frac{2^L}{4} \frac{U_j^{L**}}{2^L} \frac{2^L - U_j^{L**}}{2^L} \right)^{1/2} |\xi_j^L| \\ &\leq 3 + \sqrt{\frac{1+\epsilon}{4}} 2^{(L+j-N)/2} |\xi_j^L| \text{ pour tout } M(L) < j < N\end{aligned}$$

(pour $j = N$, on remarque $U_N^{L-1**} = U_N^{L**} / 2$)

Le choix de l'échelle impose $2^{L+M(L)-N} \geq 0.57(x + 64 \log n)$. On en déduit

$$\sum_{j>M(L)} 2^{-(L+j-N)} \leq \alpha = \frac{1}{0.57(x + 64 \log n)}$$

En utilisant alors $(a+b)^2 \leq 11a^2 + (11/10)b^2$, on obtient $\pi_i' \leq \pi_i$; $i = 2, 3$ avec:

$$\pi_i \leq \mathbf{P} \left(N - M(L) + 11\lambda_i w_i \alpha + \frac{11(1+\epsilon)\lambda_i}{40} \sum_{j>M(L)} (\xi_j^L)^2 \geq (r/3)(x + 64 \log n) \right)$$

avec $w_2 = 9, w_3 = 1$. On choisit les coefficients α_i ; $i = 1, 2, 3$ de manière à avoir:

$$\lambda = 0.785 = \lambda_1 = \frac{11(1+\epsilon)\lambda_2}{40} = \frac{11(1+\epsilon)\lambda_3}{40}$$

On applique le lemme 9 à π_1, π_2, π_3 avec $N - M(L) \leq \log n / \log 2$:

$$\begin{aligned}P_2^k(l_0, L) &\leq \left[1 + \exp\left(\frac{144\alpha}{1+\epsilon}\right) + \exp\left(\frac{16\alpha}{1+\epsilon}\right) \right] \\ &\quad \times \exp\left(-\frac{2r}{15\lambda}x - \log n \left(\frac{2 \times 64r}{15\lambda} - \frac{2}{5\lambda \log 2} - \frac{\log 5}{2 \log 2}\right)\right)\end{aligned}$$

En remplaçant $r, \lambda, \alpha, \epsilon$ par leur valeur numérique, on obtient finalement:

$$P_2^k(l_0, L) \leq 3.433 \exp\left(-\frac{x}{17} - 2 \log n\right) \quad (22)$$

Finalement, en reportant 22 dans 8, on obtient:

$$P \leq 4.6 \cdot 10^{-3} \exp\left(-\frac{x}{32}\right) + 3.433 (N - L_0) 2^{N-L_0} 2^L 2^{N-L-M(L)} \exp\left(-\frac{x}{17} - 2 \log n\right)$$

En utilisant:

$$\begin{cases} N - L_0 \leq \log n / \log 2 \\ 2^{N-L-M(L)} \leq 1 / [0.57(x + 64 \log n)] \\ 2^{-L_0} \leq 1 / [1.14(x + 64 \log n)] \\ 2^{N+L} \leq 4n^2 \end{cases}$$

on obtient:

$$\begin{aligned} P &\leq 4.6 \times 10^{-3} \exp\left(-\frac{x}{32}\right) + 8.5 \times 10^{-4} \exp\left(-\frac{x}{17}\right) \\ &\leq 5.5 \times 10^{-3} \exp\left(-\frac{x}{32}\right) \end{aligned} \quad (23)$$

En réunissant 1, 2 et 23 on a:

$$\begin{aligned} P &\leq 7 \times 10^{-3} \exp(-x/41) && \text{lorsque } x + 64 \log n \leq n/12 \\ P &\leq 0.35 \exp(-x/20) && \text{lorsque } x + 64 \log n > n/12 \end{aligned}$$

Finalement, on a le résultat:

$$P \leq 0.35 \exp\left(-\frac{x}{41}\right)$$

4 Preuve du théorème 2

4.1 Construction des variables

Soit n un entier supérieur ou égal à 3 et N tel que $2^{N-1} < n \leq 2^N$. A partir d'une suite de mouvements browniens indépendants $B_1, \dots, B_n$ on construit un processus de même loi que $(1_{X_1 \leq t}, \dots, 1_{X_n \leq t})$ tel que l'inégalité du théorème 2 soit vérifiée. Mais, contrairement à la section 3, il faut que lorsque n augmente, par exemple $2^{Nd} \leq n \leq 2^{Nd+1}$ où d est un entier, les 2^N premiers mouvements browniens $B_1, \dots, B_{2^N}$ construisent encore $(1_{X_1 \leq t}, \dots, 1_{X_{2^N} \leq t})$.

Pour cela, l'intervalle $]0, 2^N]$ est découpé en $N - 1$ bandes $D_1, \dots, D_{N-1}$ avec:

$$\begin{aligned} D_1 &=]0, 4] \\ D_j &=]2^j, 2^{j+1}] \quad j = 2, \dots, N - 1 \end{aligned}$$

Sur chaque bande D_j ; $j \geq 2$ on construit un processus de même loi que $(1_{X_{2^{j+1}} \leq t}, \dots, 1_{X_{2^{j+1}} \leq t})$ à partir des mouvements browniens $B_{2^{j+1}}, \dots, B_{2^{j+1}}$ par la méthode de la section 3.

(Sur la bande D_1 , on construit un processus de même loi que $(1_{X_1 \leq t}, \dots, 1_{X_4 \leq t})$ à partir de $B_1, \dots, B_4$). Notons $(1_{\tilde{X}_1 \leq t}, \dots, 1_{\tilde{X}_{2^j} \leq t})$ le processus ainsi reconstruit et par abus:

$$\hat{F}_{D_1} = \frac{1}{4} \sum_{i=1}^4 1_{\tilde{X}_i \leq t}$$

$$\hat{F}_{D_j} = \frac{1}{2^j} \sum_{i=2^{j+1}}^{2^{j+1}} 1_{\tilde{X}_i \leq t} \quad \text{pour } j \geq 2$$

Conséquences

1. Le processus $(1_{\tilde{X}_1 \leq t}, \dots, 1_{\tilde{X}_{2^j} \leq t})$ a même loi que le processus $(1_{X_1 \leq t}, \dots, 1_{X_{2^j} \leq t})$ puisque l'égalité en loi est vérifiée sur chacune des bandes d'après la section 3.1 et que les bandes sont indépendantes entre elles.
2. On a les inégalités suivantes:

$$\begin{aligned} \mathbf{P} \left(\max_{2^{j+1} \leq m \leq 2^{j+1}} \sup_{0 \leq t \leq 1} \left| (m - 2^j) \sum_{i=2^{j+1}}^m 1_{\tilde{X}_i \leq t} - (m - 2^j)t - \sum_{i=2^{j+1}}^m B_i^o(t) \right| \right. \\ \left. \geq \log(2^j)(x + 64 \log(2^j)) \right) \leq 0.35 \exp(-x/32) \end{aligned} \quad (24)$$

$$\begin{aligned} \mathbf{P} \left(\max_{1 \leq m \leq 4} \sup_{0 \leq t \leq 1} \left| m \sum_{i=1}^m 1_{\tilde{X}_i \leq t} - mt - \sum_{i=1}^m B_i^o(t) \right| \geq \log(4)(x + 64 \log(4)) \right) \\ \leq 0.35 \exp(-x/32) \end{aligned} \quad (25)$$

$$\mathbf{P} \left(\sup_{0 \leq t \leq 1} \left| 2^j \hat{F}_{D_j}(t) - 2^j t - \sum_{i=2^{j+1}}^{2^{j+1}} B_i^o(t) \right| \geq x + 12 \log(2^j) \right) \leq 2 \exp(-x/6) \quad (26)$$

$$\mathbf{P} \left(\sup_{0 \leq t \leq 1} \left| 4 \hat{F}_{D_1}(t) - 4t - \sum_{i=1}^4 B_i^o(t) \right| \geq x + 12 \log(4) \right) \leq 2 \exp(-x/6) \quad (27)$$

Les relations (24) et (25) sont des conséquences directes du lemme 8, les relations (26) et (27) sont obtenues en remarquant que sur chaque bande la première étape (Section 3-Descente conditionnelle-Etape 0) correspond à la construction hongroise utilisée par Komlos, Major et Tusnady [6] et Bretagnolle et Massart [2] pour démontrer le théorème 1.

Il suffit de contrôler la probabilité:

$$\mathcal{P} = \mathbf{P}\left(\max_{1 \leq m \leq 2^N} \sup_{0 \leq t \leq 1} \left| m \sum_{i=1}^m \mathbf{1}_{\tilde{X}_i \leq t} - mt - \sum_{i=1}^m B_i^o(t) \right| \geq \log n(x + 76 \log n)\right)$$

Remarque: Si l'espace de probabilité est "suffisamment" riche, l'inégalité vérifiée sur le processus reconstruit $(\mathbf{1}_{\tilde{X}_1 \leq t}, \dots, \mathbf{1}_{\tilde{X}_n \leq t})$ l'est aussi sur le processus de départ d'après le lemme de Skohorod [9]. En effet, sur un espace de probabilité assez riche on peut trouver une suite de mouvements browniens indépendants $\bar{B}_1, \dots, \bar{B}_n$ tels que le processus $(\mathbf{1}_{\tilde{X}_1 \leq t}, \dots, \mathbf{1}_{\tilde{X}_n \leq t}, B_1, \dots, B_n)$ ait même loi que le processus $(\mathbf{1}_{X_1 \leq t}, \dots, \mathbf{1}_{X_n \leq t}, \bar{B}_1, \dots, \bar{B}_n)$. (Remarquons que dans l'énoncé du théorème 2, la suite de mouvements browniens $B_1, \dots, B_n$ correspond en fait à la suite $\bar{B}_1, \dots, \bar{B}_n$ définie ci dessus!)

4.2 Contrôle de $\mathcal{P}$

On note:

$$\Phi_m(t) = m \sum_{i=1}^m \mathbf{1}_{X_i \leq t} - mt - \sum_{i=1}^m B_i^o(t)$$

$$\Phi_0(t) = 0$$

On définit $\pi(m)$ par:

$$\begin{aligned} \pi(m) &= 0 & \text{si } 1 \leq m \leq 4 \\ \pi(m) &= 2^j & \text{si } 2^j + 1 \leq m \leq 2^{j+1} \quad j = 2, \dots, N-1 \end{aligned}$$

Lorsque $m \geq 5$, on définit $j(m)$ par:

$$\pi(m) = 2^{j(m)}$$

On note:

$$\begin{aligned} \tilde{\Phi}_1 &= \Phi_4 \\ \tilde{\Phi}_k &= \Phi_{2^{k+1}} - \Phi_{2^k} \quad 2 \leq k \leq N-1 \end{aligned}$$

Alors pour tout $m \geq 5$, on a:

$$|\Phi_m| \leq |\Phi_m - \Phi_{\pi(m)}| + \sum_{k=1}^{j(m)-1} |\tilde{\Phi}_k|$$

Et pour $1 \leq m \leq 4$:

$$|\Phi_m| \leq |\Phi_m - \Phi_{\pi(m)}|$$

On obtient:

$$\mathcal{P} \leq P_1 + P_2 \tag{28}$$

$$\text{où } P_1 = \mathbf{P}\left(\max_{1 \leq m \leq 2^N} \sup_{0 \leq t \leq 1} |\Phi_m(t) - \Phi_{\pi(m)}(t)| \geq \log n(\alpha_1 x + 64 \log n)\right)$$

$$P_2 = \mathbf{P}\left(\max_{5 \leq m \leq 2^N} \sup_{0 \leq t \leq 1} \sum_{k=1}^{j(m)-1} |\tilde{\Phi}_k(t)| \geq \log n(\alpha_2 x + 12 \log n)\right)$$

avec $\alpha_1 = 0.787$ et $\alpha_2 = 0.213$.

Contrôle de P_1

$$P_1 \leq \sum_{j=1}^{N-1} \mathbf{P} \left(\max_{m \in D_j \cap \mathbb{N}} \sup_{0 \leq t \leq 1} |\Phi_m(t) - \Phi_{\pi(m)}(t)| \geq \log n (\alpha_1 x + 64 \log n) \right)$$

Lorsque $j \geq 2$, on écrit:

$$\log n = \log(2^j) + \log(n/2^j)$$

et lorsque $j = 1$:

$$\log n = \log 4 + \log(n/4)$$

Les inégalités (24) et (25) donnent directement:

$$P_1 \leq 0.35 \sum_{j=2}^{N-1} \omega_j \exp\left(-\frac{\alpha_1 x}{32} - 2 \log\left(\frac{n}{2^j}\right)\right)$$

avec $\omega_2 = 2$ et $\omega_j = 1$ si $j \geq 2$.

On obtient:

$$\begin{aligned} P_1 &\leq 0.35 \exp(-x/41) \sum_{j=2}^{N-1} \omega_j \frac{4^j}{n^2} \\ &\leq 0.882 \exp(-x/41) \end{aligned} \tag{29}$$

Contrôle de P_2

Puisque $j(m)$ est croissante de m et que $j(2^N) = N - 1$, on a:

$$P_2 \leq \mathbf{P}((N - 2) \max_{1 \leq k \leq N-2} \sup_{0 \leq t \leq 1} |\tilde{\Phi}_k(t)| \geq \log n (\alpha_2 x + 12 \log n))$$

On utilise $(\log n)/(N - 2) > \log 2$, la décomposition:

$$\begin{aligned} \log n &= \log 2^k + \log(n/2^k) & \text{si } k \geq 2 \\ \log n &= \log 4 + \log(n/4) & \text{si } k = 1 \end{aligned}$$

et les inégalités (26) et (27) pour obtenir:

$$\begin{aligned} P_2 &\leq \sum_{k=1}^{N-2} \mathbf{P} \left(\sup_{0 \leq t \leq 1} |\tilde{\Phi}_k(t)| \geq (\log 2)(\alpha_2 x + 12 \log n) \right) \\ &\leq 2 \sum_{k=2}^{N-2} \omega_k \exp\left(-\frac{\alpha_2 (\log 2) x}{6} - 2(\log 2) \log\left(\frac{n}{2^k}\right)\right) \end{aligned}$$

avec $\omega_2 = 2$ et $\omega_k = 1$ pour $k \neq 2$.

On obtient:

$$\begin{aligned} P_2 &\leq 2 \exp(-x/41) \sum_{k=2}^{N-2} \omega_k \frac{(4^{\log 2})^k}{n^{2 \log 2}} \\ &\leq 2.373 \exp(-x/41) \end{aligned} \tag{30}$$

Les relations (28), (29), (30) donnent le résultat:

$$\mathcal{P} \leq 3.26 \exp(-x/41)$$

Appendice : preuve du Lemme 2 et de ses cas particuliers

Préliminaires

Les notations sont celles de l'introduction .

Remarques sur $H(n, np, np')$

1. $\mathbb{E}(H(n, np, np')) = np p'$
2. $\text{Var}(H(n, np, np')) = (n/n - 1) (np p' q q')$
3. Propriétés de symétrie
 - (a) $H(n, np, np') = H(n, np', np)$
 - (b) $H(n, nq, np') = np' - H(n, np, np')$
 - (c) $H(n, np, nq') = np - H(n, np, np')$
 - (d) $H(n, nq, nq') = nq - np' + H(n, np, np')$

On notera X et σ pour X_n et σ_n .

A Cas σ^2 petit : $\sigma^2 \leq \sigma_0^2$

$$-\min(np p', nq q') = \frac{-\sigma^2}{\max(pp', qq')} \leq X \leq \frac{\sigma^2}{\max(p'q, pq')} = \min(np'q, npq')$$

$$\text{et } \begin{cases} \max(p'q, pq') \geq (p'q + pq')/2 = (1 - \delta\delta')/4 \geq \eta/4 \\ \max(pp', qq') \geq (pp' + qq')/2 = (1 + \delta\delta')/4 \geq \eta/4 \end{cases}$$

donc $|X| \leq M$ avec $M = 4\sigma_0^2/\eta$

$$\text{et } |X - \sigma Y| \leq M + \sigma |Y| \leq \begin{cases} M + Y^2 & \text{si } |Y| \geq \sigma \\ M + \sigma_0^2 & \text{sinon} \end{cases}$$

B Cas σ^2 grand : $\sigma^2 \geq \sigma_0^2$

Dans toute la suite, on note ξ les valeurs de $N - np p'$ prises par X (variable hypergéométrique recentrée). Par conséquent $\max(-np p', -nq q') \leq \xi \leq \min(np'q, npq')$.

B.1 Quelques lemmes techniques et leurs conséquences

Lemme 10 *Encadrement de $\log \mathbf{P}(X = \xi)$.*

Soit $\psi(\xi) = -\log(\sqrt{2\pi}\sigma) - \log \mathbf{P}(X = \xi)$. Pour $0 \leq \xi \leq \lambda\sigma^2$, $0 < \lambda < 1$, on a:

$$\psi(\xi) = \theta(\xi) + \rho(\xi), \quad \text{avec } \theta(\xi) = \xi^2/2\sigma^2 + \xi\delta\delta'/2\sigma^2 - \xi^3\delta\delta'/6\sigma^4$$

$$\text{et } \begin{cases} \rho(\xi) \leq 1/6(1-\lambda)\sigma^2 + -\xi^2 S_2/4(1+\lambda)^2\sigma^4 + \xi^4 S_3/12(1-\lambda)^3\sigma^6 \\ \rho(\xi) \geq -\xi^2 S_2/4(1-\lambda)^2\sigma^4 + \xi^4 S_3/12(1+\lambda)^3\sigma^6 \end{cases}$$

$$\text{où } S_k = (p^k + q^k)(p'^k + q'^k), \quad \delta = p - q \text{ et } \delta' = p' - q'$$

Preuve:

En développant suivant la formule de Stirling, on obtient sur $0 \leq \xi \leq \lambda\sigma^2$

$$P(X = \xi) = CS(\xi, n, p, p')a(\xi)$$

où $a(\xi)$ et $CS(\xi)$ sont définis par:

- $a(\xi) = \frac{1}{\sqrt{2\pi\sigma}} \exp(-\alpha(\xi) - \beta(\xi))$
- $CS(\xi) = \gamma_{np}\gamma_{nq}\gamma_{np'}\gamma_{nq'}/(\gamma_n \prod_{i=1}^4 \gamma_{n\alpha_i + (-1)^i \xi})$

$$\text{avec } \alpha_1 = p'q \quad \alpha_2 = pp' \quad \alpha_3 = pq' \quad \alpha_4 = qq'$$

$$\text{et } \begin{cases} \alpha(\xi) = \sum_{i=1}^4 (n\alpha_i + (-1)^i \xi) \log(1 + (-1)^i \xi/(n\alpha_i)) \\ \beta(\xi) = 1/2 \sum_{i=1}^4 \log(1 + (-1)^i \xi/(n\alpha_i)) \end{cases}$$

En utilisant d'une part la décroissance de γ_a , d'autre part $1 \leq \gamma_a \leq 1 + \frac{1^+}{12a}$ et les huit inégalités du type: $np \leq n$ et $npp' + \xi \leq np'$ (nous n'écrivons que les deux premières), on obtient:

$$\prod_{i=1}^4 \frac{1}{1 + \frac{1^+}{12(n\alpha_i + (-1)^i \xi)}} \leq CS(\xi) \leq 1$$

En utilisant $\log(1+x) \leq x$ et $(n\alpha_i + (-1)^i \xi)^{-1} \leq (n\alpha_i(1-\lambda))^{-1}$ sur $0 \leq \xi \leq \lambda\sigma^2$, on obtient:

$$0 \leq -\log CS \leq \frac{1}{6(1-\lambda)\sigma^2}$$

Par la formule de Taylor, on développe $\alpha(\xi)$ à l'ordre 4 et $\beta(\xi)$ à l'ordre 2, on contrôle les restes et le résultat s'ensuit.

Lemme 11 *Majoration de $r(\xi) = \mathbf{P}(X = \xi + 1)/\mathbf{P}(X = \xi)$.*

1. $r(\xi) \leq \begin{cases} \exp(-\xi/\sigma^2) & \text{si } \delta\delta' \leq 0 \\ \exp(-\xi(1 - \delta\delta')/2\sigma^2) & \text{si } \delta\delta' > 0 \end{cases}$
2. $r(\xi) \leq \exp(-\xi/\sigma^2 + \xi^2\delta\delta'/2\sigma^4) \quad \forall \delta\delta'$

Preuve:

$$r(\xi) = \frac{(np'q - \xi)(npq' - \xi)}{(npp' + \xi + 1)(nqq' + \xi + 1)} \leq \exp(\Delta)$$

avec $\Delta = \log(1 - \xi/np'q) + \log(1 - \xi/npq') - \log(1 + \xi/npp') - \log(1 + \xi/nqq')$

La majoration à l'ordre 2 des logarithmes donne immédiatement 2 et 1 dans le cas $\delta\delta' \leq 0$:

$$\Delta \leq -\frac{\xi}{\sigma^2} + \frac{\delta\delta'\xi^2}{2\sigma^4}$$

Pour 1 dans le cas $\delta\delta' > 0$, on remarque:

$$\log(1 + \frac{\xi}{npp'}) \geq -\log(1 - \frac{\xi}{n \min(p, p')}) \text{ et } \log(1 + \frac{\xi}{nqq'}) \geq -\log(1 - \frac{\xi}{n \min(q, q')})$$

donc $\Delta \leq -\xi(1 - qq' \min(q, q') - pp' \min(p, p'))/\sigma^2$
et $1 - qq' \min(q, q') - pp' \min(p, p') \geq p'q + pq' = (1 - \delta\delta')/2$

Conséquences des lemmes

- Conséquence 1: Pour tout $\sigma^2 \geq 4$ et $0 \leq \xi \leq 1$, on a:

$$\log(\sqrt{2\pi}\sigma\mathbf{P}(X = \xi)) \leq \begin{cases} 0 & \text{si } \delta\delta' > 0 \\ -\delta\delta'/2\sigma^2 & \text{si } \delta\delta' \leq 0 \end{cases}$$

- Conséquence 2: Pour tout $\xi > 1$ on a:

$$\mathbf{P}(X \geq \xi) \leq \frac{\exp(-\omega(\xi - 1)/\sigma^2)}{1 - \exp(-\omega(\xi - 1)/\sigma^2)} \mathbf{P}(X = \xi - 1)$$

$$\text{avec } \omega = \begin{cases} 1 & \text{si } \delta\delta' \in [-1 + \eta, 0] \\ (1 - \delta\delta')/2 & \text{si } \delta\delta' \in]0, 1 - \eta] \end{cases}$$

- Conséquence 3: On note $\phi(y) = \exp(-y^2/2\sigma^2)/\sqrt{2\pi}\sigma$.

1. Si $\delta\delta' \leq 0$ alors pour tout $\xi \geq 0$ on a:

$$\frac{\mathbf{P}(X = \xi + 1)}{\mathbf{P}(X = \xi)} \leq \exp\left(\frac{-2\Lambda + 1}{2\sigma^2}\right) \frac{\phi(\xi + 1 - \Lambda)}{\phi(\xi - \Lambda)}$$

2. Si $\delta\delta' > 0$, on pose

$$\Lambda(\xi) = \Lambda + \frac{\delta\delta'}{2} \left(\frac{\xi + l}{\sigma} \right)^2$$

Alors pour tous $l \geq 0$ et $\xi \geq 0$ on a :

$$\frac{\mathbf{P}(X = \xi + 1)}{\mathbf{P}(X = \xi)} \leq \exp \left(\frac{-2\Lambda + 1}{2\sigma^2} \right) \frac{\phi(\xi + 1 - \Lambda(\xi))}{\phi(\xi - \Lambda(\xi))}$$

Preuves: Pour la conséquence 1, on utilise le lemme 10 avec $\lambda = 1/4$:

$$2\sigma^2 \log(\sqrt{2\pi}\sigma\mathbf{P}(X = \xi)) \leq f(\xi) = -\xi^2 - \delta\delta'\xi + \frac{\delta\delta'\xi^3}{3\sigma^2} + \frac{8\xi^2}{9\sigma^2}$$

$f'(0)$ est du signe de $-\delta\delta'$, $f'(1) \leq 0$ pour $\sigma^2 \geq 4$ et $f''(\xi) \leq 0$ pour $\sigma^2 \geq 4$ et $0 \leq \xi \leq 1$. Si $\delta\delta' > 0$ alors $f(\xi) \leq 0$ pour $0 \leq \xi \leq 1$ et si $\delta\delta' \leq 0$, le maximum de $f(\xi)$ est atteint entre 0 et 1, comme $-\xi^2(1 - 8\xi/9\sigma^2) \leq 0$ on obtient $f(\xi) \leq -\delta\delta'\xi \leq -\delta\delta'$ pour $0 \leq \xi \leq 1$.

La conséquence 2 s'obtient en majorant par une série géométrique (lemme 11 (1)).

Pour la conséquence 3, on utilise

$$\frac{\phi(\xi + 1 - \Lambda)}{\phi(\xi - \Lambda)} = \exp\left(-\frac{\xi}{\sigma^2} + \frac{2\Lambda(\xi) - 1}{2\sigma^2}\right)$$

$$\text{avec } \Lambda(\xi) = \begin{cases} \Lambda & \text{si } \delta\delta' \leq 0 \\ \Lambda + (\delta\delta'/2)((\xi + l)/\sigma)^2 & \text{si } \delta\delta' > 0 \end{cases}$$

et le lemme 11 : avec (1) si $\delta\delta' \leq 0$
avec (2) si $\delta\delta' > 0$

Lemme 12 *Minoration de $r(\xi) = \mathbf{P}(X = \xi + 1)/\mathbf{P}(X = \xi)$.*

Si $|\delta\delta'| \leq 1 - \eta = \theta$ et $\xi + 1 \leq \lambda\sigma^2$ avec $\lambda < 1$ alors

$$\log(r(\xi)) \geq -\lambda - \frac{(1 + \theta)^2}{4}(-\log(1 - \lambda) - \lambda)$$

En effet:

$$r(\xi) = \frac{(np'q - \xi)(npq' - \xi)}{(np'p + \xi + 1)(nqq' + \xi + 1)} \geq \frac{(np'q - (\xi + 1))(npq' - (\xi + 1))}{(np'p + \xi + 1)(nqq' + \xi + 1)}$$

En posant $\xi' = \xi + 1$, $u = \xi'/(np'q)$, $v = \xi'/(npq')$, $w = \xi'/(np'p)$, $s = \xi'/(nqq')$, on obtient:

$$\log(r(\xi)) \geq \log(1 - u)(1 - v) - \log(1 + w)(1 + s)$$

En utilisant $\begin{cases} \text{si } u \leq \tilde{u} < 1 & \log(1-u) \geq -u - \alpha u^2 \text{ avec } \alpha = (-\log(1-\tilde{u}) - \tilde{u})/\tilde{u}^2 \\ \log(1+u) \leq u \end{cases}$ on obtient:

$$\begin{aligned} \log(r(\xi)) &\geq -\frac{\xi'}{\sigma^2} - \alpha(u^2 + v^2) \\ \text{avec } \alpha &= \frac{-\log(1-\lambda) - \lambda}{\lambda^2} \\ \text{et } u^2 + v^2 &= \frac{\xi'^2}{\sigma^4} ((p'q)^2 + (pq')^2) \end{aligned}$$

Or si $|\delta\delta'| \leq 1 - \eta = \theta$ alors $(p'q)^2 + (pq')^2 \leq (1 + \theta)^2/4$ d'où le résultat.

B.2 Inégalité linéaire

Dans toute cette section, on notera ξ_s , $s \in \mathbf{R}^+$ l'entier recentré par npp' tel que $s - 1 < \xi_s \leq s$. On rappelle que $|\delta\delta'| < 1 - \eta$ avec $0 < \eta < 1$.

En utilisant les symétries de la loi hypergéométrique, il suffit de vérifier:

$$X \leq \alpha + \sigma Y + \beta |\delta\delta'| Y^2 \quad \text{pour } X = \Phi_{n,np,np'}^{-1} \circ \Phi(Y) - npp' \text{ et } X \geq 0$$

• On commence par montrer:

$$\mathbf{P}(X \in [\xi, \tilde{\xi}]) \leq \begin{cases} \mathbf{P}(\sigma Y \in [\xi - S, \xi_1]) & \text{si } npp' \notin \mathbf{N} \\ \mathbf{P}(\sigma Y \in [\xi - S, 0]) & \text{si } npp' \in \mathbf{N} \end{cases}$$

$$\mathbf{P}(X \geq \xi) \leq \mathbf{P}(\sigma Y \geq \xi - \Delta(\xi)) \text{ si } \tilde{\xi} + 1 \leq \xi \leq \min(np'q, npq')$$

$$\begin{aligned} \text{Avec pour } \delta\delta' \leq 1/8 \quad S = 3 \quad \text{et } \tilde{\xi} &= \begin{cases} \xi_2 & \text{si } npp' \notin \mathbf{N} \\ 1 & \text{sinon} \end{cases} \\ \text{pour } \delta\delta' > 1/8 \quad S = 4 \quad \text{et } \tilde{\xi} &= \begin{cases} \xi_3 & \text{si } npp' \notin \mathbf{N} \\ 2 & \text{sinon} \end{cases} \end{aligned}$$

$$\text{et } \Delta(\xi) = \begin{cases} 2 & \text{si } \delta\delta' \leq 0 \\ \alpha(\delta\delta') + (\delta\delta'\xi^2)/(2\sigma^2) & \text{si } \delta\delta' > 0 \text{ où } \alpha(\delta\delta') = \begin{cases} 1.835 & \text{si } \delta\delta' \leq 1/8 \\ 2.9 & \text{si } \delta\delta' > 1/8 \end{cases} \end{cases}$$

On pose :

$$\begin{cases} D(\xi) = \log \phi(\xi + 1 - \Delta(\xi + 1)) - \log \mathbf{P}(X = \xi) + \log(1 - \Delta(\xi + 1) + \Delta(\xi)) \\ \phi(y) = \exp(-y^2/2\sigma^2)/\sqrt{2\pi}\sigma \\ \Delta'(\xi) = \Delta(\xi) - 1 \end{cases}$$

Si $0 \leq \xi \leq \tilde{\xi}$ Il suffit de montrer:

$$\mathbf{P}(\xi \leq X \leq \tilde{\xi}) \leq \mathbf{P}(\sigma Y \in [0, \tilde{\xi} - \xi + 2])$$

et d'utiliser ensuite:

$$\begin{aligned} \mathbf{P}(\sigma Y \in [0, \tilde{\xi} - \xi + 2]) &\leq \begin{cases} \mathbf{P}(\sigma Y \in [\xi - (\tilde{\xi} - \xi_1 + 2), \xi_1]) & \text{si } npp' \notin \mathbf{N} \\ \mathbf{P}(\sigma Y \in [\xi - (\tilde{\xi} + 2), 0]) & \text{si } npp' \in \mathbf{N} \end{cases} \\ &= \begin{cases} \mathbf{P}(\sigma Y \in [\xi - S, \xi_1]) & \text{si } npp' \notin \mathbf{N} \\ \mathbf{P}(\sigma Y \in [\xi - S, 0]) & \text{si } npp' \in \mathbf{N} \end{cases} \end{aligned}$$

1. Si $\delta\delta' \leq 1/8$, la conséquence 1 donne:

$$\mathbf{P}(\xi \leq X \leq \tilde{\xi}) \leq \frac{K \exp(1/2\sigma^2)}{\sqrt{2\pi}\sigma} \text{ avec } K = \tilde{\xi} - \xi + 1 \in \{1, 2\}$$

Pour $\sigma^2 \geq \sigma_0^2$, on doit vérifier:

$$K \exp(1/2\sigma_0^2) \leq \int_0^{K+1} \exp(-t^2/2\sigma_0^2) dt$$

$$\text{soit } \frac{K \exp(1/2\sigma_0^2)}{\sqrt{2\pi}\sigma_0} + \frac{1}{2} \leq \Phi\left(\frac{K+1}{\sigma_0}\right)$$

En lisant les valeurs sur une table de loi normale, cette inégalité fonctionne dès que $\sigma_0^2 \geq 4.5$.

2. Si $\delta\delta' > 1/8$, la conséquence 1 donne:

$$\mathbf{P}(\xi \leq X \leq \tilde{\xi}) \leq \frac{K}{\sqrt{2\pi}\sigma} \text{ avec } K = \tilde{\xi} - \xi + 1 \in \{1, 2, 3\}$$

et l'inégalité à vérifier s'écrit pour $\sigma^2 \geq \sigma_0^2$:

$$\frac{K}{\sqrt{2\pi}\sigma_0} + \frac{1}{2} \leq \Phi\left(\frac{K+1}{\sigma_0}\right)$$

qui fonctionne dès que $\sigma_0^2 \geq 9$.

Si $\tilde{\xi} + 1 \leq \xi \leq \xi_{\sigma^2/2}$, on montre:

$$\mathbf{P}(X = \xi) \leq \mathbf{P}(\sigma Y \in [\xi - \Delta(\xi); \xi + 1 - \Delta(\xi + 1)])$$

Pour $\xi \leq \xi_{\sigma^2/2}$, la fonction $\xi - \Delta(\xi)$ est croissante. De plus, l'inégalité

$$\tilde{\xi} + 1 - \Delta(\tilde{\xi} + 1) \geq \begin{cases} \xi_1 & \text{si } npp' \notin \mathbf{N} \\ 0 & \text{sinon} \end{cases}$$

est vraie sans conditions si $\delta\delta' \leq 0$, avec la condition $\sigma^2 \geq 3.5$ si $\delta\delta' \in]0, 1/8]$ et avec la condition $\sigma^2 \geq 80$ si $\delta\delta' > 1/8$. Le raccordement avec les $\xi \in [0, \tilde{\xi}]$ est possible et nous sommes ramenés à vérifier $D(\xi) \geq 0$. On utilise la majoration de $P(X = \xi)$ du lemme 10 en remarquant $S_2 \leq (1 + (\delta\delta')^2)/2$.

1. Si $-1 + \eta \leq \delta\delta' \leq 0$, on obtient

$$\begin{aligned} D(\xi) &\geq -\frac{\xi^2}{2\sigma^2} + \frac{\xi}{\sigma^2} - \frac{1}{2\sigma^2} + \frac{\xi^2}{2\sigma^2} + \frac{\delta\delta'\xi}{2\sigma^2} - \frac{S_2\xi^2}{\sigma^4} \\ &\geq \frac{1+\eta}{2} \frac{\xi}{\sigma^2} + (-1 + \frac{\eta}{2}) \left(\frac{\xi}{\sigma^2}\right)^2 - \frac{1}{2\sigma^2} \end{aligned}$$

Cette fonction est concave de ξ/σ^2 , elle est positive en $\xi/\sigma^2 = 2/\sigma^2$ et en $\xi/\sigma^2 = 1/2$ dès que $\sigma^2 \geq 4/(4\eta - \eta^2)$ (en particulier, pour $|\delta\delta'| \leq 1/8$, $\sigma^2 \geq 4.5$ convient).

2. Si $0 < \delta\delta' \leq 1 - \eta$, on obtient

$$\begin{aligned} D(\xi) &\geq \xi\Delta'(\xi+1)/\sigma^2 - (\Delta'(\xi+1))^2/2\sigma^2 + \delta\delta'\xi/2\sigma^2 - \delta\delta'\xi^3/6\sigma^4 \\ &\quad - S_2\xi^2/\sigma^4 + \log(1 - (\delta\delta'(2\xi+1)/2\sigma^2)) \end{aligned}$$

On utilise

- (a) $\xi\Delta'(\xi+1)/\sigma^2 - (\Delta'(\xi+1))^2/2\sigma^2 \geq \xi\Delta'(\xi+1)/2\sigma^2$
car $\xi - \Delta(\xi) \geq 0$ pour $\tilde{\xi} + 1 \leq \xi \leq \xi_{\sigma^2/2} + 1$ et $\sigma^2 \geq 4.5$
- (b) $\log(1 - (\delta\delta'/2\sigma^2)(2\xi+1)) \geq \begin{cases} \log(1 - (5/2)(\delta\delta'\xi/2\sigma^2)) & \text{si } \delta\delta' \leq 1/8 \\ \log(1 - (7/3)(\delta\delta'\xi/2\sigma^2)) & \text{si } \delta\delta' > 1/8 \end{cases}$
- (c) Si $\delta\delta'\xi/2\sigma^2 \leq r$ alors $\delta\delta'\xi/2\sigma^2 + \log(1 - \lambda(\delta\delta'\xi/2\sigma^2)) \geq -\tau(r)\delta\delta'\xi/2\sigma^2$
avec $-\tau(r) = 1 + (1/r)\log(1 - \lambda r)$

On obtient pour $\delta\delta'\xi/2\sigma^2 \leq r$:

$$\begin{aligned} D(\xi) &\geq \frac{\xi}{2\sigma^2} \left\{ \alpha(\delta\delta') - 1 + \frac{\delta\delta'}{2} \left(\frac{\xi+1}{\sigma}\right)^2 - \frac{\delta\delta'\xi^2}{3\sigma^2} - (1 + \delta\delta')\frac{\xi}{\sigma^2} - \tau(r)\delta\delta' \right\} \\ &\geq \frac{\xi}{2\sigma^2} \left\{ \frac{\delta\delta'\xi^2}{6\sigma^2} + \alpha(\delta\delta') - 1 - \frac{\xi}{\sigma^2} - \tau(r)\delta\delta' \right\} \end{aligned}$$

- Si $\delta\delta' \leq 1/8$ alors $r = 1/32$ entraîne $\tau(r)\delta\delta' \leq 0.21$, comme $\xi/\sigma^2 \leq 1/2$ on obtient $D(\xi) \geq 0$.

- Si $\delta\delta' > 1/8$ on sépare deux cas:

Soit $\xi^2/6\sigma^2 - 0.45\tau(1/4) \geq 0$, alors avec $\xi/\sigma^2 \leq 1/2$ on obtient $\alpha(\delta\delta') = 2.9 \geq 1 + \xi/\sigma^2 + 0.55\tau(1/4)$.

Soit $\xi^2/6\sigma^2 - 0.45\tau(1/4) < 0$ ce qui entraîne $\xi/\sigma^2 \leq 0.214$ et $\alpha(\delta\delta') = 2.9 \geq 1 + 0.214 + \tau(0.107)$ pour $\sigma^2 \geq 150$.

Si $\xi_{\sigma^2/2} + 1 \leq \xi \leq \min(np'q, npq')$

Lorsque $\delta\delta' \leq 1/8$, la fonction $\xi - \Delta(\xi)$ est croissante jusqu'en $\min(np'q, npq')$ de manière évidente si $\delta\delta' \leq 0$ et en utilisant $\xi \leq 4\sigma^2/\eta$ si $\delta\delta' \in]0, 1/8]$. On montre alors:

$$\mathbf{P}(X = \xi) \leq \mathbf{P}(\sigma Y \in [\xi - \Delta(\xi); \xi + 1 - \Delta(\xi + 1)])$$

Dans le cas $\delta\delta' > 1/8$, on montre directement:

$$\mathbf{P}(X \geq \xi) \leq \mathbf{P}(\sigma Y \geq \xi - \Delta(\xi))$$

1. Si $-1 + \eta \leq \delta\delta' \leq 1/8$ on a:

$$R = \frac{\mathbf{P}(\sigma Y \in [\xi - \Delta(\xi); \xi + 1 - \Delta(\xi + 1)])}{\mathbf{P}(X = \xi)} \geq (1 - \Delta(\xi + 1) + \Delta(\xi)) \frac{\phi(\xi - \Delta(\xi + 1))}{\mathbf{P}(X = \xi)}$$

La conséquence 3 avec $\Lambda = 1$ pour $\delta\delta' \leq 0$ et $\Lambda = \alpha(\delta\delta') - 1 = 0.835$ pour $\delta\delta' > 0$ donne:

$$R \geq (1 - \Delta(\xi + 1) + \Delta(\xi)) \exp\left(\frac{2\Lambda - 1}{2\sigma^2}(\xi - \xi_{\sigma^2/2})\right) \frac{\phi(\xi_{\sigma^2/2} + 1 - \Delta(\xi_{\sigma^2/2} + 1))}{\mathbf{P}(X = \xi_{\sigma^2/2})}$$

(a) Si $\delta\delta' \leq 0$, on obtient $R \geq \exp D(\xi_{\sigma^2/2}) \geq 1$ pour $\sigma^2 \geq 4.5$ d'après le cas précédent.

(b) Si $\delta\delta' > 0$, on obtient en posant $K = 2\Lambda - 1 = 0.67$:

$$R \geq g(\xi) \frac{\phi(\xi_{\sigma^2/2} + 1 - \Delta(\xi_{\sigma^2/2} + 1))}{\mathbf{P}(X = \xi_{\sigma^2/2})}$$

$$\text{où } g(\xi) = \left(1 - \frac{\delta\delta'}{2\sigma^2}(2\xi + 1)\right) \exp\left(\frac{K\xi}{2\sigma^2} - \frac{K}{4}\right)$$

Comme $g(\xi)$ est une fonction croissante de ξ pour $\delta\delta' \leq 1/8$ et $\sigma^2 \geq 4.5$, on obtient $R \geq \exp D(\xi_{\sigma^2/2}) \geq 1$ d'après le cas précédent.

2. Si $1/8 < \delta\delta' \leq 1 - \eta$, on note

$$R = \frac{\mathbf{P}(\sigma Y \geq \xi - \Delta(\xi))}{\mathbf{P}(X \geq \xi)}$$

La fonction $\xi - \Delta(\xi)$ est concave de ξ (linéaire si $\delta\delta' \leq 0$) et $\xi_{\sigma^2/2} - \Delta(\xi_{\sigma^2/2}) \geq 3\sigma$ pour σ^2 assez grand ($\sigma^2 \geq 50$ convient). Soit ξ_d le plus grand entier recentré par npq' tel que $\xi - \Delta(\xi) \geq 3\sigma$.

En supposant qu'en $\xi = \xi_d$ l'on ait $R \geq 1$, alors pour tout $\xi > \xi_d$:

$$\begin{aligned} \mathbf{P}(X \geq \xi) &\leq \mathbf{P}(X \geq \xi_d) \leq \mathbf{P}(\sigma Y \geq \xi_d - \Delta(\xi_d)) \\ &\leq \mathbf{P}(\sigma Y \geq 3\sigma) \\ &\leq \mathbf{P}(\sigma Y \geq \xi - \Delta(\xi)) \end{aligned}$$

On peut donc se restreindre aux ξ tels que $\xi_{\sigma^2/2} + 1 \leq \xi \leq \xi_d$.

En utilisant l'inégalité gaussienne

$$P(Y \geq y) \geq \frac{y^2 - 1}{y^3} \frac{\exp(-y^2/2)}{\sqrt{2\pi}}$$

avec $y = (\xi - \Delta(\xi))/\sigma$, ce qui entraîne $y^2 - 1 \geq 8y^2/9$, on a :

$$P(\sigma Y \geq \xi - \Delta(\xi)) \geq \frac{8\sigma^2}{9(\xi - 1)} \phi(\xi - \Delta(\xi)) \quad (31)$$

La conséquence 2 avec $\omega = (1 - \delta\delta')/2$ et 31 donnent:

$$R \geq Q\left(\frac{\xi - 1}{\sigma^2}\right) \frac{\phi(\xi - 1 - \Delta'(\xi))}{P(X = \xi - 1)} \quad (32)$$

$$\text{où } Q\left(\frac{\xi - 1}{\sigma^2}\right) = \frac{8\sigma^2}{9(\xi - 1)} (\exp(\omega(\xi - 1)/\sigma^2) - 1)$$

La conséquence 3(2) avec $\Lambda = \alpha(\delta\delta') - 1 = 1.9$ et les relations $0 \leq 3\sigma - 1 \leq \xi - 2 - \Delta'(\xi) \leq \xi - 2 - \Delta'(\xi - 1)$ pour $\xi_{\sigma^2/2} \leq \xi \leq \xi_d$ permettent d'obtenir les inégalités suivantes:

$$\frac{\phi(\xi - 1 - \Delta'(\xi))}{P(X = \xi - 1)} \geq \frac{\phi(\xi - 2 - \Delta'(\xi))}{P(X = \xi - 2)} \geq \frac{\phi(\xi - 2 - \Delta'(\xi - 1))}{P(X = \xi - 2)} \quad (33)$$

En descendant de proche en proche avec les inégalités 33, l'inégalité 32 devient:

$$R \geq Q\left(\frac{\xi - 1}{\sigma^2}\right) \frac{\phi(\xi_{\sigma^2/2} + 1 - \Delta(\xi_{\sigma^2/2} + 1))}{P(X = \xi_{\sigma^2/2})}$$

Q est une fonction croissante de $(\xi - 1)/\sigma^2 \geq (1/2) - (1/\sigma^2) = 1/2^+$. La majoration de $P(X = \xi)$ du lemme 10 avec $S_2 \leq (1 + \delta\delta')/2$ et la relation 2a qui reste valide pour $\xi \leq \xi_d$ donnent:

$$\begin{aligned} R &\geq \frac{16^+}{9} (\exp(\omega/2^+) - 1) \exp \left\{ \frac{\xi_{\sigma^2/2}}{2\sigma^2} \left(\frac{\delta\delta' \xi_{\sigma^2/2}^2}{6\sigma^2} + \delta\delta' + \alpha(\delta\delta') - 1 - \frac{\xi_{\sigma^2/2}}{\sigma^2} \right) \right\} \\ &\geq \frac{16^+}{9} \left(\exp\left(\frac{1 - \delta\delta'}{4^+}\right) - 1 \right) \exp \left\{ \delta\delta' \left(\frac{\sigma^2}{96^+} + \frac{1}{4^+} \right) + \frac{7}{20^+} \right\} \end{aligned}$$

avec $16^+ = 8 \times 2^+$, $4^+ = 2 \times 2^+$, $96^+ = 12 \times 2^{+3}$, et $20^+ = 5 \times 2^+$.

Cette fonction est croissante de $\delta\delta'$ pour $\sigma^2 \geq 96^+/(1 - \delta\delta')$ et est supérieure à 1 en $\delta\delta' = 1/8$ dès que $\sigma^2 \geq 1700/(1 - \delta\delta')$.

• Appliquons la méthode de transformation par quantile décrite dans le livre de Csörgö et Revesz [3] pages 133-134 à $X = \Phi_{n, np, np'}^{-1} \circ \Phi(Y) - np p'$ et $X \geq 0$. On obtient:

1. Si $\delta\delta' \in [-1 + \eta, 0]$ et $\sigma^2 \geq 4.5$:

$$\begin{aligned} X &\leq \sigma Y + \begin{cases} 2 & \text{si } X < 2 \\ 3 & \text{sinon} \end{cases} \\ &\leq \sigma Y + 3 \end{aligned}$$

2. Si $\delta\delta' \in]0, 1/8]$ et $\sigma^2 \geq 4.5$:

$$\begin{aligned} X &\leq \sigma Y + 3 \quad \text{si } X < 2 \\ X &\leq \sigma Y + 1.835 + \frac{\delta\delta'}{2} \left(\frac{X}{\sigma}\right)^2 \quad \text{si } X \geq 2 \end{aligned} \quad (34)$$

3. Si $\delta\delta' \in]1/8, 1 - \eta]$ et $\sigma^2 \geq 1700/\eta$:

$$\begin{aligned} X &\leq \sigma Y + 4 \quad \text{si } X < 3 \\ X &\leq \sigma Y + 2.9 + \frac{\delta\delta'}{2} \left(\frac{X}{\sigma}\right)^2 \quad \text{si } X \geq 3 \end{aligned} \quad (35)$$

On transforme 34 et 35 en inégalités quadratiques:

– Pour 34, en utilisant $X/\sigma^2 \leq 4/(1 - \delta\delta')$ on a:

$$\frac{5X}{7} \leq X(1 - \frac{2\delta\delta'}{1 - \delta\delta'}) \leq X(1 - \frac{\delta\delta'X}{2\sigma^2}) \leq \sigma Y + 1.835$$

En reportant dans 34 et en utilisant $(a + b)^2 \leq 2(a^2 + b^2)$ on obtient pour $\sigma^2 \geq 4.5$:

$$X \leq \sigma Y + 3 + \delta\delta' \left(\frac{7}{5}\right)^2 Y^2 \quad (36)$$

– Pour 35, si $X \leq \xi_{\sigma^2/2}$, on a:

$$\frac{3X}{4} \leq X(1 - \frac{\delta\delta'}{4}) \leq \sigma Y + 2.9$$

et on obtient, pour $\sigma^2 \geq 1700/\eta$:

$$X \leq \sigma Y + 4 + \delta\delta' \left(\frac{4}{3}\right)^2 Y^2$$

Si $X \geq \xi_{\sigma^2/2}$ l'inégalité 35 en $\sigma^2/2$ entraîne $Y \geq (3\sigma/8) - (2.9/\sigma)$. Comme Y est une fonction croissante de X et $X/\sigma \leq 4\sigma/\eta \leq 4Y/(0.37\eta)$, on obtient :

$$X \leq \sigma Y + 4 + \delta\delta' \frac{59}{\eta^2} Y^2 \quad (37)$$

B.3 Inégalité quadratique

On rappelle que dans toute la suite : $X = \Phi_{(n,np,np')}^{-1} \circ \Phi(Y) - np'$

En utilisant les propriétés de symétrie de $H(n,np,np')$, il suffit de montrer qu'il existe des constantes c et d telles que:

$$|X - \sigma Y| \leq c + dY^2 \quad \text{sur} \quad X \geq 0$$

1. D'après l'inégalité linéaire, $0 \leq X \leq c + \sigma Y + dY^2$ est immédiate

- Si $|\delta\delta'| \leq 1 - \eta$ et $\sigma^2 \geq 1700$, d'après 37 $c = 4$ et $d = |\delta\delta'|59/(\eta)^2$
- Si $|\delta\delta'| \leq 1/8$ et $\sigma^2 \geq 4.5$, d'après 36 $c = 3$ et $d = |\delta\delta'|(7/5)^2$

2. Il reste à montrer qu'il existe des constantes c et d telles que:

$$-c - dY^2 + \sigma Y \leq X \quad \text{sur} \quad X \geq 0 \quad (38)$$

Preuve de (38) :

- Sur $X > (\sigma^2/(4d)) - c$, l'inégalité 38 est immédiate en remarquant:

$$-c - dY^2 + \sigma Y \leq \frac{\sigma^2}{4d} - c \leq X$$

En particulier si $\sigma^2 < 4cd$ l'inégalité 38 est toujours vérifiée.

- On se place donc sur $\sigma^2 \geq 4cd$ et $0 \leq X \leq (\sigma^2/(4d)) - c$.

Il suffit alors, d'après les techniques de plongement (c'est à dire la transformation par quantile décrite dans le livre de Csörgö et Revesz [3] pages 133-134) de montrer que pour tout $\xi \in [0, (\sigma^2/(4d)) - c]$:

$$\begin{aligned} P(X \geq \xi) &\geq P\left(\sigma Y \geq \frac{\sigma^2}{2d}(1 - \sqrt{1-x})\right) \\ \text{où } x &= \frac{4d(\xi + c - 1)}{\sigma^2} \end{aligned} \quad (39)$$

Suivant les différentes valeurs de ξ , nous montrerons directement l'inégalité 39 ou bien nous montrerons tout d'abord l'inégalité point par point suivante :

$$P(X = \xi) \geq P_G(\xi) \quad (40)$$

$$\text{où} \quad P_G(\xi) = P\left(\sigma Y \in \left[\frac{\sigma^2}{2d}(1 - \sqrt{1-x}); \frac{\sigma^2}{2d}(1 - \sqrt{1-x'})\right]\right)$$

$$\text{avec} \quad \begin{cases} x = 4d(\xi + c')/\sigma^2 & \text{où } c' = c - 1 \\ x' = 4d(\xi + c)/\sigma^2 \end{cases}$$

1. Majoration de $\mathbf{P}_G(\xi)$ sur un intervalle $[x_0, x_1]$

Sur $x_0 \leq x \leq x' \leq x_1$ avec $\begin{cases} x_0 = (1/2) + (4d^2)/(\sigma^2) - (1/2)\sqrt{1 - (16d^2)/(\sigma^2)} \\ x_1 = (1/2) + (4d^2)/(\sigma^2) + (1/2)\sqrt{1 - (16d^2)/(\sigma^2)} \end{cases}$
on a :

$$\begin{aligned} \log \mathbf{P}_G(\xi) &\leq f_1(x) \\ \text{avec } f_1(x) &= -\log(\sqrt{2\pi}\sigma) - \frac{1}{2}\log(1-x) - \frac{\sigma^2}{8d^2}(1-\sqrt{1-x})^2 \end{aligned} \quad (41)$$

En effet par le changement de variable $u = \sigma(1 - \sqrt{1-t})/(2d)$ on obtient

$$\begin{aligned} \mathbf{P}_G(\xi) &= \int_x^{x'} \frac{\sigma}{2d\sqrt{2\pi}} g(u) du \\ \text{où } g(u) &= \frac{1}{\sqrt{1-u}} \exp\left(-\frac{\sigma^2}{8(d)^2}(1-\sqrt{1-u})^2\right) \end{aligned}$$

Or dès que $\sigma^2 \geq 16d^2$ g est décroissante sur $x_0 \leq t \leq x_1$ d' où le résultat.

2. Comparaison d'une majoration du membre de gauche de 39 et de la majoration 41

On utilise l'inégalité $\mathbf{P}(Y \geq y) \leq (\exp(-y^2/2))/(y\sqrt{2\pi})$ et on obtient:

$$\begin{aligned} \mathbf{P}\left(Y \geq \frac{\sigma}{2d}(1-\sqrt{1-x})\right) &\leq f_2(x) \\ \text{où } f_2(x) &= -\log(\sigma\sqrt{2\pi}) - \log\left(\frac{1-\sqrt{1-x}}{2d}\right) - \frac{\sigma^2}{8d^2}(1-\sqrt{1-x})^2 \end{aligned}$$

La seule modification par rapport à 41 est le $-(1/2)\log(1-x)$ qui devient $-\log((1-\sqrt{1-x})/(2d))$.

Donc

$$f_2(x) \leq f_1(x) \Leftrightarrow x \geq \tilde{x} \text{ avec } \tilde{x} = \frac{4d(1+d)}{(2d+1)^2} \quad (42)$$

3. Etude de l'inégalité 40

Pour avoir la majoration de $\mathbf{P}_G(\xi)$ dès le début, il faut que $4d(c-1)/\sigma^2 \geq x_0$, soit la condition notée C_1 : $c > (d+1)^2$

On suppose donc C_1 vérifiée.

En utilisant la minoration de $\mathbf{P}(X = \xi)$ où $\xi \leq \lambda\sigma^2$ (lemme 10, ici $\lambda \leq 1/(4d)$), la majoration de $\mathbf{P}_G(\xi)$ (41) et en supposant $|\delta\delta'| \leq \theta$, il suffit de vérifier la positivité de $A + B + C$ avec:

$$\begin{aligned} A &= \frac{\sigma^2(1-\sqrt{1-x})^2}{8(d)^2} \\ B &= \frac{1}{2}\log(1-x) \end{aligned}$$

$$C = -\frac{\xi\theta}{2\sigma^2} - \frac{\xi^2}{2\sigma^2} - \frac{\xi^3\theta}{6\sigma^4} + \frac{\xi^2 S_2}{4\sigma^4(1+\lambda)^2} - \frac{\xi^4 S_3}{(12\sigma^6(1-\lambda)^3)} + \frac{1}{(6(1-\lambda)\sigma^2)}$$

En utilisant les inégalités suivantes:

$$\begin{aligned} 2 - x - 2\sqrt{1-x} &\geq \frac{x^2}{4} + \frac{x^3}{8} + \frac{5x^4}{64} \\ \text{si } x \leq \beta < 1 & B \geq -\frac{x}{2} - \frac{x^2}{4} - Tx^3 \quad \text{où} \quad T \leq \frac{-\log(1-\beta) - \beta - \beta^2/2}{2\beta^3} \quad (43) \\ \frac{1}{6(1-\lambda)\sigma^2} &\leq \frac{1}{6\sigma^2} - \frac{\xi}{6\sigma^4(1-\lambda)} \\ S_2 &\geq 1/4 S_3 \leq (1+3\theta^2)/4 \end{aligned}$$

Sachant que $x = 4d\eta/\sigma^2$ où $\eta = \xi + c'$, en regroupant les termes en $1/\sigma^2$, en $1/\sigma^4$, en $1/\sigma^6$, on obtient $A + B + C \geq \tau_1 + \tau_2 + \tau_3$ avec

$$\begin{aligned} \tau_1 &= \frac{1}{2\sigma^2} (\eta^2 - \xi^2 - \xi\theta - \frac{1}{3} \underline{4d\eta}) \\ \tau_2 &= \frac{1}{\sigma^4} \left(d\eta^3 - \underline{4d^2\eta^2} + \frac{\xi^2}{16(1+\lambda^2)} - \frac{\xi^3\theta}{6} - \frac{\xi}{6(1-\lambda)} \right) \\ \tau_3 &= \frac{1}{\sigma^6} \left(\frac{5d^2\eta^4}{2} - \underline{T(4d\eta)^3} - \frac{\xi^4(1+3\theta^2)}{48(1-\lambda)^3} \right) \end{aligned}$$

les termes soulignés proviennent de $(1/2)\log(1-x)$

On impose que dans τ_3 le terme en ξ^4 soit positif, ce qui impose la condition suivante sur d :

$$\frac{5d^2}{2} \geq \frac{(1+3\theta^2) \times d^3}{48(d-1/4)^3} \quad (44)$$

- Dans le cas général $\theta = 1$ donc 44 donne $d = 0.51$.
- Si $\theta = 1/8$ on obtient avec 44 $d = 0.41$.

On traitera en détail le cas $\theta = 1/8$ puisqu'il est utilisé dans la preuve du théorème 2 (le cas général se fait de la même manière). En développant τ_1 et τ_2 avec $\theta = 1/8$ et $\lambda = 1/(4d)$, on obtient:

$$\begin{aligned} \sigma^2 \tau_1 &= \frac{1}{2} \left(\left(2c' - \underline{4d} - \frac{1}{8} \right) \xi + c'^2 \underline{4dc'} - \frac{1}{3} \right) \\ \sigma^4 \tau_2 &= \xi^3 \left(d - \frac{1}{48} \right) + \xi^2 \left(3dc' - \underline{4d^2} + \frac{d^2}{(4d+1)^2} \right) \\ &\quad + \xi \left(3dc'^2 - \underline{8d^2c'} - \frac{4d}{6(4d-1)} \right) + dc'^3 - \underline{4d^2c'^2} \end{aligned}$$

on rappelle que les termes soulignés proviennent de $(1/2)\log(1-x)$
d étant égal à 0.4, on choisit $c' = 2$ de telle sorte que r_1 et r_2 soient totalement positifs (la condition C_1 est alors vérifiée). En utilisant $\sigma^2 \geq 4cd$, on obtient:

$$\sigma^6(r_1 + r_2 + r_3) \geq \frac{5d^2}{2}(\eta^4 - \xi^4) - (4d)^3 T \eta^3 + (4cd)\sigma^4 r_2 + (4cd)^2 \sigma^2 r_1$$

En développant cette équation et en regroupant les termes suivant leur degré en ξ , on obtient:

$$\frac{\sigma^6}{d^2}(r_1 + r_2 + r_3) \geq \xi^3(31.39 - 26.4T) + \xi^2(83 - 157.44T) + \xi(293.7 - 315T) + (85.12 - 209.92T)$$

En imposant que cette équation soit totalement positive, on obtient $T \leq 0.4053$ où T est défini en 43 ce qui impose $\beta = 0.73$. En résumé :

$$\begin{aligned} \text{si } \theta = 1/8 \quad d = 0.41 \quad c = 3 \quad x \leq 0.73 \text{ alors} \\ \log P(X = \xi) \geq f_1(x) \end{aligned} \quad (45)$$

4. Fin de la démonstration dans le cas $\theta = 1/8$

On rappelle $d = 0.41$ et $c = 3$ et donc $\sigma^2 \geq 4.92$.

- (a) Si $0.6982 = \tilde{x}_1 \leq x \leq \beta = 0.73$, d'après la remarque 42 (ici $\tilde{x} \leq 0.6982$) et le résumé (45) l'inégalité 39 est vérifiée, en effet:

$$P(X = \xi) \geq f_1(x) \geq f_2(x) \geq P\left(Y \geq \frac{\sigma}{2d}(1 - \sqrt{1-x})\right)$$

- (b) Si $x < \tilde{x}_1 = 0.6982$ et sous la condition supplémentaire de décroissance de la fonction g (c'est à dire : $x' \leq (1/2) + (4d^2/\sigma^2) + (1/2)\sqrt{1 - (16d^2/\sigma^2)}$ noté inégalité a), on obtient l'inégalité 40 point par point.

- Si $\sigma^2 \geq 5.8$ alors l'inégalité a est vérifiée pour tout $x < \tilde{x}_1$ d'où 40
- Si $4.92 \leq \sigma^2 < 5.8$ l'inégalité a est vérifiée si $x < 0.639$ d'où 40
- Si $4.92 \leq \sigma^2 < 5.8$ et $0.639 \leq x < \tilde{x}_1$

En utilisant la minoration du rapport $r(\xi) = P_{\xi+1}/P_\xi$ (lemme 10) avec $\lambda = 1/(4d)$ soit $\log(P_{\xi+1}/P_\xi) \geq -0.715$, on obtient sachant que $\xi \leq \sigma^2/4d - 3$:

$$\log(P(X \geq \xi)) \geq \log(P(X = \xi)) + 0.612 \quad (46)$$

En utilisant de plus la majoration de $P\left(Y \geq \frac{\sigma}{2d}(1 - \sqrt{1-x})\right)$ et en minorant $\log((1 - \sqrt{1-x})/2d)$ par $\log((1 - \sqrt{1-0.639})/2d)$, on obtient:

$$\begin{aligned} \log(P(X \geq \xi)) - \log P\left(Y \geq \frac{\sigma}{2d}(1 - \sqrt{1-x})\right) &\geq \\ r_1 + r_2 + r_3 - \frac{1}{2}\log(1-x) + \log\left(\frac{1 - \sqrt{1-x}}{2d}\right) + 0.612 &\geq \end{aligned}$$

$$\begin{aligned}\frac{1}{2\sigma^2}((4 - 1/8)\xi + 4 - 1/3) + \log\left(\frac{1 - \sqrt{1 - 0.639}}{2d}\right) + 0.612 &\geq \\ \frac{1}{2 \times 5.8}(4 - 1/3) + \log\left(\frac{1 - \sqrt{1 - 0.639}}{2d}\right) + 0.612 &\geq 0\end{aligned}$$

(c) Si $x > \beta = 0.73$, de la même manière

$$\log(\mathbf{P}(X \geq \xi)) - \log \mathbf{P}\left(Y \geq \frac{\sigma}{2d}(1 - \sqrt{1 - x})\right) \geq \log\left(\frac{1 - \sqrt{1 - 0.73}}{2d}\right) + 0.612 \geq 0$$

5. Dans le cas général où $\theta = 1$ par le même procédé on obtient $d = 0.51$ et $c = 4$.

En résumé

- Dans le cas général, si $|\delta\delta'| \leq 1 - \eta$ et $\sigma^2 \geq 1700$,

$$|X - \sigma Y| \leq 4 + \min(0.51, \delta\delta' \frac{59}{\eta^2}) Y^2$$

- Si $|\delta\delta'| \leq 1/8$ et $\sigma^2 \geq 4.5$,

$$|X - \sigma Y| \leq 3 + 0.41 Y^2$$

B.4 L'inégalité linéaire n'est pas améliorable dans le cas $|\delta\delta'| \geq \rho > 0$

En utilisant les propriétés de symétrie de $H(n, np, np')$, il suffit de montrer que l'inégalité n'est pas améliorable si $\delta\delta' \geq \rho > 0$ et $X \geq 0$.

On a vu que pour tout $A \geq 4$ il existe B tel que

$$\mathbf{P}(X \geq \xi) \leq \mathbf{P}\left(\sigma Y \geq \xi - A - B\delta\delta' \frac{\xi^2}{\sigma^2}\right) \quad (47)$$

Montrons que pour ξ de l'ordre de $\sigma^{4/3} \log \sigma$, B doit être supérieur à $1/6$.

Soit ξ le plus grand entier recentré inférieur à $\sigma^{4/3} \log \sigma$.

En utilisant la minoration de $\mathbf{P}(X = \xi)$ et $\mathbf{P}(Y \geq y) \leq 1/(\sqrt{2\pi}y) \exp(-y^2/2)$ il faut pour que 47 soit vérifiée avoir l'inégalité:

$$\begin{aligned}\frac{\xi^2}{2\sigma^2} - \frac{\xi^3 \delta\delta'}{6\sigma^4} + \log \sigma + \epsilon(\sigma) &> \frac{y^2}{2} + \log y \\ \text{où } y = \frac{\xi}{\sigma} - \frac{A}{\sigma} - B\delta\delta' \frac{\xi^2}{\sigma^3} \text{ et } \lim_{\sigma \rightarrow \infty} \epsilon(\sigma) &= 0\end{aligned}$$

En utilisant la définition de ξ , on obtient:

$$\frac{\sigma^{2/3}(\log \sigma)^2}{2} + \delta\delta'(B - \frac{1}{6})(\log \sigma)^3 + \epsilon'(\sigma) + \log \sigma \geq \frac{\sigma^{2/3}(\log \sigma)^2}{2} + \frac{\log \sigma}{3} + \log(\log \sigma)$$

Si $\delta\delta' \geq \rho > 0$, ceci entraine $B \geq 1/6$. Donc l'inégalité linéaire n'est pas améliorable.

References

- [1] BENETT G., 1962. *Probability inequalities for the sum of independants random variables*. J. AM. Statis.Assoc., 57, 33-45
- [2] BRETAGNOLLE J. & MASSART P., 1989. *Hungarian constructions from the non asymptotic view point*. Annals of probability, Vol 17, 239-256
- [3] CSÖRGÖ M. & REVESZ P., 1981. *Strong approximations in probability and statistics*. Academic Press, New York, 133-134
- [4] DVORETZKY A. & KIEFER J.C & WOLFOWITZ J., 1956. *Asymptotic minimax character of the sample distribution function and of the classical multinomial estimator*. Ann. Math. Stat., 33, 642-669
- [5] KIEFER J., 1969. *On the deviations in the Skorohod-Strassen approximation scheme*. Z. Wahrschein. Verw. Geb., 13, 321-332
- [6] KOMLÓS J. & MAJOR P. & TUSNÁDY G., 1975. *An approximation of partial sums of independent RV'- and the sample DF. I*. Z. warschein. Verw. Geb., 32, 111-131
- [7] MASSART P., 1990. *The tight constant of the Dvorestsky-Kiefer-Wolfowitz inequality*. A paraitre dans Annals of probability
- [8] SHORACK G.R. & WELLNER J.A., 1986. *Empirical processes with applications to statistics*. Wiley & Sons.
- [9] SKOROHOD A.V., 1976. *On a representation of random variables*. Th. Proba. Appl., 628-632
- [10] WELLNER J., 1978. *Limit theorems for the ratio of the empirical distribution function to the true distribution function*. Z .warschein. Verw. Geb., 45, 73-88

CHAPITRE 2

***APPROXIMATION FORTE DE SOUS FONCTIONS
DE REPARTITION EMPIRIQUE***

Approximation forte d'un vecteur de sous processus empiriques

Résumé

Soient X une variable aléatoire de fonction de répartition $F(t) = \mathbf{P}(X \leq t)$ avec F continue et R une variable aléatoire à valeur dans $1, \dots, m$. On définit les sous fonctions de répartition $F^j(t) = \mathbf{P}(X \leq t, R = j)$ et pour un n -échantillon du couple (X, R) les sous fonctions de répartition empirique:

$$\hat{F}_n^j(t) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{X_i \leq t} \mathbf{1}_{R_i=j}$$

On note $\alpha_n^j = \sqrt{n}(\hat{F}_n^j - F^j)$ les sous ponts empiriques associés. L'étude de la convergence et de la vitesse de convergence de statistiques sur des modèles censurés à plusieurs populations nécessite l'approximation forte du vecteur $(\alpha_n^j); j = 1, \dots, m$ par un processus gaussien vectoriel. Burke, Csörgö et Horváth [3] ont démontré en 1981 l'existence d'un tel processus $(W_n^j); j = 1, \dots, m$ et ont effectué un contrôle en probabilité de $\max_{1 \leq j \leq m} \|\alpha_n^j - W_n^j\|_\infty$ avec une borne exponentielle dépendant de m pour le terme d'erreur. Une autre construction permet d'obtenir un processus gaussien $(\tilde{W}_n^j); j = 1, \dots, m$ tel que le contrôle en probabilité de $\max_{1 \leq j \leq m} \|\alpha_n^j - \tilde{W}_n^j\|_\infty$ donne pour le terme d'erreur une borne exponentielle avec des constantes explicites et sans dépendance de m . La clé de cette approximation est le théorème d'invariance de Komlos, Major et Tusnady [4] pour des fonctions de répartition classiques, ainsi que le résultat de Bretagnolle et Massart [2] pour l'évaluation des constantes. On utilise la même construction hongroise que dans ces deux articles, mais cette construction est réalisée conditionnellement à la variable multinomiale $(\sum_{i=1}^n \mathbf{1}_{R_i=j}); j = 1, \dots, m$.

1 Introduction

Modèle On suppose que toutes les variables sont définies sur un même espace de probabilité Ω suffisamment riche pour assurer l'existence simultanée sur Ω de tous les processus utilisés.

Soient $X_1, \dots, X_n$ des variables aléatoires indépendantes et de même loi F :

$$F(t) = \mathbf{P}(X \leq t)$$

On note $\hat{F}_n$ la fonction de répartition empirique:

$$\hat{F}_n(t) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{X_i \leq t}$$

Soit R une variable discrète prenant ses valeurs dans $\{1, \dots, m\}$. On note:

$$N_j = \sum_{i=1}^n \mathbf{1}_{R_i=j}$$

$$p_j = \mathbf{E}(N_j/n) = \mathbf{P}(R = j)$$

La variable $N = (N_1, \dots, N_m)$ suit la loi multinomiale $\mathcal{M}_m(n, p_1, \dots, p_m)$.

Certains problèmes statistiques (par exemple l'étude de modèles avec censure aléatoire) conduisent à considérer les sous fonctions de répartition empirique:

$$\hat{F}_n^j(t) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{X_i \leq t, R_i=j}$$

Notons F^j les sous fonctions de répartition associées que nous supposons continues:

$$F^j(t) = \mathbf{E}(\hat{F}_n^j(t)) = \mathbf{P}(X \leq t, R = j)$$

L'approximation forte du processus $\hat{F}_n - F$ par un pont brownien est maintenant bien connue (voir Komlos, Major et Tusnady [4] et Bretagnolle et Massart [2] pour l'évaluation des constantes). Celle du processus vectoriel $(\hat{F}_n^j - F^j)_{j=1}^m$ a été traitée par Burke, Csörgö et Horváth [3]. Afin d'améliorer leur résultat, nous formalisons différemment le problème.

Conditionnellement à $\{N = k\}$ où

$$k \in \mathcal{K} = \{(k_1, \dots, k_m); k_j \in \mathbf{N}; \sum_{j=1}^m k_j = n\} \quad (1)$$

le processus $(n\hat{F}_n^j)_{j=1}^m$ a même loi qu'un processus V_k :

$$V_k = (k_1 \hat{U}_{k_1}^1, \dots, k_m \hat{U}_{k_m}^m)$$

tel que:

1. V_k est indépendant de N .
2. $\hat{U}_{k_j}^j$ est une k_j -fonction de répartition empirique de la fonction de répartition $u_j(t) = \mathbf{P}(X \leq t/R = j)$ que l'on a supposée continue.
3. Les composantes de V_k sont mutuellement indépendantes.

Remarque 1: L'assertion 1 n'est vérifiée que si Ω est suffisamment riche.

Remarque 2: L'assertion 1 entraîne l'égalité en loi de V_N et du processus vectoriel $(n\hat{F}_n^j)_{j=1}^m$.

Remarque 3: La remarque 2 entraîne:

$$\mathbf{E}(V_N) = \mathbf{E}((n\hat{F}_n^j)_{j=1}^m) = (np_1u_1, \dots, np_mu_m)$$

Notations

Pour toute fonction $\phi : \mathbf{R}^+ \rightarrow \mathbf{R}$, on pose

$$\|\phi\|_\infty = \sup_{t \in [0; +\infty]} |\phi(t)|$$

On note B le mouvement brownien sur $[0; 1]$. Pour toute fonction $f : \mathbf{R}^+ \rightarrow [0; 1]$ telle que $f(0) = 0$, $f(1) = 1$ on définit le processus $B^o(f)$ par

$$B^o(f(t)) = B(f(t)) - f(t)B(1)$$

Approximation forte d'un vecteur de sous processus empiriques

Si l'espace de probabilité Ω est suffisamment riche, on a le théorème suivant:

Theoreme 1 *Pour tout entier $n \geq 3$ et pour tout entier $m \geq 1$, il existe un vecteur gaussien centré $Z = (Z_1, \dots, Z_m)$, de matrice de covariance égale à la matrice de covariance de $N/\sqrt{n}$ et un processus $B = (B_1, \dots, B_m)$ où les B_j sont des mouvements browniens mutuellement indépendants et indépendants de Z et de N tels que pour tout x positif:*

1.

$$\mathbf{P}(\max_{1 \leq j \leq m} \|n(\hat{F}_n^j - p_j u_j) - \sqrt{np_j} B_j^o(u_j) - \sqrt{n} u_j Z_j\|_\infty \geq x + 40 \log n) \leq 7 \exp(-x/20)$$

2.

$$\mathbf{P}(\max_{1 \leq j \leq m} |N_j - np_j - \sqrt{n} Z_j| \geq x + 24 \log n) \leq 2 \exp(-x/12)$$

Cas d'un modèle censuré à plusieurs populations

On considère n individus indépendants répartis sur m populations disjointes. Soient T et C les variables aléatoires représentant la survie et la censure, et R une variable aléatoire à valeur dans $1, \dots, m$ qui indique la population d'origine.

Soit $(X_i, d_i, R_i); i = 1, \dots, n$ le n -échantillon obtenu avec $X = \min(T, C)$, $d = \mathbf{1}_{T \leq C}$ et $R_i = j$ si le i^{eme} individu provient de la j^{eme} population. On suppose que X a une

fonction de répartition continue.

Pour tout $j \in 1, \dots, m$, on note:

$$N_j^1 = \sum_{i=1}^n \mathbf{1}_{d_i=1, R_i=j} \quad N_j^0 = \sum_{i=1}^n \mathbf{1}_{d_i=0, R_i=j}$$

$$p_j^1 = \mathbf{P}(d = 1, R = j) \quad p_j^0 = \mathbf{P}(d = 0, R = j)$$

$$u_j^1 = \mathbf{P}(X \leq t/d = 1, R = j) \quad u_j^0 = \mathbf{P}(X \leq t/d = 0, R = j)$$

Le vecteur $N = (N_j^1, N_j^0); j = 1, \dots, m$ suit une loi multinomiale $M_{2m}(n, p_1^1, p_1^0, \dots, p_m^1, p_m^0)$.

Soit le processus vectoriel de $\mathbf{R}^{2m}$: $(n\hat{F}_j^1, n\hat{F}_j^0); j = 1, \dots, m$ où $\hat{F}_j^\epsilon; j = 1, \dots, m; \epsilon = 0, 1$ est la sous fonction de répartition empirique:

$$\hat{F}_j^\epsilon(t) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{X_i \leq t, d_i=\epsilon, R_i=j}$$

Le processus $(n\hat{F}_j^1, n\hat{F}_j^0)_{j=1}^m$ conditionné par $\{N = k\}$, où

$$k \in \{(k_1, \dots, k_{2m}); k_j \in \mathbf{N}; \sum_{j=1, \dots, 2m} k_j = n\}$$

a même loi qu'un processus V_k :

$$V_k = (k_1 \hat{U}_{k_1}^1, \dots, k_m \hat{U}_{k_{2m}}^{2m})$$

tel que:

1. $\hat{U}_{k_l}^l$ est une k_l -fonction de répartition empirique de la loi u_l^1 si l est pair et de la loi u_l^0 si l est impair.
2. Les composantes de V_k sont mutuellement indépendantes.

Si l'espace de probabilité Ω est suffisamment riche, on peut choisir V_k indépendant de N et appliquer le théorème 1:

Application du théorème 1 *Approximation forte du vecteur des sous empiriques d'un modèle censuré à plusieurs populations:*

Pour tous entiers $n \geq 3$ et $m \geq 1$, il existe un vecteur gaussien centré $Z = (Z_j^1, Z_j^0); j = 1, \dots, m$ avec

$$\begin{aligned} \mathbf{E}(Z_j^\epsilon)^2 &= p_j^\epsilon(1 - p_j^\epsilon) \\ \mathbf{E}(Z_j^\epsilon Z_{j'}^{\epsilon'}) &= -p_j^\epsilon p_{j'}^{\epsilon'} \end{aligned}$$

et un processus $B = (B_j^1, B_j^0); j = 1, \dots, m$ où les B_j^ϵ sont des mouvements browniens mutuellement indépendants et indépendants de Z et de N tels que pour tout x positif:

$$\mathbf{P}\left(\max_{\epsilon=0,1} \max_{1 \leq j \leq m} \|\hat{F}_j^\epsilon - np_j^\epsilon u_j^\epsilon - \sqrt{np_j^\epsilon}(B_j^{\epsilon^0}(u_j^\epsilon)) - \sqrt{n}u_j^\epsilon Z_j^\epsilon\|_\infty \geq x + 40 \log n\right) \leq 7 \exp(-x/20)$$

Remarque 4: on obtient de même l'approximation forte pour un modèle censuré à une seule population ($\epsilon \in \{0, 1\}, j = 1$) ou pour un modèle non censuré à plusieurs populations ($\epsilon = 1, j = 1, \dots, m$).

2 Résultats d'approximation forte et inégalités exponentielles

Les notations sont celles de l'introduction. Le théorème 2 est dû à Komlos, Major et Tusnady [4] et à Bretagnolle et Massart [2] pour l'évaluation des constantes.

Theoreme 2 (Approximation forte d'une fonction de répartition empirique) *Si l'espace de probabilité Ω est suffisamment riche et si la fonction de répartition F est continue, pour tout entier $n \geq 2$, il existe un mouvement brownien B tel que pour tout x positif:*

$$\mathbf{P} \left(\| n(\hat{F}_n - F) - \sqrt{n}B^o(F) \|_\infty \geq x + 12 \log n \right) \leq 2 \exp(-x/6)$$

Le lemme 1 est dû à Bennett [1] et à Wellner[6]. Une preuve en est donnée dans le livre de Shorack et Wellner[5].

Lemme 1 *Soit Z une binomiale de moyenne m . Alors pour tout x positif et tout signe ϵ on a:*

$$\mathbf{P}(\epsilon(Z - m) \geq x) \leq \exp \left(-mh \left(\frac{\epsilon x}{m} \right) \right)$$

$$\begin{aligned} \text{où } h(t) &= (1+t) \log(1+t) - t \quad \text{pour } t > -1 \\ &= +\infty \quad \text{pour } t \leq -1 \end{aligned}$$

Comme $h(t) \leq h(-t)$, on a immédiatement:

$$\mathbf{P}(|Z - m| \geq x) \leq 2 \exp(-mh(x/m))$$

On rappelle l'inégalité de Kolmogorov:

Lemme 3 *Pour tout x positif:*

$$\mathbf{P} \left(\sup_{t \in [0,1]} |B^o(t)| \geq x \right) \leq 2 \exp(-2x^2)$$

Le lemme 4 donne l'approximation forte d'une variable multinomiale $\mathcal{M}_m(n, p_1, \dots, p_m)$ avec, pour le contrôle en probabilité du terme d'erreur, une borne exponentielle qui ne dépend pas des paramètres.

Lemme 4 *Soit $N = (N_1, \dots, N_m)$ une variable multinomiale $\mathcal{M}_m(n, p_1, \dots, p_m)$ définie sur un espace de probabilité Ω . Si Ω est suffisamment riche alors pour tout entier $n \geq 2$ il existe un vecteur gaussien centré $Z = (Z_1, \dots, Z_m)$ de matrice de covariance égale à la matrice de covariance de $N/\sqrt{n}$ tel que pour tout x positif:*

$$\mathbf{P} \left(\max_{1 \leq j \leq m} |N_j - np_j - \sqrt{n}Z_j| \geq x + 24 \log n \right) \leq 2 \exp(-x/12)$$

Preuve du lemme 4: Soit $I = [a ; b]$ $a \leq b$ un intervalle de $[0 ; 1]$. Pour toute fonction $f : [0 ; 1] \rightarrow \mathbf{R}$, on note $f(I) = f(b) - f(a)$ et $|I| = b - a$. Soient

$$\begin{cases} I_1 = [0 ; p_1] \\ I_j = [p_1 + \dots + p_{j-1} ; p_1 + \dots + p_j] \quad j = 2, \dots, m \end{cases}$$

La variable multinomiale N a même loi que le vecteur $(n\hat{G}_n(I_1), \dots, n\hat{G}_n(I_m))$ où $\hat{G}_n$ est une n -fonction de répartition empirique de la loi uniforme sur $[0 ; 1]$. Dès que $n \geq 2$, d'après le théorème 2, il existe un mouvement brownien B tel que pour tout x positif:

$$\begin{aligned} & \mathbf{P} \left(\max_{1 \leq j \leq m} |n(\hat{G}_n(I_j) - |I_j|) - \sqrt{n}B^o(I_j)| \geq x + 24 \log n \right) \\ & \leq \mathbf{P} \left(2 \sup_{0 \leq t \leq 1} |n(\hat{G}_n(t) - t) - \sqrt{n}B^o(t)| \geq x + 24 \log n \right) \\ & = 2 \exp(-x/12) \end{aligned}$$

On prend pour Z une version du vecteur $B^o(I_1), \dots, B^o(I_m)$. $\square$

3 Preuve du théorème 1

On rappelle que $\mathcal{K}$ défini en (1) est l'ensemble des valeurs prises par N .

3.1 Construction des variables

D'après le lemme 4, il existe un vecteur gaussien centré $Z = (Z_1, \dots, Z_m)$ tel que $\mathbf{E}(Z_i^2) = p_i(1 - p_i)$ et $\mathbf{E}(Z_i Z_j) = -p_i p_j$ pour $i \neq j$ tel que pour tout x positif:

$$\mathbf{P} \left(\max_{1 \leq j \leq m} |N_j - np_j - \sqrt{n}Z_j| \geq x + 24 \log n \right) \leq 2 \exp(-x/12) \quad (2)$$

Soit $\tilde{B} = (\tilde{B}_1, \dots, \tilde{B}_m)$ où les $\tilde{B}_j$ sont des mouvements browniens mutuellement indépendants et indépendants de Z et de N .

Sur chaque évènement $\{N = k\}$ où $k \in \mathcal{K}$, on obtient par la construction hongroise pour les fonctions de répartition empirique du théorème 2 une version V_k de $(n\hat{F}_n^j)_{j=1}^m$:

$$V_k = (k_1 \hat{U}_{k_1}^1, \dots, k_m \hat{U}_{k_m}^m)$$

où $\hat{U}_{k_j}^j$ est une k_j -fonction de répartition empirique de la loi u_j , et $\hat{U}_{k_j}^j$ est construite à partir de $\tilde{B}_j$.

Conséquences:

1. Les processus V_N et $(n\hat{F}_n^j)_{j=1}^m$ sont égaux en loi.
2. Pour tout $k \in \mathcal{K}$, V_k est indépendant de N .

3. Pour tout entier $k_j \in 2, \dots, n$ et pour tout x positif, on a:

$$\mathbf{P} \left(\| k_j(U_{k_j}^j - u_j) - \sqrt{k_j}(\tilde{B}_j^o(u_j)) \|_\infty \geq x + 12 \log k_j \right) \leq 2 \exp(-x/6) \quad (3)$$

La conséquence 1 entraîne l'existence d'un processus B tel que le vecteur de processus $((n\hat{F}_n^j)_{j=1}^m, N, B, Z)$ soit égal en loi au vecteur de processus $(V_N, N, \tilde{B}, Z)$. En particulier, B a même loi que $\tilde{B}$ et est indépendant de Z et N .

La probabilité à majorer s'écrit:

$$\begin{aligned} P &= \mathbf{P} \left(\max_{1 \leq j \leq m} \| n(\hat{F}_n^j - p_j u_j) - \sqrt{np_j}(B_j^o(u_j)) - \sqrt{n}u_j Z_j \|_\infty \geq x + 40 \log n \right) \\ &= \mathbf{P} \left(\max_{1 \leq j \leq m} \| N_j U_{N_j}^j - np_j u_j - \sqrt{np_j}(\tilde{B}_j^o(u_j)) - \sqrt{n}u_j Z_j \|_\infty \geq x + 40 \log n \right) \end{aligned}$$

Remarque 5: Le processus $N_j U_{N_j}^j$ ne dépend que de N_j et $\tilde{B}_j$ et est donc indépendant de $\tilde{B}_l$ pour $j \neq l$. De même, $n\hat{F}_n^j$ est indépendant de B_l pour $j \neq l$.

Remarque 6: Dans ce qui précède, l'espace de probabilité Ω est supposé suffisamment riche pour assurer l'existence simultanée de $(n\hat{F}_n^j)_{j=1}^m$, de Z , de $\tilde{B}$ et B indépendants de Z et de N .

3.2 Contrôle de la probabilité P

Soient $\lambda_1 = 9/29, \lambda_2 = 18/29$ et $\lambda_3 = 2/29$. On majore la probabilité P :

$$P \leq P_1 + P_2 + P_3$$

$$\begin{aligned} \text{avec } P_1 &= \mathbf{P}(\max_{1 \leq j \leq m} \| N_j(U_{N_j}^j - u_j) - \sqrt{N_j}(\tilde{B}_j^o(u_j)) \|_\infty \geq \lambda_1 x + 12 \log n) \\ P_2 &= \mathbf{P}(\max_{1 \leq j \leq m} | N_j - np_j - \sqrt{n}Z_j | \geq \lambda_2 x + 24 \log n) \\ P_3 &= \mathbf{P}(\max_{1 \leq j \leq m} \| (\sqrt{N_j} - \sqrt{np_j})(\tilde{B}_j^o(u_j)) \|_\infty \geq \lambda_3 x + 4 \log n) \end{aligned}$$

Contrôle de P_1 :

$$P_1 \leq \sum_{k \in K} \sum_{j=1}^m \mathbf{P} \left(\| N_j(U_{N_j}^j - u_j) - \sqrt{N_j}(\tilde{B}_j^o(u_j)) \|_\infty \geq \lambda_1 x + 12 \log n, N = k \right)$$

Pour tout $k \in K$, on définit $I_k = \{j \in \{1, \dots, m\}; k_j \geq 2\}$ et $\bar{I}_k = \{1, \dots, m\} \setminus I_k$. En utilisant l'indépendance de N avec $(U_{k_j}^j, \tilde{B}_j); j = 1, \dots, m$ on obtient:

$$P_1 \leq \max_{k \in K} \left(\sum_{j \in I_k} P_{k_j}^j + \sum_{j \in \bar{I}_k} P_{k_j}^j \right) \quad (4)$$

$$\text{où } P_{k_j}^j = \mathbf{P} \left(\| k_j(U_{k_j}^j - u_j) - \sqrt{k_j}(\tilde{B}_j^o(u_j)) \|_\infty \geq \lambda_1 x + 12 \log n \right)$$

- Si $j \in \bar{I}_k$ alors k_j vaut 0 ou 1. Comme $P_0^j = 0$ on a:

$$\sum_{j \in \bar{I}_k} P_{k_j}^j = \sum_{j \in \bar{I}_k^1} P_{k_j}^j \text{ où } \bar{I}_k^1 = \{j \in \{1, \dots, m\}; k_j = 1\}$$

On remarque $\text{Card}(\bar{I}_k^1) \leq n$ et $\|U_1^j\|_\infty \leq 1$ et on obtient

$$\sum_{j \in \bar{I}_k} P_{k_j}^j \leq n \max_{\bar{I}_k^1} \mathbf{P}(2 + \|B_j^2(u_j)\|_\infty \geq \lambda_1 x + 12 \log n)$$

Comme $12 \log n - 2 \geq 10 \log n$ dès que $n \geq 3$, on peut utiliser le lemme 3. Finalement,

$$\begin{aligned} \sum_{j \in \bar{I}_k} P_{k_j}^j &\leq 2n \exp(-2(\lambda_1 x + 10 \log n)^2) \\ &\leq 2\epsilon \exp(-x/20) \quad \text{avec } \epsilon = 3^{-199} \end{aligned} \quad (5)$$

- Si $j \in I_k$, on peut utiliser l'inégalité (3):

$$\begin{aligned} \sum_{j \in I_k} P_{k_j}^j &\leq \sum_{j \in I_k} 2 \exp\left(-\frac{\lambda_1 x}{6} - 2 \log\left(\frac{n}{k_j}\right)\right) \\ &\leq 2 \exp(-x/20) \end{aligned} \quad (6)$$

Les inégalités (4), (5), (6) entraînent:

$$P_1 \leq 2(1 + \epsilon) \exp(-x/20) \quad (7)$$

Contrôle de P_2 :

D'après l'inégalité (2), on a immédiatement:

$$P_2 \leq 2 \exp(-\lambda_2 x/12) \leq 2 \exp(-x/20) \quad (8)$$

Contrôle de P_3 :

Soient $r = 0.2962$ et $\alpha = 0.3335$. On définit I et $\bar{I}$ par :

$$\begin{aligned} I &= \{j \in \{1, \dots, m\}; np_j \geq r(\lambda_3 x + 4 \log n)\} \\ \bar{I} &= \{1, \dots, m\} \setminus I \end{aligned}$$

On majore la probabilité P_3 :

$$P_3 \leq Q_1 + Q_2 + Q_3 \quad (9)$$

$$\begin{aligned} \text{avec } Q_1 &= \mathbf{P} \left(\max_{j \in I} \left| \frac{N_j - np_j}{\sqrt{np_j}} \right| \parallel \tilde{B}_j^o(u_j) \parallel_\infty \geq \lambda_3 x + 4 \log n \right) \\ Q_2 &= \mathbf{P} \left(\max_{j \in I} \sqrt{N_j} \parallel \tilde{B}_j^o(u_j) \parallel_\infty \geq (1 - \alpha)(\lambda_3 x + 4 \log n) \right) \\ Q_3 &= \mathbf{P} \left(\max_{j \in I} \sqrt{np_j} \parallel \tilde{B}_j^o(u_j) \parallel_\infty \geq \alpha (\lambda_3 x + 4 \log n) \right) \end{aligned}$$

Contrôle de Q_1 :

On remarque $\text{Card}(I) \leq n/(4r \log 3)$ et on pose $\gamma = 2.66$. On a:

$$\begin{aligned} Q_1 \leq \frac{n}{4r \log 3} \max_{j \in I} &\left[\mathbf{P} \left(|N_j - np_j| \geq \sqrt{np_j \gamma (\lambda_3 x + 4 \log n)} \right) \right. \\ &\left. + \mathbf{P} \left(\parallel \tilde{B}_j^o(u_j) \parallel_\infty \geq \sqrt{(\lambda_3 x + 4 \log n)/\gamma} \right) \right] \end{aligned}$$

D'après les lemmes 2 et 3:

$$\begin{aligned} Q_1 \leq \frac{n}{2r \log 3} \max_{j \in I} &\exp \left(-np_j h \left(\sqrt{\gamma (\lambda_3 x + 4 \log n)/np_j} \right) \right) \\ &+ \exp \left(\frac{-2(\lambda_3 x + 4 \log n)}{\gamma} \right) \end{aligned}$$

Pour $t \leq \sqrt{\gamma/r}$ on a $h(t) \geq 0.282t^2$. On obtient:

$$Q_1 \leq \frac{1}{9r \log 3} \exp(-x/20) \leq 0.35 \exp(-x/20) \quad (10)$$

Contrôle de Q_2 :

On pose $\beta = 1.775$ et on définit $\bar{I}_N$ par

$$\bar{I}_N = \{j \in \{1, \dots, m\} ; N_j \neq 0\} \cap \bar{I}$$

Comme $\text{Card}(\bar{I}_N) \leq n$ et $\bar{I}_N \subset \bar{I}$ on a :

$$\begin{aligned} Q_2 \leq n \max_{j \in \bar{I}} &[\mathbf{P}(N_j \geq \beta(1 - \alpha)(\lambda_3 x + 4 \log n)) \\ &+ \mathbf{P} \left(\parallel (\tilde{B}_j^o(u_j)) \parallel_\infty \geq \sqrt{(1 - \alpha)(\lambda_3 x + 4 \log n)/\beta} \right)] \end{aligned}$$

Soit $Q_j = \mathbf{P}(N_j \geq \beta(1 - \alpha)(\lambda_3 x + 4 \log n))$. Sur $\bar{I}$, on a:

$$Q_j \leq \mathbf{P}(N_j - np_j \geq (\beta(1 - \alpha) - r)(\lambda_3 x + 4 \log n))$$

D'après le lemme 2, pour tout $j \in \bar{I}$:

$$Q_j \leq \exp \left(-np_j h \left(\frac{(\beta(1 - \alpha) - r)(\lambda_3 x + 4 \log n)}{np_j} \right) \right)$$

Pour $t \geq (\beta(1 - \alpha) - r)/r$ on a $h(t) \geq 0.847t$ et on obtient

$$Q_j \leq \exp \left(-\frac{x}{20} - 3 \log n \right)$$

Cette dernière inégalité et le lemme 3 entraînent:

$$Q_2 \leq (1/3) \exp(-x/20) \quad (11)$$

Contrôle de Q_3 :

$$Q_3 \leq \sum_{j \in \bar{I}} \left[\mathbf{1} \left(\sqrt{np_j(\lambda_3 x + 4 \log(1/p_j))} \geq \sqrt{r}(\lambda_3 x - 4 \log p_j) \right) \right. \\ \left. + \mathbf{P} \left(\| \tilde{B}_j^o(u_j) \|_\infty \geq \alpha \sqrt{(\lambda_3 x + 4 \log n)/r} \right) \right]$$

- Si $1 < np_j < r(\lambda_3 x + 4 \log n)$ alors

$$\sqrt{np_j(\lambda_3 x + 4 \log(1/p_j))} < \sqrt{r}(\lambda_3 x + 4 \log n)$$

- Si $np_j \leq 1$ alors

$$\frac{\lambda_3 x + 4 \log n}{\sqrt{\lambda_3 x + 4 \log(1/p_j)}} \geq \frac{2 \log n}{\sqrt{\log(1/p_j)}}$$

La fonction $f(x) = -x \log x$ est croissante pour $x \leq 1/e$. Ici, $p_j \leq 1/n \leq 1/e$ dès que $n \geq 3$ donc $p_j \log(1/p_j) \leq (\log n)/n \leq 4r(\log n)^2/n$.

Avec le lemme 3 et ce qui précède, on obtient:

$$Q_3 \leq \sum_{j \in \bar{I}} 2 \exp \left(-\frac{2\alpha^2}{r} (\lambda_3 x - 4 \log p_j) \right) \\ \leq 2 \exp(-x/20) \quad (12)$$

Les inégalités (9), (10), (11), (12) entraînent:

$$P_3 \leq 2.7 \exp(-x/20) \quad (13)$$

Les inégalités (7), (8), (13), entraînent:

$$P \leq 7 \exp(-x/20)$$

References

- [1] BENETT G., 1962. *Probability inequalities for the sum of independants random variables*. J. AM. Statis. Assoc., 57, 33-45
- [2] BRETAGNOLLE J. & MASSART P., 1989. *Hungarian constructions from the non asymptotic view point*. Annals of probability, Vol 17, 239-256
- [3] BURKE M.D. & CSÖRGÖ S. & HORVÁTH L., 1981. *Strong Approximations of Some Biometric Estimates under Random Censorship*. Z. Warschein. Verw. Geb., 56, 87-112
- [4] KOMLÓS J. & MAJOR P. & TUSNÁDY G., 1975. *An approximation of partial sums of independent RV'- and the sample DF. I*. Z. Warschein. Verw. Geb., 32, 111-131
- [5] SHORACK G.R. & WELLNER J.A., 1986. *Empirical processes with applications to statistics*. Wiley & Sons.
- [6] WELLNER J., 1978 *Limit theorems for the ratio of the empirical distribution function to the true distribution function*. Z. Warschein. Verw. Geb., 45, 73-88

SECONDE PARTIE

***APPLICATION DES PRINCIPES D'INVARIANCE A
LA STATISTIQUE DE MODELES CENSURES***

Etude de certaines statistiques sur un modèle censuré à deux populations

Résumé Lors de l'étude d'une variable aléatoire T , les observations $(T_1, \dots, T_n)$ peuvent être perturbées par une censure aléatoire C . On observe alors le n -échantillon:

$$(T_1 \wedge C_1, \dots, T_n \wedge C_n)$$

qui ne permet pas l'accès à la fonction de répartition empirique de T .

Deux méthodes sont utilisées pour l'étude de tels modèles: les techniques martingales et les principes d'invariance. Les propriétés des estimateurs de $P(T \geq t)$ ont été traitées par différents auteurs, par exemple et Gill [9] et [10] (approche martingale), Burke, Csörgö et Horváth [4] (principes d'invariance) ou Einmahl et Koning [8] qui associent les deux méthodes. Les principes d'invariance permettent d'obtenir l'approximation forte des statistiques par le processus gaussien associé (éventuellement en précisant la vitesse de convergence comme Burke, Csörgö et Horváth [4]) qui est un résultat plus précis que la seule convergence en loi obtenue par l'approche martingale (en particulier, si l'on n'utilise que les outils martingales, lors de la construction de tests le comportement asymptotique des statistiques n'est connu que sous l'hypothèse nulle).

Nous nous intéressons ici à la comparaison de la loi de la variable T entre deux populations A et B lorsque l'on dispose des deux séries d'observations:

$$\begin{aligned} & (T_1^A \wedge C_1^A, \dots, T_{n_1}^A \wedge C_{n_1}^A) \\ & (T_1^B \wedge C_1^B, \dots, T_{n_2}^B \wedge C_{n_2}^B) \end{aligned}$$

et nous adoptons la formalisation du modèle de Cox: dans les calculs, T^A et T^B peuvent être considérés comme des temps de survie qui suivent des lois exponentielles de paramètre θ_A et θ_B . A notre connaissance, les statistiques utilisées (statistique du logrank, statistique de Gehan ou statistique de Prentice) n'ont été l'objet que de l'approche martingale qui n'établit la convergence en loi vers une variable gaussienne que sous l'hypothèse d'égalité ou de contiguïté des paramètres, sans préciser la vitesse de convergence. (Par exemple, Gill [9] ou Harrington D. et Fleming T. [11]).

Les principes d'invariance donnent l'approximation forte des statistiques temporelles de Gehan et du logrank par le processus gaussien correspondant, avec une erreur d'ordre $(\log n)/\sqrt{n}$ lorsque l'on arrête les observations à une date fixe, dans le cas contraire cette erreur est d'ordre $(\log^2 n)/\sqrt{n}$ pour la statistique du logrank et reste d'ordre $(\log n)/\sqrt{n}$ pour la statistique de Gehan.

La clé de la démonstration est l'approximation forte de sous fonctions de répartition empirique, ainsi que le développement de certaines intégrales où elles interviennent. Ces intégrales ne sont pas restreintes à des compacts, les observations ne sont donc pas forcément arrêtées à une date fixe et l'on peut utiliser toute l'information.

Ce développement peut être intéressant en lui-même pour des problèmes autres que la comparaison des temps de survie entre deux populations. En particulier, il permet d'obtenir l'approximation forte et la vitesse de convergence du gradient du logarithme de la vraisemblance de Cox, à partir desquelles on retrouve celles de l'estimateur du maximum de vraisemblance de Cox. Nous appliquons aussi ce développement aux estimateurs de la variance des statistiques considérées ce qui nous autorise, sous l'hypothèse nulle, à remplacer la variance par son estimateur. Sous l'hypothèse nulle, le résultat standard de convergence en loi de la statistique normalisée par l'estimateur de sa variance vers une gaussienne centrée réduite en découle, avec une précision supplémentaire: la vitesse de convergence.

Enfin, ce développement reste valide lorsque le nombre de populations est différent de deux et lorsque le modèle n'est pas censuré, et ceci pour des modèles censurés généraux autres que le modèle de Cox.

1 Introduction

1.1 modèle

Nous considérons une version simplifiée du modèle de Cox. Soit $(T_i, C_i, \delta_i)_{i=1}^n$ un n -échantillon indépendant où T est une variable de survie, C une variable de censure et δ une variable qui indique la population d'origine et que nous appelons covariable.

Nous supposons que le nombre de populations est égal à deux et que la covariable δ ne dépend pas du temps:

$$\delta_i = \begin{cases} 1 & \text{si la } i^{\text{eme}} \text{ observation provient d'une population A} \\ 0 & \text{si la } i^{\text{eme}} \text{ observation provient d'une population B} \end{cases}$$

Ce modèle est un modèle de Cox si:

- $P(T \geq t/\delta) = (h_0(t))^{\exp(\beta_0 \delta)}$
- C et T sont indépendantes conditionnellement à δ

Nous supposons h_0 continue. On notera:

$$\begin{aligned} P(C \geq t/\delta = 1) &= C_1(t) \\ P(C \geq t/\delta = 0) &= C_2(t) \end{aligned}$$

et on supposera C_1 et C_2 dérivables sur l'intérieur du support de la loi de C .

On observe $(X_i, d_i, \delta_i)_{i=1}^n$ où $X = T \wedge C$ et $d = \mathbf{1}_{T \leq C}$ et l'on s'intéresse à la vitesse de convergence vers la variable gaussienne correspondante de certaines statistiques sur un tel échantillon.

1.2 Cas des statistiques de rang

Les applications de la section 4 (Comparaison des temps de survie entre les deux populations) et de la section 5 (Estimateur du maximum de vraisemblance de Cox) concerneront des statistiques de rang. Sont ainsi dénommées les statistiques qui n'utilisent que les observations $(n_{1i}, n_{2i}, d_i, \delta_i)_{i=1}^n$ avec:

$$\begin{aligned} \mathcal{R}_{1i} &= \{j; X_j \geq X_i; \delta_j = 1\} & n_{1i} &= \text{Card}(\mathcal{R}_{1i}) \\ \mathcal{R}_{2i} &= \{j; X_j \geq X_i; \delta_j = 0\} & n_{2i} &= \text{Card}(\mathcal{R}_{2i}) \\ \mathcal{R}_i &= \mathcal{R}_{1i} \cup \mathcal{R}_{2i} & n_i &= n_{1i} + n_{2i} \end{aligned}$$

Les ensembles $\mathcal{R}_{1i}$ et $\mathcal{R}_{2i}$ désignent les individus encore non observés juste avant la date X_i dans la première et seconde population, nous les appelons ensemble des individus à risque en X_i dans la première et seconde population.

Autrement dit, on ne retient de l'information contenue dans les variables $(X_i)_{i=1}^n$ que l'ordre dans lequel elles apparaissent et on en oublie les valeurs observées.

L'ordre étant préservé par tous les changements de temps croissants, on peut supposer:

$$\begin{aligned} S_1(t) &= \mathbf{P}(T \geq t/\delta = 1) = \exp(-\theta_0 t) \quad t > 0 \\ S_2(t) &= \mathbf{P}(T \geq t/\delta = 0) = \exp(-t) \quad t > 0 \\ &\text{avec } \theta_0 = \exp(\beta_0) \end{aligned}$$

1.3 Notations

Soient les quatre sous fonctions de répartition empirique:

$$\hat{F}_1(t) = \left(\sum_{i=1}^n \mathbf{1}_{X_i \leq t, d_i = 1, \delta_i = 1} \right) / n$$

$$\hat{F}_2(t) = \left(\sum_{i=1}^n \mathbf{1}_{X_i \leq t, d_i = 0, \delta_i = 1} \right) / n$$

$$\hat{F}_3(t) = \left(\sum_{i=1}^n \mathbf{1}_{X_i \leq t, d_i = 1, \delta_i = 0} \right) / n$$

$$\hat{F}_4(t) = \left(\sum_{i=1}^n \mathbf{1}_{X_i \leq t, d_i = 0, \delta_i = 0} \right) / n$$

Et les quatre sous queues de répartition empirique:

$$\hat{G}_1(t) = \left(\sum_{i=1}^n \mathbf{1}_{X_i \geq t, d_i = 1, \delta_i = 1} \right) / n$$

$$\hat{G}_2(t) = \left(\sum_{i=1}^n \mathbf{1}_{X_i \geq t, d_i = 0, \delta_i = 1} \right) / n$$

$$\hat{G}_3(t) = \left(\sum_{i=1}^n \mathbf{1}_{X_i \geq t, d_i = 1, \delta_i = 0} \right) / n$$

$$\hat{G}_4(t) = \left(\sum_{i=1}^n \mathbf{1}_{X_i \geq t, d_i = 0, \delta_i = 0} \right) / n$$

Ces notations ne sont pas classiques (les sous fonctions de répartition empirique sont continues à droite et les sous queues de répartition empirique sont continues à gauche) mais les statistiques sur les modèles censurés les emploient simultanément.

On note:

$$\begin{aligned} F_j(t) &= \mathbf{E}(\hat{F}_j(t)) \\ G_j(t) &= \mathbf{E}(\hat{G}_j(t)) \end{aligned}$$

Nous numérotons les classes disjointes de 1 à 4:

- $A_1 = \{d = 1, \delta = 1\}$ est l'évènement 'non censuré dans la première population'
- $A_2 = \{d = 0, \delta = 1\}$ est l'évènement 'censuré dans la première population'
- $A_3 = \{d = 1, \delta = 0\}$ est l'évènement 'non censuré dans la seconde population'
- $A_4 = \{d = 0, \delta = 0\}$ est l'évènement 'censuré dans la seconde population'

Les probabilités de ces différentes classes sont:

$$p_j = \mathbf{P}(A_j) \quad (1)$$

On note $q_j = 1 - p_j$ et l'on suppose:

$$\omega = \inf_j p_j > 0$$

Remarque: Si ω est égal à zéro, il faut réduire le nombre de classes disjointes considérées à celles dont la probabilité est non nulle, par exemple si le modèle n'est pas censuré ou s'il ne comporte qu'une seule population ou encore si une seule des deux populations subit une censure.

Le vecteur $N = (N_1, \dots, N_4)$ avec

$$N_j = n\hat{F}_j(\infty)$$

suit une loi multinômiale $\mathcal{M}(n, p_1, \dots, p_4)$. Remarquons que:

$$\hat{G}_j(t) = N_j - \hat{F}_j(t^-)$$

Les lois conditionnelles sont notées:

$$\mu_j(t) = \mathbf{P}(X \leq t/A_j)$$

ce qui entraîne:

$$\begin{aligned} F_j(t) &= \mathbf{E}\hat{F}_j(t) = p_j\mu_j(t) \\ G_j(t) &= \mathbf{E}\hat{G}_j(t) = p_j(1 - \mu_j(t)) \end{aligned}$$

Si l'on note:

$$\eta_j = \begin{cases} 1 & \text{si } j \in \{1, 3\} \\ 0 & \text{si } j \in \{2, 4\} \end{cases} \quad \eta'_j = \begin{cases} 1 & \text{si } j \in \{1, 2\} \\ 0 & \text{si } j \in \{3, 4\} \end{cases}$$

on peut écrire F_j et G_j sous la forme:

$$F_j(t) \text{ ou } G_j(t) = \int_{I_F \text{ ou } I_G} \mathbf{P}(\delta = \eta'_j) [\eta_j C_{2-\eta'_j} + (1 - \eta_j) S_{2-\eta'_j}] \, d[-\eta_j S_{2-\eta'_j} - (1 - \eta_j) C_{2-\eta'_j}]$$

avec $I_F =]0; t[$ ou $I_G =]t; \infty[$

1.4 Notation p.co

Si $(X_n)_n$ est une suite de variables aléatoires et $(a_n)_n$ une suite de réels positifs, la notation:

$$|X_n| \leq a_n \quad \text{p.co}$$

signifie qu'il existe une suite positive $(U_n)_n$ telle que $\sum_n U_n < \infty$ et:

$$\mathbf{P}(|X_n| > a_n) \leq U_n$$

En particulier, $|X_n| \leq A_n$ p.co entraîne:

- $\mathbf{P}(|X_n| > a_n) \rightarrow 0$ quand $n \rightarrow \infty$
- Pour presque tout événement ω , il existe $n_0(\omega)$ tel que pour tout $n \geq n_0$:

$$|X_n(\omega)| \leq a_n$$

2 Principe d'invariance et inégalités exponentielles

On note U_n la n -fonction de répartition empirique de la loi uniforme sur $]0; 1[$ et V_n une n -sous fonction de répartition empirique de la loi uniforme sur $]0; 1[$:

$$V_N = U_N/n$$

où N suit une loi binômiale $\mathcal{B}(n, p)$ avec $p \geq \omega > 0$. On suppose de plus que kU_k est indépendant de N pour tout entier $k \in \{0, \dots, n\}$.

On désigne par B le mouvement brownien sur $\mathbf{R}^+$ et par B° le pont brownien:

$$B^\circ(t) = B(t) - tB(1) \quad t \in [0; 1]$$

Par analogie, on note U_n° le pont empirique:

$$U^\circ(t) = U_n(t) - t \quad t \in [0; 1]$$

Enfin, si $f : \mathbf{R}^+ \rightarrow \mathbf{R}$ on pose:

$$\|f\|_\infty = \sup_{t \in \mathbf{R}^+} |f(t)|$$

et si $f : [0; 1] \rightarrow \mathbf{R}$ on pose:

$$\|f\|_\infty = \sup_{t \in [0; 1]} |f(t)|$$

2.1 Principe d'invariance

Conditionnellement à $N = (n_1, n_2, n_3, n_4)$, le vecteur $(n_j \hat{F}_n^j)_{j=1}^4$ a même loi qu'un processus $(n_j \hat{H}_{n_j}^j)_{j=1}^4$ tel que:

- $\hat{H}_{n_j}^j$ est une vraie fonction de répartition empirique de la loi μ_j
- Les composantes $n_j \hat{H}_{n_j}^j$ sont mutuellement indépendantes et sont indépendantes de N

D'après Castelle [5], on a le théorème:

Theoreme 1 *Pour tout entier $n \geq 3$, il existe un vecteur gaussien centré $Z = (Z_1, \dots, Z_4)$, de matrice de covariance égale à la matrice de corrélation de $N/\sqrt{n}$ et un processus $B = (B_1, \dots, B_4)$ où les B_j sont des mouvements browniens mutuellement indépendants et indépendants de Z et de N tels que pour tout x positif l'on ait simultanément:*

1.

$$P(\max_{1 \leq j \leq 4} |N_j - np_j - \sqrt{np_j q_j} Z_j| \geq x + 24 \log n) \leq 2 \exp(-x/12)$$

2.

$$P(\max_{1 \leq j \leq m} \|n(\hat{F}_n^j - F_j) - \sqrt{np_j}(B_j^o(\mu_j)) - \sqrt{np_j q_j} \mu_j Z_j\|_\infty \geq x + 40 \log n) \leq 7 \exp(-x/20)$$

Remarque 1 Par construction (voir Castelle [5], Construction des variables) B_j^o est indépendant de $n(\hat{F}_n^i - F_i)$ pour $i \neq j$.

Remarque 2 Le vecteur Z et le processus B dépendent de n .

2.2 Inégalités exponentielles pour les processus et sous processus empiriques

L'inégalité suivante est due à Dvoretzky, Kiefer et Wolfowitz [7] et à Massart [12] pour l'évaluation des constantes.

Lemme 1 *Pour tout u positif*

$$P(n \|U_n^o\|_\infty \geq u) \leq 2 \exp(-2u^2/n)$$

Le lemme 2 est dû à Bennett [2] et à Wellner [14]. Une preuve en est donnée dans le livre de Shorack et Wellner [13] (inégalité 11.1.1 page 440).

Lemme 2 *Soit X une variable binômiale de moyenne m . Alors pour tout x positif et tout signe ϵ*

$$P(\epsilon(X - m) \geq x) \leq \exp(-mh(\frac{\epsilon x}{m}))$$

$$\text{où } h(t) = \begin{cases} (1+t) \log(1+t) - t & \text{si } t > -1 \\ +\infty & \text{sinon} \end{cases}$$

Comme $h(-t) \geq h(t)$ il vient immédiatement:

Lemme 3 $P(|X - m| \geq x) \leq 2 \exp(-mh(x/m))$

On trouve de nouveau l'inégalité suivante dans le livre de Shorack et Wellner [13] (inégalité 11.1.2 page 444):

Lemme 4 Pour tout $b \in]0; 1/2[$ et tout x positif

$$P(n \sup_{s \in [0; b]} |U_n^o(s)| \geq x) \leq 2 \exp\left(-nb(1-b)h\left(\frac{x}{nb}\right)\right)$$

Remarque 3 La fonction h possède les propriétés:

- Si $t \leq t_0$ alors $h(t) \geq (h(t_0)/t_0^2) t^2$
- Si $t \geq t_0$ alors $h(t) \geq (h(t_0)/t_0) t$

Le lemme 5 est une conséquence des lemmes 1 et 4.

Lemme 5 Pour tout $b \in]0; 1[$ et tout u positif

$$P(n \sup_{s \in [b; 1]} \left| \frac{U_n^o(s)}{\sqrt{s}} \right| \geq u) \leq \left(\frac{4}{b} \vee 6\right) \exp\left(-\frac{1}{10} \min\left(\frac{u^2}{n}, u\sqrt{b}\right)\right)$$

Preuve du lemme 5:

- 1^{er} cas: $b \leq 0.05264$ Soit $I_k = [s_k, s_{k+1}[$ une subdivision de $[b; 1]$ en intervalles de longueur b , et α un réel de $]0; 1[$.

$$\begin{aligned} P &\leq \frac{1}{b} \max_k P(n \sup_{s \in I_k} |U_n^o(s)| \geq u\sqrt{s_k}) \\ &\leq \frac{1}{b} \left[P(n \sup_{s \in [0; b]} |U_n^o(s)| \geq \alpha u\sqrt{b}) \right. \\ &\quad \left. + \max_k P(n | U_n^o(s_k)| \geq (1-\alpha)u\sqrt{s_k}) \right] \end{aligned}$$

En utilisant les lemmes 3 et 4:

$$P \leq \frac{2}{b} \left[\exp\left(-nb(1-b)h\left(\frac{\alpha u}{n\sqrt{b}}\right)\right) + \exp\left(-ns_k h\left(\frac{(1-\alpha)u}{n\sqrt{s_k}}\right)\right) \right]$$

On choisit $\alpha = 0.507$, on sépare les cas $u/(n\sqrt{b}) > 1$, $u/(n\sqrt{b}) < 1$, $u/(n\sqrt{s_k}) > 1$, $u/(n\sqrt{s_k}) < 1$ et on utilise la remarque 3 avec le choix $t_0 = 1$ pour obtenir le résultat.

- 2^{eme} cas: $b \geq 0.05264$ D'après le lemme 1:

$$P \leq P(n \sup_{s \in [b; 1]} |U_n^o(s)| \geq u\sqrt{b}) \leq 6 \exp(-u^2/(10n)) \quad \square$$

Inégalités exponentielles pour les sous empiriques

Les lemmes 6 et 7 sont les équivalents des lemmes 1 et 5 pour les sous empiriques:

Lemme 6 *Pour tout u positif*

$$P_1 = P(n \sup_{s \in [0;1]} |V_n(s) - ps| \geq u) \leq 8 \exp \left(-\frac{1}{6} \min(u, u^2/n) \right)$$

Lemme 7 *Pour tout $b \in]0;1[$, tout $\alpha \in]0;1[$ et tout u positif*

$$\begin{aligned} P_2 &= P(n \sup_{s \in [b;1]} \left| \frac{V_n(s) - ps}{\sqrt{s}} \right| \geq u) \\ &\leq \left(\frac{4}{b} \vee 6 \right) \exp \left(-\frac{1}{10} \min \left(\frac{\alpha^2 u^2}{n}, \alpha u \sqrt{b} \right) \right) + 2 \exp \left(-h(1-\alpha) \min \left(\frac{u^2}{n}, u \right) \right) \end{aligned}$$

Preuve des lemmes 6 et 7: On remarque:

$$n |V_n(s) - ps| \leq N |U_N^o(s)| + |s| |N - np|$$

On obtient:

$$\begin{aligned} P_1 &\leq P(N \|U_N^o\|_\infty \geq \alpha_1 u) + P(|N - np| \geq (1 - \alpha_1)u) \\ P_2 &\leq P(N \sup_{s \in [b;1]} \left| \frac{U_N^o(s)}{\sqrt{s}} \right| \geq \alpha_2 u) + P(|N - np| \geq (1 - \alpha_2)u) \end{aligned}$$

On choisit $\alpha_1 = 0.31$. L'indépendance de kU_k^o et de N , les lemmes 1,3,5 et la remarque 3 donnent alors le résultat.

Conséquences des lemmes 6 et 7

Lemme 8

$$\| \hat{G}_n^j - G_j \|_\infty \leq \sqrt{\frac{7 \log n}{n}} \quad p.co$$

Lemme 9 *On suppose $n \geq 2$. Soit β_n tel que $G_j(\beta_n) \geq (3p_j \log n)/(n \log 3)$. Alors:*

$$\begin{aligned} \sup_{t \in [0; \beta_n]} \left| \frac{\hat{G}_n^j(t) - G_j(t)}{\sqrt{G_j(t)}} \right| &\leq \frac{14}{\sqrt{p_j}} \left(\frac{\log n}{n} \right)^{1/2} & p.co \\ \sup_{t \in [0; \beta_n]} \left| \frac{\hat{G}_n^j(t) - G_j(t)}{G_j(t)} \right| &\leq \frac{9}{p_j} & p.co \end{aligned}$$

Remarque: Pour $n \geq 2$, on a $(3 \log n)/(n \log 3) \leq 1$, ce qui assure l'existence de β_n .

Preuve des lemmes 8 et 9: On montre le lemme 9 à partir du lemme 7, le lemme 8 s'obtient de la même manière à partir du lemme 6.

Sur $N_j = n_j$, le processus $n\hat{G}_n^j$ a même loi que le processus $n_j\hat{H}_{n_j}^j$, qui est une vraie queue empirique de la loi μ_j , et que l'on choisit indépendant de N_j . Avec le changement de variable $s = 1 - \mu_j(t)$, on a:

$$\begin{aligned} \mathbf{P} &= \mathbf{P} \left(\sup_{t \in [0; \beta_n]} \left| \frac{\hat{G}_n^j(t) - G_j(t)}{\sqrt{G_j(t)}} \right| \geq 14 \left(\frac{\log n}{np_j} \right)^{1/2} \right) \\ &= \mathbf{P} \left(\sup_{t \in [0; \beta_n]} \left| \frac{\hat{H}_{N_j}^j(t) - nG_j(t)}{\sqrt{G_j(t)}} \right| \geq 14 \left(\frac{n \log n}{p_j} \right)^{1/2} \right) \\ &\leq \mathbf{P} \left(\sup_{s \in [b_n; 1]} \left| \frac{V_n(s) - p_j s}{\sqrt{s}} \right| \geq 14 \left(\frac{\log n}{n} \right)^{1/2} \right) \end{aligned}$$

avec $b_n = (3 \log n)/(n \log 3)$. On choisit $\alpha = 0.89$ dans le lemme 7 et finalement:

$$\mathbf{P} \leq \left(\frac{4n}{3} \vee 6 \right) \exp(-2.05 \log n) + 2 \exp(-\min(1.14 \log n, 0.08 \sqrt{n \log n}))$$

qui est le terme général d'une série convergente.

2.3 Inégalités pour les processus gaussiens

Nous rappelons tout d'abord l'inégalité de grandes déviations de la loi normale centrée réduite (lemme 10) et l'inégalité classique de Kolmogorov (lemme 11):

Lemme 10 *Soit Z une variable de loi normale centrée réduite. Alors pour tout x positif:*

$$\mathbf{P}(|Z| \geq x) \leq 2 \exp(-x^2/2)$$

Lemme 11 *Pour tout x positif:*

$$\mathbf{P} \left(\sup_{0 \leq t \leq 1} |B^o(t)| \geq x \right) \leq 2 \exp(-2x^2)$$

L'inégalité suivante est due à Bretagnolle et Massart ([3], inégalité 2.4).

Lemme 12 *Pour tout $b \in]0, 1/2]$ et tout x positif:*

$$\mathbf{P} \left(\sup_{t \in [0, b]} |B^o(t)| \geq x \right) \leq 2 \exp(-x^2/(2b(1-b)))$$

3 Etude générale de certains processus construits sur des données censurées

Nous nous plaçons dans le cadre d'un modèle censuré à deux populations (et donc quatre classes disjointes) et nous considérons les statistiques temporelles composées de termes de la forme:

$$S_i^n(t) = \int_0^t \Phi(\hat{G}_1, \hat{G}_2, \hat{G}_3, \hat{G}_4) d\hat{F}_i \quad i = 1, \dots, 4 \quad t \in]0; \infty]$$

où Φ est une fonction à valeurs réelles bornée sur $[0, 1]^4$. Sont concernées en particulier les statistiques du logrank et de Gehan.

L'approximation (à n fixé) du processus S_i^n par un processus gaussien à l'aide des principes d'invariance n'est possible que si le processus S_i^n a été préalablement approché par un processus qui soit fonction linéaire des sous ponts empiriques $\hat{F}_l - F_l$; $l = 1, \dots, 4$ et $\hat{G}_k - G_k$; $k = 1, \dots, 4$. Il s'agit donc de trouver un tel processus τ_i^n et de contrôler en probabilité la déviation uniforme $\|S_i^n - \tau_i^n\|_\infty$.

Nous devons également envisager le cas où la fonction Φ dépend d'un paramètre $\theta > 0$:

$$S_i^n(t, \theta) = \int_0^t \Phi_\theta(\hat{G}_1, \hat{G}_2, \hat{G}_3, \hat{G}_4) d\hat{F}_i \quad i = 1, \dots, 4 \quad t \in]0; \infty]$$

où Φ_θ est une fonction à valeurs réelles bornée sur $[0, 1]^4$ et dépendant continûment du paramètre θ .

En particulier, les dérivées premières et secondes du logarithme de la vraisemblance de Cox sont fonction d'un paramètre. Pour étudier l'estimateur du maximum de vraisemblance de Cox (section 5) nous devons donc majorer la déviation uniforme en θ :

$$\sup_{\theta \in C} |S_i^n(\infty, \theta) - \tau_i^n(\infty, \theta)|$$

où C est un compact de $\mathbf{R}^+$.

3.1 Etude du processus S_i^n

On ne traite que le cas $\hat{F}_i = \hat{F}_1$ car les conditions sur les $\hat{F}_i$ sont symétriques. On pose:

$$S^n(t) = \int_0^t \Phi(\hat{G}_1, \hat{G}_2, \hat{G}_3, \hat{G}_4) d\hat{F}_1$$

On note:

$$\begin{aligned} \hat{H} &= \Phi(\hat{G}_1, \hat{G}_2, \hat{G}_3, \hat{G}_4) & H &= \Phi(G_1, G_2, G_3, G_4) \\ \hat{G} &= \hat{G}_1 + \hat{G}_2 + \hat{G}_3 + \hat{G}_4 & G &= G_1 + G_2 + G_3 + G_4 \end{aligned}$$

(G et $\hat{G}$ sont les vraies queue et queue empirique)

Dans toute la suite, on note:

$$\begin{aligned} r_n &= (\log n)/n \\ \omega &= \min(p_1, p_2, p_3, p_4) \quad \text{où } p_j \text{ est défini en (1)} \end{aligned}$$

La statistique S^n se développe en:

$$S^n = S_1 + S_2^n + T_1^n + T_2^n$$

$$\text{avec } \begin{cases} S_1(t) = \int_0^t H \, dF_1 \\ S_2^n(t) = \int_0^t H \, d(\hat{F}_1 - F_1) \\ T_1^n(t) = \int_0^t (\hat{H} - H) \, dF_1 \\ T_2^n(t) = \int_0^t (\hat{H} - H) \, d(\hat{F}_1 - F_1) \end{cases}$$

Dés que Φ est dérivable, on peut linéariser $\hat{H} - H$:

$$\hat{H}(s) - H(s) = \left(\sum_{j=1}^4 \phi_j(s)(\hat{G}_j(s) - G_j(s)) \right) + u_n(s) \quad (2)$$

$$\text{avec } \phi_j(s) = \frac{d\Phi}{dx_j}(G_1(s), \dots, G_4(s))$$

Soient $(\beta_n)_n$ et $(M_n(t))_n$ les deux suites définies par:

$$\begin{aligned} G_1(\beta_n) &= 3p_1 r_n / \log 3 \\ M_n(t) &= \inf(\beta_n, t) \end{aligned} \quad (3)$$

Remarque: L'existence de β_n est assurée pour tout $n \geq 2$ car $3r_n / \log 3 \leq 1$.

En utilisant (2), on obtient:

$$S^n = S_1 + S_2^n + S_3^n + \Delta^n + R_1^n + R_2^n + R_3^n \quad (4)$$

$$\text{avec } \begin{cases} S_1(t) = \int_0^t H \, dF_1 \\ S_2^n(t) = \int_0^t H \, d(\hat{F}_1 - F_1) \\ S_3^n(t) = \sum_{j=1}^4 \int_0^t \phi_j(\hat{G}_j - G_j) \, dF_1 \\ \Delta^n(t) = \sum_{j=1}^4 \int_0^{M_n(t)} \phi_j(\hat{G}_j - G_j) \, d(\hat{F}_1 - F_1) \\ R_1^n(t) = \int_{M_n(t)}^t (\hat{H} - H) \, d(\hat{F}_1 - F_1) \\ R_2^n(t) = \int_0^{M_n(t)} u_n \, d\hat{F}_1 \\ R_3^n(t) = \int_{M_n(t)}^t u_n \, dF_1 \end{cases}$$

Dans la relation (4), S_1 est le terme de centrage, S_2^n et S_3^n sont des termes principaux et Δ^n , R_1^n , R_2^n , R_3^n sont des termes d'erreur. Pour majorer $\| S^n - S_1^n - S_2^n - S_3^n \|_\infty$, on contrôle $\| \Delta^n \|_\infty$ et $\| R_i^n \|_\infty$; $i = 1, 2, 3$. Les termes $|R_2^n|$ et $|R_3^n|$ dépendent de la fonction u_n et se majoreront au cas par cas.

Theoreme 2 Soient les conditions A_1, A_2, A_3 :

- A_1 Φ de classe C^2
 A_2 Les fonctions $\epsilon_1^j = |\phi_j|$ G sont bornées
 A_3 Les fonctions $\epsilon_2^j = |\phi_j'|$ $G^2/|G'|$ sont bornées

On pose $\epsilon_1 = \max_j \epsilon_1^j$ et $\epsilon_2 = \max_j \epsilon_2^j$, on rappelle que $(M_n(t))_n$ est définie en (3) et on pose:

$$I_n(t) = \int_0^{M_n(t)} \frac{\epsilon_2 |G'|}{G} ds$$

Lorsque les conditions A_1, A_2, A_3 sont vérifiées on a:

$$\left\| \frac{R_1^n + \Delta^n}{K_1 + K_2 I_n} \right\|_\infty \leq r_n \quad p.co$$

où K_1 et K_2 sont majorées par $A(\|\epsilon_1\|_\infty + \|\epsilon_2\|_\infty)/\omega$, A étant une constante absolue.

La preuve du théorème 2 se trouve dans la section 3.3.

Remarque 4: De manière immédiate:

$$\begin{aligned} I_n(t) \leq \|\epsilon_2\|_\infty \log \frac{1}{G(M_n(t))} &\leq \|\epsilon_2\|_\infty \min(\log \frac{1}{G(\beta_n)}, \log \frac{1}{G(t)}) \\ &\leq \|\epsilon_2\|_\infty (1 + \log(1/\omega)) \min(\log n, \log \frac{1}{G(t)}) \end{aligned}$$

3.2 Etude de $S_1^n(t, \theta)$ lorsque θ varie dans un compact de $\mathbf{R}^+$

Cette partie est presque identique à la précédente (section 3.1: Etude du processus S_1^n). Nous ne détaillerons que les arguments sur lesquels repose l'uniformité en θ .

Les conditions sur les sous fonctions de répartition empirique étant symétriques, on étudie $S_1^n(t, \theta)$. On pose:

$$S^n(t, \theta) = \int_0^t \Phi_\theta(\hat{G}_1, \hat{G}_2, \hat{G}_3, \hat{G}_4) d\hat{F}_1$$

Les notations sont les mêmes que dans la section 3.1 mais l'on rajoute un paramètre θ :

$$\begin{aligned} \hat{H}(s, \theta) &= \Phi_\theta(\hat{G}_1, \hat{G}_2, \hat{G}_3, \hat{G}_4) \\ H(s, \theta) &= \Phi_\theta(G_1, G_2, G_3, G_4) \end{aligned}$$

La linéarisation (2) devient:

$$\hat{H}(s, \theta) - H(s, \theta) = \left(\sum_{j=1}^4 \phi_j(s, \theta) (\hat{G}_j(s) - G_j(s)) \right) + u_n(s, \theta)$$

Le même développement qu'en (4) entraîne:

$$S^n(t, \theta) = S_1(t, \theta) + S_2^n(t, \theta) + S_3^n(t, \theta) + \Delta^n(t, \theta) \\ + R_1^n(\infty, \theta) + R_2^n(\infty, \theta) + R_3^n(\infty, \theta)$$

où les différents termes sont les mêmes qu'en (4) après avoir remplacé $H(s)$, $\hat{H}(s)$, $\phi_j(s)$ et $u_n(s)$ par $H(s, \theta)$, $\hat{H}(s, \theta)$, $\phi_j(s, \theta)$ et $u_n(s, \theta)$.

Pour contrôler la déviation:

$$\sup_{\theta \in C} \sup_{t > 0} |S^n(t, \theta) - S_1^n(t, \theta) - S_2^n(t, \theta) - S_3^n(t, \theta)|$$

où C est un compact de $\mathbf{R}^+$, il faut majorer les termes d'erreur $\sup_{\theta \in C} \sup_{t > 0} |\Delta^n(\infty, \theta)|$ et $\sup_{\theta \in C} \sup_{t > 0} |R_i^n(\infty, \theta)|$; $i = 1, 2, 3$. Les termes $|R_2^n|$ et $|R_3^n|$ dépendent de la fonction u_n , leur majoration est différente pour chaque application.

Theoreme 3 Soient les conditions A'_1 , A'_2 , A'_3 , A'_4 :

A'_1 Φ de classe C^2

A'_2 Pour tout compact C de $\mathbf{R}^+$ les fonctions $\epsilon_1^j(s, \theta) = |\phi_j(s, \theta)| G(s)$ vérifient:
 $\sup_{\theta \in C} \sup_{s > 0} \epsilon_1^j(s, \theta) \leq K(C)$
 où $K(C)$ est une constante qui dépend du compact C .

A'_3 Pour tout compact C de $\mathbf{R}^+$ les fonctions $\epsilon_2^j(s, \theta) = |[(d\phi_j)/(ds)](s, \theta) \times G^2/G'|$ vérifient:
 $\sup_{\theta \in C} \sup_{s > 0} \epsilon_2^j(s, \theta) \leq K(C)$
 où $K(C)$ est une constante qui dépend du compact C

A'_4 Pour tout compact C de $\mathbf{R}^+$:
 $\sup_{\theta \in C} \sup_{s > 0} |[(d\phi_j)/(d\theta)](s, \theta) \times G(s)| \leq A(C)$
 où $A(C)$ est une constante qui dépend du compact C

On note:

$$\epsilon_1(s, \theta) = \max_j \epsilon_1^j(s, \theta) \quad \epsilon_2(s, \theta) = \max_j \epsilon_2^j(s, \theta)$$

Pour tout compact C de $\mathbf{R}^+$, on pose:

$$\|\epsilon_1(C)\|_\infty = \sup_{\theta \in C} \sup_{s > 0} \epsilon_1(s, \theta)$$

$$\|\epsilon_2(C)\|_\infty = \sup_{\theta \in C} \sup_{s > 0} \epsilon_2(s, \theta)$$

Lorsque les conditions A'_1 , A'_2 , A'_3 , A'_4 sont vérifiées, on a:

$$\sup_{\theta \in C} \sup_{t > 0} \left| \frac{R_1^n(t, \theta) + \Delta^n(t, \theta)}{K_1(C) + K_2(C) \inf(1 + \log(1/G(t)), \log n)} \right| \leq \frac{\log n}{n} \quad p.co$$

où $K_1(C)$ et $K_2(C)$ sont majorées par

$$\frac{A(C)}{\|\epsilon_1(C)\|_\infty} \times B \times (\|\epsilon_1(C)\|_\infty + \|\epsilon_2(C)\|_\infty) \frac{\log \omega^{-1}}{\omega}$$

$A(C)$ étant la constante de la condition A'_4 et B une constante absolue.

La preuve du théorème 3 se trouve dans la section 3.3.

Commentaire: Comme il sera clair dans la preuve des théorèmes 2 et 3 seules sont nécessaires les hypothèses S_1, S_2, C_1, C_2 dérivables sur l'intérieur du support des lois et l'hypothèse de non dépendance du temps de la covariable δ . De plus, ces théorèmes sont valides quelque soit le nombre de classes disjointes considérées, ce qui permet une généralisation à un nombre quelconque de populations, ou encore de retrouver les résultats pour des modèles non censurés.

3.3 Preuve des théorèmes 2 et 3

Pour la suite $(M_n(t))_n$ définie en (3) les lemmes 8 et 9 entraînent:

$$\sup_{t>0} \sup_{s \in [0; M_n(t)]} \left| \frac{\hat{G}_1(s) - G_1(s)}{\sqrt{G_1(s)}} \right| \leq \frac{14\sqrt{r_n}}{\sqrt{\omega}} \quad \text{p.co} \quad (5)$$

$$\sup_{t>0} \sup_{s \in [0; M_n(t)]} \left| \frac{\hat{G}_1(s) - G_1(s)}{G_1(s)} \right| \leq \frac{9}{\omega} \quad \text{p.co} \quad (6)$$

3.3.1 Preuve du théorème 2

Etude du terme R_1^n :

$$\forall t > 0 \quad |R_1^n(t)| \leq \begin{cases} 0 & \text{si } t \leq \beta_n \\ 2 \max(\|\Phi\|_\infty) \int_{\beta_n}^\infty d(\hat{F}_1 + F_1) & \text{si } t > \beta_n \end{cases}$$

On a:

$$\int_{\beta_n}^\infty d(\hat{F}_1 + F_1) \leq |\hat{G}_1(\beta_n) - G_1(\beta_n)| + 2G_1(\beta_n)$$

Le lemme 2 entraîne:

$$\begin{aligned} P &= P(|\hat{G}_1(\beta_n) - G_1(\beta_n)| \geq 3r_n) \\ &\leq 2 \exp(-nG_1(\beta_n)h(3r_n/G_1(\beta_n))) \end{aligned}$$

Comme $3r_n/G_1(\beta_n) \geq \log 3$, avec la remarque 3:

$$\begin{aligned} P &\leq 2 \exp(-nG_1(\beta_n) \frac{h(\log 3)}{\log 3} \frac{3r_n}{G_1(\beta_n)}) \\ &\leq 2 \exp(-1.24 \log n) \end{aligned}$$

Etude du terme Δ^n : On sépare les cas $j \neq 1$ (terme croisé) et $j = 1$ (terme diagonal).

Etude du terme croisé: On prend par exemple $j = 2$ et on note:

$$\Delta C^n(t) = \int_0^{M_n(t)} \phi_2(\hat{G}_2 - G_2) d(\hat{F}_1 - F_1)$$

D'après le théorème 1:

$$\| \hat{G}_2 - G_2 - \frac{\sqrt{np_2}(B_2^o(\mu_2))}{\sqrt{n}} - \sqrt{\frac{q_2}{p_2}} G_2 \frac{Z_2}{\sqrt{n}} \|_{\infty} \leq 98r_n \quad \text{p.co}$$

D'après la remarque 1, B_2 est indépendant de Z_2 et de $\hat{F}_1 - F_1$. On obtient:

$$\| \Delta C \|_{\infty} \leq \| \Delta C_1^n \|_{\infty} + \| \Delta C_2^n \|_{\infty} + \| \Delta C_3^n \|_{\infty} \quad \text{p.co} \quad (7)$$

$$\text{avec } \begin{cases} |\Delta C_1^n(t)| = 98r_n \int_0^{M_n(t)} |\phi_2| d(\hat{F}_1 + F_1) \\ |\Delta C_2^n(t)| = \sqrt{q_2/p_2} (|Z_2|/\sqrt{n}) \left| \int_0^{M_n(t)} G_2 \phi_2 d(\hat{F}_1 - F_1) \right| \\ |\Delta C_3^n(t)| = \sqrt{p_2} \left| \int_0^{M_n(t)} (\phi_2(B_2^o(\mu_2)))/\sqrt{n} d(\hat{F}_1 - F_1) \right| \end{cases}$$

Contrôle de ΔC_1 :

Pour tout $t > 0$, après une intégration par parties, on a:

$$|\Delta C_1^n(t)| \leq 98r_n \left\{ \left| \int_0^{M_n(t)} \phi_2 (\hat{G}_1 + G_1) ds \right| + \int_0^{M_n(t)} (\hat{G}_1 + G_1) |\phi_2'| ds \right\}$$

En utilisant $\hat{G}_1 + G_1 = \hat{G}_1 - G_1 + 2G_1$, on obtient:

$$\forall t > 0 \quad |\Delta C_1^n(t)| \leq 98r_n (2 \| \epsilon_1 \|_{\infty} + I_n(t)) \left(\sup_{t>0} \sup_{s \in [0; M_n(t)]} \left| \frac{\hat{G}_1(s) - G_1(s)}{G_1(s)} \right| + 2 \right)$$

Finalement, avec l'inégalité (6):

$$\left\| \frac{\Delta C_1^n}{K(2 \| \epsilon_1 \|_{\infty} + I_n)} \right\|_{\infty} \leq r_n \quad \text{p.co} \quad (8)$$

avec $K = 980/\omega$ qui convient.

Contrôle de ΔC_2 :

Posons:

$$J^n(t) = \left| \int_0^{M_n(t)} G_2 \phi_2 d(\hat{F}_1 - F_1) \right|$$

Après une intégration par parties, pour tout $t > 0$, on a:

$$J^n(t) \leq \underbrace{\left| \epsilon_1 (\hat{G}_1 - G_1) \right|_0^{M_n(t)}}_{J_1^n(t)} + \underbrace{\left| \int_0^{M_n(t)} (\hat{G}_1 - G_1) (G_2' \phi_2 + G_2 \phi_2') ds \right|}_{J_2^n(t)}$$

Le lemme 8 entraine:

$$\| J_1^n(t) \|_\infty \leq 6 \| \epsilon_1 \|_\infty \sqrt{r_n} \quad \text{p.co}$$

En utilisant l'inégalité (5), on obtient:

$$\begin{aligned} \| J_2^n \|_\infty &\leq (\| \epsilon_1 \|_\infty + \| \epsilon_2 \|_\infty) \left(\left\| \int_0^t \frac{|G'|}{\sqrt{G}} ds \right\|_\infty \right) \sup_{t>0} \sup_{s \in [0; M_n(t)]} \left| \frac{\hat{G}_1(s) - G_1(s)}{\sqrt{G_1(s)}} \right| \\ &\leq K \sqrt{r_n} \quad \text{p.co} \end{aligned}$$

avec $K = 14(\| \epsilon_1 \|_\infty + \| \epsilon_2 \|_\infty)/\sqrt{\omega}$ qui convient.

D'autre part, avec la transformée de Cramer de la gaussienne centrée réduite (lemme 10):

$$\mathbb{P}\left(\left| \frac{Z_2}{\sqrt{n}} \right| \geq 2\sqrt{r_n}\right) \leq 2 \exp(-2 \log n)$$

on obtient $\sqrt{q_2/p_2} |Z_2| / \sqrt{n} \leq 2\sqrt{r_n/\omega}$ p.co. Finalement:

$$\| \Delta C_2^n \|_\infty \leq K r_n \quad \text{p.co} \quad (9)$$

avec $K = (37 \| \epsilon_1 \|_\infty + 28 \| \epsilon_2 \|_\infty)/\omega$ qui convient.

Contrôle de ΔC_3^n :

Par l'égalité en loi $\mathcal{L}(B_2^o(\mu_2)) = \mathcal{L}(B_2^o(1 - \mu_2))$ et l'indépendance de B_2 et de $\hat{F}_1$, on a $\mathcal{L}(\Delta C_3^n) = \mathcal{L}(\Delta C_3^{*n})$ où:

$$\Delta C_3^{*n}(t) = \int_0^{M_n(t)} \sqrt{p_2} \phi_2 \frac{(B_2^o(1 - \mu_2))}{\sqrt{n}} d(\hat{F}_1 - F_1)$$

En décomposant $B_2^o(1 - \mu_2(t)) = B_2(1 - \mu_2(t)) - (1 - \mu_2(t))B_2(1)$, on obtient:

$$\| \Delta C_3^{*n} \|_\infty \leq \| \Delta \tilde{C}_3^n \|_\infty + \| \Delta \tilde{\tilde{C}}_3^n \|_\infty$$

$$\text{avec } \begin{cases} \Delta \tilde{C}_3^n(t) = (B_2(1)/\sqrt{np_2}) \int_0^{M_n(t)} \phi_2 G_2 d(\hat{F}_1 - F_1) \\ \Delta \tilde{\tilde{C}}_3^n(t) = [\int_0^{M_n(t)} \sqrt{p_2} \phi_2 (B_2(1 - \mu_2)) d(\hat{F}_1 - F_1)]/\sqrt{n} \end{cases}$$

Le terme $\Delta \tilde{\tilde{C}}_3^n$ est du même type que ΔC_2^n , donc:

$$\| \Delta \tilde{\tilde{C}}_3^n \|_\infty \leq K r_n \quad \text{p.co} \quad (10)$$

avec $K = (37 \| \epsilon_1 \|_\infty + 28 \| \epsilon_2 \|_\infty)/\omega$ qui convient.

Conditionnellement à $\hat{F}_1$, $\sqrt{n} \Delta \tilde{C}_3^n$ est un processus gaussien centré et de variance:

$$\sigma_n^2(t) = 2 \int_0^{M_n(t)} G_2(s) V_s dV_s - \int_0^{M_n(t)} G_2(s) d(\sum_{0 < u \leq s} (\Delta V_u)^2)$$

$$\text{avec } V_s = \int_0^s \phi_2 d(\hat{F}_1 - F_1)$$

et où $\Delta V_u = V_u - V_{u-}$ désigne le saut du processus V au point u .

Par la formule d'intégration:

$$(V_s)^2 = 2 \int_0^s V_u dV_u - \sum_{0 < u \leq s} (\Delta V_u)^2$$

et une intégration par parties, on obtient:

$$\| \sigma_n^2 \|_\infty \leq \| \bar{\sigma}_n^2(t) \|_\infty + \| \bar{\bar{\sigma}}_n^2 \|_\infty \quad (11)$$

$$\text{avec } \begin{cases} \bar{\sigma}_n^2(t) = | [G_2(s) V_s^2]_0^{M_n(t)} | \\ \bar{\bar{\sigma}}_n^2(t) = | \int_0^{M_n(t)} V_s^2 d(G_2(s)) | \end{cases}$$

Pour évaluer ces deux termes, on étudie V_s^2 . Posons:

$$J(s) = \int_0^s \frac{\epsilon_2 |G'|}{G^{3/2}} du$$

Pour tout $s \in [0; M_n(t)]$, une intégration par parties donne:

$$\begin{aligned} |V_s| &\leq |[\phi_2(\hat{G}_1 - G_1)]_0^s| + \left| \int_0^s (\hat{G}_1 - G_1) \phi_2' du \right| \\ &\leq \left(\sup_{s \in [0; M_n(t)]} \sup_{u \in [0; s]} \left| \frac{\hat{G}_1(u) - G_1(u)}{\sqrt{G_1(u)}} \right| \right) (2 \| \epsilon_1 \|_\infty^{1/2} \sqrt{\phi_2(s)} + J(s)) \end{aligned}$$

Avec l'inégalité (5):

$$\forall s \in [0; M_n(t)] \quad |V_s| \leq \frac{14\sqrt{r_n}}{\sqrt{\omega}} (K \sqrt{\phi_2(s)} + J(s)) \quad \text{p.co}$$

avec $K = 2 \| \epsilon_1 \|_\infty^{1/2}$ qui convient.

On remarque:

$$J(s) \leq 2 \| \epsilon_2 \|_\infty / \sqrt{G(s)} \quad (12)$$

On obtient:

$$\| \bar{\sigma}_n^2 \|_\infty \leq K r_n \quad \text{p.co} \quad (13)$$

avec $K = 4 \times 14^2 (\| \epsilon_1 \|_\infty + \| \epsilon_2 \|_\infty)^2 / \omega$ qui convient.

En utilisant $(a + b)^2 \leq 2(a^2 + b^2)$, on a:

$$\forall t > 0 \quad \bar{\bar{\sigma}}_n^2(t) \leq r_n \left[K_1 \int_0^{M_n(t)} |\phi_2| d(-G_2) + K_2 \int_0^{M_n(t)} J^2 d(-G_2) \right] \quad \text{p.co}$$

avec $K_1 = 8 \times 14^2 \| \epsilon_1 \|_\infty / \omega$ et $K_2 = 2 \times 14^2 / \omega$ qui conviennent.

On obtient facilement:

$$\int_0^{M_n(t)} |\phi_2| \, d(-G_2) \leq 2 \|\epsilon_1\|_\infty + I_n(t)$$

Une intégration par parties et la relation (12) donnent:

$$\int_0^{M_n(t)} J^2 \, d(-G_2) \leq 4 \|\epsilon_2\|_\infty^2 + 2 \int_0^{M_n(t)} G_2 |J| \, dJ$$

En remarquant que pour tout $s \in [0; M_n(t)]$:

$$J(s) \leq I_n(s)/\sqrt{G(s)}$$

on obtient:

$$\int_0^{M_n(t)} G_2 |J| \, dJ \leq (I_n(t))^2$$

Finalement,

$$\forall t > 0 \quad \bar{\sigma}_n^2(t) \leq r_n \left(K_3 + K_1 I_n(t) + K_4 (I_n(t))^2 \right) \quad \text{p.co} \quad (14)$$

avec $K_3 = 2K_1 \|\epsilon_1\|_\infty + 4K_2 \|\epsilon_2\|_\infty^2$ et $K_4 = 2K_2$ qui conviennent.

En réunissant (11), (13), (14) et en utilisant $|ab| \leq (a^2 + b^2)/2$ on a:

$$\left\| \frac{\sigma_n^2}{K' + K''(I_n)^2} \right\|_\infty \leq r_n \quad \text{p.co} \quad (15)$$

avec $\begin{cases} K' = 14^2(28 \|\epsilon_1\|_\infty^2 + 16 \|\epsilon_2\|_\infty^2)/\omega \\ K'' = 8 \times 14^2/\omega \end{cases}$ qui conviennent.

On pose $\lambda_n(t) = \sqrt{n} r_n (K' + K''(I_n(t))^2)^{1/2}$ et $\rho_n(t) = \lambda_n(t)/\sqrt{\log n}$. Pour contrôler $\Delta \tilde{C}_3^n$, il suffit de montrer que la probabilité $\mathbf{P}^n$, avec:

$$\mathbf{P}^n = \mathbf{P}(\sqrt{n} \left\| \frac{\Delta \tilde{C}_3^n}{\lambda_n} \right\|_\infty \geq \frac{1.76}{\sqrt{\omega}})$$

est majorée par le terme général d'une série convergente. On a:

$$\mathbf{P}^n \leq \mathbf{P}_1^n + \mathbf{P}_2^n \quad (16)$$

avec $\begin{cases} \mathbf{P}_1^n = \mathbf{P}(\|\sigma_n/\rho_n\|_\infty \geq 1) \\ \mathbf{P}_2^n = \mathbf{P}(\|\sqrt{n} \Delta \tilde{C}_3^n / \lambda_n\|_\infty \geq 1.76/\sqrt{\omega}, \|\sigma_n/\rho_n\|_\infty \leq 1) \end{cases}$

D'après (15), $\mathbf{P}_1^n$ est majorée par le terme général d'une série convergente:

$$\mathbf{P}_1^n \leq U_n \quad \text{avec} \quad \sum_n U_n < \infty \quad (17)$$

Nous étudions P_2^n en deux étapes: la première pour montrer l'inégalité à t fixé et la seconde pour contrôler les oscillations et obtenir l'uniformité en t . Il suffit de montrer que P_2^n est majorée par le terme général d'une série convergente. La majoration obtenue sera de nouveau utilisée dans la preuve du théorème 3 et afin de ne pas reprendre deux fois les calculs, nous montrons le résultat plus fort:

$$P_2^n \leq V_n \quad \text{avec} \quad \sum_n \sqrt{n} V_n < \infty$$

1ere étape Contrôle à t fixé.

Nous posons:

$$\pi_n(t) = P^{\hat{F}_1} \left(\left| \frac{\sqrt{n} \Delta \tilde{C}_3^n(t)}{\lambda_n(t)} \right| \geq 2.3, \left\| \frac{\sigma_n}{\rho_n} \right\|_\infty \leq 1 \right)$$

où $P^{\hat{F}_1}$ désigne la probabilité conditionnelle à $\hat{F}_1$.

Conditionnellement à $\hat{F}_1$, $\sqrt{n} \Delta \tilde{C}_3^n(t) / \sigma_n(t)$ est une gaussienne centrée réduite, dont on connaît la transformée de Cramer (lemme 10):

$$\begin{aligned} \pi_n(t) &\leq P^{\hat{F}_1} \left(\left| \frac{\sqrt{n} \Delta \tilde{C}_3^n(t)}{\sigma_n(t)} \right| \geq \frac{2.3 \lambda_n(t)}{\rho_n(t)} \right) \\ &\leq 2 \exp(-2.6 \log n) \end{aligned}$$

Le résultat ne dépendant pas de la valeur prise par $\hat{F}_1$, on a:

$$P \left(\left| \frac{\sqrt{n} \Delta \tilde{C}_3^n}{\lambda_n} \right| \geq 2.3, \left\| \frac{\sigma_n}{\rho_n} \right\|_\infty \leq 1 \right) \leq \frac{2}{n^{2.6}} \quad (18)$$

2eme étape Uniformité en t

Pour tout $t > \beta_n$, on a:

$$\frac{\Delta \tilde{C}_3^n(t)}{\lambda_n(t)} = \frac{\Delta \tilde{C}_3^n(\beta_n)}{\lambda_n(\beta_n)}$$

La probabilité à contrôler est majorée:

$$P_2^n \leq P \left(\sup_{0 < t \leq \beta_n} \left| \frac{\sqrt{n} \Delta \tilde{C}_3^n(t)}{\lambda_n(t)} \right| \geq 2.3 + \frac{0.133}{\sqrt{\omega}}, \left\| \frac{\sigma_n}{\rho_n} \right\|_\infty \leq 1 \right)$$

Soit $\tilde{\beta}_n$ tel que $F(\tilde{\beta}_n)$ soit le plus petit multiple de r_n qui soit inférieur ou égal à $F_1(\beta_n)$ ($\tilde{\beta}$ peut être égal à zéro). On découpe l'intervalle $]0, \beta_n]$ en N intervalles disjoints $I_1, \dots, I_N$ tels que:

$$\begin{aligned} I_j &=]t_{j-1}, t_j] \\ t_0 &= 0 \quad t_{N-1} = \tilde{\beta}_n \quad t_N = \beta_n \\ F(I_j) &= r_n \quad \text{pour } j = 1, \dots, N-1 \end{aligned}$$

On a $F(I_j) \leq r_n$ pour $j = 1, \dots, N$ et $N \leq 1/r_n < n$.

En remarquant de plus:

$$\lambda_n(t) > \sqrt{K'} \sqrt{n} r_n \geq 74 \|\epsilon_1\|_\infty \sqrt{n} r_n / \sqrt{\omega}$$

où K' était définie en (15), on obtient:

$$\begin{aligned} P_2^n &\leq \mathbf{P}\left(\sup_{0 < t \leq \beta_n} \left| \frac{\Delta \tilde{C}_3^n(t)}{\lambda_n(t)} \right| \geq 2.3 + \frac{0.133}{\sqrt{\omega}}, \|\frac{\sigma_n}{\rho_n}\|_\infty \leq 1\right) \\ &\leq n \left\{ \max_{t_j} P(t_j) + \max_{I_j} Q(I_j) \right\} \end{aligned} \quad (19)$$

$$\text{avec } \begin{cases} P(t_j) = \mathbf{P}(|\sqrt{n} \Delta \tilde{C}_3^n(t_j)| \geq 2.3 \lambda_n(t_j), \|\sigma_n / \rho_n\|_\infty \leq 1) \\ Q(I_j) = \mathbf{P}(\sup_{t \in I_j} |\Delta \tilde{C}_3^n(t_j) - \Delta \tilde{C}_3^n(t)| \geq 9.842 \|\epsilon_1\|_\infty r_n / \omega) \end{cases}$$

Le terme $P(t_j)$ est majoré en (18):

$$P(t_j) \leq 2/n^{2.6} \quad (20)$$

De manière immédiate,

$$\sup_{t \in I_j} |\Delta \tilde{C}_3^n(t_j) - \Delta \tilde{C}_3^n(t)| \leq \sup_{t \in I_j} \left| \frac{\sqrt{p_2} \phi_2(t) (B_2(1 - \mu_2(t)))}{\sqrt{n}} \right| \times (\hat{F} + F)(I_j)$$

En utilisant $\mathcal{L}(\sqrt{p_2} B_2(1 - \mu_2)) = \mathcal{L}(B_2(G_2))$ on obtient:

$$Q(I_j) \leq Q_1(I_j) + Q_2(I_j) \quad (21)$$

$$\text{avec } \begin{cases} Q_1(I_j) = \mathbf{P}(\|\epsilon_1\|_\infty \sup_{t \in I_j} |B_2(G_2(t)) / (\sqrt{n} G(t))| \geq \|\epsilon_1\|_\infty 1.929 / \omega) \\ Q_2(I_j) = \mathbf{P}((\hat{F} + F)(I_j) \geq 5.1 r_n) \end{cases}$$

Le lemme 2 et la remarque 3 entraînent:

$$\begin{aligned} Q_2(I_j) &= \mathbf{P}((\hat{F} - F)(I_j) \geq 3.1 r_n) \\ &\leq \exp(-n F(I_j) h(\frac{3.1 r_n}{F(I_j)})) \\ &\leq \exp(-n r_n h(3.1)) \leq n^{-2.6} \end{aligned} \quad (22)$$

D'autre part on a:

$$\sup_{t \in I_j} \left| \frac{B_2(G_2(t))}{\sqrt{n} G(t)} \right| \leq \sup_{t \in I_j} \left| \frac{B_2(G_2(t)) - B_2(G_2(t_j))}{\sqrt{n} (3 p_1 r_n / \log 3)} \right| + \left| \frac{B_2(G_2(t_j))}{\sqrt{n} G_2(t_j)} \right|$$

La loi des accroissements du mouvement brownien sur un intervalle ne dépend que de la taille de cet intervalle. Comme la variation de G_2 sur tout intervalle est inférieure à celle de G on a:

$$Q_1(I_j) \leq R_1 + R_2 + R_3$$

$$\text{avec } \begin{cases} R_1 = \mathbf{P}(\sup_{t \in [0, r_n]} |B_2^\circ(t)| \geq 1.929\alpha_1(3/\log 3)((\log n)/\sqrt{n})) \\ R_2 = \mathbf{P}(|B_2(1)| \geq 1.929\alpha_2(3/\log 3)\sqrt{n}) \\ R_3 = \mathbf{P}(|B_2(1)| \geq 1.929\alpha_3\sqrt{3/(\omega \log 3)}\sqrt{\log n}) \end{cases}$$

où $\alpha_1 = 0.433$, $\alpha_2 = 0.061$, $\alpha_3 = 0.506$. Les lemmes 10 et 12 entraînent:

$$Q_1(I_j) \leq 4n^{-2.6} + 2 \exp(-n/20) \quad (23)$$

Les relations (21), (22), (23) entraînent:

$$Q(I_j) \leq 5n^{-2.6} + 2 \exp(-n/20) \quad (24)$$

Les relations (19), (20), (24) entraînent:

$$P_2^n \leq V_n \quad \text{avec} \quad \sum_n \sqrt{n} V_n < \infty \quad (25)$$

Les relations (16), (17), (25) entraînent:

$$\left\| \frac{\sqrt{n} \Delta \tilde{C}_3^n}{\lambda_n} \right\|_\infty \leq \frac{1.76}{\sqrt{\omega}} \quad \text{p.co}$$

En utilisant $\sqrt{a+b} \leq \sqrt{a} + \sqrt{b}$ on obtient:

$$\left\| \frac{\Delta \tilde{C}_3^n}{K' + K'' I_n} \right\|_\infty \leq r_n \quad \text{p.co} \quad (26)$$

$$\text{avec } \begin{cases} K' = (131 \|\epsilon_1\|_\infty + 99 \|\epsilon_2\|_\infty)/\omega \\ K'' = 70/\omega \end{cases} \quad \text{qui conviennent.}$$

En réunissant (10) et (26):

$$\left\| \frac{\Delta C_3^n}{K''' + K'' I_n} \right\|_\infty \quad \text{p.co} \quad (27)$$

avec $K''' = (168 \|\epsilon_1\|_\infty + 127 \|\epsilon_2\|_\infty)/\omega$ qui convient.

En réunissant (7), (8), (9), (27):

$$\left\| \frac{\Delta C^n}{K_1 + K_2 I_n} \right\|_\infty \leq r_n \quad \text{p.co} \quad (28)$$

avec $\begin{cases} K_1 = (2165 \|\epsilon_1\|_\infty + 155 \|\epsilon_2\|_\infty)/\omega \\ K_2 = 1050/\omega \end{cases}$ qui conviennent.

Etude du terme diagonal: On note:

$$\Delta D^n(t) = \int_0^{M_n(t)} \phi_1(\hat{G}_1 - G_1) d(\hat{F}_1 - F_1)$$

Par la formule d'intégration:

$$(V_s)^2 = 2 \int_0^s V_{u-} dV_u + \sum_{0 < u \leq s} (\Delta V_s)^2$$

et une intégration par parties, on obtient:

$$\|\Delta D^n\|_\infty \leq \|\Delta D_1^n\|_\infty + \|\Delta D_2^n\|_\infty + \|\Delta D_3^n\|_\infty \quad (29)$$

$$\text{avec } \begin{cases} |\Delta D_1^n(t)| = \int_0^{M_n(t)} |\phi_1| d(\sum_{0 < u \leq s} (\Delta \hat{G}_1(u))^2) \\ |\Delta D_2^n(t)| = |[(\hat{G}_1 - G_1)^2 \phi_1]_0^{M_n(t)}| \\ |\Delta D_3^n(t)| = \int_0^{M_n(t)} (\hat{G}_1 - G_1)^2 d\phi_1 \end{cases}$$

Le calcul des sauts donne:

$$\sum_{0 < u \leq s} (\Delta \hat{G}_1(u))^2 = \hat{F}_1(s)/n$$

On obtient:

$$\Delta D_1^n(t) = \frac{1}{n} \int_0^{M_n(t)} |\phi_1| d\hat{F}_1$$

ΔD_1^n est du même type que ΔC_1^n puisque ϕ_1 vérifie les mêmes hypothèses que ϕ_2 . Donc:

$$\left\| \frac{\Delta D_1^n}{K(2 \|\epsilon_1\|_\infty + I_n)} \right\|_\infty \leq r_n \quad \text{p.co} \quad (30)$$

avec $K = 14/\omega$ qui convient.

Avec l'inégalité (5) on obtient facilement:

$$\|\Delta D_2^n\|_\infty \leq 2K \|\epsilon_1\|_\infty r_n \quad \text{p.co} \quad (31)$$

$$\left\| \frac{\Delta D_3^n}{I_n} \right\|_\infty \leq K r_n \quad \text{p.co}$$

avec $K = 14^2/\omega$ qui convient.

En réunissant (29), (30), (31):

$$\left\| \frac{\Delta D^n}{K_1 + K_2 I_n} \right\|_\infty \leq r_n \quad \text{p.co} \quad (32)$$

avec $K_1 = 420 \|\epsilon_1\|_\infty / \omega$ et $K_2 = 210/\omega$ qui conviennent.

3.3.2 Preuve du théorème 3

Cette preuve est essentiellement la même que celle du théorème 2. Nous ne reprenons que le passage où les hypothèses A_1' , A_2' , A_3' ne permettent pas une généralisation immédiate. (preuve du théorème 2 - Etude de Δ^n - Etude du terme croisé- Contrôle de ΔC_3^n)

Nous adaptons la preuve du théorème 2 et nous considérons le processus:

$$\Delta \tilde{C}_3^n(t, \theta) = [\int_0^{M_n(t)} \sqrt{p_2} \phi_2(s, \theta) (B_2(1 - \mu_2(s))) d(\hat{F}_1 - F_1)(s)] / \sqrt{n}$$

Conditionnellement à $\hat{F}_1$, $(\sqrt{n} \Delta \tilde{C}_3^n(t, \theta))_{(t, \theta)}$ est un processus gaussien centré dont la variance $\sigma_n^2(t, \theta)$ vérifie une relation analogue à la relation (15):

$$\sup_{\theta \in \mathcal{C}} \sup_{t > 0} \frac{\sigma_n^2(t, \theta)}{K'(\mathcal{C}) + K''(\mathcal{C})(\inf(1 + \log(1/G(t)), \log n))^2} \leq r_n \quad \text{p.co} \quad (33)$$

$$\text{avec } \begin{cases} K'(\mathcal{C}) = 14^2 (28 \|\epsilon_1(\mathcal{C})\|_\infty^2 + 16 \|\epsilon_2(\mathcal{C})\|_\infty^2) / \omega \\ K''(\mathcal{C}) = 8 \times 14^2 \|\epsilon_2(\mathcal{C})\|_\infty^2 \times (1 + \log \omega^{-1})^2 / \omega \end{cases} \text{ qui conviennent.}$$

Notons:

$$\begin{aligned} \rho_n(t, \mathcal{C}) &= \sqrt{r_n} [K'(\mathcal{C}) + K''(\mathcal{C})(\inf(1 + \log(1/G(t)), \log n))^2]^{1/2} \\ \lambda_n(t, \mathcal{C}) &= \sqrt{\log n} \rho_n(t, \mathcal{C}) \end{aligned}$$

Afin d'obtenir une majoration de

$$\sup_{\theta \in \mathcal{C}} \sup_{t > 0} |\sqrt{n} \Delta \tilde{C}_3^n(t, \theta) / \lambda_n(t, \mathcal{C})|$$

nous contrôlons la probabilité:

$$P^n = \mathbf{P}(\sup_{\theta \in \mathcal{C}} \sup_{t > 0} |\frac{\sqrt{n} \Delta \tilde{C}_3^n(t, \theta)}{\lambda_n(t, \mathcal{C})}| \geq \frac{1.76 + B(\mathcal{C})}{\sqrt{\omega}})$$

où $B(\mathcal{C})$ est une constante qui sera identifiée par la suite. De manière immédiate:

$$P^n \leq P_1^n + P_2^n \quad (34)$$

$$\text{avec } \begin{cases} P_1^n = \mathbf{P}(\sup_{\theta \in \mathcal{C}} \sup_{t > 0} \sigma_n(t, \theta) / \rho_n(t, \mathcal{C}) \geq 1) \\ P_2^n = \mathbf{P}(\sup_{\theta \in \mathcal{C}} \sup_{t > 0} |\sqrt{n} \sqrt{n} \Delta \tilde{C}_3^n(\infty, \theta) / \lambda_n(t, \mathcal{C})| \geq (1.76 + B(\mathcal{C})) / \sqrt{\omega} \\ \quad \sup_{\theta \in \mathcal{C}} \sup_{t > 0} \sigma_n(t, \theta) / \rho_n(t, \mathcal{C}) \leq 1) \end{cases}$$

D'après (33), P_1^n est majorée par le terme général d'une série convergente:

$$P_1^n \leq U_n \quad \text{avec } \sum_n U_n < \infty \quad (35)$$

Pour contrôler P_2^n , nous discrétisons le paramètre θ sur la grille: $\theta_1 < \theta_2 < \dots < \theta_N$ avec $\theta_{j+1} - \theta_j = \sqrt{r_n}$. Notons $|\mathcal{C}|$ la taille du compact $\mathcal{C}$. On a:

$$P_2^n \leq \frac{|\mathcal{C}|}{\sqrt{r_n}} \max_k \{P_k^n + Q_k^n\} \quad (36)$$

avec:

$$\begin{aligned} P_k^n &= \mathbf{P}(\sup_{t>0} |\frac{\sqrt{n}\Delta\tilde{C}_3^n(t, \theta_k)}{\lambda_n(t, \mathcal{C})}| \geq \frac{1.76}{\sqrt{\omega}}, \sup_{\theta \in \mathcal{C}} \sup_{t>0} \frac{\sigma_n(t, \theta)}{\rho_n(t, \mathcal{C})} \leq 1) \\ Q_k^n &= \mathbf{P}(\sup_{t>0} \sup_{\theta \in [\theta_k, \theta_{k+1}]} |\frac{\sqrt{n}(\Delta\tilde{C}_3^n(t, \theta) - \Delta\tilde{C}_3^n(t, \theta_k))}{\lambda_n(t, \mathcal{C})}| \geq \frac{B(\mathcal{C})}{\sqrt{\omega}}) \end{aligned}$$

Le terme P_k^n a été contrôlé dans la preuve du lemme 15. D'après la relation (25):

$$P_k^n \leq V_n \quad \text{avec} \quad \sum_n \sqrt{n}V_n < \infty \quad (37)$$

Pour contrôler Q_k^n , nous remarquons tout d'abord:

$$\lambda_n(t, \mathcal{C})/\sqrt{n} \geq r_n \sqrt{K'(\mathcal{C})} \geq 74 \|\epsilon_1(\mathcal{C})\|_\infty / \sqrt{\omega}$$

puis nous contrôlons les oscillations:

$$\begin{aligned} \sup_{t>0} \sup_{\theta \in [\theta_k, \theta_{k+1}]} |\Delta\tilde{C}_3^n(t, \theta) - \Delta\tilde{C}_3^n(t, \theta_k)| \leq \\ 2\sqrt{r_n} \sup_{t>0} |\frac{\sqrt{p_2}(B_2(1 - \mu_2(t)))}{\sqrt{n}}| \times \sup_{s \in [0, \beta_n]} \sup_{\theta, \theta' \in \mathcal{C}; \theta \neq \theta'} |\frac{\phi_2(s, \theta) - \phi_2(s, \theta')}{\theta - \theta'}| \end{aligned}$$

L'hypothèse $\mathbf{A}'_4$ et la relation $G(\beta_n) \geq 2\omega r_n$ entraînent alors:

$$Q_k^n \leq \mathbf{P}(\sup_{t>0} |\frac{\sqrt{p_2}(B_2(1 - \mu_2(t)))}{\sqrt{n}}| \geq 74\sqrt{r_n} \|\epsilon_1(\mathcal{C})\|_\infty \frac{B(\mathcal{C})}{A(\mathcal{C})})$$

où $A(\mathcal{C})$ est la constante de l'hypothèse $\mathbf{A}'_4$. En utilisant alors l'égalité en loi:

$$\mathcal{L}(\sup_{t>0} |\sqrt{p_2}(B_2(1 - \mu_2(t)))|) = \mathcal{L}(\sup_{t>0} |B_2(G_2(t))|)$$

on obtient:

$$Q_k^n \leq \mathbf{P}(|B_2(1)| + \sup_{s \in [0, 1]} |B_2^o(s)| \geq 74\sqrt{r_n} \|\epsilon_1(\mathcal{C})\|_\infty \times \frac{B(\mathcal{C})}{A(\mathcal{C})} \times \sqrt{n})$$

En choisissant $B(\mathcal{C}) \geq 3A(\mathcal{C})/(74 \|\epsilon_1(\mathcal{C})\|_\infty)$, les lemmes 10 et 11 entraînent:

$$Q_k^n \leq 4/n^2 \quad (38)$$

Les relations (36), (37), (38) entraînent:

$$P_2^n \leq V_n \quad \text{avec} \quad \sum_n V_n < \infty \quad (39)$$

Les relations (34), (35), (39) entraînent:

$$\sup_{\theta \in \mathcal{C}} \sup_{t > 0} \left| \frac{\sqrt{n} \Delta \tilde{C}_3^n(t, \theta)}{\lambda_n(t, \mathcal{C})} \right| \leq \frac{1.76 + B(\mathcal{C})}{\sqrt{\omega}} \quad \text{p.co}$$

En utilisant $\sqrt{a+b} \leq \sqrt{a} + \sqrt{b}$ on obtient:

$$\sup_{\theta \in \mathcal{C}} \sup_{t > 0} \left| \frac{\Delta \tilde{C}_3^n(t, \theta)}{K'(\mathcal{C}) + K''(\mathcal{C}) \inf(1 + \log(1/G(t)), \log n)} \right| \leq r_n \quad \text{p.co}$$

avec $\begin{cases} K'(\mathcal{C}) = (1.76 + B(\mathcal{C}))(75 \|\epsilon_1(\mathcal{C})\|_\infty + 56 \|\epsilon_2(\mathcal{C})\|_\infty)/\omega \\ K''(\mathcal{C}) = 198(1.76 + B(\mathcal{C})) \|\epsilon_2(\mathcal{C})\|_\infty (\log \omega^{-1})/\omega \end{cases}$ qui conviennent.

4 Application à la comparaison des temps de survie entre deux populations

Avec les notations de la section 2 et en particulier de la section 1.2 (Cas des statistiques de rang) les statistiques utilisées lors des tests de comparaison des temps de survie entre les deux populations sont de la forme:

$$\hat{S}(t, \theta) = \frac{1}{n} \sum_{i=1}^n \alpha_i d_i \mathbf{1}(X_i \leq t) \left(\delta_i - \frac{n_{1i}\theta}{n_{1i}\theta + n_{2i}} \right)$$

où $\theta > 0$ est un paramètre, $t \in]0; \infty]$ est le temps auquel on arrête les observations et $(\alpha_i)_{i=1}^n$ sont des pondérations. On rappelle que θ_0 est le vrai paramètre de la loi.

Lorsque $\alpha_i = 1$; $i = 1, \dots, n$, la statistique obtenue est égale au gradient divisé par n du logarithme de la vraisemblance de Cox au point $\log(\theta)$. Elle s'écrit:

$$\hat{S}_1(t, \theta) = \frac{1}{n} \sum_{i=1}^n d_i \mathbf{1}(X_i \leq t) \left(\delta_i - \frac{n_{1i}\theta}{n_{1i}\theta + n_{2i}} \right)$$

Cette statistique peut être utilisée pour tester " $\theta_0 = \theta$ ".

La statistique du logrank est un cas particulier de la précédente: $\hat{S}_2(t) = \hat{S}_1(t, 1)$ et peut être utilisée pour tester l'égalité des deux paramètres, soit " $\theta_0 = 1$ ".

Pour la statistique de Gehan, on prend $\theta = 1$ et comme pondérations:

$$\alpha_i = \frac{n_{1i} + n_{2i}}{n}$$

On obtient:

$$\hat{S}_3(t) = \frac{1}{n} \sum_{i=1}^n d_i \mathbf{1}(X_i \leq t) \left(\frac{\delta_i n_{2i} - (1 - \delta_i) n_{1i}}{n} \right)$$

Cette statistique peut également être utilisée pour tester " $\theta_0 = 1$ ".

4.1 Propriété martingale

Afin d'approcher les variances de ces statistiques, on emploie un argument martingale. On note $\mathcal{T}(Y)$ la tribu engendrée par la variable aléatoire Y et n_1, n_2 les cardinaux (aléatoires) de la première et de la seconde population:

$$n_1 = \sum_{i=1}^n \delta_i \quad n_2 = n - n_1$$

Soient $\tilde{X}_k; k = 1, \dots, n$ les observations classées:

$$\tilde{X}_1 < \dots < \tilde{X}_k < \dots < \tilde{X}_n$$

et $I(k)$ l'antirang de la k^{eme} observation:

$$X_{I(k)} = \tilde{X}_k$$

On note:

$$\tilde{d}_k = d_{I(k)} \quad \tilde{\delta}_k = \delta_{I(k)} \quad \tilde{\alpha}_k = \alpha_{I(k)}$$

Lemme 13 La variable $\hat{S}(t, \theta_0)$ est la variable terminale de la martingale $\mathcal{T}_k$ adaptée:

$$M_k = \sum_{j=1}^k (\tilde{\alpha}_j \tilde{d}_j \tilde{\delta}_j \mathbf{1}_{\tilde{X}_j \leq t}) - \mathbf{E}_{\theta_0}^{\mathcal{T}_{j-1}}(\tilde{\alpha}_j \tilde{d}_j \tilde{\delta}_j \mathbf{1}_{\tilde{X}_j \leq t})$$

avec $\mathcal{T}_{k-1} = \mathcal{T}_0 \cap (\mathcal{T}_{k-1}^* \cup \mathcal{T}(\tilde{d}_k, \mathbf{1}_{(\tilde{X}_k \leq M)})$

$$\text{où } \begin{cases} \mathcal{T}_0 = \mathcal{T}(n_1, n_2) \\ \mathcal{T}_{k-1}^* = \mathcal{T}(\tilde{d}_j, \tilde{\delta}_j, \mathbf{1}_{(\tilde{X}_j \leq t)}; 1 \leq j \leq k-1) \end{cases}$$

Preuve du lemme 13: Avec les notations de la section 1.2, les ensembles des individus à risque dans la première et dans la seconde population peuvent se récrire:

$$\begin{aligned} \mathcal{R}_k^1 &= \{I(j); j \geq k; \tilde{\delta}_j = 1\} \\ \mathcal{R}_k^2 &= \{I(j); j \geq k; \tilde{\delta}_j = 0\} \end{aligned}$$

On rappelle $\text{card}(\mathcal{R}_k^1) = n_{1k}$ et $\text{card}(\mathcal{R}_k^2) = n_{2k}$. Il suffit de calculer $\mathcal{E}$ avec:

$$\mathcal{E} = \mathbf{E}_{\theta_0}^{\mathcal{T}_{k-1}}(\tilde{\alpha}_k \tilde{d}_k \tilde{\delta}_k \mathbf{1}_{(\tilde{X}_k \leq t)})$$

On note:

$$\mathbf{P}_k = \mathbf{P}(\tilde{\delta}_k = 1 / \mathcal{T}_0 \cap \mathcal{T}_{k-1}^* \cap \{\tilde{d}_k = 1\} \cap \{\mathbf{1}_{(\tilde{X}_k \leq t)} = 1\})$$

En utilisant les propriétés de l'espérance conditionnelle, on a:

$$\begin{aligned} \mathcal{E} &= \tilde{\alpha}_k \tilde{d}_k \mathbf{1}_{(\tilde{X}_k \leq t)} \mathbf{E}_{\theta_0}^{\mathcal{T}_{k-1}}(\tilde{\delta}_k) \\ &= \tilde{\alpha}_k \tilde{d}_k \mathbf{1}_{(\tilde{X}_k \leq t)} \mathbf{P}_k \end{aligned}$$

Soit S l'ensemble de tous les événements possibles de $\mathcal{T}_0 \cap \mathcal{T}_{k-1}^*$. Alors:

$$\mathbf{P}_k = \sum_{s \in S} \frac{\mathbf{P}(\{\tilde{\delta}_k = 1\} \cap \{\tilde{d}_k = 1\} \cap \{\mathbf{1}(\tilde{X}_k \leq t) = 1\} \cap s)}{\mathbf{P}(\{\tilde{d}_k = 1\} \cap \{\mathbf{1}(\tilde{X}_k \leq t) = 1\} \cap s)}$$

On note:

$$\begin{aligned}\mathbf{P}_k^1 &= \mathbf{P}(\{\tilde{\delta}_k = 1\} \cap \{\tilde{d}_k = 1\} \cap \{\mathbf{1}(\tilde{X}_k \leq t) = 1\} \cap s) \\ \mathbf{P}_k^2 &= \mathbf{P}(\{\tilde{\delta}_k = 0\} \cap \{\tilde{d}_k = 1\} \cap \{\mathbf{1}(\tilde{X}_k \leq t) = 1\} \cap s)\end{aligned}$$

On obtient:

$$\mathbf{P}_k = \sum_{s \in S} \frac{\mathbf{P}_k^1}{\mathbf{P}_k^1 + \mathbf{P}_k^2} \mathbf{1}_s \quad (40)$$

Avec les notations de la section 2, on a:

$$\begin{aligned}\mathbf{P}(X \in [a, b], \delta = 1, d = 1) &= \theta_0 \mathbf{P}(\delta = 1) \int_a^b C_1 S_1 \, dt \\ \mathbf{P}(X \in [a, b], \delta = 1) &= \mathbf{P}(\delta = 1) [C_1 S_1]_a^b \\ \mathbf{P}(X \in [a, b], \delta = 0, d = 1) &= \mathbf{P}(\delta = 0) \int_a^b C_2 S_2 \, dt \\ \mathbf{P}(X \in [a, b], \delta = 0) &= \mathbf{P}(\delta = 0) [C_2 S_2]_a^b\end{aligned}$$

On peut alors calculer les probabilités $\mathbf{P}_k^1$ et $\mathbf{P}_k^2$:

- Si $n_{1k} = 0$ alors $\mathbf{P}_k^1 = 0$
Si $n_{1k} \geq 1$ alors:

$$\mathbf{P}_k^1 = \theta_0 C_{n_{k-1}}^{n_{1k}-1} \int \square \int dt_1 \dots dt_{k-1} \int_{t_{k-1}}^t (a_1 C_1 S_1)^{n_{1k}} (a_2 C_2 S_2)^{n_{2k}} dt$$

- Si $n_{2k} = 0$ alors $\mathbf{P}_k^2 = 0$
Si $n_{2k} \geq 1$ alors:

$$\mathbf{P}_k^2 = C_{n_{k-1}}^{n_{1k}} \int \square \int dt_1 \dots dt_{k-1} \int_{t_{k-1}}^t (a_1 C_1 S_1)^{n_{1k}} (a_2 C_2 S_2)^{n_{2k}} dt$$

où $a_1 = \mathbf{P}(\delta = 1)$, $a_2 = \mathbf{P}(\delta = 0)$ et où $\square$ représente l'évènement s .

En reportant dans (40), on obtient:

$$\mathbf{P}_k = \frac{\theta_0 n_{1k}}{\theta_0 n_{1k} + n_{2k}} \quad \square$$

Conséquence du lemme 13

On peut écrire:

$$\hat{S}(t, \theta_0) = \frac{1}{n} \sum_{j=1}^n \xi_j$$

avec $\xi_j = \tilde{\alpha}_j \tilde{d}_j \tilde{\delta}_j (\mathbf{1}_{\tilde{X}_j \leq t}) - \mathbf{E}_{\theta_0}^{\tau_{j-1}}(\tilde{\alpha}_j \tilde{d}_j \tilde{\delta}_j \mathbf{1}_{\tilde{X}_j \leq t})$

On considère la variable aléatoire $\hat{\Gamma}^2(t, \theta)$:

$$\begin{aligned}\hat{\Gamma}^2(t, \theta_0) &= \frac{1}{n} \sum_{j=1}^n \text{Var}_{\theta_0}^{\tau_{j-1}}(\xi_j) \\ &= \frac{1}{n} \sum_{i=1}^n \alpha_i^2 d_i \mathbf{1}_{(X_i \leq t)} \frac{\theta_0 n_{1i} n_{2i}}{(\theta_0 n_{1i} + n_{2i})^2}\end{aligned}$$

On a:

$$\mathbf{E}_{\theta_0}(\hat{\Gamma}^2(t, \theta_0)) = \text{Var}_{\theta_0}(\sqrt{n} \hat{S}(t, \theta_0)) \quad (41)$$

4.2 Statistique de Gehan

On définit la variable $\hat{\Gamma}_3^2(t)$:

$$\hat{\Gamma}_3^2(t) = \frac{1}{n} \sum_{i=1}^n d_i \mathbf{1}_{(X_i \leq t)} \frac{n_{1i}}{n} \frac{n_{2i}}{n}$$

Theoreme 4 1. *Sous θ_0 , on peut construire une suite $(w_3^n(t, \theta_0))_n$ de processus gaussiens centrés et de variance $\sigma_3^2(t, \theta_0)$ tels que pour tout t positif:*

$$\hat{S}_3(t) = m_3(t, \theta_0) + \frac{w_3^n(t, \theta_0)}{\sqrt{n}} + \rho_n(t, \theta_0) \quad (42)$$

avec:

$$\sup_{t>0} |\rho_n(t, \theta_0)| \leq K \frac{\log n}{n} \quad p.co$$

où $K = K(\theta_0, \omega)$ est une constante et:

$$m_3(t, \theta_0) = \mathbf{P}(\delta = 0) \mathbf{P}(\delta = 1) (\theta_0 - 1) \int_0^t C_1 S_1 C_2 S_2 \, ds$$

$$\sup_{t>0} \sigma_3^2(t, \theta_0) \leq K'$$

où $K' = K'(\theta_0, \omega)$ est une constante.

Sous l'hypothèse $\theta_0 = 1$ on a:

$$\begin{aligned}\sigma_3^2(t, 1) &= \text{Var}_{\theta_0=1}(w_3^n(t, \theta_0)) \\ &= \mathbf{P}(\delta = 0) \mathbf{P}(\delta = 1) \int_0^t C_1 S_1 C_2 S_2 [\mathbf{P}(\delta = 1) C_1 S_1 + \mathbf{P}(\delta = 0) C_2 S_2] \, ds\end{aligned} \quad (43)$$

2. Sous θ_0 , on peut construire une suite $(v_3^n(t, \theta_0))_n$ de processus gaussiens centrés et de variance $\tau_3^2(t, \theta_0)$ tels que pour tout t positif:

$$\hat{\Gamma}_3^2(t) = \gamma_3^2(t, \theta_0) + \frac{v_3^n(t, \theta_0)}{\sqrt{n}} + \bar{\rho}_n(t, \theta_0) \quad (44)$$

avec:

$$\sup_{t>0} |\bar{\rho}_n(t, \theta_0)| \leq \bar{K} \frac{\log n}{n} \quad p.co$$

où $\bar{K} = \bar{K}(\theta_0, \omega)$ est une constante et:

$$\gamma_3^2(t, \theta_0) = \mathbf{P}(\delta = 0)\mathbf{P}(\delta = 1) \int_0^t C_1 S_1 C_2 S_2 [\theta_0 \mathbf{P}(\delta = 1) C_1 S_1 + \mathbf{P}(\delta = 0) C_2 S_2] ds$$

$$\sup_{t>0} \tau_3^2(t, \theta_0) \leq \bar{K}'$$

où $\bar{K}' = \bar{K}'(\theta_0, \omega)$ est une constante.

3. Sous l'hypothèse $\theta_0 = 1$, pour tout t positif on a:

$$\frac{\hat{S}_3(t)}{\hat{\Gamma}_3(t)} = \frac{w_3^n(t, 1)}{\sigma_3(t, 1)\sqrt{n}} + \bar{\rho}_n(t) \quad (45)$$

avec:

$$|\sigma_3^2(t, 1)\bar{\rho}_n(t)| \leq \bar{\bar{K}} \frac{\log n}{n} \quad p.co$$

où $\bar{\bar{K}} = \bar{\bar{K}}(\theta_0, \omega)$ est une constante ne dépendant pas de t .

Formulons quelques commentaires. Tout d'abord, les processus gaussiens $w_3^n(t, \theta_0)$ et $v_3^n(t, \theta_0)$ possèdent la propriété suivante:

$$\sup_{t>0} |w_3^n(t, \theta_0)| \leq A\sqrt{\log n} \quad p.co \quad (46)$$

$$\sup_{t>0} |v_3^n(t, \theta_0)| \leq A'\sqrt{\log n} \quad p.co \quad (47)$$

où $A = A(\theta_0)$ et $A' = A'(\theta_0)$ sont des constantes.

Nous n'avons pas explicité les variances de ces processus. Néanmoins, il est possible de les calculer. Par exemple, on peut exprimer $\sigma_3^2(t, \theta_0)$ comme une somme d'intégrales en calculant la variance de $w_3^n(t, \theta_0)$ sous θ_0 dans l'expression (54).

La preuve de ce théorème se trouve dans la section 4.4.

4.3 Gradient du logarithme de la vraisemblance de Cox et statistique du logrank

On définit la variable $\hat{\Gamma}_1^2(t, \theta)$:

$$\hat{\Gamma}_1^2(t, \theta) = \frac{1}{n} \sum_{i=1}^n d_i \mathbf{1}(X_i \leq t) \frac{\theta n_{1i} n_{2i}}{(\theta n_{1i} + n_{2i})^2}$$

Theoreme 5 1. *Sous θ_0 , on peut construire une suite $(w_1^n(t, \theta, \theta_0))_n$ de processus gaussiens centrés et de variance $\sigma_1^2(t, \theta, \theta_0)$ tels que pour tout t positif et pour tout compact C de $\mathbf{R}^+$:*

$$\hat{S}_1(t, \theta) = m_1(t, \theta, \theta_0) + \frac{w_1^n(t, \theta, \theta_0)}{\sqrt{n}} + \rho_n(t, \theta, \theta_0) \quad (48)$$

avec:

$$\sup_{\theta \in C} \sup_{t > 0} \left| \frac{\rho_n(t, \theta, \theta_0)}{\inf(1 + \log(1/G(t)), \log n)} \right| \leq K \frac{\log n}{n} \quad p.co$$

où $K = K(C, \theta_0, \omega)$ est une constante et:

$$m_1(t, \theta, \theta_0) = \mathbf{P}(\delta = 0)\mathbf{P}(\delta = 1)(\theta_0 - \theta) \int_0^t \frac{C_1 S_1 C_2 S_2}{\theta \mathbf{P}(\delta = 1) C_1 S_1 + \mathbf{P}(\delta = 0) C_2 S_2} ds$$

$$\sup_{\theta \in C} \sup_{t > 0} \sigma_1^2(t, \theta, \theta_0) \leq K'$$

où $K' = K'(C, \theta_0, \omega)$ est une constante.

Sous l'hypothèse $\theta_0 = \theta$, on a:

$$\begin{aligned} \sigma_1^2(t, \theta, \theta) &= \text{Var}_{\theta_0=\theta}(w_1^n(t, \theta, \theta_0)) \\ &= \mathbf{P}(\delta = 0)\mathbf{P}(\delta = 1) \int_0^t \frac{\theta C_1 S_1 C_2 S_2}{\theta \mathbf{P}(\delta = 1) C_1 S_1 + \mathbf{P}(\delta = 0) C_2 S_2} ds \end{aligned}$$

2. *Sous θ_0 , on peut construire une suite de processus gaussiens $(v_1^n(t, \theta, \theta_0))_n$ centrés et de variance $\gamma_1^2(t, \theta, \theta_0)$ tels que pour tout t positif et pour tout compact C de $\mathbf{R}^+$:*

$$\hat{\Gamma}_1^2(t, \theta) = \gamma_1^2(t, \theta, \theta_0) + \frac{v_1^n(t, \theta, \theta_0)}{\sqrt{n}} + \bar{\rho}_n(t, \theta, \theta_0) \quad (49)$$

avec:

$$\sup_{\theta \in C} \sup_{t > 0} \left| \frac{\bar{\rho}_n(t, \theta, \theta_0)}{\inf(1 + \log(1/G(t))), \log n)} \right| \leq \bar{K} \frac{\log n}{n} \quad p.co$$

où $\bar{K} = \bar{K}(C, \theta_0, \omega)$ est une constante et:

$$\gamma_1^2(t, \theta, \theta_0) = \mathbf{P}(\delta = 0)\mathbf{P}(\delta = 1) \int_0^t \frac{\theta C_1 S_1 C_2 S_2 \{\theta_0 \mathbf{P}(\delta = 1) C_1 S_1 + \mathbf{P}(\delta = 0) C_2 S_2\}}{\{\theta \mathbf{P}(\delta = 1) C_1 S_1 + \mathbf{P}(\delta = 0) C_2 S_2\}^2} ds$$

$$\sup_{\theta \in \mathcal{C}} \sup_{t > 0} \tau_1^2(t, \theta, \theta_0) \leq \bar{K}'$$

où $\bar{K}' = \bar{K}'(\mathcal{C}, \theta_0, \omega)$ est une constante.

3. Sous l'hypothèse $\theta_0 = \theta$, pour tout t positif et pour tout compact $\mathcal{C}$ de $\mathbf{R}^+$ on a :

$$\frac{\hat{S}_1(t, \theta)}{\hat{\Gamma}_1(t, \theta)} = \frac{w_1^n(t, \theta, \theta)}{\sigma_1(t, \theta, \theta)\sqrt{n}} + \bar{\rho}_n(t, \theta)$$

avec :

$$\sup_{\theta \in \mathcal{C}} \left| \frac{\sigma_1^2(t, \theta, \theta) \bar{\rho}_n(t, \theta)}{\inf(1 + \log(1/G(t)), \log n)} \right| \leq \bar{K} \frac{\log n}{n} \quad p.co$$

où $\bar{K} = \bar{K}(\mathcal{C}, \omega)$ est une constante ne dépendant pas de t .

Les commentaires du théorème 4 sont encore valables: pour tout compact $\mathcal{C}$ de $\mathbf{R}^+$, les processus gaussiens $w_1^n(t, \theta_0)$ et $v_1^n(t, \theta_0)$ possèdent la propriété:

$$\begin{aligned} \sup_{\theta \in \mathcal{C}} \sup_{t > 0} |w_1^n(t, \theta, \theta_0)| &\leq A\sqrt{\log n} \quad p.co \\ \sup_{\theta \in \mathcal{C}} \sup_{t > 0} |v_1^n(t, \theta, \theta_0)| &\leq A'\sqrt{\log n} \quad p.co \end{aligned} \tag{50}$$

où $A = A(\mathcal{C}, \theta_0)$ et $A' = A'(\mathcal{C}, \theta_0)$ sont des constantes.

De plus, il est possible d'explicitier les variances de ces processus, par exemple on peut exprimer $\sigma_1^2(t, \theta, \theta_0)$ comme une somme d'intégrales en calculant la variance de $w_1^n(t, \theta, \theta_0)$ sous θ_0 dans l'expression (87).

Theoreme 6 *L'approximation gaussienne de la statistique du logrank s'obtient en remplaçant θ par 1 dans le théorème 5.*

La preuve du théorème 5 se trouve dans la section 4.4.

4.4 Preuve des théorèmes 4 et 5

Preuve du théorème 4

On note:

$$\begin{aligned} \hat{G}_A &= \hat{G}_1 + \hat{G}_2 & \hat{G}_B &= \hat{G}_3 + \hat{G}_4 \\ G_A &= G_1 + G_2 & G_B &= G_3 + G_4 \end{aligned}$$

On écrit $\hat{S}_3(t)$ sous forme intégrale:

$$\hat{S}_3(t) = \int_0^t \hat{G}_B d\hat{F}_1 - \int_0^t \hat{G}_A d\hat{F}_3$$

Nous utilisons la méthode décrite dans la section 3. Dans la décomposition (2) on trouve:

$$\phi_j = 1 \quad j = 1, 2, 3, 4 \quad \text{et} \quad u_n = 0$$

Les conditions $\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3$ sont vérifiées avec:

$$\epsilon_1 = G \quad \text{et} \quad \epsilon_2 = 0$$

En développant la statistique comme en (4) on obtient:

$$|R_2^n| = |R_3^n| = 0, \quad I_n = 0$$

D'après le théorème 2:

$$\|R_1^n\|_\infty + \|\Delta^n\|_\infty \leq K_1 r_n \quad \text{p.co}$$

ce qui entraîne:

$$\|\hat{S}_3(\cdot) - m_3(\cdot, \theta_0) - \frac{\hat{w}_3(\cdot, \theta_0)}{\sqrt{n}}\|_\infty \leq K_1 r_n \quad \text{p.co} \quad (51)$$

avec:

$$m_3(t, \theta_0) = \mathbf{P}(\delta = 0)\mathbf{P}(\delta = 1)(\theta_0 - 1) \int_0^t C_1 S_1 C_2 S_2 \, ds$$

et:

$$\begin{aligned} \hat{w}_3(t, \theta_0) &= I(\hat{F}_i - F_i; i = 1, 2, 3, 4, \hat{G}_j - G_j; j = 1, 2, 3, 4) \\ &= \int_0^t \sqrt{n}(\hat{F}_3 - F_3) \, dG_A - \int_0^t \sqrt{n}(\hat{F}_1 - F_1) \, dG_B \\ &\quad - \int_0^t \sqrt{n}(\hat{G}_A - G_A) \, dF_3 + \int_0^t \sqrt{n}(\hat{G}_B - G_B) \, dF_1 \end{aligned}$$

D'après le théorème 1, on a simultanément:

$$\|\hat{F}_i - F_i - \frac{\sqrt{p_i}(B_i^o(\mu_i))}{\sqrt{n}} - \sqrt{\frac{q_i}{p_i}} F_i \frac{Z_i}{\sqrt{n}}\|_\infty \leq 61 r_n \quad \text{p.co} \quad (52)$$

$$\|\hat{G}_j - G_j - \frac{\sqrt{p_j}(B_j^o(\mu_j))}{\sqrt{n}} - \sqrt{\frac{q_j}{p_j}} G_j \frac{Z_j}{\sqrt{n}}\|_\infty \leq 98 r_n \quad \text{p.co}$$

Et finalement:

$$\|\frac{\hat{w}_3(\cdot, \theta_0)}{\sqrt{n}} - \frac{w_3^n(\cdot, \theta_0)}{\sqrt{n}}\|_\infty \leq K_2 r_n \quad \text{p.co} \quad (53)$$

où $w_3^n(t, \theta_0)$ est le processus gaussien obtenu en remplaçant dans l'expression de $\hat{w}_3(t, \theta_0)$ les sous ponts empiriques par les processus gaussiens associés:

$$w_3^n(t, \theta_0) = I\left(\frac{\sqrt{p_i}(B_i)^o(\mu_i)}{\sqrt{n}} - \sqrt{\frac{q_i}{p_i}} F_i \frac{Z_i}{\sqrt{n}}, \frac{\sqrt{p_j}(B_j)^o(\mu_j)}{\sqrt{n}} - \sqrt{\frac{q_j}{p_j}} G_j \frac{Z_j}{\sqrt{n}}\right) \quad (54)$$

$$i = 1, 2, 3, 4 \quad j = 1, 2, 3, 4$$

On en déduit:

$$\| \hat{S}_3(\cdot) - m_3(\cdot, \theta_0) - \frac{w_3^n(\cdot, \theta_0)}{\sqrt{n}} \|_{\infty} \leq K_3 r_n \quad \text{p.co}$$

Remarque: la dépendance en n de la variable $w_3^n(t, \theta_0)$ provient de la remarque 2 du théorème 1.

La variance du processus $w_3^n(t, \theta_0)$ est difficilement calculable. Néanmoins,

$$|\mathbb{E}(\sqrt{p_i}(B_i^o(\mu_i)) - \sqrt{p_i q_i} \nu_i Z_i)(\sqrt{p_j}(B_j^o(\mu_j)) - \sqrt{p_j q_j} \nu_j Z_j)| \leq 2 \quad (55)$$

$$\nu_i \in \{\mu_i, 1 - \mu_i\} \quad ; \quad \nu_j \in \{\mu_j, 1 - \mu_j\}$$

On en déduit:

$$\sup_{t \geq 0} \sigma_3^2(t, \theta_0) < \infty$$

L'assertion (42) est démontrée. Le lemme 8 et la relation (53) entraînent immédiatement:

$$\| \frac{w_1^n(\cdot, \theta_0)}{\sqrt{n}} \|_{\infty} \leq A(\theta_0) \sqrt{r_n} \quad \text{p.co}$$

ce qui prouve la relation (46).

On écrit $\hat{\Gamma}_3^2(t)$ sous forme intégrale:

$$\hat{\Gamma}_3^2(t) = \int_0^t \hat{G}_B \hat{G}_A \, d(\hat{F}_1 + \hat{F}_3)$$

La méthode décrite dans la section 3 permet également d'étudier ce processus. On décompose $\hat{G}_B \hat{G}_A - G_A G_B$ suivant (2):

$$\hat{G}_B \hat{G}_A - G_A G_B = G_A (\hat{G}_B - G_B) G_B (\hat{G}_A - G_A) + (\hat{G}_A - G_A) (\hat{G}_B - G_B)$$

On obtient:

$$\phi_1 = \phi_2 = G_B \quad \phi_3 = \phi_4 = G_A \quad u_n = (\hat{G}_A - G_A)(\hat{G}_B - G_B)$$

Les conditions $\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3$ sont vérifiées avec:

$$\epsilon_1 = \max(G_A G, G_B G) \quad \text{et} \quad \epsilon_2 = \max(|G'_A| G^2 / |G'|, |G'_B| G^2 / |G'|)$$

Nous développons alors la statistique comme en (4). En remarquant:

$$\| I_n \|_{\infty} \leq \int_0^{\infty} |G'| |G| \, ds \leq 1/2$$

le théorème 2 entraîne:

$$\| R_1^n \|_{\infty} + \| \Delta^n \|_{\infty} \leq K_1 r_n \quad \text{p.co}$$

Il reste à contrôler R_2^n et R_3^n avec:

$$\begin{aligned} |R_2^n(t)| &= \int_0^{M_n(t)} |(\hat{G}_A - G_A)(\hat{G}_B - G_B)| \, d\hat{F}_1 \\ |R_3^n(t)| &= \int_{M_n(t)}^t |(\hat{G}_A - G_A)(\hat{G}_B - G_B)| \, dF_1 \end{aligned}$$

(les termes en $\hat{F}_3$ et F_3 sont du même type)

D'après le lemme 8, on a immédiatement $\|R_2^n\|_\infty \leq 7r_n$ p.co. Si $t < \beta_n$ alors $|R_3^n(t)| = 0$, dans le cas contraire:

$$|R_3^n(t)| \leq 4 \int_{\beta_n}^\infty dF_1 \leq 11r_n$$

On conclut comme précédemment et l'on obtient l'assertion (44) et la relation (47).

Les inégalités suivantes sont des conséquences de (42) et de (44):

$$\|\mathbf{E}_{\theta_0}(\hat{S}_3(\cdot)) - m_3(\cdot, \theta_0)\|_\infty \leq K_1 r_n \quad (56)$$

$$\|\mathbf{Var}_{\theta_0}((\sqrt{n}\hat{S}_3(\cdot)) - \sigma_3^2(\cdot, \theta_0))\|_\infty \rightarrow 0 \quad \text{quand } n \rightarrow \infty \quad (57)$$

$$\|\mathbf{E}_{\theta_0}\hat{\Gamma}_3^2(\cdot) - \gamma_3^2(\cdot, \theta_0)\|_\infty \leq K_2 r_n \quad (58)$$

Preuve de (56), (57), (58)

Soit A_n l'évènement:

$$A_n = \left\{ \left\| \hat{S}_3(\cdot) - m_3(\cdot, \theta_0) - \frac{w_3^n(\cdot, \theta_0)}{\sqrt{n}} \right\|_\infty > Kr_n \right\}$$

où K est la constante de (42) et soit $\bar{A}_n$ l'évènement complémentaire. D'après (42):

$$nP(A_n) \rightarrow 0 \quad \text{quand } n \rightarrow \infty \quad (59)$$

Posons:

$$\sigma^*(\theta_0) = \sup_{t>0} \sigma(t, \theta_0)$$

Les inégalités suivantes sont des propriétés immédiates des variables gaussiennes et de l'inégalité (59):

$$\mathbf{E} \left(\left| \frac{w_3(t, \theta_0)}{\sqrt{n}} \right| \mathbf{1}_{A_n} \right) \leq \frac{K'\sigma^*(\theta_0)}{\sqrt{n}} P(A_n) \sqrt{\log \frac{1}{P(A_n)}} \leq K \frac{\sqrt{\log n}}{n^{3/2}} \quad (60)$$

$$\mathbf{E} \left(\left(\frac{w_3(t, \theta_0)}{\sqrt{n}} \right)^2 \mathbf{1}_{A_n} \right) \leq \frac{K'(\sigma^*(\theta_0))^2}{n} P(A_n) \log \frac{1}{P(A_n)} \leq K \frac{\log n}{n^2} \quad (61)$$

où K est une constante ne dépendant pas de t .
En utilisant le résultat (42) et les majorations:

$$\begin{aligned} \|\hat{S}_3\|_\infty &\leq 1 \\ \|\mathbf{m}_3(\cdot, \theta_0)\|_\infty &\leq \mu(\theta_0) = (1 + \theta_0)\mathbf{P}(\delta = 0)\mathbf{P}(\delta = 1) \int_0^\infty C_1 S_1 C_2 S_2 \, ds \end{aligned}$$

on obtient:

$$\begin{aligned} |\mathbf{E}_{\theta_0}(\hat{S}_3(t)) - \mathbf{m}_3(t, \theta_0)| &\leq \mathbf{E}_{\theta_0} \left[\left| \hat{S}_3(t) - \mathbf{m}_3(t, \theta_0) - \frac{w_3^n(t, \theta_0)}{\sqrt{n}} \right| (\mathbf{1}_{A_n} + \mathbf{1}_{\bar{A}_n}) \right] \\ &\leq K r_n + (1 + \mu(\theta_0))\mathbf{P}(A_n) + \mathbf{E}_{\theta_0} \left(\left| \frac{w_3^n(t, \theta_0)}{\sqrt{n}} \right| \mathbf{1}_{A_n} \right) \end{aligned}$$

Les inégalités (59) et (60) entraînent:

$$|\mathbf{E}_{\theta_0}(\hat{S}_3(t)) - \mathbf{m}_3(t, \theta_0)| \leq K r_n$$

où K ne dépend pas de t , soit:

$$\|\mathbf{E}_{\theta_0}(\hat{S}_3(\cdot)) - \mathbf{m}_3(\cdot, \theta_0)\|_\infty \leq K r_n$$

L'inégalité (56) est démontrée, l'inégalité (58) s'obtient par les mêmes arguments.
Une conséquence directe de (42) et (56) est:

$$\|\hat{S}_3(\cdot) - \mathbf{E}_{\theta_0}(\hat{S}_3(\cdot)) - \frac{w_3^n(\cdot, \theta_0)}{\sqrt{n}}\|_\infty \leq K r_n \quad \text{p.co} \quad (62)$$

Pour démontrer (57), nous décomposons:

$$|\text{Var}(\sqrt{n}\hat{S}_3(t)) - \sigma_3^2(t, \theta_0)| \leq T_1^n + T_2^n \quad (63)$$

avec:

$$\begin{aligned} T_1^n &= n \mathbf{E}_{\theta_0} \left| \hat{S}_3(t) - \mathbf{E}_{\theta_0}(\hat{S}_3(t)) - \frac{w_3^n(t, \theta_0)}{\sqrt{n}} \right|^2 \\ T_2^n &= 2n \mathbf{E}_{\theta_0} \left\| \frac{w_3^n(t, \theta_0)}{\sqrt{n}} \right\| \left\| \hat{S}_3(t) - \mathbf{E}_{\theta_0}(\hat{S}_3(t)) - \frac{w_3^n(t, \theta_0)}{\sqrt{n}} \right\| \end{aligned}$$

Remarquons:

$$\|\mathbf{E}_{\theta_0}(\hat{S}_3)\|_\infty \leq 1$$

Soit A_n l'évènement:

$$A_n = \left\{ \left\| \hat{S}_3(\cdot) - \mathbf{E}_{\theta_0}(\hat{S}_3(\cdot)) - \frac{w_3^n(\cdot, \theta_0)}{\sqrt{n}} \right\|_\infty > K r_n \right\}$$

où K est la constante de (62) et soit $\bar{A}_n$ l'évènement complémentaire. Les relations (59), (60) et (61) sont encore vraies et avec les mêmes arguments que dans la preuve de l'inégalité (56) on obtient:

$$T_1^n \leq n[(Kr_n)^2 + 2P(A_n) + E_{\theta_0}\left(\left(\frac{w_3^n(t, \theta_0)}{\sqrt{n}}\right)^2 \mathbf{1}_{A_n}\right)]$$

Les relations (59) et (61) entraînent:

$$T_1^n \rightarrow 0 \quad \text{quand } n \rightarrow \infty \quad (64)$$

De même,

$$T_2^n \leq 2n[r_n E_{\theta_0}\left(\left|\frac{w_3^n(t, \theta_0)}{\sqrt{n}}\right|\right) + 2E_{\theta_0}\left(\left|\frac{w_3^n(t, \theta_0)}{\sqrt{n}}\right| \mathbf{1}_{A_n}\right) + E_{\theta_0}\left(\left(\frac{w_3^n(t, \theta_0)}{\sqrt{n}}\right)^2 \mathbf{1}_{A_n}\right)]$$

La relation évidente :

$$\|E_{\theta_0}\left(\left|\frac{w_3^n(\cdot, \theta_0)}{\sqrt{n}}\right|\right)\|_{\infty} \leq \frac{2\sigma^*(\theta_0)}{\sqrt{n}}$$

et les inégalités (60), (61) entraînent:

$$T_2^n \leq K \frac{\log n}{\sqrt{n}} \quad (65)$$

où K ne dépend pas de t , soit en réunissant (63), (64), (65):

$$\|\text{Var}(\sqrt{n}\hat{S}_3(\cdot)) - \sigma_3^2(\cdot, \theta_0)\|_{\infty} \rightarrow 0 \quad \text{quand } n \rightarrow \infty$$

□

Lorsque $\theta_0 = 1$, l'égalité (41) entraîne:

$$\begin{aligned} \|\gamma_3^2(\cdot, 1) - \sigma_3^2(\cdot, 1)\|_{\infty} &\leq \|\gamma_3^2(\cdot, 1) - E_{\theta_0=1}(\hat{\Gamma}_3^2(\cdot))\|_{\infty} \\ &\quad + \|\text{Var}_{\theta_0=1}((\sqrt{n}\hat{S}_3(\cdot)) - \sigma_3^2(\cdot, \theta_0))\|_{\infty} \end{aligned}$$

D'après (57) et (58) le membre de droite tend vers 0 quand $n \rightarrow \infty$. Le membre de gauche ne dépendant pas de n , on en déduit:

$$\gamma_3^2(t, 1) = \sigma_3^2(t, 1)$$

L'égalité (43) est démontrée.

Il reste à prouver la troisième partie du théorème 4. Par hypothèse on prend $\theta_0 = 1$. Rappelons:

$$m_3(t, 1) = 0$$

$$\sigma_3^2(t, 1) = \gamma_3^2(t, 1)$$

On note T_n le terme devant être contrôlé:

$$T_n(t) = \left| \frac{\hat{S}_3(t)}{\hat{\Gamma}_3(t)} - \frac{w_3^n(t, 1)}{\sigma_3(t, 1)\sqrt{n}} \right|$$

La relation (42) devient:

$$\left| \hat{S}_3(t) - \frac{w_3^n(t, 1)}{\sqrt{n}} \right| \leq K_1 r_n \quad \text{p.co} \quad (66)$$

Avec la relation (44) et le lemme 10 on obtient:

$$\begin{aligned} |\hat{\Gamma}_3^2(t) - \sigma_3^2(t, 1)| &\leq \left| \hat{\Gamma}_3^2(t) - \sigma_3^2(t, 1) - \frac{v_3^n(t, 1)}{\sqrt{n}} \right| + \left| \frac{v_3^n(t, 1)}{\sqrt{n}} \right| \\ &\leq K_1 r_n + 2(\sup_t \tau_3(t, 1))\sqrt{r_n} \quad \text{p.co} \\ &\leq K_2 r_n \quad \text{p.co} \end{aligned} \quad (67)$$

Les inégalités (66), (67) et de nouveau le lemme 10 entraînent:

$$\begin{aligned} T_n(t) &\leq \frac{1}{\hat{\Gamma}_3(t)} \left| \hat{S}_3(t) - \frac{w_3^n(t, 1)}{\sqrt{n}} \right| + \left| \frac{w_3^n(t, 1)}{\sigma_3^2(t, 1)\sqrt{n}} \right| |\hat{\Gamma}_3^2(t) - \sigma_3^2(t, 1)| \\ &\leq \frac{K_3 r_n}{\hat{\Gamma}_3(t)\sigma_3(t, 1)} \quad \text{p.co} \end{aligned}$$

où K_3 est une constante qui ne dépend pas de t .

La relation (67), le fait qu'à t fixé il existe n tel que $\sigma_3^2(t, 1)/2 \geq K_2 r_n$ où K_2 est la constante de (67) et l'inégalité:

$$\mathbf{P}(\hat{\Gamma}_3^2(t) \leq \frac{\sigma_3^2(t, 1)}{2}) \leq \mathbf{P}(|\hat{\Gamma}_3^2(t) - \sigma_3^2(t, 1)| \geq \frac{\sigma_3^2(t, 1)}{2})$$

entraînent:

$$\hat{\Gamma}_3^2(t) \geq \frac{\sigma_3^2(t, 1)}{2} \quad \text{p.co}$$

ce qui achève la preuve du théorème 4.

Preuve du théorème 5

On note:

$$\begin{aligned} \hat{H}_A &= \frac{\hat{G}_A}{\theta \hat{G}_A + \hat{G}_B} & \hat{H}_B &= \frac{\hat{G}_B}{\theta \hat{G}_A + \hat{G}_B} \\ H_A &= \frac{G_A}{\theta G_A + G_B} & H_B &= \frac{G_B}{\theta G_A + G_B} \end{aligned}$$

où $\hat{G}_A, \hat{G}_B, G_A, G_B$ sont les mêmes que dans la preuve du théorème 4. On écrit $\hat{S}_1(t, \theta)$ sous forme intégrale:

$$\hat{S}_1(t, \theta) = \int_0^t \hat{H}_B d\hat{F}_1 - \theta \int_0^t \hat{H}_A d\hat{F}_3$$

Nous appliquons la méthode décrite dans la section 3 au terme $\int_0^t \hat{H}_B d\hat{F}_1$ (l'autre terme étant du même type). Pour obtenir le développement (4), on linéarise $\hat{H}_B - H_B$ suivant (2):

$$\begin{aligned}\hat{H}_B - H_B &= \theta \hat{H}_A \frac{\hat{G}_B - G_B}{\theta G_A + G_B} - \theta \hat{H}_B \frac{\hat{G}_A - G_A}{\theta G_A + G_B} \\ &= \theta \hat{H}_A \frac{\hat{G}_B - G_B}{\theta G_A + G_B} - \theta H_B \frac{\hat{G}_A - G_A}{\theta G_A + G_B} + u_n\end{aligned}$$

avec:

$$u_n = \theta(\hat{H}_A - H_A) \frac{\hat{G}_B - G_B}{\theta G_A + G_B} - \theta(\hat{H}_B - H_B) \frac{\hat{G}_A - G_A}{\theta G_A + G_B} \quad (68)$$

$$\begin{aligned}&= -\theta \hat{H}_A \left(\frac{\hat{G}_B - G_B}{\theta G_A + G_B} \right)^2 - \theta^2 \hat{H}_B \left(\frac{\hat{G}_A - G_A}{\theta G_A + G_B} \right)^2 \\ &\quad + \theta(1 - \theta) \hat{H}_A \hat{H}_B \left(\frac{\hat{G}_A - G_A}{\theta G_A + G_B} \right) \left(\frac{\hat{G}_B - G_B}{\theta G_A + G_B} \right)\end{aligned} \quad (69)$$

On trouve:

$$\begin{aligned}\phi_1 &= \phi_2 = -\theta G_B / (\theta G_A + G_B)^2 \\ \phi_3 &= \phi_4 = \theta G_A / (\theta G_A + G_B)^2\end{aligned}$$

On vérifie $\| \epsilon_1(C) \|_\infty \leq \sup_{\theta \in C} \max(\theta, 1/\theta)$, $\| \epsilon_2(C) \|_\infty \leq 3 \sup_{\theta \in C} \max(\theta^2, 1/\theta^2)$ et les conditions $\mathbf{A}'_1$, $\mathbf{A}'_2$, $\mathbf{A}'_3$ sont satisfaites. La condition $\mathbf{A}'_4$ est également satisfaite avec $A(C) = \sup_{\theta \in C} \max(1, 1/\theta)$. On développe alors la statistique comme en (4). Pour tout compact C de $\mathbb{R}^+$, le théorème 3 entraîne:

$$\sup_{\theta \in C} \sup_{t > 0} \left| \frac{R_1^n(t, \theta) + \Delta^n(t, \theta)}{\inf(1 + \log(1/G(t)), \log n)} \right| \leq K(C) r_n \quad \text{p.co} \quad (70)$$

où $K(C)$ est une constante qui dépend du compact C .

Remarque: Toutes les constantes considérées dépendent de ω aussi nous ne le précisons plus.

Il reste à contrôler R_2^n et R_3^n .

Contrôle de R_2^n

En utilisant l'égalité (69), $\| \theta \hat{H}_A \|_\infty, \| \hat{H}_B \|_\infty \leq 1$ ainsi que les relations $|ab| \leq (a^2 + b^2)/2$ et $(a - b)^2 \leq a^2 + b^2$ on obtient:

$$\forall t > 0 \quad |R_2^n(t, \theta)| \leq K_1(\theta) \int_0^{M_n(t)} \left[\left(\frac{\hat{G}_A - G_A}{\theta G_A + G_B} \right)^2 + \left(\frac{\hat{G}_B - G_B}{\theta G_A + G_B} \right)^2 \right] d\hat{F}_1$$

où $K_1(\theta)$ est une constante dépendant continûment de θ .

En utilisant $\theta G_A + G_B \geq G/\lambda(\theta) \leq G_A/\lambda(\theta)$ où $\lambda(\theta) = \max(1, 1/\theta)$, pour tout t positif on a:

$$|R_2^n(t, \theta)| \leq K_2(\theta) \left[\int_0^{M_n(t)} \left(\frac{\hat{G}_A - G_A}{G} \right)^2 d(\hat{F}_1) + \int_0^{M_n(t)} \left(\frac{\hat{G} - G}{G} \right)^2 d(\hat{F}_1) \right] \quad (71)$$

où $K_2(\theta) = K_1(\theta)\lambda(\theta)$. Les deux termes du membre de droite se contrôlent de la même façon.

Soit $\Psi \in \{G_A, G\}$ et les suites $(\alpha_n)_n$, $(Q_n)_n$ définies par:

$$\Psi(\alpha_n) = (3\pi r_n)/\log 3 \text{ où } \pi = \begin{cases} p_A = p_1 + p_2 & \text{si } \Psi = G_A \\ 1 & \text{si } \Psi = G \end{cases} \quad (72)$$

$$Q_n(t) = \min(\alpha_n, t)$$

On étudie l'intégrale $J^n(t)$ avec:

$$\begin{aligned} J^n(t) &= \int_0^{M_n(t)} \left(\frac{\hat{\Psi} - \Psi}{G} \right)^2 d(\hat{F}_1) \\ &\leq \begin{cases} J_1^n(t) + J_2^n(t) & \text{si } Q_n(t) < M_n(t) \\ J_2^n(t) & \text{si } Q_n(t) \geq M_n(t) \end{cases} \end{aligned}$$

$$\text{où } \begin{cases} J_1^n(t) = \int_0^{Q_n(t)} \left(\frac{\hat{\Psi} - \Psi}{\sqrt{\Psi}} \right)^2 \frac{d(\hat{F}_1)}{G} \\ J_2^n(t) = \int_{Q_n(t)}^{M_n(t)} \left(\frac{\hat{\Psi} - \Psi}{G} \right)^2 d(-\hat{\Psi})^- \end{cases}$$

Le terme J_1^n se contrôle avec le lemme 9 et une intégration par parties. Pour tout t positif on a:

$$\begin{aligned} J_1^n(t) &\leq \left(\sup_{t>0} \sup_{s \in [0; Q_n(t)]} \left| \frac{\hat{\Psi} - \Psi}{\sqrt{\Psi}} \right| \right)^2 \left(\sup_{t>0} \sup_{s \in [0; Q_n(t)]} \left| \frac{\hat{\Psi} - \Psi}{\Psi} \right| + 1 \right) \\ &\quad \times \left(2 + \log \left(\frac{1}{G(M_n(t))} \right) \right) \end{aligned} \quad (73)$$

On obtient :

$$\left\| \frac{J_1^n(\cdot)}{2 + \log(1/G(M_n(\cdot)))} \right\|_\infty \leq K_2 r_n \quad \text{p.co}$$

avec $K_2 = 14^2 \times 9/\omega^2$ qui convient.

En utilisant la relation $(a + b)^2 \leq 2(a^2 + b^2)$ et le fait que $Q_n(t) < M_n(t)$ entraîne $Q_n(t) = \alpha_n$, on a:

$$\|J_2^n\|_\infty \leq 2 \left[\left(\frac{\hat{\Psi}(\alpha_n)}{G(\beta_n)} \right)^2 + 1 \right] \hat{\Psi}(\alpha_n)$$

De même que dans la preuve du lemme 10:

$$P(|\hat{\Psi}(\alpha_n) - \Psi(\alpha_n)| \geq 3r_n) \leq 2 \exp(-1.24 \log n)$$

Finalement,

$$\|J_2^n\|_\infty \leq \frac{74r_n}{\omega^2} \quad \text{p.co} \quad (74)$$

Soit, en réunissant (73) et (74):

$$\left\| \frac{J^n}{\inf(1 + \log(1/G), \log n)} \right\|_\infty \leq Kr_n \quad \text{p.co} \quad (75)$$

où K est une constante ne dépendant pas de θ .

Les relations (71) et (75) entraînent que pour tout compact $\mathcal{C}$ de $\mathbb{R}^+$:

$$\sup_{\theta \in \mathcal{C}} \sup_{t > 0} \left| \frac{R_2^n(t, \theta)}{\inf(1 + \log(1/G(t)), \log n)} \right| \leq K(\mathcal{C})r_n \quad \text{p.co} \quad (76)$$

où $K(\mathcal{C})$ est une constante dépendant du compact $\mathcal{C}$.

Contrôle de R_3^n

En utilisant l'égalité (68), $\|\theta \hat{H}_A\|_\infty, \|\theta H_A\|_\infty, \|\hat{H}_B\|_\infty, \|H_B\|_\infty \leq 1$ ainsi que la relation $|a - b| \leq |a| + |b|$, on obtient:

$$|u_n(s, \theta)| \leq K_1(\theta) \left[\left| \frac{\hat{G}(s) - G(s)}{\theta G_A(s) + G_B(s)} \right| + \left| \frac{\hat{G}_A(s) - G_A(s)}{\theta G_A(s) + G_B(s)} \right| \right]$$

où $K_1(\theta)$ est une constante dépendant continûment de θ .

On obtient:

$$|u_n(s, \theta)| \leq K_2(\theta) \left[\left| \frac{\hat{G}(s) - G(s)}{G(s)} \right| + \left| \frac{\hat{G}_A(s) - G_A(s)}{G_A(s)} \right| \right]$$

et pour tout t positif:

$$|R_3^n(t, \theta)| \leq K_2(\theta) \left[\int_{M_n(t)}^t \left| \frac{\hat{G}(s) - G(s)}{G(s)} \right| dF_1 + \int_{M_n(t)}^t \left| \frac{\hat{G}_A(s) - G_A(s)}{G_A(s)} \right| dF_1 \right] \quad (77)$$

où $K_2(\theta) = K_1(\theta) \max(1, 1/\theta)$ dépend continûment de θ .

Les deux termes du membre de droite se contrôlent de la même façon. Soit $\Psi \in \{G_A, G\}$. On remarque que $\beta_n \geq t$ entraîne $R_3^n = 0$ donc le terme à étudier est majoré par:

$$J^n = \int_{\beta_n}^\infty \left| \frac{\hat{\Psi} - \Psi}{\Psi} \right| dF_1$$

Soit $(\alpha_n)_n$ la suite définie en (72). On obtient:

$$J^n \leq \begin{cases} J_1^n + J_2^n & \text{si } \alpha_n > \beta_n \\ J_2^n & \text{si } \alpha_n \leq \beta_n \end{cases}$$

avec

$$\begin{cases} J_1^n = \int_{\beta_n}^{\alpha_n} \left| \frac{\hat{\Psi} - \Psi}{\Psi} \right| dF_1 \\ J_2^n = \int_{\alpha_n}^{\infty} \left| \frac{\hat{\Psi} - \Psi}{\Psi} \right| d(-\Psi) \end{cases}$$

Le lemme 9 entraine:

$$\begin{aligned} J_1^n &\leq \left(\sup_{s \in [0; \alpha_n]} \left| \frac{\hat{\Psi} - \Psi}{\Psi} \right| \right) \int_{\beta_n}^{\infty} dF_1 \\ &\leq 25r_n/\omega \quad \text{p.co} \end{aligned} \quad (78)$$

On décompose le terme J_2^n :

$$J_2^n \leq J_3^n + \int_{\alpha_n}^{\infty} d(-\Psi) \leq J_3^n + (3r_n/\log 3)$$

avec:

$$J_3^n = \int_{\alpha_n}^{\infty} \frac{\hat{\Psi}}{\Psi} d(-\Psi)$$

Or $\hat{\Psi}(t) = (\sum_{i=1}^n \mathbf{1}_{\epsilon_i X_i \geq t})/n$ avec $\epsilon_i = 1$ si $\Psi = G$ et $\epsilon_i = \delta_i$ si $\Psi = G_A$. Donc

$$J_3^n = \frac{1}{n} \sum_{i=1}^n Z_i$$

$$\text{avec } Z_i = \int_{\alpha_n}^{\infty} \frac{\mathbf{1}_{\epsilon_i X_i \geq t}}{\Psi(t)} d(-\Psi(t))$$

Comme $\mathbf{P}(\epsilon_i X_i \geq t) = \Psi(t)$, on obtient $\mathbf{E}(Z_i) = 3\pi r_n/\log 3$, ce qui entraine:

$$\mathbf{P}(J_3^n \geq Kr_n) \leq \exp(-nh(Kr_n)) \quad (80)$$

où h est la transformée de Cramer de Z_i :

$$h(a) = \sup_s \{as - \log \mathbf{E}[\exp(sZ_i)]\} \quad (81)$$

On minore la fonction h :

$$\begin{aligned} Z_i &= \int_{\alpha_n}^{\epsilon_i X_i \vee \alpha_n} \frac{d(-\Psi)}{\Psi} = \log \left(\frac{\Psi(\alpha_n)}{\Psi(\epsilon_i X_i \vee \alpha_n)} \right) \\ \mathbf{E}[\exp(sZ_i)] &= \mathbf{P}(\epsilon_i X_i < \alpha_n) + \mathbf{E} \left[\left(\frac{\Psi(\alpha_n)}{\Psi(\epsilon_i X_i)} \right)^s \mathbf{1}_{\epsilon_i X_i \geq \alpha_n} \right] \\ &= 1 - \Psi(\alpha_n) + \int_{\alpha_n}^{\infty} \left(\frac{\Psi(\alpha_n)}{\Psi(t)} \right)^s d(-\Psi)(t) \end{aligned}$$

Le changement de variable $u = \Psi(t)$ entraine:

$$E[\exp(sZ_i)] = \begin{cases} 1 + (s/(1-s))(3\pi r_n / \log 3) & \text{si } s < 1 \\ +\infty & \text{sinon} \end{cases}$$

En reportant cette égalité dans (81) avec $\log(1+x) \leq x$, on a:

$$h(Kr_n) \geq \sup_{s < 1} \left\{ Kr_n s - \frac{s}{1-s} \frac{3r_n}{\log 3} \right\}$$

On prend $K = 8$ et $s = 0.5$ et on reporte dans (80) pour obtenir:

$$J_3^n \leq 8r_n \quad \text{p.co} \quad (82)$$

En réunissant (78), (79), (82) on a:

$$J^n \leq 31r_n / \omega \quad \text{p.co} \quad (83)$$

Les relations (77) et (83) entraînent que pour tout compact C de $\mathbf{R}^+$:

$$\sup_{\theta \in C} \sup_{t > 0} |R_3^n(t, \theta)| \leq K(C)r_n \quad \text{p.co} \quad (84)$$

où $K(C)$ est une constante dépendant du compact C . Les relations (70), (76) et (84) entraînent que pour tout compact C de $\mathbf{R}^+$:

$$\sup_{\theta \in C} \sup_{t > 0} \left| \frac{\Delta^n(t, \theta) + R_1^n(t, \theta) + R_2^n(t, \theta) + R_3^n(t, \theta)}{\inf(1 + \log(1/G(t)), \log n)} \right| \leq K(C)r_n \quad \text{p.co}$$

Finalement:

$$\sup_{\theta \in C} \sup_{t > 0} \left| \frac{\hat{S}_1(t, \theta) - m_1(t, \theta, \theta_0) - (\hat{w}_1(t, \theta, \theta_0)/\sqrt{n})}{\inf(1 + \log(1/G(t)), \log n)} \right| \leq K(C)r_n \quad \text{p.co} \quad (85)$$

avec:

$$m_1(t, \theta, \theta_0) = P(\delta = 0)P(\delta = 1)(\theta_0 - \theta) \int_0^t \frac{C_1 S_1 C_2 S_2}{\theta P(\delta = 1) C_1 S_1 + P(\delta = 0) C_2 S_2} ds$$

et

$$\begin{aligned} \hat{w}_1(t, \theta, \theta_0) &= J_\theta(\hat{F}_i - F_i; i = 1, 2, 3, 4, \hat{G}_j - G_j; j = 1, 2, 3, 4) \\ &= \int_0^t \sqrt{n}(\hat{F}_3 - F_3) d(\theta H_A) - \int_0^t \sqrt{n}(\hat{F}_1 - F_1) d(\theta H_B) \\ &\quad + \theta \int_0^t \left[\frac{G_A(\hat{G}_B - G_B) - G_B(\hat{G}_A - G_A)}{(\theta G_A + G_B)^2} \right] d(F_1 + F_3) \end{aligned}$$

D'après (52):

$$\sup_{\theta \in C} \sup_{t > 0} \left| \frac{w_1^n(t, \theta, \theta_0)}{\sqrt{n}} - \frac{\hat{w}_1(t, \theta, \theta_0)}{\sqrt{n}} \right| \leq K(C)r_n \quad \text{p.co} \quad (86)$$

où $K(C)$ est une constante dépendant du compact C et où $w_1^n(t, \theta, \theta_0)$ est le processus gaussien obtenu en remplaçant dans l'expression de $\hat{w}_1(t, \theta, \theta_0)$ les sous ponts empiriques par les processus gaussiens associés:

$$w_1^n(t, \theta, \theta_0) = J_\theta \left(\frac{\sqrt{p_i}(B_i^o(\mu_j))}{\sqrt{n}} - \sqrt{\frac{q_i}{p_i}} F_i \frac{Z_i}{\sqrt{n}}, \frac{\sqrt{p_j}(B_j^o(\mu_j))}{\sqrt{n}} - \sqrt{\frac{q_j}{p_j}} G_j \frac{Z_j}{\sqrt{n}} \right) \quad (87)$$

$$i = 1, 2, 3, 4 \quad j = 1, 2, 3, 4$$

Remarque: La dépendance en n de la variable $w_1^n(t, \theta, \theta_0)$ provient de la remarque 2 du théorème 1.

L'argument déjà utilisé (55) entraîne:

$$\sup_{\theta \in C} \sup_{t > 0} \sigma_1^2(t, \theta, \theta_0) < K(C) \quad (88)$$

où $K(C)$ est une constante dépendant du compact C . Les relations (85), (86) et (88) démontrent l'assertion (48). On obtient la relation (50) à partir du lemme 8 et de la relation (86).

On écrit $\hat{\Gamma}_1^2(t, \theta)$ sous forme intégrale:

$$\hat{\Gamma}_1^2(t, \theta) = \int_0^t \theta \hat{H}_A \hat{H}_B d(\hat{F}_1 + \hat{F}_3)$$

Lors de l'étude de $\hat{S}_1(t, \theta)$, on a obtenu:

$$\begin{aligned} \hat{H}_B - H_B &= \phi_1^B (\hat{G}_A - G_A) + \phi_3^B (\hat{G}_B - G_B) + u_B^n \\ \hat{H}_A - H_A &= \phi_1^A (\hat{G}_A - G_A) + \phi_3^A (\hat{G}_B - G_B) + u_A^n \end{aligned}$$

avec:

$$\begin{aligned} \phi_1^B &= -\theta G_B / (\theta G_A + G_B)^2 & \phi_3^B &= \theta G_A / (\theta G_A + G_B)^2 \\ \phi_1^A &= G_B / (\theta G_A + G_B)^2 & \phi_3^A &= -G_A / (\theta G_A + G_B)^2 \end{aligned}$$

On pose:

$$\begin{aligned} \phi_1 &= \phi_2 = H_A \phi_1^B + H_B \phi_1^A \\ \phi_3 &= \phi_4 = H_A \phi_3^B + H_B \phi_3^A \\ u_n &= H_A u_B^n + H_B u_A^n + (\hat{H}_A - H_A)(\hat{H}_B - H_B) \end{aligned}$$

On linéarise $\hat{H}_A \hat{H}_B - H_A H_B$ suivant (2) pour obtenir la décomposition (4):

$$\begin{aligned} \hat{H}_A \hat{H}_B - H_A H_B &= H_A (\hat{H}_B - H_B) + H_B (\hat{H}_A - H_A) + (\hat{H}_A - H_A)(\hat{H}_B - H_B) \\ &= \phi_1 (\hat{G}_A - G_A) + \phi_3 (\hat{G}_B - G_B) + u^n \end{aligned}$$

Soit $\lambda(\theta) = \max(1, 1/\theta)$. On a:

$$\sup_{t>0} \max(|H_A(t, \theta)|, |H_B(t, \theta)|) \leq \lambda(\theta) \quad (89)$$

$$\sup_{t>0} \max(|\frac{dH_A}{dt}(t, \theta) \times \frac{G(t)}{G'(t)}|, |\frac{dH_B}{dt}(t, \theta) \times \frac{G(t)}{G'(t)}|) \leq 2\theta\lambda(\theta) \quad (90)$$

$$\sup_{t>0} \max(|\frac{dH_A}{d\theta}(t, \theta)|, |\frac{dH_B}{d\theta}(t, \theta)|) \leq (\lambda(\theta))^2 \quad (91)$$

Avec (89), (90), (91) et les majorations déjà obtenues lors de l'étude de $\hat{S}_1(t, \theta)$ pour les expressions:

$$\sup_{\theta \in \mathcal{C}} \sup_{t>0} \max_{C \in \{A, B\}} (|\phi_1^C G|, |\phi_3^C G|) \quad (92)$$

$$\sup_{\theta \in \mathcal{C}} \sup_{t>0} \max_{C \in \{A, B\}} (|\frac{d\phi_1^C}{dt}(t, \theta) \times \frac{G^2(t)}{G'(t)}|, |\frac{d\phi_3^C}{dt}(t, \theta) \times \frac{G^2(t)}{G'(t)}|) \quad (93)$$

$$\sup_{\theta \in \mathcal{C}} \sup_{t>0} \max_{C \in \{A, B\}} (|\frac{d\phi_1^C}{d\theta}(t, \theta) \times G(t)|, |\frac{d\phi_3^C}{d\theta}(t, \theta) \times G(t)|) \quad (94)$$

on trouve:

$$\begin{aligned} \|\epsilon_1(\mathcal{C})\|_\infty &\leq 2 \sup_{\theta \in \mathcal{C}} \max(\theta, \theta^{-2}) \\ \|\epsilon_2(\mathcal{C})\|_\infty &\leq 10 \sup_{\theta \in \mathcal{C}} \max(\theta^2, \theta^{-3}) \\ A(\mathcal{C}) &= 4 \sup_{\theta \in \mathcal{C}} \max(\theta, \theta^{-3}) \end{aligned}$$

et les conditions A'_1, A'_2, A'_3, A'_4 sont satisfaites.

On développe alors la statistique comme en (4). L' inégalité du type (70) est encore valable. Il reste à contrôler R_2^{*n} et R_3^{*n} avec:

$$R_2^{*n} \leq R_2'^n + R_2''^n \quad R_3^{*n} \leq R_3'^n + R_3''^n$$

$$\text{où } \left\{ \begin{array}{l} R_2'^n(t, \theta) = \int_0^{M_n(t)} |H_A u_B + H_B u_A| \, d\hat{F}_1 \\ R_2''^n = \int_0^{M_n(t)} |(\hat{H}_A - H_A)(\hat{H}_B - H_B)| \, dF_1 \\ R_3'^n = \int_{M_n(t)}^t |H_A u_B + H_B u_A| \, dF_1 \\ R_3''^n = \int_{M_n(t)}^t |(\hat{H}_A - H_A)(\hat{H}_B - H_B)| \, dF_1 \end{array} \right.$$

(Les termes en $\hat{F}_3$ et F_3 sont du même type)

Comme $\sup_{t>0} \max(\hat{H}_A(t, \theta), H_A(t, \theta), \hat{H}_B(t, \theta), H_B(t, \theta)) \leq \lambda(\theta)$, les termes $R_2'^n$ et $R_3'^n$

se contrôlent comme les termes R_2^n et R_3^n lors de l'étude de $\hat{S}_1(t, \theta)$. Posons $\lambda(C) = \sup_{\theta \in C} \max(1, \theta^{-1})$. On a immédiatement:

$$\sup_{\theta \in C} \sup_{t > 0} R_3^{''n}(t, \theta) \leq 4\lambda(C)G_1(\beta_n) \leq 11\lambda(C)r_n \quad \text{p.co}$$

En utilisant les égalités:

$$\begin{aligned} \hat{H}_B - H_B &= \theta[\hat{H}_A(\hat{G}_B - G_B) - \hat{H}_B(\hat{G}_A - G_A)]/(\theta G_A + G_B) \\ \hat{H}_A - H_A &= [\hat{H}_B(\hat{G}_A - G_A) - \hat{H}_A(\hat{G}_B - G_B)]/(\theta G_A + G_B) \end{aligned}$$

ainsi que les relations $|ab| \leq (a^2 + b^2)/2$, $(a + b)^2 \leq 2(a^2 + b^2)$ et l'inégalité

$$\sup_{\theta \in C} \sup_{t > 0} \max(\hat{H}_A(t, \theta), \hat{H}_B(t, \theta)) \leq \lambda(C)$$

on obtient:

$$\forall t > 0 \quad |R_2^{''n}(t, \theta)| \leq K_1(C) \int_0^{M_n(t)} \left[\left(\frac{\hat{G}_A - G_A}{\theta G_A + G_B} \right)^2 + \left(\frac{\hat{G} - G}{\theta G_A + G_B} \right)^2 \right] d\hat{F}_1$$

où $K_1(C)$ est une constante dépendant du compact C . Ce terme est identique au terme R_2^n de l'étude de $\hat{S}_1(t, \theta)$.

Finalement, on obtient:

$$\sup_{\theta \in C} \sup_{t > 0} \left| \frac{R_2^{*n}(t) + R_3^{*n}(t)}{\inf(1 + \log(1/G(t)), \log n)} \right| \leq K(C)r_n \quad \text{p.co}$$

où $K(C)$ est une constante dépendant du compact C .

Le reste de la preuve est identique à celle du théorème 4.

5 Application à l'estimateur du maximum de vraisemblance de Cox

Avec les notations de l'introduction et en particulier de la section 1.2 (Cas des statistiques de rang) la pseudo vraisemblance de Cox est définie par:

$$\begin{aligned} V_n(\beta) &= \prod_{i=1}^n \left(\frac{\exp(\beta \delta_i)}{\sum_{j \in \mathcal{R}_i} \exp(\beta \delta_j)} \right)^{d_i} \\ &= \exp \left(\sum_{i=1}^n d_i (\delta_i \beta - \log(n_{1i} \exp \beta + n_{2i})) \right) \end{aligned}$$

Nous notons:

$$L_n(\beta) = \log(V_n(\beta)) = \sum_{i=1}^n d_i (\delta_i \beta - \log(n_{1i} \exp(\beta) + n_{2i}))$$

$$\text{grad} L_n(\beta) = \sum_{i=1}^n d_i \left(\delta_i - \frac{n_{1i} \exp(\beta)}{n_{1i} \exp(\beta) + n_{2i}} \right)$$

Soit $\hat{\beta}_n$ l'estimateur du maximum de vraisemblance de Cox qui vérifie:

$$\begin{aligned} L_n(\hat{\beta}_n) &= \max_{\beta} L_n(\beta) \\ \text{grad} L_n(\hat{\beta}_n) &= 0 \end{aligned}$$

Nous donnerons également les résultats pour l'estimateur:

$$\hat{\theta}_n = \exp(\hat{\beta}_n)$$

Cet estimateur maximise la fonction $L_n(\log \theta)$ et annule la fonction $\text{grad} L_n(\log \theta)$.

Consistance des estimateurs

Andersen et Gill [1] ont montré la convergence en probabilité de $\hat{\beta}_n$ vers β_0 , la convergence presque sûre en est une conséquence directe par des arguments de concavité. (voir également [1], appendice 2). La convergence de $\hat{\theta}_n$ vers θ_0 est immédiate. Nous utiliserons ici le résultat dans sa version presque sûre, soit:

Lemme 14

$$\begin{aligned} \lim_{n \rightarrow \infty} |\hat{\beta}_n - \beta_0| &= 0 \quad \mathbf{P}_{\beta_0} \text{ p.s.} \\ \lim_{n \rightarrow \infty} |\exp(\hat{\beta}_n) - \exp(\beta_0)| &= 0 \quad \mathbf{P}_{\beta_0} \text{ p.s.} \end{aligned}$$

Le lemme 15 précise la vitesse de convergence de ces estimateurs.

Lemme 15

$$\begin{aligned} \overline{\lim}_{n \rightarrow \infty} \sqrt{(n/\log n)} |\hat{\beta}_n - \beta_0| &\leq K \quad \mathbf{P}_{\beta_0} \text{ p.s.} \\ \overline{\lim}_{n \rightarrow \infty} \sqrt{(n/\log n)} |\exp(\hat{\beta}_n) - \exp(\beta_0)| &\leq K' \quad \mathbf{P}_{\beta_0} \text{ p.s.} \end{aligned}$$

où $K = K(\beta_0)$ et $K' = K'(\beta_0)$ sont des constantes.

Approximation forte des estimateurs

Theoreme 7 Sous β_0 , on peut construire une suite W^n de variables gaussiennes centrées et de variance $(\sigma^2(\beta_0))^{-1}$ telle que;

$$\begin{aligned} \overline{\lim}_{n \rightarrow \infty} \frac{n}{\log^2 n} \left| \hat{\beta}_n - \beta_0 - \frac{W^n(\beta_0)}{\sqrt{n}} \right| &\leq K \quad \mathbf{P}_{\beta_0} \text{ p.s.} \\ \overline{\lim}_{n \rightarrow \infty} \frac{n}{\log^2 n} \left| \exp(\hat{\beta}_n) - \exp(\beta_0) - \frac{W^n(\beta_0)}{\exp(\beta_0)\sqrt{n}} \right| &\leq K' \quad \mathbf{P}_{\beta_0} \text{ p.s.} \end{aligned}$$

où $K = K(\beta_0)$ et $K' = K'(\beta_0)$ sont des constantes et:

$$\sigma^2(\beta_0) = \int_0^\infty \frac{\mathbf{P}(\delta = 0)\mathbf{P}(\delta = 1)\exp(\beta_0)C_1S_1C_2S_2}{\exp(\beta_0)\mathbf{P}(\delta = 1)C_1S_1 + \mathbf{P}(\delta = 0)C_2S_2} dt$$

Preuve du lemme 15

Sous β_0 , d'après le théorème 5, on peut construire une suite $(w_1^n(\infty, \exp(\beta), \exp(\beta_0)))_n$ de processus gaussiens centrés et de variance $\sigma_1^2(\infty, \exp(\beta), \exp(\beta_0))$ tels que:

$$\frac{\text{grad} L_n(\beta)}{n} = m_1(\infty, \exp(\beta), \exp(\beta_0)) + \frac{w_1^n(\infty, \exp(\beta), \exp(\beta_0))}{\sqrt{n}} + \rho_n(\infty, \exp(\beta), \exp(\beta_0)) \quad (95)$$

avec

$$m_1(\infty, \exp(\beta), \exp(\beta_0)) = \mathbf{P}(\delta = 0)\mathbf{P}(\delta = 1) \int_0^\infty \frac{(\exp(\beta_0) - \exp(\beta))C_1 S_1 C_2 S_2}{\exp(\beta)\mathbf{P}(\delta = 1)C_1 S_1 + \mathbf{P}(\delta = 0)C_2 S_2} ds$$

et pour tout compact $\mathcal{C}$:

$$\sup_{\beta \in \mathcal{C}} |\rho_n(\infty, \exp(\beta), \exp(\beta_0))| \leq K_1(\mathcal{C}) \frac{\log^2 n}{n} \quad \text{p.co} \quad (96)$$

Rappelons la relation (50):

$$\sup_{\beta \in \mathcal{C}} \left| \frac{w_1^n(\infty, \exp(\beta), \exp(\beta_0))}{\sqrt{n}} \right| \leq K_2(\mathcal{C}) \sqrt{\frac{\log n}{n}} \quad \text{p.co} \quad (97)$$

En réunissant (95), (96), (97) on obtient:

$$\sup_{\beta \in \mathcal{C}} \left| \frac{\text{grad} L_n(\beta)}{n} - m_1(\infty, \exp(\beta), \exp(\beta_0)) \right| \leq K_3(\mathcal{C}) \sqrt{r_n} \quad \text{p.co} \quad (98)$$

Par définition, $\hat{\beta}_n$ annule $\text{grad} L_n(\beta)$. D'autre part $m_1(\infty, \exp(\beta), \exp(\beta_0))$ ne s'annule qu'en $\beta = \beta_0$ et la fonction dérivée est strictement négative et bornée inférieurement sur le compact $\mathcal{C}$. Posons:

$$m(\mathcal{C}) = \inf_{\beta \in \mathcal{C}} \left| \frac{dm_1}{d\beta}(\infty, \exp(\beta), \exp(\beta_0)) \right|$$

Dés que $\hat{\beta}_n \in \mathcal{C}$ on a:

$$\begin{aligned} m(\mathcal{C})|\hat{\beta}_n - \beta_0| &\leq |m_1(\infty, \exp(\hat{\beta}_n), \exp(\beta_0)) - m_1(\infty, \exp(\beta_0), \exp(\beta_0))| \\ &\leq \sup_{\beta \in \mathcal{C}} \left| \frac{\text{grad} L_n(\beta)}{n} - m_1(\infty, \exp(\beta), \exp(\beta_0)) \right| \\ &\leq K_3(\mathcal{C}) \sqrt{r_n} \quad \text{p.co} \end{aligned} \quad (99)$$

où $K_3(\mathcal{C})$ est la constante de (98). En prenant $\mathcal{C} = [\beta_0 - 1, \beta_0 + 1]$ et $K(\mathcal{C}) = K_3(\mathcal{C})/m(\mathcal{C})$ on a:

$$\mathbf{1}(|\hat{\beta}_n - \beta_0| \geq K(\mathcal{C})\sqrt{r_n}) \leq \mathbf{1}(\hat{\beta}_n \notin \mathcal{C}) + \mathbf{1}(|\hat{\beta}_n - \beta_0| \geq K(\mathcal{C})\sqrt{r_n}, \hat{\beta}_n \in \mathcal{C})$$

Le lemme 14 et la relation (99) entraînent alors le résultat.

Preuve du théorème 7

Lorsque $\beta = \beta_0$, la relation (95) devient:

$$\frac{\text{grad} L_n(\beta)}{n} = \frac{w_1^n(\infty, \exp(\beta_0), \exp(\beta_0))}{\sqrt{n}} + \rho_n(\infty, \exp(\beta_0), \exp(\beta_0)) \quad (100)$$

avec:

$$|\rho_n(\infty, \exp(\beta_0), \exp(\beta_0))| \leq K_1 r_n(\log n) \quad \text{p.co}$$

et l'on connaît la forme explicite de $\sigma_1^2(\infty, \exp(\beta_0), \exp(\beta_0))$ (théorème 5). Posons:

$$\begin{aligned} \sigma^2(\beta_0) &= \sigma_1^2(\infty, \exp(\beta_0), \exp(\beta_0)) \\ &= \mathbf{P}(\delta = 0)\mathbf{P}(\delta = 1) \int_0^\infty \frac{\exp(\beta_0)C_1S_1C_2S_2}{\exp(\beta_0)\mathbf{P}(\delta = 1)C_1S_1 + \mathbf{P}(\delta = 0)C_2S_2} ds \end{aligned}$$

et:

$$W^n(\beta_0) = \frac{w_1^n(\infty, \exp(\beta_0), \exp(\beta_0))}{\sigma^2(\beta_0)}$$

Soit $C_n = [\beta_0 - K_2\sqrt{r_n}, \beta_0 + K_2\sqrt{r_n}]$ où K_2 est la constante du lemme 15:

$$\lim_{n \rightarrow \infty} \mathbf{1}(\hat{\beta}_n \notin C_n) = 0 \quad \mathbf{P}_{\beta_0} \text{ p.s.} \quad (101)$$

Avec le développement de Taylor d'ordre 1:

$$\frac{\text{grad} L_n(\beta_0)}{n} = -\frac{1}{n} \left[\frac{d^2 L_n}{d\beta^2}(\beta_n^*) \right] (\hat{\beta}_n - \beta_0)$$

on définit β_n^* strictement compris entre $\hat{\beta}_n$ et β_0 . On obtient:

$$\begin{aligned} \mathbf{1}(|\hat{\beta}_n - \beta_0 - \frac{W^n(\beta_0)}{\sqrt{n}}| \geq K \frac{\log^2 n}{n}) \\ \leq \mathbf{1}(\beta_n^* \notin C_n) + \mathbf{1}(|\hat{\beta}_n - \beta_0 - \frac{W^n(\beta_0)}{\sqrt{n}}| \geq K \frac{\log^2 n}{n}, \beta_n^* \in C_n) \end{aligned} \quad (102)$$

La relation (101) entraîne:

$$\lim_{n \rightarrow \infty} \mathbf{1}(\beta_n^* \notin C_n) = 0 \quad \mathbf{P}_{\beta_0} \text{ p.s.} \quad (103)$$

Notons:

$$P_n = \mathbf{P}_{\beta_0}(|\hat{\beta}_n - \beta_0 - \frac{W^n(\beta_0)}{\sqrt{n}}| \geq K \frac{\log^2 n}{n}, \beta_n^* \in C_n)$$

Sous β_0 d'après le théorème 5 on peut construire une suite $v_1^n(\infty, \exp(\beta), \exp(\beta_0))$ de processus gaussiens centrés et de variance $\gamma_1^2(\infty, \exp(\beta), \exp(\beta_0))$ tels que:

$$\frac{-1}{n} \frac{d^2 L_n}{d\beta^2}(\beta) = \gamma_1^2(\infty, \exp(\beta), \exp(\beta_0)) + \frac{v_1^n(\infty, \exp(\beta), \exp(\beta_0))}{\sqrt{n}} + \bar{\rho}_n(\infty, \exp(\beta), \exp(\beta_0)) \quad (104)$$

avec:

$$\begin{aligned} & \gamma_1^2(\infty, \exp(\beta), \exp(\beta_0)) \\ &= \mathbf{P}(\delta = 0) \mathbf{P}(\delta = 1) \int_0^1 \frac{\exp(\beta) C_1 S_1 C_2 S_2 \{ \exp(\beta_0) \mathbf{P}(\delta = 1) C_1 S_1 + \mathbf{P}(\delta = 0) C_2 S_2 \}}{\{ \exp(\beta) \mathbf{P}(\delta = 1) C_1 S_1 + \mathbf{P}(\delta = 0) C_2 S_2 \}^2} ds \end{aligned}$$

et pour tout compact $\mathcal{C}$:

$$\sup_{\beta \in \mathcal{C}} |\bar{\rho}_n(\infty, \exp(\beta), \exp(\beta_0))| \leq K_3(\mathcal{C}) \frac{\log^2 n}{n} \quad \text{p.co} \quad (105)$$

Rappelons la seconde inégalité de (50):

$$\sup_{\beta \in \mathcal{C}} \left| \frac{v_1^n(\infty, \exp(\beta), \exp(\beta_0))}{\sqrt{n}} \right| \leq K_4(\mathcal{C}) \sqrt{\frac{\log n}{n}} \quad \text{p.co} \quad (106)$$

En réunissant (104), (105), (106) on obtient:

$$\sup_{\beta \in \mathcal{C}} \left| \frac{-1}{n} \left[\frac{d^2 L_n}{d\beta^2}(\beta) \right] - \gamma_1^2(\infty, \exp(\beta), \exp(\beta_0)) \right| \leq K_5(\mathcal{C}) \sqrt{\tau_n} \quad \text{p.co} \quad (107)$$

et en particulier:

$$\forall \beta \in \mathcal{C} \quad \frac{-1}{n} \left[\frac{d^2 L_n}{d\beta^2}(\beta) \right] \geq \frac{\gamma_1^2(\infty, \exp(\beta), \exp(\beta_0))}{2} \quad \text{p.co} \quad (108)$$

D'autre part, on a:

$$\sup_{\beta \in \mathcal{C}_n} |\gamma_1^2(\infty, \exp(\beta), \exp(\beta_0)) - \sigma^2(\beta_0)| \leq K_6 \sqrt{\tau_n} \quad (109)$$

où K_6 dépend de la constante K_2 . Posons maintenant $\mathcal{C} = [\beta_0 - K_2, \beta_0 + K_2]$. Les relations (108) et (109) entraînent:

$$\forall \beta \in \mathcal{C}_n \quad \frac{-1}{n} \left[\frac{d^2 L_n}{d\beta^2}(\beta) \right] \geq \frac{\sigma^2(\beta_0)}{4} \quad \text{p.co} \quad (110)$$

La relation (110) entraîne:

$$\begin{aligned} P_n &= \mathbf{P}_{\beta_0} \left(\frac{-1}{n} \left[\frac{d^2 L_n}{d\beta^2}(\beta_n^*) \right] |\hat{\beta}_n - \beta_0 - \frac{W^n(\beta_0)}{n}| \geq \frac{-1}{n} \left[\frac{d^2 L_n}{d\beta^2}(\beta_n^*) \right] \frac{K \log^2 n}{n}, \beta_n^* \in \mathcal{C}_n \right) \\ &\leq Q_n + U_n \end{aligned} \quad (111)$$

avec $\sum_n U_n < \infty$ et:

$$Q_n = \mathbf{P}_{\beta_0} \left(\left| \frac{-1}{n} \left[\frac{d^2 L_n}{d\beta^2}(\beta_n^*) \right] \hat{\beta}_n - \beta_0 - \frac{W^n(\beta_0)}{n} \right| \geq \frac{\sigma^2(\beta_0)}{4} \frac{K \log^2 n}{n}, \beta_n^* \in \mathcal{C}_n \right)$$

On obtient:

$$Q_n = Q_n^1 + Q_n^2 + Q_n^3 \quad (112)$$

avec:

$$Q_1^n = \mathbf{P}_{\beta_0} \left(\left| \frac{\text{grad} L_n(\beta_0)}{n} - \frac{\sigma^2(\beta_0) W^n(\beta_0)}{\sqrt{n}} \right| \geq \frac{\sigma^2(\beta_0)}{12} \frac{K \log^2 n}{n} \right)$$

$$Q_2^n = \mathbf{P}_{\beta_0} \left(\left| \frac{W^n(\beta_0)}{\sqrt{n}} \right| \left| \sigma^2(\beta_0) - \gamma_1^2(\infty, \exp(\beta_n^*), \exp(\beta_0)) \right| \geq \frac{\sigma^2(\beta_0)}{12} \frac{K \log^2 n}{n}, \beta_n^* \in \mathcal{C}_n \right)$$

$$Q_3^n = \mathbf{P}_{\beta_0} \left(\left| \frac{W^n(\beta_0)}{\sqrt{n}} \right| \left| \frac{-1}{n} \left[\frac{d^2 L_n}{d\beta^2}(\beta_n^*) \right] - \gamma_1^2(\infty, \exp(\beta_n^*), \exp(\beta_0)) \right| \geq \frac{\sigma^2(\beta_0)}{12} \frac{K \log^2 n}{n}, \beta_n^* \in \mathcal{C} \right)$$

En choisissant $K \geq 12K_1/\sigma^2(\beta_0)$, la relation (100) entraine:

$$Q_1^n \leq V_n \quad \text{avec} \quad \sum_n V_n < \infty$$

En choisissant $K \geq 24 \max(K_6, K_5(C))/\sigma^3(\beta_0)$, le lemme 10 et les relations (107), (109) entraînent:

$$Q_2^n + Q_3^n \leq R_n \quad \text{avec} \quad \sum_n R_n < \infty$$

En reportant les deux dernières inégalités dans (112) et (111) on obtient:

$$\lim_{n \rightarrow \infty} \mathbf{1} \left(\left| \hat{\beta}_n - \beta_0 - \frac{W^n(\beta_0)}{n} \right| \geq K \frac{\log^2 n}{n}, \beta_n^* \in \mathcal{C}_n \right) = 0 \quad \mathbf{P}_{\beta_0} \text{ p.s.} \quad (113)$$

On conclut avec (102), (103) et (113).

Avec le développement de Taylor:

$$\frac{\text{grad} L_n(\beta_0)}{n} = \frac{1}{\exp(\beta_n^*)} \left(\frac{-1}{n} \left[\frac{d^2 L_n}{d\beta^2}(\beta_n^*) \right] (\exp(\hat{\beta}_n) - \exp(\beta_0)) \right)$$

où β_n^* est strictement compris entre $\hat{\beta}_n$ et β_0 , on obtient de même l'approximation forte de $\exp(\hat{\beta}_n)$.

References

- [1] ANDERSEN P.K. and GILL R.D. (1982) *Cox's regression model for counting processes: a large sample study*. Annals of Statistics, vol 10, 4, 1100-1120.
- [2] BENNETT G. (1962) *Probability inequalities for the sum of independants random variables*. J. AM. Statis. Assoc., 57, 33-45.
- [3] BRETAGNOLLE J. and MASSART P. (1989) *Hungarian construction from the non asymptotic view point*. Annals of Probability, vol 10, 239-256.
- [4] BURKE M.D., CSÖRGÖ S., HORVÁTH L. (1981) *Strong Approximations of Some Biometric Estimates under Random Censorship*. Z. Warschein. Verw. Geb., 56, 87-112.
- [5] CASTELLE N. *Approximation forte d'un vecteur de processus empiriques*. Voir 1ère partie, chapitre deux de cette thèse.
- [6] COX D.R and OAKES D. (1984) *Analysis of Survival Data*. Chapman & Hall.
- [7] DVORETZKY A., KIEFER J.C. and WOLFOWITZ J. (1956) *Asymptotic minimax character of the sample distribution function and of the classical multinomial estimator*. Ann. Math. Stat., 33, 642-669.
- [8] EINMAHL J.H.J. and KONING A.J. (1989) *Limits theorems for a general weighted process under random censoring with applications*. Medical Informatics and Statistics Report, 24. (University of Limburg)
- [9] GILL R.D. (1980) *Censoring and Stochastic Integrals*. Mathematical Centre Tracts 124. Amsterdam: Mathematische Centre.
- [10] GILL R.D. (1983) *Large sample behaviour of the product-limit estimator on the whole line*. Annals of Statistics, vol 11, 49-58.
- [11] HARRINGTON D. and FLEMING T. (1982) *A class of rank test procedures for censored survival data*. Biometrika, 69, 3, 553-566.
- [12] MASSART P., 1990. *The tight constant of the Dvoretzky-Kiefer-Wolfowitz inequality*. A paraitre dans Annals of probability.
- [13] SHORACK G.R. & WELLNER J.A. (1986) *Empirical processes with applications to statistics*. Wiley & Sons.
- [14] WELLNER J. (1978) *Limit theorems for the ratio of the empirical distribution function to the true distribution function*. Z. Warschein. Verw. Geb., 45, 73-88.