Un modèle de Markov caché en assurance et Estimation de frontière et de point terminal

Gilles Stupfler

INSTITUT DE RECHERCHE MATHÉMATIQUE AVANCÉE
Université de Strasbourg et CNRS (UMR 7501)
7, rue René Descartes
F-67084 Strasbourg CEDEX

Un modèle de Markov caché en assurance et Estimation de frontière et de point terminal

Gilles Stupfler

Thèse soutenue publiquement le 10 novembre 2011 devant le jury composé de :

Hansjoerg Albrecher	Rapporteur
Jean-Noël Bacro	Examinateur
Laurent Gardes	Examinateur
Irène Gijbels	Rapporteur
Stéphane Girard	Directeur de thèse
Armelle Guillou	Directeur de thèse
Stéphane Loisel	Examinateur

Mots-clés : processus de Poisson, processus de Markov, estimateur du maximum de vraisemblance, consistance, algorithme EM, estimation de point terminal, méthode des moments d'ordre élevé, normalité asymptotique, estimation de frontière.

Classification mathématique : 62F10, 62F12, 62G05, 62G20, 62G32, 62M05, 62P05.

Remerciements

Je tiens tout d'abord à remercier mes deux directeurs de thèse, Stéphane Girard et Armelle Guillou. Jamais avares de conseils avisés, toujours attentifs, critiques aussi bien sur la forme que sur le fond, disponibles, patients et précis, rien de tout cela n'aurait été possible sans eux. Merci infiniment !

Je suis très honoré que Hansjoerg Albrecher et Irène Gijbels aient accepté de rapporter ma thèse. Je suis aussi reconnaissant à Jean-Noël Bacro et Laurent Gardes pour leur participation à mon jury. Merci à eux pour l'intérêt qu'ils ont porté à mon travail et pour leurs remarques constructives.

Je suis également heureux de compter Stéphane Loisel, avec qui j'ai eu l'opportunité de travailler pendant cette thèse, parmi les membres de mon jury. Merci beaucoup pour les nombreuses discussions que nous avons eues ensemble, les précieux coups de main et l'accueil toujours chaleureux à Lyon.

Un grand merci à Alexandre You pour son aide dans la construction et l'exploitation des divers jeux de données que j'ai été amené à utiliser. La partie de ce travail sur les modèles de Markov cachés a beaucoup bénéficié de son expertise.

Ma thèse a été l'occasion de travailler au sein de l'équipe de Statistique de l'IRMA. Merci à tous leurs membres et en particulier à Frédéric, Myriam et aux deux Nicolas.

Mille mercis à Anne-Laure Fougères et Cécile Mercadier pour m'avoir permis de faire un exposé à la conférence EVA 2011 à Lyon. Merci à El-Hadji, Emilie, Jonathan et Ludovic d'en avoir fait une semaine vraiment pas comme les autres.

Il ne serait pas possible de travailler à l'université sans eux : merci aux personnels administratifs de l'université de Strasbourg, et en particulier à Claudine Bernard, Claudine Bonnin et Simone Thouvenin pour leur aide sans faille. Merci également à tous les agents de la bibliothèque pour leur disponibilité et leur efficacité ainsi qu'à Alexis Palaticky et Alain Sartout pour leurs coups de main informatiques et autres déblocages d'imprimantes récalcitrantes.

Pour les pauses thé, les nombreuses discussions à propos de tout et de rien que nous avons eues ensemble et parce que sans lui le 416 ne serait pas le 416, merci à mon co-bureau Antoine. Merci également aux doctorants du premier étage, en particulier à Alex pour les sorties ciné et Auguste et Aurélien pour les matches de foot.

Parce qu'ils me soutiennent et me subissent aussi, merci à ma famille : mes parents, mon petit frère, tante Claudine et oncle Serge, tante Soeur, Niki et Anne-Marie.

Et enfin, pour les soirées, les week-ends, les balades, les bières, les concerts, les délires, les discussions à coeur ouvert, les coups de coeur et les coups de gueule, les fous rires et les larmes et tous ces moments inoubliables vécus ensemble qui ont fait que ces années ont passé si vite, merci à mes amis Adrien, Ambroise, Anaïs, Christophe, Daniel, Hichem, Marie, Maurice, Médéric, Olivier, Philippe, Sandra et Sinan.

A Jacques

Publications et conférences

Articles soumis :

Guillou, A., Loisel, S., Stupfler, G. (2011). Estimation of the parameters of a Markov-modulated loss process in insurance. (Chapitre 1)
Disponible sur HAL : http://hal.archives-ouvertes.fr/hal-00589696/fr.

Girard, S., Guillou, A., Stupfler, G. (2011). Estimating an endpoint with high order moments. (Chapitre 2)
A paraître dans Test, disponible sur HAL : http://hal.archives-ouvertes.fr/inria-00596979/fr.

Girard, S., Guillou, A., Stupfler, G. (2011). Frontier estimation with kernel regression on high order moments. (Chapitre 3)
En révision, disponible sur HAL : http://hal.archives-ouvertes.fr/hal-00499369/fr.

Conférence internationale :

7th Conference on Extreme Value Analysis, Probabilistic and Statistical Models and their Applications, 27 juin-1er juillet 2011, Lyon. Titre de l'exposé : Estimating an endpoint using high order moments.

Table des matières

Introduction

Cette thèse est divisée en deux parties. Dans un premier temps (Chapitre 1), on se concentre sur la présentation d'un nouveau modèle, basé sur un processus de Poisson à modulation markovienne, pour décrire un processus de pertes en assurance. On expose également une technique pour l'estimation des paramètres de ce modèle. Dans une seconde partie indépendante, on considère un problème de théorie des valeurs extrêmes : l'estimation du point terminal (Chapitre 2) ou de la frontière (Chapitre 3) d'une distribution. On décrit une méthode d'estimation utilisant des moments d'ordre élevé de l'échantillon. Avant de présenter les résultats obtenus, on rappelle quelques éléments de base nécessaires pour ces travaux.

Première partie : Processus de pertes à modulation markovienne en assurance

Contexte général

La plupart des applications de la théorie des probabilités utilisent les processus stochastiques. Parmi eux, on peut distinguer les processus à temps continu, qui sont utilisés notamment pour

- l'actuariat, lorsqu'on étudie les réserves d'une compagnie d'assurances ;
- la finance, pour le calcul du cours d'une action sur les marchés financiers ;
- les télécommunications, pour modéliser le nombre de requêtes reçues par un serveur ;
- la biologie, afin d'étudier l'évolution d'une population au cours du temps.

Dans la première partie de ce travail, on s'intéresse à une généralisation de ce qu'on peut considérer comme étant l'exemple le plus simple de processus à temps continu : le processus de Poisson homogène. Ce processus est défini de la façon suivante (voir Çinlar [25]) :

Définition. Un *processus de Poisson homogène* est un processus à temps continu $(N(t))_{t\geq 0}$, à valeurs réelles et non identiquement nul, tel que

- $N(0) = 0$ et les trajectoires de N sont des fonctions en escalier continues à droite ;
- les sauts de N sont égaux à 1 presque sûrement ;
- N est à *accroissements indépendants* : pour tous $0 < s < t$, la variable aléatoire $N(t) - N(s)$ est indépendante de $\{N(u),\, u \leq s\}$;
- N est à *accroissements stationnaires* : pour tous $0 < s < t$, la variable aléatoire $N(t) - N(s)$ a la même loi que la variable aléatoire $N(t - s)$.

Le résultat suivant résume quelques propriétés fondamentales de ces processus. On renvoie le lecteur à [25], Théorème 1.9 p. 74 et Théorème 2.5 p. 80 pour plus de détails.

Proposition. *Soit N un processus de Poisson homogène. Alors il existe $\lambda > 0$ tel que :*

- *Pour tout $t > 0$, $N(t)$ suit une loi de Poisson de paramètre λt.*
- *En définissant les instants de saut de N par*

$$\tau_0 = 0, \quad \textit{et} \ \ \forall\, n \in \mathbb{N}, \ \ \tau_{n+1} = \inf\{t > \tau_n \,|\, N(t) > n\},$$

  alors $(\tau_{n+1} - \tau_n)_{n \geq 0}$ est une suite de variables aléatoires indépendantes de même loi exponentielle de paramètre λ.

Le réel λ est appelé intensité du processus de Poisson homogène N.

Un processus de Poisson homogène possède donc de bonnes propriétés théoriques, ce qui explique qu'il constitue une excellente approximation pour modéliser un grand nombre de phénomènes. Par exemple, les bases de la théorie du risque, posées par Lundberg en 1903, s'appuyaient sur des processus de Poisson homogènes pour représenter le processus d'arrivée des sinistres: pour davantage d'informations sur ce sujet, nous renvoyons le lecteur à Klugman *et al.* [74] et Rolski *et al.* [94]. Une telle hypothèse permet d'obtenir un grand nombre de résultats et rend le modèle facile d'utilisation. Cependant, les propriétés d'un tel modèle peuvent paraître trop fortes à certains égards. Par exemple, il est assez rare que les temps $\tau_{n+1} - \tau_n$ séparant la survenue de deux évènements soient indépendants de même loi exponentielle. De plus, on peut imaginer que le comportement de N soit influencé par des facteurs externes, de sorte que l'hypothèse de stationnarité ne soit plus vérifiée. Une première tentative pour introduire de l'irrégularité dans le modèle est d'utiliser une définition généralisée du processus de Poisson :

Définition. Un *processus de Poisson* (peut-être inhomogène) est un processus à temps continu $(N(t))_{t \geq 0}$, à valeurs réelles et non identiquement nul, tel que

- $N(0) = 0$ et les trajectoires de N sont des fonctions en escalier continues à droite ;

- les sauts de N sont égaux à 1 presque sûrement ;

- N est à *accroissements indépendants* : pour tous $0 < s < t$, la variable aléatoire $N(t) - N(s)$ est indépendante de $\{N(u),\, u \leq s\}$.

Plus simplement, un tel processus sera appelé *processus de Poisson.* Sa loi est connue (voir [25], Proposition 7.9 p. 97) :

Proposition. *Soient N un processus de Poisson et $t \mapsto \mu(t)$ la fonction définie par $\mu(t) = \mathbb{E}(N(t))$. On suppose que la fonction μ est continue. Alors il existe une fonction positive et localement intégrable λ telle que pour tous $0 < s < t$, $N(t) - N(s)$ suit une loi de Poisson de paramètre*

$$\mu(t) - \mu(s) = \int_s^t \lambda(u)\, du.$$

La fonction λ est appelée intensité du processus de Poisson N.

L'intensité d'un processus de Poisson varie avec le temps : remarquons qu'un tel processus est un processus de Poisson homogène si et seulement si son intensité est constante. Ceci rend possible la construction de modèles plus réalistes : supposons par exemple que l'on souhaite modéliser le nombre de tempêtes affectant une région donnée du monde. La probabilité d'une tempête dépend de la saison dans laquelle on se trouve : partant de ce constat, on peut construire un processus N dont le taux de saut sera élevé pour certaines saisons et faible pour d'autres. Un tel processus n'est par conséquent pas stationnaire et bien sûr, les temps entre les sauts ne suivent pas une loi exponentielle. Ceci étant, les accroissements de N sont encore indépendants. De plus, on peut vouloir construire une dépendance stochastique du taux de saut en des paramètres environnementaux, comme la météorologie ou le contexte économique, par exemple. Dans le modèle précédent, on ne tient notamment pas compte de phénomènes climatologiques comme El Niño ou La Niña, dont on sait qu'ils augmentent la probabilité d'évènements extrêmes comme les tempêtes, mais ne surviennent pas de façon régulière dans l'année. Pour intégrer ce type de dépendance, on peut par exemple considérer un processus de Poisson à modulation markovienne : avant de définir ce type de processus, on rappelle le concept de dépendance markovienne.

Définition. Une *chaîne de Markov* (homogène) sur un espace d'états fini $\{1, \ldots, r\}$, $r \in \mathbb{N} \setminus \{0\}$ est un processus à temps discret (X_n), à valeurs dans $\{1, \ldots, r\}$, tel que

- pour tout $n \in \mathbb{N} \setminus \{0\}$, sachant X_n, X_{n+1} est indépendante de $(X_i)_{i \leq n-1}$;

- il existe une matrice carrée P de taille r telle que

$$\forall\, n \in \mathbb{N}, \quad \forall\, i,\, j \in \{1, \ldots, r\}, \quad \mathbb{P}(X_{n+1} = j \,|\, X_n = i) = P_{ij}.$$

La matrice P est appelée *matrice de transition* de X et la loi π sur $\{1, \ldots, r\}$ définie par

$$\forall\, i \in \{1, \ldots, r\}, \quad \pi_i = \mathbb{P}(X_0 = i)$$

est appelée *loi initiale* de X. Enfin, une chaîne de Markov X avec matrice de transition P est dite *irréductible* si pour tous i et j, il existe $n \in \mathbb{N} \setminus \{0\}$ tel que $(P^n)_{ij} > 0$.

En d'autres termes, la dépendance markovienne se traduit par une indépendance du futur du processus et de son passé conditionnellement à son présent. L'analogue en temps continu d'une chaîne de Markov est appelé processus de Markov :

Définition. Un *processus de Markov* (homogène) sur un espace d'états fini $\{1, \ldots, r\}$, $r \in \mathbb{N} \setminus \{0\}$ est un processus à temps continu $(J(t))_{t \geq 0}$, à valeurs dans $\{1, \ldots, r\}$, tel que

- pour tous $n \in \mathbb{N} \setminus \{0\}$ et $0 < t_1 < \ldots < t_n < t_{n+1}$, sachant $J(t_n)$, $J(t_{n+1})$ est indépendante de $(J(t_i))_{i \leq n-1}$ et sa loi ne dépend que de l'accroissement $t_{n+1} - t_n$. Autrement dit, pour toute fonction $f : \{1, \ldots, r\} \to [0, \infty)$,

$$\mathbb{E}(f(J(t_{n+1})) \,|\, J(t_i),\, 1 \leq i \leq n) = \mathbb{E}(f(J(t_{n+1} - t_n)) \,|\, J(0))\,;$$

- les trajectoires de J sont des fonctions en escalier presque sûrement continues à droite ;
- quand $t \to 0$,

$$\forall\, i, j \in \{1, \ldots, r\}, \quad \mathbb{P}(J(t) = j \,|\, J(0) = i) \to \mathbb{1}_{\{i=j\}}.$$

La loi π définie sur $\{1, \ldots, r\}$ par

$$\forall\, i \in \{1, \ldots, r\}, \quad \pi_i = \mathbb{P}(J(0) = i)$$

est appelée *loi initiale* de J. Le processus J est dit *irréductible* si

$$\forall\, t > 0, \quad \forall\, i, j \in \{1, \ldots, r\}, \quad \mathbb{P}(J(t) = j \,|\, J(0) = i) > 0.$$

Les conditions de la définition précédente assurent qu'il existe une matrice $L = (\ell_{ij})$, appelée *générateur* de J, dont les éléments non diagonaux sont positifs et telle que la somme des éléments de chaque ligne de L vaut 0, de sorte que les temps de saut de J suivent une loi exponentielle, de paramètre $-\ell_{ii} > 0$ quand J est dans l'état i ; dans ce cas, J va dans l'état $j \neq i$ avec probabilité $-\ell_{ij}/\ell_{ii}$. L caractérise donc complètement le comportement de J ; de plus, pour tout $t \geq 0$, si $P(t)$ est la matrice carrée de taille r dont les éléments sont les $\mathbb{P}(J(t) = j \,|\, J(0) = i)$, $i, j \in \{1, \ldots, r\}$, on a $P(t) = \exp(tL)$.

Revenons à notre cadre de travail : grossièrement, un processus de Poisson à modulation markovienne (MMPP, pour Markov-modulated Poisson process) est un processus de Poisson dont l'intensité est contrôlée par un processus de Markov sous-jacent. Plus précisément :

Définition. Un *processus de Poisson à modulation markovienne* est un processus à temps continu $(J(t),\, N(t))_{t\geq 0}$ tel que

- J est un processus de Markov irréductible sur un espace d'états fini $\{1,\, \ldots,\, r\}$, $r \in \mathbb{N} \setminus \{0\}$, dont le générateur est noté L et la loi initiale est notée π ;
- il existe $\lambda_1,\, \ldots,\, \lambda_r \geq 0$ tels que sachant J, le processus N est un processus de Poisson ayant pour intensité $t \mapsto \lambda_{J(t)}$.

Lorsque $(J,\, N)$ est un MMPP, le processus N est un cas particulier des processus de Poisson doublement stochastiques : on renvoie le lecteur à Grandell [51] pour plus d'informations. Les MMPPs servent notamment à modéliser :

- le processus de réserve d'une compagnie d'assurances (Asmussen [5]) ;
- des phénomènes environnementaux, comme la pollution de l'air (Davison et Ramesh [28]) ;
- des réseaux de communication (Heffes et Lucantoni [64], Kawashima et Saito [71]).

Dans la suite, on se donne un MMPP $(J,\, N)$, on note τ_n l'instant du $n-$ième saut de N, $J_n = J(\tau_n)$ l'état de J au moment de ce saut et $Y_n = \tau_n - \tau_{n-1}$ le temps écoulé entre le $(n-1)-$ième et le $n-$ième saut de N, avec par convention $Y_0 = 0$. Enfin, soit Λ la matrice diagonale dont les éléments sont les λ_j, $1 \leq j \leq r$. Dans ce cadre, les temps Y_n ne sont ni indépendants, ni identiquement distribués, puisque l'intensité de N dépend de l'état de J. Ceci étant, la structure de dépendance associée à un tel processus est également markovienne :

Proposition. *Le processus $(J_n,\, Y_n)$ est un processus à renouvellement markovien (MRP, pour Markov renewal process). Autrement dit :*

- *le processus (J_n) est une chaîne de Markov irréductible à espace d'états fini $\{1,\, \ldots,\, r\}$, $r \in \mathbb{N} \setminus \{0\}$;*
- *pour tout $n \in \mathbb{N} \setminus \{0\}$, conditionnellement à J_n, le couple $(J_{n+1},\, Y_{n+1})$ est indépendant de $((J_i,\, Y_i)_{i\leq n-1},\, Y_n)$.*

Pour plus de détails sur les MRPs, on renvoie le lecteur à [25]. Le comportement d'un MMPP à l'instant t dépend donc seulement de l'état de J à l'instant t et au moment du dernier saut de N : ceci s'interprète comme une structure de dépendance markovienne, qui provient du caractère markovien de J.

Afin de pouvoir développer des techniques d'inférence pour de tels processus, il est important de savoir calculer leur loi. Pour cela, la première étape consiste à déterminer l'expression de la densité conditionnelle de Y_1 sachant l'état initial (voir par exemple Meier-Hellstern [85]) :

Proposition. *Posons, pour tous $t > 0$ et $i, j \in \{1, \ldots, r\}$,*

$$f_{ij}(t)\, dt := \mathbb{P}(Y_1 \in dt,\ J(t) = j \mid J(0) = i).$$

Alors, si $f(t)$ est la matrice carrée de taille r dont les éléments sont les $f_{ij}(t)$, on a

$$f(t) = \exp(t(L - \Lambda))\Lambda.$$

Avant d'aller plus loin, remarquons qu'en intégrant f, on obtient la matrice de transition de (J_n) :

$$P = (\Lambda - L)^{-1}\, \Lambda.$$

D'après Rydén [97], cette matrice possède une unique loi stationnaire π, telle que π_i est strictement positif si et seulement si λ_i l'est. Si la loi initiale de J est exactement π, alors le processus (J_n, Y_n) est stationnaire et par conséquent

$$\forall\, n \in \mathbb{N} \setminus \{0\}, \quad \forall\, y > 0, \quad \mathbb{P}(Y_n \in dy,\, J_n = j \mid J_{n-1} = i) = f_{ij}(y)\, dy.$$

Cette identité est l'argument clé pour calculer la loi du processus (J_n, Y_n), grâce au résultat suivant (voir [25], Proposition 1.10 p. 314) :

Proposition. *Soit (J_n, Y_n) un MRP. Supposons que pour tous i, j, il existe une fonction positive et localement intégrable f_{ij} telle que*

$$\forall\, n \in \mathbb{N} \setminus \{0\}, \quad \forall\, y > 0, \quad \mathbb{P}(Y_n \leq y,\, J_n = j \mid J_{n-1} = i) = \int_0^y f_{ij}(u)\, du.$$

Alors, pour tous j_0, $(j_1, y_1), \ldots, (j_n, y_n)$, on a

$$\mathbb{P}\left(J_0 = j_0,\, \bigcap_{i=1}^{n} \{J_i = j_i,\, Y_i \in dy_i\}\right) = \pi_{j_0} \left(\prod_{i=1}^{n} f_{j_{i-1},\, j_i}(y_i)\right) dy_1 \cdots dy_n.$$

Ce résultat entraîne que la vraisemblance de $(y_1, \ldots, y_n)$ est

$$\forall\, y_1, \ldots, y_n \in \mathbb{R}^d, \quad \mathcal{L}(y_1, \ldots, y_n) = \sum_{j_0, \ldots, j_n} \pi_{j_0} \left(\prod_{i=1}^{n} f_{j_{i-1},\, j_i}(y_i)\right). \tag{1}$$

La complexité de cette expression est exponentielle en n, ce qui est bien trop élevé du point de vue du temps de calcul. En fait, en notant $f(y)$ la matrice carrée de taille r dont les éléments sont les $f_{ij}(y)$ et $\mathbf{1}$ le vecteur colonne de taille r n'ayant que des 1, la somme (1) se réécrit

$$\forall\, y_1, \ldots, y_n \in \mathbb{R}^d, \quad \mathcal{L}(y_1, \ldots, y_n) = \pi \left(\prod_{i=1}^{n} f(y_i)\right) \mathbf{1}.$$

Par conséquent, pour calculer la densité de $(Y_1, \ldots, Y_n)$, il suffit de calculer un produit de n matrices carrées de taille r. La complexité de cette opération est alors quadratique en n, ce qui est acceptable en pratique.

Une autre façon de voir un MMPP consiste à poser $J'_n = (J_{n-1}, J_n)$ et à considérer (J', Y) comme un modèle de Markov caché, dont la définition est donnée ci-dessous :

Définition. Un *modèle de Markov caché* (HMM, pour hidden Markov model) est un processus à temps discret (X_n, Y_n) tel que

- le processus X est une chaîne de Markov à espace d'états fini ;
- pour tout $n \in \mathbb{N} \setminus \{0\}$, sachant X_n, Y_n est indépendante de $(X_i, Y_i)_{i \leq n-1}$ et a pour densité f_{X_n}.

Que ce soit pour les MRPs ou les HMMs, la chaîne de Markov sous-jacente est vue comme un processus de données manquantes, qui influence le comportement du processus Y. La principale différence entre ces deux façons de voir un MMPP réside dans l'interprétation du modèle : dans le cas d'un MRP, on considère que la chaîne de Markov évolue en même temps que le processus Y (à cause de la dépendance de Y_n en J_{n-1} et J_n), alors que pour un HMM, la chaîne de Markov évolue en premier, suivi du processus observé Y.

Les HMMs ont été introduits en 1966 par Baum et Petrie [9] et sont utilisés dans la modélisation d'une grande variété de phénomènes, notamment :

- en biosciences : modélisation de l'ADN (Churchill [24]) ;
- en reconnaissance vocale (Juang et Rabiner [70]) ;
- en économétrie : modélisation de la volatilité (Kim *et al.* [72]) ;
- en climatologie : pour modéliser l'occurrence des jours de pluie (Smith [101]).

On renvoie le lecteur à Cappé *et al.* [18] pour un ouvrage récent sur le sujet.

La plupart des méthodes d'inférence statistique pour les modèles de Markov caché se basent sur l'étude de la vraisemblance : pour des HMMs classiques, la consistance de l'estimateur du maximum de vraisemblance (EMV) a été prouvée par Leroux [78], la preuve s'inspirant d'une méthode développée par Baum et Petrie [9]. La normalité asymptotique de cet estimateur a été obtenue par Bickel *et al.* [12]. Pour les MMPPs, la consistance de l'EMV a été obtenue par Rydén [97], la preuve s'inspirant de celle de Leroux dans le cas des HMMs, mais avec des hypothèses plus faibles. Le calcul de l'EMV dans ce cas se fait en général en utilisant l'algorithme de Nelder-Mead (voir Press *et al.* [92]), un algorithme de quasi-Newton ou un algorithme EM (voir Baum *et al.* [10]). On rappelle brièvement ici le principe de l'algorithme EM : pour plus de détails à ce sujet, voir par exemple McLachlan et Krishnan [84].

Supposons que $(X_n, Y_n)_{n \geq 1}$ est une suite de variables aléatoires dont la distribution est décrite par un paramètre $\Phi \in \mathbb{R}^d$. On suppose que le processus Y est observé et que le processus (à valeurs discrètes) X est manquant : la vraisemblance du modèle est alors

$$\mathcal{L}((Y_i)_{1 \leq i \leq n}, \Phi) = \sum_{X_1, \dots, X_n} \mathcal{L}((X_i, Y_i)_{1 \leq i \leq n}, \Phi)$$

où $\mathcal{L}((X_i,\, Y_i)_{1\le i\le n},\, \Phi)$ est appelée la *vraisemblance complète* du modèle. L'EMV du paramètre Φ est défini comme

$$\widehat{\Phi}_n = \underset{\Phi}{\operatorname{argmax}}\, \mathcal{L}((Y_i)_{1\le i\le n},\, \Phi).$$

Cependant, la vraisemblance est, dans la plupart des cas, impossible à maximiser directement en pratique. On ne peut donc pas calculer l'EMV de façon classique. L'algorithme EM est une procédure itérative qui consiste à maximiser la vraisemblance complète du modèle conditionnellement au processus observé, plutôt que de travailler directement avec la vraisemblance du processus. L'algorithme est le suivant :

Expectation step (E step): Calculer l'espérance conditionnelle de la log-vraisemblance complète sachant les données observées, sous le paramètre courant $\Phi^{(t)}$:

$$Q\left(\Phi \;\middle|\; \Phi^{(t)}\right) = \mathbb{E}_{\Phi^{(t)}}\left[\ln \mathcal{L}((X_i,\, Y_i)_{1\le i\le n},\, \Phi) \,|\, Y_i,\ 1 \le i \le n\right].$$

Maximisation step (M step): Trouver un paramètre $\Phi^{(t+1)}$ qui maximise cette quantité :

$$\Phi^{(t+1)} = \underset{\Phi}{\operatorname{argmax}}\, Q\left(\Phi \;\middle|\; \Phi^{(t)}\right).$$

Le résultat suivant (voir Dempster *et al.* [30]) montre que cet algorithme fait croître la vraisemblance à chaque étape :

Théorème. *Soit* $\left(\Phi^{(t)}\right)_{t\ge 0}$ *une suite de paramètres construite à l'aide de l'algorithme EM. Alors*

$$\forall\, t \in \mathbb{N}, \quad \mathcal{L}\left((Y_i)_{1\le i\le n},\, \Phi^{(t+1)}\right) \ge \mathcal{L}\left((Y_i)_{1\le i\le n},\, \Phi^{(t)}\right).$$

Un des principaux avantages de l'algorithme EM est qu'il ne requiert pas de calculer les dérivées de la log-vraisemblance, contrairement à la majorité des méthodes d'optimisation standard (comme l'algorithme du gradient ou la méthode de Newton). Cependant, il est possible que l'étape M de l'algorithme soit non explicite, ce qui impose d'utiliser un algorithme d'optimisation classique pour calculer $\Phi^{(t)}$. Il se trouve heureusement que ce n'est pas le cas pour les MMPPs, voir Rydén [100]. De plus, la convergence de la suite $\left(\Phi^{(t)}\right)$ vers l'EMV n'est pas garantie : en effet, la suite ainsi construite pourrait converger vers un point-selle de la log-vraisemblance. Cette sensibilité de l'algorithme EM à l'estimation initiale peut être résolue en pratique en calculant une première estimation proche de l'EMV (ce qui aura pour effet de diminuer le nombre d'itérations nécessaires à la convergence de l'algorithme), ou en démarrant l'algorithme à partir de plusieurs points différents.

Pour d'autres méthodes basées sur la vraisemblance, on peut citer Deng et Mark [31], Holst [66] et Meier-Hellstern [85]. Parmi les techniques n'utilisant pas la vraisemblance, citons une méthode d'estimation des moments et des fonctions de covariance (voir Gusella [54] et Rossiter [96]), ou la maximisation d'une vraisemblance "découpée", introduite par Rydén dans [98, 99], étudiée ensuite par Vandekerkhove [104] dans le cadre de mélanges de modèles de Markov cachés.

Un processus de pertes à modulation markovienne

Dans [82], Loisel suggère que la corrélation entre différentes branches d'une compagnie d'assurances peut provenir de chocs communs et d'une modulation par un processus d'environnement markovien. Dans le Chapitre 1, nous introduisons et nous étudions un processus de pertes à modulation markovienne, de la forme (J, N, S), où (J, N) est un MMPP et $S = (S_1, \ldots, S_n)$ est un processus de pertes (éventuellement multivarié) : autrement dit, les trajectoires des S_k sont des fonctions en escalier dont les sauts sont positifs. Le comportement de S est supposé être contrôlé par (J, N) de la façon suivante : les S_k ne peuvent sauter qu'à un instant de saut de N et si N saute lorsque J est dans l'état i, alors un saut des processus $S_{k_1}, \ldots, S_{k_p}$ se produit avec probabilité $p(i, e)$, où $e = \{k_1, \ldots, k_p\}$. Dans ce cas, la valeur X du saut a pour loi $\mathbb{P}_{\theta(i,e)}$, où $(\mathbb{P}_\theta)_{\theta\in\Theta}$ est un modèle statistique paramétrique. Enfin, les sauts de S sont indépendants sachant (J, N). Ce modèle est un modèle à chocs communs (voir Lindskog et McNeil [80]). Ses paramètres sont :

- les éléments ℓ_{ij} du générateur L de J ;
- les intensités de saut λ_i de N ;
- les probabilités $p(i, e)$;
- les paramètres $\theta(i, e)$.

En supposant disposer des données suivantes :

- le nombre d'états r de J ;
- la connaissance des processus N et S jusqu'à un certain temps T, qu'on suppose être un instant de saut de N,

on démontre la consistance de l'EMV du paramètre global Φ du modèle. Dans ce but, on adapte la preuve de Rydén [97] de la consistance de l'EMV pour un MMPP. On donne un algorithme EM, adapté de [100], pour calculer l'EMV, ainsi qu'une procédure d'estimation initiale. On utilise ce procédé sur des jeux de données réelles en assurance dans le cas univarié. Enfin, on réalise une étude sur simulations dans le cas d'un processus S multivarié.

Deuxième partie : Méthode des moments d'ordre élevé et estimation d'un point terminal

Le contexte de la théorie des valeurs extrêmes

La théorie des valeurs extrêmes a pour but de modéliser et de décrire la survenue et l'intensité d'évènements dits "rares". Les champs d'application de cette théorie sont très variés, notamment :

- la climatologie : évènements climatiques extrêmes (précipitations, températures, chutes de neige), modélisation des grands feux de forêt (Alvarado *et al.* [3]) ;
- l'hydrologie : crues consécutives à des pluies torrentielles (crue de l'Ouvèze à Vaison-La-Romaine, 1992) ; aux Pays-Bas, digues menacées par l'effet conjoint des grandes marées et des conditions climatiques en Mer du Nord (inondation de 1953) ;
- l'assurance : survenue de sinistres d'intensité exceptionnelle (ouragan Katrina en 2005, importants incendies en risques industriels, sinistres graves en responsabilité civile automobile) qui peuvent avoir des conséquences négatives sur les résultats et la solvabilité des organismes d'assurance ;
- la finance : fortes variations du cours d'actifs financiers, gestion du risque opérationnel des banques (crise des subprimes, fin des années 2000).

L'approche standard en théorie des probabilités place l'accent sur le comportement en moyenne et la variabilité autour de la moyenne, par le biais d'outils probabilistes comme par exemple la loi des grands nombres ou le théorème central limite. La particularité de la théorie des valeurs extrêmes est qu'elle se concentre sur la queue des distributions engendrant les divers phénomènes extrêmes étudiés. Les évènements considérés étant rares, le statisticien dispose de peu d'observations pour analyser les données. Un des problèmes principaux en théorie des valeurs extrêmes consiste à concevoir des techniques permettant d'estimer de manière efficace les quantiles des lois de probabilité étudiées : pour $p \in [0, 1]$, le *quantile d'ordre* $1 - p$ d'une fonction de répartition F associée à la variable aléatoire X est le réel x_p tel que

$$x_p = \sup\{x \in \mathbb{R} \,|\, F(x) < 1 - p\}.$$

x_p représente la première valeur de X qui sera dépassée avec probabilité inférieure à p. Dans notre contexte, les quantiles d'intérêt sont ceux pour lesquels p est “petit”. En climatologie par exemple, X peut représenter un niveau de retour de pluie journalier. On cherche alors à estimer un niveau atteint “rarement”, par exemple, en moyenne une fois tous les 10 ans.

Dans la deuxième partie de ce travail, on s'intéresse plus particulièrement au cas où le *point terminal*

$$\theta = \sup\{x \in \mathbb{R} \,|\, F(x) < 1\}$$

de F est fini. La variable aléatoire X associée à F a alors un support borné à droite et θ peut être vu comme un quantile de F pour $p = 0$. Cette situation se présente par exemple lorsqu'on étudie certains phénomènes environnementaux (Neves et Pereira [89]). Notre but est de présenter une approche nouvelle permettant d'estimer θ.

Avant de décrire notre méthode, rappelons dans un premier temps quelques éléments de la théorie des valeurs extrêmes nécessaires à notre étude.

Cadre univarié

Le cadre de la théorie des valeurs extrêmes est le suivant. Supposons que (X_n) est une suite de copies indépendantes d'une variable aléatoire X. Soient $X_{1,\,n} \leq X_{2,\,n} \leq \ldots \leq X_{n,\,n}$ les statistiques d'ordre du $n-$échantillon $(X_1, \ldots, X_n)$. On s'intéresse au comportement asymptotique du maximum $X_{n,\,n}$. Soit F la fonction de répartition de X. Si θ est le point terminal (peut-être infini) de F, une application directe du lemme de Borel-Cantelli donne $X_{n,\,n} \xrightarrow{\text{p.s.}} \theta$, ce qui n'est pas un résultat très précis, notamment du point de vue de la vitesse de convergence. On s'intéresse par conséquent au comportement asymptotique des maxima renormalisés

$$Y_n := \frac{X_{n,\,n} - b_n}{a_n} \tag{2}$$

où (a_n) et (b_n) sont deux suites réelles, avec $a_n > 0$. Les deux premières questions essentielles de la théorie des valeurs extrêmes sont alors :

- de trouver les limites possibles (en loi) de la suite (Y_n) ;
- de donner des conditions sur F pour qu'il existe des suites (a_n) et (b_n) telles que la suite (Y_n) possède une limite en loi.

La réponse à la première question a été donnée par Fisher et Tippett en 1928 [38] et Gnedenko en 1943 [49] sous la forme du résultat suivant :

Théorème (Fisher et Tippett [38], Gnedenko [49]) *Supposons que la suite (Y_n) définie en (2) converge en loi vers une variable Y, supposée non constante, de fonction de répartition G. Alors il existe a, b, $\gamma \in \mathbb{R}$ tels que $G(x) = G_\gamma(ax+b)$, où*

$$G_\gamma(x) := \begin{cases} \exp\left(-(1+\gamma x)^{-1/\gamma}\right) & si\ \gamma \neq 0\ \ et\ 1+\gamma x > 0, \\ \exp\left(-\exp(-x)\right) & si\ \gamma = 0. \end{cases}$$

Dans ce cas, on dit que F appartient au domaine d'attraction de G_γ et on note $F \in \mathcal{D}(G_\gamma)$.

Les distributions G_γ sont appelées distributions GEV (pour Generalised Extreme-Value). Par définition de G_γ, on peut distinguer trois types de distributions GEV, suivant la valeur du paramètre γ :

- le cas $\gamma > 0$, où $1 - G_\gamma$ est à support borné à gauche et à queue lourde :

$$1 - G_\gamma(x) = (\gamma x)^{-1/\gamma}(1 + \mathrm{o}(1)) \quad \text{quand } x \to \infty\,;$$

- le cas $\gamma < 0$, où $1 - G_\gamma$ est à support borné à droite ;
- le cas $\gamma = 0$, où $1 - G_\gamma$ est à décroissance exponentielle.

Ceci conduit à distinguer ces trois domaines, via la définition suivante.

Définition. Avec les notations du Théorème de Fisher-Tippett-Gnedenko :

- L'ensemble $\bigcup_{\gamma>0} \mathcal{D}(G_\gamma)$ est appelé *domaine d'attraction de Fréchet.*
- L'ensemble $\bigcup_{\gamma<0} \mathcal{D}(G_\gamma)$ est appelé *domaine d'attraction de Weibull.*
- L'ensemble $\mathcal{D}(G_0)$ est appelé *domaine d'attraction de Gumbel.*

La question est alors de caractériser les fonctions F appartenant à chaque domaine d'attraction. La réponse est fournie par le théorème suivant, dû à Gnedenko [49] et de Haan [55] :

Théorème (Gnedenko [49], de Haan [55]) *Soit $\gamma \in \mathbb{R}$. La fonction de répartition F ayant pour point terminal θ appartient à $\mathcal{D}(G_\gamma)$ si et seulement si*

- *pour $\gamma > 0$: θ est infini et*
$$\forall\, x > 0, \quad \frac{1 - F(tx)}{1 - F(t)} \to x^{-1/\gamma} \quad \textit{quand } t \to \infty\,;$$
- *pour $\gamma < 0$: θ est fini et*
$$\forall\, x > 0, \quad \frac{1 - F\left[\theta - \dfrac{1}{tx}\right]}{1 - F\left[\theta - \dfrac{1}{t}\right]} \to x^{1/\gamma} \quad \textit{quand } t \to \infty\,;$$
- *pour $\gamma = 0$: on ne peut rien dire à propos de θ et il existe une fonction f strictement positive telle que*
$$\forall\, x > 0, \quad \frac{1 - F(t + xf(t))}{1 - F(t)} \to e^{-x} \quad \textit{quand } t \uparrow \theta.$$

Le réel γ est appelé *indice des valeurs extrêmes de X.*

Ce résultat nous permet de donner quelques exemples de distributions appartenant aux différents domaines d'attraction :

- Domaine d'attraction de Fréchet : Fréchet, Cauchy, Burr type III, Burr type XII, Pareto, Log-Gamma, Student.
- Domaine d'attraction de Weibull : Uniforme, Beta, Reverse Burr.
- Domaine d'attraction de Gumbel : Weibull, Gamma, Normale, Log-normale.

En notant $\overline{F} = 1 - F$ la fonction de survie associée à F, on a alors

- pour $\gamma > 0$: $F \in \mathcal{D}(G_\gamma)$ si et seulement si θ est infini et
$$\forall\, x > 0, \quad \frac{\overline{F}(tx)}{\overline{F}(t)} \to x^{-1/\gamma} \quad \text{quand } t \to \infty\,;$$

- pour $\gamma < 0$: $F \in \mathcal{D}(G_\gamma)$ si et seulement si θ est fini et, en notant $\overline{G} : x \mapsto \overline{F}(\theta - 1/x)$,

$$\forall\, x > 0, \quad \frac{\overline{G}(tx)}{\overline{G}(t)} \to x^{1/\gamma} \quad \text{quand } t \to \infty.$$

Les fonctions $\overline{F}$ et $\overline{G}$ ci-dessus font ainsi partie de la classe des fonctions à variation régulière, dont on rappelle la définition ci-dessous :

Définition. Une fonction L mesurable strictement positive à l'infini est dite *à variation régulière d'indice* $\alpha \in \mathbb{R}$ si

$$\forall\, x > 0, \quad \frac{L(tx)}{L(t)} \to x^{\alpha} \quad \text{quand } t \to \infty.$$

Si $\alpha = 0$, on dit que L est *à variation lente.* L'ensemble des fonctions à variation régulière d'indice α est noté RV_α.

Il est clair que si L est une fonction à variation lente et si $\alpha \in \mathbb{R}$, alors la fonction f définie par $f(x) = x^\alpha\, L(x)$ pour tout $x > 0$ appartient à RV_α. En fait, RV_α est exactement composé de toutes ces fonctions :

Proposition. *Soient $\alpha \in \mathbb{R}$ et $f \in \mathrm{RV}_\alpha$. Alors il existe une fonction à variation lente L telle que*

$$\forall\, x > 0, \quad f(x) = x^\alpha\, L(x).$$

Ce résultat montre que l'étude des fonctions à variation régulière se ramène à celle des fonctions à variation lente. Parmi les fonctions à variation lente, on peut citer

- les fonctions possédant une limite strictement positive à l'infini ;
- les fonctions de la forme $f : x \mapsto |\ln x|^\beta$, $\beta \in \mathbb{R}$;
- les fonctions L telles que

$$\exists\, M > 0, \quad \forall\, x \geq M, \quad L(x) = C + Dx^{-\beta}(1 + \delta(x)) \tag{3}$$

  où C, $\beta > 0$, $D \in \mathbb{R} \setminus \{0\}$ et δ est une fonction mesurable tendant vers 0 à l'infini. L'ensemble de ces fonctions L est appelé *classe de Hall.*

Pour de plus amples informations sur la théorie des fonctions à variation régulière, on renvoie le lecteur à Bingham *et al.* [14].

En utilisant la notion de fonction à variation régulière, on obtient un critère pratique pour tester l'appartenance aux domaines d'attraction de Weibull ou de Fréchet :

Théorème (Gnedenko [49], de Haan [55]) *Soient $\gamma \neq 0$, F une fonction de répartition avec point terminal θ et $\overline{F}$ la fonction de survie associée. Alors F appartient à $\mathcal{D}(G_\gamma)$ si et seulement si*

- *pour* $\gamma > 0$ *:* θ *est infini et* $\overline{F} \in \mathrm{RV}_{-1/\gamma}$ *;*

- *pour* $\gamma < 0$ *:* θ *est fini et* $x \mapsto \overline{F}(\theta - 1/x) \in \mathrm{RV}_{1/\gamma}$.

On dégage deux corollaires importants de ce résultat :

- dans le cas $\gamma > 0$, la fonction de survie $\overline{F}$ est à décroissance polynomiale ;

- dans le cas $\gamma < 0$, la fonction de survie $\overline{F}$ se comporte comme un polynôme au voisinage du point terminal.

De plus, les domaines d'attraction de Fréchet et Weibull sont étroitement liés : si $\theta \in \mathbb{R}$ est un réel fixé et si X est une variable aléatoire dont la fonction de répartition appartient au domaine d'attraction de Fréchet, alors la fonction de répartition de $\theta - 1/X$ appartient au domaine d'attraction de Weibull. Réciproquement, si Y est une variable aléatoire dont la fonction de répartition appartient au domaine d'attraction de Weibull, avec point terminal θ, alors la fonction de répartition de $(\theta - Y)^{-1}$ appartient au domaine d'attraction de Fréchet.

Notre but étant d'exposer une nouvelle méthode d'estimation du point terminal θ, on s'intéresse en particulier au cas où F appartient au domaine d'attraction de Weibull. On présente rapidement ici quelques approches existantes pour ce problème :

L'estimateur du maximum :

L'estimateur du maximum est l'outil le plus ancien disponible pour estimer un point terminal : c'est l'estimateur $\widehat{\theta}_n^{Max} = X_{n,n}$. Ses propriétés asymptotiques sont bien connues (voir de Haan et Ferreira [57]) ; en particulier, il est presque sûrement convergent quelle que soit la distribution de X. Notons que de nombreux raffinements de cette méthode ont été étudiés, voir par exemple Cooke [26], de Haan [56], Miller [88] et Robson et Whitlock [93].

L'estimateur du maximum est particulièrement adapté lorsque les données à exploiter sont obtenues sous forme de maxima, par exemple lorsqu'on travaille sur des maxima de températures ou de précipitations en climatologie. Cependant, cette approche est critiquable dans le sens où elle n'utilise qu'une seule valeur de l'échantillon de données disponibles. La déperdition d'information est ainsi très importante. De plus, les performances de cet estimateur se dégradent fortement lorsque l'indice des valeurs extrêmes tend vers 0.

L'estimateur du maximum de vraisemblance de Hall (Hall [58]) :

L'estimateur du maximum de vraisemblance de Hall est une solution $\widehat{\theta}_n^H$ de l'équation

$$r_n \left\{ \frac{1}{\sum_{j=1}^{r_n-1} \ln(1+\xi_{j,n}(\theta))} - \frac{1}{\sum_{j=1}^{r_n-1} \xi_{j,n}(\theta)} \right\} = 1, \quad \theta > X_{n,n}$$

où r_n est une *suite intermédiaire* d'entiers, c'est-à-dire telle que $1 < r_n < n$, $r_n \to \infty$ et $r_n/n \to 0$ quand $n \to \infty$ et

$$\xi_{j,\,n}(\theta) = \frac{X_{n-j+1,\,n} - X_{n-r_n+1,\,n}}{\theta - X_{n-j+1,\,n}}.$$

Les propriétés asymptotiques de cet estimateur ont été étudiées par Hall [58], dans le cas particulier où L appartient à la classe de Hall (3). Il est assez instable et n'est pas explicite. Il a été récemment amélioré par Li et Peng [79].

L'idée d'utiliser un nombre croissant de statistiques d'ordre de l'échantillon a ensuite été plus largement développée dans le cadre de l'approche Peaks Over Thresholds (POT), via l'approximation de la loi des excès au-delà d'un seuil élevé par des GPD (Generalised Pareto Distributions). Plus précisément, soient $u < \theta$ et F_u la *fonction de répartition des excès* définie par

$$F_u(x) = \mathbb{P}(X - u \leq x \,|\, X > u) = \frac{F(u+x) - F(u)}{1 - F(u)}.$$

F_u décrit la loi de X sachant $\{X > u\}$. On a alors le résultat suivant, dû à Balkema et de Haan [8] et Pickands [91]:

Théorème (Balkema et de Haan [8], Pickands [91]) *Soit $\gamma \in \mathbb{R}$. La fonction de répartition F avec point terminal associé θ appartient à $\mathcal{D}(G_\gamma)$ si et seulement s'il existe une fonction σ strictement positive telle que la loi des excès F_u puisse être uniformément approchée par une GPD $G_{\gamma,\,\sigma}$, c'est-à-dire*

$$\sup_{0<x<\theta-u} \left| F_u(x) - G_{\gamma,\,\sigma(u)}(x) \right| \to 0$$

quand $u \to \theta$, où

$$G_{\gamma,\,\sigma}(x) := \begin{cases} 1 - \left(1 + \gamma \dfrac{x}{\sigma}\right)^{-1/\gamma} & si\ \gamma \neq 0\ et\ 1 + \gamma \dfrac{x}{\sigma} > 0, \\ 1 - e^{-x/\sigma} & si\ \gamma = 0. \end{cases}$$

Ce théorème montre que lorsque le seuil u est suffisamment grand, la loi des excès au-delà de u peut être approchée par une GPD de paramètres γ et σ. On présente maintenant quelques estimateurs dont la construction se base sur ce résultat. Soit (k_n) une suite intermédiaire d'entiers. La construction générique est la suivante : si $(\widehat{\gamma}_n, \widehat{\sigma}_n)$ sont des estimateurs de (γ, σ) tels que $(\widehat{\gamma}_n, \widehat{\sigma}_n, X_{n-k_n,\,n})$ est asymptotiquement normal, alors la variable aléatoire

$$\widehat{\theta} := X_{n-k_n,\,n} - \frac{\widehat{\sigma}_n}{\widehat{\gamma}_n}$$

est un estimateur asymptotiquement normal de θ, voir [57], Théorème 4.5.1 et Corollaire 4.5.2 p. 146.

Il suffit donc de donner des estimateurs de l'indice des valeurs extrêmes γ et du paramètre d'échelle σ. Citons, de manière non exhaustive, quelques approches possibles :

L'estimateur des moments (Dekkers *et al.* [29]) :

Pour $j \in \{1, 2\}$, on note

$$M_n^{(j)} = \frac{1}{k_n} \sum_{i=0}^{k_n - 1} \left(\ln X_{n-i,\, n} - \ln X_{n-k_n,\, n}\right)^j .$$

$M_n^{(1)}$ est l'estimateur de Hill de γ, dont les propriétés sont données dans Hill [65]. L'estimateur des moments de γ présenté dans [29] est alors

$$\widehat{\gamma}_n^M := M_n^{(1)} + \widehat{\gamma}_-, \quad \widehat{\gamma}_- = 1 - \frac{1}{2}\left[1 - \frac{\left(M_n^{(1)}\right)^2}{M_n^{(2)}}\right]^{-1} .$$

L'estimateur du paramètre d'échelle est

$$\widehat{\sigma}_n^M := X_{n-k_n,\, n}\, M_n^{(1)}(1 - \widehat{\gamma}_-).$$

En ce qui concerne les propriétés asymptotiques du couple $\left(\widehat{\gamma}^M,\, \widehat{\sigma}^M\right)$, voir [57]. Une version plus performante de l'estimateur dans le cas $\gamma < -1/2$ a été étudiée par Aarssen et de Haan [1] : elle est donnée par

$$\widehat{\theta}_n^{M,\,*} = X_{n-k_1,\, n} - \frac{X_{n-k_1,\, n}\, M_{n,\, k_1}^{(1)}(1 - \widehat{\gamma}_-)}{\widehat{\gamma}_n^M}$$

où $k_1 \in \{1, 2, \ldots, n-1\}$ est un seuil *fixe*, indépendant de n et

$$M_{n,\, k_1}^{(1)} = \frac{1}{k_1} \sum_{i=0}^{k_1 - 1} \left(\ln X_{n-i,\, n} - \ln X_{n-k_1,\, n}\right) .$$

Estimateur des moments pondérés (Hosking et Wallis [67]) :

On définit les moments pondérés d'une variable aléatoire Y de fonction de répartition F comme étant

$$\forall\, p,\, r,\, s > 0, \quad M_{p,\, r,\, s} := \mathbb{E}\left(Y^p\, [F(Y)]^r\, [1 - F(Y)]^s\right)$$

(voir Greenwood *et al.* [52]). D'après [67], si Y a une distribution de Pareto généralisée de paramètres $(\mu,\, \sigma)$,

$$\forall\, s > \gamma - 1, \quad M_{1,\, 0,\, s} = \frac{\sigma}{(s+1)(s+1-\gamma)}.$$

En notant

$$P_n := \frac{1}{k_n} \sum_{j=0}^{k_n - 1} (X_{n-j,\, n} - X_{n-k_n,\, n}),$$

$$Q_n := \frac{1}{k_n} \sum_{j=0}^{k_n - 1} \frac{j}{k_n} (X_{n-j,\, n} - X_{n-k_n,\, n}),$$

P_n et Q_n sont respectivement les moments empiriques correspondant à $M_{1,\, 0,\, 0}$ et $M_{1,\, 0,\, 1}$; les estimateurs de γ et σ sont alors

$$\widehat{\gamma}_n^{PWM} := \frac{P_n - 4Q_n}{P_n - 2Q_n},$$

$$\widehat{\sigma}_n^{PWM} := \frac{2P_nQ_n}{P_n - 2Q_n}.$$

La normalité asymptotique du couple $\left(\widehat{\gamma}^{PWM}, \widehat{\sigma}^{PWM}\right)$ a été établie dans [67]. Ces estimateurs ont depuis été généralisés par Diebolt *et al.* [33, 34].

L'estimateur du maximum de vraisemblance (Smith [102]) :
La log-vraisemblance d'un échantillon de variables GPD de taille k_n étant

$$\mathcal{L}(Y_1, \ldots, Y_{k_n} ; \gamma, \sigma) = -k_n \ln \sigma - \left[\frac{1}{\gamma} + 1\right] \sum_{j=1}^{k_n} \ln\left[1 + \frac{\gamma}{\sigma} Y_j\right],$$

le couple d'estimateurs du maximum de vraisemblance de (γ, σ) est

$$\left(\widehat{\gamma}_n^{MLE}, \widehat{\sigma}_n^{MLE}\right) = \underset{(\gamma,\sigma)}{\operatorname{argmax}} \, \mathcal{L}(Y_1, \ldots, Y_{k_n} ; \gamma, \sigma).$$

On renvoie le lecteur à Drees *et al.* [35] pour l'étude détaillée de ces estimateurs.

Pour conclure ce tour d'horizon des méthodes d'estimation de point terminal, citons Athreya et Fukuchi [7] et Loh [81] pour des méthodes bootstrap, Goldenshluger et Tsybakov [50] pour l'estimation de point terminal en présence d'un bruit, Hall et Wang [61] pour une méthode de minimisation et Hall et Wang [62] pour une approche bayésienne.

Point terminal : le cas avec covariable

Dans ce paragraphe, on se concentre sur l'estimation en présence d'une covariable. Plus précisément, on suppose disposer de copies indépendantes (X_i, Y_i) d'un couple (X, Y), où X est à valeurs dans $\mathbb{R}^d$ et Y est à valeurs dans $[0, \infty)$. On suppose que X est à densité f dont le support est un compact $\Omega \subset \mathbb{R}^d$ et que, conditionnellement à $X = x$, la loi de Y est à support borné $[0, g(x)]$. La fonction g est appelée *frontière* du support de (X, Y). La question analogue à celle de l'estimation d'un point terminal est ici d'estimer la frontière g en un point $x \in \Omega$ quelconque. En économétrie par exemple, dans le cadre d'études de productivité, X désigne la production d'une unité et Y représente son coût de fonctionnement (Cazals *et al.* [19]).

Du point de vue de l'estimation, on ne peut pas utiliser directement les techniques développées dans le cas univarié, puisque la probabilité qu'une des X_i au moins soit égale à un point x fixé est nulle. Une première idée est de chercher comment estimer f elle-même. Un estimateur très populaire pour ce problème est l'estimateur de Parzen-Rosenblatt [90, 95], défini par :

$$\forall\, x \in \mathbb{R}^d, \quad \widehat{f}_n(x) = \frac{1}{n} \sum_{i=1}^{n} K_{h_n}(x - X_i)$$

où (h_n) est une suite de réels positifs qui tend vers 0, appelée *fenêtre* et

$$\forall\, u \in \mathbb{R}^d, \quad K_h(u) = \frac{1}{h^d} K\left(\frac{u}{h^d}\right)$$

où K est une densité de probabilité sur $\mathbb{R}^d$, généralement à support compact, qu'on appelle *noyau*. En d'autres termes, pour estimer la densité f en un point $x \in \mathbb{R}^d$, on calcule la moyenne pondérée des observations dans un voisinage du point x, chacune étant affectée d'un poids selon leur distance à ce point. Cette technique possède d'excellentes propriétés tant théoriques ([90, 95]) que pratiques. Elle suggère de plus que dans le cadre d'un problème avec covariable ayant une loi à densité, il peut être utile de tenir compte des observations au voisinage du point considéré.

Comme dans le cadre univarié, de nombreuses méthodes ont été proposées pour ce problème d'estimation. Nous en citons (de manière non exhaustive) quelques-unes ici.

Méthode utilisant la théorie des valeurs extrêmes :

En premier lieu, on peut penser à adapter l'estimateur du maximum au contexte avec covariable, en considérant le maximum des observations au voisinage d'un point. Ce travail a été effectué par Geffroy [40] dans le cas où $d = 1$ et $\Omega = [0, 1]$: soient (k_n) une suite strictement croissante d'entiers et $I_{n,r} = [(r-1)/k_n, r/k_n]$ pour $r \in \{1, \ldots, k_n\}$. On pose

$$\forall r \in \{1, \ldots, k_n\}, \quad U_{n,r} = \max\{Y_i \mid X_i \in I_{n,r}\}.$$

L'estimateur de Geffroy de g est alors

$$\forall x \in \mathbb{R}^d, \quad \widehat{g}_n^{Ge}(x) = \sum_{r=1}^{k_n} U_{n,r} \, \mathbb{1}_{\{x \in I_{n,r}\}}.$$

Tout comme son analogue dans le cas univarié, cet estimateur possède des propriétés théoriques satisfaisantes (voir [40]), mais ses performances s'avèrent malheureusement assez mauvaises en pratique lorsqu'on dispose de peu de points au voisinage de la frontière, comme par exemple lorsque la loi conditionnelle de Y sachant $\{X = x\}$ est une distribution GEV avec un indice des valeurs extrêmes proche de 0. De plus, il nécessite de subdiviser *a priori* le support de X. Signalons que des généralisations de cet estimateur ont été proposées pour en améliorer les performances, notamment par Girard et Jacob [45].

L'approche utilisant la théorie des valeurs extrêmes a également donné lieu à l'étude d'autres procédés, notamment lorsque les observations sont des réalisations d'un processus ponctuel : parmi les approches développées dans ce cadre, citons Gardes [39], Girard et Jacob [43, 44] (en utilisant une méthode de projection décrite par Jacob et Suquet [68]) et Menneteau [86]. Des techniques générales d'étude de ces estimateurs dans ce contexte sont proposées par Girard et Menneteau [48].

Méthode par polynômes locaux (cas $d = 1$) :

Cette méthode consiste à approcher g par une courbe polynomiale par morceaux. La technique générale est la suivante : on subdivise le support Ω de f en sous-ensembles disjoints $\Omega_1, \ldots, \Omega_m$ et on fixe des points $u_i \in \Omega_i$ pour tout $i \in \{1, \ldots, m\}$. Un estimateur localement polynomial de g de degré k est de la forme

$$\forall\, x \in \mathbb{R}^d, \quad \widehat{g}_n^P(x) = \sum_{i=1}^{m} \sum_{j=0}^{k} \widehat{a}_{i,\,j} (x - u_i)^j \, \mathbb{1}_{\{x \in \Omega_i\}} \,;$$

les $\widehat{a}_{i,\,j}$ sont les coefficients des polynômes utilisés pour l'interpolation de g. En général, ce sont les solutions d'un problème de minimisation : voir par exemple Hall et Park [59], Hall *et al.* [60], Härdle *et al.* [63], Knight [75], Korostelev et Tsybakov [77]. Ces estimateurs sont appréciés pour leurs bonnes propriétés théoriques ; les défauts principaux de cette approche sont la nécessité d'imposer un partage *a priori* du domaine Ω et que les estimateurs obtenus peuvent être non explicites.

Il existe encore d'autres méthodes, comme les méthodes d'optimisation linéaire (Girard *et al.* [42]). Signalons enfin que dans le cas d'une covariable réelle ($d = 1$) et lorsque la frontière g est une fonction croissante, le point $g(x)$ peut s'interpréter comme le point terminal de la distribution de Y sachant $\{X \leq x\}$. Ce type de situation est notamment courant en économétrie. Des résultats dans ce cadre particulier ont été obtenus par Deprins *et al.* [32], Farrell [36], Gijbels *et al.* [41], Korostelev *et al.* [76] ou Aragon *et al.* [4], Cazals *et al.* [19], Daouia et Simar [27] pour des estimateurs robustes.

Méthode des moments d'ordre élevé

La seconde partie de ce travail se concentre sur des estimateurs utilisant la méthode des moments d'ordre élevé. Cette méthode a été introduite, dans le cadre multivarié, par Girard et Jacob [46] qui ont considéré l'estimateur

$$\forall\, x \in \mathbb{R}^d, \quad \widehat{g}_n^{GJ}(x) = \left[(p_n + 1) \frac{\displaystyle\sum_{i=1}^{n} Y_i^{p_n} \, K_{h_n}(x - X_i)}{\displaystyle\sum_{i=1}^{n} K_{h_n}(x - X_i)} \right]^{1/p_n}$$

où (p_n) est une suite de réels positifs tendant vers l'infini. L'utilisation de la fenêtre (h_n) permet comme précédemment de choisir les observations pour lesquelles la covariable est proche de x ; la nouveauté vient de l'utilisation de la puissance $p_n \to \infty$, qui permet de donner (exponentiellement) plus de poids aux observations Y_i proches de $g(x)$. La procédure conserve par conséquent toute l'information donnée par l'échantillon, en pénalisant les points éloignés de la frontière. Cet estimateur est consistant quelle que soit la loi de Y sachant X et il est asymptotiquement normal dans le cas où Y sachant X suit une loi uniforme. Il possède également un comportement satisfaisant en pratique. Cependant, lorsque la loi de Y sachant X n'est pas uniforme, cet estimateur n'est *pas* asymptotiquement normal (Chapitre 3, Théorème 3.2). Signalons que la méthode des moments d'ordre élevé a aussi été employée par Girard et Jacob dans [47] dans le cadre d'une estimation par polynômes locaux, mais la loi asymptotique de l'estimateur n'est pas connue.

La méthode des moments d'ordre élevé donnant de bons résultats en pratique, on souhaite trouver un estimateur défini sous forme explicite, utilisant cette approche, qui soit asymptotiquement normal dans un cadre moins restrictif. Pour ce faire, on considère d'abord le cas univarié, en introduisant l'estimateur

$$\frac{1}{\widehat{\theta}_n} = \frac{1}{ap_n}\left[((a+1)p_n+1)\frac{\widehat{\mu}_{(a+1)p_n}}{\widehat{\mu}_{(a+1)p_n+1}} - (p_n+1)\frac{\widehat{\mu}_{p_n}}{\widehat{\mu}_{p_n+1}}\right]$$

où $a > 0$ et

$$\widehat{\mu}_{p_n} = \frac{1}{n}\sum_{i=1}^{n} X_i^{p_n}$$

qui est l'estimateur empirique classique du moment $\mu_{p_n} = \mathbb{E}(X^{p_n})$. Dans le cas multivarié, l'estimateur considéré est

$$\forall\, x \in \mathbb{R}^d, \quad \frac{1}{\widehat{g}_n(x)} = \frac{1}{ap_n}\left[((a+1)p_n+1)\frac{\widehat{\mu}_{(a+1)p_n}(x)}{\widehat{\mu}_{(a+1)p_n+1}(x)} - (p_n+1)\frac{\widehat{\mu}_{p_n}(x)}{\widehat{\mu}_{p_n+1}(x)}\right]$$

où

$$\widehat{\mu}_{p_n}(x) = \frac{1}{n}\sum_{i=1}^{n} Y_i^{p_n}\, K_{h_n}(x - X_i)$$

est l'estimateur empirique à noyau du moment conditionnel $m_{p_n}(x) = \mathbb{E}(Y^{p_n} \,|\, X = x)$: c'est la généralisation au cas avec covariable de l'estimateur qu'on considère dans le cas univarié.

Ces deux estimateurs sont obtenus par des considérations sur les moments de l'échantillon. Nos résultats sont les suivants :

- Dans le Chapitre 2, on montre que l'estimateur des moments d'ordre élevé dans le cadre univarié est consistant sans hypothèse paramétrique supplémentaire sur la loi de X. La normalité asymptotique de l'estimateur est également établie en supposant que X est positive et sa loi appartient au domaine d'attraction de Weibull avec une condition classique de régularité sur la fonction à variation lente perturbatrice, ou lorsque cette fonction appartient à la classe de Hall (3) sans hypothèse supplémentaire. On s'inspire ensuite de la technique utilisée pour construire un estimateur du point terminal lorsque X n'est plus nécessairement positive et on démontre ses propriétés asymptotiques. On compare enfin, sur simulations, les performances de nos estimateurs avec celles d'estimateurs classiques.

- Dans le Chapitre 3, on montre que l'estimateur des moments d'ordre élevé dans le cadre avec covariable est consistant sans hypothèse paramétrique supplémentaire sur la loi conditionnelle de Y sachant X. La normalité asymptotique de l'estimateur est établie en supposant que la loi conditionnelle de Y sachant X est positive et appartient à la classe de Hall (3). Les performances de notre estimateur sont comparées avec celles d'autres estimateurs classiques sur simulations.

Part I

Markov-modulated loss processes in insurance

1 Estimation of the parameters of a Markov-modulated loss process in insurance

1.1 Model, assumptions and notation

We briefly recall the context of our study. We consider an MMPP (J, N), where J is an irreducible continuous-time Markov process with transition intensity matrix L on the state space $\{1, \ldots, r\}$, where $r \in \mathbb{N} \setminus \{0\}$, and N is a univariate counting process such that, when J is in state i, N is a Poisson process with intensity λ_i. It is assumed that (J, N) is nontrivial: there exists $i \in \{1, \ldots, r\}$ such that $\lambda_i > 0$. We further consider a loss process $S = (S_1, \ldots, S_n)$ (namely, the sample paths of the S_k are step functions with nonnegative increments) whose behaviour is driven by N in the following sense: assume that the S_k can only jump when N does, and that if N jumps at time t when J is in state i, then a simultaneous jump of the processes $S_{k_1}, \ldots, S_{k_p}$ at time t occurs with probability $p(i, e)$ where $e = \{k_1, \ldots, k_p\}$ is a subset of $\{1, \ldots, n\}$. We then assume that the random variables E_s, such that the S_k with $k \in E_s$ (and only these) jumped at the time of the sth jump of N, are independent given the process (J, N). Finally, assume that the value of the jump $X_s = S(\tau_s) - S(\tau_{s-1})$, where (τ_j) is the sequence of the times when N jumps, has a distribution $\mathbb{P}_{\theta(i, e)}$, where $(\mathbb{P}_\theta)_{\theta \in \Theta}$ is a parametric statistical model, that is

$$\mathbb{P}(S(\tau_s) - S(\tau_{s-1}) = x \,|\, J(\tau_s) = i, E_s = e) = \mathbb{P}_{\theta(i, e)}(\forall\, m,\ m \in e \Rightarrow X_m = x_m)$$

where X is a random vector with distribution $\mathbb{P}_{\theta(i, e)}$, with clearly $x_m = 0$ if $m \notin e$. Note that this model can be seen as a common shock model as in [80]: it is assumed that given the process (J, N) and the sequence (E_s), the process $(S(\tau_s))$ has independent increments.

The context of our work is the following: let us assume that the process S has been observed up to time T, so that the available data is:

- The number r of states of J;
- The full knowledge of the processes N and S between time 0 and time T, both assumed to be times when N jumps.

The goal is to estimate the unknown parameters of the model, namely:

- The elements ℓ_{ij} of the transition intensity matrix L of J;
- The jump intensities λ_i of N;
- The probabilities $p(i,\, e)$;
- The parameters $\theta(i,\, e)$.

Notice that the process J is not observed, which induces technical difficulties. For the sake of brevity, we let Φ be the global parameter of the model. The distribution of the process with parameter Φ is then denoted by $\mathbb{P}_\Phi$.

1.2 Asymptotic properties of the maximum likelihood estimator

Our aim is to estimate the parameters with a maximum likelihood estimator (MLE). The available data is:

- The values $0 < t_1 < \ldots < t_k = T$ of the τ_i, *i.e.* the times when N jumps (equivalently, the inter-event times $y_1,\, \ldots,\, y_k$);
- $e_1,\, \ldots,\, e_k$ the successive values of the E_i;
- $x_1,\, \ldots,\, x_k$ the successive values of the X_i.

Let now

$$\begin{aligned} f_{ij}(t,\, \Phi)\, dt &:= \mathbb{P}_\Phi(\tau_1 \in dt,\; J(t) = j \,|\, J(0) = i) \\ \overline{F}_{ij}(t,\, \Phi) &:= \mathbb{P}_\Phi(\tau_1 > t,\; J(t) = j \,|\, J(0) = i). \end{aligned}$$

Setting $\Lambda = \mathrm{diag}(\lambda_i,\, 1 \le i \le r)$, recall that

$$f(t,\, \Phi) = \exp(t(L(\Phi) - \Lambda(\Phi)))\Lambda(\Phi), \quad \overline{F}(t,\, \Phi) = \exp(t(L(\Phi) - \Lambda(\Phi))).$$

Let then

$$\begin{aligned} p(\cdot,\, e,\, \Phi) &= \mathrm{diag}(p(i,\, e,\, \Phi),\, 1 \le i \le r), \\ \mathbb{P}_{\theta(\cdot,\, e,\, \Phi)}(X = x) &= \mathrm{diag}(\mathbb{P}_{\theta(i,\, e,\, \Phi)}(X = x),\, 1 \le i \le r), \end{aligned}$$

and in matrix notation

$$\begin{aligned}
\forall\, e \subset \{1,\, \dots,\, n\},\ e \neq \varnothing,\quad g(t,\, e,\, x,\, \Phi) &= f(t,\, \Phi)\, p(\cdot,\, e,\, \Phi)\, \mathbb{P}_{\theta(\cdot,\, e,\, \Phi)}(X = x) \\
g(t,\, \varnothing,\, x,\, \Phi) &= f(t,\, \Phi)\, p(\cdot,\, \varnothing,\, \Phi)\, \mathbb{1}_{\{x=0\}}.
\end{aligned}$$

With these notations, the $(i,\, j)$th element of the matrix $g(t,\, e,\, x,\, \Phi)$ is

$$\begin{aligned}
\forall\, e \subset \{1,\, \dots,\, n\},\ e \neq \varnothing,\quad g_{ij}(t,\, e,\, x,\, \Phi) &= f_{ij}(t,\, \Phi)\, p(j,\, e,\, \Phi)\, \mathbb{P}_{\theta(j,\, e,\, \Phi)}(X = x) \\
g_{ij}(t,\, \varnothing,\, x,\, \Phi) &= f_{ij}(t,\, \Phi)\, p(j,\, \varnothing,\, \Phi)\, \mathbb{1}_{\{x=0\}}.
\end{aligned}$$

It is now sufficient to specify the initial distribution of J to compute the likelihood of the observations. Denote by $P(\Phi)$ the transition matrix of the discrete-time Markov chain $(J_i = J(\tau_i))$: integrating f, one gets

$$P(\Phi) = (\Lambda(\Phi) - L(\Phi))^{-1}\Lambda(\Phi).$$

According to [97], $P(\Phi)$ has a unique stationary distribution $\pi(\Phi)$ and we have, if $a(\Phi)$ is the only stationary distribution of the continuous-time process $(J(t))_{t\geq 0}$ and $\mathbf{1}$ is the column vector of size r with all entries equal to 1,

$$\pi(\Phi) = \frac{a(\Phi)\Lambda(\Phi)}{a(\Phi)\Lambda(\Phi)\mathbf{1}}.$$

From now on, we shall assume that the initial distribution of J is $\pi(\Phi)$; the process $(J_i,\, Y_i,\, E_i,\, X_i)$ is then $\mathbb{P}_\Phi-$stationary, because the bivariate process $(J_i,\, Y_i)$ is a stationary Markov renewal process. Thus, the likelihood of the observed data under the distribution $\mathbb{P}_\Phi$ is

$$\mathcal{L}((y_i,\, e_i,\, x_i)_{1\leq i\leq k},\, \Phi) = \pi(\Phi)\left(\prod_{i=1}^{k} g(y_i,\, e_i,\, x_i,\, \Phi)\right)\mathbf{1}.$$

Assuming now that we know the states $j_i = J(t_i)$, the complete likelihood of the data is

$$\mathcal{L}((j_i)_{0\leq i\leq k},\, (y_i,\, e_i,\, x_i)_{1\leq i\leq k},\, \Phi) = \pi_{j_0}(\Phi)\left(\prod_{i=1}^{k} g_{j_{i-1},\, j_i}(y_i,\, e_i,\, x_i,\, \Phi)\right).$$

To give a result on the strong consistency of the MLE, we first need some notations: for an arbitrary parameter Φ, denote by F_Φ the set of all parameters Φ' such that for all e

$$(\forall\, j \in \{1,\, \dots,\, r\},\ \lambda_j(\Phi)\, p(j,\, e,\, \Phi) = 0) \Leftrightarrow (\forall\, j \in \{1,\, \dots,\, r\},\ \lambda_j(\Phi')\, p(j,\, e,\, \Phi') = 0).$$

F_Φ can be thought of as the set of the elements Φ' such that a simultaneous jump of the processes $S_{k_1},\, \dots,\, S_{k_q}$ is a.s. impossible under the law $\mathbb{P}_\Phi$ if and only if it is a.s. impossible under the law $\mathbb{P}_{\Phi'}$. Write further $\Phi \sim \Phi'$ whenever $(Y_i,\, E_i,\, X_i)$ has the same law under $\mathbb{P}_\Phi$ and under $\mathbb{P}_{\Phi'}$.

We finally write down the hypotheses we need to state our main result:

(A_1) For all $e \neq \varnothing$, the distributions $\mathbb{P}_{\theta(\cdot,\, e)}$ have the same support, with no atom at 0.

(A_2) For all $e \neq \varnothing$ and all Φ, Φ', there exists a neighborhood G of Φ' such that for every subset $G_{\Phi'}$ of G and all $i, j \in \{1, \ldots, r\}$,

$$\int \left| \ln \sup_{\varphi \in G_{\Phi'}} \mathbb{P}_{\theta(i,\, e,\, \varphi)}(m \in e \Rightarrow X_m = x_m) \right| \mathbb{P}_{\theta(j,\, e,\, \Phi)}(m \in e \Rightarrow X_m = x_m)\, dx < \infty.$$

(A_3) For all $e \neq \varnothing$, all $i \in \{1, \ldots, r\}$ and all x, $\varphi \mapsto \mathbb{P}_{\theta(i,\, e,\, \varphi)}(m \in e \Rightarrow X_m = x_m)$ is a continuous function.

This allows us to state our main result:

Theorem 1.1. *Assume that $(A_1 - A_3)$ hold. Let Φ_0 be the true value of the parameter, and let C be a compact set of F_{Φ_0} such that $\Phi_0 \in C$. Let $\widehat{\Phi}_k$ be the MLE for Φ_0 on C, computed with k observations. Then if $O \subset C$ is an open set in F_{Φ_0} containing the equivalence class of Φ_0 modulo $\sim$, one has $\widehat{\Phi}_k \in C$ a.s. for k large enough.*

Proof of Theorem 1.1. We closely follow the proof of Theorem 1 in [97]: pick $\Phi' \in F_{\Phi_0}$ such that $\Phi' \nsim \Phi_0$. Lemma 1.8 implies that there exists $\varepsilon > 0$ such that $H(\Phi_0, \Phi') < H(\Phi_0, \Phi_0) - 2\varepsilon$. Now, with the notations of Lemma 1.3, Lemma 1.5 entails that there exists $N \in \mathbb{N} \setminus \{0\}$ with

$$\left| \frac{1}{N} \mathbb{E}_{\Phi_0}(q_{0N}(\Phi')) - H(\Phi_0, \Phi') \right| < \varepsilon$$

so that

$$\frac{1}{N} \mathbb{E}_{\Phi_0}(q_{0N}(\Phi')) < H(\Phi_0, \Phi_0) - \varepsilon.$$

We then pick a neighborhood G of Φ' in F_{Φ_0} given by Lemma 1.3; in particular, for every subset $G_{\Phi'}$ of G containing Φ',

$$\mathbb{E}_{\Phi_0} \left| \ln \sup_{\varphi \in G_{\Phi'}} q_{0N}(\varphi) \right| < \infty.$$

Letting B_r be the open ball centered at Φ' with radius r, the continuity of q_{0N} gives, as $r \to 0$:

$$\ln \sup_{\varphi \in G \cap B_r} q_{0N}(\varphi) \to \ln q_{0N}(\Phi').$$

Set now $A_r = \left\{ \sup_{\varphi \in G \cap B_r} q_{0N}(\varphi) \leq 1 \right\}$, and let A_r^c denote the complement of A_r. Notice that

$$\left| \ln \sup_{\varphi \in G \cap B_r} q_{0N}(\varphi) \right| = -\ln \left[\sup_{\varphi \in G \cap B_r} q_{0N}(\varphi) \right] 1\!\!1_{A_r} + \ln \left[\sup_{\varphi \in G \cap B_r} q_{0N}(\varphi) \right] 1\!\!1_{A_r^c}$$

which entails

$$\left| \ln \sup_{\varphi \in G \cap B_r} q_{0N}(\varphi) \right| \leq |\ln q_{0N}(\Phi')| + \left| \ln \sup_{\varphi \in G} q_{0N}(\varphi) \right|.$$

We can then use the dominated convergence theorem to get a neighborhood $G_{\Phi'} \subset G$ of Φ' in F_{Φ_0} such that

$$\frac{1}{N} \mathbb{E}_{\Phi_0} \left| \ln \sup_{\varphi \in G_{\Phi'}} q_{0N}(\varphi) \right| \leq \frac{1}{N} \mathbb{E}_{\Phi_0}(\ln q_{0N}(\Phi')) + \frac{\varepsilon}{2} < H(\Phi_0, \Phi_0) - \frac{\varepsilon}{2}.$$

Now, because $\left(Z_{st} = \ln \sup_{\varphi \in G_{\Phi'}} q_{st}(\varphi)\right)$ is $\mathbb{P}_{\Phi_0}-$subadditive and ergodic, Kingman's theorem (see Theorem 1.2) implies that there exists a finite constant $H(\Phi_0,\, \Phi',\, G_{\Phi'})$ such that

$$\frac{1}{k}\mathbb{E}_{\Phi_0}\left[\ln \sup_{\varphi \in G_{\Phi'}} q_{0k}(\varphi)\right] \to H(\Phi_0,\, \Phi',\, G_{\Phi'})$$

and

$$\frac{1}{k}\ln \sup_{\varphi \in G_{\Phi'}} q_{0k}(\varphi) \to H(\Phi_0,\, \Phi',\, G_{\Phi'}) \quad \mathbb{P}_{\Phi_0} - \text{a.s.}$$

as $k \to \infty$. Theorem 1.1 in [73] entails

$$H(\Phi_0,\, \Phi',\, G_{\Phi'}) \leq \frac{1}{N}\mathbb{E}_{\Phi_0}\left[\ln \sup_{\varphi \in G_{\Phi'}} q_{0N}(\varphi)\right] < H(\Phi_0,\, \Phi_0) - \frac{\varepsilon}{2};$$

putting

$$p_{st}(\varphi \,|\, J(0) = i) = \mathcal{L}((Y_i,\, E_i,\, X_i)_{s+1 \leq i \leq t},\, \varphi \,|\, J(0) = i)$$

and remarking that for all $\varphi \in G_{\Phi'}$

$$\begin{aligned} q_{0k}(\varphi) &= \left(\sum_{i \in \mathcal{C}(\varphi)} \pi_i(\varphi)\right) \max_{i \in \mathcal{C}(\varphi)} p_{0k}(\varphi \,|\, J(0) = i) \\ &\geq \sum_{i \in \mathcal{C}(\varphi)} \pi_i(\varphi) p_{0k}(\varphi \,|\, J(0) = i) \\ &= p_{0k}(\varphi), \end{aligned}$$

one gets $\ln \sup_{\varphi \in G_{\Phi'}} p_{0k}(\varphi) - \ln \sup_{\varphi \in G_{\Phi'}} q_{0k}(\varphi) \leq 0$ and thus

$$\limsup_{k \to \infty} \left\{\frac{1}{k}\ln \sup_{\varphi \in G_{\Phi'}} p_{0k}(\varphi)\right\} \leq H(\Phi_0,\, \Phi',\, G_{\Phi'}) < H(\Phi_0,\, \Phi_0) - \frac{\varepsilon}{2}.$$

Cover now the compact set $O^c \cap C$ by the $G_{\Phi'_i}$, $1 \leq i \leq d$. We have

$$\sup_{\varphi \in O^c} \{\ln p_{0k}(\varphi) - \ln p_{0k}(\Phi_0)\} \leq \max_{1 \leq i \leq d} \left\{\ln \sup_{\varphi \in G_{\Phi'_i}} p_{0k}(\varphi) - \ln p_{0k}(\Phi_0)\right\} \to -\infty$$

with $\mathbb{P}_{\Phi_0}-$probability 1 as $k \to \infty$. This shows that necessarily $\widehat{\Phi}_k \in C$ a.s. for k large enough, and completes the proof. ■

Notice that since our model is not identifiable, any convergence result has to be stated modulo $\sim$. In that sense, this result is the best possible one.

Under some additional assumptions, one can apply the asymptotic normality theorem in [12] in order to obtain the one of our estimator. This result is rather technical: we refer the reader to Guillou *et al.* [53] for details.

1.3 An EM algorithm to compute the MLE

We now give an EM algorithm, adapted from [100], allowing us to compute the MLE in our context. Recall the available data:

- The values $0 < t_1 < \ldots < t_k = T$ of the τ_i, *i.e.* the times when N jumps (equivalently, the inter-event times $y_1, \ldots, y_k$);
- $e_1, \ldots, e_k$ the successive values of the E_i;
- $x_1, \ldots, x_k$ the successive values of the X_i.

We want to estimate

- The elements ℓ_{ij} of the transition intensity matrix L of J;
- The jump intensities λ_i of N;
- The probabilities $p(i, e)$;
- The parameters $\theta(i, e)$.

We let $0 < u_1 < \ldots < u_m < T$ be the jump times of J in the time interval $[0, T]$, $u_0 = 0$ and $u_{m+1} = T$; let further s_i be the state of J on the interval $[u_{i-1}, u_i)$, $\Delta u_i = u_i - u_{i-1}$ and z_i be the number of jumps of N in the interval $[u_{i-1}, u_i)$.

Recall that, if N' is a homogeneous Poisson process, then given $\{N'(t) = n\}$, the event times of N' in the interval $[0, t]$ are uniformly distributed. Consequently, Bayes' formula implies that the complete likelihood of the data is

$$\begin{aligned}
\mathcal{L}^c &= \pi_{s_1}\left[\prod_{i=1}^{m} \frac{\ell_{s_i, s_{i+1}}}{-\ell_{s_i, s_i}}(-\ell_{s_i, s_i}\exp(\ell_{s_i, s_i}\Delta u_i))\right]\exp(\ell_{s_{m+1}, s_{m+1}}\Delta u_{m+1}) \\
&\times \left[\prod_{i=1}^{m+1} \frac{(\lambda_{s_i}\Delta u_i)^{z_i}}{z_i!}\exp(-\lambda_{s_i}\Delta u_i)\frac{z_i!}{(\Delta u_i)^{z_i}}\right] \\
&\times \prod_{i=1}^{r}\left[\prod_{\substack{e\subset\{1,\ldots,n\}\\ e\neq\varnothing}} p(i, e)^{\mathrm{card}(A(i, e))}\prod_{j\in A(i, e)}\mathbb{P}_{\theta(i, e)}(\forall\, m\in e,\ X_m = x_{m, j})\right] p(i, \varnothing)^{\mathrm{card}(A(i, \varnothing))}
\end{aligned}$$

where $A(i, e) = \{j \in \{1, \ldots, k\} \mid J(t_j) = i,\ e_j = e\}$ stands for the set of the jump times of N when the S_m with $m \in e$ (and only these) jump and J is in state i; $A(i, \varnothing)$ stands for the set of the jump times of N when none of the S_m jumps and J is in state i.

From that identity, we deduce that the complete log-likelihood is

$$\begin{aligned}\ln \mathcal{L}^c &= \sum_{i=1}^r 1\!\!1_{\{X(0)=i\}} \ln \pi_i + \sum_{i=1}^r T_i\, \ell_{ii} + \sum_{i=1}^r \sum_{\substack{j=1\\ j\neq i}}^r m_{ij}(T)\, \ln \ell_{ij} + \sum_{i=1}^r (n_i\, \ln \lambda_i - \lambda_i\, T_i)\\ &+ \sum_{i=1}^r \sum_{e\subset\{1,\,\dots,\,n\}} \mathrm{card}(A(i,\,e))\, \ln p(i,\,e)\\ &+ \sum_{i=1}^r \sum_{\substack{e\subset\{1,\,\dots,\,n\}\\ e\neq\varnothing}} \sum_{j=1}^k \ln \mathbb{P}_{\theta(i,\,e)}(\forall\, m\in e,\ X_m = x_{m,\,j})\, 1\!\!1_{\{j\in A(i,\,e)\}}\end{aligned}$$

where

- $T_i = \displaystyle\int_0^T 1\!\!1_{\{J(u)=i\}}\, du$ is the time spent by the process J in state i until time T;
- $m_{ij}(T) = \mathrm{card}\{s\,|\,0 < s \leq T,\ J(s_-) = i,\ J(s) = j\}$ is the number of jumps from state i to state j of the process J;
- $n_i = \displaystyle\sum_{j=1}^k 1\!\!1_{\{J(t_j)=i\}}$ is the number of events that occurred when J is in state i.

The M step. We now compute the conditional expectation of $\ln \mathcal{L}^c(\Phi)$ under a parameter φ, given the event $\{N(u),\, S(u),\, 0 \leq u \leq T\}$: one has

$$\begin{aligned}\ln \widehat{\mathcal{L}^c(\Phi)} &= \sum_{i=1}^r \widehat{1\!\!1_{\{X(0)=i\}}} \ln \pi_i + \sum_{i=1}^r \widehat{T}_i\, \ell_{ii} + \sum_{i=1}^r \sum_{\substack{j=1\\ j\neq i}}^r \widehat{m_{ij}(T)}\, \ln \ell_{ij} + \sum_{i=1}^r \left(\widehat{n_i}\, \ln \lambda_i - \lambda_i\, \widehat{T}_i\right)\\ &+ \sum_{i=1}^r \sum_{e\subset\{1,\,\dots,\,n\}} \widehat{\mathrm{card}(A(i,\,e))}\, \ln p(i,\,e)\\ &+ \sum_{i=1}^r \sum_{\substack{e\subset\{1,\,\dots,\,n\}\\ e\neq\varnothing}} \sum_{j=1}^k \ln \mathbb{P}_{\theta(i,\,e)}(\forall\, m\in e,\ X_m = x_{m,\,j})\, \widehat{1\!\!1_{\{j\in A(i,\,e)\}}}\end{aligned}$$

where $\widehat{Z} = \mathbb{E}_\varphi(Z\,|\,N(u),\, S(u),\, 0 \leq u \leq T)$.

For T large enough, the first term may be neglected; recalling that

$$\ell_{ii} = -\sum_{\substack{j=1\\ j\neq i}}^r \ell_{ij},\quad p(i,\,\varnothing) = 1 - \sum_{\substack{e\subset\{1,\,\dots,\,n\}\\ e\neq\varnothing}} p(i,\,e),\quad \mathrm{card}(A(i,\,\varnothing)) = n_i - \sum_{\substack{e\subset\{1,\,\dots,\,n\}\\ e\neq\varnothing}} \mathrm{card}(A(i,\,e)),$$

one gets, for all $i,\, j \in \{1,\,\dots,\,r\}$ and $i \neq j$, the identities

$$\widehat{p}(i,\,e) = \frac{\widehat{\mathrm{card}(A(i,\,e))}}{\widehat{n_i}},\quad \widehat{\ell_{ij}} = \frac{\widehat{m_{ij}(T)}}{\widehat{T_i}},\quad \widehat{\lambda}_i = \frac{\widehat{n_i}}{\widehat{T_i}},$$

$$\sum_{j=1}^k \frac{\partial}{\partial\theta(i,\,e)} \ln \mathbb{P}_{\theta(i,\,e)}(\forall\, m\in e,\ X_m = x_{m,\,j})\Big|_{\theta(i,\,e)=\widehat{\theta}(i,\,e)} \widehat{1\!\!1_{\{j\in A(i,\,e)\}}} = 0,$$

where $\widehat{p}(i,\,e)$, $\widehat{\ell_{ij}}$ and $\widehat{\lambda_i}$ are the desired estimators, and the last set of equations is to be solved taking the properties of the statistical model $(\mathbb{P}_\theta)$ into account.

The E step. According to Lemma 1.9, if $A(e) = \bigcup_{i=1}^{r} A(i,\, e) = \{j \in \{1,\, \ldots,\, k\} \,|\, e_j = e\}$, then

$$\begin{aligned}
\widehat{T_i} &= \int_0^T \frac{\mathbb{P}_\varphi(J(v) = i,\, N(u),\, S(u),\, 0 \le u < v)}{\mathbb{P}_\varphi(N(u),\, S(u),\, 0 \le u \le T)} \\
&\quad \times \mathbb{P}_\varphi(N(u),\, S(u),\, v \le u \le T \,|\, J(v) = i)\, dv, \\
\widehat{n_i} &= \sum_{j=1}^{k} \frac{\mathbb{P}_\varphi(J(t_j) = i,\, N(u),\, S(u),\, 0 \le u \le T)}{\mathbb{P}_\varphi(N(u),\, S(u),\, 0 \le u \le T)}, \\
\widehat{1\!\!1_{\{j \in A(i,\, e)\}}} &= 1\!\!1_{\{j \in A(e)\}}\, \mathbb{P}_\varphi(J(t_j) = i \,|\, N(u),\, S(u),\, 0 \le u \le T), \\
\widehat{\mathrm{card}(A_i(e))} &= \sum_{j=1}^{k} \widehat{1\!\!1_{\{j \in A(i,\, e)\}}} = \sum_{j=1}^{k} 1\!\!1_{\{j \in A(e)\}}\, \mathbb{P}_\varphi(J(t_j) = i \,|\, N(u),\, S(u),\, 0 \le u \le T), \\
\widehat{m_{ij}(T)} &= \ell_{ij}(\varphi) \int_0^T \frac{\mathbb{P}_\varphi(J(v) = i,\, N(u),\, S(u),\, 0 \le u < v)}{\mathbb{P}_\varphi(N(u),\, S(u),\, 0 \le u \le T)} \\
&\quad \times \mathbb{P}_\varphi(N(u),\, S(u),\, v \le u \le T \,|\, J(v) = j)\, dv.
\end{aligned}$$

Let w_i be the column vector of size r with all entries except the ith equal to 0, and its ith entry equal to 1. Firstly,

$$\mathbb{P}_\varphi(N(u),\, S(u),\, 0 \le u < v,\, J(v) = i) = \pi(\varphi) \left(\prod_{j=1}^{N(v)} g(y_j,\, e_j,\, x_j,\, \varphi) \right) \overline{F}(v - t_{N(v)},\, \varphi)\, w_i.$$

Secondly, if w_i^t is the transpose of w_i,

$\mathbb{P}(N(u),\, S(u),\, v \le u \le T,\, \varphi \,|\, J(v) = i)$

$$= w_i^t\, g(t_{N(v)+1} - v,\, e_{N(v)+1},\, x_{N(v)+1},\, \varphi) \left(\prod_{j=N(v)+2}^{k} g(y_j,\, e_j,\, x_j,\, \varphi) \right) \mathbf{1},$$

and finally

$\mathbb{P}_\varphi(J(t_j) = i,\, N(u),\, S(u),\, 0 \le u \le T)$

$$= \pi(\varphi) \left(\prod_{p=1}^{j} g(y_p,\, e_p,\, x_p,\, \varphi) \right) w_i\, w_i^t \left(\prod_{p=j+1}^{k} g(y_p,\, e_p,\, x_p,\, \varphi) \right) \mathbf{1}.$$

θ is generally estimated with a numerical (e.g. quasi-Newton) method.

Procedure. Here, we describe a way of implementing our algorithm, by induction on $\ell \in \mathbb{N}$. Define, if Φ_ℓ is the parameter estimate at step ℓ,

- $G_\ell(0) = \pi(\Phi_\ell)$ and $\forall\ 0 \le q \le k-1,\ G_\ell(q+1) = G_\ell(q)\, g(y_{q+1},\, e_{q+1},\, x_{q+1},\, \Phi_\ell)$;
- $D_\ell(k) = \mathbf{1}$ and $\forall\ 0 \le q \le k-1,\ D_\ell(k-q-1) = g(y_{k-q},\, e_{k-q},\, x_{k-q},\, \Phi_\ell)\, D_\ell(k-q)$.

Set then $A_{ij}(\Phi_\ell) = B_i(\cdot,\, \Phi_\ell) = C_i(\Phi_\ell) = 0$ and do, for all $q \in \mathbb{N}$ such that $1 \le q \le k$,

$$\begin{aligned} A_{ij}(\Phi_\ell) &\leftarrow& A_{ij}(\Phi_\ell) + \int_{t_{q-1}}^{t_q} G_\ell(q-1)\, \overline{F}(t - t_{q-1},\, \Phi_\ell)\, w_i\, w_j^t\, g(t_q - t,\, e_q,\, x_q,\, \Phi_\ell)\, D_\ell(q)\, dt, \\ B_i(q,\, \Phi_\ell) &\leftarrow& G_\ell(q)\, w_i\, w_i^t\, D_\ell(q), \\ C_i(\Phi_\ell) &\leftarrow& C_i(\Phi_\ell) + B_i(q,\, \Phi_\ell). \end{aligned}$$

The estimates at step $\ell + 1$ are then

$$\widehat{p}(i,\, e) = \frac{\sum_{j=1}^{k} \mathbb{1}_{\{j \in A(e)\}}\, B_i(j,\, \Phi_\ell)}{C_i(\Phi_\ell)},\ \widehat{\ell_{ij}} = \ell_{ij}(\Phi_\ell)\, \frac{A_{ij}(\Phi_\ell)}{A_{ii}(\Phi_\ell)},\ \widehat{\lambda}_i = \frac{C_i(\Phi_\ell)}{A_{ii}(\Phi_\ell)},$$

and the $\widehat{\theta}(i,\, e)$ that maximise the functionals

$$\theta \mapsto \sum_{j=1}^{k} \ln \mathbb{P}_\theta(\forall\, m \in e,\ X_m = x_{m,\,j})\, B_i(j,\, \Phi_\ell)\, \mathbb{1}_{\{j \in A(e)\}}.$$

1.4 *A posteriori* reconstruction of the states

Once the parameters of the model are estimated, it can be interesting to estimate the successive states of the Markov chain (J_i). To this end, we can adapt the procedure described in [85]: consider the complete log-likelihood as a function of the missing data:

$$(j_0,\, \dots,\, j_k) \mapsto \ln\left(\pi_{j_0}\left(\widehat{\Phi}\right)\right) + \sum_{i=1}^{k} \ln g_{j_{i-1},\, j_i}\left(y_i,\, e_i,\, x_i,\, \widehat{\Phi}\right).$$

An estimator of $(j_0,\, \dots,\, j_k)$ is then a $(k+1)-$tuple $\left(\widehat{j_0},\, \dots,\, \widehat{j_k}\right)$ which maximises this functional. Such an estimator has excellent properties, see Caliebe [17]. From a practical point of view, one may reconstruct the states using the Viterbi algorithm (see [105]), namely:

1. Set $V_j = 0$ and $C_j = [j]$ for all $j \in \{1,\, \dots,\, r\}$, and $q = 1$.
2. If $q \ge k+1$, go to step 6. Otherwise, set
$$\alpha_{i,\,j}^{(q)} = \ln g_{ij}\left(y_{k-q+1},\, e_{k-q+1},\, x_{k-q+1},\, \widehat{\Phi}\right).$$
3. For all $i,\ j \in \{1,\, \dots,\, r\}$, compute $\beta_{i,\,j}^{(q)} = \alpha_{i,\,j}^{(q)} + V_j$ and an index $j_i^{(q)}$ such that $\beta_{i,\,j_i^{(q)}}^{(q)} = \max_{j \in \{1,\, \dots,\, r\}} \beta_{i,\,j}^{(q)}$.
4. For all $i \in \{1,\, \dots,\, r\}$, replace V_i by $\beta_{i,\,j_i^{(q)}}^{(q)}$ and C_i by $[j_i^{(q)},\, C_i]$.
5. Replace q by $q+1$ and go back to step 2.
6. Find an index i such that $V_i = \max_{j \in \{1,\, \dots,\, r\}} V_j$.

An estimate of the states is then the sequence $\left(\widehat{j_0},\, \dots,\, \widehat{j_k}\right) = C_i$.

1.5 Numerical illustrations

1.5.1 Computing a first estimate

Providing a first estimate for an iterative algorithm is usually a daunting task. Here, we describe a procedure, adapted from the one given in [85], which worked fairly well in our examples:

1. Compute the average of the inter-event times $\widehat{\lambda^*} = k/T$, and mobile averages of the inter-event times y_j, denoted by z_j (for the first and last times of the observed sample, put $z_j = y_j$).

2. Set $\widehat{J}(\cdot) = 0$; pick $q_1 \leq 1 < q_2 < \ldots < q_{r-1}$. For all $j \in \{1, \ldots, k\}$:

   (a) if $z_j > \dfrac{1}{q_1 \widehat{\lambda^*}}$, set $\widehat{J}(t_j) = 1$;

   (b) for all $i \in \{1, \ldots, r-2\}$, if $\dfrac{1}{q_{i+1}\widehat{\lambda^*}} < z_j \leq \dfrac{1}{q_i \widehat{\lambda^*}}$, set $\widehat{J}(t_j) = i + 1$;

   (c) if $z_j \leq \dfrac{1}{q_{r-1}\widehat{\lambda^*}}$, set $\widehat{J}(t_j) = r$.

3. Compute $\widehat{n_i} = \displaystyle\sum_{j=1}^{k-1} \mathbb{1}_{\left\{\widehat{J}(t_j)=i\right\}}$ for $i \in \{1, \ldots, r\}$.

4. Compute, for all $i, j \in \{1, \ldots, r\}$

$$\widehat{P}_{ij} = \frac{1}{\widehat{n_i}} \sum_{\ell=2}^{k} \mathbb{1}_{\left\{\widehat{J}(t_{\ell-1})=i,\, \widehat{J}(t_\ell)=j\right\}},$$

   which is the first estimate of P_{ij}, the probability that the Markov chain (J_i) jumps from state i to state j.

5. Calculate, for all $i \in \{1, \ldots, r\}$, $\widehat{\pi_i} = \dfrac{1}{k}\left(\widehat{n_i} + \mathbb{1}_{\left\{\widehat{J}(t_k)=i\right\}}\right)$, the first estimate of π_i.

6. Let $\lambda^* = \sum_{i=1}^{r} \lambda_i\, a_i$ be the average jump rate of N; thanks to the identities

$$\forall\, i \in \{1, \ldots, r\}, \quad \lambda_i = \lambda^*\, \pi_i\, a_i^{-1} \quad \text{and} \quad L = \Lambda(\mathrm{Id} - P^{-1}),$$

   and under the constraint $\sum_{i=1}^{r} a_i = 1$, consider L and Λ as functions of $a_1, \ldots, a_{r-1}$, and, given $\widehat{\lambda^*}, \widehat{\pi_1}, \ldots, \widehat{\pi_r}, \widehat{P}, y_1, \ldots, y_k$ and $\widehat{J}$, maximise the complete likelihood with respect to the parameters $a_1, \ldots, a_{r-1}$: let $\widehat{a_1}, \ldots, \widehat{a_{r-1}}$ be the estimates obtained this way.

7. For all $i \in \{1, \ldots, r\}$, compute $\widehat{\lambda_i} = \widehat{\lambda^*}\, \widehat{\pi_i}\, \widehat{a_i}^{-1}$, let

$$\widehat{\Lambda} = \mathrm{diag}\left(\widehat{\lambda_i},\, 1 \leq i \leq r\right) \quad \text{and} \quad \widehat{L} = \widehat{\Lambda}\left(\mathrm{Id} - \widehat{P}^{-1}\right).$$

   These are rough estimates for Λ and L.

8. Use $\widehat{L}$ and $\widehat{\Lambda}$ as initial values for an EM algorithm to provide estimates for L and Λ (see [100]), which we denote by $\overline{L}$ and $\overline{\Lambda}$. Compute the corresponding stationary distributions $\overline{a}$ and $\overline{\pi}$.

9. Perform a state reconstruction of J with the Viterbi algorithm using $\overline{L}$ and $\overline{\Lambda}$, and let $\overline{J}$ be the process obtained this way.

10. For all $i \in \{1, \ldots, r\}$, calculate $\overline{n_i} = \sum_{j=1}^{k-1} \mathbb{1}_{\{\overline{J}(t_j)=i\}}$.

11. For all $e \subset \{1, \ldots, n\}$, compute

$$\overline{p}(i, e) = \frac{1}{\overline{n_i}} \sum_{j=1}^{k-1} \mathbb{1}_{\{\overline{J}(t_j)=i\}} \mathbb{1}_{\{\forall m \in \{1, \ldots, n\},\ S_m(t_j) - S_m(t_{j-1}) > 0 \Leftrightarrow m \in e\}}$$

which is the initial estimate of $p(i, e)$.

12. For all $i = 1, \ldots, r$ and $e \neq \emptyset$, consider the x_j such that $\overline{J}(t_j) = i$ and $e_j = e$ as realisations of independent and identically distributed random variables with parameter $\theta(i, e)$, and estimate $\theta(i, e)$ with a standard method (maximum likelihood method for instance).

This procedure is adapted in the particular case when $\lambda_1 < \ldots < \lambda_r$ strongly differ, which shall be the case in our numerical study below.

1.5.2 A non-life insurance example

We now use our algorithm on a real set of non-life insurance data. From January 2004 to November 2009, 594 accidents corresponding to blazes causing industrial damage or losses were observed. The days of these events were recorded, and so were, if necessary, the compensations for the victims; the processes N and S obtained this way are shown on Figure 1.1–1.2. This situation corresponds to the case $n = 1$ of our model. We finally choose $r = 2$, which is justified by the fact that the MLE, computed only for L and Λ with $r = 3$ sets all parameters corresponding to the third state to 0. Before modeling the claims themselves, the parameters of this model are

- ℓ_{12} and ℓ_{21}, the jump rates of the hidden Markov process J;
- λ_1 and λ_2, the jump intensities of the shock counting process N;
- $p(1, 1)$ and $p(2, 1)$, the probabilities that, when an accident happens, the insurance firm has to compensate.

As for the claim sizes, a quick analysis of the data shows that some claims have a small size and a few others are very large, which prevents us from modeling the situation by a log-Normal, Gamma or Generalised Pareto distribution (GPD). In actuarial statistics, one may either try to separate

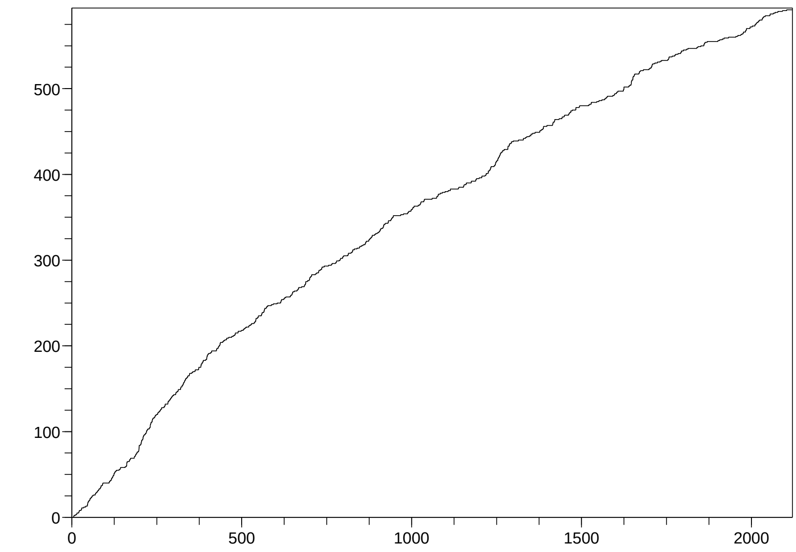


Figure 1.1: The counting process N

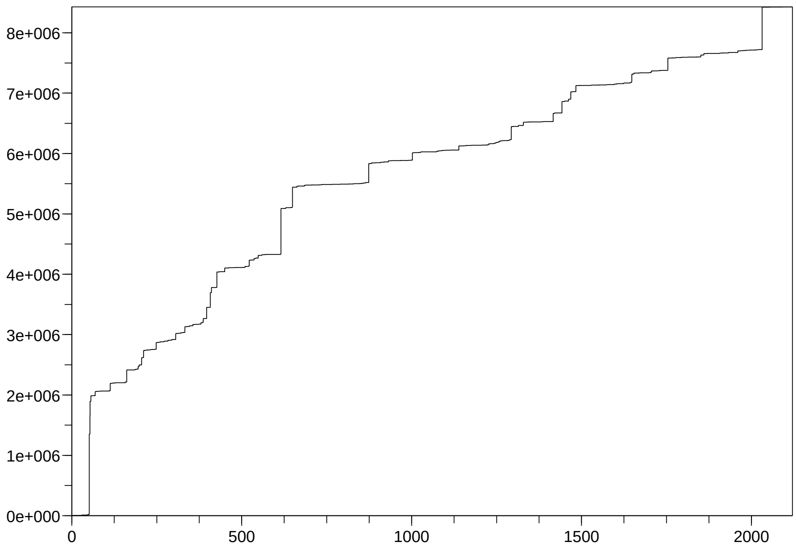


Figure 1.2: The loss process S

so-called attritional claims and large claims thanks to some threshold as in many Solvency II partial internal models, or deal directly with a mixture of distributions, or with a distribution that looks like Lognormal or Gamma distributions for small values and gets more and more Pareto-type for large values, like the Champernowne distribution (see Champernowne [20, 21] and Johnson *et al.* [69]). Another possibility is to use a classical kernel density estimator after transforming the data (see Buch-Larsen *et al.* [15]). Here, we use a mixture of a light-tailed and a heavy-tailed distribution, namely a Gamma distribution and a GPD. $\mathbb{P}_\theta$ then has density

$$x \mapsto q\,\frac{(bx)^{a-1}}{\Gamma(a)}\,be^{-bx}\,\mathbb{1}_{\{x>0\}} + (1-q)\,\frac{1}{\sigma}\left(1+\frac{\xi(x-\mu)}{\sigma}\right)^{-1-1/\xi}\mathbb{1}_{\{x>\mu\}}$$

where a, b, σ, $\xi > 0$, $0 < q < 1$ and $\mu = 49.33$ is the minimal (observed) claim size (the unit is the euro).

Consequently, the parameters to be estimated are ℓ_{12}, ℓ_{21}, λ_1, λ_2, $p(1, 1)$, $p(2, 1)$, a_1, a_2, b_1, b_2, σ_1, σ_2, ξ_1, ξ_2, q_1 and q_2.

Estimating the parameters with the EM algorithm, with a quasi-Newton algorithm to estimate the parameters a_i, b_i, σ_i, ξ_i and q_i during the M step gives the following results:

$$\widehat{L} = \begin{pmatrix} -0.0065 & 0.0065 \\ 0.0018 & -0.0018 \end{pmatrix}, \ \widehat{\Lambda} = \begin{pmatrix} 0.462 & 0 \\ 0 & 0.214 \end{pmatrix},$$

$$\widehat{p}(\cdot, 1) = \begin{pmatrix} 0.963 & 0 \\ 0 & 0.947 \end{pmatrix}, \ \widehat{p}(\cdot, 0) = \begin{pmatrix} 0.037 & 0 \\ 0 & 0.053 \end{pmatrix},$$

$$\widehat{a} = \begin{pmatrix} 4.52 \\ 4.14 \end{pmatrix}, \ \widehat{b} = \begin{pmatrix} 0.011 \\ 0.0073 \end{pmatrix}, \ \widehat{\sigma} = \begin{pmatrix} 1145 \\ 1216 \end{pmatrix}, \ \widehat{\xi} = \begin{pmatrix} 1.45 \\ 1.31 \end{pmatrix}, \ \widehat{q} = \begin{pmatrix} 0.230 \\ 0.335 \end{pmatrix}.$$

Theoretically, the claim sizes thus have infinite means in both states. This means that the tail of the claim size distribution is very heavy. However, reinsurance mechanisms and other guarantees may enable the insurer to provide insurance coverage of those risks up to some high threshold level. A further analysis then shows that

1. Sojourn times in state 1 are on average 3.5 times shorter than in state 2;
2. There are more accidents when J is in state 1 than in state 2;
3. Because $\widehat{p}(1, 1)$ is slightly greater than $\widehat{p}(2, 1)$, these accidents cause more losses to the insurance firm;
4. Losses in state 1 are more likely to be heavy-tailed than in state 2.

An *a posteriori* reconstruction of the states of J is given in Figure 1.3.

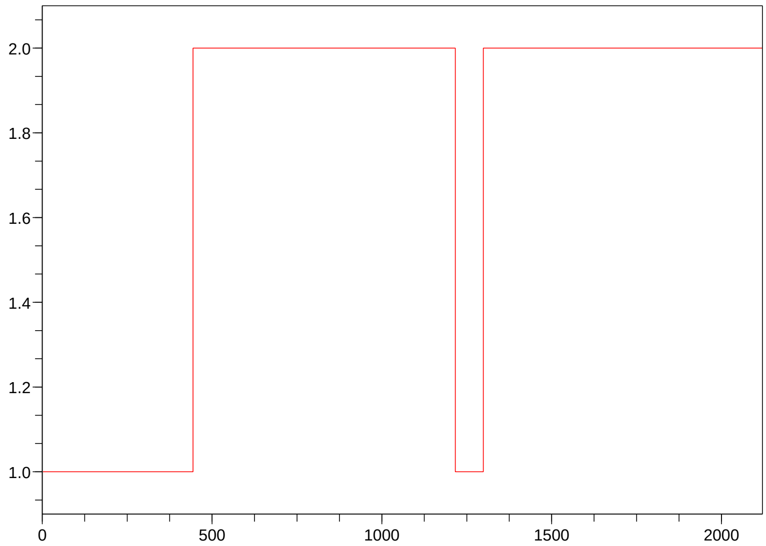


Figure 1.3: *A posteriori* reconstruction of the states of J

1.5.3 A life insurance data set

Let us now present an application in the life insurance field. From January 2006 to July 2010, 1507 closures of savings accounts (also called surrenders) were observed. The months of these events were recorded, along with the amount of money withdrawn. Early surrenders can be regarded as claims for the insurance company in some cases, because it corresponds to a drop in future business, and because sometimes the insurer has been unable to charge all the fees (that are often partly paid by the policyholder at each time period and not upfront) before the surrender. Surrender risk is complex: tax and penalty relief, interest rate levels, competition between insurance companies, as well as other factors are at stake. For a review on surrender triggers, the interested reader might consult Loisel and Milhaud [83] or Milhaud *et al.* [87]. In the present data study, we are interested in the big picture in a quite stable regime (and not in prediction of future surrender rates): in the considered period, the portfolio seems to have been pretty stable, mainly sensitive to external competition (which is difficult to observe in practice). We assume that conditionally with respect to the state of the environment, the probability for one policyholder to surrender his contract does not depend on the amount of savings. To set a precise date for the kth surrender, we draw a uniform random variable and add it to the month of this event to obtain an exact date. Here, the "claims" are the amounts of money withdrawn; the processes N and S are represented on Figure 1.4–1.5, the unit of time being the month. Again, this situation fits the case $n = 1$ of our model; we use a

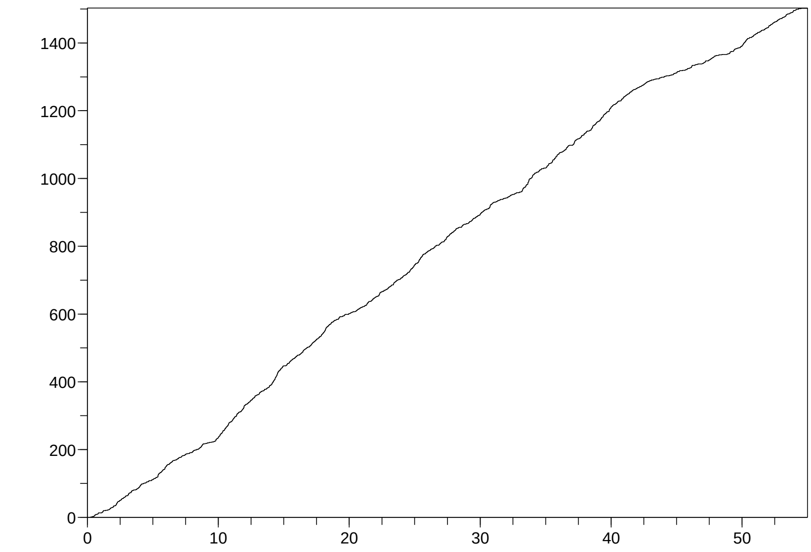


Figure 1.4: The counting process N

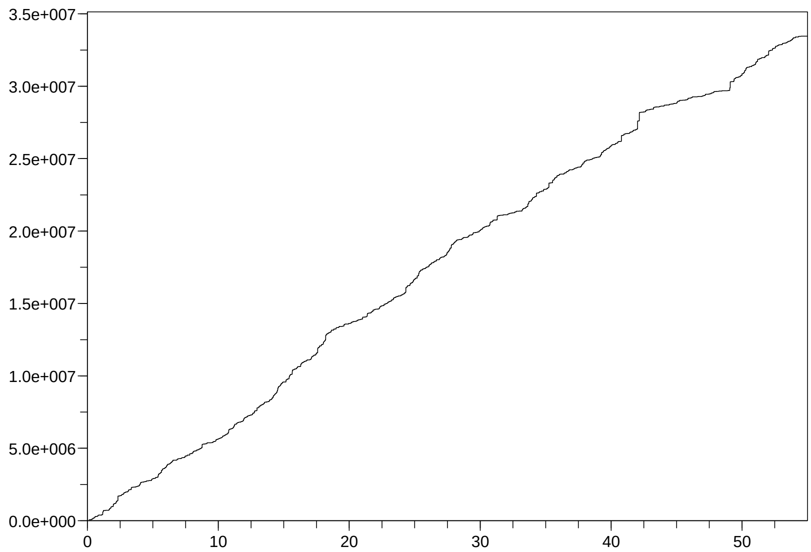


Figure 1.5: The process S representing the cumulative amount of money withdrawn

two-state model for this situation, so that the parameters are

- ℓ_{12} and ℓ_{21}, the jump rates of the hidden Markov process J;
- λ_1 and λ_2, the jump intensities of the shock counting process N.

Note that in this example, there is no need to estimate $p(1, 1)$ and $p(2, 1)$. In state 1, we use a mixture of a light-tailed and a heavy-tailed distribution, namely a Weibull distribution and a GPD, the density of $\mathbb{P}_\theta$ then being

$$x \mapsto q\,\frac{a}{b}\left[\frac{x-\mu}{b}\right]^{a-1} e^{-((x-\mu)/b)^a}\, \mathbb{1}_{\{x>\mu\}} + (1-q)\,\frac{1}{\sigma}\left(1+\frac{\xi(x-\mu)}{\sigma}\right)^{-1-1/\xi} \mathbb{1}_{\{x>\mu\}}$$

where a, b, σ, $\xi > 0$, $0 < q < 1$ and $\mu = 1.1$ is the minimal (observed) amount (the unit is the euro). In state 2, we fit a GPD, whose density is

$$x \mapsto \frac{1}{\sigma}\left(1+\frac{\xi(x-\mu)}{\sigma}\right)^{-1-1/\xi} \mathbb{1}_{\{x>\mu\}} \tag{1.1}$$

where σ, $\xi > 0$ and again $\mu = 1.1$. Of course, surrender amounts are not completely independent at the microscopic level as each policyholder has a certain balance on his savings account that is known at a precise date. We are aware that in theory, the X_i are not independent and identically distributed in each state, but in practice there are enough policyholders and enough randomness in the surrendered amounts for this assumption to be acceptable in practice at the macroscopic level in each state of the environment (this is supported by statistical tests).

Consequently, the parameters to be estimated are ℓ_{12}, ℓ_{21}, λ_1, λ_2, a, b, σ_1, σ_2, ξ_1, ξ_2 and q.

Estimating the parameters with the EM algorithm, with a quasi-Newton algorithm to estimate the parameters a, b, σ_i, ξ_i and q during the M step gives the following results:

$$\widehat{L} = \begin{pmatrix} -0.254 & 0.254 \\ 0.373 & -0.373 \end{pmatrix},\ \widehat{\Lambda} = \begin{pmatrix} 34.2 & 0 \\ 0 & 17.4 \end{pmatrix},$$
$$\widehat{a} = 1.65,\ \widehat{b} = 9141,\ \widehat{\sigma} = \begin{pmatrix} 22350 \\ 14591 \end{pmatrix},\ \widehat{\xi} = \begin{pmatrix} 0.17 \\ 0.40 \end{pmatrix},\ \widehat{q} = 0.306.$$

An *a posteriori* reconstruction of the states of J is shown in Figure 1.6.

Note that results show that during some fierce competition periods, surrender rates become more important (they double from one state to the other). In the state where surrender rates are higher, the surrendered amount fitted distribution is composed of a light-tailed part and a heavy-tailed part, whereas for smaller surrender rates, this distribution does not incorporate any light-tailed part. This suggests that policies with smaller facial amounts are more sensitive to changes in the environment. Once again, here, the heavy-tailed part must be regarded as a statistical fit, and the tail would have to be cut at an appropriate level *a posteriori*.

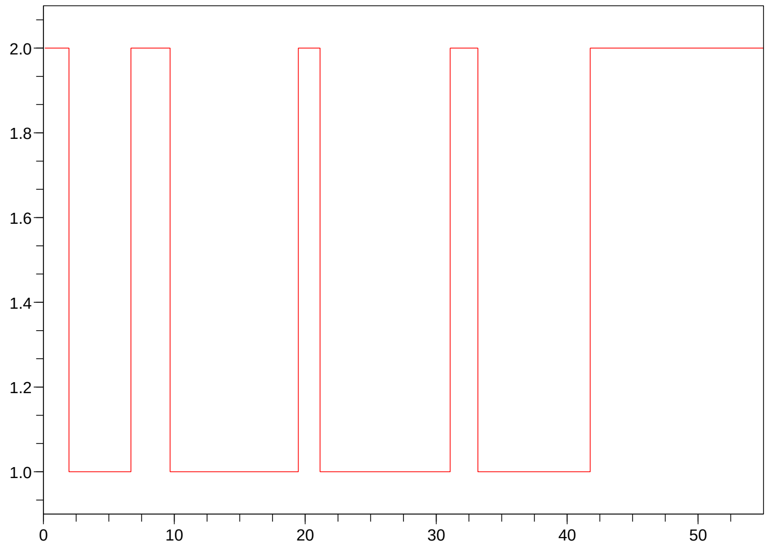


Figure 1.6: *A posteriori* reconstruction of the states of J

1.5.4 Simulations in the multivariate setting

Motivation

One of the main purposes of insurance is risk diversification and mutualisation: the law of large numbers and the central limit theorem often apply in practice when independence between individual risks is not too unrealistic. For example, this works fairly well for motor insurance portfolios (without motor liability insurance) at the national level. However, when it comes to hurricane risks or earthquake risks, individual risks are only conditionally independent with respect to the occurrence or not of such events in the country. This correlation makes it difficult to diversify those risks at the national level, and one often uses reinsurance: risks are then diversified at the global level (floods in Australia, tsunamis in Asia, hurricanes in the East Coast of North America, earthquakes in Japan, Monte Carlo and San Francisco, storms in Europe for instance). Nevertheless, those risks are not really independent, as some (often ignored) correlation factors are present. Even if they are geographically scattered, meteorological phenomena like the El Niño-La Niña Southern Oscillation (ENSO) may simultaneously influence claim occurrence and severity in those different zones. For example, it is now accepted that the probabilities of severe floods in Australia, strong snowstorms in North America and hurricanes on the US East Coast increase during La Niña episodes, while other kinds of events are more likely during El Niño episodes. To build a model for ENSO and to understand all its impacts on different areas of the world is far beyond the scope of this paper. Of course, ENSO is observed and can be (partly) measured, its behaviour is not really Markovian and claim arrival processes feature seasonality. There are certainly other kinds of unobserved environment processes that jointly modulate claim processes in different regions of the world. In our illustrative example, we just imagine that some unobserved Markov process influences claim frequencies in three regions A ($k = 1$), B ($k = 2$) and C ($k = 3$). Regions A and B are assumed to be close to each other, so that common shocks (events that simultaneously cause claims in both regions) are possible. In our example, phase changes are more frequent than for the ENSO cycle. We simulate the corresponding multivariate risk process, and we check whether it would be possible or not for us to estimate the parameters of the model and to re-build the states of the environment modulating process (without observing it of course).

A model with 2 states of the environment

We first assume that $r = 2$: in state 1, claims are less frequent and less severe in the three zones, and common shocks are not present, *i.e.* $p(1, e) = 0$ if $\mathrm{Card}(e) \geq 2$. In state 2, claims are more likely and more severe in average, and common shocks are possible for zones A and B, *i.e.* $p\,(2, \{1, 2\}) > 0$. Take $\lambda_1 = 20$, $\lambda_2 = 200$, $p(1, \{1\}) = p(1, \{2\}) = 0.3$, $p(1, \{3\}) = 0.4$, $p(2, \{1\}) = p(2, \{2\}) = 0.2$, $p(2, \{3\}) = 0.4$ and $p\,(2, \{1, 2\}) = 0.2$. The univariate claim severity distributions are chosen to be GP distributed as in (1.1), with the parameters being

$$\begin{aligned}
\mu(\{1\}) = \mu(\{2\}) = \mu(\{3\}) = 1,\\
\sigma(1, \{1\}) = \sigma(1, \{2\}) = \sigma(1, \{3\}) = 1,\\
\sigma(2, \{1\}) = \sigma(2, \{2\}) = \sigma(2, \{3\}) = 20,\\
\xi(1, \{1\}) = \xi(1, \{2\}) = \xi(1, \{3\}) = 1/2,\\
\xi(2, \{1\}) = \xi(2, \{2\}) = \xi(2, \{3\}) = 2.
\end{aligned}$$

Univariate claims are therefore more severe in average and in the tail for state 2 for all three lines. As far as the bivariate claims in state 2 are concerned, we model them by a bivariate GPD as in Cai and Tan [16] and Chiragiev and Landsman [22]; namely, their density has the form

$$(x, y) \mapsto \frac{\alpha(\alpha+1)}{\sigma_1\,\sigma_2}\left(1 + \frac{x-\mu_1}{\sigma_1} + \frac{x-\mu_2}{\sigma_2}\right)^{-\alpha-2} \mathbb{1}_{\{x>\mu_1\}}\,\mathbb{1}_{\{y>\mu_2\}}$$

where α, μ_1, μ_2, σ_1, $\sigma_2 > 0$, and we choose

$$\mu(\{1, 2\}) = \begin{pmatrix} 3 \\ 3 \end{pmatrix}, \quad \sigma(2, \{1, 2\}) = \begin{pmatrix} 30 \\ 20 \end{pmatrix}, \quad \alpha(2, \{1, 2\}) = 2.$$

Assume that we observe the multivariate claim process during 30 years, and that the average time spent in state 1 (before switching to state 2) is 1 year, while the average time spent in state 2 (before switching to state 1) is 3 months. Namely, $\ell_{12} = 1$ and $\ell_{21} = 4$.

The estimate of $\mu(\{e\})$, $e \neq \varnothing$ is chosen as the vector of the minima of the claims arising when a shock affects simultaneously the lines $L_{k_1}, \ldots, L_{k_p}$, with $e = \{k_1, \ldots, k_p\}$. Results are given below:

$$\widehat{\ell_{12}} = 1.064,\ \widehat{\ell_{21}} = 3.891,$$
$$\widehat{\lambda_1} = 21.21,\ \widehat{\lambda_2} = 195.7,$$
$$\widehat{p}(1, \{1\}) = 0.340,\ \widehat{p}(1, \{2\}) = 0.276,\ \widehat{p}(1, \{3\}) = 0.384,$$
$$\widehat{p}(2, \{1\}) = 0.227,\ \widehat{p}(2, \{2\}) = 0.182,\ \widehat{p}(2, \{3\}) = 0.394,$$
$$\widehat{p}\,(2, \{1, 2\}) = 0.197,$$
$$\widehat{\mu}(\{1\}) = 1.002,\ \widehat{\mu}(\{2\}) = 1.000,\ \widehat{\mu}(\{3\}) = 1.004,$$
$$\widehat{\sigma}(1, \{1\}) = 0.950,\ \widehat{\sigma}(1, \{2\}) = 1.393,\ \widehat{\sigma}(1, \{3\}) = 0.999,$$
$$\widehat{\sigma}(2, \{1\}) = 18.22,\ \widehat{\sigma}(2, \{2\}) = 19.18,\ \widehat{\sigma}(2, \{3\}) = 24.83,$$
$$\widehat{\xi}(1, \{1\}) = 0.552,\ \widehat{\xi}(1, \{2\}) = 0.507,\ \widehat{\xi}(1, \{3\}) = 0.493,$$
$$\widehat{\xi}(2, \{1\}) = 2.206,\ \widehat{\xi}(2, \{2\}) = 2.220,\ \widehat{\xi}(2, \{3\}) = 1.888,$$
$$\widehat{\mu}(\{1, 2\}) = \begin{pmatrix} 3.142 \\ 3.040 \end{pmatrix},\ \widehat{\sigma}(2, \{1, 2\}) = \begin{pmatrix} 25.98 \\ 18.06 \end{pmatrix},\ \widehat{\alpha}(2, \{1, 2\}) = 1.79.$$

The estimation procedure yields fairly good results and the states are correctly retrieved, see Figure 1.9. Of course, if the observation period was shorter, or if the phase change intensities were smaller, then it would be impossible to estimate transition rates accurately.

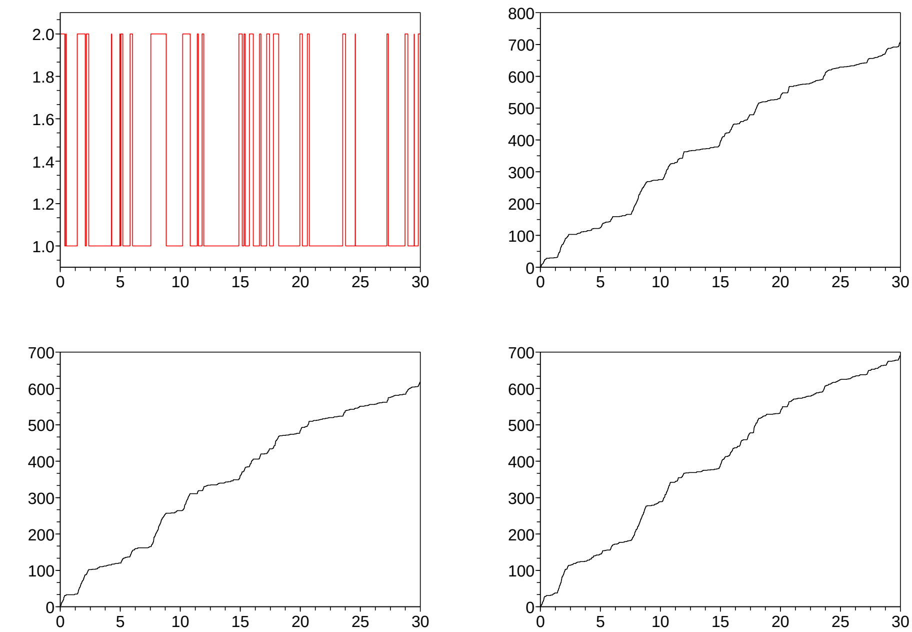


Figure 1.7: The counting processes: top left: the true process J, top right: the counting process related to S_1, bottom left: the counting process related to S_2, bottom right: the counting process related to S_3

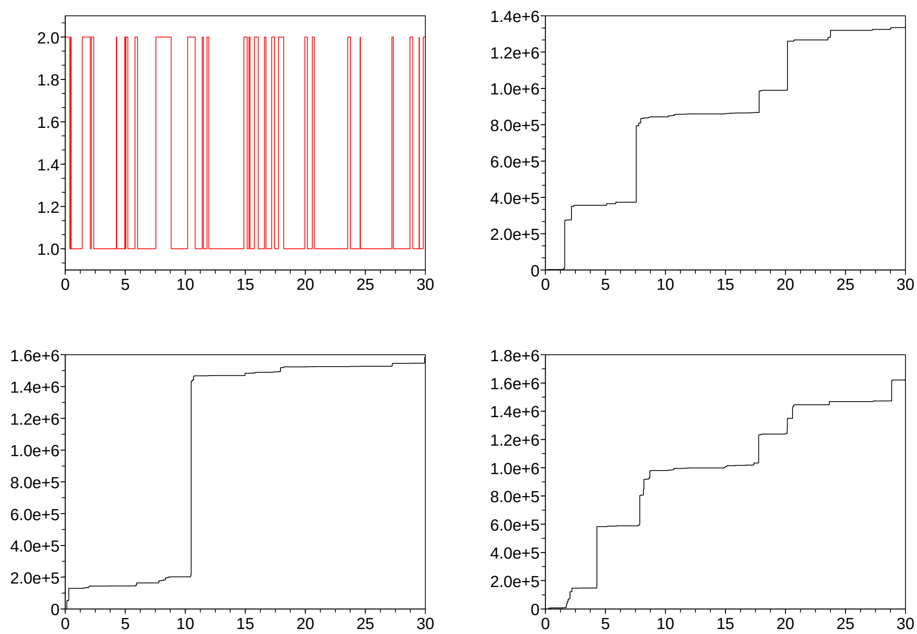


Figure 1.8: The loss processes S_k, top left: the true process J, top right: the loss process S_1, bottom left: the loss process S_2, bottom right: the loss process S_3

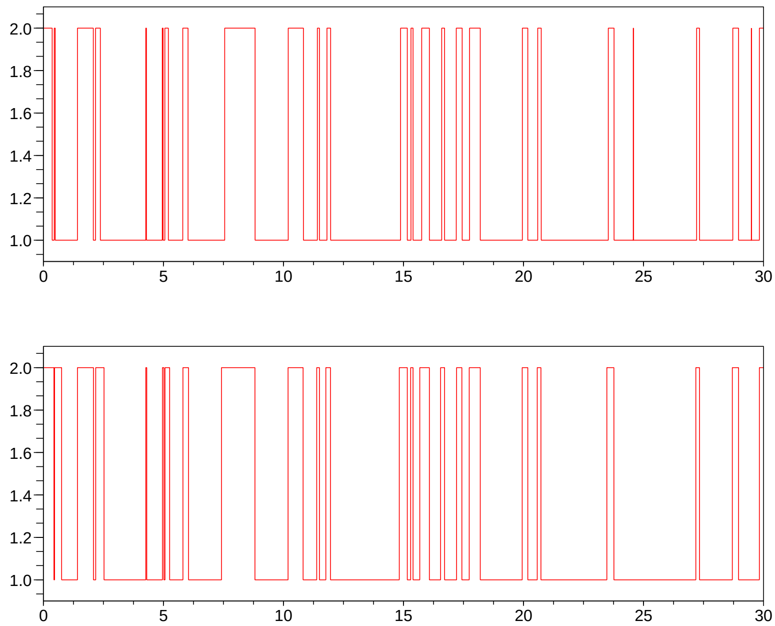


Figure 1.9: Reconstruction of the hidden Markov process J: top: the true process J, bottom: the reconstructed process $\widehat{J}$

A model with 3 states of the environment

We now assume that $r = 3$ and that common shocks are not present, *i.e.* for $i \in \{1, 2, 3\}$, $p(i, e) = 0$ if $\text{Card}(e) \geq 2$. In state 1, claims are not very frequent and not very severe in the three zones. In state 2, claims are more likely and more severe in average for the three zones. State 3 corresponds to exceptional conditions that favor extremely severe claims for zones A and B but protect zone C. Take $\lambda_1 = 20$, $\lambda_2 = 200$, $\lambda_3 = 1000$, $p(1, \{1\}) = p(1, \{2\}) = 0.3$, $p(1, \{3\}) = 0.4$, $p(2, \{1\}) = p(2, \{2\}) = 0.3$, $p(2, \{3\}) = 0.4$, $p(3, \{1\}) = p(3, \{2\}) = 0.45$ and $p(3, \{3\}) = 0.1$. The claim severity distributions are once again modeled by GP distributions, with

$$\mu(\{1\}) = \mu(\{2\}) = \mu(\{3\}) = 1,$$
$$\sigma(1, \{1\}) = \sigma(1, \{2\}) = \sigma(1, \{3\}) = 1,$$
$$\sigma(2, \{1\}) = \sigma(2, \{2\}) = \sigma(2, \{3\}) = 20,$$
$$\sigma(3, \{1\}) = \sigma(3, \{2\}) = 200, \ \sigma(3, \{3\}) = 0.5,$$
$$\xi(1, \{1\}) = \xi(1, \{2\}) = \xi(1, \{3\}) = 1/4,$$
$$\xi(2, \{1\}) = \xi(2, \{2\}) = \xi(2, \{3\}) = 1/2,$$
$$\xi(2, \{1\}) = \xi(2, \{2\}) = 1, \ \xi(2, \{3\}) = 1/3,$$

These parameters are chosen so that claims for zone C in state 3 are very small compared to those for zones A and B. Assume that we observe the multivariate claim process during 30 years, that the average time spent in state 1 (before switching to another state) is 1 year (resp. 3 months for state 2, 1 month for state 3), and that jumps from state 1 to state 3 or from state 3 to state 1 are a.s. impossible. Assume finally that when one leaves state 2, the probability to go to state 1 is 2/3. The intensity transition parameters are then $\ell_{12} = 1$, $\ell_{13} = 0$, $\ell_{21} = 8/3$, $\ell_{23} = 4/3$, $\ell_{31} = 0$, $\ell_{32} = 12$.

Again, the estimate of $\mu(\{i\})$, $i \in \{1, 2, 3\}$ is chosen as the minimum of the claims affecting line i. The results are the following:

$$
\begin{gathered}
\widehat{\ell_{12}} = 1.691,\ \widehat{\ell_{13}} = 0,\ \widehat{\ell_{21}} = 2.513,\ \widehat{\ell_{23}} = 1.288,\ \widehat{\ell_{31}} = 0,\ \widehat{\ell_{32}} = 10.76,\\
\widehat{\lambda_1} = 27.44,\ \widehat{\lambda_2} = 198.3,\ \widehat{\lambda_3} = 976.3,\\
\widehat{p}(1, \{1\}) = 0.289,\ \widehat{p}(1, \{2\}) = 0.332,\ \widehat{p}(1, \{3\}) = 0.379,\\
\widehat{p}(2, \{1\}) = 0.306,\ \widehat{p}(2, \{2\}) = 0.298,\ \widehat{p}(2, \{3\}) = 0.396,\\
\widehat{p}(3, \{1\}) = 0.448,\ \widehat{p}(3, \{2\}) = 0.444,\ \widehat{p}(3, \{3\}) = 0.109,\\
\widehat{\mu}(\{1\}) = 1.003,\ \widehat{\mu}(\{2\}) = 1.001,\ \widehat{\mu}(\{3\}) = 1.000,\\
\widehat{\sigma}(1, \{1\}) = 1.013,\ \widehat{\sigma}(1, \{2\}) = 1.065,\ \widehat{\sigma}(1, \{3\}) = 1.016,\\
\widehat{\sigma}(2, \{1\}) = 19.17,\ \widehat{\sigma}(2, \{2\}) = 19.85,\ \widehat{\sigma}(2, \{3\}) = 20.83,\\
\widehat{\sigma}(3, \{1\}) = 191.9,\ \widehat{\sigma}(3, \{2\}) = 191.2,\ \widehat{\sigma}(3, \{3\}) = 0.472,\\
\widehat{\xi}(1, \{1\}) = 0.356,\ \widehat{\xi}(1, \{2\}) = 0.298,\ \widehat{\xi}(1, \{3\}) = 0.251,\\
\widehat{\xi}(2, \{1\}) = 0.504,\ \widehat{\xi}(2, \{2\}) = 0.437,\ \widehat{\xi}(2, \{3\}) = 0.433,\\
\widehat{\xi}(2, \{1\}) = 0.957,\ \widehat{\xi}(2, \{2\}) = 0.948,\ \widehat{\xi}(2, \{3\}) = 0.443.
\end{gathered}
$$

Once again, results are correct because we have enough environment process changes during our observation period, see Figure 1.12. Results are slightly less accurate than in the 2-dimensional case, for example regarding λ_1. Note that although results would be completely inaccurate for large numbers of lines or numbers of states of the environment, estimation and reconstruction results are acceptable for 3 lines and 3 states of the environment.

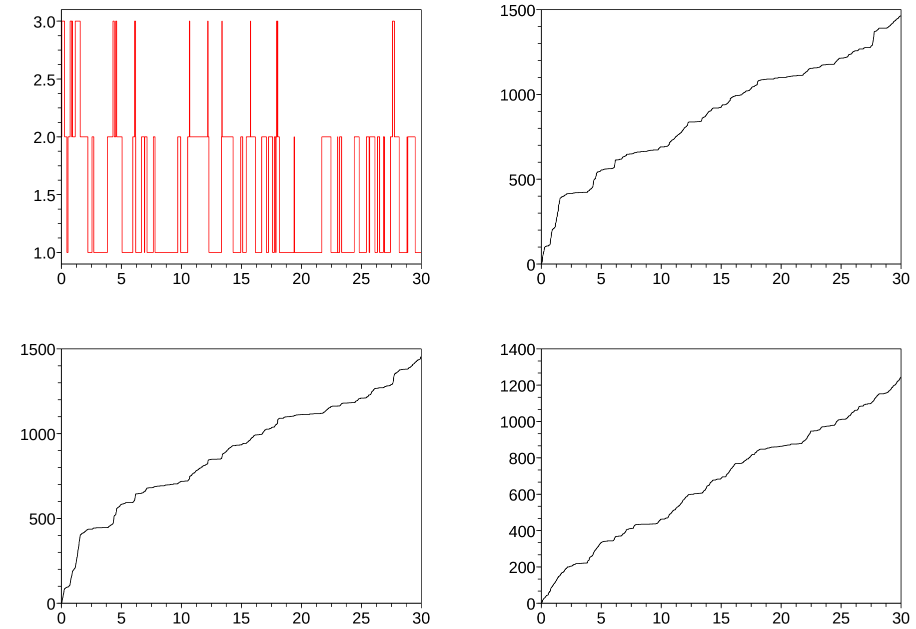


Figure 1.10: The counting processes: top left: the true process J, top right: the counting process related to S_1, bottom left: the counting process related to S_2, bottom right: the counting process related to S_3

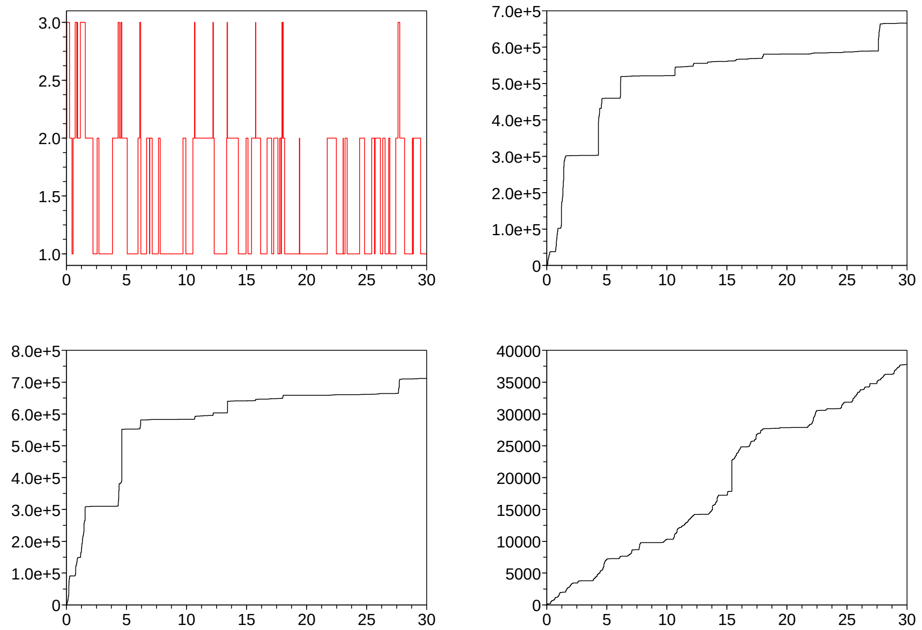


Figure 1.11: The loss processes S_k, top left: the true process J, top right: the loss process S_1, bottom left: the loss process S_2, bottom right: the loss process S_3

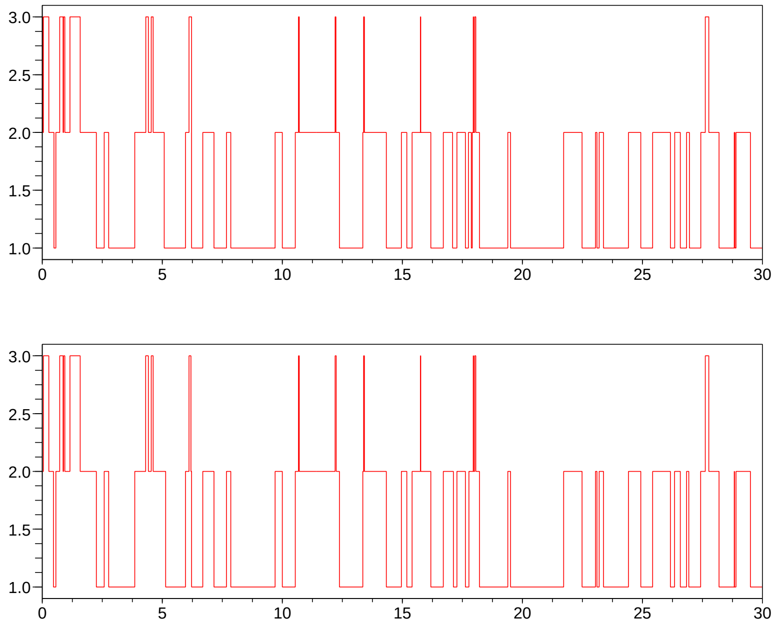


Figure 1.12: Reconstruction of the hidden Markov process J: top: the true process J, bottom: the reconstructed process $\widehat{J}$

1.6 Appendix A: Auxiliary results

We first state three technical lemmas in order to adapt the intermediate results of [97]:

Lemma 1.1. *Let $(X, \mathcal{M}, \mu)$ be a finite measure space. Assume that f and g are two Borel non-negative functions on X such that $\ln f$ and $\ln g$ are $\mu-$integrable. Then $\ln(f+g)$ is $\mu-$integrable.*

Let $g_i = \sum_{j=1}^{r} g_{ij}$. The next lemma is an integrability result on $\ln g_i$.

Lemma 1.2. *For all Φ and $\Phi' \in F_\Phi$, set $\mathcal{C}(\Phi') = \{i \in \{1, \ldots, r\} \,|\, \pi_i(\Phi') > 0\}$. Then, for all $i \in \mathcal{C}(\Phi')$, one has*

$$\mathbb{E}_\Phi \left[\max_{i \in \mathcal{C}(\Phi')} |\ln g_i(Y_1, E_1, X_1, \Phi')| \right] < \infty.$$

Before stating our third lemma, we recall Kingman's result on subadditive processes (see Theorems 1.5 and 1.7 in [73]).

Theorem 1.2 (Kingman [73]) *Let (W_{st}) be a real-valued double-index process. Assume that (W_{st}) is subadditive, namely*

- $\forall\, s < u < t, \quad W_{st} \leq W_{su} + W_{ut}$;
- (W_{st}) *is stationary relative to the shift* $W_{st} \to W_{s+1,\,t+1}$;

- $\forall\, t, \quad \mathbb{E}|W_{0t}| < \infty$;
- $\exists\, A \in \mathbb{R}, \quad \forall\, t, \quad \mathbb{E}(W_{0t}) \geq -At$.

Then there exists an integrable random variable ξ such that

(i) $\dfrac{W_{0t}}{t} \to \xi$ *a.s. and in L^1 as $t \to \infty$;*

(ii) If the $\sigma-$algebra invariant relative to the shift in 2. is a.s. trivial, then ξ is a.s. constant.

Using this formalism, we get another result on our process:

Lemma 1.3. *Pick an arbitrary parameter Φ, and set for all $\Phi' \in F_\Phi$*

$$q_{st}(\Phi') = \max_{j \in \mathcal{C}(\Phi')} \mathcal{L}((Y_i,\, E_i,\, X_i)_{s+1 \leq i \leq t},\, \Phi' \,|\, J(0) = j).$$

Then there exists a neighborhood G of Φ' in F_Φ such that for all $G_{\Phi'} \subset G$ containing Φ', the process $\Big(Z_{st} = \ln \sup_{\varphi \in G_{\Phi'}} q_{st}(\varphi)\Big)$ is $\mathbb{P}_\Phi-$subadditive and ergodic.

We now write analogues of the lemmas of [97], in order to show our main theorem. Note that, using the finite dimensional distributions of the stationary process $(J_i,\, Y_i,\, E_i,\, X_i)$, it is possible to extend the sequences $(J_i)_{i \geq 0}$ and $(Y_i,\, E_i,\, X_i)_{i \geq 1}$ to doubly infinite sequences $(J_i)_{i \in \mathbb{Z}}$ and $(Y_i,\, E_i,\, X_i)_{i \in \mathbb{Z}}$. Because $(J_i,\, Y_i)$ is an MRP, it holds that

$$\mathcal{L}(Y_1,\, E_1,\, X_1,\, \Phi \,|\, (Y_i,\, E_i,\, X_i)_{-n \leq i \leq 0}) = \sum_{j=1}^{r} \mathbb{P}_\Phi(X_0 = j \,|\, (Y_i,\, E_i,\, X_i)_{-n \leq i \leq 0})\, g_j(Y_1,\, E_1,\, X_1,\, \Phi).$$

Remark that, letting $\mathcal{F}_n$ be the $\sigma-$field $\mathcal{F}_n = \sigma((Y_i,\, E_i,\, X_i)_{-n \leq i \leq 0})$, the sequence $(\mathcal{F}_n)_{n \geq 0}$ is nondecreasing. A martingale theorem of Lévy (see Theorem 35.5 p. 358 in Billingsley [13]) then yields, for all $j \in \{1,\, \ldots,\, r\}$,

$$\mathbb{P}_\Phi(X_0 = j \,|\, (Y_i,\, E_i,\, X_i)_{-n \leq i \leq 0}) \to \mathbb{P}_\Phi(X_0 = j \,|\, (Y_i,\, E_i,\, X_i)_{i \leq 0}) \quad \mathbb{P}_\Phi - \text{a.s.}$$

as $n \to \infty$, which in turn gives a definition of $\mathcal{L}(Y_1,\, E_1,\, X_1,\, \Phi \,|\, (Y_i,\, E_i,\, X_i)_{i \leq 0})$.

Lemma 1.4. *For all Φ, $\ln \mathcal{L}(Y_1,\, E_1,\, X_1,\, \Phi \,|\, (Y_i,\, E_i,\, X_i)_{i \leq 0})$ is $\mathbb{P}_\Phi-$integrable. Let*

$$H(\Phi) = \mathbb{E}_\Phi(-\ln \mathcal{L}(Y_1,\, E_1,\, X_1,\, \Phi \,|\, (Y_i,\, E_i,\, X_i)_{i \leq 0})).$$

Then, as $n \to \infty$,

(i) $\dfrac{1}{n}\mathbb{E}_\Phi \ln \mathcal{L}((Y_i,\, E_i,\, X_i)_{1 \leq i \leq n},\, \Phi) \to -H(\Phi)$;

(ii) $\dfrac{1}{n} \ln \mathcal{L}((Y_i,\, E_i,\, X_i)_{1 \leq i \leq n},\, \Phi) \to -H(\Phi) \;\; \mathbb{P}_\Phi-$*a.s.*

Lemma 1.5. *For all Φ and all $\Phi' \in F_\Phi$, there exists a finite real number $H(\Phi,\, \Phi')$ such that, as $n \to \infty$,* *(i)* $\frac{1}{n}\,\mathbb{E}_\Phi \ln \mathcal{L}((Y_i,\, E_i,\, X_i)_{1\le i\le n},\, \Phi') \to H(\Phi,\, \Phi')$;

(ii) $\frac{1}{n}\,\mathbb{E}_\Phi \left[\max_{j\in\mathcal{C}(\Phi')} \mathcal{L}((Y_i,\, E_i,\, X_i)_{1\le i\le n},\, \Phi' \,|\, J(0)=j)\right] \to H(\Phi,\, \Phi')$;

(iii) $\frac{1}{n} \ln \mathcal{L}((Y_i,\, E_i,\, X_i)_{1\le i\le n},\, \Phi') \to H(\Phi,\, \Phi')$ $\mathbb{P}_\Phi-a.s.$;

(iv) $\frac{1}{n} \max_{j\in\mathcal{C}(\Phi')} \mathcal{L}((Y_i,\, E_i,\, X_i)_{1\le i\le n},\, \Phi' \,|\, J(0)=j) \to H(\Phi,\, \Phi')$ $\mathbb{P}_\Phi-a.s.$

To adapt Lemma 6 of [97], define

$$\Omega = \{y_i,\, e_i,\, x_i,\, u_1(i-1),\, \dots,\, u_r(i-1)\}_{i\ge 1}$$

where $y_i \in [0,\, \infty)$, $e_i \in \{0,\, 1\}^n$, $x_i \in [0,\, \infty)^n$, $u_j(i-1) \in [0,\, 1]$. Ω is equipped with its Borel $\sigma-$algebra $\mathcal{B}$ (to define a topology on Ω, use a weighted sum of metrics so that Ω is complete and separable).

Define by induction on $k \in \mathbb{N}$

$$u_j(0,\, \Phi) = \pi_j(\Phi), \quad u_j(k,\, \Phi) = \frac{\sum_i g_{ij}(y_k,\, e_k,\, x_k,\, \Phi)\, u_i(k-1,\, \Phi)}{\sum_i g_i(y_k,\, e_k,\, x_k,\, \Phi)\, u_i(k-1,\, \Phi)},$$

so that Bayes' formula yields $u_j(k,\, \Phi) = \mathbb{P}_\Phi(J_k = j \,|\, (y_i,\, e_i,\, x_i)_{1\le i\le k})$.

Let now $\mathcal{P}_n$ be the set of all subsets of $\{1,\, \dots,\, n\}$ and define, for all Borel sets $B \subset [0,\, \infty)^k \times \mathcal{P}_n^k \times [0,\, \infty)^{nk} \times [0,1]^{rk}$,

$$\mathbb{P}_{\Phi,\, \Phi'}(B) = \mathbb{P}_\Phi((y_i,\, e_i,\, x_i,\, u(i-1,\, \Phi'))_{1\le i\le k} \in B).$$

We let $\mathcal{S} : \Omega \to \Omega$, $(y_i,\, e_i,\, x_i,\, u(i-1,\, \Phi)) \mapsto (y_{i+1},\, e_{i+1},\, x_{i+1},\, u(i,\, \Phi))$ be the standard shift transformation, and

$$\widetilde{\mathbb{P}}^{(N)}_{\Phi,\, \Phi'} = \frac{1}{N}\sum_{i=0}^{N-1} \mathbb{P}_{\Phi,\, \Phi'} \circ \mathcal{S}^{-i}.$$

Finally, we let Y_i, E_i, X_i and $U(i-1,\, \Phi')$ be the corresponding coordinate mappings.

Lemma 1.6. *For all Φ and $\Phi' \in F_\Phi$, there exist an increasing sequence of integers (N_j) and a probability measure $\widetilde{\mathbb{P}}_{\Phi,\, \Phi'}$ on Ω such that*

(i) For all k, the law of $(Y_i,\, E_i,\, X_i,\, U(i-1,\, \Phi'))_{1\le i\le k}$ under the probability $\widetilde{\mathbb{P}}^{(N_j)}_{\Phi,\, \Phi'}$ converges weakly to the law of $(Y_i,\, E_i,\, X_i,\, U(i-1,\, \Phi'))_{1\le i\le k}$ under the probability $\widetilde{\mathbb{P}}_{\Phi,\, \Phi'}$;

(ii) The process $(Y_i,\, E_i,\, X_i,\, U(i-1,\, \Phi'))$ is $\widetilde{\mathbb{P}}_{\Phi,\, \Phi'}-$stationary;

(iii) The process $(Y_i,\, E_i,\, X_i)$ has the same law under $\widetilde{\mathbb{P}}_{\Phi,\, \Phi'}$ and under $\mathbb{P}_\Phi$.

The analogue of Lemma 7 in Rydén gives a representation of $H(\Phi,\, \Phi')$:

Lemma 1.7. *Let Φ be an arbitrary parameter and $\Phi' \in F_\Phi$. Then*

$$H(\Phi,\,\Phi') = \widetilde{\mathbb{E}}_{\Phi,\,\Phi'}\left[\ln\left(\sum_{i=1}^{r} U_i(0,\,\Phi')\, g_i(Y_1,\,E_1,\,X_1,\,\Phi')\right)\right].$$

For all Φ and $\Phi' \in F_\Phi$, set now $K(\Phi,\,\Phi') = H(\Phi,\,\Phi) - H(\Phi,\,\Phi')$: this real number is well defined and finite. We state a result similar to Lemma 8 of [97]:

Lemma 1.8. *Pick Φ and $\Phi' \in F_\Phi$. Then $K(\Phi,\,\Phi') \geq 0$, and $(K(\Phi,\,\Phi') = 0) \Leftrightarrow (\Phi \sim \Phi')$.*

We finally write a result that solves the E step of the EM algorithm in Section 1.3:

Lemma 1.9. *Retain the notations of Section 1.3, and let*

$$A(e) = \bigcup_{i=1}^{r} A(i,\,e) = \{j \in \{1,\,\ldots,\,k\} \,|\, e_j = e\}.$$

Then

$$\begin{aligned}
\widehat{T_i} &= \int_0^T \frac{\mathbb{P}_\varphi(J(v) = i,\, N(u),\, S(u),\, 0 \leq u < v)}{\mathbb{P}_\varphi(N(u),\, S(u),\, 0 \leq u \leq T)} \\
&\qquad \times \mathbb{P}_\varphi(N(u),\, S(u),\, v \leq u \leq T \,|\, J(v) = i)\, dv, \\
\widehat{n_i} &= \sum_{q=1}^{k} \frac{\mathbb{P}_\varphi(J(t_q) = i,\, N(u),\, S(u),\, 0 \leq u \leq T)}{\mathbb{P}_\varphi(N(u),\, S(u),\, 0 \leq u \leq T)}, \\
\widehat{1\!\!1_{\{j \in A(i,\,e)\}}} &= 1\!\!1_{\{j \in A(e)\}}\, \mathbb{P}_\varphi(J(t_j) = i \,|\, N(u),\, S(u),\, 0 \leq u \leq T), \\
\mathrm{card}\widehat{(A(i,\,e))} &= \sum_{j=1}^{k} 1\!\!1_{\{j \in A(e)\}}\, \mathbb{P}_\varphi(J(t_j) = i \,|\, N(u),\, S(u),\, 0 \leq u \leq T), \\
\widehat{m_{ij}(T)} &= \ell_{ij}(\varphi) \int_0^T \frac{\mathbb{P}_\varphi(J(v) = i,\, N(u),\, S(u),\, 0 \leq u < v)}{\mathbb{P}_\varphi(N(u),\, S(u),\, 0 \leq u \leq T)} \\
&\qquad \times \mathbb{P}_\varphi(N(u),\, S(u),\, v \leq u \leq T \,|\, J(v) = j)\, dv.
\end{aligned}$$

1.7 Appendix B: Proofs

Proof of Lemma 1.1. Start by writing

$$\int_X |\ln(f+g)|\, d\mu = \int_{\{f+g \geq 1\}} \ln(f+g)\, d\mu - \int_{\{f+g<1\}} \ln(f+g)\, d\mu.$$

Now $f \leq f+g$, so that $-\ln(f+g)\, 1\!\!1_{\{f+g<1\}} \leq -\ln f\, 1\!\!1_{\{f+g<1\}}$. Therefore

$$\int_X |\ln(f+g)|\, d\mu \leq \int_{\{f+g \geq 1\}} \ln(f+g)\, d\mu + \int_X |\ln f|\, d\mu.$$

Thus, it is enough to prove that $\displaystyle\int_{\{f+g \geq 1\}} \ln(f+g)\, d\mu < \infty$.

Notice that

$$\forall\, x,\, y \geq 2, \quad \ln\left(\frac{1}{x} + \frac{1}{y}\right) \leq 0$$

which leads to

$$\forall\, x,\, y \geq 2, \quad \ln(x+y) \leq \ln x + \ln y.$$

Consequently

$$\begin{aligned}
\int_{\{f+g\geq 1\}} \ln(f+g)\, d\mu &= \int_{\{f+g\geq 1,\, f\geq 2,\, g\geq 2\}} \ln(f+g)\, d\mu + \int_{\{f+g\geq 1,\, f<2,\, g<2\}} \ln(f+g)\, d\mu \\
&+ \int_{\{f+g\geq 1,\, f\geq 2,\, g<2\}} \ln(f+g)\, d\mu + \int_{\{f+g\geq 1,\, f<2,\, g\geq 2\}} \ln(f+g)\, d\mu \\
&\leq \int_{\{f+g\geq 1,\, f\geq 2,\, g\geq 2\}} [\ln f + \ln g]\, d\mu + \mu(X)\, 2\ln 2 \\
&+ \int_{\{f+g\geq 1,\, f\geq 2,\, g<2\}} \ln(f+2)\, d\mu + \int_{\{f+g\geq 1,\, f<2,\, g\geq 2\}} \ln(g+2)\, d\mu \\
&\leq 2\int_X |\ln f|\, d\mu + 2\int_X |\ln g|\, d\mu + \mu(X)\, 4\ln 2
\end{aligned}$$

which gives $\int_{\{f+g\geq 1\}} \ln(f+g)\, d\mu < \infty$ and ends the proof of this lemma. ■

Proof of Lemma 1.2. First, note that

$$\mathbb{E}_\Phi \left[\max_{i\in\mathcal{C}(\Phi')} |\ln g_i(Y_1,\, E_1,\, X_1,\, \Phi')| \right] \leq \sum_{i\in\mathcal{C}(\Phi')} \mathbb{E}_\Phi |\ln g_i(Y_1,\, E_1,\, X_1,\, \Phi')|;$$

we shall then prove that for all $\Phi' \in F_\Phi$ and $i \in \mathcal{C}(\Phi')$, $\mathbb{E}_\Phi |\ln g_i(Y_1,\, E_1,\, X_1,\, \Phi')| < \infty$. Write

$$\mathbb{E}_\Phi |\ln g_i(Y_1,\, E_1,\, X_1,\, \Phi')| = \sum_e \int |\ln g_i(y,\, e,\, x,\, \Phi')|\, \mathcal{L}(dy,\, e,\, dx,\, \Phi).$$

Using the equality $g_i = \sum_j g_{ij}$ and Lemma 1.1, it is enough to show that for all i and all j, e such that $\lambda_j(\Phi')\, p(j,\, e,\, \Phi') > 0$, the function

$$(y,\, x) \mapsto \ln \left[\exp(y(L(\Phi') - \Lambda(\Phi')))_{i,\, j}\, \lambda_j(\Phi')\, p(j,\, e,\, \Phi')\, \mathbb{P}_{\theta(j,\, e,\, \Phi')}(m \in e \Rightarrow X_m = x_m) \right]$$

is $\mathcal{L}(\cdot,\, e,\, \cdot,\, \Phi)-$integrable. The hypothesis on the statistical model $(\mathbb{P}_\theta)$ first gives, for all i, j,

$$\int \left| \ln \mathbb{P}_{\theta(j,\, e,\, \Phi')}(m \in e \Rightarrow X_m = x_m) \right| \mathbb{P}_{\theta(i,\, e,\, \Phi)}(m \in e \Rightarrow X_m \in dx_m) < \infty.$$

Because J is a $\mathbb{P}_{\Phi'}-$irreducible Markov process, for all $j \neq i$, one has

$$\forall\, y > 0, \quad \exp(y(L(\Phi') - \Lambda(\Phi')))_{i,\, j} = \frac{1}{\lambda_j(\Phi')} \mathbb{P}_{\Phi'}(J(y) = j,\, Y_1 \in dy \,|\, J(0) = i) > 0$$

and therefore, since for a given matrix A

$$\forall\, y \in \mathbb{R}, \quad \exp(yA) = \sum_{n=0}^{\infty} \frac{A^n}{n!} y^n,$$

there exists an integer $n(i,\, j)$ such that $(L(\Phi') - \Lambda(\Phi'))_{i,\, j}^{n(i,\, j)} > 0$, which we pick to be minimal among the integers satisfying this property. Consequently, in a neighborhood of 0,

$$\exp(y(L(\Phi') - \Lambda(\Phi')))_{i,\, j} = (L(\Phi') - \Lambda(\Phi'))_{i,\, j}^{n(i,\, j)} \frac{y^{n(i,\, j)}}{n(i,\, j)!} + \mathrm{O}(y^{n(i,\, j)+1})$$

so that the functions

$$y \mapsto \ln\left[\exp(y(L(\Phi') - \Lambda(\Phi')))_{i,\,j}\,\lambda_j(\Phi')\,p(j,\,e,\,\Phi')\right]\mathcal{L}(y,\,e,\,\Phi)$$

are integrable in a neighborhood of 0.

Using a corollary of Perron-Frobenius' theorem (see e.g. Appendix A.4.8 in Asmussen [6]) entails that there exist a matrix A with positive entries and two real numbers $\mu > \delta > 0$ such that, as $y \to \infty$,

$$\exp(y(L(\Phi') - \Lambda(\Phi'))) = \exp(-\delta y)\,A + \mathrm{O}(\exp(-\mu y));$$

thus, for all $i,\, j \in \{1,\,\dots,\,r\}$, there exists $a_{i,\,j} > 0$ such that, as $y \to \infty$,

$$\exp(y(L(\Phi') - \Lambda(\Phi')))_{i,\,j} = \exp(-\delta y)\,a_{i,\,j} + \mathrm{O}(\exp(-\mu y)).$$

Therefore, the functions

$$y \mapsto \ln\left[\exp(y(L(\Phi') - \Lambda(\Phi')))_{i,\,j}\,\lambda_j(\Phi')\,p(j,\,e,\,\Phi')\right]\mathcal{L}(y,\,e,\,\Phi)$$

are integrable in a neighborhood of ∞. Finally, Lemma 1.2 follows. ■

Proof of Lemma 1.3. A proof similar to the one of [97] shows that the first and second hypothesis are satisfied for any G, and that the process is ergodic. To show the third one, let G be an arbitrary subset of F_Φ containing Φ'. Set $B = \left\{\sup_{\varphi\in G} q_{0t}(\varphi) \le 1\right\}$. The inequality

$$q_{0t}(\varphi) \le \prod_{i=1}^{t} q_{i-1,\,i}(\varphi)$$

and the stationarity of the process $(Y_i,\,E_i,\,X_i)$ together imply

$$\begin{aligned}\mathbb{E}_\Phi\left|\ln\sup_{\varphi\in G} q_{0t}(\varphi)\right| &= -\mathbb{E}_\Phi\left[\ln\sup_{\varphi\in G} q_{0t}(\varphi)\mathbb{1}_B\right] + \mathbb{E}_\Phi\left[\ln\sup_{\varphi\in G} q_{0t}(\varphi)\mathbb{1}_{B^c}\right]\\ &\le \mathbb{E}_\Phi|\ln q_{0t}(\Phi')| + t\,\mathbb{E}_\Phi\left|\ln\sup_{\varphi\in G} q_{01}(\varphi)\right|.\end{aligned}$$

Now

$$q_{0t}(\Phi') = \max_{k_0\in\mathcal{C}(\Phi')}\sum_{k_1,\,\dots,\,k_t\in\mathcal{C}(\Phi')}\prod_{s=1}^{t} g_{k_{s-1},\,k_s}(Y_s,\,E_s,\,X_s,\,\Phi'),$$

so that, because $\left|\ln\max_{1\le i\le n} a_i\right| \le \max_{1\le i\le n}|\ln a_i|$ for all real numbers $(a_i)_{1\le i\le n}$,

$$\begin{aligned}|\ln q_{0t}(\Phi')| &\le \max_{k_0\in\mathcal{C}(\Phi')}\left|\ln\left[\sum_{k_1,\,\dots,\,k_t\in\mathcal{C}(\Phi')}\prod_{s=1}^{t} g_{k_{s-1},\,k_s}(Y_s,\,E_s,\,X_s,\,\Phi')\right]\right|\\ &\le \sum_{k_0\in\mathcal{C}(\Phi')}\left|\ln\left[\sum_{k_1,\,\dots,\,k_t\in\mathcal{C}(\Phi')}\prod_{s=1}^{t} g_{k_{s-1},\,k_s}(Y_s,\,E_s,\,X_s,\,\Phi')\right]\right|.\end{aligned}$$

To show that $\mathbb{E}_\Phi |\ln q_{0t}(\Phi')| < \infty$, we notice that applying Lemma 1.1 and using the properties of the logarithm function, it is enough to show that for all s, e_s and $k_{s-1},\, k_s \in \mathcal{C}(\Phi')$ such that $\lambda_{k_s}(\Phi')\, p(k_s,\, e_s,\, \Phi') > 0$,

$$\int |\ln g_{k_{s-1},\, k_s}(y_s,\, e_s,\, x_s,\, \Phi')|\, \mathcal{L}((dy_j,\, e_j,\, dx_j)_{1\le j\le t},\, \Phi) < \infty,$$

which was already shown in the proof of Lemma 1.2. It is therefore enough to show that

$$\mathbb{E}_\Phi \left| \ln \sup_{\varphi\in G} q_{01}(\varphi) \right| < \infty.$$

Put then $C = \left\{\sup_{\varphi\in G} q_{01}(\varphi) \le 1\right\}$ and write

$$\begin{aligned} \mathbb{E}_\Phi \left| \ln \sup_{\varphi\in G} q_{01}(\varphi) \right| &= -\mathbb{E}_\Phi \left[\ln \sup_{\varphi\in G} q_{01}(\varphi) \mathbb{1}_C \right] + \mathbb{E}_\Phi \left[\ln \sup_{\varphi\in G} q_{01}(\varphi) \mathbb{1}_{C^c} \right] \\ &\le \mathbb{E}_\Phi |\ln q_{01}(\Phi')| + \mathbb{E}_\Phi \left[\ln \sup_{\varphi\in G} q_{01}(\varphi) \mathbb{1}_{C^c} \right]. \end{aligned}$$

Because $\mathbb{E}_\Phi |\ln q_{01}(\Phi')| < \infty$, it is sufficient to prove that there exists a neighborhood G of Φ' in F_Φ such that for every subset $G_{\Phi'}$ of G containing Φ',

$$\mathbb{E}_\Phi \left[\ln \sup_{\varphi\in G_{\Phi'}} q_{01}(\varphi) \mathbb{1}_{C^c} \right] < \infty.$$

Since

$$q_{01}(\varphi) = \max_{i\in\mathcal{C}(\varphi)} g_i(Y_1,\, E_1,\, X_1,\, \varphi) \le \sum_{i=1}^{r} g_i(Y_1,\, E_1,\, X_1,\, \varphi)$$

one has

$$\sup_{\varphi\in G} q_{01}(\varphi) \le \sum_{i,\, j=1}^{r} \sup_{\varphi\in G} g_{ij}(Y_1,\, E_1,\, X_1,\, \varphi).$$

Lemma 1.1 therefore shows that it is enough to find a neighborhood G of Φ' in F_Φ such that for all j, e such that $\lambda_j(\Phi')\, p(j,\, e,\, \Phi') > 0$ and every subset $G_{\Phi'}$ of G containing Φ', we have

$$\forall\, i \in \{1,\, \ldots,\, r\}, \quad \int \left| \ln \sup_{\varphi\in G_{\Phi'}} g_{ij}(y,\, e,\, x,\, \varphi) \right| \mathcal{L}(dy,\, e,\, dx,\, \Phi) < \infty.$$

The hypothesis on the statistical model $(\mathbb{P}_\theta)$ and the fact that J has a finite state space together imply that there exists a neighborhood G_2 of Φ' such that for all i, j and every subset $G_{\Phi'}$ of G_2 containing Φ',

$$\int \left| \ln \sup_{\varphi\in G_{\Phi'}} \mathbb{P}_{\theta(i,\, e,\, \varphi)}(m \in e \Rightarrow X_m = x_m) \right| \mathbb{P}_{\theta(j,\, e,\, \Phi)}(m \in e \Rightarrow X_m = x_m)\, dx < \infty;$$

it is finally enough to find, for every e, a neighborhood G_1 of Φ' such that for every subset $G_{\Phi'}$ of G_1 containing Φ',

$$\int \left| \ln \sup_{\varphi\in G_{\Phi'}} \exp(y(L(\varphi) - \Lambda(\varphi)))_{i,\, j}\, \lambda_j(\varphi)\, p(j,\, e,\, \varphi) \right| \mathcal{L}(dy,\, e,\, \Phi) < \infty.$$

Pick then a neighborhood G_1 of Φ' in F_Φ such that

- $\forall\, j, \quad \sup_{\varphi \in G} |\lambda_j(\varphi) - \lambda_j(\Phi')| < \lambda_j(\Phi')1\!\!1_{\{\lambda_j(\Phi')>0\}}$;
- $\forall\, e,\, j, \quad \sup_{\varphi \in G} |p(j,\, e,\, \varphi) - p(j,\, e,\, \Phi')| < p(j,\, e,\, \Phi')1\!\!1_{\{p(j,\, e,\, \Phi')>0\}}$,

so that if $\varphi \in G = G_1 \cap G_2$, the parameters $\lambda_j(\varphi)$ and $p(j,\, e,\, \varphi)$ are bounded for all j and e, and if one of the parameters is positive under $\mathbb{P}_{\Phi'}$, then it is also under $\mathbb{P}_\varphi$. Set

$$M = \sup_{\varphi \in G} \max_{j,\, e} \lambda_j(\varphi)\, p(j,\, e,\, \varphi).$$

For all $\varphi \in G$, the inequalities

$$\varepsilon\, \exp(y(L(\varphi) - \Lambda(\varphi)))_{i,\, j} \leq \exp(y(L(\varphi) - \Lambda(\varphi)))_{i,\, j}\, \lambda_j(\varphi)\, p(j,\, e,\, \varphi) \leq M$$

hold, if $\varepsilon > 0$ is defined by

$$\varepsilon = \inf_{\varphi \in G} \{\lambda_k(\varphi)\, p(k,\, e,\, \varphi),\, k,\, e \text{ such that } \lambda_k(\Phi')\, p(k,\, e,\, \Phi') > 0\}.$$

Therefore, for all $\varphi \in G$,

$$\varepsilon\, \exp(y(L(\varphi) - \Lambda(\varphi)))_{i,\, j} \leq \sup_{\psi \in G} \exp(y(L(\psi) - \Lambda(\psi)))_{i,\, j}\, \lambda_j(\psi)\, p(j,\, e,\, \psi) \leq M.$$

Especially,

$$\ln(\varepsilon\, \exp(y(L(\Phi') - \Lambda(\Phi')))_{i,\, j}) \leq \ln \left[\sup_{\psi \in G} \exp(y(L(\psi) - \Lambda(\psi)))_{i,\, j}\, \lambda_j(\psi)\, p(j,\, e,\, \psi) \right] \leq \ln M$$

and because the function

$$y \mapsto \ln(\varepsilon\, \exp(y(L(\Phi') - \Lambda(\Phi')))_{i,\, j})\, \mathcal{L}(y,\, e,\, \Phi)$$

is integrable, the result is proven.

To show that the fourth requirement is met, notice that $Z_{0t} \geq \ln q_{0t}(\Phi')$. Furthermore,

$$q_{0t}(\Phi') = \max_{k_0 \in \mathcal{C}(\Phi')} \sum_{k_1,\, \dots,\, k_t \in \mathcal{C}(\Phi')} \prod_{s=1}^{t} g_{k_{s-1},\, k_s}(Y_s,\, E_s,\, X_s,\, \Phi'),$$

Using the definition of F_Φ, we see that we may focus on the indices e such that there exists an index $j \in \{1,\, \dots,\, r\}$ with $\lambda_j(\Phi')\, p(j,\, e,\, \Phi') > 0$. Let I be the set of these indices.

Put then $\gamma = \min\{\lambda_j(\Phi')\, p(j,\, e,\, \Phi'),\, j,\, e \text{ such that } \lambda_j(\Phi')\, p(j,\, e,\, \Phi') > 0\}$. We set indices $e_j \in I$ for all $j \in \{1,\, \dots,\, t\}$. There exist $k_1(e_1),\, \dots,\, k_t(e_t)$ such that

$$\forall\, j \in \{1,\, \dots,\, t\}, \quad \lambda_{k_j(e_j)}(\Phi')\, p(k_j(e_j),\, e_j,\, \Phi') \geq \gamma > 0.$$

Since

$$\begin{aligned} q_{0t}(\Phi') &= \max_{k_0 \in \mathcal{C}(\Phi')} \sum_{k_1,\, \dots,\, k_t \in \mathcal{C}(\Phi')} \prod_{s=1}^{t} g_{k_{s-1},\, k_s}(Y_s,\, E_s,\, X_s,\, \Phi') \\ &\geq \max_{k_0 \in \mathcal{C}(\Phi')} \prod_{s=1}^{t} g_{k_{s-1}(E_{s-1}),\, k_s(E_s)}(Y_s,\, E_s,\, X_s,\, \Phi') \end{aligned}$$

and

$$g_{k_{s-1}(e_{s-1}),\, k_s(e_s)}(y_s,\, e_s,\, x_s,\, \Phi')$$

$$\geq \gamma \exp(y_s(L(\Phi') - \Lambda(\Phi')))_{k_{s-1}(e_{s-1}),\, k_s(e_s)}\, \mathbb{P}_{\theta(k_s(e_s),\, e_s,\, \Phi')}(m \in e_s \Rightarrow X_m = x_{m,\, s})$$

we get, for all $k_0 = k_0(e_0) \in \mathcal{C}(\Phi')$,

$$\begin{aligned} \mathbb{E}_\Phi \ln q_{0t}(\Phi') \quad &\geq \sum_{\substack{e_s \in I \\ 1 \leq s \leq t}} \int \ln \exp(y_s(L(\Phi') - \Lambda(\Phi')))_{k_{s-1}(e_{s-1}),\, k_s(e_s)}\, \mathcal{L}((dy_j,\, e_j,\, dx_j)_{1 \leq j \leq t},\, \Phi) \\ &+ \sum_{\substack{e_s \in I \\ 1 \leq s \leq t}} \int \ln \mathbb{P}_{\theta(k_s(e_s),\, e_s,\, \Phi')}(m \in e_s \Rightarrow X_m = x_{m,\, s})\, \mathcal{L}((dy_j,\, e_j,\, dx_j)_{1 \leq j \leq t},\, \Phi) \\ &+ tC \ln \gamma \end{aligned}$$

where $C = \operatorname{card} I$. We can eliminate the condition $e_j \in I$ because the terms with $e_j \notin I$ do not contribute, and thus

$$\begin{aligned} \mathbb{E}_\Phi \ln q_{0t}(\Phi') \quad &\geq \sum_{\substack{e_s \\ 1 \leq s \leq t}} \int \ln \exp(y_s(L(\Phi') - \Lambda(\Phi')))_{k_{s-1}(e_{s-1}),\, k_s(e_s)}\, \mathcal{L}(dy_s,\, e_{s-1},\, e_s,\, \Phi) \\ &+ \sum_{\substack{e_s \\ 1 \leq s \leq t}} \int \ln \mathbb{P}_{\theta(k_s(e_s),\, e_s,\, \Phi')}(m \in e_s \Rightarrow X_m = x_{m,\, s})\, \mathcal{L}(e_s,\, dx_s,\, \Phi) \\ &+ tC \ln \gamma. \end{aligned}$$

The stationarity of the process $(Y_i,\, E_i,\, X_i)$ yields

$$\left| \sum_{\substack{e_s \\ 1 \leq s \leq t}} \int \ln \mathbb{P}_{\theta(k_s(e_s),\, e_s,\, \Phi')}(m \in e_s \Rightarrow X_m = x_{m,\, s})\, \mathcal{L}(e_s,\, dx_s,\, \Phi) \right|$$

$$\leq t \sum_{e_1} \int \left| \ln \mathbb{P}_{\theta(k_1(e_1),\, e_1,\, \Phi')}(m \in e_1 \Rightarrow X_m = x_{m,1}) \right| \mathcal{L}(e_1,\, dx_1,\, \Phi) < \infty$$

and

$$\left| \sum_{\substack{e_s \\ 2 \leq s \leq t}} \int \ln \exp(y_s(L(\Phi') - \Lambda(\Phi')))_{k_{s-1}(e_{s-1}),\, k_s(e_s)}\, \mathcal{L}(dy_s,\, e_{s-1},\, e_s,\, \Phi) \right|$$

$$\leq (t-1) \sum_{e_1,\, e_2} \int \left| \ln \exp(y_2(L(\Phi') - \Lambda(\Phi')))_{k_1(e_1),\, k_2(e_2)} \right| \mathcal{L}(dy_2,\, e_1,\, e_2,\, \Phi) < \infty.$$

The term with $s = 1$ is also finite, which completes the proof of our result. ■

Proof of Lemma 1.4. To show Lemma 1.4, notice that applying Lemma 1.2 yields

$$\mathbb{E}_\Phi \left[\max_{i \in \mathcal{C}(\Phi)} |\ln g_i(Y_1,\, E_1,\, X_1,\, \Phi)| \right] < \infty$$

and adapt the proof of Lemma 4 in [97]. ■

Proof of Lemma 1.5. Retaining the notation of Lemma 1.3, the process $(\ln q_{st}(\Phi'))$ is subadditive and ergodic; the proof is then similar to the proof of Lemma 5 in [97] and is therefore omitted. ■

Proof of Lemma 1.6. The argument follows the proof of Lemma 6 in [97] and is omitted. ■

Proof of Lemma 1.7. The proof is the same as the proof of Lemma 7 in [97]. ■

Proof of Lemma 1.8. The proof is an easy adaptation of the proof of Lemma 8 in [97]. ■

Proof of Lemma 1.9. Since $\{N(u), S(u), v \leq u \leq T\}$ and $\{N(u), S(u), 0 \leq u < v\}$ are independent given $\{J(v) = i\}$,

$$\begin{aligned}
\widehat{T_i} &= \int_0^T \mathbb{P}_\varphi(J(v) = i \,|\, N(u), S(u), 0 \leq u \leq T)\, dv \\
&= \int_0^T \frac{\mathbb{P}_\varphi(J(v) = i, N(u), S(u), 0 \leq u \leq T)}{\mathbb{P}_\varphi(N(u), S(u), 0 \leq u \leq T)}\, dv \\
&= \int_0^T \frac{\mathbb{P}_\varphi(J(v) = i, N(u), S(u), 0 \leq u < v)}{\mathbb{P}_\varphi(N(u), S(u), 0 \leq u \leq T)} \\
&\qquad \times \mathbb{P}_\varphi(N(u), S(u), v \leq u \leq T \,|\, J(v) = i)\, dv.
\end{aligned}$$

To get $\widehat{n_i}$, write

$$\begin{aligned}
\widehat{n_i} &= \sum_{q=1}^{k} \mathbb{P}_\varphi(J(t_q) = i \,|\, N(u), S(u), 0 \leq u \leq T) \\
&= \sum_{q=1}^{k} \frac{\mathbb{P}_\varphi(J(t_q) = i, N(u), S(u), 0 \leq u \leq T)}{\mathbb{P}_\varphi(N(u), S(u), 0 \leq u \leq T)}.
\end{aligned}$$

Then

$$\widehat{\mathbb{1}_{\{j \in A(i, e)\}}} = \mathbb{1}_{\{j \in A(e)\}}\, \mathbb{P}_\varphi(J(t_j) = i \,|\, N(u), S(u), 0 \leq u \leq T).$$

Consequently, we have

$$\widehat{\mathrm{card}(A(i, e))} = \sum_{j=1}^{k} \widehat{\mathbb{1}_{\{j \in A(i, e)\}}} = \sum_{j=1}^{k} \mathbb{1}_{\{j \in A(e)\}}\, \mathbb{P}_\varphi(J(t_j) = i \,|\, N(u), S(u), 0 \leq u \leq T).$$

We finish by showing how to compute $\widehat{m_{ij}(T)}$. First, if (U_j) stands for the sequence of the jump times of J,

$$\begin{aligned}
m_{ij}(T) &= \mathrm{card}\{s \,|\, 0 < s \leq T, J(s_-) = i, J(s) = j\} \\
&= \mathrm{card}\{Q \in \mathbb{N} \setminus \{0\} \,|\, J(U_{Q-1}) = i, J(U_Q) = j, U_Q \leq T\} \\
&= \sum_{Q=1}^{\infty} \mathbb{1}_{\{J(U_{Q-1}) = i, J(U_Q) = j, U_Q \leq T\}}
\end{aligned}$$

and therefore

$$\begin{aligned}
\widehat{m_{ij}(T)} &= \sum_{Q=1}^{\infty} \mathbb{P}_\varphi(J(U_{Q-1}) = i,\, J(U_Q) = j,\, U_Q \le T \,|\, N(u),\, S(u),\, 0 \le u \le T) \\
&= \sum_{Q=1}^{\infty} \frac{\mathbb{P}_\varphi(J(U_{Q-1}) = i,\, J(U_Q) = j,\, U_Q \le T,\, N(u),\, S(u),\, 0 \le u \le T)}{\mathbb{P}_\varphi(N(u),\, S(u),\, 0 \le u \le T)} \\
&= \int_0^T \sum_{Q=1}^{\infty} \frac{\mathbb{P}_\varphi(J(U_{Q-1}) = i,\, J(v) = j,\, U_Q \in\, dv,\, N(u),\, S(u),\, 0 \le u \le T)}{\mathbb{P}_\varphi(N(u),\, S(u),\, 0 \le u \le T)}\, dv.
\end{aligned}$$

Now, using independence given $\{J(v) = j\}$, one has for all $Q \in \mathbb{N} \setminus \{0\}$,

$$\begin{aligned}
&\mathbb{P}_\varphi(J(U_{Q-1}) = i,\, J(v) = j,\, U_Q \in\, dv,\, N(u),\, S(u),\, 0 \le u \le T) \\
&\qquad = \mathbb{P}_\varphi(N(u),\, S(u),\, v \le u \le T \,|\, J(v) = j) \\
&\qquad \times \mathbb{P}_\varphi(J(U_{Q-1}) = i,\, J(v) = j,\, U_Q \in\, dv,\, N(u),\, S(u),\, 0 \le u < v).
\end{aligned}$$

Notice that for all $Q \in \mathbb{N} \setminus \{0\}$,

$$\begin{aligned}
\mathbb{P}_\varphi(J(U_Q) = j,\, U_Q - U_{Q-1} \in dv \,|\, J(U_{Q-1}) = i) &= \frac{\ell_{ij}(\varphi)}{-\ell_{ii}(\varphi)}(-\ell_{ii}(\varphi) \exp(\ell_{ii}(\varphi)\, v))\, dv \\
\mathbb{P}_\varphi(U_Q - U_{Q-1} > v \,|\, J(U_{Q-1}) = i) &= \exp(\ell_{ii}(\varphi)\, v)
\end{aligned}$$

so that

$$\mathbb{P}_\varphi(J(U_Q) = j,\, U_Q - U_{Q-1} \in dv \,|\, J(U_{Q-1}) = i) = \ell_{ij}(\varphi)\, \mathbb{P}_\varphi(U_Q - U_{Q-1} > v \,|\, J(U_{Q-1}) = i)\, dv.$$

Write then

$$\begin{aligned}
&\mathbb{P}_\varphi(J(U_{Q-1}) = i,\, J(U_Q) = j,\, U_Q \in\, dv,\, N(u),\, S(u),\, 0 \le u < v) \\
&= \int \sum_{j_0,\, \dots,\, j_{Q-2}} \mathbb{P}_\varphi\Bigg(\bigcap_{q=0}^{Q-2} \{J(s_q) = j_q\},\, \bigcap_{q=0}^{Q-1} \{U_q \in ds_q\},\, J(U_{Q-1}) = i,\, J(U_Q) = j, \\
&\qquad\qquad U_Q \in\, dv,\, N(u),\, S(u),\, 0 \le u < v \Bigg).
\end{aligned}$$

Given the event

$$\left\{ \bigcap_{q=0}^{Q-2} \{J(s_q) = j_q\},\, \bigcap_{q=0}^{Q-1} \{U_q \in ds_q\},\, J(U_{Q-1}) = i,\, J(U_Q) = j,\, U_Q \in\, dv \right\},$$

the distribution of $\{N(u),\, S(u),\, 0 \le u < v\}$ is equal to the distribution of $\{N(u),\, S(u),\, 0 \le u < v\}$ given

$$\left\{ \bigcap_{q=0}^{Q-2} \{J(s_q) = j_q\},\, \bigcap_{q=0}^{Q-1} \{U_q \in ds_q\},\, J(U_{Q-1}) = i,\, U_Q > v \right\},$$

because the jump intensities of N and the shock probabilities in the interval $[0,\, v)$ are determined by the values of J in $[0,\, v)$. Consequently, using Bayes' formula,

$$\mathbb{P}_\varphi(J(U_{Q-1}) = i,\, J(U_Q) = j,\, U_Q \in\, dv,\, N(u),\, S(u),\, 0 \le u < v)$$

$$= \int \sum_{j_0,\,\ldots,\,j_{Q-2}} \ell_{ij}(\varphi)\,\mathbb{P}_\varphi\left(\bigcap_{q=0}^{Q-2}\{J(s_q)=j_q\},\,\bigcap_{q=0}^{Q-1}\{U_q\in ds_q\},\,J(U_{Q-1})=i,\,U_Q>v,\right.$$
$$\left. N(u),\,S(u),\,0\le u<v\right)dv$$

which yields

$$\mathbb{P}_\varphi(J(U_{Q-1})=i,\,J(U_Q)=j,\,U_Q\in\,dv,\,N(u),\,S(u),\,0\le u<v)$$
$$=\ell_{ij}(\varphi)\,\mathbb{P}_\varphi(J(U_{Q-1})=i,\,U_{Q-1}\le v,\,U_Q>v,\,N(u),\,S(u),\,0\le u<v)\,dv,$$

thus implying that

$$\sum_{Q=1}^{\infty}\mathbb{P}_\varphi(J(U_{Q-1})=i,\,J(U_Q)=j,\,U_Q\in\,dv,\,N(u),\,S(u),\,0\le u<v)$$
$$=\ell_{ij}(\varphi)\,\mathbb{P}_\varphi(J(v)=i,\,N(u),\,S(u),\,0\le u<v)\,dv.$$

From that, we deduce

$$\begin{aligned}\widehat{m_{ij}(T)} \quad &= \quad \ell_{ij}(\varphi)\int_0^T \frac{\mathbb{P}_\varphi(J(v)=i,\,N(u),\,S(u),\,0\le u<v)}{\mathbb{P}_\varphi(N(u),\,S(u),\,0\le u\le T)}\\ &\qquad \times\mathbb{P}_\varphi(N(u),\,S(u),\,v\le u\le T\mid J(v)=j)\,dv\end{aligned}$$

which completes the proof. ■

Part II

Endpoints, frontiers and high order moments

2 Endpoint estimation with high order moments

2.1 The problem

In this chapter, we let $(X_1, \ldots, X_n)$ be n independent copies of a random variable X, with support bounded to the right. Namely, we assume that the right endpoint $\theta = \sup\{x \in \mathbb{R} \,|\, F(x) < 1\}$ of the cumulative distribution function F of X is finite. We address the problem of estimating the real number θ with a high order moments method. In Section 2.2, we examine the case where X is positive; that hypothesis shall be dropped in Section 2.3. In Section 2.5, we present some particular examples where our assumptions hold. A simulation study is proposed in Section 2.6 to illustrate the efficiency of our estimators, and to compare them with estimators of the endpoint estimation literature. Auxiliary results are postponed to Appendix A and proven in Appendix B.

2.2 The positive case

In this section, we assume that X has a bounded support $[0, \theta]$, and we introduce an estimator of θ using high order moments of the variable of interest X. More precisely, the estimator is given by

$$\frac{1}{\widehat{\theta}_n} = \frac{1}{ap_n}\left[((a+1)p_n+1)\frac{\widehat{\mu}_{(a+1)p_n}}{\widehat{\mu}_{(a+1)p_n+1}} - (p_n+1)\frac{\widehat{\mu}_{p_n}}{\widehat{\mu}_{p_n+1}}\right] \tag{2.1}$$

where (p_n) is a nonrandom positive sequence such that $p_n \to \infty$, $a > 0$ and

$$\widehat{\mu}_{p_n} := \frac{1}{n}\sum_{i=1}^{n} X_i^{p_n}$$

is the classical moment estimator of $\mu_{p_n} := \mathbb{E}(X^{p_n})$. From a practical point of view, taking high order moments gives more weight to observations close to θ. From a theoretical point of view, the estimator $\widehat{\theta}_n$ converges in probability to θ without any parametric assumption on the distribution of X, see Section 2.2.2. The asymptotic normality of the estimator is established in Section 2.2.3 when the cumulative distribution function F of X belongs to the Weibull max-domain of attraction.

2.2.1 Construction of the estimator

To justify the introduction of our estimator (2.1), pick θ, $\alpha > 0$ and let first Y be a random variable with survival function $\overline{G}$ defined by $\overline{G}(y) = (1 - y/\theta)^\alpha$ for all $y \in [0, \theta]$. We get, for all $p \geq 1$,

$$M_p := \mathbb{E}(Y^p) = p \int_0^\infty y^{p-1}\, \overline{G}(y)\, dy = \alpha\, \theta^p\, B(p+1,\, \alpha) \tag{2.2}$$

where $B(x,\, y) = \displaystyle\int_0^1 t^{x-1}\,(1-t)^{y-1}\, dt$ is the Beta function. This yields, for all $p \geq 1$,

$$\frac{M_p}{M_{p+1}} = \frac{1}{\theta}\left[1 + \frac{\alpha}{p+1}\right] \tag{2.3}$$

leading to, for every nonrandom sequence (p_n) of real numbers ≥ 1 and every $a > 0$,

$$\frac{1}{\theta} = \frac{1}{ap_n}\left[((a+1)p_n + 1)\frac{M_{(a+1)p_n}}{M_{(a+1)p_n+1}} - (p_n+1)\frac{M_{p_n}}{M_{p_n+1}}\right].$$

Using the above ideas, we shall then define our estimator in two steps. First, the moment M_{p_n} is replaced by the true moment μ_{p_n}, we assume that $p_n \to \infty$ and we set

$$\frac{1}{\Theta_n} := \frac{1}{ap_n}\left[((a+1)p_n + 1)\frac{\mu_{(a+1)p_n}}{\mu_{(a+1)p_n+1}} - (p_n+1)\frac{\mu_{p_n}}{\mu_{p_n+1}}\right].$$

Second, μ_{p_n} is estimated by the corresponding empirical moment $\widehat{\mu}_{p_n}$; plugging $\widehat{\mu}_{p_n}$ in $1/\Theta_n$ yields the estimator (2.1) of $1/\theta$.

2.2.2 Consistency

In this section, we state and prove the consistency of our estimator in a nonparametric context. The only hypothesis is

(A_0) X is positive and the endpoint $\theta = \sup\{x \geq 0 \,|\, F(x) < 1\}$ of X is finite.

To this end, the first step is to prove a result similar to (2.3) for μ_{p_n}.

Proposition 2.1. *Assume that* (A_0) *holds. Then* $\mu_p/\mu_{p+1} \to 1/\theta$ *as* $p \to \infty$.

Proof. Let $Z = X/\theta$. Then Z is a positive random variable with endpoint equal to 1, and

$$\theta\frac{\mu_p}{\mu_{p+1}} = \frac{\mathbb{E}(Z^p)}{\mathbb{E}(Z^{p+1})}.$$

Applying Lemma 2.2 gives the result. ■

The second step consists in showing that μ_{p_n} can be replaced by its empirical counterpart $\widehat{\mu}_{p_n}$. To this end, let us recall a weak law of large numbers for triangular arrays of nonnegative random variables (see e.g. Corollary 2 p. 358 in Chow and Teicher [23]):

Theorem 2.2 (A weak law of large numbers) *Let (Y_{nj}) be an infinite double array of nonnegative random variables. Assume that*

- *For all $n \in \mathbb{N} \setminus \{0\}$, the Y_{nj}, $1 \le j \le n$ are independent;*
- *For all $n \in \mathbb{N} \setminus \{0\}$, $\sum_{j=1}^{n} \mathbb{E}(Y_{nj}) = 1$;*
- *For all $\varepsilon > 0$, $\sum_{j=1}^{n} \mathbb{E}\left(Y_{nj}\, \mathbb{1}_{\{Y_{nj} \ge \varepsilon\}}\right) \to 0$ as $n \to \infty$.*

Then $\sum_{j=1}^{n} Y_{nj} \xrightarrow{\mathbb{P}} 1$.

We have the following result:

Proposition 2.3. *Assume that (A_0) holds. If $n\,\mu_{p_n}/\theta^{p_n} \to \infty$, then $\widehat{\mu}_{p_n}/\mu_{p_n} \xrightarrow{\mathbb{P}} 1$ as $n \to \infty$.*

Proof. Let $Y_{nj} := X_j^{p_n}/(n\,\mu_{p_n})$ for $1 \le j \le n$. The desired result is then tantamount to $\sum_{j=1}^{n} Y_{nj} \xrightarrow{\mathbb{P}} 1$. Notice next that for all n, the Y_{nj}, $1 \le j \le n$ are positive independent random variables, and $\sum_{j=1}^{n} \mathbb{E}(Y_{nj}) = 1$. Using Theorem 2.2, it is enough to show that

$$\forall\, \varepsilon > 0, \quad \sum_{j=1}^{n} \mathbb{E}\left(Y_{nj}\, \mathbb{1}_{\{Y_{nj} \ge \varepsilon\}}\right) \to 0$$

as $n \to \infty$. The Y_{nj}, $1 \le j \le n$ being identically distributed, it is equivalent to prove that

$$\forall\, \varepsilon > 0, \quad \frac{\theta^{p_n}}{\mu_{p_n}} \mathbb{E}\left(Z^{p_n}\, \mathbb{1}_{\{Z^{p_n} \ge \varepsilon n\, \mu_{p_n}/\theta^{p_n}\}}\right) \to 0$$

where $Z = X/\theta \in [0,\,1]$ almost surely. Since $n\,\mu_{p_n}/\theta^{p_n} \to \infty$, we get, for sufficiently large n

$$\frac{\theta^{p_n}}{\mu_{p_n}} \mathbb{E}\left(Z^{p_n}\, \mathbb{1}_{\{Z^{p_n} \ge \varepsilon n\, \mu_{p_n}/\theta^{p_n}\}}\right) = 0,$$

and the result is proven. ■

Theorem 2.4. *Assume that (A_0) holds. If $n\,\mu_{(a+1)p_n}/\theta^{(a+1)p_n} \to \infty$, then $\widehat{\theta}_n \xrightarrow{\mathbb{P}} \theta$ as $n \to \infty$.*

Proof. Remark first that

$$\frac{\mu_{(a+1)p_n}}{\theta^{(a+1)p_n}} = (a+1)p_n \int_0^1 y^{(a+1)p_n - 1}\, \overline{F}(\theta y)\, dy \le (a+1)p_n \int_0^1 y^{p_n - 1}\, \overline{F}(\theta y)\, dy = (a+1)\frac{\mu_{p_n}}{\theta^{p_n}}$$

and use the hypothesis to get $n\,\mu_{p_n}/\theta^{p_n} \to \infty$. Second, Lemma 2.2 entails $n\,\mu_{p_n+1}/\theta^{p_n+1} \to \infty$ and $n\,\mu_{(a+1)p_n+1}/\theta^{(a+1)p_n+1} \to \infty$ as $n \to \infty$. We can then apply Proposition 2.3 to rewrite the frontier estimator as

$$\frac{1}{\widehat{\theta}_n} = \frac{1}{a p_n}\left[((a+1)p_n + 1)\frac{\mu_{(a+1)p_n}}{\mu_{(a+1)p_n+1}}(1 + o_{\mathbb{P}}(1)) - (p_n+1)\frac{\mu_{p_n}}{\mu_{p_n+1}}(1 + o_{\mathbb{P}}(1))\right].$$

Using Proposition 2.1 yields $\mu_{p_n}/\mu_{p_n+1} \to 1/\theta$ and $\mu_{(a+1)p_n}/\mu_{(a+1)p_n+1} \to 1/\theta$ as $n \to \infty$. Replacing in the above equality, the conclusion follows. ■

2.2.3 Asymptotic normality

We now examine the asymptotic normality of our estimator. Before stating our results, we write a technical hypothesis in the regular variation framework: for an arbitrary $b \in [0, 1]$, we shall say that a slowly varying function L satisfies hypothesis (H_b) if:

(H_b) One has

$$\forall\, x > b, \quad L(x) = c\, \exp\left(\int_1^x \frac{\eta(t)}{t}\, dt\right),$$

where c is a positive constant and η is a bounded Borel function tending to 0 at ∞, continuously derivable on (b, ∞), ultimately monotonic and non identically 0, such that $|\eta'|$ is regularly varying and there exists $\nu \leq 0$ with

$$x\, \frac{\eta'(x)}{\eta(x)} \to \nu \qquad \text{as } x \to \infty.$$

The equality $L(x) = c\, \exp\left(\int_1^x \eta(t)\, t^{-1}\, dt\right)$ is the Karamata representation for normalised slowly varying functions (see Theorem 1.3.1 p. 12 in [14]):

Theorem 2.5 (Karamata) *Let L be a slowly varying function. Then there exist an ultimately positive Borel function C having a finite limit $c > 0$ at ∞ and a Borel function η, converging to 0 at ∞, such that*

$$\forall\, x > 0, \quad L(x) = C(x)\, \exp\left(\int_1^x \frac{\eta(t)}{t}\, dt\right).$$

The function L is said to be normalised when C is a positive constant.

Under (H_b), the function $|\eta|$ is ultimately nonincreasing and regularly varying with index ν, and the function $x \mapsto x\, |\eta'(x)|$ is regularly varying with index ν. In the extreme-value framework, ν is referred to as the second order parameter and (H_b) is a second order condition. Finally, let us note that (H_b) entails $x\, \eta'(x) = \mathrm{O}(\eta(x))$, so that $x\, \eta'(x) \to 0$ as $x \to \infty$.

We can now introduce the additional assumptions needed for our study:

(A_1) $\forall\, x \in [0, \theta]$, $\overline{F}(x) = (1 - x/\theta)^{\alpha}\, L\left((1 - x/\theta)^{-1}\right)$ where θ, $\alpha > 0$ and L is a slowly varying function.

(A_2) L satisfies hypothesis (H_1).

Note that in the general context of extreme-value theory, (A_1) entails that the distribution belongs to the Weibull max-domain of attraction with extreme-value index $-1/\alpha$.

We first show that (2.3) still holds, up to an error term, when M_p is replaced by μ_p.

Proposition 2.6. *Assume that (A_1) and (A_2) hold. Then,*

$$\frac{\mu_p}{\mu_{p+1}} = \frac{M_p}{M_{p+1}} + \mathrm{O}\left(\frac{|\eta(p)|}{p}\right).$$

Proof. Considering the change of variables $x = \dfrac{1}{p}(1 - y/\theta)^{-1}$ in (2.2) yields

$$M_p = p^{-\alpha}\,\theta^p\,\Gamma(\alpha+1)\,R_M(p)$$

with $\Gamma(x) = \displaystyle\int_0^\infty t^{x-1}\,e^{-t}\,dt$ being the Gamma function and

$$R_M(p) = 1 + \frac{I_1\,E_1(p) + I_2\,E_2(p)}{\Gamma(\alpha+1)},$$

where I_1, I_2, $E_1(p)$ and $E_2(p)$ are defined in Lemma 2.9 by

$$\begin{aligned} E_1(p) &= \frac{1}{I_1}\int_0^1 f_p(x)\,x^{-\alpha-2}\,dx - 1, & I_1 &= \int_1^\infty y^\alpha\,e^{-y}\,dy, \\ E_2(p) &= \frac{1}{I_2}\int_1^\infty g_p(x)\,x^{-\alpha-2}\,dx - 1, & I_2 &= \int_0^1 y^\alpha\,e^{-y}\,dy, \end{aligned}$$

and where f_p, g_p are the functions introduced in Lemma 2.8:

$$\begin{aligned} \forall\,x \in (0,\,1], \quad f_p(x) &= \left(1 - \frac{1}{p}\right)^{-\alpha-1}\left(1 + \frac{1}{(p-1)x}\right)^{-\alpha-2}\left(1 - \frac{1}{(p-1)x+1}\right)^{p-1}, \\ \forall\,x \in [1,\,\infty), \quad g_p(x) &= \left(1 - \frac{1}{px}\right)^{p-1}. \end{aligned}$$

The same change of variables yields

$$\mu_p = p\int_0^\infty y^{p-1}\,\overline{F}(y)\,dy = p^{-\alpha}\,\theta^p\,L(p)\,\Gamma(\alpha+1)\,\left[R_M(p) + R_\delta(p)\right] \tag{2.4}$$

with

$$R_\delta(p) = \frac{I_1\,\delta_1(p) + I_2\,\delta_2(p)}{\Gamma(\alpha+1)}$$

where $\delta_1(p)$ and $\delta_2(p)$ are defined in Lemmas 2.10 and 2.11 by

$$\begin{aligned} \delta_1(p) &= \frac{1}{I_1}\int_0^1 f_p(x)\left[\frac{L_1((p-1)x)}{L_1(p-1)} - x\right] x^{-\alpha-3}\,dx, & L_1(x) &= x\,L(x+1), \\ \delta_2(p) &= \frac{1}{I_2}\int_1^\infty g_p(x)\left[\frac{L_2(px)}{L_2(p)} - \frac{1}{x}\right] x^{-\alpha-1}\,dx, & L_2(x) &= \frac{L(x)}{x}. \end{aligned}$$

Since $\displaystyle\int_p^{p+1}\frac{\eta(t)}{t}\,dt = \mathrm{O}\left(\frac{|\eta(p)|}{p}\right)$, one clearly has

$$\frac{\mu_p}{\mu_{p+1}} - \frac{M_p}{M_{p+1}} = \frac{1}{\theta}\left[1 + \frac{1}{p}\right]^\alpha\left[\frac{R_M(p) + R_\delta(p)}{R_M(p+1) + R_\delta(p+1)} - \frac{R_M(p)}{R_M(p+1)}\right] + \mathrm{O}\left(\frac{|\eta(p)|}{p}\right), \tag{2.5}$$

and it is straightforward that

$$\frac{R_M(p) + R_\delta(p)}{R_M(p+1) + R_\delta(p+1)} - \frac{R_M(p)}{R_M(p+1)} = \frac{R_\delta(p)\,R_M(p+1) - R_\delta(p+1)\,R_M(p)}{[R_M(p+1) + R_\delta(p+1)]\,R_M(p+1)}.$$

Furthermore

$$R_\delta(p)\,R_M(p{+}1) - R_\delta(p{+}1)\,R_M(p) = [R_\delta(p) - R_\delta(p{+}1)]\,R_M(p{+}1) - R_\delta(p{+}1)\,[R_M(p) - R_M(p{+}1)];$$

Lemmas 2.9, 2.10 and 2.11 entail that $R_M \to 1$ and $R_\delta \to 0$ as $p \to \infty$, and

$$\begin{aligned} R_\delta(p+1) &= \mathrm{O}\left(|\eta(p)|\left(1+\mathcal{L}(p)\right)\right), \\ R_M(p) - R_M(p+1) &= \mathrm{O}\left(\frac{1}{p^2}\right), \\ R_\delta(p) - R_\delta(p+1) &= \mathrm{O}\left(\frac{|\eta(p)|}{p}\right), \end{aligned}$$

where $\mathcal{L}$ is a slowly varying function. Consequently,

$$\frac{R_M(p)+R_\delta(p)}{R_M(p+1)+R_\delta(p+1)} - \frac{R_M(p)}{R_M(p+1)} = \mathrm{O}\left(\frac{|\eta(p)|}{p} + \frac{|\eta(p)|\left(1+\mathcal{L}(p)\right)}{p^2}\right) = \mathrm{O}\left(\frac{|\eta(p)|}{p}\right),$$

and replacing in (2.5) yields the desired result. ■

Applying Proposition 2.6 enables us to control the bias term introduced when M_{p_n} is replaced by μ_{p_n}:

$$\frac{1}{\Theta_n} = \frac{1}{\theta} + \mathrm{O}\left(\frac{|\eta(p_n)|}{p_n}\right). \tag{2.6}$$

To study the random term, we shall use Lyapounov's central limit theorem, which we recall below (see e.g. Theorem 27.3 p. 312 in Billingsley [13]):

Theorem 2.7 (Lyapounov) *Let (Y_{nj}) be an infinite double array of centered random variables. Assume that there exists $\delta > 0$ such that $Y_{nj} \in L^{2+\delta}$ for all n and j, and that*

- *For all $n \in \mathbb{N} \setminus \{0\}$, the Y_{nj}, $1 \le j \le n$ are independent;*
- *Setting $S_n := \sum_{j=1}^{n} Y_{nj}$ and $\sigma_n^2 = \mathrm{Var}(S_n)$, one has*

$$\sum_{j=1}^{n} \frac{\mathbb{E}|Y_{nj}|^{2+\delta}}{\sigma_n^{2+\delta}} \to 0 \quad \text{as } n \to \infty.$$

Then $S_n/\sigma_n \xrightarrow{d} \mathcal{N}(0,1)$.

Turning to the random term, the following result holds:

Theorem 2.8. *Assume that (A_1) and (A_2) hold. If $n\, p_n^{-\alpha}\, L(p_n) \to \infty$ then*

$$v_n \left(\frac{\widehat{\theta}_n}{\Theta_n} - 1\right) \xrightarrow{d} \mathcal{N}\left(0,\, V(\alpha,\, a)\right) \quad \text{as } n \to \infty,$$

with $v_n = \sqrt{n\, L(p_n)}\, p_n^{-\alpha/2+1}$ and

$$V(\alpha,\, a) = \frac{\alpha+1}{a^2\, \Gamma(\alpha)} \left[2^{-\alpha-2} - 2\frac{(a+1)^{\alpha+1}}{(a+2)^{\alpha+2}} + 2^{-\alpha-2}(a+1)^{\alpha}\right].$$

Proof. Our goal is to prove that the sequence of random variables $\left(\xi_n^{(1)}\right)$ defined by

$$\xi_n^{(1)} = \frac{\theta}{\sqrt{V(\alpha,\, a)}}\, v_n \left(\frac{1}{\widehat{\theta}_n} - \frac{1}{\Theta_n}\right)$$

converges in distribution to a standard Gaussian random variable, Theorem 2.8 then being a simple consequence of this result.

The first step consists in using Lemma 2.13i) in order to linearise $\xi_n^{(1)}$:

$$\begin{aligned}\xi_n^{(1)} &= u_{n,a}^{(1)}\left[\zeta_n^{(1,1)} + \left(\frac{\mu_{p_n+1}}{\widehat{\mu}_{p_n+1}} - 1\right)\zeta_n^{(1,2)} + \left(1+\frac{ap_n}{p_n+1}\right)\left(\frac{\mu_{(a+1)p_n+1}}{\widehat{\mu}_{(a+1)p_n+1}} - 1\right)\zeta_n^{(1,3)}\right](1+\mathrm{o}(1))\\ &= u_{n,a}^{(1)}\left[\zeta_n^{(1,1)} + \mathrm{o}_{\mathbb{P}}\left(\zeta_n^{(1,2)}\right) + \mathrm{o}_{\mathbb{P}}\left(\zeta_n^{(1,3)}\right)\right](1+\mathrm{o}(1)),\end{aligned}$$

in view of Proposition 2.3. Thus, to conclude the proof, it is enough to show that

$$u_{n,a}^{(1)}\,\zeta_n^{(1,1)} \xrightarrow{d} \mathcal{N}(0,1), \tag{2.7a}$$

$$u_{n,a}^{(1)}\,\zeta_n^{(1,2)} \xrightarrow{d} \mathcal{N}(0,\,C_2), \tag{2.7b}$$

$$u_{n,a}^{(1)}\,\zeta_n^{(1,3)} \xrightarrow{d} \mathcal{N}(0,\,C_3), \tag{2.7c}$$

where C_2 and C_3 are suitable constants. Note that in fact, (2.7b) and (2.7c) are stronger than what is necessary, but their proofs are similar to (2.7a). Let us then write $\zeta_n^{(1,1)} = \sum_{k=1}^n S_{n,k}^{(1,1)}$, where

$$\begin{aligned}S_{n,k}^{(1,1)} &= \frac{1}{n}\left\{A_n^{(1)}\right\}^t\left[X_k^{p_n},\ X_k^{p_n+1},\ X_k^{(a+1)p_n},\ X_k^{(a+1)p_n+1}\right]^t,\\ A_n^{(1)} &= \left[a_{n,0}^{(1,1)},\ a_{n,1}^{(1,1)},\ a_{n,2}^{(1,1)},\ a_{n,3}^{(1,1)}\right]^t,\\ a_{n,0}^{(1,1)} &= -1,\\ a_{n,1}^{(1,1)} &= \frac{\mu_{p_n}}{\mu_{p_n+1}},\\ a_{n,2}^{(1,1)} &= \left(1+\frac{ap_n}{p_n+1}\right)\frac{\mu_{p_n+1}}{\mu_{(a+1)p_n+1}},\\ a_{n,3}^{(1,1)} &= -\left(1+\frac{ap_n}{p_n+1}\right)\frac{\mu_{p_n+1}\,\mu_{(a+1)p_n}}{\mu^2_{(a+1)p_n+1}},\end{aligned}$$

with A^t standing for the transposed matrix of A. Since the $S_{n,k}^{(1,1)}$, $1\le k\le n$ are independent, identically distributed and centered random variables, in order to use Theorem 2.7, it remains to prove that

$$\frac{\mathbb{E}\left|S_{n,1}^{(1,1)}\right|^3}{\sqrt{n}\left[\mathrm{Var}\left(S_{n,1}^{(1,1)}\right)\right]^{3/2}} \to 0 \tag{2.8}$$

as $n\to\infty$, which requires to control $\mathrm{Var}\left(S_{n,1}^{(1,1)}\right)$ and $\mathbb{E}\left|S_{n,1}^{(1,1)}\right|^3$.

To compute an equivalent for $\mathrm{Var}\left(S_{n,1}^{(1,1)}\right)$, remark that $\mathrm{Var}\left(S_{n,1}^{(1,1)}\right) = \frac{1}{n^2}\left\{A_n^{(1)}\right\}^t M_n^{(1)}\,A_n^{(1)}$ with

$$M_n^{(1)} = \begin{bmatrix} \mu_{2p_n} & \mu_{2p_n+1} & \mu_{(a+2)p_n} & \mu_{(a+2)p_n+1}\\ \mu_{2p_n+1} & \mu_{2p_n+2} & \mu_{(a+2)p_n+1} & \mu_{(a+2)p_n+2}\\ \mu_{(a+2)p_n} & \mu_{(a+2)p_n+1} & \mu_{(2a+2)p_n} & \mu_{(2a+2)p_n+1}\\ \mu_{(a+2)p_n+1} & \mu_{(a+2)p_n+2} & \mu_{(2a+2)p_n+1} & \mu_{(2a+2)p_n+2}\end{bmatrix}.$$

Let us rewrite that as

$$\begin{aligned}\mathrm{Var}\left(S_{n,1}^{(1,1)}\right) &= \frac{1}{n^2}\left[w(1,\,1,\,p_n) - 2\left(1+\frac{ap_n}{p_n+1}\right)\frac{\mu_{p_n+1}}{\mu_{(a+1)p_n+1}}\,w(1,\,a+1,\,p_n)\right.\\ &\quad\left. + \left(1+\frac{ap_n}{p_n+1}\right)^2\frac{\mu_{p_n+1}^2}{\mu_{(a+1)p_n+1}^2}\,w(a+1,\,a+1,\,p_n)\right]\end{aligned}$$

where, by analogy with Lemma 2.14,

$$w(s,\,t,\,p_n) = \left[-1,\ \frac{\mu_{sp_n}}{\mu_{sp_n+1}}\right]\begin{bmatrix}\mu_{(s+t)p_n} & \mu_{(s+t)p_n+1}\\ \mu_{(s+t)p_n+1} & \mu_{(s+t)p_n+2}\end{bmatrix}\left[-1,\ \frac{\mu_{tp_n}}{\mu_{tp_n+1}}\right]^t.$$

We now use (2.4) together with Lemmas 2.9, 2.10, 2.11 and 2.14 to get

$$w(s,\,t,\,p_n) = \frac{\Gamma(\alpha+1)\,\alpha(\alpha+1)}{(s+t)^{\alpha+2}}\,\theta^{(s+t)p_n}\,p_n^{-\alpha-2}\,L(p_n)\,(1+\mathrm{o}(1)).$$

Taking into account that

$$\left(1+\frac{ap_n}{p_n+1}\right)\frac{\mu_{p_n+1}}{\mu_{(a+1)p_n+1}} = \frac{(a+1)^{\alpha+1}}{\theta^{ap_n}}(1+\mathrm{o}(1)), \tag{2.9}$$

we get

$$\mathrm{Var}\left(S_{n,1}^{(1,1)}\right) = a^2\,\Gamma^2(\alpha+1)\,V(\alpha,\,a)\,\frac{\theta^{2p_n}\,p_n^{-\alpha-2}\,L(p_n)}{n^2}\,(1+\mathrm{o}(1)).$$

Since $n\,p_n^{-\alpha}\,L(p_n)\to\infty$, it now suffices to prove that

$$\mathbb{E}\left|S_{n,1}^{(1,1)}\right|^3 = \mathrm{O}\left(n^{-3}\,\theta^{3p_n}\,p_n^{-\alpha-3}\,L(p_n)\right).$$

To this aim, let us introduce $Z_1 = X_1/\theta$: Hölder's inequality leads to

$$\begin{aligned}\frac{\mathbb{E}\left|S_{n,1}^{(1,1)}\right|^3}{n^{-3}\,\theta^{3p_n}} &\le 4\,\mathbb{E}\left|Z_1^{p_n}\left[a_{n,0}^{(1,1)} + a_{n,1}^{(1,1)}\,\theta\,Z_1\right]\right|^3\\ &+ 4\,\mathbb{E}\left|Z_1^{(a+1)p_n}\left[a_{n,2}^{(1,1)}\,\theta^{ap_n} + a_{n,3}^{(1,1)}\,\theta^{ap_n+1}\,Z_1\right]\right|^3.\end{aligned}$$

The random variable Z_1 has survival function $\overline{F}_1(x) = (1-x)^\alpha\,L\left((1-x)^{-1}\right)$; setting

$$\begin{aligned}H_{n,0}^{(1,1)}(z) &= -1,\\ H_{n,1}^{(1,1)}(z) &= \alpha z,\\ H_{n,2}^{(1,1)}(z) &= \left(1+\frac{ap_n}{p_n+1}\right)\theta^{ap_n}\,\frac{\mu_{p_n+1}}{\mu_{(a+1)p_n+1}},\\ H_{n,3}^{(1,1)}(z) &= -\left(1+\frac{ap_n}{p_n+1}\right)\theta^{ap_n}\,\frac{\mu_{p_n+1}}{\mu_{(a+1)p_n+1}}\,\frac{\alpha z}{a+1},\end{aligned}$$

some straightforward computations show that there exist two sequences of Borel functions $\left(\chi_n^{(1,1)}\right)$ and $\left(\chi_n^{(1,2)}\right)$ uniformly converging to 0 on $[0,\,1]$ such that for all $z\in[0,\,1]$,

$$\begin{aligned}a_{n,0}^{(1,1)} + a_{n,1}^{(1,1)}\,\theta z &= H_{n,0}^{(1,1)}(z)(1-z) + \frac{H_{n,1}^{(1,1)}(z)+\chi_n^{(1,1)}(z)}{p_n},\\ a_{n,2}^{(1,1)}\,\theta^{ap_n} + a_{n,3}^{(1,1)}\,\theta^{ap_n+1}z &= H_{n,2}^{(1,1)}(z)(1-z) + \frac{H_{n,3}^{(1,1)}(z)+\chi_n^{(1,2)}(z)}{p_n}.\end{aligned}$$

Recalling (2.9), we obtain that the $H_{n,j}^{(1,1)}$, $j \in \{0, 1, 2, 3\}$ are Borel uniformly bounded functions on [0, 1], so that we can use Lemma 2.15 twice to obtain the desired bound for $\mathbb{E}\left|S_{n,1}^{(1,1)}\right|^3$. Finally, applying Theorem 2.7 concludes the proof of (2.7a).

Proofs of (2.7b) and (2.7c) are completely similar since $\zeta_n^{(1,2)}$ and $\zeta_n^{(1,3)}$ can be rewritten as

$$\begin{aligned}
\zeta_n^{(1,2)} &= \sum_{k=1}^{n} S_{n,k}^{(1,2)} \quad \text{with} \quad S_{n,k}^{(1,2)} = \frac{1}{n}\left[a_{n,0}^{(1,2)},\ a_{n,1}^{(1,2)}\right]\left[X_k^{p_n},\ X_k^{p_n+1}\right]^t, \\
\zeta_n^{(1,3)} &= \sum_{k=1}^{n} S_{n,k}^{(1,3)} \quad \text{with} \quad S_{n,k}^{(1,3)} = \frac{1}{n}\left[a_{n,0}^{(1,3)},\ a_{n,1}^{(1,3)}\right]\left[X_k^{(a+1)p_n},\ X_k^{(a+1)p_n+1}\right]^t
\end{aligned}$$

where

$$\begin{aligned}
a_{n,0}^{(1,2)} &= -1, \\
a_{n,1}^{(1,2)} &= \frac{\mu_{p_n}}{\mu_{p_n+1}}, \\
a_{n,0}^{(1,3)} &= \frac{\mu_{p_n+1}}{\mu_{(a+1)p_n+1}}, \\
a_{n,1}^{(1,3)} &= -\frac{\mu_{p_n+1}\,\mu_{(a+1)p_n}}{\mu^2_{(a+1)p_n+1}}.
\end{aligned}$$

Applying Lemma 2.15 with

$$\begin{aligned}
H_{n,0}^{(1,2)}(z) &= -1, \\
H_{n,1}^{(1,2)}(z) &= \alpha z, \\
H_{n,0}^{(1,3)}(z) &= \theta^{ap_n}\,\frac{\mu_{p_n+1}}{\mu_{(a+1)p_n+1}}, \\
H_{n,1}^{(1,3)}(z) &= -\theta^{ap_n}\,\frac{\mu_{p_n+1}}{\mu_{(a+1)p_n+1}}\,\frac{\alpha z}{a+1}
\end{aligned}$$

yields $\mathbb{E}\left|S_{n,1}^{(1,j)}\right|^3 = \mathrm{O}\left(n^{-3}\,\theta^{3p_n}\,p_n^{-\alpha-3}\,L(p_n)\right)$, $j \in \{2, 3\}$. Using Theorem 2.7 then allows us to complete the proof of Theorem 2.8. ■

The asymptotic normality of $\widehat{\theta}_n$ centered on the true endpoint θ is then a consequence of (2.6) and Theorem 2.8:

Theorem 2.9. *Assume that* (A_1) *and* (A_2) *hold. If* $n\,p_n^{-\alpha}\,L(p_n) \to \infty$ *and* $n\,p_n^{-\alpha}\,L(p_n)\,\eta^2(p_n) \to 0$, *then*

$$v_n\left(\frac{\widehat{\theta}_n}{\theta} - 1\right) \xrightarrow{d} \mathcal{N}\left(0,\, V(\alpha,\, a)\right) \quad \text{as } n \to \infty,$$

with the notations of Theorem 2.8.

Proof. Notice that

$$\frac{\widehat{\theta}_n}{\theta} - 1 = \frac{\Theta_n}{\theta}\left[\frac{\widehat{\theta}_n}{\Theta_n} - 1\right] + \left[\frac{\Theta_n}{\theta} - 1\right].$$

Since $\Theta_n/\theta \to 1$ as $n \to \infty$, Theorem 2.8 and Slutsky's lemma yield

$$\frac{\Theta_n}{\theta}\, v_n \left[\frac{\widehat{\theta}_n}{\Theta_n} - 1\right] \xrightarrow{d} \mathcal{N}\left(0,\, V(\alpha,\, a)\right) \quad \text{as } n \to \infty.$$

Recalling (2.6), we get, using the hypothesis $n\, p_n^{-\alpha}\, L(p_n)\, \eta^2(p_n) \to 0$,

$$v_n \left[\frac{\Theta_n}{\theta} - 1\right] \to 0$$

as $n \to \infty$. Using once again Slutsky's lemma completes the proof. ■

2.2.4 A special case: the Hall model

Here, we focus on the particular case of the Hall model (3): namely,

$$\exists\, M > 0, \quad \forall\, x > M, \quad L(x) = C + Dx^{-\beta}(1 + \delta(x))$$

where C, $\beta > 0$, $D \in \mathbb{R} \setminus \{0\}$ and δ is a bounded Borel function going to 0 at ∞. We state analogues of Theorems 2.8 and 2.9; the proofs are based on those of these results, and we only show the main differences between the general case and the Hall case.

Note first that putting

$$\widetilde{\delta}(x) = \begin{cases} x^{\beta}\, \dfrac{L(x) - C}{D} - 1 & \text{if } 1 \leq x \leq M \\ \delta(x) & \text{if } x > M \end{cases}$$

and noting that L is locally bounded away from 0 and ∞ on $[1,\, \infty)$ because $\overline{F}$ is a survival function, $\widetilde{\delta}$ is then a bounded Borel function going to 0 at ∞. Moreover,

$$\forall\, x > 1, \quad L(x) = C + Dx^{-\beta}\left(1 + \widetilde{\delta}(x)\right).$$

Consequently, we may restrict ourselves to the case $M = 1$. We therefore write

$$\mu_p = C\, \alpha\, \theta^p\, B(p + 1,\, \alpha)\, (1 + \varepsilon(p))$$

where the error term is

$$\varepsilon(p) = \frac{D}{C}\, \frac{B(p,\, \alpha + \beta + 1)}{B(p,\, \alpha + 1)} \left[1 + \frac{\displaystyle\int_0^1 x^{p-1}\, (1 - x)^{\alpha+\beta}\, \delta\left((1 - x)^{-1}\right)\, dx}{B(p,\, \alpha + \beta + 1)}\right]. \tag{2.10}$$

We are now ready to state our results in that context. Note that, contrary to the general case, there is no regularity hypothesis on the function L. The first theorem is an analogue of Theorem 2.8.

Theorem 2.10. *In the case of the Hall model (3), assume that $n\, p_n^{-\alpha} \to \infty$ as $n \to \infty$. Then*

$$w_n \left(\frac{\widehat{\theta}_n}{\Theta_n} - 1\right) \xrightarrow{d} \mathcal{N}\left(0,\, \frac{V(\alpha,\, a)}{C}\right) \quad \text{as } n \to \infty$$

with $w_n = \sqrt{n}\, p_n^{-\alpha/2+1}$.

Proof of Theorem 2.10. Using Tricomi and Erdélyi [103] and Lemma 2.16iii) entails

$$(n,\, x) \mapsto (p_n + x)^{\beta} \frac{B(p_n + x,\, \alpha + \beta + 1)}{B(p_n + x,\, \alpha + 1)} \in \mathcal{E}_2.$$

Recalling (2.10), applying Lemmas 2.16i) and 2.17 leads to

$$(n,\, x) \mapsto (p_n + x)^{\beta}\, \varepsilon(p_n + x) \in \mathcal{E}_2.$$

With the help of the equality

$$\frac{\mu_p}{\mu_{p+1}} = \frac{1}{\theta}\left[1 + \frac{\alpha}{p+1}\right]\left[1 + \frac{\varepsilon(p) - \varepsilon(p+1)}{1 + \varepsilon(p+1)}\right],$$

the proof is then a straightforward adaptation of the one of Theorem 2.8. ■

Because $(n,\, x) \mapsto (p_n + x)^{\beta}\, \varepsilon(p_n + x) \in \mathcal{C}_2$, one has, as $n \to \infty$

$$p_n^{\beta+1}\{\varepsilon(p_n + 1) - \varepsilon(p_n)\} \to 0.$$

This is the key to the proof of the asymptotic normality of $\widehat{\theta}_n$ centered on the parameter θ:

Theorem 2.11. *In the case of the Hall model (3), assume that* $n\, p_n^{-\alpha} \to \infty$ *and that* $n\, p_n^{-\alpha-2\beta} \to 0$ *as* $n \to \infty$. *Then*

$$w_n\left(\frac{\widehat{\theta}_n}{\theta} - 1\right) \xrightarrow{d} \mathcal{N}\left(0,\, \frac{V(\alpha,\, a)}{C}\right) \quad \text{as } n \to \infty$$

with the notations of Theorem 2.10.

Proof. The proof is an immediate adaptation of that of Theorem 2.9. ■

To conclude this section, let us finally draw a comparison between the hypotheses of Theorem 2.9 and Theorem 2.11. Elementary computations show that in the case of the Hall model, if δ is a twice derivable bounded function, with $x\,\delta'(x) \to 0$ and $x^2\,\delta''(x) \to 0$ as $x \to \infty$, then L is twice continously derivable on $(1,\, \infty)$, with

$$\eta(y) := y\frac{L'(y)}{L(y)} = Dy^{-\beta}\frac{y\,\delta'(y) - \beta(1 + \delta(y))}{C + Dy^{-\beta}(1 + \delta(y))}$$

and

$$\eta'(y) = \frac{\eta(y)}{y}[\beta - \eta(y)] + Dy^{-\beta}\frac{(1-\beta)\delta'(y) + y\,\delta''(y)}{C + Dy^{-\beta}(1 + \delta(y))}$$

so that hypothesis (A_2) holds. Moreover,

$$\eta(p) = -\beta\,\frac{D}{C}\,p^{-\beta}\,(1 + \mathrm{o}(1))$$

so that the hypotheses of Theorem 2.9 on the rate of divergence of (p_n) translate to $n\, p_n^{-\alpha} \to \infty$ and $n\, p_n^{-\alpha-2\beta} \to 0$ as $n \to \infty$, which are the hypotheses of Theorem 2.11.

2.3 The general case

2.3.1 The framework

In this section, we no longer assume that X is positive. It is therefore impossible to use moments of the variable of interest X, since $|X|$ could have an infinite mean. To solve this problem, we consider the random variables $(Y_1, \ldots, Y_n)$ such that $Y_k = e^{X_k}$. Y has a bounded support $[0, \varphi]$ where $\varphi = e^{\theta}$. It is enough to provide an estimate $\widetilde{\varphi}_n$ of φ; we introduce an estimator using high moments of the variable of interest Y. More precisely, the estimator is given by

$$\widetilde{\varphi}_n := \left[\frac{\widetilde{\mathfrak{m}}_{p_n}}{\widetilde{\mathfrak{m}}_{p_n+1}} \frac{\widetilde{\mathfrak{m}}_{(a+1)p_n+a+1}}{\widetilde{\mathfrak{m}}_{(a+1)p_n}} \right]^{1/a} \tag{2.11}$$

where (p_n) is a nonrandom positive sequence such that $p_n \to \infty$, $a > 0$ and

$$\widetilde{\mathfrak{m}}_{p_n} := \frac{1}{n} \sum_{i=1}^{n} Y_i^{p_n} = \frac{1}{n} \sum_{i=1}^{n} e^{p_n X_i}$$

is the classical empirical counterpart of $\mathfrak{m}_{p_n} := \mathbb{E}(Y^{p_n}) = \mathbb{E}\left(e^{p_n X}\right)$. Our estimator of θ is then

$$\widetilde{\theta}_n := \ln \widetilde{\varphi}_n = \frac{1}{a} \left\{ \ln \left[\frac{\widetilde{\mathfrak{m}}_{p_n}}{\widetilde{\mathfrak{m}}_{p_n+1}} \right] - \ln \left[\frac{\widetilde{\mathfrak{m}}_{(a+1)p_n}}{\widetilde{\mathfrak{m}}_{(a+1)p_n+a+1}} \right] \right\}. \tag{2.12}$$

We shall show the consistency of our estimator without any parametric hypothesis on the cumulative distribution function of X: we only assume that

(B_0) The endpoint $\theta = \sup\{x \in \mathbb{R} \,|\, F(x) < 1\}$ of X is finite.

We further prove that our estimator is asymptotically Gaussian when the cumulative distribution function of X belongs to the Weibull max-domain of attraction.

2.3.2 Construction of the estimator

To justify the introduction of our estimator, we first state an analogue of Proposition 2.1.

Proposition 2.12. *Assume that* (B_0) *holds. Then for all* $u \geq 1$, *one has* $\mathfrak{m}_p/\mathfrak{m}_{p+u} \to 1/\varphi^u$ *as* $p \to \infty$.

Proof. Like in the proof of Proposition 2.1, let $Z = Y/\varphi$ and write

$$\varphi^u \frac{\mathfrak{m}_p}{\mathfrak{m}_{p+u}} = \frac{\mathbb{E}(Z^p)}{\mathbb{E}(Z^{p+u})}.$$

Applying Lemma 2.2 to the random variable Z completes the proof. ■

Picking $a > 0$, if (p_n) is a nonrandom positive sequence such that $p_n \to \infty$, one then has

$$\varphi = \Phi_n \left(1 + \mathrm{o}(1)\right) \quad \text{where} \quad \Phi_n = \left[\frac{\mathfrak{m}_{p_n}}{\mathfrak{m}_{p_n+1}} \frac{\mathfrak{m}_{(a+1)p_n+a+1}}{\mathfrak{m}_{(a+1)p_n}} \right]^{1/a}. \tag{2.13}$$

Using this idea, our estimator is defined by replacing the true moment $\mathfrak{m}_{p_n}$ with its empirical counterpart $\widetilde{\mathfrak{m}}_{p_n}$ in the expression of Φ_n, thus yielding the estimator (2.11) of φ, and the corresponding estimate of θ is $\widetilde{\theta}_n = \ln \widetilde{\varphi}_n$ as in (2.12).

2.3.3 Asymptotic properties

Before stating the consistency of our estimator, we give a straightforward analogue of Proposition 2.3:

Proposition 2.13. *Assume that (B_0) holds. If $n\,\mathfrak{m}_{p_n}/\varphi^{p_n} \to \infty$, then $\widetilde{\mathfrak{m}}_{p_n}/\mathfrak{m}_{p_n} \xrightarrow{\mathbb{P}} 1$ as $n \to \infty$.*

The consistency of our estimator is now an easy consequence of Proposition 2.13:

Theorem 2.14. *Assume that (B_0) holds. If $n\,\mathfrak{m}_{(a+1)p_n}/\varphi^{(a+1)p_n} \to \infty$, then $\widetilde{\theta}_n \xrightarrow{\mathbb{P}} \theta$ as $n \to \infty$.*

Proof of Theorem 2.14. The same ideas that were used in the proof of Theorem 2.4 yield $\mathfrak{m}_{(a+1)p_n}/\varphi^{(a+1)p_n} \leq (a+1)\mathfrak{m}_{p_n}/\varphi^{p_n}$, and therefore $n\,\mathfrak{m}_{p_n}/\varphi^{p_n} \to \infty$. Use now Lemma 2.2 to obtain $n\,\mathfrak{m}_{p_n+1}/\varphi^{p_n+1} \to \infty$ and $n\,\mathfrak{m}_{(a+1)p_n+a+1}/\varphi^{(a+1)p_n+a+1} \to \infty$ as $n \to \infty$. We can then apply Proposition 2.3 to get $\widetilde{\varphi}_n/\Phi_n \xrightarrow{\mathbb{P}} 1$. Recalling (2.13), we obtain $\widetilde{\varphi}_n \xrightarrow{\mathbb{P}} \varphi$: taking logarithms, the conclusion follows. ■

To get the limiting distribution of our estimator, we need some more hypotheses:

(B_1) $\forall\, x < \theta$, $\overline{F}(x) = (\theta - x)^\alpha\, L\left((\theta - x)^{-1}\right)$ where $\theta \in \mathbb{R}$, $\alpha > 0$ and L is a slowly varying function.

(B_2) L satisfies hypothesis (H_0).

Note that (B_1) and (B_2) are the exact analogues of (A_1) and (A_2) if the positivity assumption on X is dropped. In particular, hypothesis (B_1) holds if and only if F belongs to the Weibull max-domain of attraction.

We now compute an asymptotic expansion of $\mathfrak{m}_p/\mathfrak{m}_{p+u}$ in that particular context, which is the key to our analysis below:

Proposition 2.15. *Assume that (B_1) and (B_2) hold. Then, with previous notations,*

$$\forall\, u \geq 1, \quad \frac{\mathfrak{m}_p}{\mathfrak{m}_{p+u}} = \frac{1}{\varphi^u}\left[1 + \frac{u}{p}\right]^\alpha + \mathrm{O}\left(\frac{|\eta(p)|}{p}\right).$$

Proof of Proposition 2.15. Using the change of variables $x = \dfrac{1}{p}(\ln\varphi - \ln y)^{-1}$ in the equality

$$\mathfrak{m}_p = p\int_0^\varphi y^{p-1}\,\overline{F}(\ln y)\,dy$$

entails

$$\mathfrak{m}_p = p^{-\alpha}\,\varphi^p\,L(p)\,[\Gamma(\alpha+1) + I_1\,\varepsilon_1(p) + I_2\,\varepsilon_2(p)] \tag{2.14}$$

where

$$\begin{aligned} \varepsilon_1(p) &:= \frac{1}{I_1}\int_0^1 e^{-1/x}\,x^{-\alpha-3}\left[\frac{L_1(px)}{L_1(p)} - x\right]dx, & L_1(x) &= x\,L(x),\\ \varepsilon_2(p) &:= \frac{1}{I_2}\int_1^\infty e^{-1/x}\,x^{-\alpha-1}\left[\frac{L_2(px)}{L_2(p)} - \frac{1}{x}\right]dx, & L_2(x) &= \frac{L(x)}{x}. \end{aligned}$$

Putting

$$\tau(p,\,u) := \frac{I_1[\varepsilon_1(p) - \varepsilon_1(p+u)] + I_2[\varepsilon_2(p) - \varepsilon_2(p+u)]}{\Gamma(\alpha+1) + I_1\,\varepsilon_1(p+u) + I_2\,\varepsilon_2(p+u)},$$

one has

$$\forall\, u \geq 1, \quad \frac{\mathfrak{m}_p}{\mathfrak{m}_{p+u}} = \frac{1}{\varphi^u}\left[1 + \frac{u}{p}\right]^\alpha \exp\left[-\int_p^{p+u}\frac{\eta(t)}{t}\,dt\right][1 + \tau(p,\,u)].$$

Besides, Lemmas 2.10 and 2.11 give $\tau(p,\,u) = \mathrm{O}\left(\frac{|\eta(p)|}{p}\right)$; because $\int_p^{p+u}\frac{\eta(t)}{t}\,dt = \mathrm{O}\left(\frac{|\eta(p)|}{p}\right)$, one clearly has

$$\frac{\mathfrak{m}_p}{\mathfrak{m}_{p+u}} - \frac{1}{\varphi^u}\left[1 + \frac{u}{p}\right]^\alpha = \mathrm{O}\left(\frac{|\eta(p)|}{p}\right)$$

which ends the proof. ■

Applying Proposition 2.15 enables us to control the bias term introduced when φ is replaced by Φ_n:

$$\Phi_n = \varphi + \mathrm{O}\left(\frac{|\eta(p_n)|}{p_n}\right). \tag{2.15}$$

We now turn to the random term. We have:

Theorem 2.16. *Assume that* (B_1) *and* (B_2) *hold. If* $n\,p_n^{-\alpha}\,L(p_n) \to \infty$, *then*

$$v_n\,(\widetilde{\varphi}_n^a - \Phi_n^a) \xrightarrow{d} \mathcal{N}\left(0,\,\varphi^{2a}\,a^2\,V(\alpha,\,a)\right) \quad \text{as } n \to \infty$$

with the notations of Theorem 2.8.

Proof of Theorem 2.16. We closely follow the proof of Theorem 2.8: our goal is to prove that the sequence of random variables $\left(\xi_n^{(2)}\right)$ defined by

$$\xi_n^{(2)} = \frac{\varphi^{-a}}{a\sqrt{V(\alpha,\,a)}}\,v_n\,(\widetilde{\varphi}_n^a - \Phi_n^a)$$

converges in distribution to a standard Gaussian random variable, Theorem 2.16 then being a simple consequence of this result.

Lemma 2.13ii) allows us to linearise $\xi_n^{(2)}$:

$$\begin{aligned} \xi_n^{(2)} &= u_{n,\,a}^{(2)}\left[\zeta_n^{(2,\,1)} + \left(\frac{1}{\varphi^{a+1}}\,\frac{\widetilde{\mathfrak{m}}_{(a+1)p_n+a+1}}{\widetilde{\mathfrak{m}}_{(a+1)p_n}}\,\frac{\mathfrak{m}_{p_n+1}}{\widetilde{\mathfrak{m}}_{p_n+1}} - 1\right)\zeta_n^{(2,\,2)} + \left(\frac{\mathfrak{m}_{(a+1)p_n}}{\widetilde{\mathfrak{m}}_{(a+1)p_n}} - 1\right)\zeta_n^{(2,\,3)}\right]\\ &\times\ (1 + \mathrm{o}(1)). \end{aligned}$$

Proposition 2.3 and Lemma 2.12 then lead to the equality

$$\xi_n^{(2)} = u_{n,a}^{(2)} \left[\zeta_n^{(2,1)} + o_{\mathbb{P}}\left(\zeta_n^{(2,2)}\right) + o_{\mathbb{P}}\left(\zeta_n^{(2,3)}\right)\right](1 + o(1)).$$

Thus, to conclude the proof, it is enough to show that

$$u_{n,a}^{(2)} \zeta_n^{(2,1)} \xrightarrow{d} \mathcal{N}(0,1), \tag{2.16a}$$

$$u_{n,a}^{(2)} \zeta_n^{(2,2)} \xrightarrow{d} \mathcal{N}(0, C_2), \tag{2.16b}$$

$$u_{n,a}^{(2)} \zeta_n^{(2,3)} \xrightarrow{d} \mathcal{N}(0, C_3), \tag{2.16c}$$

where C_2 and C_3 are suitable constants. Let us then write $\zeta_n^{(2,1)} = \sum_{k=1}^n S_{n,k}^{(2,1)}$, where

$$\begin{aligned}
S_{n,k}^{(2,1)} &= \frac{1}{n} \left\{A_n^{(2)}\right\}^t \left[Y_k^{p_n},\ Y_k^{p_n+1},\ Y_k^{(a+1)p_n},\ Y_k^{(a+1)p_n+a+1}\right]^t, \\
A_n^{(2)} &= \left[a_{n,0}^{(2,1)},\ a_{n,1}^{(2,1)},\ a_{n,2}^{(2,1)},\ a_{n,3}^{(2,1)}\right]^t, \\
a_{n,0}^{(2,1)} &= 1, \\
a_{n,1}^{(2,1)} &= -\frac{\mathfrak{m}_{p_n}}{\mathfrak{m}_{p_n+1}}, \\
a_{n,2}^{(2,1)} &= -\frac{1}{\varphi^{a+1}} \frac{\mathfrak{m}_{p_n}\, \mathfrak{m}_{(a+1)p_n+a+1}}{\mathfrak{m}_{(a+1)p_n}^2}, \\
a_{n,3}^{(2,1)} &= \frac{1}{\varphi^{a+1}} \frac{\mathfrak{m}_{p_n}}{\mathfrak{m}_{(a+1)p_n}}.
\end{aligned}$$

Because the $S_{n,k}^{(2,1)}$, $1 \le k \le n$ are independent, identically distributed and centered random variables, we shall prove that

$$\frac{\mathbb{E}\left|S_{n,1}^{(2,1)}\right|^3}{\sqrt{n}\left[\operatorname{Var}\left(S_{n,1}^{(2,1)}\right)\right]^{3/2}} \to 0$$

as $n \to \infty$, and use Theorem 2.7.

To compute an equivalent for $\operatorname{Var}\left(S_{n,1}^{(2,1)}\right)$, remark that $\operatorname{Var}\left(S_{n,1}^{(2,1)}\right) = \frac{1}{n^2}\left\{A_n^{(2)}\right\}^t M_n^{(2)} A_n^{(2)}$ with

$$M_n^{(2)} = \begin{bmatrix}
\mathfrak{m}_{2p_n} & \mathfrak{m}_{2p_n+1} & \mathfrak{m}_{(a+2)p_n} & \mathfrak{m}_{(a+2)p_n+a+1} \\
\mathfrak{m}_{2p_n+1} & \mathfrak{m}_{2p_n+2} & \mathfrak{m}_{(a+2)p_n+1} & \mathfrak{m}_{(a+2)p_n+a+2} \\
\mathfrak{m}_{(a+2)p_n} & \mathfrak{m}_{(a+2)p_n+1} & \mathfrak{m}_{(2a+2)p_n} & \mathfrak{m}_{(2a+2)p_n+a+1} \\
\mathfrak{m}_{(a+2)p_n+a+1} & \mathfrak{m}_{(a+2)p_n+a+2} & \mathfrak{m}_{(2a+2)p_n+a+1} & \mathfrak{m}_{(2a+2)p_n+2a+2}
\end{bmatrix}.$$

Rewrite that as

$$\begin{aligned}
\operatorname{Var}\left(S_{n,1}^{(2,1)}\right) &= \frac{1}{n^2}\Bigg[w(1,\,1,\,1,\,1,\,p_n) - \frac{2}{\varphi^{a+1}} \frac{\mathfrak{m}_{p_n}\,\mathfrak{m}_{(a+1)p_n+a+1}}{\mathfrak{m}_{(a+1)p_n}^2} w(1,\,a+1,\,1,\,a+1,\,p_n) \\
&\quad + \frac{1}{\varphi^{2(a+1)}} \frac{\mathfrak{m}_{p_n}^2\,\mathfrak{m}_{(a+1)p_n+a+1}^2}{\mathfrak{m}_{(a+1)p_n}^4} w(a+1,\,a+1,\,a+1,\,a+1,\,p_n)\Bigg]
\end{aligned}$$

where, as in Lemma 2.14,

$$w(s,\,t,\,u,\,v,\,p_n) = \left[1,\ -\frac{\mathfrak{m}_{sp_n}}{\mathfrak{m}_{sp_n+u}}\right] \left[\begin{array}{cc} \mathfrak{m}_{(s+t)p_n} & \mathfrak{m}_{(s+t)p_n+v} \\ \mathfrak{m}_{(s+t)p_n+u} & \mathfrak{m}_{(s+t)p_n+u+v} \end{array}\right] \left[1,\ -\frac{\mathfrak{m}_{tp_n}}{\mathfrak{m}_{tp_n+v}}\right]^t.$$

We now use the expansion (2.14) and apply Lemmas 2.9, 2.10, 2.11 and 2.14 to obtain

$$w(s,\,t,\,u,\,v,\,p_n) = \frac{\Gamma(\alpha+1)\,\alpha(\alpha+1)\,uv}{(s+t)^{\alpha+2}}\,\varphi^{(s+t)p_n}\,p_n^{-\alpha-2}\,L(p_n)\,(1+\mathrm{o}(1))$$

so that, thanks to the equivalent

$$\frac{1}{\varphi^{a+1}}\,\frac{\mathfrak{m}_{p_n}\,\mathfrak{m}_{(a+1)p_n+a+1}}{\mathfrak{m}^2_{(a+1)p_n}} = \frac{(a+1)^\alpha}{\varphi^{ap_n}}(1+\mathrm{o}(1)),$$

we get

$$\mathrm{Var}\left(S^{(2,\,1)}_{n,\,1}\right) = a^2\,\Gamma^2(\alpha+1)\,V(\alpha,\,a)\,\frac{\varphi^{2p_n}\,p_n^{-\alpha-2}\,L(p_n)}{n^2}\,(1+\mathrm{o}(1)).$$

To control $\mathbb{E}\left|S^{(2,\,1)}_{n,\,1}\right|^3$, write $Z_1 = Y_1/\varphi$ and use Hölder's inequality to get

$$\begin{aligned} \frac{\mathbb{E}\left|S^{(2,\,1)}_{n,\,1}\right|^3}{n^{-3}\,\varphi^{3p_n}} \quad &\le \quad 4\,\mathbb{E}\left|Z_1^{p_n}\left[a^{(2,\,1)}_{n,\,0} + a^{(2,\,1)}_{n,\,1}\,\varphi\,Z_1\right]\right|^3 \\ &+ \quad 4\,\mathbb{E}\left|Z_1^{(a+1)p_n}\left[a^{(2,\,1)}_{n,\,2}\,\varphi^{ap_n} + a^{(2,\,1)}_{n,\,3}\,\varphi^{ap_n+a+1}\,Z_1^{a+1}\right]\right|^3; \end{aligned}$$

Z_1 has survival function $\overline{G}_1$ defined by $\overline{G}_1(x) = (-\ln x)^\alpha\,L\left((-\ln x)^{-1}\right) \quad \forall\,x \in (0,\,1)$. Setting

$$\begin{aligned} H^{(2,\,1)}_{n,\,0}(z) &= 1, \\ H^{(2,\,1)}_{n,\,1}(z) &= -\alpha z, \\ H^{(2,\,1)}_{n,\,2}(z) &= -\varphi^{ap_n}\,\frac{\mathfrak{m}_{p_n}}{\mathfrak{m}_{(a+1)p_n}}\,\frac{1-z^{a+1}}{1-z}, \\ H^{(2,\,1)}_{n,\,3}(z) &= \alpha\,\varphi^{ap_n}\,\frac{\mathfrak{m}_{p_n}}{\mathfrak{m}_{(a+1)p_n}}, \end{aligned}$$

some more easy computations show that there exist two sequences of Borel functions $\left(\chi_n^{(2,\,1)}\right)$ and $\left(\chi_n^{(2,\,2)}\right)$ uniformly converging to 0 on $[0,\,1]$ such that for all $z \in [0,\,1]$,

$$\begin{aligned} a^{(2,\,1)}_{n,\,0} + a^{(2,\,1)}_{n,\,1}\,\varphi z &= H^{(2,\,1)}_{n,\,0}(z)(1-z) + \frac{H^{(2,\,1)}_{n,\,1}(z) + \chi_n^{(2,\,1)}(z)}{p_n}, \\ a^{(2,\,1)}_{n,\,2}\,\varphi^{ap_n} + a^{(2,\,1)}_{n,\,3}\,\varphi^{ap_n+a+1}z^{a+1} &= H^{(2,\,1)}_{n,\,2}(z)(1-z) + \frac{H^{(2,\,1)}_{n,\,3}(z) + \chi_n^{(2,\,2)}(z)}{p_n}. \end{aligned}$$

Recalling that, as $n \to \infty$,

$$\varphi^{ap_n}\,\frac{\mathfrak{m}_{p_n}}{\mathfrak{m}_{(a+1)p_n}} \to (a+1)^\alpha,$$

we obtain that the $H^{(2,\,1)}_{n,\,j}$, $j \in \{0,\,1,\,2,\,3\}$ are Borel uniformly bounded functions on $[0,\,1]$, so that we can use Lemma 2.15 twice to obtain $\mathbb{E}\left|S^{(2,\,1)}_{n,\,1}\right|^3 = \mathrm{O}\left(n^{-3}\,\varphi^{3p_n}\,p_n^{-\alpha-3}\,L(p_n)\right)$. Finally, applying Theorem 2.7 finishes the proof of (2.16a).

Proofs of (2.16b) and (2.16c) are completely similar since $\zeta_n^{(2,\,2)}$ and $\zeta_n^{(2,\,3)}$ can be rewritten as

$$\zeta_n^{(2,\,2)} = \sum_{k=1}^{n} S_{n,\,k}^{(2,\,2)} \quad \text{with} \quad S_{n,\,k}^{(2,\,2)} = \frac{1}{n}\left[a_{n,\,0}^{(2,\,2)},\; a_{n,\,1}^{(2,\,2)}\right]\left[Y_k^{p_n},\; Y_k^{p_n+1}\right]^t,$$

$$\zeta_n^{(2,\,3)} = \sum_{k=1}^{n} S_{n,\,k}^{(2,\,3)} \quad \text{with} \quad S_{n,\,k}^{(2,\,3)} = \frac{1}{n}\left[a_{n,\,0}^{(2,\,3)},\; a_{n,\,1}^{(2,\,3)}\right]\left[Y_k^{(a+1)p_n},\; Y_k^{(a+1)p_n+a+1}\right]^t$$

where

$$\begin{aligned} a_{n,\,0}^{(2,\,2)} &= a_{n,\,0}^{(2,\,1)}, \\ a_{n,\,1}^{(2,\,2)} &= a_{n,\,1}^{(2,\,1)}, \\ a_{n,\,0}^{(2,\,3)} &= a_{n,\,2}^{(2,\,1)}, \\ a_{n,\,1}^{(2,\,3)} &= a_{n,\,3}^{(2,\,1)}. \end{aligned}$$

Applying Lemma 2.15 with

$$\begin{aligned} H_{n,\,0}^{(2,\,2)}(z) &:= H_{n,\,0}^{(2,\,1)}(z), \\ H_{n,\,1}^{(2,\,2)}(z) &:= H_{n,\,1}^{(2,\,1)}(z), \\ H_{n,\,0}^{(2,\,3)}(z) &:= H_{n,\,2}^{(2,\,1)}(z), \\ H_{n,\,1}^{(2,\,3)}(z) &:= H_{n,\,3}^{(2,\,1)}(z) \end{aligned}$$

yields $\mathbb{E}\left|S_{n,\,1}^{(2,\,j)}\right|^3 = \mathrm{O}\left(n^{-3}\,\varphi^{3p_n}\,p_n^{-\alpha-3}\,L(p_n)\right)$, $j \in \{2,\,3\}$. Using Theorem 2.7 then allows us to complete the proof of Theorem 2.16. ■

Recalling (2.15) makes it possible to derive the asymptotic normality of $\widetilde{\varphi}_n^a$ centered on φ^a:

Theorem 2.17. *Assume that* (B_1) *and* (B_2) *hold. If* $n\,p_n^{-\alpha}\,L(p_n) \to \infty$ *and* $n\,p_n^{-\alpha}\,L(p_n)\,\eta^2(p_n) \to 0$*, then*

$$v_n\left(\widetilde{\varphi}_n^a - \varphi^a\right) \xrightarrow{d} \mathcal{N}\left(0,\, \varphi^{2a}\,a^2\,V(\alpha,\,a)\right) \quad \textit{as } n \to \infty.$$

Proof. Notice that

$$\widetilde{\varphi}_n^a - \varphi^a = \left(\widetilde{\varphi}_n^a - \Phi_n^a\right) - \left(\varphi^a - \Phi_n^a\right).$$

Theorem 2.16 yields

$$v_n\left(\widetilde{\varphi}_n^a - \Phi_n^a\right) \xrightarrow{d} \mathcal{N}\left(0,\, \varphi^{2a}\,a^2\,V(\alpha,\,a)\right) \quad \text{as } n \to \infty.$$

Besides, (2.15) and the hypothesis $n\,p_n^{-\alpha}\,L(p_n)\,\eta^2(p_n) \to 0$ entail

$$v_n\left(\varphi^a - \Phi_n^a\right) \to 0$$

as $n \to \infty$. Using Slutsky's lemma completes the proof. ■

Since $\widetilde{\theta}_n = \ln \widetilde{\varphi}_n$, the delta method thus yields the asymptotic normality of $\widetilde{\theta}_n$:

Theorem 2.18. *Assume that* (B_1) *and* (B_2) *hold. If* $n\, p_n^{-\alpha}\, L(p_n) \to \infty$ *and* $n\, p_n^{-\alpha}\, L(p_n)\, \eta^2(p_n) \to 0$*, then*

$$v_n \left(\widetilde{\theta}_n - \theta \right) \xrightarrow{d} \mathcal{N}\left(0,\, V(\alpha,\, a)\right) \quad \text{as } n \to \infty.$$

It can be noted that, contrary to the estimator $\widehat{\theta}_n$ in the positive case, the asymptotic distribution of $\widetilde{\theta}_n - \theta$ does not depend on θ.

We finally state the analogue of Theorem 2.18 in the case of the Hall model. The proof is similar to that of Theorem 2.11 and is omitted.

Theorem 2.19. *In the case of the Hall model (3), assume that* $n\, p_n^{-\alpha} \to \infty$ *and* $n\, p_n^{-\alpha-2\beta} \to 0$ *as* $n \to \infty$*. Then*

$$w_n \left(\widetilde{\theta}_n - \theta \right) \xrightarrow{d} \mathcal{N}\left(0,\, \frac{V(\alpha,\, a)}{C}\right) \quad \text{as } n \to \infty$$

with the notations of Theorem 2.10.

2.4 Optimal convergence rates

We conclude the theoretical study of our estimators by comparing the convergence rate of the high order moments estimator to that of the extreme-value moment estimator of Aarssen and de Haan [1], which we shall consider in our numerical experiments below. We begin by considering the latter estimator:

1. If $\alpha \geq 2$, then the extreme-value moment estimator has the rate of convergence

$$w_n = \frac{\sqrt{k}}{a(n/k)}$$

provided $k = k(n) \to \infty$, $k/n \to 0$, and if, setting $U = (1/\overline{F})^{\leftarrow}$ where $\leftarrow$ denotes the left-continuous inverse, $\ln U$ satisfies the second order condition (see [57], p. 103)

$$\frac{\dfrac{\ln U(tx) - \ln U(t)}{a(t)/U(t)} - \dfrac{x^\gamma - 1}{\gamma}}{Q(t)} \to \frac{1}{\rho}\left[\frac{x^{\gamma+\rho} - 1}{\gamma + \rho} - \frac{x^\gamma - 1}{\gamma}\right]$$

for all $x > 0$ as $t \to \infty$, where $\gamma = -1/\alpha$ is the extreme-value index of $\overline{F}$, $\rho = -\nu\gamma$ is assumed to be different from γ, a is a regularly varying function with index γ, Q has a constant sign at infinity and converges to 0 at infinity: moreover, it is required that $\sqrt{k}\, Q(n/k) \to \lambda$, $\lambda \in \mathbb{R}$. Here, we shall assume that $\nu > -1$, so that Q is regularly varying with index ρ (see [57], p. 103). In that case, up to a slowly varying factor, the rate of convergence (w_n) is $n^{-\gamma}\, k^{\gamma+1/2}$. As a consequence, (w_n) increases as k increases; since the largest sequence k satisfying the condition $\sqrt{k}\, Q(n/k) \to \lambda$, $\lambda \in \mathbb{R}$ has order $n^{\rho/(\rho-1/2)}$, the rate of convergence (w_n) has order $n^{(1-\nu)/(\alpha-2\nu)}$.

2. If $\alpha < 2$, *i.e.* if the extreme-value index γ of $\overline{F}$ satisfies $\gamma < -1/2$, provided the second order condition for $\ln U$ and the requirements on k above hold, then the extreme-value moment estimator and the maximum estimator have the same rate of convergence $n^{1/\alpha}$ (see [57], Corollary 1.2.4 p. 21 and Remark 4.5.5 p. 147).

We now turn to the high order moments estimator: up to the slowly varying factor $\sqrt{L(p_n)}$, the rate of convergence is $v_n = \sqrt{n}\, p_n^{-\alpha/2+1}$, where (p_n) satisfies $n\, p_n^{-\alpha}\, L(p_n) \to \infty$ and $n\, p_n^{-\alpha}\, L(p_n)\, \eta^2(p_n) \to 0$. Once again, we shall consider the cases $\alpha \geq 2$ and $\alpha < 2$ separately:

1. If $\alpha \geq 2$, then the smaller the sequence (p_n) is, the higher the rate (v_n) is. The constraint on (p_n) is therefore the condition $n\, p_n^{-\alpha}\, L(p_n)\, \eta^2(p_n) \to 0$. Since $|\eta|$ is regularly varying with index ν, this condition is essentially $n\, p_n^{2\nu-\alpha} \to 0$: the smallest possible sequence (p_n) satisfying this requirement has order $n^{1/(\alpha-2\nu)}$. Consequently, (v_n) has order $n^{(1-\nu)/(\alpha-2\nu)}$.

2. If now $\alpha < 2$, then the rate (v_n) increases as (p_n) increases: the constraint on (p_n) is the condition $n\, p_n^{-\alpha}\, L(p_n) \to \infty$. The largest possible sequence (p_n) satisfying this condition has order $n^{1/\alpha}$, which yields a rate (v_n) with order $n^{1/\alpha}$.

Both estimators therefore essentially have the same optimal rates of convergence.

2.5 Examples

In this section, we highlight some cases where our hypotheses hold. Since $\eta(x) = x\, L'(x)/L(x)$, one can see that either $(A_1 - A_2)$ or $(B_1 - B_2)$ are satisfied in the general context of:

1. The Hall model (3), if δ is a bounded, twice continuously derivable function such that $\delta(x) \to 0$, $x\, \delta'(x) \to 0$ and $x^2\, \delta''(x) \to 0$ as $x \to \infty$. In that case, $\nu = -\beta < 0$.

2. The case $L(x) = f(\ln x)$, where f is a rational function with positive coefficients. Here, $\nu = 0$.

Let us now focus on four particular distributions that are also used for the numerical experiments of Section 2.6.

2.5.1 The positive case

We first give two examples in the case of positive random variables. Both of them have an endpoint $\theta = 1$; the first distribution has survival function

$$\overline{F}(x) = \left[1 + \left(\frac{1}{x} - 1\right)^{-\tau_1}\right]^{-\tau_2}, \quad x \in (0,\, 1), \tag{2.17}$$

with $\tau_1,\,\tau_2>0$. Remark that, if X is distributed from (2.17), then $X=1-1/(1+Y)$ where Y has a Burr$(1,\,\tau_1,\,\tau_2)$ type XII distribution as in Beirlant *et al.* [11]: namely, Y has survival function

$$\overline{G}(y)=(1+y^{\tau_1})^{-\tau_2}\,,\quad y\geq 0.$$

It is straightforward to check that (A_1) holds with $\alpha=\tau_1\,\tau_2$ and

$$\forall\, y\geq 1,\quad L(y)=\left[\frac{y^{\tau_1}}{1+(y-1)^{\tau_1}}\right]^{\tau_2}.$$

L is clearly infinitely derivable on $(1,\,\infty)$ and one readily obtains

$$\forall\, y>1,\quad \eta(y):=y\,\frac{L'(y)}{L(y)}=\tau_1\,\tau_2\,\frac{1-(y-1)^{\tau_1-1}}{1+(y-1)^{\tau_1}}.$$

As a result, η is bounded, continuously derivable on $(1,\,\infty)$, ultimately monotonic and non identically 0. Straightforward computations show that $|\eta'|$ is regularly varying, and

$$y\,\frac{\eta'(y)}{\eta(y)}=-y\left[(\tau_1-1)\frac{(y-1)^{\tau_1-2}}{1-(y-1)^{\tau_1-1}}+\tau_1\,\frac{(y-1)^{\tau_1-1}}{1+(y-1)^{\tau_1}}\right]\to-\min(\tau_1,\,1)<0,$$

as $y\to\infty$. Thus (A_2) holds with $\nu=-\min(\tau_1,\,1)$. Note that one can also show that L belongs to the Hall class (3).

The second considered distribution has survival function

$$\overline{F}(x)=\frac{1}{\Gamma(b)}\int_{-\ln(1-x)}^{\infty}(\lambda t)^{b-1}\,\lambda e^{-\lambda t}\,dt,\quad x\in(0,\,1),\tag{2.18}$$

with $b>1$ and $\lambda>0$. Here, when X is distributed from (2.18), it can be rewritten as $X=1-e^{-Y}$ where Y is Gamma$(b,\,\lambda)$ distributed: the survival function of Y is

$$\overline{G}(y)=\frac{1}{\Gamma(b)}\int_{y}^{\infty}(\lambda t)^{b-1}\,\lambda e^{-\lambda t}\,dt,\quad y\geq 0.$$

In this example, (A_1) holds with $\alpha=\lambda$, and for all $y>1$

$$L(y)=\frac{\lambda^{b-1}}{\Gamma(b)}[\ln y]^{b-1}[1+\delta(y)],\tag{2.19}$$

$$\delta(y)=\frac{1}{\lambda^{b-1}\,y^{-\lambda}\,[\ln y]^{b-1}}\left[\int_{\ln y}^{\infty}(\lambda t)^{b-1}\,\lambda e^{-\lambda t}\,dt\right]-1=(b-1)\int_{1}^{\infty}u^{b-2}\,e^{-\lambda(u-1)\ln y}\,du.\tag{2.20}$$

The function δ is infinitely derivable on $(1,\,\infty)$ and goes to 0 at ∞. Therefore, L is slowly varying and infinitely derivable on $(1,\,\infty)$. (2.19) and (2.20) together give

$$\eta(y):=y\,\frac{L'(y)}{L(y)}=\lambda\left[1-\frac{\lambda^{b-1}\,y^{-\lambda}\,[\ln y]^{b-1}}{\displaystyle\int_{\ln y}^{\infty}(\lambda t)^{b-1}\,\lambda e^{-\lambda t}\,dt}\right];$$

as a consequence, η is bounded in a neighborhood of 1. Besides, (2.19) and (2.20) yield

$$\begin{aligned}\eta(y)&=\frac{b-1}{\ln y}+\frac{y\,\delta'(y)}{1+\delta(y)}\\&=\frac{b-1}{\ln y}-\lambda(b-1)\int_{1}^{\infty}(u-1)\,u^{b-2}\,e^{-\lambda(u-1)\ln y}\,du\,(1+\mathrm{o}(1))\\&=\frac{b-1}{\ln y}+\mathrm{o}\left(\frac{1}{\ln y}\right),\end{aligned}$$

so that η is bounded and positive in a neighborhood of ∞. Finally, remarking that

$$\frac{d}{dy}\left[y\,\delta'(y)\right] = \frac{\lambda^2(b-1)}{y}\int_1^\infty (u-1)^2\,u^{b-2}\,e^{-\lambda(u-1)\ln y}\,du = \mathrm{o}\left(\frac{1}{y[\ln y]^2}\right)$$

it follows that

$$\eta'(y) = \frac{(1-b)}{y[\ln y]^2}(1+\mathrm{o}(1)),$$

entailing that η is ultimately nonincreasing, that $|\eta'|$ is regularly varying and that $y\,\eta'(y)/\eta(y) \to 0$ as $y \to \infty$. As a conclusion, (A_2) holds with $\nu = 0$.

2.5.2 The general case

Both of the next two examples we consider are negative distributions with endpoint $\theta = 0$. Their construction is an analogue of that of the examples in Section 2.5.1; the first one has survival function

$$\overline{F}(x) = \left[\frac{1}{1+(-x)^{-\tau_1}}\right]^{\tau_2}, \quad x < 0 \tag{2.21}$$

with τ_1, $\tau_2 > 0$, that is, $X = -1/Y$ where Y has a Burr(1, τ_1, τ_2) type XII distribution. It is easy to show that (B_1) holds, with $\alpha = \tau_1\,\tau_2$ and

$$\forall\, y > 0, \quad L(y) = \left[\frac{1}{1+y^{-\tau_1}}\right]^{\tau_2}.$$

L is infinitely derivable on $(0,\,\infty)$, and one has

$$\forall\, y > 0, \quad \eta(y) := y\,\frac{L'(y)}{L(y)} = \frac{\tau_1\,\tau_2}{1+y^{\tau_1}}.$$

As a consequence, η is bounded, continuously derivable on $(0,\,\infty)$, monotonic and non identically 0. Direct computations show that

$$\forall\, y > 0, \quad y\,\frac{\eta'(y)}{\eta(y)} = \frac{-\tau_1}{1+y^{-\tau_1}} \to -\tau_1 < 0,$$

so that $|\eta'|$ is regularly varying and therefore (B_2) is satisfied, with $\nu = -\tau_1$. It would again be easy to check that L belongs to the Hall class (3).

Our second example in this section has survival function

$$\overline{F}(x) = \frac{1}{\Gamma(b)}\int_{\ln(1-1/x)}^\infty (\lambda t)^{b-1}\,\lambda e^{-\lambda t}\,dt, \quad x < 0 \tag{2.22}$$

with $b > 1$ and $\lambda > 0$, which is tantamount to $X = -1/(e^Y - 1)$ where Y is Gamma(b, λ) distributed. Here, (B_1) is satisfied with $\alpha = \lambda$ and

$$\forall\, y > 0, \quad L(y) = \frac{\lambda^{b-1}}{\Gamma(b)}\left(1+y^{-1}\right)^{-\lambda}\left[\ln(y+1)\right]^{b-1}\left[1+\delta(y+1)\right]$$

with δ as in (2.20). It follows that (B_2) holds with $\nu = 0$.

2.6 Numerical experiments

In this section, we shall examine the performances of our estimators on samples with size $n = 500$ on sixteen situations obtained by considering the models (2.17), (2.18), (2.21) and (2.22) with four different sets of parameters, see the first columns of Tables 1–2. In the first eight examples, which are based on models (2.17) and (2.18), we consider the high order moments estimator $\widehat{\theta}_n$; in the last eight situations, based on Examples (2.21) and (2.22), the estimator $\widetilde{\theta}_n$ is chosen. We let $p_n = n^{1/\alpha}/\ln\ln n$ in order to satisfy the assumptions in Theorems 2.9 and 2.18, and a set $\mathcal{A} = \{0.1, 0.2, 0.3, \ldots, 25\}$ of different values of a is tested. In each of the sixteen situations, $N = 1000$ replications of the sample are generated, and we compare our estimators to the extreme-value moment estimator of [1] with a fixed threshold $k_1 = 1$, which depends on a parameter $k \in \{2, 3, \ldots, n-1\}$, and to the (naive) maximum estimator. Note that for Examples (2.21) and (2.22), since the extreme-value moment estimator can only be computed for random variables with a positive endpoint, the random variables $Y_i = e^{X_i}$ were used. For each estimator, depending on a parameter p, the procedure is the following: the average L^1-error is computed

$$E(p) = \frac{1}{N}\sum_{j=1}^{N} |\varepsilon(j, p)|, \quad \text{where} \quad \varepsilon(j, p) = \widehat{\theta}^{(j,p)} - \theta$$

with $\widehat{\theta}^{(j,p)}$ being the estimator computed on the jth replication with parameter p and θ being the endpoint of the chosen distribution. Then, the "optimal" value of p is retained: $p^\star = \operatorname{argmin}\{E(p), p \in \mathcal{P}\}$, where $\mathcal{P}$ is the set of tested values of p. Remark that, since the maximum estimator does not depend on any parameter, the associated function E is constant. Numerical results are summarised in Tables 1 and 2, where "optimal" errors are displayed.

In the upper part of Table 1, it appears that, for the distribution (2.17), all three estimators perform worse as $|\nu|$ decreases. This phenomenon can be explained since ν drives the bias of most extreme-value estimators. For instance, when $|\nu|$ is small, η converges slowly to 0 and Proposition 2.6 shows that the approximation error of μ_p/μ_{p+1} by M_p/M_{p+1} is large. The same type of behaviour is observed in Example (2.21), see the upper part of Table 2. Besides, the lower part of Table 1 shows that, for the distribution (2.18), when α increases, the average L^1-errors increase as well, since the simulated points are getting more and more distant from the endpoint. This phenomenon also happens for Example (2.22), see the lower part of Table 2. Let us highlight that, in all the considered situations, our estimators yield slightly better (optimal) results than the maximum estimator and the extreme-value moment estimator.

To further compare the behaviour of the estimators in the "optimal" case, boxplots of the associated errors $\varepsilon(j, p^*)$ are displayed on Figures 2.1–2.2 and Figures 2.7–2.8. Clearly, the maximum as well as our estimators underestimate the endpoint. However, the errors associated to our estimators are smaller than the error of the maximum. Besides, our estimators have a smaller variance than both the maximum estimator and the extreme-value moment estimator.

A graphical comparison on both models of the functions E associated to the four estimators is proposed on Figures 2.3–2.6 and Figures 2.9–2.12. On models (2.18) and (2.22), the shapes of the curves associated to our estimators and to the extreme-value moment estimator are similar, see Figures 2.5–2.6 and Figures 2.11–2.12. On the contrary, it appears on Figures 2.3–2.4 and Figures 2.9–2.10 that, on models (2.17) and (2.21), the functions E associated to the extreme-value moment estimator and our estimators have very different shapes, even though they have similar minima. The error associated to the extreme-value moment estimator is very sensitive to the choice of the parameter k whereas the error functions associated to our estimators are stable for a large panel of a values.

Distribution	Maximum	Extreme-value moment estimator	High order moments estimator $\widehat{\theta}_n$
$1 - \dfrac{1}{1+\mathrm{Burr}(1,\,\tau_1,\,\tau_2)}$			
$(\tau_1,\,\tau_2) = (1,\,1)$ $\Rightarrow (\alpha,\,\nu) = (1,\,-1)$	$2.0 \cdot 10^{-3}$	$1.8 \cdot 10^{-3}$	$1.7 \cdot 10^{-3}$
$(\tau_1,\,\tau_2) = (5/6,\,6/5)$ $\Rightarrow (\alpha,\,\nu) = (1,\,-5/6)$	$2.0 \cdot 10^{-3}$	$2.1 \cdot 10^{-3}$	$1.7 \cdot 10^{-3}$
$(\tau_1,\,\tau_2) = (2/3,\,3/2)$ $\Rightarrow (\alpha,\,\nu) = (1,\,-2/3)$	$2.1 \cdot 10^{-3}$	$2.0 \cdot 10^{-3}$	$1.8 \cdot 10^{-3}$
$(\tau_1,\,\tau_2) = (1/2,\,2)$ $\Rightarrow (\alpha,\,\nu) = (1,\,-1/2)$	$2.3 \cdot 10^{-3}$	$2.4 \cdot 10^{-3}$	$2.0 \cdot 10^{-3}$
$1 - \exp(-\mathrm{Gamma}(b,\,\lambda))$			
$(b,\,\lambda) = (2,\,1)$ $\Rightarrow (\alpha,\,\nu) = (1,\,0)$	$2.3 \cdot 10^{-4}$	$1.9 \cdot 10^{-4}$	$1.9 \cdot 10^{-4}$
$(b,\,\lambda) = (2,\,5/4)$ $\Rightarrow (\alpha,\,\nu) = (5/4,\,0)$	$1.1 \cdot 10^{-3}$	$9.2 \cdot 10^{-4}$	$8.4 \cdot 10^{-4}$
$(b,\,\lambda) = (2,\,5/3)$ $\Rightarrow (\alpha,\,\nu) = (5/3,\,0)$	$5.7 \cdot 10^{-3}$	$4.5 \cdot 10^{-3}$	$3.9 \cdot 10^{-3}$
$(b,\,\lambda) = (2,\,5/2)$ $\Rightarrow (\alpha,\,\nu) = (5/2,\,0)$	$3.0 \cdot 10^{-2}$	$2.1 \cdot 10^{-2}$	$1.8 \cdot 10^{-2}$

Table 2.1: Mean L^1-errors associated to the estimators in Examples (2.17) and (2.18).

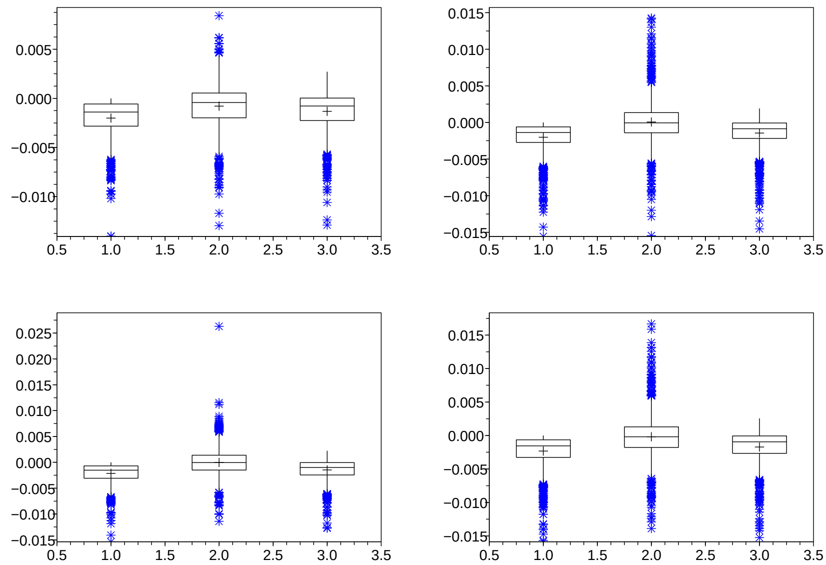


Figure 2.1: Boxplots of $\varepsilon(j,\, p^*)$ on model (2.17). Left: maximum estimator, middle: extreme-value moment estimator, right: high order moments estimator $\widehat{\theta}_n$. Top left: $(\tau_1,\, \tau_2) = (1,\, 1)$; top right: $(\tau_1,\, \tau_2) = (5/6,\, 6/5)$; bottom left: $(\tau_1,\, \tau_2) = (2/3,\, 3/2)$; bottom right: $(\tau_1,\, \tau_2) = (1/2,\, 2)$.

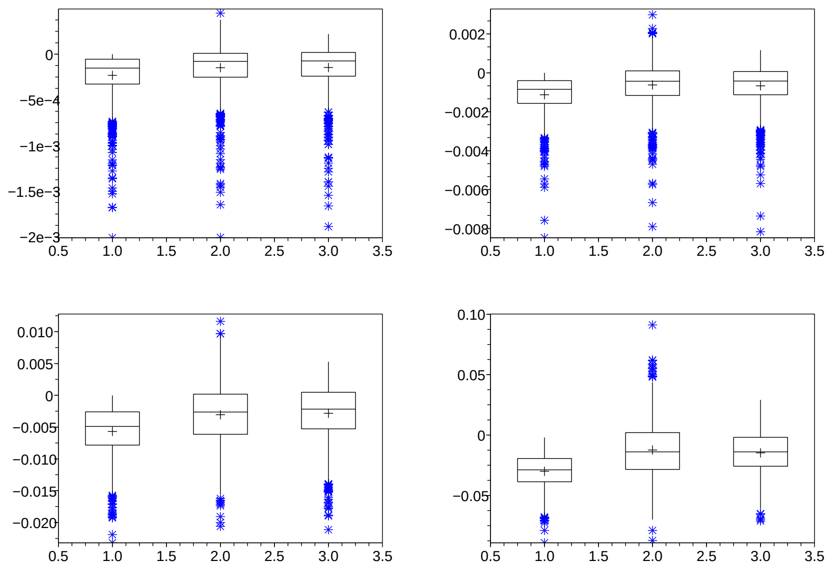


Figure 2.2: Boxplots of $\varepsilon(j,\, p^*)$ on model (2.18). Left: maximum estimator, middle: extreme-value moment estimator, right: high order moments estimator $\widehat{\theta}_n$. Top left: $(b,\, \lambda) = (2,\, 1)$; top right: $(b,\, \lambda) = (2,\, 5/4)$; bottom left: $(b,\, \lambda) = (2,\, 5/3)$; bottom right: $(b,\, \lambda) = (2,\, 5/2)$.

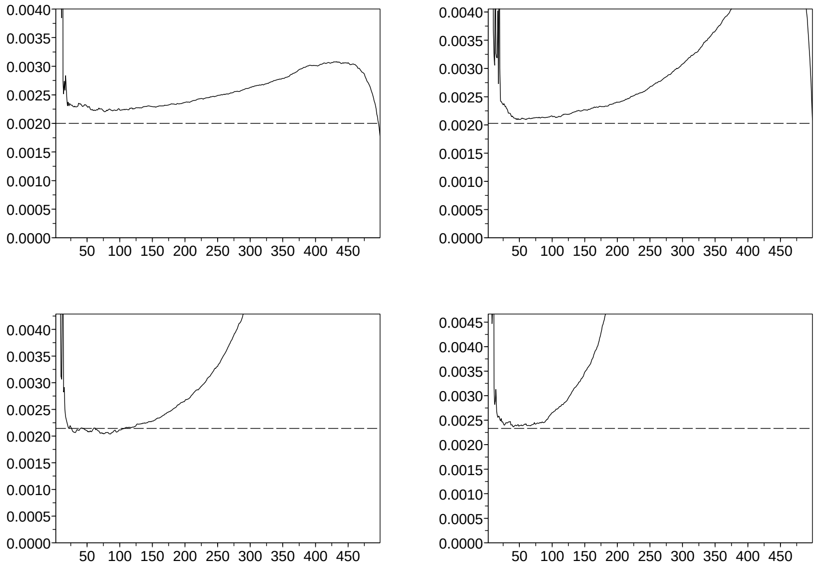


Figure 2.3: Comparison of the maximum and extreme-value moment estimators on model (2.17). Horizontally: parameter k, vertically: error E, dashed line: maximum estimator, solid line: extreme-value moment estimator. Top left: $(\tau_1, \tau_2) = (1, 1)$; top right: $(\tau_1, \tau_2) = (5/6, 6/5)$; bottom left: $(\tau_1, \tau_2) = (2/3, 3/2)$; bottom right: $(\tau_1, \tau_2) = (1/2, 2)$.

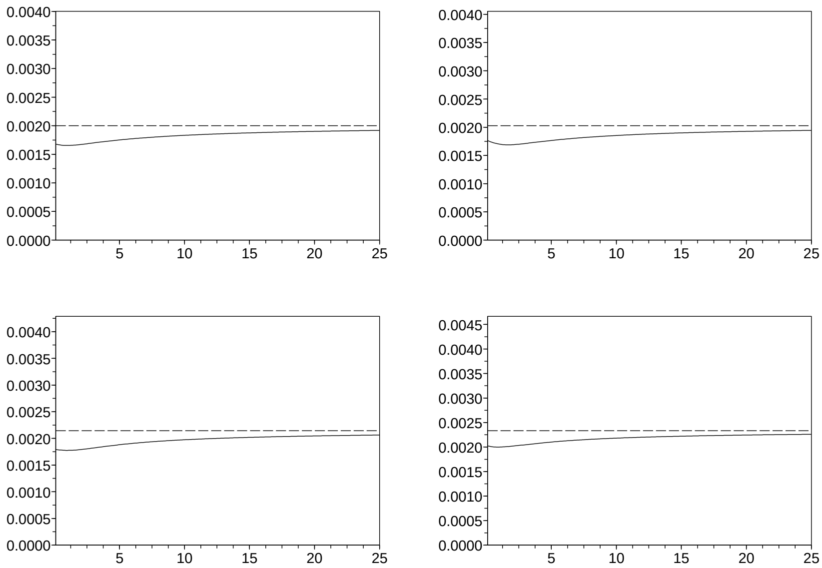


Figure 2.4: Comparison of the maximum estimator and high order moments estimator $\widehat{\theta}_n$ on model (2.17). Horizontally: parameter a, vertically: error E, dashed line: maximum estimator, solid line: high order moments estimator $\widehat{\theta}_n$. Top left: $(\tau_1, \tau_2) = (1, 1)$; top right: $(\tau_1, \tau_2) = (5/6, 6/5)$; bottom left: $(\tau_1, \tau_2) = (2/3, 3/2)$; bottom right: $(\tau_1, \tau_2) = (1/2, 2)$.

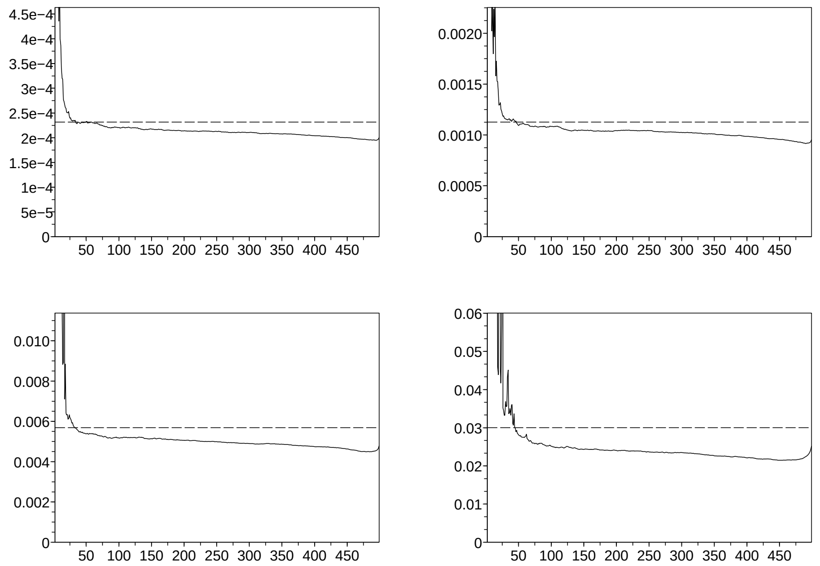


Figure 2.5: Comparison of the maximum and extreme-value moment estimators on model (2.18). Horizontally: parameter k, vertically: error E, dashed line: maximum estimator, solid line: extreme-value moment estimator. Top left: $(b, \lambda) = (2, 1)$, top right: $(b, \lambda) = (2, 5/4)$, bottom left: $(b, \lambda) = (2, 5/3)$, bottom right: $(b, \lambda) = (2, 5/2)$.

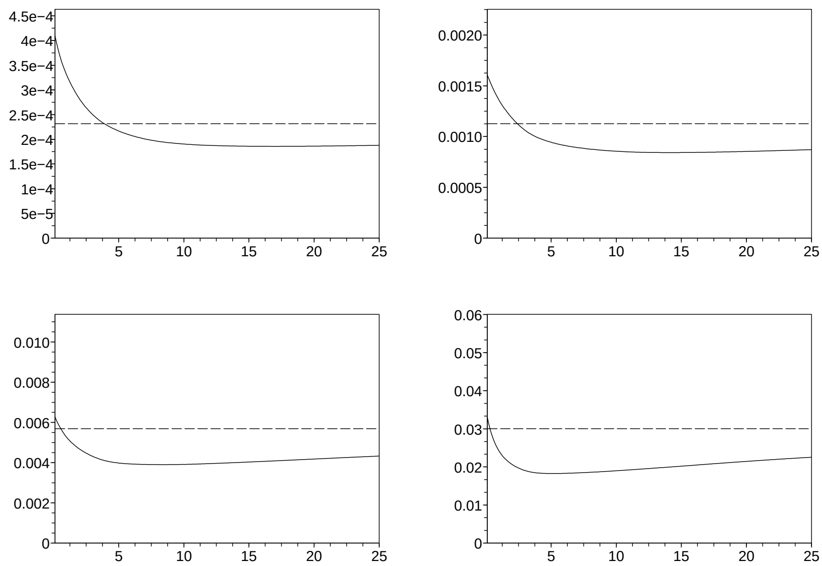


Figure 2.6: Comparison of the maximum estimator and high order moments estimator $\widehat{\theta}_n$ on model (2.18). Horizontally: parameter a, vertically: error E, dashed line: maximum estimator, solid line: high order moments estimator $\widehat{\theta}_n$. Top left: $(b, \lambda) = (2, 1)$, top right: $(b, \lambda) = (2, 5/4)$, bottom left: $(b, \lambda) = (2, 5/3)$, bottom right: $(b, \lambda) = (2, 5/2)$.

Distribution	Maximum	Extreme-value moment estimator	High order moments estimator $\widetilde{\theta}_n$
$-1/\text{Burr}(1,\, \tau_1,\, \tau_2)$			
$(\tau_1,\, \tau_2) = (1,\, 1)$ $\Rightarrow (\alpha,\, \nu) = (1,\, -1)$	$2.0 \cdot 10^{-3}$	$2.1 \cdot 10^{-3}$	$1.7 \cdot 10^{-3}$
$(\tau_1,\, \tau_2) = (5/6,\, 6/5)$ $\Rightarrow (\alpha,\, \nu) = (1,\, -5/6)$	$2.0 \cdot 10^{-3}$	$2.0 \cdot 10^{-3}$	$1.7 \cdot 10^{-3}$
$(\tau_1,\, \tau_2) = (2/3,\, 3/2)$ $\Rightarrow (\alpha,\, \nu) = (1,\, -2/3)$	$2.2 \cdot 10^{-3}$	$2.0 \cdot 10^{-3}$	$1.8 \cdot 10^{-3}$
$(\tau_1,\, \tau_2) = (1/2,\, 2)$ $\Rightarrow (\alpha,\, \nu) = (1,\, -1/2)$	$2.3 \cdot 10^{-3}$	$2.3 \cdot 10^{-3}$	$2.0 \cdot 10^{-3}$
$-1/(\exp(\text{Gamma}(b,\, \lambda)) - 1)$			
$(b,\, \lambda) = (2,\, 1)$ $\Rightarrow (\alpha,\, \nu) = (1,\, 0)$	$2.3 \cdot 10^{-4}$	$2.2 \cdot 10^{-4}$	$1.9 \cdot 10^{-4}$
$(b,\, \lambda) = (2,\, 5/4)$ $\Rightarrow (\alpha,\, \nu) = (5/4,\, 0)$	$1.1 \cdot 10^{-3}$	$9.2 \cdot 10^{-4}$	$8.5 \cdot 10^{-4}$
$(b,\, \lambda) = (2,\, 5/3)$ $\Rightarrow (\alpha,\, \nu) = (5/3,\, 0)$	$5.7 \cdot 10^{-3}$	$4.6 \cdot 10^{-3}$	$4.1 \cdot 10^{-3}$
$(b,\, \lambda) = (2,\, 5/2)$ $\Rightarrow (\alpha,\, \nu) = (5/2,\, 0)$	$3.1 \cdot 10^{-2}$	$2.4 \cdot 10^{-2}$	$2.3 \cdot 10^{-2}$

Table 2.2: Mean L^1-errors associated to the estimators in Examples (2.21) and (2.22).

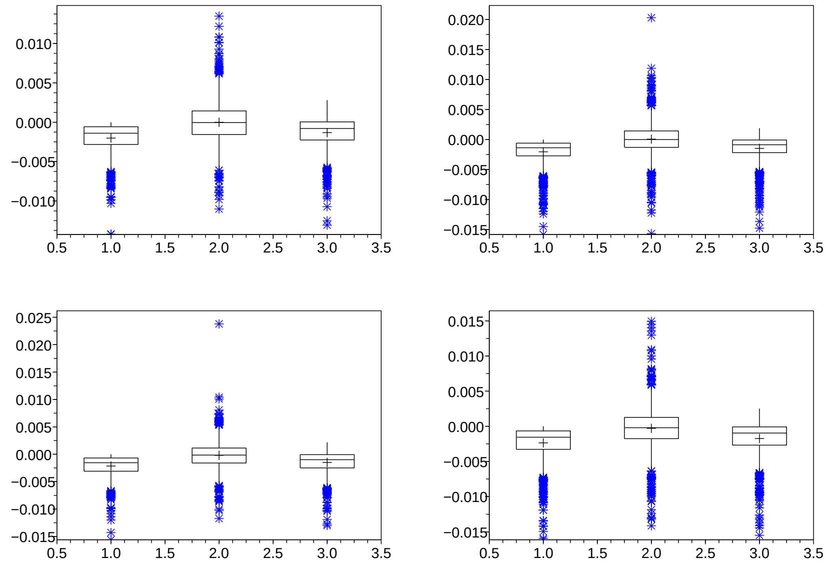


Figure 2.7: Boxplots of $\varepsilon(j,\, p^*)$ on model (2.21). Left: maximum estimator, middle: extreme-value moment estimator, right: high order moments estimator $\widetilde{\theta}_n$. Top left: $(\tau_1,\, \tau_2) = (1,\, 1)$; top right: $(\tau_1,\, \tau_2) = (5/6,\, 6/5)$; bottom left: $(\tau_1,\, \tau_2) = (2/3,\, 3/2)$; bottom right: $(\tau_1,\, \tau_2) = (1/2,\, 2)$.

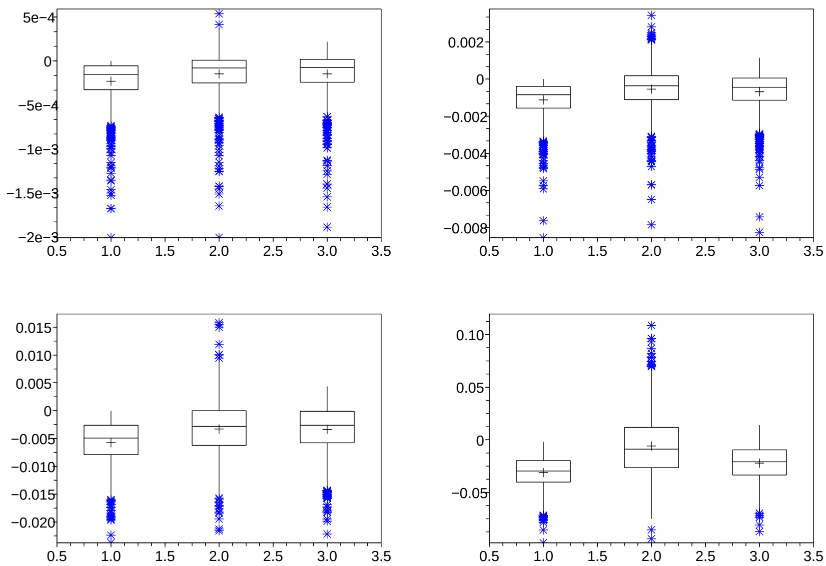


Figure 2.8: Boxplots of $\varepsilon(j,\, p^*)$ on model (2.22). Left: maximum estimator, middle: extreme-value moment estimator, right: high order moments estimator $\widetilde{\theta}_n$. Top left: $(b,\, \lambda) = (2,\, 1)$; top right: $(b,\, \lambda) = (2,\, 5/4)$; bottom left: $(b,\, \lambda) = (2,\, 5/3)$; bottom right: $(b,\, \lambda) = (2,\, 5/2)$.

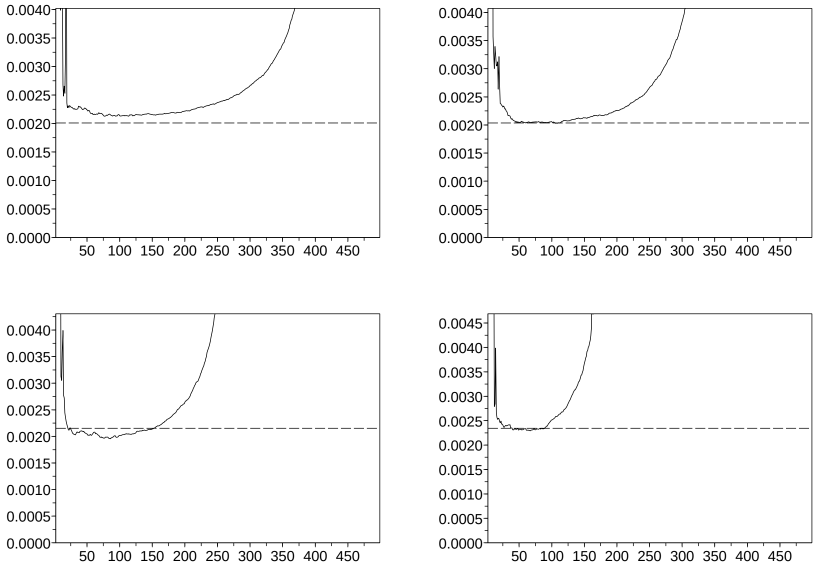


Figure 2.9: Comparison of the maximum and extreme-value moment estimators on model (2.21). Horizontally: threshold k, vertically: error E, dashed line: maximum estimator, solid line: extreme-value moment estimator. Top left: $(\tau_1, \tau_2) = (1, 1)$; top right: $(\tau_1, \tau_2) = (5/6, 6/5)$; bottom left: $(\tau_1, \tau_2) = (2/3, 3/2)$; bottom right: $(\tau_1, \tau_2) = (1/2, 2)$.

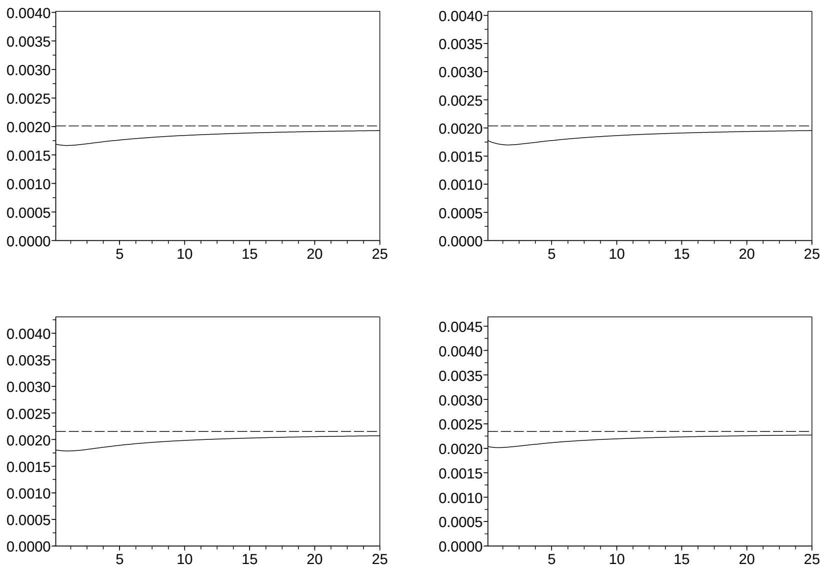


Figure 2.10: Comparison of the maximum estimator and high order moments estimator $\widetilde{\theta}_n$ on model (2.21). Horizontally: parameter a, vertically: error E, dashed line: maximum estimator, solid line: high order moments estimator $\widetilde{\theta}_n$. Top left: $(\tau_1, \tau_2) = (1, 1)$; top right: $(\tau_1, \tau_2) = (5/6, 6/5)$; bottom left: $(\tau_1, \tau_2) = (2/3, 3/2)$; bottom right: $(\tau_1, \tau_2) = (1/2, 2)$.

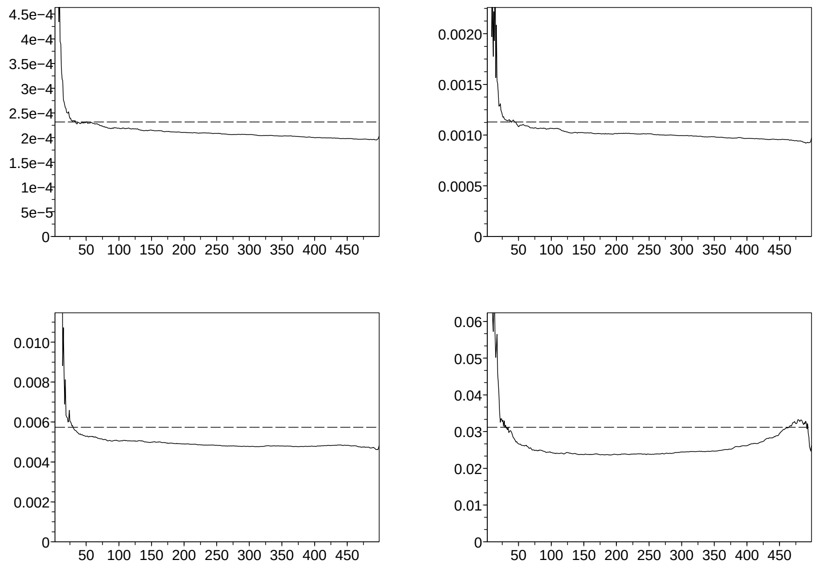


Figure 2.11: Comparison of the maximum and extreme-value moment estimators on model (2.22). Horizontally: threshold k, vertically: error E, dashed line: maximum estimator, solid line: extreme-value moment estimator. Top left: $(b, \lambda) = (2, 1)$; top right: $(b, \lambda) = (2, 5/4)$; bottom left: $(b, \lambda) = (2, 5/3)$; bottom right: $(b, \lambda) = (2, 5/2)$.

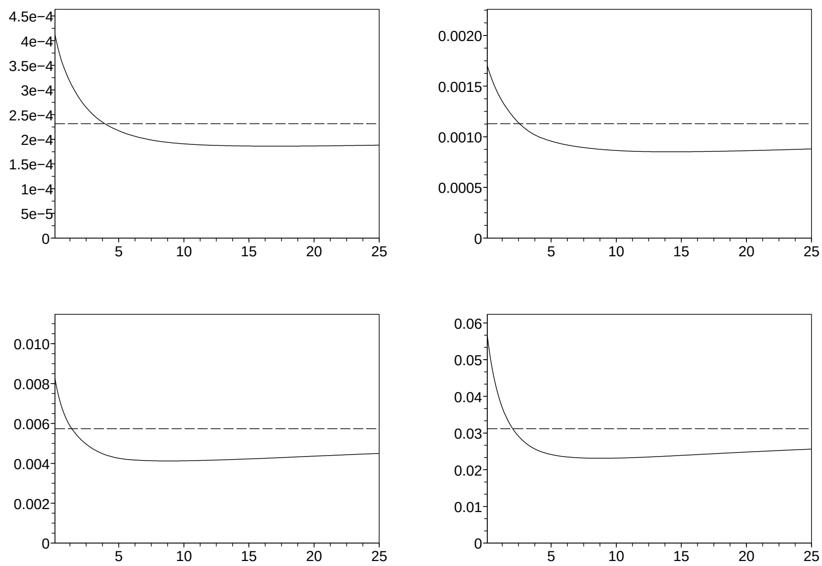


Figure 2.12: Comparison of the maximum estimator and high order moments estimator $\widetilde{\theta}_n$ on model (2.22). Horizontally: parameter a, vertically: error E, dashed line: maximum estimator, solid line: high order moments estimator $\widetilde{\theta}_n$. Top left: $(b, \lambda) = (2, 1)$; top right: $(b, \lambda) = (2, 5/4)$; bottom left: $(b, \lambda) = (2, 5/3)$; bottom right: $(b, \lambda) = (2, 5/2)$.

2.7 Appendix A: Auxiliary results

The first lemma is a technical result we shall need several times in the proofs.

Lemma 2.1. *Let f be a positive bounded Borel function on $(0,\,1)$. Then*

$$\forall\,\delta\in(0,\,1),\qquad \int_0^1 x^p\,f(x)\,dx=\int_{1-\delta}^1 x^p\,f(x)\,dx\,(1+\mathrm{o}(1))\quad \text{as } p\to\infty.$$

The next result deals with the behaviour of the high order moment $\mathbb{E}(Z^p)$, $p\to\infty$, when Z is a positive random variable with right endpoint equal to 1.

Lemma 2.2. *Let Z be a positive random variable with right endpoint equal to 1. Then for all $u\geq 1$, $\mathbb{E}(Z^p)/\mathbb{E}(Z^{p+u})\to 1$ as $p\to\infty$.*

Before making further progress, we state a classical uniform convergence result in the regular variation framework.

Lemma 2.3. *Pick $\alpha\neq 0$ and let $f\in\mathrm{RV}_\alpha$.*

(i) If $\alpha<0$, then for all $b>0$, $\displaystyle\sup_{x\in[b,\,\infty)}\left|\frac{f(tx)}{f(t)}-x^\alpha\right|\to 0$ *as $t\to\infty$.*

(ii) If $\alpha>0$, then for all $b>0$, $\displaystyle\sup_{x\in(0,\,b]}\left|\frac{f(tx)}{f(t)}-x^\alpha\right|\to 0$ *as $t\to\infty$.*

The next lemma draws some consequences from the property (H_b) on the function η.

Lemma 2.4. *Pick $b\in[0,\,1]$ and let η be a continuously derivable function on $(b,\,\infty)$ such that $|\eta'|$ is regularly varying with index $\nu-1$, where $\nu\leq 0$, and $x\,\eta'(x)\to 0$ as $x\to\infty$. Then, for all $u\geq 1$,*

(i) $\displaystyle t\sup_{x\geq 1}|\eta(tx)-\eta((t+u)x)|\to 0$ *as $t\to\infty$.*

(ii) For all $q>-\nu$, $\displaystyle t\sup_{x\in(0,\,1]}x^q\,|\eta(tx+b)-\eta((t+u)x+b)|\to 0$ *as $t\to\infty$.*

Before proceeding, let us introduce some more notations. For all $k\in\mathbb{R}$, let P_k be the set of collections of Borel functions (f_p) on $(0,\,1]$ such that

- $\exists\,p_k\geq 1,\quad \exists\,C_k\geq 0,\quad \forall\,p\geq p_k,\quad \forall\,x\in(0,\,1],\quad |f_p(x)|\leq C_k\,x^k,$
- $\forall\,u\geq 1,\quad \exists\,p_k\geq 1,\quad \exists\,C_k\geq 0,\quad \forall\,p\geq p_k,\quad \forall\,x\in(0,\,1],\quad p^2|f_{p+u}-f_p|(x)\leq C_k\,x^k,$
- $\forall\,u,\,v\geq 1,\quad \forall\,x\in(0,\,1],\quad p^2|f_{p+u+v}-f_{p+v}-[f_{p+u}-f_p]|(x)\to 0$ as $p\to\infty$.

Let $P=\displaystyle\bigcap_{k\geq 0}P_k$. Besides, let U be the set of collections of Borel functions (f_p) on $[1,\,\infty)$ such that

- $\displaystyle\sup_{x\geq 1}|f_p(x)|=\mathrm{O}(1)$ as $p\to\infty$,
- $\forall\,u\geq 1,\quad \displaystyle p^2\sup_{x\geq 1}|f_{p+u}-f_p|(x)=\mathrm{O}(1)$ as $p\to\infty$,

- $\forall\, u,\, v \geq 1, \quad p^2 \sup_{x\geq 1} |f_{p+u+v} - f_{p+v} - [f_{p+u} - f_p]|(x) \to 0$ as $p \to \infty$.

These sets will reveal useful to study the asymptotic properties of our estimators since they are based on increments of sequences of functions. A stability property of the set P is given in the next lemma.

Lemma 2.5. *Let (f_p), (g_p) be two collections of Borel functions. If for some $k \in \mathbb{R}$, $(f_p) \in P_k$ and $(g_p) \in P$, then $(f_p\, g_p) \in P$.*

We now give a continuity property of some integral transforms defined on P and U.

Lemma 2.6. *Let $(f_p) \in P$, $(g_p) \in U$ and (u_p), (v_p) be two collections of Borel functions such that there exist four Borel functions f, g, u, v, with f and u (resp. g and v) being defined on $(0, 1]$ (resp. $[1, \infty)$), $f_p(x) \to f(x)$ for all $x \in (0, 1]$,*

$$\sup_{x\geq 1} |g_p(x) - g(x)| \to 0, \quad \sup_{0<x\leq 1} |u_p(x) - u(x)| \to 0 \text{ and } \sup_{x\geq 1} |v_p(x) - v(x)| \to 0 \text{ as } p \to \infty.$$

Assume further that u and v are bounded. Then, for all $k > 1$,

$$\begin{aligned} \int_0^1 x^{-k}\, f_p(x)\, u_p(x)\, dx &\to \int_0^1 x^{-k}\, f(x)\, u(x)\, dx, \\ \int_1^\infty x^{-k}\, g_p(x)\, v_p(x)\, dx &\to \int_1^\infty x^{-k}\, g(x)\, v(x)\, dx \end{aligned}$$

as $p \to \infty$.

The following lemma provides sufficient conditions on collections of functions to belong to the previous sets.

Lemma 2.7. *Let (f_p), (g_p) be two collections of Borel functions. Assume that there exist Borel functions F_i and bounded Borel functions G_i, $0 \leq i \leq 2$, such that*

$$\forall\, x \in (0, 1], \quad p^2 \left| f_p(x) - \sum_{k=0}^{2} p^{-k} F_k(x) \right| \to 0 \text{ as } p \to \infty,$$

$$p^2 \sup_{x\geq 1} \left| g_p(x) - \sum_{k=0}^{2} p^{-k} G_k(x) \right| \to 0 \text{ as } p \to \infty.$$

Then, for all $x \in (0, 1]$ and $u, v \geq 1$, $p^2\, |f_{p+u+v} - f_{p+v} - [f_{p+u} - f_p]|(x) \to 0$ as $p \to \infty$, and $(g_p) \in U$.

We are now in position to exhibit two particular elements of P and U:

Lemma 2.8. *Let (f_p) and (g_p), $p \geq 1$ be two collections of Borel functions defined by*

$$\begin{aligned} \forall\, x \in (0, 1], \quad f_p(x) &= \left(1 - \frac{1}{p}\right)^{-\alpha-1} \left(1 + \frac{1}{(p-1)x}\right)^{-\alpha-2} \left(1 - \frac{1}{(p-1)x+1}\right)^{p-1}, \\ \forall\, x \in [1, \infty), \quad g_p(x) &= \left(1 - \frac{1}{px}\right)^{p-1}. \end{aligned}$$

Then $(f_p) \in P$, $(g_p) \in U$ *and*

$$\forall\, x \in (0,\, 1], \quad f_p(x) \to e^{-1/x} \text{ and } \sup_{x \geq 1} \left| g_p(x) - e^{-1/x} \right| \to 0 \text{ as } p \to \infty. \tag{2.23}$$

Lemmas 2.9, 2.10 and 2.11 are the key tools for establishing precise expansions of the moments M_p, μ_p and $\mathfrak{m}_p$.

Lemma 2.9. *Let* $(f_p) \in P$ *and* $(g_p) \in U$ *such that (2.23) holds and define*

$$\begin{aligned} E_1(p) &= \frac{1}{I_1} \int_0^1 f_p(x)\, x^{-\alpha-2}\, dx - 1, & I_1 &= \int_1^\infty y^\alpha\, e^{-y}\, dy, \\ E_2(p) &= \frac{1}{I_2} \int_1^\infty g_p(x)\, x^{-\alpha-2}\, dx - 1, & I_2 &= \int_0^1 y^\alpha\, e^{-y}\, dy. \end{aligned}$$

Then, for all $i \in \{1,\, 2\}$ *and all* $u,\, v \geq 1$,

(i) $E_i(p) \to 0$ *as* $p \to \infty$,

(ii) $p^2(E_i(p+u) - E_i(p)) = \mathrm{O}(1)$,

(iii) $p^2(E_i(p+u+v) - E_i(p+v) - [E_i(p+u) - E_i(p)]) \to 0$ *as* $p \to \infty$.

Lemma 2.10. *Pick* $b \in [0,\, 1]$. *Let* $(f_p) \in P$ *such that* $f_p(x) \to e^{-1/x}$ *as* $p \to \infty$ *for all* $x \in (0,\, 1]$ *and define*

$$\delta_1(p) = \frac{1}{I_1} \int_0^1 f_p(x) \left[\frac{\ell_1((p-b)x)}{\ell_1(p-b)} - x \right] x^{-\alpha-3}\, dx, \quad \ell_1(x) = x\, \ell(x+b),$$

where ℓ *is a slowly varying function. Then*

(i) $\delta_1(p) \to 0$ *as* $p \to \infty$.

Moreover, if ℓ *satisfies* (H_b), *then for all* $u,\, v \geq 1$,

(ii) There exists a slowly varying function $\mathcal{L}$ *such that* $\delta_1(p) = \mathrm{O}\left(|\eta(p)|\, \mathcal{L}(p)\right)$,

(iii) $\delta_1(p+u) - \delta_1(p) = \mathrm{O}\left(\frac{|\eta(p)|}{p} \right)$,

(iv) $p^2(\delta_1(p+u+v) - \delta_1(p+v) - [\delta_1(p+u) - \delta_1(p)]) \to 0$ *as* $p \to \infty$.

Lemma 2.11. *Let* $(g_p) \in U$ *such that* $\sup_{x \geq 1} \left| g_p(x) - e^{-1/x} \right| \to 0$ *as* $p \to \infty$ *and define*

$$\delta_2(p) = \frac{1}{I_2} \int_1^\infty g_p(x) \left[\frac{\ell_2(px)}{\ell_2(p)} - \frac{1}{x} \right] x^{-\alpha-1}\, dx, \quad \ell_2(x) = \frac{\ell(x)}{x},$$

where ℓ *is a slowly varying function. Then*

(i) $\delta_2(p) \to 0$ *as* $p \to \infty$.

Moreover, if ℓ *satisfies* (H_b) *for some* $b \in [0,\, 1]$, *then for all* $u,\, v \geq 1$, *(ii)* $\delta_2(p) = \mathrm{O}\left(|\eta(p)|\right)$,

(iii) $\delta_2(p+u) - \delta_2(p) = \mathrm{O}\left(\dfrac{|\eta(p)|}{p}\right)$,

(iv) $p^2(\delta_2(p+u+v) - \delta_2(p+v) - [\delta_2(p+u) - \delta_2(p)]) \to 0$ *as* $p \to \infty$.

The following result gives an equivalent of the high order moment $\mathbb{E}(Z^p)$, $p \to \infty$ when Z is a positive random variable whose cumulative distribution function satisfies a parametric requirement.

Lemma 2.12. *Let Z be a random variable with survival function $z \mapsto f \circ g(z)$, $z \in (0,\,1)$, where f is regularly varying with index $-\alpha < 0$, say $f(x) = x^{-\alpha}\,L(x)$ where L is a slowly varying function, and $(1-x)g(x) \to 1$, $x \to 1$. Then, as $p \to \infty$,*

$$\mathbb{E}(Z^p) = p^{-\alpha}\,L(p)\,\Gamma(\alpha+1)\,(1+\mathrm{o}(1)).$$

The lemma below gives linearisations of the sequences of random variables $\left(\xi_n^{(1)}\right)$ and $\left(\xi_n^{(2)}\right)$ appearing in the proofs of Theorems 2.8 and 2.16:

Lemma 2.13. *Let (p_n) be a positive sequence tending to ∞.*

(i) Let $\nu_p^{(1)} = \widehat{\mu}_p - \mu_p$. If (A_1) is satisfied, then

$$\xi_n^{(1)} = u_{n,\,a}^{(1)}\left[\zeta_n^{(1,\,1)} + \left(\frac{\mu_{p_n+1}}{\widehat{\mu}_{p_n+1}} - 1\right)\zeta_n^{(1,\,2)} + \left(1 + \frac{a p_n}{p_n+1}\right)\left(\frac{\mu_{(a+1)p_n+1}}{\widehat{\mu}_{(a+1)p_n+1}} - 1\right)\zeta_n^{(1,\,3)}\right](1+\mathrm{o}(1)),$$

where

$$\begin{aligned} \zeta_n^{(1,\,1)} &= \zeta_n^{(1,\,2)} + \left[1 + \frac{a p_n}{p_n+1}\right]\zeta_n^{(1,\,3)}, \\ \textit{with}\quad \zeta_n^{(1,\,2)} &= -\nu_{p_n}^{(1)} + \frac{\mu_{p_n}}{\mu_{p_n+1}}\,\nu_{p_n+1}^{(1)}, \\ \zeta_n^{(1,\,3)} &= \frac{\mu_{p_n+1}}{\mu_{(a+1)p_n+1}}\left[\nu_{(a+1)p_n}^{(1)} - \frac{\mu_{(a+1)p_n}}{\mu_{(a+1)p_n+1}}\,\nu_{(a+1)p_n+1}^{(1)}\right] \\ \textit{and}\quad u_{n,\,a}^{(1)} &= \frac{1}{a\,\Gamma(\alpha+1)}\sqrt{\frac{1}{V(\alpha,\,a)}}\,\frac{p_n^{\alpha}\,v_n}{\theta^{p_n}\,L(p_n)}. \end{aligned}$$

(ii) Let $\nu_p^{(2)} = \widetilde{\mathfrak{m}}_p - \mathfrak{m}_p$. If (B_1) is satisfied, then

$$\begin{aligned} \xi_n^{(2)} &= u_{n,\,a}^{(2)}\left[\zeta_n^{(2,\,1)} + \left(\frac{1}{\varphi^{a+1}}\,\frac{\widetilde{\mathfrak{m}}_{(a+1)p_n+a+1}}{\widetilde{\mathfrak{m}}_{(a+1)p_n}}\,\frac{\mathfrak{m}_{p_n+1}}{\widetilde{\mathfrak{m}}_{p_n+1}} - 1\right)\zeta_n^{(2,\,2)} + \left(\frac{\mathfrak{m}_{(a+1)p_n}}{\widetilde{\mathfrak{m}}_{(a+1)p_n}} - 1\right)\zeta_n^{(2,\,3)}\right] \\ &\times\ (1+\mathrm{o}(1)) \end{aligned}$$

where

$$\begin{aligned} \zeta_n^{(2,\,1)} &= \zeta_n^{(2,\,2)} + \zeta_n^{(2,\,3)}, \\ \textit{with}\quad \zeta_n^{(2,\,2)} &= \nu_{p_n}^{(2)} - \frac{\mathfrak{m}_{p_n}}{\mathfrak{m}_{p_n+1}}\,\nu_{p_n+1}^{(2)}, \\ \zeta_n^{(2,\,3)} &= \frac{1}{\varphi^{a+1}}\,\frac{\mathfrak{m}_{p_n}\,\mathfrak{m}_{(a+1)p_n+a+1}}{\mathfrak{m}_{(a+1)p_n}^2}\left[-\nu_{(a+1)p_n}^{(2)} + \frac{\mathfrak{m}_{(a+1)p_n}}{\mathfrak{m}_{(a+1)p_n+a+1}}\,\nu_{(a+1)p_n+a+1}^{(2)}\right] \\ \textit{and}\quad u_{n,\,a}^{(2)} &= \frac{1}{a\,\Gamma(\alpha+1)}\sqrt{\frac{1}{V(\alpha,\,a)}}\,\frac{p_n^{\alpha}\,v_n}{\varphi^{p_n}\,L(p_n)}. \end{aligned}$$

The next result is a key element of the proofs of Theorems 2.8 and 2.16.

Lemma 2.14. *Let Z be a bounded random variable such that*

$$\forall\, p \geq 1, \quad \mathbb{E}(Z^p) = p^{-\alpha}\, L(p)\, \Gamma(\alpha+1)\, [1+\varepsilon(p)]$$

where $\alpha > 0$, L is a slowly varying function satisfying condition (H_b) for some $b \in [0,\, 1]$, and ε is a function tending to 0 at ∞ such that for all $u,\, v \geq 1$

- *$p(\varepsilon(p+u) - \varepsilon(p)) \to 0$ as $p \to \infty$;*
- *$p^2(\varepsilon(p+u+v) - \varepsilon(p+v) - [\varepsilon(p+u) - \varepsilon(p)]) \to 0$ as $p \to \infty$.*

Let

$$w(s,\, t,\, u,\, v,\, p) = \left[-1,\ \frac{\mathbb{E}(Z^{sp})}{\mathbb{E}(Z^{sp+u})}\right] \begin{bmatrix} \mathbb{E}\left(Z^{(s+t)p}\right) & \mathbb{E}\left(Z^{(s+t)p+v}\right) \\ \mathbb{E}\left(Z^{(s+t)p+u}\right) & \mathbb{E}\left(Z^{(s+t)p+u+v}\right) \end{bmatrix} \left[-1,\ \frac{\mathbb{E}(Z^{tp})}{\mathbb{E}(Z^{tp+v})}\right]^t.$$

Then, as $p \to \infty$,

$$w(s,\, t,\, u,\, v,\, p) = \frac{\Gamma(\alpha+1)\, \alpha(\alpha+1)\, uv}{(s+t)^{\alpha+2}}\, p^{-\alpha-2}\, L(p)\, (1+\mathrm{o}(1)).$$

The next lemma of this section provides an asymptotic bound of the third-order moments appearing in the proofs of Theorems 2.8 and 2.16.

Lemma 2.15. *Let $m \in \mathbb{N}$, $(H_{n,\,j})$, $0 \leq j \leq m$ be sequences of Borel uniformly bounded functions on $(0,\, 1)$ and (p_n) be a positive sequence tending to ∞. Write further*

$$\forall\, z \in (0,\, 1), \quad h_n(z) = \sum_{j=0}^{m} \frac{H_{n,\,j}(z)}{p_n^j} (1-z)^{m-j}.$$

Next, let Z be a random variable with survival function $z \mapsto f \circ g(z)$, $z \in (0,\, 1)$, where f is regularly varying with index $-\alpha < 0$, say $f(x) = x^{-\alpha}\, L(x)$ where L is a slowly varying function, and $(1-x)g(x) \to 1$, $x \to 1$. Then

$$\mathbb{E}|Z^{p_n}\, h_n(Z)|^3 = \mathrm{O}\left(p_n^{-\alpha-3m}\, L(p_n)\right).$$

The final two lemmas are dedicated to the study of the particular case of the Hall model. Before stating these results, let us introduce some additional tools: let $\mathcal{C}_2 \subset \mathcal{F}(\mathbb{N} \times [0,\, \infty),\, \mathbb{R})$ be the set of all sequences of functions $u : \mathbb{N} \times [0,\, \infty) \to \mathbb{R}$, denoted by $u_n(x)$, such that u meets the following requirements:

(Q_1) For all $x \in [0,\, \infty)$, $(u_n(x))$ is a bounded sequence.

(Q_2) For all $0 \leq x < y$, $p_n[u_n(y) - u_n(x)] \to 0$ as $n \to \infty$.

(Q_3) For all $0 \leq x < y$ and $z > 0$, $p_n^2[u_n(y+z) - u_n(y) - \{u_n(x+z) - u_n(x)\}] \to 0$ as $n \to \infty$.

Set now

$$\begin{aligned}\mathcal{D} &= \{u \in \mathcal{F}(\mathbb{N}\times[0,\,\infty),\,\mathbb{R})\,|\,\forall\, x\in[0,\,\infty), \quad u_n(x)\to 0 \text{ as } n\to\infty\},\\ \mathcal{E} &= \{u \in \mathcal{F}(\mathbb{N}\times[0,\,\infty),\,\mathbb{R})\,|\,\forall\, x\in[0,\,\infty), \quad \exists\, M(x)>0, \quad \forall\, n\in\mathbb{N}, \quad u_n(x)\geq M(x)\},\end{aligned}$$

and $\mathcal{D}_2 = \mathcal{D}\cap\mathcal{C}_2$, $\mathcal{E}_2 = \mathcal{E}\cap\mathcal{C}_2$. Lemma 2.16 gives some properties of the sets $\mathcal{C}_2$, $\mathcal{D}_2$ and $\mathcal{E}_2$.

Lemma 2.16. *The sets $\mathcal{C}_2$, $\mathcal{D}_2$ and $\mathcal{E}_2$ have the following properties:*

(i) *$\mathcal{C}_2$ and $\mathcal{D}_2$ are sub-algebras of $\mathcal{F}(\mathbb{N}\times[0,\,\infty),\,\mathbb{R})$.*

(ii) *$\mathcal{E}_2$ is closed under division.*

(iii) *Let $u\in\mathcal{F}(\mathbb{N}\times[0,\,\infty),\,\mathbb{R})$ such that there exist $q\in\mathbb{N}\setminus\{0\}$, $\alpha_1<\ldots<\alpha_q\in[0,\,2]$ and bounded real sequences $(\delta_{n,\,1}),\,\ldots,\,(\delta_{n,\,q})$ with*

$$\forall\, x\in[0,\,\infty), \quad u_n(x) = 1+\sum_{j=1}^{q}\frac{\delta_{n,\,j}}{(p_n+x)^{\alpha_j}} + \mathrm{o}\left(\frac{1}{p_n^2}\right).$$

Then $u\in\mathcal{E}_2$.

The last lemma is an essential element of the proof of Theorem 2.10.

Lemma 2.17. *Let $\alpha>0$ and δ be a bounded Borel function on $[0,\,1]$, going to 0 at 1. Set*

$$q(p_n) = \frac{\displaystyle\int_0^1 x^{p_n}\,(1-x)^{\alpha}\,\delta(x)\,dx}{B(p_n+1,\,\alpha+1)}.$$

Then $(n,\,x)\mapsto q(p_n+x)\in\mathcal{D}_2$.

2.8 Appendix B: Proofs

Proof of Lemma 2.1. Pick $\delta>0$: we write

$$\int_0^1 x^p f(x)\,dx = \int_{1-\delta}^1 x^p\,f(x)\,dx\left[1+\frac{\displaystyle\int_0^{1-\delta} x^p\,f(x)\,dx}{\displaystyle\int_{1-\delta}^1 x^p\,f(x)\,dx}\right].$$

Notice then that

$$0\leq\frac{\displaystyle\int_0^{1-\delta} x^p\,f(x)\,dx}{\displaystyle\int_{1-\delta}^1 x^p\,f(x)\,dx}\leq\frac{(1-\delta)\,\sup\limits_{(0,\,1)} f}{\displaystyle\int_{1-\delta}^1\left[\frac{x}{1-\delta}\right]^p f(x)\,dx}\leq\frac{(1-\delta)\,\sup\limits_{(0,\,1)} f}{\displaystyle\left[\frac{1-\delta/2}{1-\delta}\right]^p\int_{1-\delta/2}^1 f(x)\,dx}.$$

Since $\left[\dfrac{1-\delta/2}{1-\delta}\right]^p\to\infty$ as $p\to\infty$, we get the desired result. ■

Proof of Lemma 2.2. Let $\overline{G}$ be the survival function of Z and $I_p = \mathbb{E}(Z^p)/p$. Lemma 2.1 yields

$$\forall\, \varepsilon > 0, \quad I_p = \int_{1-\varepsilon}^{1} y^{p-1}\,\overline{G}(y)\,dy\,(1 + \mathrm{o}(1)). \tag{2.24}$$

In view of

$$\forall\, \varepsilon > 0, \quad 1 \leq \frac{\displaystyle\int_{1-\varepsilon}^{1} y^{p-1}\,\overline{G}(y)\,dy}{\displaystyle\int_{1-\varepsilon}^{1} y^{p+u-1}\,\overline{G}(y)\,dy} \leq \frac{1}{(1-\varepsilon)^u}$$

and (2.24), one thus has $I_p/I_{p+u} \to 1$ as $p \to \infty$ and Lemma 2.2 is proven. ■

Proof of Lemma 2.3. See [14], Theorem 1.5.2 p. 22. ■

Proof of Lemma 2.4. Let us consider (i) and (ii) separately.

(i) Let $t,\, x \geq 1$. The mean value theorem shows that there exists $h_1(t,\, x) \in (0,\, u)$ such that

$$\begin{aligned} t\,|\eta(tx) - \eta((t+u)x)| &= u\,\frac{t}{t + h_1(t,\, x)}\,|[(t + h_1(t,\, x))x]\;\eta'\,[(t + h_1(t,\, x))x]| \\ &\leq u\,\sup_{y \geq t} |y\,\eta'(y)| \to 0 \end{aligned}$$

uniformly in $x \geq 1$, as $t \to \infty$.

(ii) Since the function $x \mapsto \eta(x + b)$ is continuously derivable on $(0,\, \infty)$, we shall only consider the case $b = 0$ without loss of generality. Pick $t \geq 1$ and $x \in (0,\, 1]$: applying the mean value theorem again shows that there exists $h_2(t,\, x) \in (0,\, u)$ such that

$$\begin{aligned} tx^q\;|\eta(tx) - \eta((t+u)x)| &= x^q\,u\,\frac{t}{t + h_2(t,\, x)}\,|[(t + h_2(t,\, x))x]\;\eta'\,[(t + h_2(t,\, x))x]| \\ &\leq u\,\left|\left[(t + h_2(t,\, x))x^{q+1}\right]\;\eta'\,[(t + h_2(t,\, x))x]\right|. \end{aligned}$$

Now, for all $h \in (0,\, u)$, one has

$$(t+h)x^{q+1}\,\eta'((t+h)x) = \left\{\frac{[(t+h)x]^{q+1}\;\eta'((t+h)x)}{(t+h)^{q+1}\,\eta'(t+h)}\right\}(t+h)\,\eta'(t+h).$$

Since $x \mapsto x^{q+1}\,|\eta'(x)|$ is regularly varying with index $q + \nu > 0$, Lemma 2.3 yields

$$\sup_{\substack{x \in (0,\, 1] \\ h \in (0,\, u)}} \left| \frac{[(t+h)x]^{q+1}\;\eta'((t+h)x)}{(t+h)^{q+1}\,\eta'(t+h)} - x^{q+\nu} \right| \to 0$$

as $t \to \infty$. Using the hypothesis $x\,\eta'(x) \to 0$ as $x \to \infty$ then gives

$$(t+h)x^{q+1}\,\eta'((t+h)x) \to 0$$

as $t \to \infty$, uniformly in $x \in (0,\, 1]$ and $h \in (0,\, u)$, which concludes the proof of Lemma 2.4. ■

Proof of Lemma 2.5. This result easily follows from the identities

$$\begin{aligned}(fg)_{p+u} - (fg)_p &= f_{p+u}\,[g_{p+u} - g_p] + g_p\,[f_{p+u} - f_p],\\ (fg)_{p+u+v} - (fg)_{p+v} - [(fg)_{p+u} - (fg)_p] &= [f_{p+u+v} - f_{p+v} - [f_{p+u} - f_p]]\,g_{p+u+v}\\ &+ [f_{p+u} - f_p]\,[g_{p+u+v} - g_{p+u}]\\ &+ [f_{p+v} - f_p]\,[g_{p+u} - g_p]\\ &+ f_{p+v}\,[g_{p+u+v} - g_{p+v} - [g_{p+u} - g_p]]\end{aligned}$$

and from the properties of (f_p) and (g_p). ■

Proof of Lemma 2.6. Remark that, for p large enough,

$$\forall\, x \in (0,\,1], \quad x^{-k}\,|f_p(x)|\,|u_p(x)| \le C_k\{|u(x)| + r(x)\}$$

where r is a bounded Borel function on $(0,\,1]$. The upper bound is an integrable function on $(0,\,1]$, so that the dominated convergence theorem yields

$$\int_0^1 x^{-k}\,f_p(x)\,u_p(x)\,dx \to \int_0^1 x^{-k}\,f(x)\,u(x)\,dx$$

as $p \to \infty$, which proves the first part of the lemma.

Since v is bounded on $[1,\,\infty)$, $(g_p\,v_p)$ converges uniformly to gv on $[1,\,\infty)$. The function $x \mapsto x^{-k}$ being integrable on $[1,\,\infty)$, the dominated convergence theorem yields

$$\int_1^\infty x^{-k}\,g_p(x)\,v_p(x)\,dx \to \int_1^\infty x^{-k}\,g(x)\,v(x)\,dx$$

as $p \to \infty$, which concludes the proof of Lemma 2.6. ■

Proof of Lemma 2.7. Remark that

$$\frac{1}{p+u} - \frac{1}{p} = \mathrm{O}\left(\frac{1}{p^2}\right) \quad \text{and} \quad \frac{1}{p+u+v} - \frac{1}{p+v} - \left[\frac{1}{p+u} - \frac{1}{p}\right] = \mathrm{O}\left(\frac{1}{p^3}\right)$$

to obtain the result. ■

Proof of Lemma 2.8. – It is clear that for all $x \in (0,\,1]$, $f_p(x) \to e^{-1/x}$ as $p \to \infty$. In order to prove that $(f_p) \in P$, let us rewrite $f_p(x)$ as $f_p(x) = \sigma_p\,\varphi_p(x)\,\psi_p(x)$ where

$$\sigma_p = \left(1 - \frac{1}{p}\right)^{-\alpha-1}, \quad \varphi_p(x) = \left(1 + \frac{1}{(p-1)x}\right)^{-\alpha-2}, \quad \psi_p(x) = \left(1 - \frac{1}{(p-1)x+1}\right)^{p-1},$$

for all $x \in (0,\,1]$, and prove that $(\sigma_p) \in P_0$, $(\varphi_p) \in P_{-1}$ and $(\psi_p) \in P$. First, note that

$$\sigma_p = 1 + \frac{\alpha+1}{p} + \frac{(\alpha+1)(\alpha+2)}{2}\,\frac{1}{p^2} + \mathrm{o}\left(\frac{1}{p^2}\right)$$

so that the collection of constant functions (σ_p) lies in P_0. Second, we have

$$\forall\, p > 1, \quad \forall\, x \in (0,\, 1], \quad |\varphi_p(x)| \leq 1 \leq x^{-1}. \tag{2.25}$$

Moreover,

$$[\varphi_{p+u} - \varphi_p](x) = \varphi_p(x) \left[\left(1 - \frac{u}{p+u-1}\right)^{-\alpha-2} \left(1 - \frac{ux}{(p+u-1)x+1}\right)^{\alpha+2} - 1 \right],$$

and since

$$\forall\, x \in (0,\, 1], \quad \frac{x}{(p+u-1)x+1} \leq \frac{1}{p},$$

Taylor expansions yield, uniformly in $x \in (0,\, 1]$,

$$[\varphi_{p+u} - \varphi_p](x) = \varphi_p(x) \left[\frac{(\alpha+2)u}{p(px+1)} + \mathrm{O}\left(\frac{1}{p^2}\right) \right].$$

From this equality and (2.25), it follows that there exists a positive constant $C^{(1)}$ such that

$$p^2 |\varphi_{p+u} - \varphi_p|(x) \leq C^{(1)}\, x^{-1}. \tag{2.26}$$

Third, let $x \in (0,\, 1]$, and consider a pointwise Taylor expansion of φ_p to get

$$\varphi_p(x) = 1 - \frac{\alpha+2}{px} + \frac{\alpha+2}{p^2 x} \left(-1 + \frac{\alpha+3}{2x}\right) + \mathrm{o}\left(\frac{1}{p^2}\right).$$

Using (2.25), (2.26) and applying Lemma 2.7 therefore shows that $(\varphi_p) \in P_{-1}$.

Let $x \in (0,\, 1]$ and $\Psi_x(p) = (1 - 1/(px+1))^p$, so that $\psi_p(x) = \Psi_x(p-1)$. Since

$$\forall\, p \geq 1, \quad \frac{\Psi_x'(p)}{\Psi_x(p)} = \ln\left[1 - \frac{1}{px+1}\right] + \frac{1}{px+1}$$

and because

$$\forall\, h \in (0,\, 1), \quad \ln(1-h) \leq -h,$$

we get that $\Psi_x(p)$ is a positive nonincreasing function of p. Pick $k \geq 0$: as a consequence, for all $p \geq k+1$ and for all $x \in (0,\, 1]$, $\psi_p(x) \leq \psi_{k+1}(x)$. Remarking that $\psi_{k+1}(x) \leq k^k\, x^k$ for all $x \in (0,\, 1]$, it follows that

$$\forall\, k \geq 0, \quad \exists\, p_k \geq 1, \quad \exists\, C_k \geq 0, \quad \forall\, p \geq p_k, \quad \forall\, x \in (0,\, 1], \quad |\psi_p(x)| \leq C_k\, x^k. \tag{2.27}$$

Recall that Ψ_x is nonincreasing and write

$$|\psi_{p+u} - \psi_p|(x) = \psi_p(x) \left\{ 1 - \left[1 - \frac{1}{(p+u-1)x+1}\right]^u \left[1 + \frac{u}{p-1}\right]^{p-1} \left[1 - \frac{ux}{(p+u-1)x+1}\right]^{p-1} \right\}.$$

Taylor expansions of the logarithm function at 1 and of the exponential function at 0 imply that, uniformly in $x \in (0,\, 1]$,

$$e^u \left[1 - \frac{ux}{(p+u-1)x+1}\right]^{p-1}$$

$$= \exp\left[\frac{u}{(p+u-1)x+1}\right] \left\{1 + \frac{u^2 x}{(p+u-1)x+1} - \frac{p}{2}\left[\frac{ux}{(p+u-1)x+1}\right]^2 + \mathrm{O}\left(\frac{1}{p^2}\right)\right\}.$$

Since

$$\forall\, x \in (0,\, 1], \quad 0 \le \frac{1}{(p+u-1)x+1} \le 1,$$

applying the mean value theorem to the function $h \mapsto \left[(1-h)e^h\right]^u$ gives

$$\left|\left[1 - \frac{1}{(p+u-1)x+1}\right]^u \exp\left[\frac{u}{(p+u-1)x+1}\right] - 1\right| \le \frac{eu}{((p+u-1)x+1)^2}.$$

Writing

$$\left[1 + \frac{u}{p-1}\right]^{p-1} = e^u \left[1 - \frac{u^2}{2p} + \mathrm{O}\left(\frac{1}{p^2}\right)\right]$$

then yields, uniformly in $x \in (0,\, 1]$,

$$|\psi_{p+u} - \psi_p|(x) \le \psi_p(x) \left[\left(eu + \frac{u^2}{2p}\right) \frac{1}{((p+u-1)x+1)^2} + \mathrm{O}\left(\frac{1}{p^2}\right)\right].$$

Therefore, there exists $C^{(2)} \ge 0$ such that, for all p large enough,

$$p^2\, |\psi_{p+u} - \psi_p|\,(x) \le \psi_p(x)\, C^{(2)}\, x^{-2}.$$

Taking (2.27) into account, this entails

$$\forall\, k \ge 0, \quad \exists\, p_k \ge 1, \quad \exists\, C_k \ge 0, \quad \forall\, p \ge p_k, \quad \forall\, x \in (0,\, 1], \quad p^2 |\psi_{p+u} - \psi_p|(x) \le C_k\, x^k. \qquad (2.28)$$

A pointwise Taylor expansion of ψ_p finally gives

$$\psi_p(x) = e^{-1/x}\left[1 + \frac{1}{2px^2} + \frac{1}{p^2x^2}\left(\frac{1}{2} - \frac{1}{3x} + \frac{1}{8x^2}\right) + \mathrm{o}\left(\frac{1}{p^2}\right)\right].$$

Using (2.27), (2.28) and applying Lemma 2.7 shows that $(\psi_p) \in P$. Lemma 2.5 therefore shows that $(f_p) \in P$.

Finally, a Taylor expansion entails, as $p \to \infty$,

$$p^2 \sup_{x \ge 1}\left|g_p(x) - e^{-1/x}\left[1 + \frac{1}{px}\left(1 - \frac{1}{2x}\right) + \frac{1}{p^2x^2}\left(\frac{1}{2} - \frac{5}{6x} + \frac{1}{8x^2}\right)\right]\right| \to 0.$$

It follows that $\sup_{x \ge 1}\left|g_p(x) - e^{-1/x}\right| \to 0$ as $p \to \infty$. Lemma 2.7 then shows that $(g_p) \in U$. ■

Proof of Lemma 2.9. The lemma is a simple consequence of $(f_p) \in P$, $(g_p) \in U$, (2.23) and of the dominated convergence theorem. ■

Proof of Lemma 2.10. Before showing the lemma, notice that if ℓ is a slowly varying function satisfying (H_b), then

$$\forall\, x > 0, \quad \ell(x+b) = c_b \exp\left(\int_1^x \frac{\eta_b(t)}{t}\, dt\right), \text{ with } c_b = c\, \exp\left(\int_1^{1+b} \frac{\eta(t)}{t}\, dt\right) \text{ and } \eta_b(t) = \frac{t\, \eta(t+b)}{t+b},$$

so that the slowly varying function $x \mapsto \ell(x+b)$ satisfies (H_0); moreover, if $(f_p) \in P$, then $(f_{p+b}) \in P$. Without loss of generality, we shall then prove the results in the case $b = 0$.

(i) Let us introduce

$$\forall\, x \in (0,\, 1], \quad Q_p^{(1)}(x) = \frac{\ell_1(px)}{\ell_1(p)} - x$$

so that

$$I_1\, \delta_1(p) = \int_0^1 f_p(x)\, Q_p^{(1)}(x)\, x^{-\alpha-3}\, dx.$$

Since ℓ_1 is regularly varying with index 1, Lemma 2.3 entails $Q_p^{(1)}(x) \to 0$ uniformly in $x \in (0,\, 1]$ as $p \to \infty$. Applying Lemma 2.6 yields $\delta_1(p) \to 0$ as $p \to \infty$.

(ii) In the following, we let p be large enough so that $|\eta|$ is nonincreasing in $[p,\, \infty)$. Pick $s > 1 - \nu$ and let $Q_p^{(1,\,1)}(x) := x^s\, Q_p^{(1)}(x)$. Using the ideas of the proof of Lemma 2.6, one has

$$|I_1\, \delta_1(p)| = \left| \int_0^1 f_p(x)\, Q_p^{(1,\,1)}(x)\, x^{-\alpha-s-3}\, dx \right| = \mathrm{O}\left(\sup_{0<x\leq 1} x^{-1} \left| Q_p^{(1,\,1)}(x) \right| \right).$$

Introducing

$$R_p^{(1,\,1)}(x) := \int_{px}^{p} \frac{\eta(t)}{t}\, dt,$$

(A_2) and the well-known inequality $|e^y - 1| \leq |y| e^{|y|}$ for all $y \in \mathbb{R}$ yield

$$\begin{aligned} \sup_{0<x\leq 1} x^{-1} \left| Q_p^{(1,\,1)}(x) \right| &= \mathrm{O}\left(\sup_{0<x\leq 1} \left\{ x^s \left| 1 - \exp\left(-R_p^{(1,\,1)}(x) \right) \right| \right\} \right) \\ &= \mathrm{O}\left(\sup_{0<x\leq 1} \left\{ x^s \left| R_p^{(1,\,1)}(x) \right| \exp \left| R_p^{(1,\,1)}(x) \right| \right\} \right). \qquad (2.29) \end{aligned}$$

Letting $\widetilde{\eta}(t) = t^{1-\nu}\, \eta(t)$ and using the change of variables $t = wx$, we get

$$\begin{aligned} x^s \left| R_p^{(1,\,1)}(x) \right| &= x^{s-(1-\nu)} \left| \int_p^{p/x} \widetilde{\eta}(wx)\, \frac{dw}{w^{2-\nu}} \right| \\ &\leq x^{s-(1-\nu)} \int_p^{p/x} \left[\left| \frac{\widetilde{\eta}(wx)}{\widetilde{\eta}(w)} - x \right| + x \right] |\widetilde{\eta}(w)|\, \frac{dw}{w^{2-\nu}}. \end{aligned}$$

Remarking that $\widetilde{\eta}$ is regularly varying with index 1, Lemma 2.3 implies that for p large enough,

$$\sup_{\substack{0<x\leq 1 \\ w \geq p}} \left| \frac{\widetilde{\eta}(wx)}{\widetilde{\eta}(w)} - x \right| \leq 1.$$

We then get, for p large enough,

$$\begin{aligned} x^s \left| R_p^{(1,1)}(x) \right| &\leq 2\, x^{s-(1-\nu)} \int_p^{p/x} |\widetilde{\eta}(w)| \,\frac{dw}{w^{2-\nu}} \\ &\leq 2\, x^{s-(1-\nu)} \int_p^{p/x} |\eta(w)| \,\frac{dw}{w} \\ &\leq 2\, |\eta(p)|\, x^{s-(1-\nu)} \int_p^{p/x} \frac{dw}{w} \\ &\leq 2\, |\eta(p)| \sup_{x\in(0,\,1]} \left| x^{s-(1-\nu)} \ln x \right| \\ &= \mathrm{O}\left(|\eta(p)|\right), \end{aligned} \tag{2.30}$$

uniformly in $x \in (0,\, 1]$, since $x \mapsto x^{s-(1-\nu)} \ln x$ is bounded on $(0,\, 1]$. Let us now consider the function $\mathcal{L}(y) = \exp\left(\int_1^y |\eta(t)|\, t^{-1}\, dt\right)$. $\mathcal{L}$ is clearly slowly varying and $\exp\left|R_p^{(1,1)}(x)\right| \leq \mathcal{L}(p)$. Consequently, in view of (2.29) and (2.30), it follows that

$$\sup_{0<x\leq 1} x^{-1} \left| Q_p^{(1,1)}(x) \right| = \mathrm{O}(|\eta(p)|\, \mathcal{L}(p)), \tag{2.31}$$

and therefore $\delta_1(p) = \mathrm{O}(|\eta(p)|\, \mathcal{L}(p))$.

(iii) Keeping in mind that $s > 1 - \nu$, the following expansion holds

$$I_1 \left[\delta_1(p+u) - \delta_1(p)\right] = \int_0^1 f_p(x) \left[Q_{p+u}^{(1,1)} - Q_p^{(1,1)}\right](x)\, x^{-\alpha-s-3}\, dx \tag{2.32}$$

$$+ \int_0^1 [f_{p+u} - f_p](x)\, Q_{p+u}^{(1,1)}(x)\, x^{-\alpha-s-3}\, dx. \tag{2.33}$$

Let us first focus on (2.32). In view of (A_2), and considering

$$R_p^{(1,2)}(x) := \int_{px}^{(p+u)x} \frac{\eta(t)}{t}\, dt,$$

for $x \in (0,\, 1]$, one obtains

$$\left[Q_{p+u}^{(1,1)} - Q_p^{(1,1)}\right](x) = x^s\, \frac{\ell_1((p+u)x)}{\ell_1(p)}\, \frac{p}{p+u} \left\{\exp\left(-R_p^{(1,2)}(1)\right) - \exp\left(-R_p^{(1,2)}(x)\right)\right\}.$$

Mimicking the proof of (ii), we get for p large enough

$$\begin{aligned} x^s \left| R_p^{(1,2)}(x) \right| &\leq x^{s-(1-\nu)} \int_p^{p+u} \left[\left| \frac{\widetilde{\eta}(wx)}{\widetilde{\eta}(w)} - x \right| + x\right] |\widetilde{\eta}(w)| \,\frac{dw}{w^{2-\nu}} \\ &\leq 2\, |\eta(p)| \ln\left(1 + \frac{u}{p}\right) \\ &= \mathrm{O}\left(\frac{|\eta(p)|}{p}\right), \end{aligned} \tag{2.34}$$

uniformly in $x \in (0,\, 1]$. A Taylor expansion of the exponential function at 0 then entails

$$\exp\left(-R_p^{(1,2)}(x)\right) = 1 - \frac{1}{x^s} \left\{x^s R_p^{(1,2)}(x)\right\} \left[1 + \rho\left(R_p^{(1,2)}(x)\right)\right]$$

where ρ is locally bounded on $\mathbb{R}$. Since

$$\sup_{\substack{0<x\le 1\\ p\ge 1}} \left| R_p^{(1,2)}(x) \right| \le u \sup_{t\ge 1} |\eta(t)| \quad \Rightarrow \quad \sup_{\substack{0<x\le 1\\ p\ge 1}} \left| \rho\left(R_p^{(1,2)}(x) \right) \right| < \infty, \tag{2.35}$$

it follows that

$$\sup_{0<x\le 1} \left\{ x^s \left| \exp\left(-R_p^{(1,2)}(x) \right) - 1 \right| \right\} = \mathrm{O}\left(\frac{|\eta(p)|}{p} \right).$$

Applying Lemma 2.3 to ℓ_1 yields

$$\sup_{0<x\le 1} \left| Q_{p+u}^{(1,1)} - Q_p^{(1,1)} \right| (x) = \mathrm{O}\left(\frac{|\eta(p)|}{p} \right)$$

and consequently,

$$\int_0^1 f_p(x) \left[Q_{p+u}^{(1,1)} - Q_p^{(1,1)} \right](x)\, x^{-\alpha-s-3}\, dx = \mathrm{O}\left(\frac{|\eta(p)|}{p} \right). \tag{2.36}$$

Focusing on (2.33), for all $0 < x \le 1$, because $(f_p) \in P$, we get for all sufficiently large p

$$p^2\, x^{-\alpha-s-2}\, |f_{p+u} - f_p|(x) \le C_{\alpha+s+2}$$

which is integrable on (0, 1]. Consequently, in view of (2.31),

$$\int_0^1 [f_{p+u} - f_p](x)\, Q_{p+u}^{(1,1)}(x)\, x^{-\alpha-s-3}\, dx = \mathrm{O}\left(\frac{|\eta(p)|\, \mathcal{L}(p)}{p^2} \right) = \mathrm{O}\left(\frac{|\eta(p)|}{p} \right). \tag{2.37}$$

Collecting (2.36) and (2.37) yields $\delta_1(p+u) - \delta_1(p) = \mathrm{O}\left(\frac{|\eta(p)|}{p} \right)$.

(iv) Let $q > 1 - \nu$ and $Q_p^{(1,2)}(x) = x^{2q+1}\, Q_p^{(1)}(x)$ so that

$$I_1\, \delta_1(p) = \int_0^1 f_p(x)\, Q_p^{(1,2)}(x)\, x^{-\alpha-2q-4}\, dx$$

and the following expansion holds

$$I_1 \left\{ \delta_1(p+u+v) - \delta_1(p+v) - [\delta_1(p+u) - \delta_1(p)] \right\}$$

$$= \int_0^1 [f_{p+u} - f_p](x) \left[Q_{p+u+v}^{(1,2)} - Q_{p+u}^{(1,2)} \right](x)\, x^{-\alpha-2q-4}\, dx \tag{2.38}$$

$$+ \int_0^1 [f_{p+v} - f_p](x) \left[Q_{p+u}^{(1,2)} - Q_p^{(1,2)} \right](x)\, x^{-\alpha-2q-4}\, dx \tag{2.39}$$

$$+ \int_0^1 f_{p+v}(x) \left[Q_{p+u+v}^{(1,2)} - Q_{p+v}^{(1,2)} - \left[Q_{p+u}^{(1,2)} - Q_p^{(1,2)} \right] \right](x)\, x^{-\alpha-2q-4}\, dx \tag{2.40}$$

$$+ \int_0^1 [f_{p+u+v} - f_{p+v} - [f_{p+u} - f_p]](x)\, Q_{p+u+v}^{(1,2)}(x)\, x^{-\alpha-2q-4}\, dx. \tag{2.41}$$

Considering (2.38) and (2.39), arguments given in the proof of (iii) show that

$$\int_0^1 [f_{p+u} - f_p](x) \left[Q_{p+u+v}^{(1,2)} - Q_{p+u}^{(1,2)} \right](x)\, x^{-\alpha-2q-4}\, dx = \mathrm{o}\left(\frac{1}{p^2} \right), \tag{2.42}$$

$$\int_0^1 [f_{p+v} - f_p](x) \left[Q_{p+u}^{(1,2)} - Q_p^{(1,2)} \right](x)\, x^{-\alpha-2q-4}\, dx = \mathrm{o}\left(\frac{1}{p^2} \right). \tag{2.43}$$

Let us now focus on (2.40). From (2.34), (2.35) and (A_2), a Taylor expansion yields

$$\left[Q^{(1,2)}_{p+u+v} - Q^{(1,2)}_{p+v} - \left[Q^{(1,2)}_{p+u} - Q^{(1,2)}_{p}\right]\right](x)$$

$$= x^{2q+1}\frac{\ell_1((p+u+v)x)}{\ell_1(p)}\,\frac{p}{p+u+v}\left\{R^{(1,2)}_p(1) - R^{(1,2)}_{p+v}(1) + R^{(1,2)}_{p+v}(x) - R^{(1,2)}_p(x)\right\} + \mathrm{o}\left(\frac{1}{p^2}\right),$$

uniformly in $x \in (0, 1]$. Let $x \in (0, 1]$: we obtain

$$\begin{aligned} x^{2q+1}\left|R^{(1,2)}_{p+v}(x) - R^{(1,2)}_p(x)\right| &= x^{2q+1}\left|\int_p^{p+u} \frac{\eta((t+v)x) - \eta(tx)}{t+v} - \frac{v\,\eta(tx)}{t(t+v)}\,dt\right| \\ &\le x^q\left[\sup_{t\ge p}\{t\,|\eta((t+v)x) - \eta(tx)|\} + v\sup_{t\ge p}\{x\,|\eta(tx)|\}\right]\frac{u}{p^2}. \end{aligned}$$

Moreover, since $t \mapsto t^{q+1}\,\eta(t)$ is regularly varying with index $q+1+\nu > 0$, Lemma 2.3 yields

$$x^{q+1}\,|\eta(tx)| = \left|\frac{(tx)^{q+1}\,\eta(tx)}{t^{q+1}\,\eta(t)}\right|\,|\eta(t)| \to 0$$

uniformly in $x \in (0, 1]$ as $t \to \infty$. Lemma 2.4ii) therefore entails that, as $p \to \infty$,

$$p^2\sup_{0<x\le 1}\left|Q^{(1,2)}_{p+u+v} - Q^{(1,2)}_{p+v} - \left[Q^{(1,2)}_{p+u} - Q^{(1,2)}_{p}\right]\right|(x) \to 0.$$

The dominated convergence theorem then yields

$$\int_0^1 f_{p+v}(x)\left[Q^{(1,2)}_{p+u+v} - Q^{(1,2)}_{p+v} - \left[Q^{(1,2)}_{p+u} - Q^{(1,2)}_{p}\right]\right](x)\,x^{-\alpha-2q-4}\,dx = \mathrm{o}\left(\frac{1}{p^2}\right). \tag{2.44}$$

Let us finally consider (2.41). Since $(f_p) \in P$ and in view of the triangular inequality, we have, for p large enough,

$$p^2\,x^{-\alpha-2q-4}\,|f_{p+u+v} - f_{p+v} - [f_{p+u} - f_p]|(x) \le C_{\alpha+2q+4}.$$

Because $(f_p) \in P$, the dominated convergence theorem yields

$$\int_0^1 [f_{p+u+v} - f_{p+v} - [f_{p+u} - f_p]](x)\,Q^{(1,2)}_{p+u+v}(x)\,x^{-\alpha-2q-4}\,dx = \mathrm{o}\left(\frac{1}{p^2}\right). \tag{2.45}$$

Collecting (2.42), (2.43), (2.44) and (2.45), (iv) follows. ■

Proof of Lemma 2.11. (i) Let us introduce

$$\forall\, x \ge 1, \quad Q^{(2)}_p(x) = \frac{\ell_2(px)}{\ell_2(p)} - \frac{1}{x},$$

so that

$$I_2\,\delta_2(p) = \int_1^\infty g_p(x)\,Q^{(2)}_p(x)\,x^{-\alpha-1}\,dx.$$

Since ℓ_2 is regularly varying with index -1, using Lemma 2.3 leads to $Q^{(2)}_p(x) \to 0$ as $p \to \infty$ uniformly in $x \ge 1$. Applying Lemma 2.6 entails $\delta_2(p) \to 0$ as $p \to \infty$.

(ii) In the sequel, we let p be so large that $|\eta|$ is nonincreasing in $[p, \infty)$. Recalling that for all $y \in \mathbb{R}$, $|e^y - 1| \le |y| e^{|y|}$, we have, for all $x \ge 1$

$$\left| Q_p^{(2)}(x) \right| \le x^{-1} \left| \int_p^{px} \frac{\eta(t)}{t}\, dt \right| \exp \left| \int_p^{px} \frac{\eta(t)}{t}\, dt \right| \le |\eta(p)|\, x^{|\eta(p)|-1} \ln x. \tag{2.46}$$

Let p be so large that $|\eta(p)| \le 1$. Since $x \mapsto x^{-\alpha-1} \ln x$ is integrable on $[1, \infty)$, the arguments of the proof of Lemma 2.6 entail

$$|I_2\, \delta_2(p)| = \left| \int_1^\infty g_p(x)\, Q_p^{(2)}(x)\, x^{-\alpha-1}\, dx \right| \le |\eta(p)| \int_1^\infty g_p(x)\, x^{-\alpha-1} \ln x\, dx = \mathrm{O}\left(|\eta(p)|\right)$$

which shows (ii).

(iii) Let us remark that

$$I_2\left[\delta_2(p+u) - \delta_2(p)\right] = \int_1^\infty g_p(x) \left[Q_{p+u}^{(2)} - Q_p^{(2)} \right](x)\, x^{-\alpha-1}\, dx \tag{2.47}$$

$$+ \int_1^\infty [g_{p+u} - g_p](x)\, Q_{p+u}^{(2)}(x)\, x^{-\alpha-1}\, dx. \tag{2.48}$$

Consider first (2.47). Because (A_2) holds, setting

$$R_p^{(2)}(x) = \int_p^{p+u} \frac{\eta(tx)}{t}\, dt,$$

for all $x \ge 1$, we get

$$\left[Q_{p+u}^{(2)} - Q_p^{(2)} \right](x) = \frac{\ell_2((p+u)x)}{\ell_2(p)}\, \frac{p+u}{p} \left[\exp\left(-R_p^{(2)}(1)\right) - \exp\left(-R_p^{(2)}(x)\right) \right].$$

Since for all $x \ge 1$, we have $p\left| R_p^{(2)}(x) \right| \le u\,|\eta(p)|$, and recalling that, as $p \to \infty$

$$\sup_{x \ge 1} \left| \frac{\ell_2((p+u)x)}{\ell_2(p)} - \frac{1}{x} \right| \to 0,$$

a Taylor expansion of the exponential function at 0 yields

$$\sup_{x \ge 1} \left| Q_{p+u}^{(2)} - Q_p^{(2)} \right|(x) = \mathrm{O}\left(\frac{|\eta(p)|}{p} \right). \tag{2.49}$$

Taking into account that $(g_p) \in U$ and using (2.46), it follows that

$$\int_1^\infty [g_{p+u} - g_p](x)\, Q_{p+u}^{(2)}(x)\, x^{-\alpha-1}\, dx = \mathrm{O}\left(\frac{|\eta(p)|}{p} \right). \tag{2.50}$$

Moreover, from (2.49), the uniform convergence of (g_p) to $x \mapsto e^{-1/x}$ on $[1, \infty)$ and the dominated convergence theorem, (2.48) is controlled as

$$\int_1^\infty g_p(x) \left[Q_{p+u}^{(2)} - Q_p^{(2)} \right](x)\, x^{-\alpha-1}\, dx = \mathrm{O}\left(\frac{|\eta(p)|}{p} \right). \tag{2.51}$$

(2.50) and (2.51) lead to $\delta_2(p+u) - \delta_2(p) = \mathrm{O}\left(\dfrac{|\eta(p)|}{p} \right)$ and establish (iii).

(iv) One has

$$
\begin{aligned}
&I_2 \left\{\delta_2(p+u+v) - \delta_2(p+v) - [\delta_2(p+u) - \delta_2(p)]\right\} \\
&= \int_1^\infty [g_{p+u} - g_p](x) \left[Q^{(2)}_{p+u+v} - Q^{(2)}_{p+u}\right](x)\, x^{-\alpha-2q-4}\, dx \qquad (2.52) \\
&+ \int_1^\infty [g_{p+v} - g_p](x) \left[Q^{(2)}_{p+u} - Q^{(2)}_{p}\right](x)\, x^{-\alpha-2q-4}\, dx \qquad (2.53) \\
&+ \int_1^\infty g_{p+v}(x) \left[Q^{(2)}_{p+u+v} - Q^{(2)}_{p+v} - \left[Q^{(2)}_{p+u} - Q^{(2)}_{p}\right]\right](x)\, x^{-\alpha-2q-4}\, dx \qquad (2.54) \\
&+ \int_1^\infty [g_{p+u+v} - g_{p+v} - [g_{p+u} - g_p]](x)\, Q^{(2)}_{p+u+v}(x)\, x^{-\alpha-2q-4}\, dx \qquad (2.55)
\end{aligned}
$$

and the four terms are considered separately. Firstly, ideas similar to those developed in the proof of (iii) allow us to control (2.52) and (2.53):

$$
\int_1^\infty [g_{p+u} - g_p](x) \left[Q^{(2)}_{p+u+v} - Q^{(2)}_{p+u}\right](x)\, x^{-\alpha-2q-4}\, dx = \mathrm{o}\left(\frac{1}{p^2}\right), \qquad (2.56)
$$

$$
\int_1^\infty [g_{p+v} - g_p](x) \left[Q^{(2)}_{p+u} - Q^{(2)}_{p}\right](x)\, x^{-\alpha-2q-4}\, dx = \mathrm{o}\left(\frac{1}{p^2}\right). \qquad (2.57)
$$

Secondly, since $p\left|R^{(2)}_p(x)\right| \to 0$ as $p \to \infty$ uniformly in $x \geq 1$, (A_2) entails

$$
\begin{aligned}
&\left[Q^{(2)}_{p+u+v} - Q^{(2)}_{p+v} - \left[Q^{(2)}_{p+u} - Q^{(2)}_{p}\right]\right](x) \\
&= \frac{\ell_2((p+u+v)x)}{\ell_2(p)}\, \frac{p+u+v}{p} \left\{R^{(2)}_p(1) - R^{(2)}_{p+v}(1) + R^{(2)}_{p+v}(x) - R^{(2)}_p(x)\right\} + \mathrm{o}\left(\frac{1}{p^2}\right),
\end{aligned}
$$

uniformly in $x \geq 1$. Remarking that

$$
\begin{aligned}
\left|R^{(2)}_{p+v}(x) - R^{(2)}_p(x)\right| &= \left|\int_p^{p+u} \frac{\eta((t+v)x) - \eta(tx)}{t+v} - \frac{v\,\eta(tx)}{t(t+v)}\, dt\right| \\
&\leq \frac{u}{p^2}\left[\sup_{t\geq p}\left\{t\,|\eta(tx) - \eta((t+v)x)|\right\} + v|\eta(p)|\right],
\end{aligned}
$$

Lemma 2.4i) and Lemma 2.3 imply that

$$
p^2 \sup_{x\geq 1}\left|Q^{(2)}_{p+u+v} - Q^{(2)}_{p+v} - \left[Q^{(2)}_{p+u} - Q^{(2)}_{p}\right]\right|(x) \to 0
$$

as $p \to \infty$. The uniform convergence of (g_p) to $x \mapsto e^{-1/x}$ on $[1, \infty)$ and the dominated convergence theorem then yield the following bound for (2.54):

$$
\int_1^\infty g_{p+v}(x) \left[Q^{(2)}_{p+u+v} - Q^{(2)}_{p+v} - \left[Q^{(2)}_{p+u} - Q^{(2)}_{p}\right]\right](x)\, x^{-\alpha-1}\, dx = \mathrm{o}\left(\frac{1}{p^2}\right). \qquad (2.58)
$$

Finally, recalling that $(g_p) \in U$ and the uniform convergence of $\left(Q^{(2)}_p\right)$ to 0 on $[1, \infty)$, (2.55) is controlled as

$$
\int_1^\infty [g_{p+u+v} - g_{p+v} - [g_{p+u} - g_p]](x)\, Q^{(2)}_{p+u+v}(x)\, dx = \mathrm{o}\left(\frac{1}{p^2}\right). \qquad (2.59)
$$

Collecting (2.56), (2.57), (2.58) and (2.59), (iv) follows and the lemma is proven. ∎

Proof of Lemma 2.12. We first write

$$\begin{aligned}\mathbb{E}(Z^p) &= p\int_0^1 z^{p-1}\, f\circ g(z)\,dz\\ &= p\int_0^1 z^{p-1}\,(1-z)^\alpha\, L\left((1-z)^{-1}\right)\left\{\left[\frac{1}{(1-z)g(z)}\right]^\alpha \frac{L(g(z))}{L\left((1-z)^{-1}\right)}\right\}dz.\end{aligned}$$

Because L is a slowly varying function, Theorem 2.5 entails

$$\ln\left[\frac{L(g(z))}{L\left((1-z)^{-1}\right)}\right] = \ln\left[\frac{C(g(z))}{C\left((1-z)^{-1}\right)}\right] + \int_{(1-z)^{-1}}^{g(z)} \frac{\eta(t)}{t}\,dt$$

where C is an ultimately positive Borel function having a finite limit $c>0$ at ∞ and η is a Borel function tending to 0 at ∞. Letting I_z be the interval

$$I_z = \left[\min\left\{(1-z)^{-1},\, g(z)\right\},\, \max\left\{(1-z)^{-1},\, g(z)\right\}\right]$$

we get

$$\left|\ln\left[\frac{L(g(z))}{L\left((1-z)^{-1}\right)}\right]\right| \le \left|\ln\left[\frac{C(g(z))}{C\left((1-z)^{-1}\right)}\right]\right| + \sup_{I_z}|\eta|\;\left|\ln((1-z)g(z))\right|.$$

Letting $z\to 1$ gives $\dfrac{L(g(z))}{L\left((1-z)^{-1}\right)} \to 1$. Therefore, for all $\varepsilon>0$, there exists $\delta>0$ such that

$$\forall\, z\in[1-\delta,\,1],\quad 1-\frac{\varepsilon}{4} \le \left[\frac{1}{(1-z)g(z)}\right]^\alpha \frac{L(g(z))}{L\left((1-z)^{-1}\right)} \le 1+\frac{\varepsilon}{4}.$$

Applying Lemma 2.1 yields

$$\mathbb{E}(Z^p) = p\int_{1-\delta}^1 z^{p-1}\,(1-z)^\alpha\, L\left((1-z)^{-1}\right)\left\{\left[\frac{1}{(1-z)g(z)}\right]^\alpha \frac{L(g(z))}{L\left((1-z)^{-1}\right)}\right\}dz\,(1+\mathrm{o}(1))$$

and therefore, for all sufficiently large p,

$$1-\frac{\varepsilon}{2} \le \frac{\mathbb{E}(Z^p)}{p\displaystyle\int_{1-\delta}^1 z^{p-1}\,(1-z)^\alpha\, L\left((1-z)^{-1}\right)dz} \le 1+\frac{\varepsilon}{2}.$$

Finally, applying Lemma 2.1 once again entails that for sufficiently large p,

$$1-\varepsilon \le \frac{\mathbb{E}(Z^p)}{p\displaystyle\int_0^1 z^{p-1}\,(1-z)^\alpha\, L\left((1-z)^{-1}\right)dz} \le 1+\varepsilon$$

and thus, as $p\to\infty$,

$$\mathbb{E}(Z^p) = p\int_0^1 z^{p-1}\,(1-z)^\alpha\, L\left((1-z)^{-1}\right)dz\,(1+\mathrm{o}(1)).$$

Retaining notations of Lemma 2.8 and defining ε_1, ε_2 by

$$\begin{aligned}\varepsilon_1(p) &:= \frac{1}{I_1}\int_0^1 f_p(x)\,x^{-\alpha-3}\left[\frac{L_1((p-1)x)}{L_1(p-1)} - x\right]dx, & L_1(x) &= x\,L(x+1),\\ \varepsilon_2(p) &:= \frac{1}{I_2}\int_1^\infty g_p(x)\,x^{-\alpha-1}\left[\frac{L_2(px)}{L_2(p)} - \frac{1}{x}\right]dx, & L_2(x) &= \frac{L(x)}{x},\end{aligned}$$

the change of variables $x=(1-z)^{-1}$ gives

$$p\int_0^1 z^{p-1}\,(1-z)^\alpha\, L\left((1-z)^{-1}\right)dz = p^{-\alpha}\,L(p)\left[\Gamma(\alpha+1) + I_1\,\varepsilon_1(p) + I_2\,\varepsilon_2(p)\right].$$

Since by Lemmas 2.10 and 2.11, $\varepsilon_1(p)$ and $\varepsilon_2(p)$ go to 0 as $p\to\infty$, the result follows. ∎

Proof of Lemma 2.13. (i) Let us remark that, from Lemma 2.12,

$$\xi_n^{(1)} = \frac{\mu_{p_n+1}\, u_{n,a}^{(1)}}{p_n+1}\, a p_n \left(\frac{1}{\widehat{\theta}_n} - \frac{1}{\Theta_n}\right)(1+\mathrm{o}(1)), \tag{2.60}$$

and consider the expansion

$$\begin{aligned} a p_n \left(\frac{1}{\widehat{\theta}_n} - \frac{1}{\Theta_n}\right) &= ((a+1)p_n+1)\frac{\widehat{\mu}_{(a+1)p_n}\,\mu_{(a+1)p_n+1} - \mu_{(a+1)p_n}\,\widehat{\mu}_{(a+1)p_n+1}}{\widehat{\mu}_{(a+1)p_n+1}\,\mu_{(a+1)p_n+1}} \\ &- (p_n+1)\frac{\widehat{\mu}_{p_n}\,\mu_{p_n+1} - \mu_{p_n}\,\widehat{\mu}_{p_n+1}}{\widehat{\mu}_{p_n+1}\,\mu_{p_n+1}} \\ &=: \Delta_n^{(1,1)} - \Delta_n^{(1,2)} \end{aligned}$$

with

$$\begin{aligned} \frac{\mu_{p_n+1}}{p_n+1}\,\Delta_n^{(1,1)} &:= \left[1+\frac{a p_n}{p_n+1}\right]\frac{\mu_{p_n+1}}{\mu_{(a+1)p_n+1}}\left\{\frac{\mu_{(a+1)p_n+1}}{\widehat{\mu}_{(a+1)p_n+1}}\left[\nu_{(a+1)p_n}^{(1)} - \frac{\mu_{(a+1)p_n}}{\mu_{(a+1)p_n+1}}\,\nu_{(a+1)p_n+1}^{(1)}\right]\right\}, \\ \frac{\mu_{p_n+1}}{p_n+1}\,\Delta_n^{(1,2)} &:= \frac{\mu_{p_n+1}}{\widehat{\mu}_{p_n+1}}\left[\nu_{p_n}^{(1)} - \frac{\mu_{p_n}}{\mu_{p_n+1}}\,\nu_{p_n+1}^{(1)}\right]. \end{aligned}$$

Replacing in (2.60), (i) follows.

(ii) We first notice that, using Lemma 2.12,

$$\xi_n^{(2)} = \frac{\mathfrak{m}_{p_n+1}\, u_{n,a}^{(2)}}{\varphi^{a+1}}\,(\widetilde{\varphi}_n^a - \Phi_n^a)(1+\mathrm{o}(1)). \tag{2.61}$$

We have $\widetilde{\varphi}_n^a - \Phi_n^a = \Delta_n^{(2,1)} + \Delta_n^{(2,2)}$ where

$$\begin{aligned} \Delta_n^{(2,1)} &= \left[\frac{\widetilde{\mathfrak{m}}_{p_n}}{\widetilde{\mathfrak{m}}_{p_n+1}} - \frac{\mathfrak{m}_{p_n}}{\mathfrak{m}_{p_n+1}}\right]\frac{\widetilde{\mathfrak{m}}_{(a+1)p_n+a+1}}{\widetilde{\mathfrak{m}}_{(a+1)p_n}}, \\ \Delta_n^{(2,2)} &= \left[\frac{\widetilde{\mathfrak{m}}_{(a+1)p_n+a+1}}{\widetilde{\mathfrak{m}}_{(a+1)p_n}} - \frac{\mathfrak{m}_{(a+1)p_n+a+1}}{\mathfrak{m}_{(a+1)p_n}}\right]\frac{\mathfrak{m}_{p_n}}{\mathfrak{m}_{p_n+1}}; \end{aligned}$$

rewriting

$$\begin{aligned} \Delta_n^{(2,1)} &= \frac{1}{\mathfrak{m}_{p_n+1}}\,\frac{\mathfrak{m}_{p_n+1}}{\widetilde{\mathfrak{m}}_{p_n+1}}\,\frac{\widetilde{\mathfrak{m}}_{(a+1)p_n+a+1}}{\widehat{\mathfrak{m}}_{(a+1)p_n}}\left[\nu_{p_n}^{(2)} - \frac{\mathfrak{m}_{p_n}}{\mathfrak{m}_{p_n+1}}\,\nu_{p_n+1}^{(2)}\right], \\ \Delta_n^{(2,2)} &= \frac{\mathfrak{m}_{(a+1)p_n+a+1}}{\mathfrak{m}_{(a+1)p_n}^2}\,\frac{\mathfrak{m}_{(a+1)p_n}}{\widetilde{\mathfrak{m}}_{(a+1)p_n}}\,\frac{\mathfrak{m}_{p_n}}{\mathfrak{m}_{p_n+1}}\left[-\nu_{(a+1)p_n}^{(2)} + \frac{\mathfrak{m}_{(a+1)p_n}}{\mathfrak{m}_{(a+1)p_n+a+1}}\,\nu_{(a+1)p_n+a+1}^{(2)}\right] \end{aligned}$$

and replacing in (2.61) shows (ii). ■

Proof of Lemma 2.14. Remark first that

$$\frac{\mathbb{E}(Z^p)}{\mathbb{E}(Z^{p+u})} = \left[1+\frac{u}{p}\right]^{\alpha}\exp\left[-\int_p^{p+u}\frac{\eta(t)}{t}\,dt\right]\left[1+\frac{\varepsilon(p)-\varepsilon(p+u)}{1+\varepsilon(p+u)}\right].$$

Noticing that

$$\forall\, u \geq 1, \quad \int_p^{p+u}\frac{\eta(t)}{t}\,dt = \mathrm{O}\left(\frac{|\eta(p)|}{p}\right) = \mathrm{o}\left(\frac{1}{p}\right),$$

a Taylor expansion of the exponential function at 0 and the first hypothesis on ε yield

$$\frac{\mathbb{E}(Z^p)}{\mathbb{E}(Z^{p+u})} = \left[1+\frac{u}{p}\right]^{\alpha} + \mathrm{o}\left(\frac{1}{p}\right). \tag{2.62}$$

This makes it possible to write

$$\begin{aligned} w(s,\,t,\,u,\,v,\,p) &= \left[-1,\ \left[1+\frac{u}{sp}\right]^{\alpha}\right] \begin{bmatrix} \mathbb{E}\left(Z^{(s+t)p}\right) & \mathbb{E}\left(Z^{(s+t)p+v}\right) \\ \mathbb{E}\left(Z^{(s+t)p+u}\right) & \mathbb{E}\left(Z^{(s+t)p+u+v}\right) \end{bmatrix} \left[-1,\ \left[1+\frac{v}{tp}\right]^{\alpha}\right]^t \\ &+ \left[-1,\ \left[1+\frac{u}{sp}\right]^{\alpha}\right] \begin{bmatrix} \mathbb{E}\left(Z^{(s+t)p}\right) & \mathbb{E}\left(Z^{(s+t)p+v}\right) \\ \mathbb{E}\left(Z^{(s+t)p+u}\right) & \mathbb{E}\left(Z^{(s+t)p+u+v}\right) \end{bmatrix} \left[0,\ \mathrm{o}\left(\frac{1}{p}\right)\right]^t \\ &+ \left[0,\ \mathrm{o}\left(\frac{1}{p}\right)\right] \begin{bmatrix} \mathbb{E}\left(Z^{(s+t)p}\right) & \mathbb{E}\left(Z^{(s+t)p+v}\right) \\ \mathbb{E}\left(Z^{(s+t)p+u}\right) & \mathbb{E}\left(Z^{(s+t)p+u+v}\right) \end{bmatrix} \left[-1,\ \left[1+\frac{v}{tp}\right]^{\alpha}\right]^t \\ &+ \left[0,\ \mathrm{o}\left(\frac{1}{p}\right)\right] \begin{bmatrix} \mathbb{E}\left(Z^{(s+t)p}\right) & \mathbb{E}\left(Z^{(s+t)p+v}\right) \\ \mathbb{E}\left(Z^{(s+t)p+u}\right) & \mathbb{E}\left(Z^{(s+t)p+u+v}\right) \end{bmatrix} \left[0,\ \mathrm{o}\left(\frac{1}{p}\right)\right]^t. \end{aligned}$$

Clearly, for all $b \geq 1$, because of (2.62) and the hypothesis on Z,

$$\begin{aligned} &[0,\,1] \begin{bmatrix} \mathbb{E}\left(Z^{(s+t)p}\right) & \mathbb{E}\left(Z^{(s+t)p+v}\right) \\ \mathbb{E}\left(Z^{(s+t)p+u}\right) & \mathbb{E}\left(Z^{(s+t)p+u+v}\right) \end{bmatrix} \left[-1,\ \left[1+\frac{1}{bp}\right]^{\alpha}\right]^t \\ &\qquad= \mathbb{E}\left(Z^{(s+t)p+u+v}\right) \left\{ \left[1+\frac{1}{bp}\right]^{\alpha} - \frac{\mathbb{E}\left(Z^{(s+t)p+u}\right)}{\mathbb{E}\left(Z^{(s+t)p+u+v}\right)} \right\} \\ &\qquad= \mathrm{O}\left(p^{-\alpha-1}\,L(p)\right). \end{aligned}$$

Besides, Lemma 2.12 yields

$$[0,\,1] \begin{bmatrix} \mathbb{E}\left(Z^{(s+t)p}\right) & \mathbb{E}\left(Z^{(s+t)p+v}\right) \\ \mathbb{E}\left(Z^{(s+t)p+u}\right) & \mathbb{E}\left(Z^{(s+t)p+u+v}\right) \end{bmatrix} [0,\,1]^t = \mathbb{E}\left(Z^{(s+t)p+u+v}\right) = \mathrm{O}\left(p^{-\alpha}\,L(p)\right)$$

so that

$$\begin{aligned} w(s,\,t,\,u,\,v,\,p) &= \left[-1,\ \left[1+\frac{u}{sp}\right]^{\alpha}\right] \begin{bmatrix} \mathbb{E}\left(Z^{(s+t)p}\right) & \mathbb{E}\left(Z^{(s+t)p+v}\right) \\ \mathbb{E}\left(Z^{(s+t)p+u}\right) & \mathbb{E}\left(Z^{(s+t)p+u+v}\right) \end{bmatrix} \left[-1,\ \left[1+\frac{v}{tp}\right]^{\alpha}\right]^t \\ &+ \mathrm{o}\left(p^{-\alpha-2}\,L(p)\right). \end{aligned}$$

Set now

$$1+T(p,\,u) = \exp\left[\int_p^{p+u} \frac{\eta(t)}{t}\,dt\right] \left[1+\frac{\varepsilon(p+u)-\varepsilon(p)}{1+\varepsilon(p)}\right]$$

to obtain

$$\frac{1}{\mathbb{E}\left(Z^{(s+t)p}\right)} \begin{bmatrix} \mathbb{E}\left(Z^{(s+t)p}\right) & \mathbb{E}\left(Z^{(s+t)p+v}\right) \\ \mathbb{E}\left(Z^{(s+t)p+u}\right) & \mathbb{E}\left(Z^{(s+t)p+u+v}\right) \end{bmatrix} = M_1(s,\,t,\,u,\,v,\,p) + M_2(s,\,t,\,u,\,v,\,p)$$

where

$$M_1(s,\,t,\,u,\,v,\,p) \;=\; \begin{bmatrix} 1 & \left[1+\dfrac{v}{(s+t)p}\right]^{-\alpha} \\ \left[1+\dfrac{u}{(s+t)p}\right]^{-\alpha} & \left[1+\dfrac{u+v}{(s+t)p}\right]^{-\alpha} \end{bmatrix},$$

$$M_2(s,\,t,\,u,\,v,\,p) \;=\; \begin{bmatrix} 0 & \left[1+\dfrac{v}{(s+t)p}\right]^{-\alpha} T((s+t)p,\,v) \\ \left[1+\dfrac{u}{(s+t)p}\right]^{-\alpha} T((s+t)p,\,u) & \left[1+\dfrac{u+v}{(s+t)p}\right]^{-\alpha} T((s+t)p,\,u+v) \end{bmatrix}.$$

Because of the hypotheses on ε and since L meets the second-order condition (H_b), we get, for all $u \geq 1$,

$$T(p,\,u) = \int_p^{p+u} \frac{\eta(t)}{t}\,dt + \frac{\varepsilon(p+u)-\varepsilon(p)}{1+\varepsilon(p)} + \mathrm{o}\left(\frac{1}{p}\right) = \mathrm{o}\left(\frac{1}{p}\right)$$

and, for all $u,\,v \geq 1$,

$$\begin{aligned} T(p,\,u+v) - T(p,\,u) - T(p,\,v) \;&=\; \int_p^{p+u} \frac{\eta(t+v)-\eta(t)}{t+v} - \frac{v\,\eta(t)}{t(t+v)}\,dt \\ &+\; \frac{\varepsilon(p+u+v)-\varepsilon(p+v)-[\varepsilon(p+u)-\varepsilon(p)]}{1+\varepsilon(p)} + \mathrm{o}\left(\frac{1}{p^2}\right) \\ &=\; \mathrm{o}\left(\frac{1}{p^2}\right); \end{aligned}$$

as a consequence,

$$\left[-1,\; \left[1+\frac{u}{sp}\right]^{\alpha}\right] M_2(s,\,t,\,u,\,v,\,p) \left[-1,\; \left[1+\frac{v}{tp}\right]^{\alpha}\right]^t = \mathrm{o}\left(\frac{1}{p^2}\right)$$

and therefore

$$\frac{w(s,\,t,\,u,\,v,\,p)}{\mathbb{E}\left(Z^{(s+t)p}\right)} = \left[-1,\; \left[1+\frac{u}{sp}\right]^{\alpha}\right] M_1(s,\,t,\,u,\,v,\,p) \left[-1,\; \left[1+\frac{v}{tp}\right]^{\alpha}\right]^t + \mathrm{o}\left(\frac{1}{p^2}\right).$$

Some straightforward computations yield

$$\left[-1,\; \left[1+\frac{u}{sp}\right]^{\alpha}\right] M_1(s,\,t,\,u,\,v,\,p) \left[-1,\; \left[1+\frac{v}{tp}\right]^{\alpha}\right]^t = \frac{\alpha(\alpha+1)\,uv}{(s+t)^2\,p^2} + \mathrm{o}\left(\frac{1}{p^2}\right);$$

because L is slowly varying, $\mathbb{E}\left(Z^{(s+t)p}\right) = (s+t)^{-\alpha}\,p^{-\alpha}\,L(p)\,\Gamma(\alpha+1)\,(1+\mathrm{o}(1))$, which completes the proof. ■

Proof of Lemma 2.15. Hölder's inequality yields

$$\mathbb{E}|Z^{p_n}\,h_n(Z)|^3 \leq (m+1)^2 \left[\sum_{j=0}^{m} \frac{1}{p_n^{3j}} \sup_{\substack{[0,\,1] \\ n\in\mathbb{N}\setminus\{0\}}} |H_{n,\,j}|^3\,\mathbb{E}\left|Z^{p_n}\,(1-Z)^{m-j}\right|^3\right].$$

It is then sufficient to prove that

$$\forall\, j \in \{0,\,\ldots,\,m\},\quad \mathbb{E}\left|Z^{p_n}\,(1-Z)^{m-j}\right|^3 = \mathrm{O}\left(p_n^{-\alpha-(3m-3j)}\,L(p_n)\right).$$

An integration by parts gives

$$\mathbb{E}\left[Z^{p_n}\,(1-Z)^{m-j}\right]^3 = \int_0^1 \frac{d}{dz}\left[z^{3p_n}\,(1-z)^{3m-3j}\right]\, f\circ g(z)\,dz.$$

It is now enough to notice that, if (s_n) is a positive real sequence tending to ∞ and $d \geq 0$,

$$\int_0^1 z^{s_n}\,(1-z)^d\, f\circ g(z)\,dz = \frac{\mathbb{E}\left(W^{s_n+1}\right)}{s_n+1}$$

where W has survival function $z \mapsto (1-z)^d\, f\circ g(z)$. Applying Lemma 2.12 therefore yields

$$\mathbb{E}\left(W^{s_n+1}\right) = s_n^{-\alpha-d}\, L(s_n)\,\Gamma(\alpha+d+1)\,(1+\mathrm{o}(1)),$$

which establishes Lemma 2.15. ■

Proof of Lemma 2.16. (i) It is enough to prove that (i) holds for $\mathcal{C}_2$. It is straightforward that $\mathcal{C}_2$ is a linear subspace of $\mathcal{F}(\mathbb{N}\times[0,\,\infty),\,\mathbb{R})$; let now u and v lie in $\mathcal{C}_2$ and set $w_n(x) = u_n(x)\,v_n(x)$. Condition (Q_1) clearly holds for w. Then, one has, for all $0 \leq x < y$:

$$w_n(y) - w_n(x) = u_n(y)[v_n(y) - v_n(x)] + v_n(x)[u_n(y) - u_n(x)]$$

so that w satisfies (Q_2).

We shall use this equality to prove that w satisfies (Q_3). Pick $0 \leq x < y$ and $z > 0$: we have

$$\begin{aligned} w_n(y+z) - w_n(y) - \{w_n(x+z) - w_n(x)\} & \\ = \;& u_n(y+z)[v_n(y+z) - v_n(y)] + v_n(y)[u_n(y+z) - u_n(y)] \\ - \;& \{u_n(x+z)[v_n(x+z) - v_n(x)] + v_n(x)[u_n(x+z) - u_n(x)]\} \\ = \;& u_n(y+z)[v_n(y+z) - v_n(y) - \{v_n(x+z) - v_n(x)\}] \\ + \;& v_n(y)[u_n(y+z) - u_n(y) - \{u_n(x+z) - u_n(x)\}] \\ + \;& [u_n(y+z) - u_n(x+z)][v_n(x+z) - v_n(x)] \\ + \;& [v_n(y) - v_n(x)][u_n(x+z) - u_n(x)] \end{aligned}$$

and recall that $u,\, v \in \mathcal{C}_2$, so that w satisfies (Q_3).

(ii) Because $\mathcal{E}_2 \subset \mathcal{C}_2$ which is a sub-algebra of $\mathcal{F}(\mathbb{N}\times[0,\,\infty),\,\mathbb{R})$, it is enough to prove that if $u \in \mathcal{E}_2$, then $1/u \in \mathcal{E}_2$. Let then $w = 1/u$ and denote by f the function defined by $f(x) = 1/x$, so that $w_n(x) = f(u_n(x))$: w clearly satisfies (Q_1) and belongs to $\mathcal{E}$. Pick now $x \in [0,\,\infty)$; because the sequence $(u_n(x))$ is bounded from below by a positive constant, f is three times continuously derivable on the interval $I := [\min(u_n(0),\,u_n(x)),\,\max(u_n(0),\,u_n(x))]$, so that a Taylor formula yields

$$w_n(x) = f(u_n(x)) = \sum_{j=0}^{2} \frac{[u_n(x) - u_n(0)]^j}{j!} f^{(j)}(u_n(0)) + \frac{1}{2}\int_{u_n(0)}^{u_n(x)} (u_n(x) - t)^2\, f^{(3)}(t)\,dt.$$

Clearly, because of (i), for all $0 \le j \le 2$, the sequence

$$v^{(j)} : (n,\, x) \mapsto \frac{[u_n(x) - u_n(0)]^j}{j!} f^{(j)}(u_n(0))$$

lies in $\mathcal{C}_2$. Now, since $u \in \mathcal{C}_2$, $[u_n(x) - u_n(0)]^3 = \mathrm{O}\left(1/p_n^3\right)$: therefore,

$$\left| \int_{u_n(0)}^{u_n(x)} (u_n(x) - t)^2 \, f^{(3)}(t)\, dt \right| \le [u_n(x) - u_n(0)]^3 \sup_I |f^{(3)}| = \mathrm{O}\left(\frac{1}{p_n^3}\right)$$

so that, putting $v = \sum_{j=0}^{2} v^{(j)} \in \mathcal{C}_2$, $w_n(x) = v_n(x) + \mathrm{o}\left(1/p_n^2\right)$. This is enough to conclude that $w \in \mathcal{C}_2$, and thus $w \in \mathcal{E}_2$, which finishes the proof.

(iii) Note first that if $\beta < 0$ and $u_n(x) = (p_n + x)^\beta$, one has

$$u_n(x) = p_n^\beta \left[1 + \beta\, \frac{x}{p_n} + \frac{\beta(\beta - 1)}{2}\, \frac{x^2}{p_n^2} + \mathrm{o}\left(\frac{1}{p_n^2}\right) \right].$$

Since clearly $(n,\, x) \mapsto \dfrac{x}{p_n}$ and $(n,\, x) \mapsto \dfrac{x^2}{p_n^2}$ belong to $\mathcal{D}_2$, one gets

$$\forall\, \beta < 0, \quad (n,\, x) \mapsto (p_n + x)^\beta \in \mathcal{D}_2,$$

from which (iii) readily follows. ■

Proof of Lemma 2.17. To prove that $q(p_n) \to 0$ as $n \to \infty$, pick $\varepsilon > 0$ and $\rho > 0$ such that

$$\forall\, x \in [1 - \rho,\, 1], \quad |\delta(x)| < \frac{\varepsilon}{2}.$$

Applying Lemma 2.1 yields

$$\frac{\left| \displaystyle\int_0^1 x^{p_n} \, (1 - x)^\alpha \, \delta(x)\, dx \right|}{\displaystyle\int_0^1 x^{p_n} \, (1 - x)^\alpha \, dx} = \frac{\left| \displaystyle\int_{1-\rho}^1 x^{p_n} \, (1 - x)^\alpha \, \delta(x)\, dx \right|}{\displaystyle\int_{1-\rho}^1 x^{p_n} \, (1 - x)^\alpha \, dx} (1 + \mathrm{o}(1));$$

because

$$\left| \int_{1-\rho}^1 x^{p_n} \, (1 - x)^\alpha \, \delta(x)\, dx \right| \le \frac{\varepsilon}{2} \int_{1-\rho}^1 x^{p_n} \, (1 - x)^\alpha \, dx,$$

one gets, for all sufficiently large n,

$$\frac{\left| \displaystyle\int_0^1 x^{p_n} \, (1 - x)^\alpha \, \delta(x)\, dx \right|}{\displaystyle\int_0^1 x^{p_n} \, (1 - x)^\alpha \, dx} \le \varepsilon,$$

which shows that $q(p_n) \to 0$ as $n \to \infty$.

Next, let for every $\beta > 0$ and every bounded Borel function g on $(0,\, 1)$,

$$I_{\beta,\, g}(p_n) = \int_0^1 x^{p_n} \, (1 - x)^\beta \, g(x)\, dx.$$

Notice that

$$(p_n + x)^{\alpha+1}\, B(p_n + x + 1,\, \alpha + 1) = \frac{(p_n + x)^{\alpha+2}}{\alpha + 1} B(p_n + x,\, \alpha + 2);$$

using [103] and Lemma 2.16iii), we get

$$(n,\, x) \mapsto (p_n + x)^{\alpha+1}\, B(p_n + x + 1,\, \alpha + 1) \in \mathcal{E}_2.$$

Lemma 2.16i) and ii) then show that it is enough to prove that

$$(n,\, x) \mapsto (p_n + x)^{\alpha+1}\, I_{\alpha,\,\delta}(p_n + x) \in \mathcal{D}_2.$$

Because $(p_n + x)^{\alpha+1} = p_n^{\alpha+1}\, (1 + x/p_n)^{\alpha+1}$ and $(n,\, x) \mapsto (1 + x/p_n)^{\alpha+1} \in \mathcal{E}_2$, we deduce from Lemma 2.16i) and ii) that is sufficient to show that

$$(n,\, x) \mapsto p_n^{\alpha+1}\, I_{\alpha,\,\delta}(p_n + x) \in \mathcal{D}_2.$$

For all $t > 0$, let then $r_t : (0,\, 1) \to [0,\, \infty)$ be the map defined by

$$\forall\, y \in (0,\, 1), \quad r_t(y) = \frac{1 - y^t}{1 - y}.$$

For all $t > 0$, r_t is a bounded Borel function on $(0,\, 1)$. Moreover, for all $0 \leq x < y$ and all $z > 0$,

$$\begin{aligned} p_n[w_n(y) - w_n(x)] &= -p_n^{\alpha+2}\, I_{\alpha+1,\,\delta r_{y-x}}(p_n + x) \\ p_n^2[w_n(y + z) - w_n(y) - \{w_n(x + z) - w_n(x)\}] &= p_n^{\alpha+3}\, I_{\alpha+2,\,\delta r_{y-x} r_z}(p_n + x) \end{aligned}$$

Because for all $x,\, m \geq 0$ and every bounded Borel function S on $(0,\, 1)$ we have

$$p_n^{\alpha+m+1}\, I_{\alpha+m,\,\delta S}(p_n + x) \to 0$$

as $n \to \infty$, Lemma 2.17 is proven. ■

3 Frontier estimation with high order moments

3.1 Context

First and foremost, we briefly recall the framework of our study. Let $(X_1, Y_1), \ldots, (X_n, Y_n)$ be n independent copies of a random pair (X, Y) such that their common distribution has a support defined by $S = \{(x, y) \in \Omega \times \mathbb{R};\ 0 \leq y \leq g(x)\}$, where Ω is a compact subset of $\mathbb{R}^d$, and X has a probability density function f on $\mathbb{R}^d$. We address the problem of estimating the function g, which is called the frontier of S. Girard and Jacob's estimator [46] of g is defined by

$$\widehat{g}_n^{GJ}(x) = \left[(p_n + 1) \frac{\displaystyle\sum_{i=1}^{n} Y_i^{p_n} K_{h_n}(x - X_i)}{\displaystyle\sum_{i=1}^{n} K_{h_n}(x - X_i)} \right]^{1/p_n}$$

where (p_n) is a nonrandom sequence of positive numbers tending to ∞, (h_n) is a nonrandom sequence of positive numbers tending to 0, and $K_h(u) = h^{-d} K(u/h)$ where K is a probability density function with support included in the unit ball $\mathcal{B}$ of $\mathbb{R}^d$.

In the particular case where Y given $X = x$ is uniformly distributed, this estimator is asymptotically Gaussian (see [46]). Here, we first prove that this conclusion fails to hold if Y given $X = x$ is not uniformly distributed, see Section 3.2. In Section 3.3, we then introduce another estimator based on a kernel regression on high order moments of the data, whose asymptotic properties are studied, see Sections 3.3.2 and 3.3.3. The performances of our estimator are examined on some finite sample situations in Section 3.4. Auxiliary results are postponed to Appendix A and proofs to Appendix B.

3.2 A drawback of Girard and Jacob's estimator

Our aim here is to prove that when Y given $X = x$ is not uniformly distributed, the estimator $\widehat{g}_n^{GJ}(x)$, centered on $g(x)$, is not asymptotically Gaussian. To this end, we shall first show that the estimator

$$\widehat{g}_n^{GJ,*}(x) = \left[\frac{1}{C(x)\,\alpha(x)\,B(p_n+1,\,\alpha(x))} \frac{\sum_{i=1}^{n} Y_i^{p_n}\, K_{h_n}(x-X_i)}{\sum_{i=1}^{n} K_{h_n}(x-X_i)} \right]^{1/p_n} \tag{3.1}$$

where, as in Section 2.2.1, B is the Beta function, is asymptotically Gaussian if the cumulative distribution function $F(\cdot\,|\,x)$ of Y given $X = x$ satisfies the parametric assumption

(P) $\forall\, y \in [0,\, g(x)],\ F(y\,|\,x) = 1 - (1 - y/g(x))^{\alpha(x)}\, L\left(x,\, (1-y/g(x))^{-1}\right)$, where L is bounded on $\mathbb{R}^d \times [1,\, \infty)$ and satisfies

$$\forall\, x \in \Omega, \quad \exists\, M(x) \geq 1, \quad \forall\, z > M(x), \quad L(x,\, z) = C(x) + D(x)\, z^{-\beta(x)}\, [1 + \delta(x,\, z)]$$

where α, β and C are positive Borel functions, D is a Borel function which never vanishes, and δ is a locally bounded Borel function on $\Omega \times [1,\, \infty)$ such that for all $x \in \Omega$, $\delta(x,\, \cdot)$ goes to 0 at ∞.

Of course, one has

$$\widehat{g}_n^{GJ,*}(x) = \left[\frac{1}{(p_n+1)\, C(x)\, \alpha(x)\, B(p_n+1,\, \alpha(x))} \right]^{1/p_n} \widehat{g}_n^{GJ}(x),$$

so that Girard and Jacob's estimator $\widehat{g}_n^{GJ}(x)$ is obtained in the case $\alpha(x) = 1$ and $C(x) = 1$. Here, for all $x \in \Omega$, the function $z \mapsto L(x,\, z)$ is slowly varying and belongs to the Hall class (3). In the general context of extreme-value theory, the conditional distribution of Y given $X = x$ belongs to the Weibull max-domain of attraction with conditional extreme-value index $-1/\alpha(x)$.

We shall assume that

(K) The kernel K is bounded and its support is included in the unit ball $\mathcal{B}$ of $\mathbb{R}^d$.

Note that (K) implies that $\forall\, q \geq 1$, $\int_{\mathcal{B}} K^q(u)\, du < \infty$. The following regularity assumptions are introduced:

(A_1) f and g are locally Hölder continuous with respective exponents η_f and η_g.

(A_2) α, β, C and D are locally Hölder continuous with respective exponents η_α, η_β, η_C and η_D.

(A_3) The function $x \mapsto M(x)$ is locally bounded on Ω.

(A_4) $\forall\, x \in \Omega$, $\sup_{\|h\| \leq \varepsilon} \sup_{z \geq T} |\delta(x,\, z) - \delta(x+h,\, z)| \to 0$ as $\varepsilon \to 0$ and $T \to \infty$.

Notice that if (P) and $(A_2 - A_4)$ hold, then β, C and D are bounded away from 0 and ∞. Therefore, putting

$$\widetilde{\delta}(x,\, z) = \begin{cases} z^{\beta(x)}\, \dfrac{L(x,\, z) - C(x)}{D(x)} - 1 & \text{if } 1 \leq z \leq M(x) \\ \delta(x,\, z) & \text{if } z > M(x) \end{cases}$$

then $\widetilde{\delta}$ is a locally bounded Borel function on $\Omega \times [1, \infty)$ such that, for all $x \in \Omega$, $\widetilde{\delta}(x, \cdot)$ goes to 0 at ∞; besides, one has

$$\forall\, x \in \Omega, \quad \forall\, z > 1, \quad L(x,\, z) = C(x) + D(x)\, z^{-\beta(x)} \left(1 + \widetilde{\delta}(x,\, z)\right)$$

and $\widetilde{\delta}$ satisfies (A_4). In the parametric setting (P), we may therefore assume that M is constant equal to 1.

Let us introduce some more notations: for all $x \in \mathbb{R}^d$, we set

$$\widehat{\mu}_{p_n}(x) = \frac{1}{n} \sum_{i=1}^{n} Y_i^{p_n}\, K_{h_n}(x - X_i) \quad \text{and} \quad \widehat{f}_n(x) = \frac{1}{n} \sum_{i=1}^{n} K_{h_n}(x - X_i).$$

Note that $\widehat{\mu}_{p_n}(x)$ is the standard empirical estimator of the moment $\mu_{p_n}(x) = \mathbb{E}(Y^{p_n}\, K_{h_n}(x - X))$, itself a smoothed version of the conditional moment $m_{p_n}(x) = \mathbb{E}(Y^{p_n} \,|\, X = x)$: that is,

$$\mu_{p_n}(x) = \int_{\Omega} K_{h_n}(x - v)\, m_{p_n}(v)\, f(v)\, dv.$$

$\widehat{f}_n(x)$ is the classical kernel estimator of f (see Parzen [90]), and $\widehat{r}_n(x) = \widehat{\mu}_{p_n}(x)/\widehat{f}_n(x)$. We may then rewrite $\widehat{g}_n^{GJ,*}(x)$ as

$$\widehat{g}_n^{GJ,*}(x) = \left\{ \frac{1}{C(x)\, \alpha(x)\, B(p_n + 1,\, \alpha(x))}\, \widehat{r}_n(x) \right\}^{1/p_n}.$$

We let

$$\mathcal{R}_n(x) = \frac{\mu_{p_n}(x)}{\mathbb{E}\left(\widehat{f}_n(x)\right)} \quad \text{and} \quad \mathcal{G}_n^*(x) = \left\{ \frac{1}{C(x)\, \alpha(x)\, B(p_n + 1,\, \alpha(x))}\, \mathcal{R}_n(x) \right\}^{1/p_n}$$

be the deterministic counterparts of $\widehat{r}_n(x)$ and $\widehat{g}_n^{GJ,*}(x)$.

Finally, for any real-valued function φ on $\mathbb{R}^d$, the oscillation of φ between two points x and $x - h_n u$, $u \in \mathcal{B}$, is defined by

$$\Delta_n^{\varphi}(x,\, u) = \varphi(x - h_n u) - \varphi(x).$$

Our first asymptotic result is the asymptotic normality of $\widehat{g}_n^{GJ,*}(x)$, centered on $\mathcal{G}_n^*(x)$:

Proposition 3.1. *Assume that (P), (K) and $(A_1 - A_4)$ hold. Let $x \in \Omega$ such that $f(x) > 0$. If $n\, p_n^{-\alpha(x)}\, h_n^d \to \infty$ and $p_n\, h_n^{\eta_g} \to 0$ then*

$$v_n(x) \left(\frac{\widehat{g}_n^{GJ,*}(x)}{\mathcal{G}_n^*(x)} - 1 \right) \xrightarrow{d} \mathcal{N}\left(0,\, \frac{2^{-\alpha(x)}\, \|K\|_2^2}{f(x)\, C(x)\, \Gamma(\alpha(x) + 1)}\right), \qquad \text{as } n \to \infty$$

where $v_n(x) = \sqrt{n}\, p_n^{-\alpha(x)/2+1}\, h_n^{d/2}$ and $\|K\|_2^2 = \int_{\mathcal{B}} K^2(u)\, du$.

Proof. First, notice that

$$\widehat{g}_n^{GJ,*}(x) = g(x)\left[\frac{1}{C(x)\,\alpha(x)\,B(p_n+1,\,\alpha(x))}\,\frac{\displaystyle\sum_{i=1}^{n}\left\{\frac{Y_i}{g(x)}\right\}^{p_n} K_{h_n}(x-X_i)}{\displaystyle\sum_{i=1}^{n} K_{h_n}(x-X_i)}\right]^{1/p_n},$$

$$\mathcal{G}_n^{*}(x) = g(x)\left[\frac{1}{C(x)\,\alpha(x)\,B(p_n+1,\,\alpha(x))}\,\frac{\mathbb{E}\left(\left\{\frac{Y}{g(x)}\right\}^{p_n} K_{h_n}(x-X)\right)}{\mathbb{E}(K_{h_n}(x-X))}\right]^{1/p_n};$$

because the frontier of the random pair $(X,\,Y/g(x))$ is equal to 1 at x, it is enough to prove our result when $g(x)=1$.

We therefore start by showing that, as $n\to\infty$,

$$\frac{v_n(x)}{p_n}\left(\frac{\widehat{r}_n(x)}{\mathcal{R}_n(x)}-1\right)\xrightarrow{d}\mathcal{N}\left(0,\,\frac{2^{-\alpha(x)}\,\|K\|_2^2}{f(x)\,C(x)\,\Gamma(\alpha(x)+1)}\right). \tag{3.2}$$

To this end, write

$$\widehat{r}_n(x)-\mathcal{R}_n(x)=\frac{\mu_{p_n}(x)}{\widehat{f}_n(x)}\left[\frac{\widehat{\mu}_{p_n}(x)}{\mu_{p_n}(x)}-\frac{\widehat{f}_n(x)}{\mathbb{E}\left(\widehat{f}_n(x)\right)}\right]=\frac{1}{\sqrt{n}}\,\frac{\mu_{p_n}(x)}{\widehat{f}_n(x)}\left\{\frac{1}{\sqrt{n}}\sum_{k=1}^{n}\sigma_{n,\,k}(x)\right\} \tag{3.3}$$

where

$$\sigma_{n,\,k}(x)=\frac{Y_k^{p_n}\,K_{h_n}(x-X_k)}{\mu_{p_n}(x)}-\frac{K_{h_n}(x-X_k)}{\mathbb{E}\left(\widehat{f}_n(x)\right)}.$$

Because for all n, the $\sigma_{n,\,k}(x)$, $1\le k\le n$ are independent, identically distributed and have zero mean, to apply Theorem 2.7, it only remains to prove that

$$\frac{\mathbb{E}|\sigma_{n,\,1}(x)|^3}{\sqrt{n}[\mathrm{Var}(\sigma_{n,\,1}(x))]^{3/2}}\to 0$$

as $n\to\infty$. Since $\sigma_{n,\,1}(x)$ has zero mean, we have $\mathrm{Var}(\sigma_{n,\,1}(x))=\mathbb{E}|\sigma_{n,\,1}(x)|^2$, so that

$$\mathrm{Var}(\sigma_{n,\,1}(x))=\frac{\mathbb{E}\left(Y^{2p_n}\,K_{h_n}^2(x-X)\right)}{\mu_{p_n}^2(x)}+\frac{\mathbb{E}\left(K_{h_n}^2(x-X)\right)}{\left[\mathbb{E}\left(\widehat{f}_n(x)\right)\right]^2}-2\,\frac{\mathbb{E}\left(Y^{p_n}\,K_{h_n}^2(x-X)\right)}{\mu_{p_n}(x)\,\mathbb{E}\left(\widehat{f}_n(x)\right)}.$$

Now, for all $q\ge 1$, the change of variables $v=x-h_n u$ yields

$$\mathbb{E}\left(K_{h_n}^q(x-X)\right)=\int_\Omega K_{h_n}^q(x-v)\,f(v)\,dv=h_n^{-d(q-1)}\int_{\mathcal{B}} f(x-h_n u)\,K^q(u)\,du.$$

Using Lemma 3.1, we therefore obtain

$$\mathbb{E}\left(K_{h_n}^q(x-X)\right)=h_n^{-d(q-1)}\,f(x)\int_{\mathcal{B}} K^q(u)\,du\left[1+\mathrm{O}\left(h_n^{\eta_f}\right)\right].$$

In particular, $\mathbb{E}\left(\widehat{f}_n(x)\right)=f(x)\,(1+\mathrm{o}(1))$. Furthermore, applying Lemma 3.3 gives

$$\mathbb{E}\left(Y^{p_n}\,K_{h_n}^q(x-X)\right)=f(x)\,C(x)\,\Gamma(\alpha(x)+1)\,p_n^{-\alpha(x)}\,h_n^{-d(q-1)}\int_{\mathcal{B}} K^q(u)\,du\,(1+\mathrm{o}(1))$$

and especially

$$\mu_{p_n}(x) = f(x)\, C(x)\, \Gamma(\alpha(x)+1)\, p_n^{-\alpha(x)}\, (1 + \mathrm{o}(1)).$$

Thus

$$\mathrm{Var}(\sigma_{n,1}(x)) = \frac{2^{-\alpha(x)}\, \|K\|_2^2}{f(x)\, C(x)\, \Gamma(\alpha(x)+1)}\, p_n^{\alpha(x)}\, h_n^{-d}\, (1 + \mathrm{o}(1)). \tag{3.4}$$

It remains to control $\mathbb{E}|\sigma_{n,1}(x)|^3$. Hölder's inequality entails

$$\mathbb{E}|\sigma_{n,1}(x)|^3 \leq 4 \left[\frac{\mathbb{E}\left(K_{h_n}^3(x-X)\right)}{\left[\mathbb{E}\left(\widehat{f}_n(x)\right)\right]^3} + \frac{\mathbb{E}\left(Y^{3p_n}\, K_{h_n}^3(x-X)\right)}{\mu_{p_n}^3(x)} \right] = \mathrm{O}\left(p_n^{2\alpha(x)}\, h_n^{-2d}\right).$$

Consequently, as $n \to \infty$,

$$\frac{\mathbb{E}|\sigma_{n,1}(x)|^3}{\sqrt{n}[\mathrm{Var}(\sigma_{n,1}(x))]^{3/2}} = \mathrm{O}\left(\frac{1}{\sqrt{n}\, p_n^{-\alpha(x)/2}\, h_n^{d/2}} \right) \to 0$$

so that Theorem 2.7 gives

$$\frac{1}{\sqrt{\mathrm{Var}(\sigma_{n,1}(x))}} \left\{ \frac{1}{\sqrt{n}} \sum_{k=1}^{n} \sigma_{n,k}(x) \right\} \xrightarrow{d} \mathcal{N}(0,1).$$

Reporting in (3.3), recalling $\mathcal{R}_n(x) = \mu_{p_n}(x)/\mathbb{E}\left(\widehat{f}_n(x)\right)$ and $\mathbb{E}\left(\widehat{f}_n(x)\right) = f(x)\,(1+\mathrm{o}(1))$ yields

$$\sqrt{n} \frac{\widehat{f}_n(x)}{f(x)\, \sqrt{\mathrm{Var}(\sigma_{n,1}(x))}} \left[\frac{\widehat{r}_n(x)}{\mathcal{R}_n(x)} - 1 \right] \xrightarrow{d} \mathcal{N}(0,1).$$

Because $n\, h_n^d \to \infty$, it holds that $\widehat{f}_n(x) \xrightarrow{\mathbb{P}} f(x)$ (see [90]). Slutsky's lemma and (3.4) then yield (3.2). To conclude the proof, remark that

$$v_n(x) \ln\left(\frac{\widehat{g}_n^{GJ,*}(x)}{\mathcal{G}_n^*(x)} \right) = \frac{v_n(x)}{p_n} \ln\left(\frac{\widehat{r}_n(x)}{\mathcal{R}_n(x)} \right);$$

from (3.2), the delta-method yields

$$v_n(x) \ln\left(\frac{\widehat{g}_n^{GJ,*}(x)}{\mathcal{G}_n^*(x)} \right) \xrightarrow{d} \mathcal{N}\left(0,\, \frac{2^{-\alpha(x)}\, \|K\|_2^2}{f(x)\, C(x)\, \Gamma(\alpha(x)+1)}\right).$$

The result follows by using the delta-method once again. ■

From Proposition 3.1, we deduce that Girard and Jacob's estimator $\widehat{g}_n^{GJ}(x)$, centered on the true function $g(x)$, is *not* asymptotically normal if $Y \mid X = x$ is not uniform on $[0, g(x)]$:

Theorem 3.2. *Assume that* (P), (K) *and* $(A_1 - A_4)$ *hold. Let* $x \in \Omega$ *such that* $f(x) > 0$, *and assume that* $\alpha(x) \neq 1$ *or* $C(x) \neq 1$. *If* $n\, p_n^{-\alpha(x)}\, h_n^d \to \infty$ *and* $p_n\, h_n^{\eta_g} \to 0$, *then the decomposition*

$$\frac{\widehat{g}_n^{GJ}(x)}{g(x)} - 1 = \Xi_n(x) + w_n(x)$$

holds, where

$$v_n(x)\, \frac{\Xi_n(x)}{g(x)} \xrightarrow{d} \mathcal{N}\left(0,\, \frac{2^{-\alpha(x)}\, \|K\|_2^2}{f(x)\, C(x)\, \Gamma(\alpha(x)+1)}\right), \quad \text{as } n \to \infty$$

with $v_n(x) = \sqrt{n}\, p_n^{-\alpha(x)/2+1}\, h_n^{d/2}$ *as in Proposition 3.1 and* $(w_n(x))$ *is a deterministic sequence such that* $w_n(x) \to 0$ *and* $v_n(x)|w_n(x)| \to \infty$.

Proof. Let first $\mathcal{G}_n(x) = \{(p_n+1)\,\mathcal{R}_n(x)\}^{1/p_n}$ and remark that

$$\frac{\widehat{g}_n^{GJ}(x)}{\mathcal{G}_n(x)} = \frac{\widehat{g}_n^{GJ,*}(x)}{\mathcal{G}_n^*(x)}.$$

Consequently, Proposition 3.1 yields

$$v_n(x)\left(\frac{\widehat{g}_n^{GJ}(x)}{\mathcal{G}_n(x)} - 1\right) \xrightarrow{d} \mathcal{N}\left(0,\, \frac{2^{-\alpha(x)}\,\|K\|_2^2}{f(x)\,C(x)\,\Gamma(\alpha(x)+1)}\right), \quad \text{as } n\to\infty.$$

Let then

$$\Xi_n(x) = \mathcal{G}_n(x)\left[\frac{\widehat{g}_n^{GJ}(x)}{\mathcal{G}_n(x)} - 1\right] \quad \text{and} \quad w_n(x) = \mathcal{G}_n(x) - g(x).$$

We have

$$\widehat{g}_n^{GJ}(x) - g(x) = \Xi_n(x) + w_n(x),$$

and Slutsky's lemma yields

$$v_n(x)\,\frac{\Xi_n(x)}{g(x)} \xrightarrow{d} \mathcal{N}\left(0,\, \frac{2^{-\alpha(x)}\,\|K\|_2^2}{f(x)\,C(x)\,\Gamma(\alpha(x)+1)}\right), \quad \text{as } n\to\infty.$$

Notice now that

$$\mathcal{G}_n(x) = \{(p_n+1)\,C(x)\,\alpha(x)\,B(p_n+1,\,\alpha(x))\}^{1/p_n}\;\mathcal{G}_n^*(x),$$

which gives the representation $w_n(x) = w_n^{(1)}(x) + w_n^{(2)}(x)$, where

$$\begin{aligned} w_n^{(1)}(x) &:= \mathcal{G}_n^*(x) - g(x), \\ w_n^{(2)}(x) &:= \left[\{(p_n+1)\,C(x)\,\alpha(x)\,B(p_n+1,\,\alpha(x))\}^{1/p_n} - 1\right]\mathcal{G}_n^*(x). \end{aligned}$$

We shall first control $w_n^{(1)}(x)$. Use Lemmas 3.3 and 3.4 together with the estimation $\mathbb{E}\left(\widehat{f}_n(x)\right) = f(x) + \mathrm{O}\left(h_n^{\eta_f}\right)$ to get

$$\mathcal{R}_n(x) = g^{p_n}(x)\,C(x)\,\alpha(x)\,B(p_n+1,\,\alpha(x))\left[1 + \mathrm{O}\left(p_n\,h_n^{\eta_g} + \ln(p_n)\,h_n^{\eta_\alpha} + h_n^{\eta_f} + h_n^{\eta_C} + p_n^{-\beta(x)}\right)\right].$$

Consequently,

$$\begin{aligned} \mathcal{G}_n^*(x) &= \left\{\frac{1}{C(x)\,\alpha(x)\,B(p_n+1,\,\alpha(x))}\,\mathcal{R}_n(x)\right\}^{1/p_n} \\ &= g(x) + \mathrm{O}\left(h_n^{\eta_g} + \frac{\ln p_n}{p_n}h_n^{\eta_\alpha} + \frac{h_n^{\eta_f}}{p_n} + \frac{h_n^{\eta_C}}{p_n} + p_n^{-\beta(x)-1}\right) \end{aligned}$$

and therefore

$$w_n^{(1)}(x) = \mathrm{O}\left(h_n^{\eta_g} + \frac{\ln p_n}{p_n}h_n^{\eta_\alpha} + \frac{h_n^{\eta_f}}{p_n} + \frac{h_n^{\eta_C}}{p_n} + p_n^{-\beta(x)-1}\right). \tag{3.5}$$

We now turn to $w_n^{(2)}(x)$. [103] yields

$$(p_n+1)\,C(x)\,\alpha(x)\,B(p_n+1,\,\alpha(x)) = C(x)\,\Gamma(\alpha(x)+1)\,p_n^{-\alpha(x)+1}\left[1 + \mathrm{O}\left(\frac{1}{p_n}\right)\right]$$

and particularly

$$w_n^{(2)}(x) = \mathcal{G}_n^*(x)\left\{(1-\alpha(x))\frac{\ln p_n}{p_n} + \frac{\ln(C(x)\,\Gamma(\alpha(x)+1))}{p_n} + \mathrm{o}\left(\frac{1}{p_n}\right)\right\}. \tag{3.6}$$

Collecting (3.5) and (3.6), provided $\alpha(x) \neq 1$ or $C(x) \neq 1$, one has

$$v_n(x)\,|w_n(x)| = v_n(x)\left|w_n^{(2)}(x)\right|(1+\mathrm{o}(1))$$

and thus

$$v_n(x)\,|w_n(x)| = g(x)\frac{v_n(x)}{p_n}\,|(1-\alpha(x))\ln(p_n) + \ln(C(x)\,\Gamma(\alpha(x)+1))|\,(1+\mathrm{o}(1)) \to \infty,$$

which ends the proof. ■

3.3 An estimation procedure using high order moments

We now introduce an estimator based on a kernel regression on high order moments of the variable of interest Y. The estimator is given by

$$\frac{1}{\widehat{g}_n(x)} = \frac{1}{ap_n}\left[((a+1)p_n+1)\frac{\widehat{\mu}_{(a+1)p_n}(x)}{\widehat{\mu}_{(a+1)p_n+1}(x)} - (p_n+1)\frac{\widehat{\mu}_{p_n}(x)}{\widehat{\mu}_{p_n+1}(x)}\right] \tag{3.7}$$

where (p_n) is a nonrandom positive sequence such that $p_n \to \infty$ and $a > 0$. From a practical point of view, the use of a small window-width h_n allows to select the pairs (X_i, Y_i) such that X_i is close to x while the use of the high power p_n gives (exponentially) more weight to the Y_i close to $g(x)$. Moreover, our estimator has an explicit formulation, and the kernel regression enables us to avoid the partitioning of S. Let us also highlight that, compared to the estimator $\widehat{g}_n^{GJ,*}(x)$ suggested in (3.1), our proposition (3.7) does not require the knowledge of the conditional extreme-value index. The asymptotic properties of $\widehat{g}_n(x)$ are investigated under two different types of assumptions. The first ones are nonparametric assumptions on the conditional survival function $\overline{F}(\cdot\,|\,x)$ of Y given $X = x$: it is assumed that

(NP_1) Given $X = x$, Y is positive and has a finite right endpoint $g(x)$.

(NP_2) The conditional survival function $\overline{F}(\cdot\,|\,x)$ of Y given $X = x$ satisfies

$$\forall\, x \in \Omega, \quad \exists\, y_0 \in (0,\,1), \qquad \sup_{y\in[y_0,\,1]}\,\sup_{u\in\mathcal{B}}\left|\frac{\overline{F}(g(x-h_n u)y\,|\,x-h_n u)}{\overline{F}(g(x)y\,|\,x)} - 1\right| \to 0 \ \text{ as } \ n \to \infty.$$

We shall show in Section 3.3.2 that, under (NP_1–NP_2), the estimator $\widehat{g}_n(x)$ converges in probability to $g(x)$ without any parametric assumption nor on the distribution of X neither on the distribution of Y given $X = x$. Remark that, although our estimator $\widehat{g}_n(x)$ is based on a kernel regression, classical results do not apply (see for instance Ferraty and Vieu [37, Theorem 6.11]) since the condition $p_n \to \infty$ induces technical difficulties.

Note that (P) and $(A_1 - A_4)$ imply $(NP_1 - NP_2)$; in Section 3.3.3, the estimator is proven to be asymptotically Gaussian in the parametric framework (P). The results and their proofs are similar to what was obtained in Chapter 2.

Some simulations are proposed in Section 3.4 to illustrate the efficiency of our estimator and to compare it with some estimators of the frontier estimation literature, particularly $\widehat{g}_n^{GJ}(x)$. Auxiliary results are postponed to Appendix A and proven in Appendix B.

3.3.1 Construction of the estimator

To motivate the construction of our estimator of $g(x)$, let us first focus on the parametric setting (P) and especially on the case where $C = 1$ and $D = 0$. Let $x \in \Omega$ and consider the conditional moment

$$m_{p_n}(x) = p_n \int_0^\infty y^{p_n - 1}\, \overline{F}(y \,|\, x)\, dy = \alpha(x)\ g^{p_n}(x)\ B(p_n + 1,\, \alpha(x)).$$

Therefore, as in Section 2.2.1,

$$\frac{m_{p_n}(x)}{m_{p_n+1}(x)} = \frac{1}{g(x)} \left(1 + \frac{\alpha(x)}{p_n + 1} \right) \tag{3.8}$$

which leads to, for every nonrandom sequence (p_n) of real numbers ≥ 1 and every $a > 0$,

$$\frac{1}{g(x)} = \frac{1}{a p_n} \left[((a+1)p_n + 1) \frac{m_{(a+1)p_n}(x)}{m_{(a+1)p_n+1}(x)} - (p_n + 1) \frac{m_{p_n}(x)}{m_{p_n+1}(x)} \right].$$

On the basis of this result, the estimator of $g(x)$ is built in two steps. First, the conditional moment $m_{p_n}(x)$ is replaced by its smoothed version $\mu_{p_n}(x)$, and we set

$$\frac{1}{G_n(x)} := \frac{1}{a p_n} \left[((a+1)p_n + 1) \frac{\mu_{(a+1)p_n}(x)}{\mu_{(a+1)p_n+1}(x)} - (p_n + 1) \frac{\mu_{p_n}(x)}{\mu_{p_n+1}(x)} \right],$$

where $p_n \to \infty$. Second, $\mu_{p_n}(x)$ is estimated by the corresponding empirical moment $\widehat{\mu}_{p_n}(x)$. Plugging $\widehat{\mu}_{p_n}(x)$ in $1/G_n(x)$ leads to the expression (3.7) of the estimator $1/\widehat{g}_n(x)$ of $1/g(x)$.

3.3.2 Consistency

In this section, the consistency of $\widehat{g}_n(x)$ is established in the nonparametric context (NP). To this end, the first step is to prove that an analogue of (3.8) still holds, up to an error term, when $m_{p_n}(x)$ is replaced by $\mu_{p_n}(x)$.

Proposition 3.3. *Assume that $(NP_1 - NP_2)$, (K) and (A_1) hold. Let $x \in \Omega$ such that $f(x) > 0$. If $p_n\, h_n^{\eta_g} \to 0$, then*

$$\frac{\mu_{p_n}(x)}{\mu_{p_n+1}(x)} \to \frac{1}{g(x)} \qquad \text{as } n \to \infty.$$

This result is a straightforward consequence of Lemma 3.2i) and ii). The second step consists in showing that $\mu_{p_n}(x)$ can be replaced by its empirical counterpart $\widehat{\mu}_{p_n}(x)$. The following result holds:

Proposition 3.4. *Assume that* $(NP_1 - NP_2)$, (K) *and* (A_1) *hold. Let* $x \in \Omega$ *such that* $f(x) > 0$. *If* $n\, m_{p_n}(x)\, h_n^d / g^{p_n}(x) \to \infty$ *and* $p_n\, h_n^{\eta_g} \to 0$ *as* $n \to \infty$, *then*

$$\frac{\widehat{\mu}_{p_n}(x)}{\mu_{p_n}(x)} \xrightarrow{\mathbb{P}} 1 \qquad \text{as } n \to \infty.$$

Proof. We follow the proof of Proposition 2.3. Let, for all $1 \leq j \leq n$,

$$U_{nj} = \frac{Y_j^{p_n}\, K_{h_n}(x - X_j)}{n\, \mu_{p_n}(x)}.$$

The desired result is then tantamount to $\sum_{j=1}^n U_{nj} \xrightarrow{\mathbb{P}} 1$ as $n \to \infty$. Let us highlight that, for all n, the U_{nj}, $1 \leq j \leq n$ are nonnegative independent random variables, and $\sum_{j=1}^n \mathbb{E}(U_{nj}) = 1$. Using Theorem 2.2 makes it enough to show that

$$\forall \varepsilon > 0, \quad \sum_{j=1}^n \mathbb{E}\left(U_{nj}\, \mathbb{1}_{\{U_{nj} \geq \varepsilon\}}\right) \to 0$$

as $n \to \infty$. Remark that the U_{nj} can be rewritten as

$$U_{nj} = \frac{V_{nj}\, K_{h_n}(x - X_j)}{n\, M_{p_n}(x)}$$

where

$$V_{nj} = \frac{Y_j^{p_n}}{\sup_{u \in \mathcal{B}} g^{p_n}(x - h_n u)} \quad \text{and} \quad M_{p_n}(x) = \frac{\mu_{p_n}(x)}{\sup_{u \in \mathcal{B}} g^{p_n}(x - h_n u)}.$$

The U_{nj}, $1 \leq j \leq n$ being identically distributed, it is equivalent to prove that, for all $\varepsilon > 0$,

$$\frac{1}{M_{p_n}(x)} \mathbb{E}\left(V_{n1}\, K_{h_n}(x - X)\, \mathbb{1}_{\{V_{n1}\, K_{h_n}(x - X) \geq \varepsilon n\, M_{p_n}(x)\}}\right) \to 0.$$

Pick then $\varepsilon > 0$ and notice that

$$V_{n1}\, K_{h_n}(x - X) \geq \varepsilon n\, M_{p_n}(x) \Leftrightarrow h_n^d\, V_{n1}\, K_{h_n}(x - X) \geq \varepsilon n\, M_{p_n}(x)\, h_n^d. \tag{3.9}$$

The left-hand side of the second inequality is nonnegative and bounded by $\sup_{\mathbb{R}^d} K$. In view of Lemma 3.2ii), condition $n\, m_{p_n}(x)\, h_n^d / g^{p_n}(x) \to \infty$ is equivalent to $n\, \mu_{p_n}(x)\, h_n^d / g^{p_n}(x) \to \infty$. Besides, $p_n\, h_n^{\eta_g} \to 0$ and (3.26) in the proof of Lemma 3.2ii) imply that

$$\sup_{u \in \mathcal{B}} \frac{g^{p_n}(x - h_n u)}{g^{p_n}(x)} \to 1$$

so that $n\, M_{p_n}(x)\, h_n^d \to \infty$ as $n \to \infty$. As a consequence, the right-hand side of (3.9) goes to ∞: therefore

$$\frac{1}{M_{p_n}(x)} \mathbb{E}\left(V_{n1}\, K_{h_n}(x - X)\, \mathbb{1}_{\{h_n^d\, V_{n1}\, K_{h_n}(x - X) \geq \varepsilon n\, M_{p_n}(x)\, h_n^d\}}\right) = 0$$

eventually, and the result is proven. ∎

As a consequence of the two previous results, we have:

Theorem 3.5. *Assume that* $(NP_1 - NP_2)$, (K) *and* (A_1) *hold. Let* $x \in \Omega$ *such that* $f(x) > 0$. *If* $n\, m_{(a+1)p_n}(x)\, h_n^d / g^{(a+1)p_n}(x) \to \infty$ *and* $p_n\, h_n^{\eta_g} \to 0$, *then* $\widehat{g}_n(x) \xrightarrow{\mathbb{P}} g(x)$ *as* $n \to \infty$.

Proof. Note that, according to the proof of Theorem 2.4,

$$\frac{m_{(a+1)p_n}(x)}{g^{(a+1)p_n}(x)} \leq (a+1)\frac{m_{p_n}(x)}{g^{p_n}(x)},$$

which implies $n\, m_{p_n}(x)\, h_n^d / g^{p_n}(x) \to \infty$. Thus, Lemma 3.2i) entails

$$n\,\frac{m_{p_n+1}(x)}{g^{p_n+1}(x)}\, h_n^d \to \infty \quad \text{and} \quad n\,\frac{m_{(a+1)p_n+1}(x)}{g^{(a+1)p_n+1}(x)}\, h_n^d \to \infty \qquad \text{as } n \to \infty.$$

We can then apply Proposition 3.4 to rewrite the frontier estimator as:

$$\frac{1}{\widehat{g}_n(x)} = \frac{1}{ap_n}\left[((a+1)p_n+1)\frac{\mu_{(a+1)p_n}(x)}{\mu_{(a+1)p_n+1}(x)}(1+\mathrm{o}_{\mathbb{P}}(1)) - (p_n+1)\frac{\mu_{p_n}(x)}{\mu_{p_n+1}(x)}(1+\mathrm{o}_{\mathbb{P}}(1))\right]. \quad (3.10)$$

From Proposition 3.3, we have

$$\frac{\mu_{p_n}(x)}{\mu_{p_n+1}(x)} \to \frac{1}{g(x)} \quad \text{and} \quad \frac{\mu_{(a+1)p_n}(x)}{\mu_{(a+1)p_n+1}(x)} \to \frac{1}{g(x)}$$

as $n \to \infty$. Replacing in (3.10), the conclusion follows. ■

3.3.3 Asymptotic normality

We now establish the asymptotic distribution of $\widehat{g}_n(x)$ under the assumption (P). The parametric model enables us to compute a more precise asymptotic expansion of $\mu_{p_n}(x)/\mu_{p_n+1}(x)$ than under the nonparametric assumption, see Proposition 3.3.

Proposition 3.6. *Assume that* (P), (K) *and* $(A_1 - A_4)$ *hold. Let* $x \in \Omega$ *such that* $f(x) > 0$. *If* $p_n\, h_n^{\eta_g} \to 0$, *then*

$$\frac{\mu_{p_n}(x)}{\mu_{p_n+1}(x)} = \frac{1}{g(x)}\left[1 + \frac{\alpha(x)}{p_n+1}\right] + \mathrm{O}\left(h_n^{\eta_g} + \frac{h_n^{\eta_\alpha}}{p_n} + p_n^{-\beta(x)-1}\right).$$

Proof. Remark that, retaining notations of Lemma 3.3, we have

$$\frac{\mathcal{L}_n^{(\alpha,\,C)}(p_n+1,\,x,\,u)}{\mathcal{L}_n^{(\alpha,\,C)}(p_n,\,x,\,u)} \;=\; 1 + \frac{\Delta_n^g(x,\,u)}{g(x)} - \frac{\Delta_n^\alpha(x,\,u)}{p_n} + \mathrm{O}\left(\frac{h_n^{\eta_g}}{p_n} + \frac{h_n^{\eta_\alpha}}{p_n^2}\right)$$

uniformly in $u \in \mathcal{B}$. Using Lemma 3.3ii) with $q = 1$ then entails

$$\left\{\frac{1}{g(x)}\left[1 + \frac{\alpha(x)}{p_n+1}\right]\right\}^{-1} \frac{M_n^{(\alpha,\,C)}(1,\,p_n,\,x)}{M_n^{(\alpha,\,C)}(1,\,p_n+1,\,x)}$$

$$= 1 + \frac{\displaystyle\int_{\mathcal{B}} \mathcal{L}_n^{(\alpha,\,C)}(p_n,\,x,\,u)\left[\frac{\Delta_n^\alpha(x,\,u)}{p_n} - \frac{\Delta_n^g(x,\,u)}{g(x)}\right] K(u)\,du}{\displaystyle\int_{\mathcal{B}} \mathcal{L}_n^{(\alpha,\,C)}(p_n,\,x,\,u)\,K(u)\,du} + \mathrm{O}\left(\frac{h_n^{\eta_g}}{p_n} + \frac{h_n^{\eta_\alpha}}{p_n^2}\right).$$

From Lemma 3.3i), $\mathcal{L}_n^{(\alpha,\,C)}(p_n,\,x,\,u) \to 1$ as $n \to \infty$ uniformly in $u \in \mathcal{B}$, so that

$$\frac{\displaystyle\int_{\mathcal{B}} \mathcal{L}_n^{(\alpha,\,C)}(p_n,\,x,\,u)\left[\frac{\Delta_n^{\alpha}(x,\,u)}{p_n} - \frac{\Delta_n^{g}(x,\,u)}{g(x)}\right] K(u)\,du}{\displaystyle\int_{\mathcal{B}} \mathcal{L}_n^{(\alpha,\,C)}(p_n,\,x,\,u)\,K(u)\,du} = \mathrm{O}\left(h_n^{\eta_g} + \frac{h_n^{\eta_\alpha}}{p_n}\right).$$

Besides, applying Lemma 3.4 yields $\tau_n(1,\,p_n,\,x) = \mathrm{O}\left(p_n^{-\beta(x)-1}\right)$; use then (3.23) to obtain

$$\frac{\mu_{p_n}(x)}{\mu_{p_n+1}(x)} = \frac{1}{g(x)}\left[1 + \frac{\alpha(x)}{p_n+1}\right] + \mathrm{O}\left(h_n^{\eta_g} + \frac{h_n^{\eta_\alpha}}{p_n} + p_n^{-\beta(x)-1}\right)$$

which completes the proof of Proposition 3.6. ■

As a straightforward consequence, we obtain a control of the bias introduced by replacing $m_{p_n}(x)$ by $\mu_{p_n}(x)$. If $p_n\,h_n^{\eta_g} \to 0$, then

$$\frac{1}{G_n(x)} = \frac{1}{g(x)} + \mathrm{O}\left(h_n^{\eta_g} + \frac{h_n^{\eta_\alpha}}{p_n} + p_n^{-\beta(x)-1}\right). \tag{3.11}$$

Let us now turn to the random term:

Theorem 3.7. *Assume that (P), (K) and $(A_1 - A_4)$ hold. Let $x \in \Omega$ such that $f(x) > 0$. If $n\,p_n^{-\alpha(x)}\,h_n^d \to \infty$ and $p_n\,h_n^{\eta_g} \to 0$ then*

$$v_n(x)\left(\frac{\widehat{g}_n(x)}{G_n(x)} - 1\right) \xrightarrow{d} \mathcal{N}\left(0,\,\frac{\|K\|_2^2\,V(\alpha(x),\,a)}{f(x)\,C(x)}\right), \qquad \text{as } n \to \infty$$

where $v_n(x) = \sqrt{n}\,p_n^{-\alpha(x)/2+1}\,h_n^{d/2}$ as in Proposition 3.1, and

$$V(\alpha(x),\,a) = \frac{\alpha(x)+1}{a^2\,\Gamma(\alpha(x))}\left[2^{-\alpha(x)-2} - 2\frac{(a+1)^{\alpha(x)+1}}{(a+2)^{\alpha(x)+2}} + 2^{-\alpha(x)-2}(a+1)^{\alpha(x)}\right]$$

as in Theorem 2.8.

Proof. The proof is an adaptation of the one of Theorem 2.8. Our goal is to prove that the sequence of random variables

$$\xi_n(x) = \frac{g(x)}{\|K\|_2}\sqrt{\frac{f(x)\,C(x)}{V(\alpha(x),\,a)}}\;v_n(x)\left(\frac{1}{\widehat{g}_n(x)} - \frac{1}{G_n(x)}\right)$$

converges in distribution to a standard Gaussian random variable. The first step consists to use Lemma 3.5 in order to linearise $\xi_n(x)$:

$$\begin{aligned}\xi_n(x) &= \left[\zeta_n^{(1)}(x) + \left(\frac{\mu_{p_n+1}(x)}{\widehat{\mu}_{p_n+1}(x)} - 1\right)\zeta_n^{(2)}(x) + \left(1 + \frac{ap_n}{p_n+1}\right)\left(\frac{\mu_{(a+1)p_n+1}(x)}{\widehat{\mu}_{(a+1)p_n+1}(x)} - 1\right)\zeta_n^{(3)}(x)\right]\\ &\times\; u_{n,\,a}(x)\,(1+\mathrm{o}(1)).\end{aligned}$$

Now, Proposition 3.4 yields

$$\xi_n(x) = u_{n,\,a}(x)\left[\zeta_n^{(1)}(x) + \mathrm{o}_{\mathbb{P}}\left(\zeta_n^{(2)}(x)\right) + \mathrm{o}_{\mathbb{P}}\left(\zeta_n^{(3)}(x)\right)\right](1+\mathrm{o}(1))$$

and to conclude the proof, it is enough to establish that

$$u_{n,\,a}(x)\,\zeta_n^{(1)}(x) \xrightarrow{d} \mathcal{N}(0,1), \tag{3.12a}$$

$$u_{n,\,a}(x)\,\zeta_n^{(2)}(x) \xrightarrow{d} \mathcal{N}(0,\,C_2), \tag{3.12b}$$

$$u_{n,\,a}(x)\,\zeta_n^{(3)}(x) \xrightarrow{d} \mathcal{N}(0,\,C_3), \tag{3.12c}$$

where C_2 and C_3 are positive constants. In all the sequel, we set

$$Z_k^{(n,\,s,\,j)}(x) = Y_k^{sp_n+j}\,K_{h_n}(x - X_k),$$

so that $\mu_{sp_n+j}(x) = \mathbb{E}\left(Z^{(n,\,s,\,j)}(x)\right)$. To prove (3.12a), remark that $\zeta_n^{(1)}(x)$ can be expanded as the sum of independent, identically distributed and centered random variables: $\zeta_n^{(1)}(x) = \sum_{k=1}^n S_{n,\,k}^{(1)}(x)$ with

$$\begin{aligned}
S_{n,\,k}^{(1)}(x) &= \frac{1}{n}\left\{A_n^{(1)}(x)\right\}^t\left[Z_k^{(n,\,1,\,0)}(x),\; Z_k^{(n,\,1,\,1)}(x),\; Z_k^{(n,\,a+1,\,0)}(x),\; Z_k^{(n,\,a+1,\,1)}(x)\right]^t,\\
A_n^{(1)}(x) &= \left[a_{n,\,0}^{(1)}(x),\; a_{n,\,1}^{(1)}(x),\; a_{n,\,2}^{(1)}(x),\; a_{n,\,3}^{(1)}(x)\right]^t,\\
a_{n,\,0}^{(1)}(x) &= -1,\\
a_{n,\,1}^{(1)}(x) &= \frac{\mu_{p_n}(x)}{\mu_{p_n+1}(x)},\\
a_{n,\,2}^{(1)}(x) &= \left(1 + \frac{ap_n}{p_n+1}\right)\frac{\mu_{p_n+1}(x)}{\mu_{(a+1)p_n+1}(x)},\\
a_{n,\,3}^{(1)}(x) &= -\left(1 + \frac{ap_n}{p_n+1}\right)\frac{\mu_{p_n+1}(x)\,\mu_{(a+1)p_n}(x)}{\mu^2_{(a+1)p_n+1}(x)}.
\end{aligned}$$

In order to use Theorem 2.7, it remains to prove that

$$\frac{\mathbb{E}\left|S_{n,\,1}^{(1)}(x)\right|^3}{\sqrt{n}\left[\operatorname{Var}\left(S_{n,\,1}^{(1)}(x)\right)\right]^{3/2}} \to 0 \tag{3.13}$$

as $n \to \infty$, which requires to control $\operatorname{Var}\left(S_{n,\,1}^{(1)}(x)\right)$ and $\mathbb{E}\left|S_{n,\,1}^{(1)}(x)\right|^3$. The variance is rewritten as

$$\begin{aligned}
\operatorname{Var}\left(S_{n,\,1}^{(1)}(x)\right) = \frac{1}{n^2}\Bigg[& w(1,\,1,\,p_n)(x) - 2\left(1 + \frac{ap_n}{p_n+1}\right)\frac{\mu_{p_n+1}(x)}{\mu_{(a+1)p_n+1}(x)}\,w(1,\,a+1,\,p_n)(x)\\
&+ \left(1 + \frac{ap_n}{p_n+1}\right)^2\frac{\mu^2_{p_n+1}(x)}{\mu^2_{(a+1)p_n+1}(x)}\,w(a+1,\,a+1,\,p_n)(x)\Bigg]
\end{aligned}$$

where, in the same fashion as in Lemma 2.14,

$$w(s,\,t,\,p_n)(x) = \left[-1,\; \frac{\mu_{sp_n}(x)}{\mu_{sp_n+1}(x)}\right]\mathcal{M}_n(s,\,t)(x)\left[-1,\; \frac{\mu_{tp_n}(x)}{\mu_{tp_n+1}(x)}\right]^t$$

with $\mathcal{M}_n(s,\,t)(x)$ being the 2×2 covariance matrix defined by

$$\mathcal{M}_n(s,\,t)(x) = \begin{bmatrix} \mathbb{E}\left(Z^{(n,\,s,\,0)}(x)\,Z^{(n,\,t,\,0)}(x)\right) & \mathbb{E}\left(Z^{(n,\,s,\,0)}(x)\,Z^{(n,\,t,\,1)}(x)\right)\\ \mathbb{E}\left(Z^{(n,\,s,\,1)}(x)\,Z^{(n,\,t,\,0)}(x)\right) & \mathbb{E}\left(Z^{(n,\,s,\,1)}(x)\,Z^{(n,\,t,\,1)}(x)\right)\end{bmatrix}.$$

Since Lemma 3.3iii) provides an asymptotic expansion of the matrix $\mathcal{M}_n(s,\,t)(x)$, it is therefore sufficient to compute an asymptotic expansion of $\mu_{p_n}(x)/\mu_{p_n+1}(x)$. Using Proposition 3.6, tedious computations lead to

$$\mathrm{Var}\left(S_{n,\,1}^{(1)}(x)\right) = a^2\,\|K\|_2^2\, f(x)\, C(x)\, \Gamma^2(\alpha(x)+1)\, V(\alpha(x),\,a)\, \frac{g^{2p_n}(x)\, p_n^{-\alpha(x)-2}}{n^2\, h_n^d}\,(1+\mathrm{o}(1)). \quad (3.14)$$

Now, focusing on the third-order moment, Hölder's inequality yields

$$\begin{aligned} n^3\,\mathbb{E}\left|S_{n,\,1}^{(1)}(x)\right|^3 &\leq 4\,\mathbb{E}\left|a_{n,\,0}^{(1)}(x)\, Z_1^{(n,\,1,\,0)}(x) + a_{n,\,1}^{(1)}(x)\, Z_1^{(n,\,1,\,1)}(x)\right|^3 \\ &+ 4\,\mathbb{E}\left|a_{n,\,2}^{(1)}(x)\, Z_1^{(n,\,a+1,\,0)}(x) + a_{n,\,3}^{(1)}(x)\, Z_1^{(n,\,a+1,\,1)}(x)\right|^3. \end{aligned}$$

The next step consists in applying Lemma 3.6 to each term of the right-hand side of this inequality. To this end, let us consider the functions

$$\begin{aligned} H_{n,\,0}^{(1)}(y) &= -1, \\ H_{n,\,1}^{(1)}(y) &= \alpha(x)y, \\ H_{n,\,2}^{(1)}(y) &= \left(1+\frac{ap_n}{p_n+1}\right) g^{ap_n}(x)\,\frac{\mu_{p_n+1}(x)}{\mu_{(a+1)p_n+1}(x)}, \\ H_{n,\,3}^{(1)}(y) &= -\left(1+\frac{ap_n}{p_n+1}\right) g^{ap_n}(x)\,\frac{\mu_{p_n+1}(x)}{\mu_{(a+1)p_n+1}(x)}\,\frac{\alpha(x)y}{a+1}, \end{aligned}$$

and note that there exist two sequences of Borel functions $\left(\chi_n^{(1)}\right)$ and $\left(\chi_n^{(2)}\right)$ uniformly convergent to 0 on $[0,\,1]$ such that

$$\begin{aligned} \sup_{u\in\mathcal{B}}\left|a_{n,\,0}^{(1)}(x) + a_{n,\,1}^{(1)}(x)\, g(x-h_n u)\, y\right| &\leq \left|H_{n,\,0}^{(1)}(y)\right|(1-y) + \frac{\left|H_{n,\,1}^{(1)}(y)\right| + \chi_n^{(1)}(y)}{p_n}, \\ \sup_{u\in\mathcal{B}}\left|a_{n,\,2}^{(1)}(x) + a_{n,\,3}^{(1)}(x)\, g(x-h_n u)\, y\right| &\leq \frac{1}{g^{ap_n}(x)}\left[\left|H_{n,\,2}^{(1)}(y)\right|(1-y) + \frac{\left|H_{n,\,3}^{(1)}(y)\right| + \chi_n^{(2)}(y)}{p_n}\right]. \end{aligned}$$

Since

$$g^{ap_n}(x)\frac{\mu_{p_n+1}(x)}{\mu_{(a+1)p_n+1}(x)} \to (a+1)^{\alpha(x)}$$

as $n\to\infty$, the functions $H_{n,\,j}^{(1)}$, $j\in\{0,\,1,\,2,\,3\}$ are bounded on $[0,\,1]$, uniformly in n, and thus Lemma 3.6 entails that

$$\mathbb{E}\left|S_{n,\,1}^{(1)}(x)\right|^3 = \mathrm{O}\left(n^{-3}\, g^{3p_n}(x)\, p_n^{-\alpha(x)-3}\, h_n^{-2d}\right). \quad (3.15)$$

Combining (3.14) and (3.15), convergence (3.13) follows from the condition $n\,p_n^{-\alpha(x)}\,h_n^d\to\infty$ and therefore (3.12a) holds.

Proofs of (3.12b) and (3.12c) are similar since $\zeta_n^{(2)}(x)$ and $\zeta_n^{(3)}(x)$ can be rewritten as

$$\begin{aligned} \zeta_n^{(2)}(x) &= \sum_{k=1}^n S_{n,\,k}^{(2)}(x) \quad \text{with} \quad S_{n,\,k}^{(2)}(x) = \frac{1}{n}\sum_{k=1}^n \left[a_{n,\,0}^{(2)}(x),\; a_{n,\,1}^{(2)}(x)\right]\left[Z_k^{(n,\,1,\,0)}(x),\; Z_k^{(n,\,1,\,1)}(x)\right]^t, \\ \zeta_n^{(3)}(x) &= \sum_{k=1}^n S_{n,\,k}^{(3)}(x) \quad \text{with} \quad S_{n,\,k}^{(3)}(x) = \frac{1}{n}\sum_{k=1}^n \left[a_{n,\,0}^{(3)}(x),\; a_{n,\,1}^{(3)}(x)\right]\left[Z_k^{(n,\,a+1,\,0)}(x),\; Z_k^{(n,\,a+1,\,1)}(x)\right]^t \end{aligned}$$

where

$$\begin{aligned}
a_{n,0}^{(1)}(x) &= -1, \\
a_{n,1}^{(1)}(x) &= \frac{\mu_{p_n}(x)}{\mu_{p_n+1}(x)}, \\
a_{n,2}^{(1)}(x) &= \frac{\mu_{p_n+1}(x)}{\mu_{(a+1)p_n+1}(x)}, \\
a_{n,3}^{(1)}(x) &= -\frac{\mu_{p_n+1}(x)\,\mu_{(a+1)p_n}(x)}{\mu^2_{(a+1)p_n+1}(x)}.
\end{aligned}$$

Applying Lemma 3.6 with

$$\begin{aligned}
H_{n,0}^{(2)}(y) &= -1, \\
H_{n,1}^{(2)}(y) &= \alpha(x)y, \\
H_{n,0}^{(3)}(y) &= g^{ap_n}(x)\,\frac{\mu_{p_n+1}(x)}{\mu_{(a+1)p_n+1}(x)}, \\
H_{n,1}^{(3)}(y) &= -g^{ap_n}(x)\,\frac{\mu_{p_n+1}(x)}{\mu_{(a+1)p_n+1}(x)}\,\frac{\alpha(x)y}{a+1}
\end{aligned}$$

yields $\mathbb{E}\left|S_{n,1}^{(j)}(x)\right|^3 = \mathrm{O}\left(n^{-3}\,g^{3p_n}(x)\,p_n^{-\alpha(x)-3}\,h_n^{-2d}\right)$, $j \in \{2, 3\}$. Theorem 2.7 then gives (3.12b) and (3.12c): Theorem 3.7 is therefore established. ■

The asymptotic normality of $\widehat{g}_n(x)$ centered on the true function $g(x)$ is then readily obtained:

Theorem 3.8. *Assume that (P), (K) and $(A_1 - A_4)$ hold. Let $x \in \Omega$ such that $f(x) > 0$. If $n\,p_n^{-\alpha(x)}\,h_n^d \to \infty$, $n\,p_n^{-\alpha(x)-2\beta(x)}\,h_n^d \to 0$, $n\,p_n^{-\alpha(x)+2}\,h_n^{d+2\eta_g} \to 0$ and $n\,p_n^{-\alpha(x)}\,h_n^{d+2\eta_\alpha} \to 0$, then*

$$v_n(x)\left(\frac{\widehat{g}_n(x)}{g(x)} - 1\right) \xrightarrow{d} \mathcal{N}\left(0, \frac{\|K\|_2^2\,V(\alpha(x), a)}{f(x)\,C(x)}\right), \qquad \text{as } n \to \infty$$

with the notations of Theorem 3.7.

Proof. The proof is an analogue of that of Theorem 2.8: write

$$\frac{\widehat{g}_n(x)}{g(x)} - 1 = \frac{G_n(x)}{g(x)}\left[\frac{\widehat{g}_n(x)}{G_n(x)} - 1\right] + \left[\frac{G_n(x)}{g(x)} - 1\right].$$

Use the convergence $G_n(x)/g(x) \to 1$ as $n \to \infty$, Theorem 3.7 and Slutsky's lemma to get

$$\frac{G_n(x)}{g(x)}\,v_n(x)\left[\frac{\widehat{g}_n(x)}{G_n(x)} - 1\right] \xrightarrow{d} \mathcal{N}\left(0, \frac{\|K\|_2^2\,V(\alpha(x), a)}{f(x)\,C(x)}\right), \qquad \text{as } n \to \infty.$$

Besides, the convergences $n\,p_n^{-\alpha(x)-2\beta(x)}\,h_n^d \to 0$, $n\,p_n^{-\alpha(x)+2}\,h_n^{d+2\eta_g} \to 0$ and $n\,p_n^{-\alpha(x)}\,h_n^{d+2\eta_\alpha} \to 0$ along with (3.11) yield

$$v_n(x)\left[\frac{G_n(x)}{g(x)} - 1\right] \to 0$$

as $n \to \infty$. Using once again Slutsky's lemma completes the proof. ■

Let us note that $n\, p_n^{-\alpha(x)}\, h_n^d \to \infty$ and $n\, p_n^{-\alpha(x)+2}\, h_n^{d+2\eta_g} \to 0$ imply that $p_n\, h_n^{\eta_g} \to 0$. Besides, if we assume that α has greater regularity than g, namely $\eta_\alpha \geq \eta_g$, then the hypotheses needed to apply Theorem 3.8 can be reduced to $n\, p_n^{-\alpha(x)}\, h_n^d \to \infty$, $n\, p_n^{-\alpha(x)-2\beta(x)}\, h_n^d \to 0$ and $n\, p_n^{-\alpha(x)+2}\, h_n^{d+2\eta_g} \to 0$.

Let now $x \in \Omega$ such that $f(x) > 0$ and note that the sequences

$$h_n(x) = \varepsilon_n^{\alpha(x)-1}\, n^{-1/(d+\eta_g\alpha(x))} \quad \text{and} \quad p_n(x) = \varepsilon_n^{d+\eta_g}\, n^{\eta_g/(d+\eta_g\alpha(x))}$$

can be chosen to check the assumptions of Theorem 3.8, where (ε_n) is an arbitrary sequence of positive real numbers tending to 0 such that $n^{-\delta}\, \varepsilon_n \to 0$ for all $\delta > 0$. With such choices, the rate of convergence $v_n(x)$ of the estimator is then $n^{\eta_g/(d+\eta_g\alpha(x))}$ up to an ε_n term. In the uniform case (that is, when α is constant equal to 1), the rate of convergence of the estimator is then $n^{\eta_g/(d+\eta_g)}$, up to the factor ε_n, which is also the rate of convergence for the estimator $\widehat{g}_n^{GJ}(x)$. Let us note that this rate of convergence has been shown to be minimax in [63] for a particular class of densities with a L^1 risk.

Moreover, the rate of convergence $v_n(x)$ is the same for both our estimator and the estimator $\widehat{g}_n^{GJ,*}(x)$; note however that computing $\widehat{g}_n^{GJ,*}(x)$ requires the knowledge of the conditional extreme-value index $\alpha(x)$ and of the constant $C(x)$. From that point of view, our method is an improvement of Girard and Jacob's approach [46] since it does not demand a precise knowledge of the conditional distribution of Y given X.

3.4 A simulation study

The behaviour of the proposed frontier estimator is investigated on different situations. In particular, we examine the case $d = 1$ where X is uniformly distributed on $\Omega = [0, 1]$ and the case $d = 2$ where $X = (X_1, X_2)$ is uniformly distributed on $\Omega = [0, 1]^2$.

(i) Let us first focus on the case $d = 1$. Three frontiers are considered:

$$\begin{aligned}
g_1(x) &= \begin{cases} 1 + \exp\left(-60\,(x - 1/4)^2\right) & \text{if } x \in [0,\, 1/3]\,, \\ 1 + \exp(-5/12) & \text{if } x \in (1/3,\, 2/3]\,, \\ 1 + 5\exp(-5/12) - 6\exp(-5/12)\,x & \text{if } x \in (2/3,\, 5/6]\,, \\ 6x - 4 & \text{if } x \in (5/6,\, 1]\,, \end{cases} \\
g_2(x) &= \left[\frac{1}{10} + \sin(\pi x)\right]\left[\frac{11}{10} - \frac{1}{2}\exp\left(-64\,(x - 1/2)^2\right)\right], \\
g_3(x) &= \frac{5}{4} - 2x(1 - x).
\end{aligned}$$

Note that g_1 is continuous but not derivable at $x = 1/3$, $x = 2/3$ and $x = 5/6$ while g_2 and g_3 are infinitely derivable.

We consider three different models for the distribution of Y given $X = x$. The first one is the simple "power law" model

$$\overline{F}(y \,|\, x) = (1 - y/g(x))^{\alpha(x)}, \quad \forall\, y \in [0,\, g(x)]. \tag{3.16}$$

In this framework, we examine the cases $g = g_1$ and $g = g_2$; besides, two models for the function α are considered:

$$\alpha_1(x) = \frac{5}{4} \quad \text{and} \quad \alpha_2(x) = \frac{5}{4} + \frac{1}{2}|\cos(2\pi x)|.$$

The second model is a "contaminated" version of the previous one, namely

$$\overline{F}(y \,|\, x) = C(x)\,(1 - y/g(x))^{\alpha_2(x)} + (1 - C(x))\,(1 - y/g(x))^{\alpha_2(x)+1/4}, \quad \forall\, y \in [0,\, g(x)], \tag{3.17}$$

with $C(x) = c + \sin(2\pi x)/16$ and $c \in \{1/8,\, 3/8,\, 5/8,\, 7/8\}$. Let us highlight that the smaller c is, the larger the contamination is.

The last model we consider is

$$\overline{F}(y \,|\, x) = \frac{1}{B(a(x),\, b(x))} \int_y^{g(x)} \left[\frac{u}{g(x)}\right]^{a(x)-1} \left[1 - \frac{u}{g(x)}\right]^{b(x)-1} \frac{du}{g(x)}, \quad \forall\, y \in [0,\, g(x)], \tag{3.18}$$

where

$$a(x) = \frac{1}{2} + \frac{1}{4}\sin(\pi x), \quad b(x) = \frac{5}{4} + \beta|\cos(2\pi x)|, \quad \beta \in \{0,\, 1/2\}. \tag{3.19}$$

Namely, given $X = x$, the random variable $Y/g(x)$ is Beta$(a(x),\, b(x))$ distributed, and b is equal to either α_1 or α_2. This situation fits our parametric model (P), with $\alpha(x) = b(x)$ (see [11]). The cases $g = g_1$ and $g = g_2$ are tested.

The uniform kernel is chosen:

$$K(x) = \frac{1}{2}\mathbb{1}_{[-1,\,1]}(x)$$

with associated bandwidth $h_n^{(m)} = 2\widehat{\sigma}(X)/n^{1/(1+\alpha_\infty)}$ and $p_n^{(m)} = n^{1/(1+\alpha_\infty)}/\sqrt{\ln(n)}$, where $\widehat{\sigma}(X)$ is the empirical standard deviation of X and $\alpha_\infty = \max_\Omega \alpha$, which is finite since α is continuous and Ω is a compact subset of $\mathbb{R}$. These sequences are chosen to check the hypotheses of Theorem 3.7. Note that the multiplicative constant $\widehat{\sigma}(X)$ has been suggested by [46], whereas the constant 2 was empirically chosen. The sample size is fixed to $n = 500$.

(ii) In the case $d = 2$, we limit ourselves to a unique model

$$g(x,\, y) = 1 + \frac{3}{20}g_1(x)y, \quad \text{and} \quad \alpha(x,\, y) = \frac{5}{4} + \frac{1}{2}|\cos(2\pi x)\,\sin(2\pi y)|, \tag{3.20}$$

the kernel being

$$K(x,\, y) = \frac{1}{4}\mathbb{1}_{[-1,\,1]\times[-1,\,1]}(x,\, y),$$

with bandwidth $h_n^{(m)} = 4\sqrt{\widehat{\sigma}(X_1)\widehat{\sigma}(X_2)}/n^{1/(2+\alpha_\infty)}$ and $p_n^{(m)} = n^{1/(2+\alpha_\infty)}/\sqrt{\ln(n)}$. The sample size is fixed to $n = 1000$.

In all cases, our moment estimator is computed with $a = 15$, the constant a having been chosen after intensive simulations. Our estimator is compared to the two estimators proposed by Geffroy [40] and Girard and Jacob [46]. Let us recall that the estimator in [40] is based on the extreme values of the sample and does not involve any smoothing, as opposed to both Girard and Jacob's estimator [46] and our estimator. For Girard and Jacob's estimator, we set $h_n^{(gj)} = 4\widehat{\sigma}(X)/\sqrt{n}$ and $p_n^{(gj)} = \sqrt{n/\ln(n)}$ if $d = 1$, and $h_n^{(gj)} = 4\sqrt{\widehat{\sigma}(X_1)\widehat{\sigma}(X_2)}/n^{1/3}$ and $p_n^{(gj)} = n^{1/3}/\sqrt{\ln(n)}$ when $d = 2$. The L^1-errors associated to each estimator are computed on 500 replications of the initial sample and the minimum, maximum and mean L^1-errors are reported in Table 3.1.

It appears that, in all the considered situations, our moment estimator yields better results than both the estimators of [40] and [46]. For a fixed frontier, all the estimators perform better on the situation $\alpha(x) = \alpha_1(x)$ than on the situation $\alpha(x) = \alpha_2(x)$. This behaviour is a natural consequence of the inequality $\alpha_2(x) > \alpha_1(x)$ for all $x \in \Omega$: as $\alpha(x)$ increases, the simulated points tend to move away from the frontier $g(x)$. This phenomenon is illustrated in the case $d = 1$ on Figures 3.1–3.2 and 3.3–3.4. On each of the upper panels the best situation is represented, *i.e.* the replication that yields the smallest L^1-error for $\widehat{g}_n$ in Table 3.1. Similarly, the worst situation is depicted on the lower panels, *i.e.* the replication that yields the largest L^1-error for $\widehat{g}_n$ in Table 3.1. In all cases, $\widehat{g}_n$ is superimposed to the frontier g.

Situation	Moment estimator	Girard-Jacob estimator	Geffroy estimator
Case $d = 1$, model (3.16)			
$\alpha(x) = \alpha_1(x)$			
Frontier g_1	0.082 [0.051, 0.117]	0.089 [0.052, 0.135]	0.107 [0.058, 0.168]
Frontier g_2	0.045 [0.032, 0.070]	0.047 [0.031, 0.078]	0.050 [0.029, 0.089]
$\alpha(x) = \alpha_2(x)$			
Frontier g_1	0.109 [0.073, 0.179]	0.162 [0.093, 0.241]	0.169 [0.087, 0.248]
Frontier g_2	0.064 [0.042, 0.088]	0.067 [0.037, 0.099]	0.072 [0.041, 0.115]
Case $d = 1$, model (3.17)			
$c = 7/8$	0.055 [0.032, 0.101]	0.108 [0.070, 0.157]	0.107 [0.067, 0.174]
$c = 5/8$	0.058 [0.032, 0.101]	0.116 [0.076, 0.161]	0.112 [0.069, 0.154]
$c = 3/8$	0.063 [0.030, 0.111]	0.127 [0.083, 0.171]	0.122 [0.062, 0.177]
$c = 1/8$	0.070 [0.037, 0.136]	0.137 [0.086, 0.190]	0.131 [0.085, 0.194]
Case $d = 1$, model (3.18)			
$b(x) = \alpha_1(x)$			
Frontier g_1	0.111 [0.060, 0.168]	0.147 [0.076, 0.205]	0.149 [0.067, 0.236]
Frontier g_2	0.053 [0.032, 0.083]	0.065 [0.035, 0.102]	0.067 [0.037, 0.104]
$b(x) = \alpha_2(x)$			
Frontier g_1	0.133 [0.074, 0.212]	0.242 [0.163, 0.357]	0.235 [0.141, 0.338]
Frontier g_2	0.064 [0.043, 0.098]	0.095 [0.062, 0.140]	0.097 [0.057, 0.145]
Case $d = 2$, model (3.20)	0.036 [0.024, 0.058]	0.146 [0.105, 0.195]	0.176 [0.124, 0.213]

Table 3.1: Mean L^1-errors and [minimum, maximum] L^1-errors associated to the estimators in the different situations.

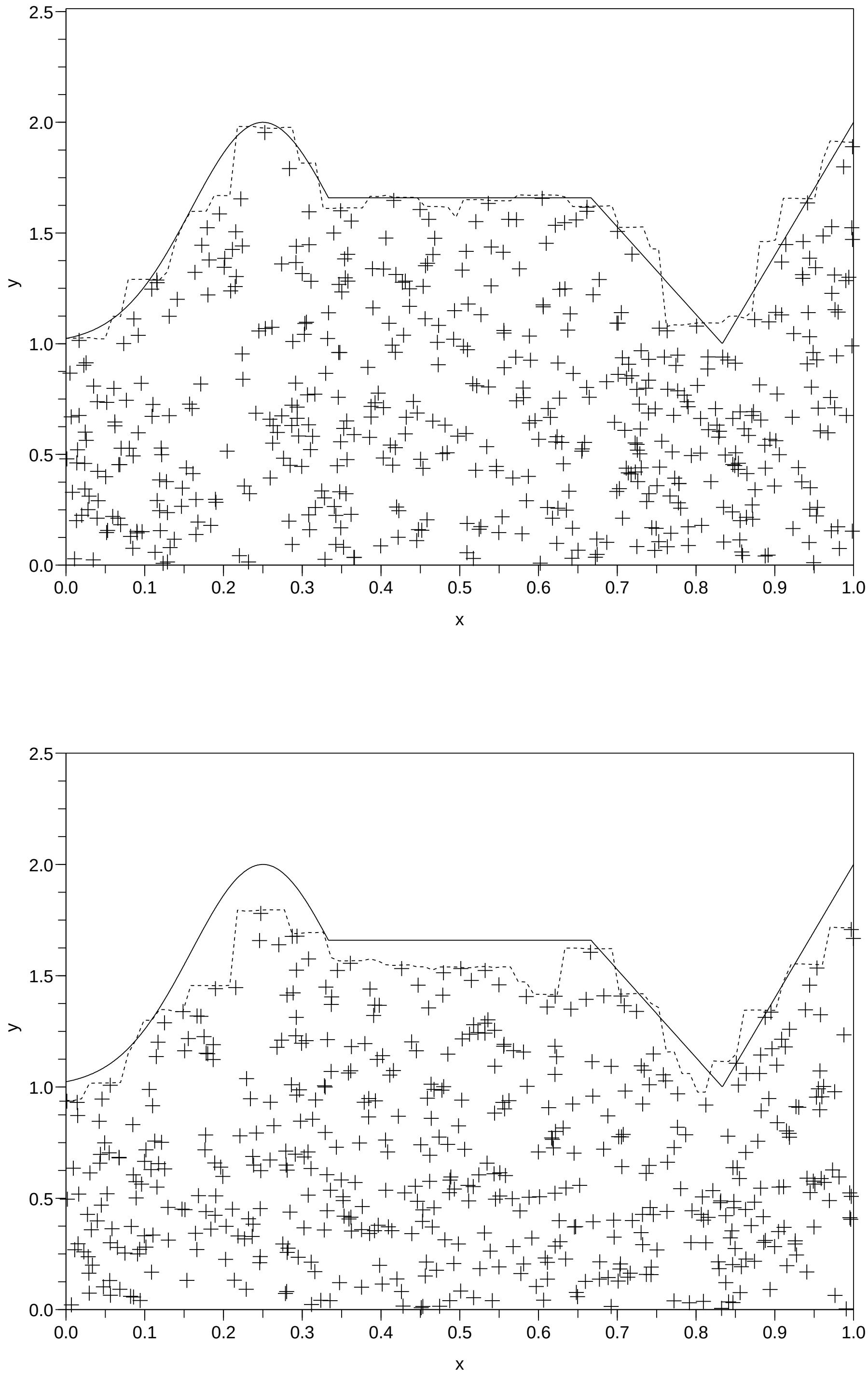
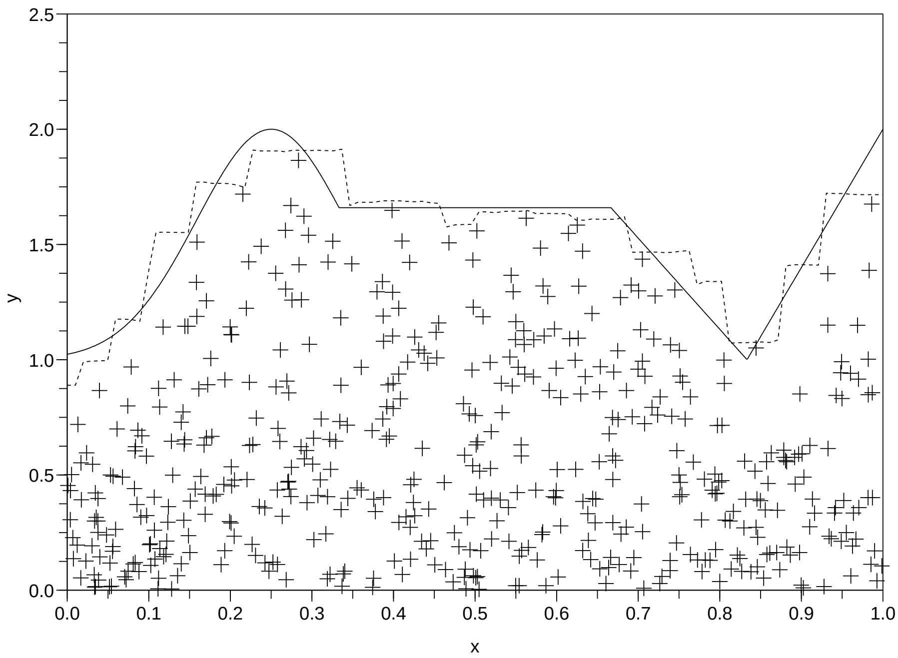


Figure 3.1: Case $d = 1$, model (3.16) and $\alpha(x) = \alpha_1(x)$: the frontier g_1 (solid line) and its moment estimate $\widehat{g}_n$ (dotted line) with $a = 15$. Top: best situation, bottom: worst situation.



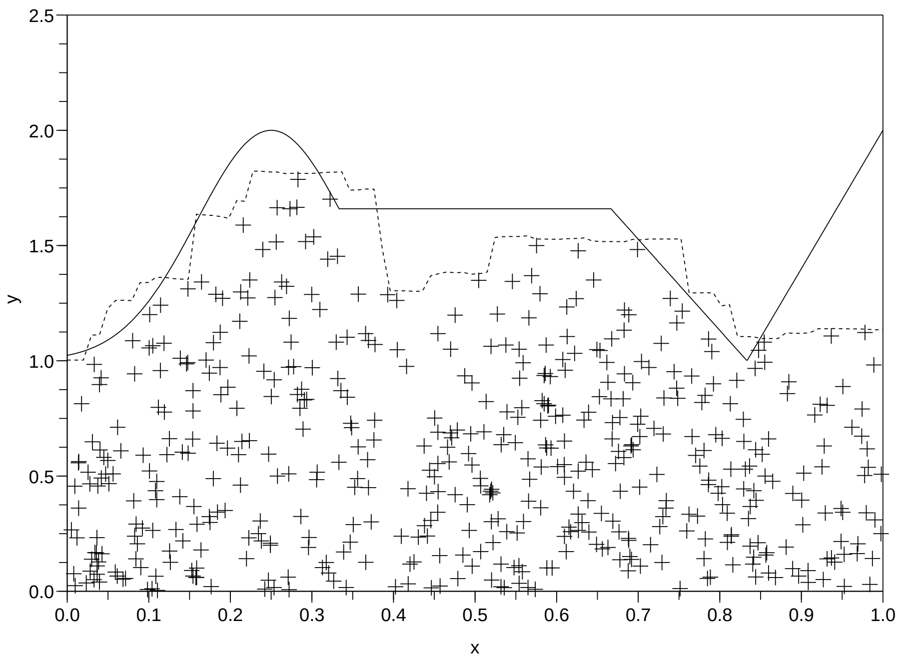


Figure 3.2: Case $d = 1$, model (3.16) and $\alpha(x) = \alpha_2(x)$: the frontier g_1 (solid line) and its moment estimate $\widehat{g}_n$ (dotted line) with $a = 15$. Top: best situation, bottom: worst situation.

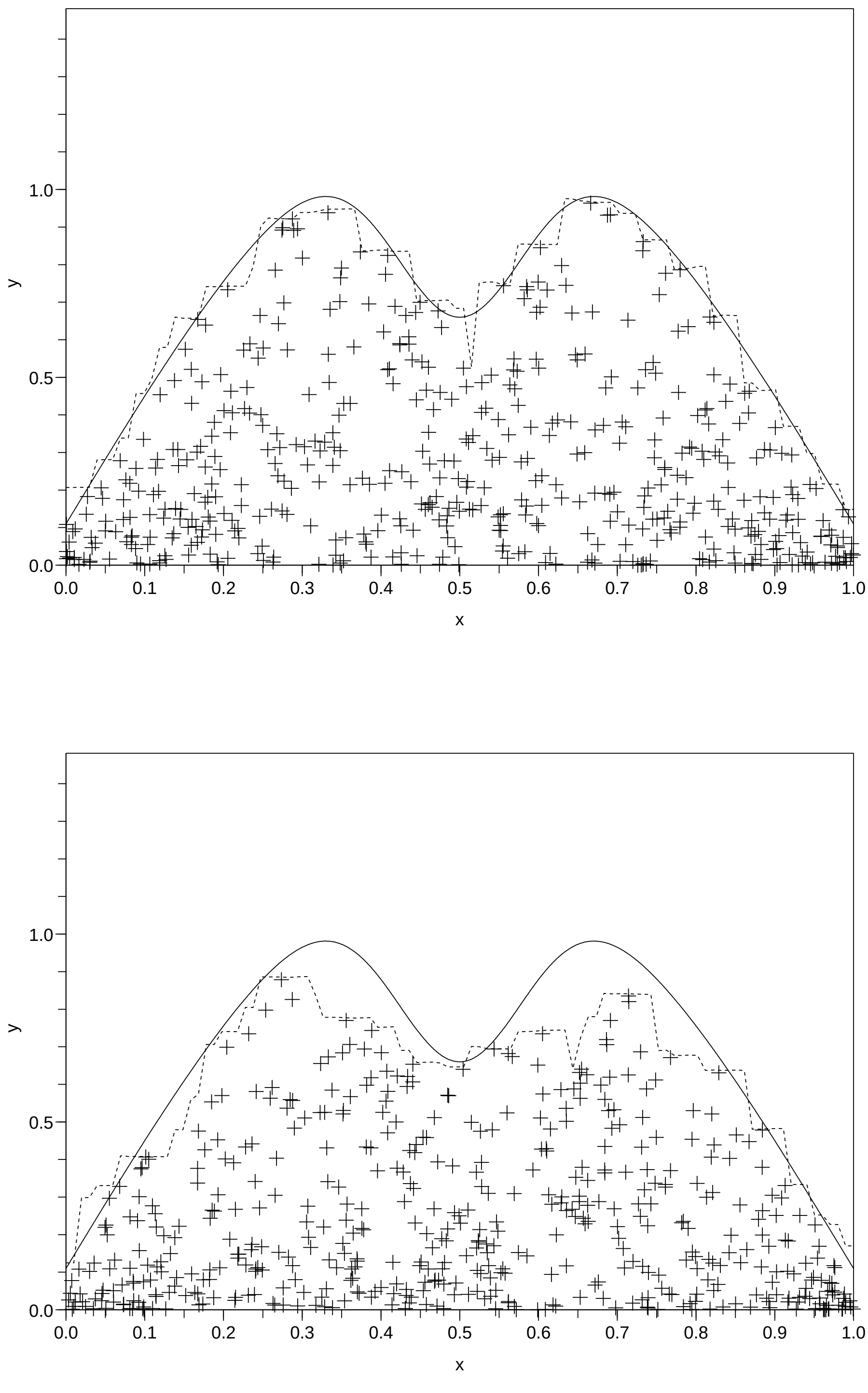
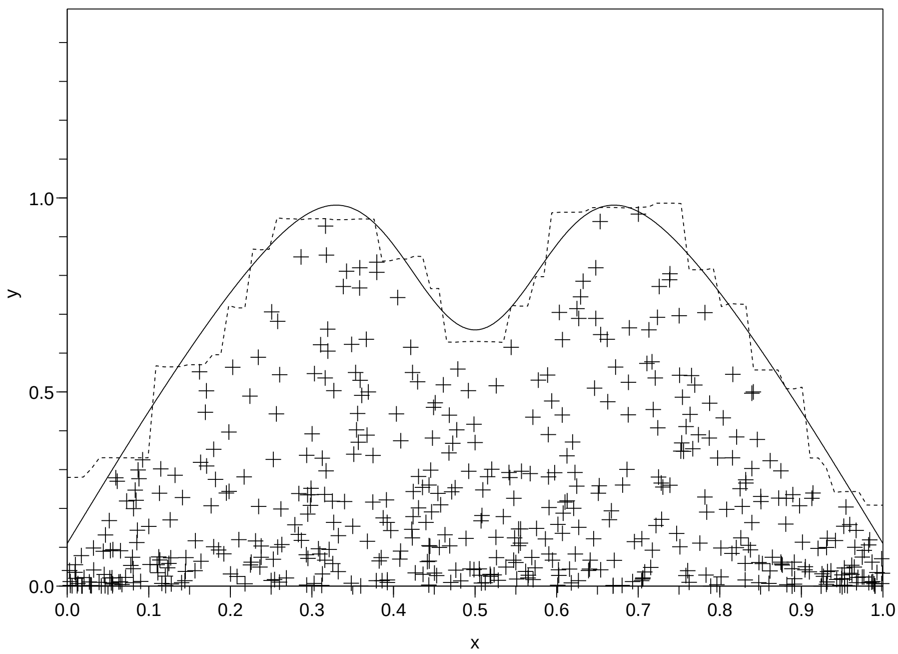


Figure 3.3: Case $d = 1$, model (3.18) and $b(x) = \alpha_1(x)$: the frontier g_2 (solid line) and its moment estimate $\widehat{g}_n$ (dotted line) with $a = 15$. Top: best situation, bottom: worst situation.



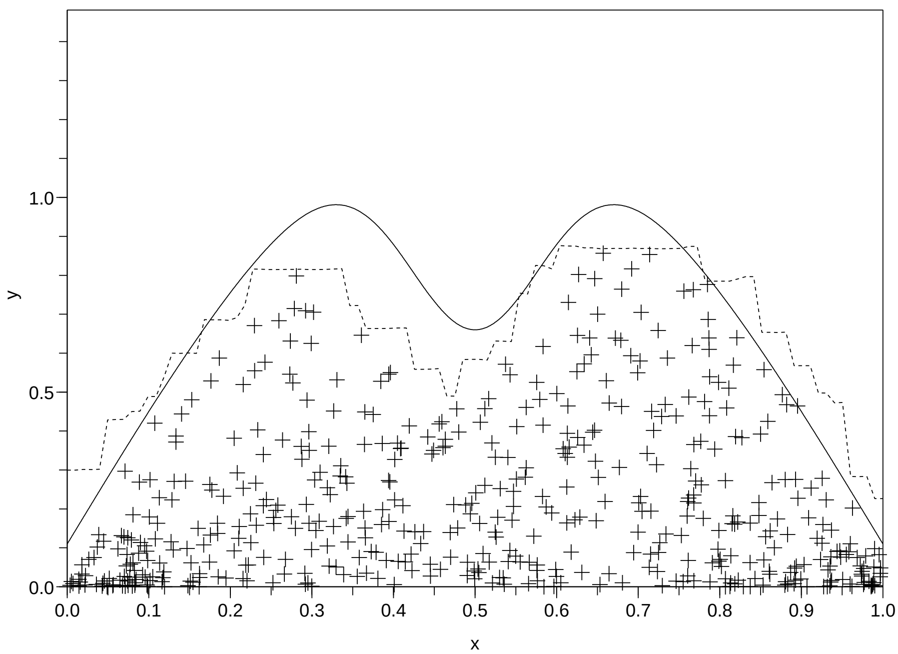


Figure 3.4: Case $d = 1$, model (3.18) and $b(x) = \alpha_2(x)$: the frontier g_2 (solid line) and its moment estimate $\widehat{g}_n$ (dotted line) with $a = 15$. Top: best situation, bottom: worst situation.

3.5 Appendix A: Auxiliary results

We first state a simple technical result we shall use several times in the proofs. Recall from Section 3.2 the notation

$$\Delta_n^{\varphi}(x,\, u) = \varphi(x - h_n u) - \varphi(x).$$

Lemma 3.1. *Let φ be a real-valued locally Hölder continuous function on $\mathbb{R}^d$ with exponent η. Then*

$$\forall\, x \in \mathbb{R}^d, \quad \sup_{u\in\mathcal{B}} |\Delta_n^{\varphi}(x,\, u)| = \mathrm{O}(h_n^{\eta}).$$

If moreover $\varphi(x) \neq 0$, then

$$\sup_{u\in\mathcal{B}} \left| \frac{\varphi(x - h_n u)}{\varphi(x)} - 1 \right| = \mathrm{O}(h_n^{\eta}).$$

The next lemma focuses on the behaviour of the moments $m_{p_n}(x)$ and $\mu_{p_n}(x)$.

Lemma 3.2. *Assume that $(NP_1 - NP_2)$, (K) and (A_1) hold. Let $x \in \Omega$ such that $f(x) > 0$. If $p_n\, h_n^{\eta_g} \to 0$, then*

(i) $m_{p_n}(x)/m_{p_n+1}(x) \to 1/g(x)$ as $n \to \infty$;

(ii) $\mu_{p_n}(x) = f(x)\, m_{p_n}(x)\,(1 + \mathrm{o}(1))$ as $n \to \infty$.

Before proceeding, we need some additional notations. Let $\overline{F}_{g,\,\gamma}(y\,|\,x) = (1 - y/g(x))^{\gamma(x)}$; we can then write

$$\overline{F}(y\,|\,x) = C(x)\,\overline{F}_{g,\,\alpha}(y\,|\,x) + D(x)\,\overline{F}_{g,\,\alpha+\beta}(y\,|\,x)\left[1 + \delta\left(x,\, (1 - y/g(x))^{-1}\right)\right].$$

Pick $x \in \mathbb{R}^d$, and set for every $q \geq 1$, every positive locally Hölder continuous function γ on $\mathbb{R}^d$ and every locally bounded Borel function W on $\mathbb{R}^d$

$$\begin{aligned} M_n^{(\gamma,\,W)}(q,\, p_n,\, x) &:= \int_{\Omega} f(v)\, W(v)\, K_{h_n}^q(x - v)\left[p_n \int_0^{\infty} y^{p_n - 1}\,\overline{F}_{g,\,\gamma}(y\,|\,v)\, dy\right] dv \\ &= h_n^{-d(q-1)} \int_{\mathcal{B}} (fWg^{p_n})(x - h_n u)\, p_n\, B(p_n,\, \gamma(x - h_n u) + 1)\, K^q(u)\, du. \end{aligned}$$

Let further

$$\mathcal{I}_{\gamma,\,R}(p_n,\, v) := \int_0^1 y^{p_n - 1}\,\overline{F}_{1,\,\gamma}(y\,|\,v)\, R\left(v,\, (1 - y)^{-1}\right)\, dy \tag{3.21}$$

for every Borel function R on $\Omega \times [1,\, \infty)$ such that for all $v \in \Omega$, $R(v,\, \cdot)$ is bounded.

With these notations,

$$\mathbb{E}\left(Y^{p_n}\, K_{h_n}^q(x - X)\right) = M_n^{(\alpha,\, C)}(q,\, p_n,\, x)[1 + \varepsilon_n(q,\, p_n,\, x)]$$

with the error term being

$$\varepsilon_n(q,\, p_n,\, x) = \frac{M_n^{(\alpha+\beta,\, D)}(q,\, p_n,\, x)}{M_n^{(\alpha,\, C)}(q,\, p_n,\, x)}\left[1 + E_n^{(\alpha+\beta,\, D)}(q,\, p_n,\, x)\right]$$

if we set

$$E_n^{(\gamma,\,W)}(q,\,p_n,\,x) = \frac{\displaystyle\int_{\mathcal{B}} (fWg^{p_n})(x-h_nu)\,p_n\,\mathcal{I}_{\gamma,\,\delta}(p_n,\,x-h_nu)\,K^q(u)\,du}{h_n^{d(q-1)}\,M_n^{(\gamma,\,W)}(q,\,p_n,\,x)}. \tag{3.22}$$

Finally, letting

$$\tau_n(q,\,p_n,\,x) = \frac{\varepsilon_n(q,\,p_n,\,x)-\varepsilon_n(q,\,p_n+1,\,x)}{1+\varepsilon_n(q,\,p_n+1,\,x)},$$

we obtain

$$\frac{\mu_{p_n}(x)}{\mu_{p_n+1}(x)} = \frac{M_n^{(\alpha,\,C)}(1,\,p_n,\,x)}{M_n^{(\alpha,\,C)}(1,\,p_n+1,\,x)}\,[1+\tau_n(1,\,p_n,\,x)]. \tag{3.23}$$

The next result of this section is technical: it provides precise expansions of $M_n^{(\gamma,\,W)}(q,\,p_n,\,x)$ for all $q \geq 1$, when $p_n\,h_n^{\eta_g} \to 0$. This result will be useful for the proofs of our next lemmas and of Theorem 3.7.

Lemma 3.3. *Let γ, W be locally Hölder continuous functions on $\mathbb{R}^d$ with respective exponents η_γ and η_W. We assume that γ is positive and W is bounded away from 0. Pick $q \geq 1$, $x \in \mathbb{R}^d$ such that $f(x) > 0$. For all $u \in \mathcal{B}$ and $n \in \mathbb{N}\setminus\{0\}$, let*

$$\begin{aligned}
\mathcal{L}_n^{(\gamma,\,W)}(p_n,\,x,\,u) &= \frac{(fW)(x-h_nu)\,\Gamma(\gamma(x-h_nu)+1)}{(fW)(x)\,\Gamma(\gamma(x)+1)}\exp\left[p_n\frac{\Delta_n^g(x,\,u)}{g(x)} - \ln(p_n)\Delta_n^\gamma(x,\,u)\right],\\
\Lambda_n^{(\gamma,\,W)}(q,\,p_n,\,x) &= h_n^{d(q-1)}\frac{M_n^{(\gamma,\,W)}(q,\,p_n,\,x)}{(fWg^{p_n})(x)}.
\end{aligned}$$

Assume that (A_1) holds. If $p_n\,h_n^{\eta_g} \to 0$, then

(i) $\mathcal{L}_n^{(\gamma,\,W)}(p_n,\,x,\,u) \to 1$ *as $n \to \infty$ uniformly in $u \in \mathcal{B}$.*

(ii) For all $q \geq 1$,

$$\begin{aligned}
\frac{\Lambda_n^{(\gamma,\,W)}(q,\,p_n,\,x)}{\gamma(x)\,B(p_n+1,\,\gamma(x))} &= \int_{\mathcal{B}} \mathcal{L}_n^{(\gamma,\,W)}(p_n,\,x,\,u)\left[1-\frac{p_n}{2}\left(\frac{\Delta_n^g(x,\,u)}{g(x)}\right)^2\right]K^q(u)\,du\\
&- \frac{1}{p_n}\int_{\mathcal{B}} \mathcal{L}_n^{(\gamma,\,W)}(p_n,\,x,\,u)\frac{\Delta_n^\gamma(x,\,u)}{2}\,[\gamma(x-h_nu)+\gamma(x)+1]\,K^q(u)\,du\\
&+ \mathrm{O}\left(\frac{h_n^{\eta_g}}{p_n}+\frac{h_n^{\eta_\gamma}}{p_n^2}\right).
\end{aligned}$$

(iii) For all $q \geq 1$,

$$\begin{aligned}
\frac{\Lambda_n^{(\gamma,\,W)}(q,\,p_n,\,x)}{\Gamma(\gamma(x)+1)\,p_n^{-\gamma(x)}} &= \int_{\mathcal{B}} \mathcal{L}_n^{(\gamma,\,W)}(p_n,\,x,\,u)\left[1-\frac{\gamma(x)(\gamma(x)+1)}{2p_n}-\frac{p_n}{2}\left(\frac{\Delta_n^g(x,\,u)}{g(x)}\right)^2\right]K^q(u)\,du\\
&- \frac{1}{p_n}\int_{\mathcal{B}} \mathcal{L}_n^{(\gamma,\,W)}(p_n,\,x,\,u)\frac{\Delta_n^\gamma(x,\,u)}{2}\,[\gamma(x-h_nu)+\gamma(x)+1]\,K^q(u)\,du\\
&+ \frac{\gamma(x)(\gamma(x)+1)(3\gamma^2(x)+7\gamma(x)+2)}{24p_n^2}\int_{\mathcal{B}} K^q(u)\,du + \mathrm{o}\left(\frac{1}{p_n^2}\right).
\end{aligned}$$

Lemma 3.3 is the key tool to a detailed study of the functions $\varepsilon_n(q,\,p_n,\,x)$, which is the focus of Lemma 3.4 below.

Lemma 3.4. *Let γ, W be locally Hölder continuous functions on $\mathbb{R}^d$. Assume that γ is positive and W is bounded away from 0. Pick $q \geq 1$ and $x \in \mathbb{R}^d$ such that $f(x) > 0$. Assume that hypotheses (A_1) and (A_4) hold, and $p_n\, h_n^{\eta_g} \to 0$ as $n \to \infty$. Then*

(i) $(n,\, t) \mapsto (p_n + t)^{\gamma(x)}\, M_n^{(\gamma,\, W)}(q,\, p_n + t,\, x) \in \mathcal{C}_2$.

(ii) $(n,\, t) \mapsto E_n^{(\gamma,\, W)}(q,\, p_n + t,\, x) \in \mathcal{D}_2$.

Our next lemma, which is similar to Lemma 2.13, consists in linearizing the random variable $\xi_n(x)$ appearing in Theorem 3.7.

Lemma 3.5. *Assume that (P), (K), $(A_1 - A_4)$ hold. Let $x \in \Omega$ such that $f(x) > 0$. If $p_n\, h_n^{\eta_g} \to 0$ then*

$$\begin{aligned} \xi_n(x) &= \left[\zeta_n^{(1)}(x) + \left(\frac{\mu_{p_n+1}(x)}{\widehat{\mu}_{p_n+1}(x)} - 1\right)\zeta_n^{(2)}(x) + \left(1 + \frac{a p_n}{p_n+1}\right)\left(\frac{\mu_{(a+1)p_n+1}(x)}{\widehat{\mu}_{(a+1)p_n+1}(x)} - 1\right)\zeta_n^{(3)}(x)\right] \\ &\times \quad u_{n,\,a}(x)\,(1 + \mathrm{o}(1)) \end{aligned}$$

where, letting $\nu_p(x) = \widehat{\mu}_p(x) - \mu_p(x)$,

$$\begin{aligned} \zeta_n^{(1)}(x) &= \zeta_n^{(2)}(x) + \left[1 + \frac{a p_n}{p_n+1}\right]\zeta_n^{(3)}(x) \\ \textit{with}\quad \zeta_n^{(2)}(x) &= -\nu_{p_n}(x) + \frac{\mu_{p_n}(x)}{\mu_{p_n+1}(x)}\,\nu_{p_n+1}(x), \\ \zeta_n^{(3)}(x) &= \frac{\mu_{p_n+1}(x)}{\mu_{(a+1)p_n+1}(x)}\,\nu_{(a+1)p_n}(x) - \frac{\mu_{p_n+1}(x)\,\mu_{(a+1)p_n}(x)}{\mu^2_{(a+1)p_n+1}(x)}\,\nu_{(a+1)p_n+1}(x) \\ \textit{and}\quad u_{n,\,a}(x) &= \frac{1}{a\|K\|_2\,\Gamma(\alpha(x)+1)}\sqrt{\frac{1}{f(x)\,C(x)\,V(\alpha(x),\,a)}\,\frac{p_n^{\alpha(x)}\,v_n(x)}{g^{p_n}(x)}}. \end{aligned}$$

Finally, the following result provides an asymptotic bound of the third-order moments appearing in the proof of Theorem 3.7.

Lemma 3.6. *Assume that (P), (K), $(A_1 - A_4)$ hold and $p_n\, h_n^{\eta_g} \to 0$ as $n \to \infty$. Let $k \in \mathbb{N}$, $(b_{n,\,j})$, $0 \leq j \leq k$ be sequences of real numbers and $x \in \mathbb{R}^d$ such that there exist $m \in \mathbb{N}$ and sequences of measurable functions $(H_{n,\,j})$, $0 \leq j \leq m$, uniformly bounded on $[0,\,1]$ with*

$$\forall\, y \in [0,\,1], \quad \sup_{u \in \mathcal{B}}\left|\sum_{j=0}^{k} b_{n,\,j}\, g^j(x - h_n u)\, y^j\right| \leq \sum_{j=0}^{m}\frac{H_{n,\,j}(y)}{p_n^j}(1-y)^{m-j}.$$

Let us consider

$$S_n(x) = \frac{1}{n}\sum_{j=0}^{k} b_{n,\,j}\, Y^{p_n+j}\, K_{h_n}(x - X).$$

Then $\mathbb{E}|S_n(x)|^3 = \mathrm{O}\left(n^{-3}\, g^{3p_n}(x)\, p_n^{-\alpha(x)-3m}\, h_n^{-2d}\right)$.

3.6 Appendix B: Proofs

Proof of Lemma 3.1. Just notice that for all sufficiently large n,

$$\forall\, x \in \mathbb{R}^d, \quad \exists\, \varepsilon_\varphi > 0, \quad \forall\, u \in \mathcal{B}, \quad |\Delta_n^\varphi(x,\, u)| \leq \varepsilon_\varphi \|h_n u\|^\eta \leq \varepsilon_\varphi\, h_n^\eta$$

to obtain the result. ■

Proof of Lemma 3.2. (i) Set $I_{p_n}(x) := \dfrac{m_{p_n}(x)}{p_n\, g^{p_n}(x)}$ and $\overline{\mathcal{F}}(y \,|\, x) = \overline{F}(g(x)y \,|\, x)$, so that

$$I_{p_n}(x) = \int_0^1 y^{p_n - 1}\, \overline{\mathcal{F}}(y \,|\, x)\, dy.$$

Pick y_0 as in (NP_2), let $\varepsilon \in (0,\, 1 - y_0)$ and $u \in \mathcal{B}$. The integral $I_{p_n}(x - h_n u)$ is rewritten as:

$$I_{p_n}(x - h_n u) \quad = \quad \int_{1-\varepsilon}^1 y^{p_n-1}\, \overline{\mathcal{F}}(y \,|\, x - h_n u)\, dy \left[1 + \frac{\displaystyle\int_0^{1-\varepsilon} y^{p_n-1}\, \overline{\mathcal{F}}(y \,|\, x - h_n u)\, dy}{\displaystyle\int_{1-\varepsilon}^1 y^{p_n-1}\, \overline{\mathcal{F}}(y \,|\, x - h_n u)\, dy} \right].$$

As in the proof of Lemma 2.2, one has

$$0 \leq \frac{\displaystyle\int_0^{1-\varepsilon} y^{p_n-1}\, \overline{\mathcal{F}}(y \,|\, x - h_n u)\, dy}{\displaystyle\int_{1-\varepsilon}^1 y^{p_n-1}\, \overline{\mathcal{F}}(y \,|\, x - h_n u)\, dy} \leq \frac{1-\varepsilon}{\left[\dfrac{1-\varepsilon/2}{1-\varepsilon}\right]^{p_n-1} \displaystyle\int_{1-\varepsilon/2}^1 \overline{\mathcal{F}}(y \,|\, x - h_n u)\, dy}.$$

Moreover

$$\sup_{u \in \mathcal{B}} \left| \int_{1-\varepsilon/2}^1 \overline{\mathcal{F}}(y \,|\, x - h_n u)\, dy - \int_{1-\varepsilon/2}^1 \overline{\mathcal{F}}(y \,|\, x)\, dy \right| \leq \int_{1-\varepsilon/2}^1 \sup_{u \in \mathcal{B}} \left| \overline{\mathcal{F}}(y \,|\, x - h_n u) - \overline{\mathcal{F}}(y \,|\, x) \right| \, dy,$$

and (NP_2) ensures that for all $y \in [1-\varepsilon,\, 1]$, the map $x \mapsto \overline{\mathcal{F}}(y \,|\, x)$ is continuous, so that, by the dominated convergence theorem,

$$\int_{1-\varepsilon/2}^1 \overline{\mathcal{F}}(y \,|\, x - h_n u)\, dy \to \int_{1-\varepsilon/2}^1 \overline{\mathcal{F}}(y \,|\, x)\, dy > 0$$

as $n \to \infty$, uniformly in $u \in \mathcal{B}$.

Since $\left[\dfrac{1-\varepsilon/2}{1-\varepsilon}\right]^{p_n-1} \to \infty$ as $n \to \infty$, we therefore get, uniformly in $u \in \mathcal{B}$,

$$I_{p_n}(x - h_n u) = \int_{1-\varepsilon}^1 y^{p_n-1}\, \overline{\mathcal{F}}(y \,|\, x - h_n u)\, dy\, (1 + \mathrm{o}(1)).$$

Because

$$\forall\, \varepsilon \in (0,\, 1 - y_0), \quad 1 \leq \frac{\displaystyle\int_{1-\varepsilon}^1 y^{p_n-1}\, \overline{\mathcal{F}}(y \,|\, x)\, dy}{\displaystyle\int_{1-\varepsilon}^1 y^{p_n}\, \overline{\mathcal{F}}(y \,|\, x)\, dy} \leq \frac{1}{1-\varepsilon}$$

one has $I_{p_n}(x)/I_{p_n+1}(x) \to 1$ as $n \to \infty$. Hence $m_{p_n}(x)/m_{1,\,p_n+1}(x) \to 1/g(x)$ as $n \to \infty$, which completes the proof of (i).

(ii) Write

$$\sup_{u\in\mathcal{B}} \left| \frac{\displaystyle\int_{1-\varepsilon}^{1} y^{p_n-1}\,\overline{\mathcal{F}}(y\,|\,x-h_n u)\,dy}{\displaystyle\int_{1-\varepsilon}^{1} y^{p_n-1}\,\overline{\mathcal{F}}(y\,|\,x)\,dy} - 1 \right| \le \sup_{u\in\mathcal{B}} \frac{\displaystyle\int_{1-\varepsilon}^{1} y^{p_n-1}\,\overline{\mathcal{F}}(y\,|\,x) \left| \frac{\overline{\mathcal{F}}(y\,|\,x-h_n u)}{\overline{\mathcal{F}}(y\,|\,x)} - 1 \right| dy}{\displaystyle\int_{1-\varepsilon}^{1} y^{p_n-1}\,\overline{\mathcal{F}}(y\,|\,x)\,dy}$$

$$\le \sup_{y\in[y_0,\,1]}\ \sup_{u\in\mathcal{B}} \left| \frac{\overline{\mathcal{F}}(y\,|\,x-h_n u)}{\overline{\mathcal{F}}(y\,|\,x)} - 1 \right| \to 0$$

as $n \to \infty$. Consequently,

$$I_{p_n}(x-h_n u) = I_{p_n}(x)\,(1+\mathrm{o}(1)) \tag{3.24}$$

uniformly in $u \in \mathcal{B}$.

Besides, (A_1) and Lemma 3.1 yield

$$\sup_{u\in\mathcal{B}} \left| \frac{f(x-h_n u)}{f(x)} - 1 \right| = \mathrm{O}\left(h_n^{\eta_f}\right) \to 0, \tag{3.25}$$

$$\text{and} \qquad p_n \sup_{u\in\mathcal{B}} \left| \frac{\Delta_n^g(x,\,u)}{g(x)} \right| = \mathrm{O}\left(p_n\, h_n^{\eta_g}\right),$$

so that the hypothesis $p_n\, h_n^{\eta_g} \to 0$ gives

$$\ln\left[\frac{g^{p_n}(x-h_n u)}{g^{p_n}(x)}\right] = p_n \ln\left[1 + \frac{\Delta_n^g(x,\,u)}{g(x)}\right] = \mathrm{O}\left(p_n\, h_n^{\eta_g}\right) \to 0 \tag{3.26}$$

uniformly in $u \in \mathcal{B}$ as $n \to \infty$. Now, recall that

$$\mu_{p_n}(x) = p_n \int_{\mathcal{B}} f(x-h_n u)\, g^{p_n}(x-h_n u)\, I_{p_n}(x-h_n u)\, K(u)\, du.$$

Collecting (3.24), (3.25) and (3.26), the dominated convergence theorem therefore gives (ii). ■

Proof of Lemma 3.3. (i) Let us introduce

$$Q_n^{(\gamma,\,W)}(x,\,u) = \frac{(fW)(x-h_n u)\,\Gamma(\gamma(x-h_n u)+1)}{(fW)(x)\,\Gamma(\gamma(x)+1)}.$$

Since f, W and γ are continuous at x and Γ is continuous on $(0,\,\infty)$, one has $Q_n^{(\gamma,\,W)}(x,\,u) \to 1$ as $n \to \infty$, uniformly in $u \in \mathcal{B}$. Moreover, since $p_n\, h_n^{\eta_g} \to 0$, we have

$$\sup_{u\in\mathcal{B}} \ln(p_n)\,|\Delta_n^\gamma(x,\,u)| = \mathrm{O}\left(h_n^{\eta_\gamma} \ln p_n\right) = \mathrm{O}\left([h_n^{\eta_g} p_n]^{\eta_\gamma/\eta_g}\ \frac{\ln p_n}{p_n^{\eta_\gamma/\eta_g}} \right) \to 0.$$

Lemma 3.1 gives

$$\sup_{u\in\mathcal{B}} p_n\,|\Delta_n^g(x,\,u)| \to 0 \qquad \text{as } n \to \infty;$$

as a conclusion, $\mathcal{L}_n^{(\gamma,\,W)}(p_n,\,x,\,u) \to 1$ as $n \to \infty$, uniformly in $u \in \mathcal{B}$.

(ii) By definition of the Beta function,

$$\frac{\Lambda_n^{(\gamma,\,W)}(q,\,p_n,\,x)}{\gamma(x)\,B(p_n+1,\,\gamma(x))} = \int_{\mathcal{B}} Q_n^{(\gamma,\,W)}(x,\,u)\,\frac{\Gamma(p_n+1+\gamma(x))}{\Gamma(p_n+1+\gamma(x-h_nu))}\,\frac{g^{p_n}(x-h_nu)}{g^{p_n}(x)}\,K^q(u)\,du. \quad (3.27)$$

Following Formula 6.1.50 p. 258 in Abramovitz and Stegun [2], simple calculations yield

$$\begin{aligned}\frac{\Gamma(p_n+1+\gamma(x))}{\Gamma(p_n+1+\gamma(x-h_nu))} &= \exp(-\ln(p_n)\Delta_n^\gamma(x,\,u))\left[1-\frac{\Delta_n^\gamma(x,\,u)}{2p_n}\,(1+\gamma(x-h_nu)+\gamma(x))\right]\\ &+ \mathrm{O}\left(\frac{h_n^{\eta_\gamma}}{p_n^2}\right) \quad (3.28)\end{aligned}$$

uniformly in $u\in\mathcal{B}$. Besides

$$\begin{aligned}\frac{g^{p_n}(x-h_nu)}{g^{p_n}(x)} &= \exp\left[p_n\ln\left(1+\frac{\Delta_n^g(x,\,u)}{g(x)}\right)\right]\\ &= \exp\left[p_n\frac{\Delta_n^g(x,\,u)}{g(x)}\right]\left[1-\frac{p_n}{2}\left(\frac{\Delta_n^g(x,\,u)}{g(x)}\right)^2\right]+\mathrm{O}\left(\frac{h_n^{\eta_g}}{p_n}\right) \quad (3.29)\end{aligned}$$

uniformly in $u\in\mathcal{B}$. Replacing (3.28) and (3.29) in (3.27) gives the first desired expansion.

(iii) According to [103], for all κ and ι, one has

$$\frac{\Gamma(x+\kappa)}{\Gamma(x+\iota)} = x^{\kappa-\iota}\left[1+\frac{\delta_1(\kappa,\,\iota)}{x}+\frac{\delta_2(\kappa,\,\iota)}{x^2}+\mathrm{o}\left(\frac{1}{x^2}\right)\right]$$

as $x\to\infty$, where

$$\begin{aligned}\delta_1(\kappa,\,\iota) &= \frac{(\kappa-\iota)(\kappa+\iota-1)}{2},\\ \delta_2(\kappa,\,\iota) &= \frac{(\kappa-\iota)(\kappa-\iota-1)(3(\kappa+\iota-1)^2-\kappa+\iota-1)}{24}.\end{aligned}$$

Consequently

$$B(p_n+1,\,\gamma(x)) = \Gamma(\gamma(x))\,p_n^{-\gamma(x)}\left[1+\frac{\delta_1(1,\,\gamma(x)+1)}{p_n}+\frac{\delta_2(1,\,\gamma(x)+1)}{p_n^2}+\mathrm{o}\left(\frac{1}{p_n^2}\right)\right].$$

Replacing in the expansion (ii) and remarking that, from (i),

$$\frac{1}{p_n^2}\int_{\mathcal{B}}\mathcal{L}_n^{(\gamma,\,W)}(p_n,\,x,\,u)\,K^q(u)\,du = \frac{1}{p_n^2}\int_{\mathcal{B}}K^q(u)\,du+\mathrm{o}\left(\frac{1}{p_n^2}\right)$$

yields (iii). ■

Proof of Lemma 3.4. (i) Like in the proof of Lemma 3.3, we have

$$\begin{aligned}&h_n^{d(q-1)}\frac{M_n^{(\gamma,\,W)}(q,\,p_n,\,x)}{\gamma(x)\,B(p_n+1,\,\gamma(x))\,(fWg^{p_n})(x)}\\ &\qquad=\int_{\mathcal{B}} Q_n^{(\gamma,\,W)}(x,\,u)\,\frac{\Gamma(p_n+1+\gamma(x))}{\Gamma(p_n+1+\gamma(x-h_nu))}\,\frac{g^{p_n}(x-h_nu)}{g^{p_n}(x)}\,K^q(u)\,du.\end{aligned}$$

Since

$$(p_n+t)^{\gamma(x)}\,B(p_n+t+1,\,\gamma(x)) = \frac{(p_n+t)^{\gamma(x)+1}}{\gamma(x)}B(p_n+t,\,\gamma(x)+1),$$

using [103] and applying Lemma 2.16iii), we obtain that

$$(n,\, t) \mapsto (p_n + t)^{\gamma(x)}\, B(p_n + t + 1,\, \gamma(x)) \in \mathcal{E}_2.$$

Consequently, it is enough to show that

$$(n,\, t) \mapsto \int_{\mathcal{B}} Q_n^{(\gamma,\, W)}(x,\, u)\, \frac{\Gamma(p_n + t + 1 + \gamma(x))}{\Gamma(p_n + t + 1 + \gamma(x - h_n u))}\, \frac{g^{p_n+t}(x - h_n u)}{g^{p_n+t}(x)}\, K^q(u)\, du$$

lies in $\mathcal{C}_2$.

Recalling (3.28) and (3.29), Taylor expansions and Lemma 2.16iii) imply that

$$\forall\, u \in \mathcal{B}, \quad \begin{cases} (n,\, t) \mapsto \exp\left[-\ln\left(1 + \dfrac{t}{p_n}\right) \Delta_n^{\gamma}(x,\, u)\right] \in \mathcal{C}_2, \\[2ex] (n,\, t) \mapsto \exp\left[t\, \dfrac{\Delta_n^{g}(x,\, u)}{g(x)}\right] \in \mathcal{C}_2; \end{cases}$$

noticing that

$$\begin{aligned} \exp\left(-\ln(p_n + t)\Delta_n^{\gamma}(x,\, u)\right) &= \exp\left(-\ln(p_n)\Delta_n^{\gamma}(x,\, u)\right)\, \exp\left[-\ln\left(1 + \frac{t}{p_n}\right) \Delta_n^{\gamma}(x,\, u)\right], \\ \exp\left[(p_n + t)\frac{\Delta_n^{g}(x,\, u)}{g(x)}\right] &= \exp\left[p_n \frac{\Delta_n^{g}(x,\, u)}{g(x)}\right]\, \exp\left[t\, \frac{\Delta_n^{g}(x,\, u)}{g(x)}\right], \end{aligned}$$

Lemma 2.16i) entails that

$$\forall\, u \in \mathcal{B}, \quad \begin{cases} (n,\, t) \mapsto \exp\left(-\ln(p_n + t)\Delta_n^{\gamma}(x,\, u)\right) \in \mathcal{C}_2, \\[2ex] (n,\, t) \mapsto \exp\left[(p_n + t)\dfrac{\Delta_n^{g}(x,\, u)}{g(x)}\right] \in \mathcal{C}_2. \end{cases}$$

Furthermore, the convergences that hold for these sequences (in the definition of $\mathcal{C}_2$) are uniform in $u \in \mathcal{B}$. Applying Lemma 2.16i) once again then yields

$$\forall\, u \in \mathcal{B}, \quad \begin{cases} (n,\, t) \mapsto \dfrac{\Gamma(p_n + t + 1 + \gamma(x))}{\Gamma(p_n + t + 1 + \gamma(x - h_n u))} \in \mathcal{C}_2, \\[2ex] (n,\, t) \mapsto \dfrac{g^{p_n+t}(x - h_n u)}{g^{p_n+t}(x)} \in \mathcal{C}_2, \end{cases}$$

and the convergences these sequences satisfy are uniform in $u \in \mathcal{B}$: (i) is proven.

(ii) To prove that for all $t \geq 0$ the sequence $\left(E_n^{(\gamma,\, W)}(q,\, p_n + t,\, x)\right)$ goes to 0 as $n \to \infty$, pick $\varepsilon > 0$ and use hypothesis (A_4) to get $\rho \in (0,\, 1)$ and $N_0 \in \mathbb{N} \setminus \{0\}$ such that

$$\forall\, n \geq N_0, \quad \sup_{u \in \mathcal{B}}\, \sup_{y \geq 1-\rho} \left|\delta\left(x - h_n u,\, (1-y)^{-1}\right)\right| \leq \frac{\varepsilon}{2}.$$

Recall (3.21) and (3.22) to write

$$E_n^{(\gamma,\, W)}(q,\, p_n + t,\, x) = \frac{\displaystyle\int_{\mathcal{B}} \left(fWg^{p_n+t}\right)(x - h_n u)\, \mathcal{I}_{\gamma,\, \delta}(p_n + t,\, x - h_n u)\, K^q(u)\, du}{\displaystyle\int_{\mathcal{B}} \left(fWg^{p_n+t}\right)(x - h_n u)\, B(p_n + t,\, \gamma(x - h_n u) + 1)\, K^q(u)\, du}.$$

Since, for $n \geq N_0$ and $u \in \mathcal{B}$,

$$\forall\, y \in [0,\, 1-\rho], \quad y^{p_n+t-1} \left|\delta\left(x - h_n u,\, (1-y)^{-1}\right)\right| \leq (1-\rho)^{p_n+t-1} \sup_{\substack{\|h\| \leq h_n \\ z \leq 1/\rho}} |\delta(x+h,\, z)|$$

and

$$\forall\, y \in [1-\rho,\, 1], \quad \left|\delta\left(x - h_n u,\, (1-y)^{-1}\right)\right| \leq \frac{\varepsilon}{2},$$

we get

$$|\mathcal{I}_{\gamma,\,\delta}(p_n + t,\, x - h_n u)| \leq (1-\rho)^{p_n+t} \sup_{\substack{\|h\| \leq h_n \\ z \leq 1/\rho}} |\delta(x+h,\, z)| + \frac{\varepsilon}{2}\, B(p_n + t,\, \gamma(x - h_n u) + 1).$$

(3.28) and [103] show that

$$\frac{B(p_n + t,\, \gamma(x - h_n u) + 1)}{\Gamma(\gamma(x) + 1)\, p_n^{-\gamma(x)-1}} \to 1$$

as $n \to \infty$, uniformly in $u \in \mathcal{B}$, so that

$$\frac{(1-\rho)^{p_n+t}}{B(p_n + t,\, \gamma(x - h_n u) + 1)} \to 0$$

as $n \to \infty$, uniformly in $u \in \mathcal{B}$. Consequently, for all n large enough, uniformly in $u \in \mathcal{B}$,

$$|\mathcal{I}_{\gamma,\,\delta}(p_n + t,\, x - h_n u)| \leq \varepsilon\, B(p_n + t,\, \gamma(x - h_n u) + 1).$$

It is now clear that for all $\varepsilon > 0$, for sufficiently large n, $\left|E_n^{(\gamma,\,W)}(q,\, p_n + t,\, x)\right| \leq \varepsilon$, which is what we wanted to prove.

To show that $(n,\, t) \mapsto E_n^{(\gamma,\,W)}(q,\, p_n + t,\, x) \in \mathcal{C}_2$, notice that

$$E_n^{(\gamma,\,W)}(q,\, p_n,\, x) = \frac{\displaystyle\int_{\mathcal{B}} (fW)(x - h_n u)\, \frac{g^{p_n}(x - h_n u)}{g^{p_n}(x)}\, \mathcal{I}_{\gamma,\,\delta}(p_n,\, x - h_n u)\, K^q(u)\, du}{\displaystyle\int_{\mathcal{B}} (fW)(x - h_n u)\, \frac{g^{p_n}(x - h_n u)}{g^{p_n}(x)}\, B(p_n,\, \gamma(x - h_n u) + 1)\, K^q(u)\, du}.$$

Using (3.28) and [103] again, Lemma 2.16iii) yields

$$(n,\, t) \mapsto (p_n + t)^{\gamma(x)+1} B(p_n + t,\, \gamma(x - h_n u) + 1) \in \mathcal{E}_2;$$

the convergences this sequence satisfies being uniform in $u \in \mathcal{B}$, we deduce that

$$(n,\, t) \mapsto (p_n + t)^{\gamma(x)+1} \int_{\mathcal{B}} (fW)(x - h_n u)\, \frac{g^{p_n+t}(x - h_n u)}{g^{p_n+t}(x)}\, B(p_n + t,\, \gamma(x - h_n u) + 1)\, K^q(u)\, du$$

is an element of $\mathcal{E}_2$. It follows from Lemma 2.16i) that it is sufficient to prove that

$$(n,\, t) \mapsto (p_n + t)^{\gamma(x)+1} \int_{\mathcal{B}} (fW)(x - h_n u)\, \frac{g^{p_n+t}(x - h_n u)}{g^{p_n+t}(x)}\, \mathcal{I}_{\gamma,\,\delta}(p_n + t,\, x - h_n u)\, K^q(u)\, du$$

lies in $\mathcal{D}_2$. Pick $u \in \mathcal{B}$; it is enough to show that

$$(n,\, t) \mapsto (p_n + t)^{\gamma(x)+1}\, \mathcal{I}_{\gamma,\,\delta}(p_n + t,\, x - h_n u)$$

belongs to $\mathcal{D}_2$ and the convergences it satisfies are uniform in $u \in \mathcal{B}$. Because

$$(p_n + t)^{\gamma(x)+1} = p_n^{\gamma(x)+1}\,(1 + t/p_n)^{\gamma(x)+1}$$

and $(n,\,t) \mapsto (1+t/p_n)^{\gamma(x)+1} \in \mathcal{E}_2$, Lemma 2.16i) and ii) make it enough to show the latter property for

$$w_n(t) = p_n^{\gamma(x)+1}\,\mathcal{I}_{\gamma,\,\delta}(p_n + t,\, x - h_n u).$$

For all $t > 0$, let $R_t : [1,\,\infty) \to [0,\,\infty)$ be the function defined by

$$\forall\, y \geq 1, \quad R_t(y) = y\left\{1 - \left[1 - \frac{1}{y}\right]^t\right\}.$$

For all $t > 0$, R_t is a bounded Borel function on $[1,\,\infty)$, and one has, for all $0 \leq t_1 < t_2$ and all $t_3 > 0$,

$$\begin{aligned} p_n[w_n(t_2) - w_n(t_1)] &= -p_n^{\gamma(x)+2}\,\mathcal{I}_{\gamma+1,\,\delta R_{t_2-t_1}}(p_n + t_1,\, x - h_n u), \\ p_n^2[w_n(t_2+t_3) - w_n(t_2) - \{w_n(t_1+t_3) - w_n(t_1)\}] &= p_n^{\gamma(x)+3}\,\mathcal{I}_{\gamma+2,\,\delta R_{t_2-t_1} R_{t_3}}(p_n + t_1,\, x - h_n u). \end{aligned}$$

Because for all $t_1,\, t_2 \geq 0$ and every bounded Borel function S on $\Omega \times [1,\,\infty)$ we have, as $n \to \infty$,

$$p_n^{\gamma(x)+t_2+1}\,\mathcal{I}_{\gamma+t_2,\,\delta S}(p_n + t_1,\, x - h_n u) \to 0$$

uniformly in $u \in \mathcal{B}$, Lemma 3.4ii) is proven. ■

Proof of Lemma 3.5. We closely follow the proof of Lemma 2.13i). Let us first remark that, from Lemma 3.3i) and iii) with $q = 1$,

$$\mu_{p_n+1}(x) = f(x)\,C(x)\,\Gamma(\alpha(x)+1)\,g^{p_n+1}(x)\,p_n^{-\alpha(x)}\,(1 + \mathrm{o}(1)),$$

leading to

$$\mu_{p_n+1}(x)\,u_{n,\,a}(x) = \frac{g(x)}{a\|K\|_2}\sqrt{\frac{f(x)\,C(x)}{V(\alpha(x),\,a)}}\,v_n(x)\,(1 + \mathrm{o}(1)), \tag{3.30}$$

and therefore

$$\xi_n(x) = \frac{\mu_{p_n+1}(x)\,u_{n,\,a}(x)}{p_n+1}\,ap_n\left(\frac{1}{\widehat{g}_n(x)} - \frac{1}{G_n(x)}\right)(1 + \mathrm{o}(1)). \tag{3.31}$$

Besides,

$$\begin{aligned} ap_n\left(\frac{1}{\widehat{g}_n(x)} - \frac{1}{G_n(x)}\right) &= ((a+1)p_n+1)\frac{\widehat{\mu}_{(a+1)p_n}(x)\,\mu_{(a+1)p_n+1}(x) - \mu_{(a+1)p_n}(x)\,\widehat{\mu}_{(a+1)p_n+1}(x)}{\widehat{\mu}_{(a+1)p_n+1}(x)\,\mu_{(a+1)p_n+1}(x)} \\ &- (p_n+1)\frac{\widehat{\mu}_{p_n}(x)\,\mu_{p_n+1}(x) - \mu_{p_n}(x)\,\widehat{\mu}_{p_n+1}(x)}{\widehat{\mu}_{p_n+1}(x)\,\mu_{p_n+1}(x)} \\ &=: \Delta_n^{(1)}(x) - \Delta_n^{(2)}(x) \end{aligned}$$

with

$$\Delta_n^{(1)}(x) := \frac{(a+1)p_n+1}{\mu_{(a+1)p_n+1}(x)} \frac{\mu_{(a+1)p_n+1}(x)}{\widehat{\mu}_{(a+1)p_n+1}(x)} \left(\nu_{(a+1)p_n}(x) - \frac{\mu_{(a+1)p_n}(x)}{\mu_{(a+1)p_n+1}(x)} \nu_{(a+1)p_n+1}(x) \right),$$
$$\Delta_n^{(2)}(x) := \frac{p_n+1}{\mu_{p_n+1}(x)} \frac{\mu_{p_n+1}(x)}{\widehat{\mu}_{p_n+1}(x)} \left(\nu_{p_n}(x) - \frac{\mu_{p_n}(x)}{\mu_{p_n+1}(x)} \nu_{p_n+1}(x) \right),$$

which leads to

$$\frac{\mu_{p_n+1}(x)}{p_n+1} \Delta_n^{(1)}(x) = \left(1 + \frac{ap_n}{p_n+1}\right) \frac{\mu_{(a+1)p_n+1}(x)}{\widehat{\mu}_{(a+1)p_n+1}(x)} \zeta_n^{(3)}(x),$$
$$\frac{\mu_{p_n+1}(x)}{p_n+1} \Delta_n^{(2)}(x) = -\frac{\mu_{p_n+1}(x)}{\widehat{\mu}_{p_n+1}(x)} \zeta_n^{(2)}(x).$$

Replacing in (3.31) concludes the proof of Lemma 3.5. ■

Proof of Lemma 3.6. Conditioning on X yields

$$\mathbb{E}|S_n(x)|^3 = \frac{1}{n^3} \int_{\mathbb{R}^d} \mathbb{E}\left[\left| \sum_{j=0}^{k} b_{n,j} \, Y^{p_n+j} \, K_{h_n}(x-v) \right|^3 \Big| X = v \right] f(v)\, dv$$
$$= \frac{1}{n^3 h_n^{2d}} \int_{\mathcal{B}} \mathbb{E}\left[\left| \sum_{j=0}^{k} b_{n,j} \, Y^{p_n+j} \right|^3 \Big| X = x - h_n u \right] K^3(u)\, f(x-h_n u)\, du.$$

Now, given $X = x - h_n u$, we have $W_n(x, u) := \dfrac{Y}{g(x-h_n u)} \leq 1$; setting

$$c_n(x) := (m+1)^2 \sup_{\substack{[0,1] \\ 0 \leq j \leq m \\ n \in \mathbb{N} \setminus \{0\}}} |H_{n,j}|^3 \sup_{u \in \mathcal{B}} \left\{ \frac{g^{3p_n}(x-h_n u)}{g^{3p_n}(x)} \right\},$$

which, according to (3.26), is a bounded sequence, Hölder's inequality entails, given $X = x - h_n u$,

$$\left| \sum_{j=0}^{k} b_{n,j} \, Y^{p_n+j} \right|^3 = g^{3p_n}(x-h_n u) \left| W_n^{p_n}(x, u) \sum_{j=0}^{k} b_{n,j} \, W_n^j(x, u) g^j(x-h_n u) \right|^3$$
$$\leq c_n(x) \, g^{3p_n}(x) \sum_{j=0}^{m} \frac{1}{p_n^{3j}} W_n^{3p_n}(x, u)(1-W_n(x, u))^{3(m-j)}.$$

It is therefore sufficient to prove that, for all $j \in \{0, \ldots, m\}$, uniformly in $u \in \mathcal{B}$,

$$\mathbb{E}(W_n^{3p_n}(x, u)(1-W_n(x, u))^{3(m-j)} \,|\, X = x - h_n u) = \mathrm{O}\left(p_n^{-\alpha(x)-(3m-3j)}\right).$$

Define $\overline{\mathcal{F}}(y \,|\, v) = \overline{F}(g(v)y \,|\, v)$ as in the proof of Lemma 3.2: using an integration by parts, one has

$\mathbb{E}(W_n^{3p_n}(x, u)(1-W_n(x, u))^{3(m-j)} \,|\, X = x - h_n u)$

$$= \int_0^1 \frac{d}{dy} \left[y^{3p_n} \, (1-y)^{3(m-j)} \right] \overline{\mathcal{F}}(y \,|\, x - h_n u)\, dy$$

since, given $X = x - h_n u$, $W_n(x,\, u)$ has survival function $\overline{\mathcal{F}}(\cdot \,|\, x - h_n u)$. To conclude, notice that if γ is a positive locally Hölder continuous function on $\mathbb{R}^d$ and (s_n) is a positive sequence tending to ∞ such that $s_n\, h_n^{\eta_g} \to 0$ as $n \to \infty$, we obtain following (3.28) and [103]

$$\int_0^1 y^{s_n}\, \overline{F}_{1,\,\gamma}(y \,|\, x - h_n u)\, dy = B(s_n + 1,\, \gamma(x - h_n u) + 1) = \mathrm{O}\left(s_n^{-\gamma(x)-1}\right)$$

uniformly in $u \in \mathcal{B}$. Since

$$\overline{\mathcal{F}}(y \,|\, v) = C(v)\, \overline{F}_{1,\,\alpha}(y \,|\, v) + D(v)\, \overline{F}_{1,\,\alpha+\beta}(y \,|\, v) \left[1 + \delta\left(v,\, (1-y)^{-1}\right)\right],$$

using Lemmas 3.3i), 3.3ii) and 3.4ii) yields

$$\mathbb{E}(W_n^{3p_n}(x,\, u)(1 - W_n(x,\, u))^{3(m-j)} \,|\, X = x - h_n u) = \mathrm{O}\left(p_n^{-\alpha(x)-(3m-3j)}\right)$$

uniformly in $u \in \mathcal{B}$, which ends the proof of Lemma 3.6. ■

Conclusion et perspectives

Dans ce travail, divisé en deux parties, on a dans un premier temps présenté un nouveau modèle de processus de pertes en assurance, ainsi qu'un procédé permettant l'estimation des paramètres de ce modèle. L'algorithme EM développé dans ce contexte a été appliqué à la fois à des données réelles dans le cas univarié et à des données simulées dans le cas multivarié, permettant ainsi de mettre en évidence l'utilité de ce processus en pratique.

Les résultats obtenus à cette occasion ouvrent différentes perspectives de travail, par exemple :

Modèle à chocs communs sur des données réelles : Le modèle présenté est un modèle avec possibilité de chocs communs entre les différentes branches d'activité. La technique d'estimation proposée a été appliquée à des données réelles, mais seulement dans le cas univarié, les essais ayant été faits sur simulations dans le cas des chocs communs. On pourrait utiliser cette méthode pour analyser des données réelles dans ce cadre.

Procédure d'estimation initiale des paramètres : L'algorithme EM utilisé étant un algorithme itératif, il nécessite l'utilisation d'une estimation initiale. La procédure décrite est adaptée dans les cas particuliers que nous avons étudiés, mais n'est pas universelle. Un thème de recherche possible est donc de développer une procédure d'estimation initiale qui soit efficace dans davantage de situations.

Saisonnalité : De nombreux phénomènes considérés en assurance, par exemple les épisodes climatiques majeurs, présentent une certaine forme de saisonnalité. Le modèle présenté n'en tient pas compte. Il serait donc intéressant d'y remédier en proposant une nouvelle procédure d'estimation et d'en examiner les performances.

Effet de tendance : Certaines catégories de sinistres considérés en assurance, comme les accidents de la route, voient leur fréquence augmenter ou diminuer en moyenne au cours du temps. En d'autres termes, les intensités de survenue des sinistres, dans chaque état du processus de Markov sous-jacent, ne sont pas constantes, mais sont des fonctions croissantes ou décroissantes du temps. On pourrait envisager de développer là aussi une technique d'estimation et l'utiliser sur des cas pratiques.

Dans un deuxième temps, on a introduit une nouvelle méthode pour estimer le point terminal d'une distribution. Les propriétés tant théoriques que pratiques des estimateurs obtenus ont été étudiées. Une généralisation à l'estimation du support d'un couple aléatoire a été proposée dans le cas avec covariable.

Cette deuxième partie ouvre elle aussi un certain nombre d'axes de recherche, dont on donne une liste non exhaustive ci-dessous.

Adaptation au cas fonctionnel : Dans le cas multivarié, notre estimateur a été défini dans le cadre d'une covariable de dimension finie. On pourrait réfléchir à la généralisation de cette construction au cas d'une covariable fonctionnelle, ce qui présente notamment un intérèt en économétrie.

Estimation de l'indice des valeurs extrêmes : En utilisant des moments d'ordre élevé, on a fourni plusieurs méthodes d'estimation du point terminal θ, dont les propriétés asymptotiques ont été étudiées dans le cadre du domaine d'attraction de Weibull. Une piste intéressante serait de développer et étudier des estimateurs de l'indice des valeurs extrêmes γ basés sur cette technique, particulièrement dans le domaine d'attraction de Fréchet, et de les comparer à des estimateurs classiques de la théorie des valeurs extrêmes.

Etude d'autres estimateurs : Dans le cas univarié, rappelons que si Y est une variable aléatoire positive ayant pour fonction de survie $\overline{G}$ avec $\overline{G}(y) = (1 - y/\theta)^\alpha$ pour tout $y \in [0,\, \theta]$, alors

$$\forall\, p \geq 1, \quad A(p) := (p+1)\frac{\mathbb{E}(Y^p)}{\mathbb{E}(Y^{p+1})} = \frac{p+1+\alpha}{\theta}.$$

Par conséquent

$$\forall\, x \in [0,\, \infty), \quad \forall\, p \geq 1, \quad A(p(x+1)) - A(p) = \frac{px}{\theta}.$$

Si $H : [0,\, \infty) \to [0,\, \infty)$ est une fonction positive à variation bornée telle que $\displaystyle\int_0^\infty x\, dH(x) < \infty$, on a donc

$$\forall\, p \geq 1, \quad \frac{1}{\theta} = \frac{1}{p}\,\frac{\displaystyle\int_0^\infty [A(p(x+1)) - A(p)]\; dH(x)}{\displaystyle\int_0^\infty x\, dH(x)}.$$

En posant

$$\widehat{\mu}_{p_n} = \frac{1}{n}\sum_{k=1}^n X_k^{p_n}$$

puis

$$\widehat{A}(p_n) = (p_n+1)\frac{\widehat{\mu}_{p_n}}{\widehat{\mu}_{p_n+1}},$$

on pourrait alors étudier des estimateurs de $1/\theta$ de la forme

$$\frac{1}{\widehat{\theta}_n} = \frac{1}{p_n} \frac{\displaystyle\int_0^\infty \left[\widehat{A}(p_n(x+1)) - \widehat{A}(p_n)\right] dH(x)}{\displaystyle\int_0^\infty x\, dH(x)}.$$

Notons que, si $H = \mathbb{1}_{[a,\,\infty)}$,

$$\frac{1}{p_n} \frac{\displaystyle\int_0^\infty \left[\widehat{A}(p_n(x+1)) - \widehat{A}(p_n)\right] dH(x)}{\displaystyle\int_0^\infty x\, dH(x)} = \frac{1}{ap_n}\left[((a+1)p_n+1)\frac{\widehat{\mu}_{(a+1)p_n}}{\widehat{\mu}_{(a+1)p_n+1}} - (p_n+1)\frac{\widehat{\mu}_{p_n}}{\widehat{\mu}_{p_n+1}}\right],$$

de sorte que ces estimateurs sont des généralisations de notre estimateur dans le cas où X est positive.

Choix des paramètres : Les estimateurs des moments d'ordre élevé présentés ici ont plusieurs paramètres : a et la suite (p_n) dans le cas univarié, ainsi que la fenêtre (h_n) dans le cas multivarié. Pour pouvoir utiliser ces estimateurs en pratique, il serait intéressant de donner des techniques de sélection de ces paramètres.

Convergence dans L^1 dans le cas multivarié : Dans le cas multivarié, on a prouvé la consistance et la normalité asymptotique ponctuelles de l'estimateur $\widehat{g}_n(x)$. Sur simulations, il apparaît que notre estimateur possède des propriétés satisfaisantes en moyenne d'ordre 1. Il serait intéressant d'essayer de donner des propriétés asymptotiques de l'erreur L^1 de notre estimateur.

Bibliography

[1] Aarssen, K., de Haan, L. (1994). On the maximal life span of humans. *Math. Popul. Stud.* **4**(4):259–281.

[2] Abramovitz, M., Stegun, I. (1965). *Handbook of Mathematical Functions*, Dover.

[3] Alvarado, E., Sandberg, D.V., Pickford, S.G. (1998). Modeling large forest fires as extreme events. *Northwest Sci.* **72**:66–75.

[4] Aragon, Y., Daouia, A., Thomas-Agnan, C. (2005). Nonparametric frontier estimation: a conditional quantile-based approach. *Econometric Theory* **21**(2):358–389.

[5] Asmussen, S. (1989). Risk theory in a Markovian environment. *Scand. Actuar. J.* **2**:69–100.

[6] Asmussen, S. (2000). *Ruin probabilities*, World Scientific.

[7] Athreya, K.B., Fukuchi, J. (1997). Confidence intervals for endpoints of a c.d.f. via bootstrap. *J. Statist. Plann. Inference* **58**:299–320.

[8] Balkema, A., de Haan, L. (1974). Residual life time at great age. *Ann. Probab.* **2**(5):792–804.

[9] Baum, L.E., Petrie, T. (1966). Statistical inference for probabilistic functions of finite-state Markov chains. *Ann. Math. Statist.* **37**(6):1554–1563.

[10] Baum, L.E., Petrie, T., Soules, G., Weiss, N. (1970). A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains. *Ann. Math. Statist.* **41**(1):164–171.

[11] Beirlant, J., Goegebeur, Y., Segers, J., Teugels, J. (2004). *Statistics of Extremes*, John Wiley and Sons.

[12] Bickel, P.J., Ritov, Y., Rydén, T. (1998). Asymptotic normality of the maximum-likelihood estimator for general hidden Markov models. *Ann. Statist.* **26**(4):1614–1635.

[13] Billingsley, P. (1979). *Probability and measure*, John Wiley and Sons.

[14] Bingham, N.H., Goldie, C.M., Teugels, J.L. (1987). *Regular Variation*, Cambridge University Press.

[15] Buch-Larsen, T., Nielsen, J.P., Guillén, M., Bolancé, C. (2005). Kernel density estimation for heavy-tailed distributions using the Champernowne transformation. *Statistics* **39**(6):503–518.

[16] Cai, J., Tan, K.S. (2007). Optimal retention for a stop-loss reinsurance under the VaR and CTE risk measure. *ASTIN Bull.* **37**(1):93–112.

[17] Caliebe, A. (2006). Properties of the maximum *a posteriori* path estimator in hidden Markov models. *IEEE Trans. Inform. Theory* **52**(1):41–51.

[18] Cappé, O., Moulines, E., Rydén, T. (2005). *Inference in Hidden Markov Models*, Springer.

[19] Cazals, C., Florens, J.-P., Simar, L. (2002). Nonparametric frontier estimation: a robust approach. *J. Econometrics* **106**(1):1–25.

[20] Champernowne, D.G. (1937). The theory of income distribution. *Econometrica* **5**(4):379–381.

[21] Champernowne, D.G. (1952). The graduation of income distributions. *Econometrica* **20**(4):591–615.

[22] Chiragiev, A., Landsman, Z. (2007). Multivariate Pareto portfolios: TCE-based capital allocation and divided differences. *Scand. Actuar. J.* **2007**(4):261–280.

[23] Chow, Y.S., Teicher, H. (1997). *Probability Theory*, Springer.

[24] Churchill, G.A. (1989). Stochastic models for heterogeneous DNA sequences. *Bull. Math. Biol.* **51**(1):79–94.

[25] Çinlar, E. (1975). *Introduction to stochastic processes*, Prentice-Hall.

[26] Cooke, P. (1979). Statistical inference for bounds of random variables. *Biometrika* **66**:367–374.

[27] Daouia, A., Simar, L. (2005). Robust nonparametric estimators of monotone boundaries. *J. Multivariate Anal.* **96**:311–331.

[28] Davison, A.C., Ramesh, N.I. (1993). A stochastic model for times of exposures to air pollution from a point source, in *Statistics for the environment, editors: V. Barnett and K.F. Turkman*, John Wiley and Sons.

[29] Dekkers, A.L.M., Einmahl, J.H.J., de Haan, L. (1989). A moment estimator for the index of an extreme-value distribution. *Ann. Statist.* **17**(4):1833–1855.

[30] Dempster, A.P., Laird, N.M., Rubin, D.B. (1977). Maximum likelihood from incomplete data via the EM algorithm (with discussion). *J. Roy. Statist. Soc. Ser. B* **39**(1):1–38.

[31] Deng, L., Mark, J.W. (1993). Parameter estimation for Markov modulated Poisson processes via the EM algorithm with time discretization. *Telecomm. Syst.* **1**(1):321–338.

[32] Deprins, D., Simar, L., Tulkens, H. (1984). Measuring labor efficiency in post offices, in P. Pestieau, M. Marchand and H. Tulkens, editors, *The Performance of Public Enterprises: Concepts and Measurements*, North Holland.

[33] Diebolt, J., Guillou, A., Rached, I. (2004). A new look at probability-weighted moments estimators. *C. R. Acad. Sci. Paris* **338**: Série I, 629–634.

[34] Diebolt, J., Guillou, A., Rached, I. (2007). Approximation of the distribution of excesses through a generalized probability-weighted moments method. *J. Statist. Plann. Inference* **137**(3):841–857.

[35] Drees, H., Ferreira, A., de Haan, L. (2004). On maximum likelihood estimation of the extreme value index. *Ann. Appl. Probab.* **14**(3):1179–1201.

[36] Farrell, M.J. (1957). The measurement of productive efficiency. *J. Roy. Statist. Soc. Ser. A* **120**:253–281.

[37] Ferraty, F., Vieu, P. (2005). *Nonparametric modelling for functional data*, Springer.

[38] Fisher, R.A., Tippett, L.H.C. (1928). Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Proc. Camb. Phil. Soc.* **24**:180–190.

[39] Gardes, L. (2002). Estimating the support of a Poisson process via the Faber-Schauder basis and extreme values. *Publ. Inst. Statist. Univ. Paris* XXXXVI:43–72.

[40] Geffroy, J. (1964). Sur un problème d'estimation géométrique. *Publ. Inst. Statist. Univ. Paris* XIII:191–210.

[41] Gijbels, I., Mammen, E., Park, B.U., Simar, L. (1999). On estimation of monotone and concave frontier functions. *J. Amer. Statist. Assoc.* **94**(445):220–228.

[42] Girard, S., Iouditski, A., Nazin, A. (2005). L_1-optimal nonparametric frontier estimation via linear programming. *Autom. Remote Control* **66**(12):2000–2018.

[43] Girard, S., Jacob, P. (2003). Extreme values and Haar series estimates of point process boundaries. *Scand. J. Stat.* **30**(2):369–384.

[44] Girard, S., Jacob, P. (2003). Projection estimates of point processes boundaries. *J. Statist. Plann. Inference* **116**(1):1–15.

[45] Girard, S., Jacob, P. (2004). Extreme values and kernel estimates of point processes boundaries. *ESAIM Probab. Stat.* **8**:150–168.

[46] Girard, S., Jacob, P. (2008). Frontier estimation via kernel regression on high power-transformed data. *J. Multivariate Anal.* **99**(3):403–420.

[47] Girard, S., Jacob, P. (2009). Frontier estimation with local polynomials and high power-transformed data. *J. Multivariate Anal.* **100**(8):1691–1705.

[48] Girard, S., Menneteau, L. (2005). Central limit theorems for smoothed extreme value estimates of Poisson point processes boundaries. *J. Statist. Plann. Inference* **135**(2):433–460.

[49] Gnedenko, B.V. (1943). Sur la distribution limite du terme maximum d'une série aléatoire. *Ann. Math.* **44**:423–453.

[50] Goldenshluger, A., Tsybakov, A. (2004). Estimating the endpoint of a distribution in the presence of additive observation errors. *Statist. Probab. Lett.* **68**:39–49.

[51] Grandell, J. (1976). *Doubly stochastic Poisson processes*, Springer.

[52] Greenwood, J.A., Landwehr, J.M., Matalas, N.C., Wallis, J.R. (1979). Probability-weighted moments: definition and relation to parameters of several distributions expressable in inverse form. *Water Resour. Res.* **15**:1049–1054.

[53] Guillou, A., Loisel, S., Stupfler, G. (2011). Asymptotic normality of the maximum likelihood estimator of a loss process, available on the webpage http://www-irma.u-strasbg.fr/ guillou/supplement.pdf.

[54] Gusella, R. (1991). Characterizing the variability of arrival processes with indexes of dispersion. *IEEE J. Sel. Areas Commun.* **9**(2):203–211.

[55] de Haan, L. (1970). *On Regular Variation and Its Applications to the Weak Convergence of Sample Extremes*, Mathematical Centre Tract 32, Mathematics Centre, Amsterdam, Holland.

[56] de Haan, L. (1981). Estimation of the minimum of a function using order statistics. *J. Amer. Statist. Assoc.* **76**:467–469.

[57] de Haan, L., Ferreira, A. (2006). *Extreme Value Theory*, Springer.

[58] Hall, P. (1982). On estimating the endpoint of a distribution. *Ann. Statist.* **10**(2):556–568.

[59] Hall, P., Park, B.U. (2004). Bandwidth choice for local polynomial estimation of smooth boundaries. *J. Multivariate Anal.* **91**(2):240–261.

[60] Hall, P., Park, B.U., Stern, S.E. (1998). On polynomial estimators of frontiers and boundaries. *J. Multivariate Anal.* **66**(1):71–98.

[61] Hall, P., Wang, J.Z. (1999). Estimating the end-point of a probability distribution using minimum-distance methods. *Bernoulli* **5**(1):177–189.

[62] Hall, P., Wang, J.Z. (2005). Bayesian likelihood methods for estimating the end point of a distribution. *J. Roy. Statist. Soc. Ser. B* **67**(5):717–729.

[63] Härdle, W., Park, B.U., Tsybakov, A. (1995). Estimation of non-sharp support boundaries. *J. Multivariate Anal.* **55**:205–218.

[64] Heffes, H., Lucantoni, D. (1986). A Markov modulated characterization of packetized voice and data traffic and related statistical multiplexer performance. *IEEE J. Sel. Areas Commun.* **4**:856–867.

[65] Hill, B.M. (1975). A simple general approach to inference about the tail of a distribution. *Ann. Statist.* **3**:1163–1174.

[66] Holst, U., Lindgren, G. (1995). Recursive estimation of parameters in Markov-modulated Poisson processes. *IEEE Trans. Commun.* **43**(11):2812–2820.

[67] Hosking, J.R.M., Wallis, J.R. (1987). Parameter and quantile estimation for the generalized Pareto distribution. *Technometrics* **29**:339–349.

[68] Jacob, P., Suquet, C. (1995). Estimating the edge of a Poisson process by orthogonal series. *J. Statist. Plann. Inference* **46**:215–234.

[69] Johnson, N.L., Kotz, S., Balakrishnan, N. (1994). *Continuous Univariate Distributions*, Vol. 1, 2nd ed., John Wiley and Sons.

[70] Juang, B.H., Rabiner, L.R. (1991). Hidden Markov models for speech recognition. *Technometrics* **33**:251–272.

[71] Kawashima, K., Saito, H. (1990). Teletraffic issues in ATM networks. *Comp. Networks ISDN Sys.* **20**:369–375.

[72] Kim, S., Shephard, N, Chib, S. (1998). Stochastic volatility: Likelihood inference and comparison with ARCH models. *Rev. Econom. Stud.* **65**:361–394.

[73] Kingman, J.F.C. (1976). Subadditive processes, in *Ecole d'Eté de Probabilités de Saint-Flour V-1975*, volume 539 of *Lecture Notes in Mathematics*, Springer.

[74] Klugman, S.A., Panjer, H.H., Willmot, G.E. (2004). *Loss models*, John Wiley and Sons.

[75] Knight, K. (2001). Limiting distributions of linear programming estimators. *Extremes* **4**(2):87–103.

[76] Korostelev, A., Simar, L., Tsybakov, A. (1995). Efficient estimation of monotone boundaries. *Ann. Statist.* **23**(2):476–489.

[77] Korostelev, A., Tsybakov, A. (1993). *Minimax theory of image reconstruction*, volume 82 of *Lecture Notes in Statistics*, Springer.

[78] Leroux, B.G. (1992). Maximum-likelihood estimation for hidden Markov models. *Stochastic Process. Appl.* **40**:127–143.

[79] Li, D., Peng, L. (2009). Does bias reduction with external estimator of second order parameter work for endpoint? *J. Statist. Plann. Inference* **139**:1937–1952.

[80] Lindskog, F., McNeil, A. (2003). Common Poisson shock models: applications to insurance and credit risk modelling. *ASTIN Bull.* **33**(2):209–238.

[81] Loh, W.Y. (1984). Estimating an endpoint of a distribution with resampling methods. *Ann. Statist.* **12**(4):1543–1550.

[82] Loisel, S. (2004). Ruin theory with K lines of business. *Proceedings of the 3rd Actuarial and Financial Day*, Brussels.

[83] Loisel, S. Milhaud, X. (2011). From deterministic to stochastic surrender risk models: impact of correlation crises on economic capital, available at http://hal.archives-ouvertes.fr/hal-00502847.

[84] McLachlan, G.J., Krishnan, T. (1997). *The EM Algorithm and Extensions*, John Wiley and Sons.

[85] Meier-Hellstern, K.S. (1987). A fitting algorithm for Markov-modulated Poisson processes having two arrival rates. *European J. Oper. Res.* **29**:970–977.

[86] Menneteau, L. (2008). Multidimensional limit theorems for smoothed extreme value estimates of point processes boundaries. *ESAIM Probab. Stat.* **12**:273–307.

[87] Milhaud, X., Loisel, S., Maume-Deschamps, V. (2010). Surrender triggers in life insurance: classification and risk predictions, available at http://hal.archives-ouvertes.fr/hal-00450003.

[88] Miller, R.G. (1964). A trustworthy jackknife. *Ann. Math. Statist.* **35**(4):1594–1605.

[89] Neves, C., Pereira, A. (2010). Detecting finiteness in the right endpoint of light-tailed distributions. *Statist. Probab. Lett.* **80**:437–444.

[90] Parzen, E. (1962). On estimation of a probability density function and mode. *Ann. Math. Statist.* **33**(3):1065–1076.

[91] Pickands, J. (1975). Statistical inference using extreme order statistics. *Ann. Statist.* **3**(1):119–131.

[92] Press, W.H., Flannery, B.P., Teukolsky, S.A., Vetterling, W.T. (1989). *Numerical Recipes*, Cambridge University Press.

[93] Robson, D.S., Whitlock, J.H. (1964). Estimation of a truncation point. *Biometrika* **51**:33–39.

[94] Rolski, T., Schmidli, H., Schmidt, V., Teugels, J. (1999). *Stochastic Processes for Insurance and Finance*, John Wiley and Sons.

[95] Rosenblatt, M. (1956). Remarks on some nonparametric estimates of a density function. *Ann. Math. Statist.* **27**(3):832–837.

[96] Rossiter, M.H. (1989). *Characterizing a random point process by a switched Poisson process*, PhD thesis, Monash University, Melbourne.

[97] Rydén, T. (1994). Parameter estimation for Markov modulated Poisson processes. *Commun. Statist. Stochastic Models* **10**(4):795–829.

[98] Rydén, T. (1994). Consistent and asymptotically normal parameter estimates for hidden Markov models. *Ann. Statist.* **22**(4):1884–1895.

[99] Rydén, T. (1995). Consistent and asymptotically normal parameter estimates for Markov modulated Poisson processes. *Scand. J. Stat.* **22**(3):295–303.

[100] Rydén, T. (1996). An EM algorithm for estimation in Markov-modulated Poisson processes. *Comput. Statist. Data Anal.* **21**(4):431–447.

[101] Smith, J.A. (1987). Statistical modeling of rainfall occurrences. *Water Resour. Res.* **23**(5):885–893.

[102] Smith, R.L. (1987). Estimating tails of probability distributions. *Ann. Statist.* **15**(3):1174–1207.

[103] Tricomi, F.G., Erdélyi, A. (1951). The asymptotic expansion of a ratio of gamma functions. *Pacific J. Math.* **1**:133–142.

[104] Vandekerkhove, P. (2005). Consistent and asymptotically normal parameter estimates for hidden Markov mixtures of Markov models. *Bernoulli* **11**(11):103–129.

[105] Viterbi, A. (1967). Error bounds for convolutional codes and an asymptotically optimum decoding algorithm. *IEEE Trans. Inform. Theory* **IT-13**(2):260–269.