



HAL
open science

Estimateur de la fonction de répartition bivariée avec censures à droite et à gauche

Agathe Guilloux, Philippe Saint-Pierre

► **To cite this version:**

Agathe Guilloux, Philippe Saint-Pierre. Estimateur de la fonction de répartition bivariée avec censures à droite et à gauche. *Annales de l'ISUP*, 2008, LII (1-2), pp.157-164. hal-03633194

HAL Id: hal-03633194

<https://hal.science/hal-03633194>

Submitted on 6 Apr 2022

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

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

Estimateur de la fonction de répartition bivariée avec censures à droite et à gauche

AGATHE GUILLOUX ET PHILIPPE SAINT-PIERRE
LSTA - Université Pierre et Marie Curie, Paris 6
175 rue du Chevaleret, 75013 PARIS, FRANCE.

agathe.guillox@upmc.fr, philippe.saint_pierre@upmc.fr

1 Introduction

Soit $T = (T_1, T_2)$ un couple de variable aléatoires (v.a.) à valeurs dans \mathbb{R}_+^2 . On note F la fonction de répartition (f.d.r.) de T . On considère que la v.a. T est censurée à gauche et à droite dans le sens où l'on observe les v.a. Y et δ définies par :

$$Y = (Y_1, Y_2) \text{ avec } Y_1 = T_1 \vee L \text{ et } Y_2 = T_2 \wedge C \\ \delta = (\delta_1, \delta_2) \text{ avec } \delta_1 = I(T_1 \geq L) \text{ et } \delta_2 = I(T_2 \leq C),$$

où (L, C) est un vecteur de v.a. positives de f.d.r. G , supposé indépendant du vecteur (T_1, T_2) . La v.a. T_1 est donc censurée à gauche alors que T_2 est censurée à droite. On cherche à estimer la f.d.r. F à partir d'un échantillon i.i.d. de la forme $(Y_i = (Y_{1,i}, Y_{2,i}), \delta_i = (\delta_{1,i}, \delta_{2,i}))$ pour $i = 1, \dots, n$.

Par exemple, dans le cas d'un patient infecté par le VIH, cette fonction pourra permettre d'étudier conjointement (i) la durée T_1 écoulée entre la consultation du patient infecté et la date de l'infection par le VIH (cette durée peut être censurée à gauche dans le cas où le patient sait uniquement qu'il n'était pas infecté à une certaine date) (ii) la durée T_2 écoulée entre la consultation et le début du SIDA (cette durée peut être censurée à droite, si le patient ne connaît pas la date exacte de son infection mais celle du résultat négatif d'un test antérieur de dépistage).

Pour autant que nous sachions, ce problème n'a pas encore été considéré dans la littérature. Drabrowska [4] a construit un estimateur de la f.d.r. bivariée F en considérant que les v.a. T_1 et T_2 étaient toutes deux censurées à droite. Ce problème a également été considéré par Pons et de Turkheim [8], Campbell et Földes [2] et Hougaard [7].

Le problème de l'estimation de la f.d.r. univariée sous censure aléatoire à droite a été intensivement étudié, nous ne citerons que Andersen *et al.* [1] pour une revue. Le même problème mais sous censure aléatoire à gauche a été beaucoup moins étudié. La plupart des auteurs (notamment Andersen *et al.* [1], Cox et Oakes [3]) proposent de "renverser" le temps, c'est-à-dire de considérer

$\tau - Y_1 = \tau - (T_1 \vee L) = (\tau - T_1) \wedge (\tau - L)$ au lieu de Y_1 , où τ est un réel positif choisi de sorte que les observations de $\tau - Y_1$ restent dans \mathbb{R}_+ . On peut se ramener au cas de la censure aléatoire à droite. Cette idée n'est par entièrement satisfaisante puisqu'elle exclut le cas où les v.a. T_1 (et Y_1) sont à support sur \mathbb{R}_+ . Une autre approche, proposée par Gomez *et al.* [5], consiste à écrire la f.d.r. F_1 de la v.a. T_1 comme la solution de l'équation intégrale de Doléans :

$$F_1(t) = F_1(\infty) - \int_t^\infty F_1(u) \frac{dH_1(u)}{H(u)} \text{ pour } t \geq 0$$

$$F_1(\infty) = 1$$

où H est la f.d.r. de la v.a. Y_1 et H_1 est définie par $H_1(t) = \mathbb{P}(Y_1 \leq t, \delta_1 = 1)$. Gomez *et al.* [5] ont montré que la solution de cette équation est donnée, pour $t \geq 0$, par :

$$F_1(t) = \prod_{s>t} \left(1 - \frac{dH_1(s)}{H(s)} \right),$$

où la notation \prod désigne le produit intégral défini par :

$$\prod_{s>t} (1 + dh(s)) = \prod_{s>t: \Delta h(s) \neq 0} (1 + \Delta h(s)) \exp\left(\int_t^\infty h^c(s) ds\right)$$

pour h fonction càdlàg à variation bornée, où h^c désigne la partie continue de h et $\Delta h(s) = h(s) - h(s-)$ (cf. Andersen *et al.* [1]).

2 Loi bivariée et fonctions de risque cumulatif

Notons \tilde{F} et \tilde{G} les fonctions définies pour tout $(s, t) \in \mathbb{R}_+$ par :

$$\begin{aligned} \tilde{F}(s, t) &= \mathbb{P}(T_1 \leq s, T_2 > t) \text{ et} \\ \tilde{G}(s, t) &= \mathbb{P}(L \leq s, C > t). \end{aligned}$$

Par définition des v.a. Y et δ et par indépendance entre (T_1, T_2) et (L, C) , les fonctions \tilde{H} , K_1 , K_2 et K_3 , définies ci-dessous, vérifient les relations suivantes pour tout $(s, t) \in \mathbb{R}_+$:

$$\tilde{H}(s, t) = \mathbb{P}(Y_1 \leq s, Y_2 > t) = \tilde{F}(s, t) \tilde{G}(s, t) \quad (1)$$

$$K_1(s, t) = \mathbb{P}(Y_1 \leq s, Y_2 > t, \delta_1 = 1, \delta_2 = 1) = - \int_0^s \int_t^\infty \tilde{G}(u, v-) \tilde{F}(du, dv)$$

$$K_2(s, t) = \mathbb{P}(Y_1 \leq s, Y_2 > t, \delta_1 = 1) = \int_0^s \tilde{G}(u, t) \tilde{F}(du, t)$$

$$K_3(s, t) = \mathbb{P}(Y_1 \leq s, Y_2 > t, \delta_2 = 1) = - \int_t^\infty \tilde{G}(s, v-) \tilde{F}(s, dv).$$

Si le couple (T_1, T_2) admet une densité f par rapport à la mesure de Lebesgue sur \mathbb{R}^2 , on a alors :

$$\begin{aligned}\tilde{F}(s, t) &= \int_0^s \int_t^\infty f(u, v) du dv \text{ soit} \\ \tilde{F}(du, dv) &= -f(u, v) du dv, \tilde{F}(du, t) = \int_t^\infty f(u, v) dv du \\ \text{et } \tilde{F}(s, dv) &= -\int_0^s f(u, v) du dv.\end{aligned}$$

Grâce à 1, on peut alors définir les fonctions de risque cumulé suivantes :

$$\begin{aligned}\Lambda_{11}(s, t) &= -\int_0^s \int_0^t \frac{\tilde{F}(du, dv)}{\tilde{F}(u, v-)} = -\int_0^s \int_0^t \frac{K_1(du, dv)}{\tilde{H}(u, v-)} \quad (2) \\ \Lambda_{10}(s, t) &= \int_0^s \frac{\tilde{F}(du, t)}{\tilde{F}(u, t)} = \int_0^s \frac{K_2(du, t)}{\tilde{H}(u, t)} \\ \Lambda_{01}(s, t) &= -\int_0^t \frac{\tilde{F}(s, dv)}{\tilde{F}(s, v-)} = -\int_0^t \frac{K_3(s, dv)}{\tilde{H}(s, v-)}.\end{aligned}$$

Les fonctions Λ_{11} , Λ_{10} , Λ_{01} sont donc estimables à l'aide des équivalents empiriques des fonctions \tilde{H} , K_1 , K_2 et K_3 . Ces versions empiriques sont calculables puisque les fonctions théoriques sont exprimées en fonction des v.a. Y et δ qui sont observables. Il reste donc à lier la fonction \tilde{F} aux fonctions Λ_{11} , Λ_{10} , Λ_{01} .

Définissons, pour (s, t) tels que $\tilde{F}(s, t) > 0$, la fonction A par $A(s, t) = \log \tilde{F}(s, t)$, alors on a :

$$\begin{aligned}\tilde{F}(s, t) &= \exp\{A(s, t)\} \\ &= \exp\left\{-\int_s^\infty \int_0^t A(du, dv) + A(\infty, t) + A(s, 0) - A(\infty, 0)\right\} \\ &= \tilde{F}(\infty, t) \tilde{F}(s, 0) \exp\left\{-\int_s^\infty \int_0^t A(du, dv)\right\}.\end{aligned}$$

La fonction $A = \log \tilde{F}$ étant une fonction à variation bornée, elle a un nombre fini de discontinuités appartenant à l'ensemble dénombrable des lignes orthogonales aux axes des coordonnées (Theorème 5.4, page 110 de Hildebrandt [6]).

Pour toute fonction bivariable Ψ , on note que /

$$\begin{aligned}\Psi(\Delta s, t) &= \Psi(s, t) - \Psi(s-, t) \\ \Psi(s, \Delta t) &= \Psi(s, t) - \Psi(s, t-) \\ \Psi(\Delta s, \Delta t) &= \Psi(s, t) - \Psi(s-, t) - \Psi(s, t-) + \Psi(s-, t-).\end{aligned}$$

Considérons les ensembles suivant,

$$\begin{aligned}E_1 &= \{(s, t) : A(s, t) < 0, A(\Delta s, t) = A(s, \Delta t) = 0\} \\ E_2 &= \{(s, t) : A(s, t) < 0, A(\Delta s, t) < 0, A(\Delta s, \Delta t) = 0\} \\ E_3 &= \{(s, t) : A(s, t) < 0, A(s, \Delta t) < 0, A(\Delta s, \Delta t) = 0\} \\ E_4 &= \{(s, t) : A(s, t) < 0, A(\Delta s, \Delta t) > 0\}.\end{aligned}$$

Les ensembles E_1 et E_4 correspondent au support, respectivement, de la composante continue et de la composante discrète de A . Les ensembles E_2 et E_3 sont les supports des composantes de A qui ont des discontinuités appartenant aux lignes orthogonales aux axes des coordonnées. D'après la décomposition de Jordan des fonctions à variations bornées dans le plan, on peut écrire

$$\begin{aligned}
 - \int_s^\infty \int_0^t A(du, dv) &= \sum_{i=1}^4 A_i(s, t), \text{ où} \\
 A_1(s, t) &= - \int_s^\infty \int_0^t I((u, v) \in E_1) A(du, dv) \\
 A_2(s, t) &= - \sum_{u>s} \int_0^t I((u, v) \in E_2) A(\Delta u, dv) \\
 A_3(s, t) &= - \sum_{v \leq t} \int_s^\infty I((u, v) \in E_3) A(du, \Delta v) \\
 A_4(s, t) &= - \sum_{u>s} \sum_{v \leq t} I((u, v) \in E_4) A(\Delta u, \Delta v).
 \end{aligned}$$

On peut montrer, par définition des fonctions de risque cumulé (voir les formules 2), que pour tout (s, t) tels que $\tilde{F}(s-, t) > 0$, on a :

$$\begin{aligned}
 \frac{\tilde{F}(s, t)}{\tilde{F}(s, t-)} &= \frac{\tilde{F}(s, t-) + \tilde{F}(s, \Delta t)}{\tilde{F}(s, t-)} = 1 - \Lambda_{01}(s, \Delta t), \\
 \frac{\tilde{F}(s-, t-)}{\tilde{F}(s, t-)} &= 1 - \Lambda_{10}(\Delta s, t-) \text{ et} \\
 \frac{\tilde{F}(s-, t)}{\tilde{F}(s, t-)} &= 1 - \Lambda_{01}(s, \Delta t) - \Lambda_{10}(\Delta s, t-) + \Lambda_{11}(\Delta s, \Delta t)
 \end{aligned}$$

Pour tout (s, t) tels que $\tilde{F}(s-, t) > 0$, en utilisant les égalités précédentes, on obtient

$$\begin{aligned}
 A_3(s, t) &= \sum_{v \leq t} \int_s^\infty I((u, v) \in E_3) \left[\frac{\tilde{F}(du, v-)}{\tilde{F}(u, v-)} - \frac{\tilde{F}(du, v)}{\tilde{F}(u, v)} \right] \\
 &= \sum_{v \leq t} \int_s^\infty I((u, v) \in E_3) \left[\frac{\tilde{F}(du, v-)\tilde{F}(u, \Delta v)}{\tilde{F}(u, v)\tilde{F}(u, v-)} - \frac{\tilde{F}(du, \Delta v)}{\tilde{F}(u, v)} \right] \\
 &= \sum_{v \leq t} \int_s^\infty I((u, v) \in E_3) \left[\frac{\tilde{F}(u, v-)}{\tilde{F}(u, v)} \left(\frac{\tilde{F}(du, v-)\tilde{F}(u, \Delta v)}{\tilde{F}(u, v-)\tilde{F}(u, v-)} - \frac{\tilde{F}(du, \Delta v)}{\tilde{F}(u, v-)} \right) \right] \\
 &= \sum_{v \leq t} \int_s^\infty I((u, v) \in E_3) \left[\frac{1}{1 - \Lambda_{01}(u, \Delta v)} [\Lambda_{11}(du, \Delta v) - \Lambda_{10}(du, v-)\Lambda_{01}(u, \Delta v)] \right]
 \end{aligned}$$

$$\begin{aligned}
 A_2(s, t) &= \sum_{u>s} \int_0^t I((u, v) \in E_2) \left[\frac{\tilde{F}(u-, dv)}{\tilde{F}(u-, v-)} - \frac{\tilde{F}(u, dv)}{\tilde{F}(u, v-)} \right] \\
 &= \sum_{u>s} \int_0^t I((u, v) \in E_2) \left[\frac{\tilde{F}(u, dv)\tilde{F}(\Delta u, v-)}{\tilde{F}(u, v-)\tilde{F}(u-, v-)} - \frac{\tilde{F}(\Delta u, dv)}{\tilde{F}(u-, v-)} \right] \\
 &= \sum_{u>s} \int_0^t I((u, v) \in E_2) \left[\frac{1}{1 - \Lambda_{10}(\Delta u, v-)} [\Lambda_{11}(\Delta u, dv) - \Lambda_{10}(\Delta u, v-)\Lambda_{01}(u, dv)] \right]
 \end{aligned}$$

$$\begin{aligned}
 A_4(s, t) &= \sum_{u>s} \sum_{v \leq t} I((u, v) \in E_4) \left[\log \left(\frac{\tilde{F}(u-, v)}{\tilde{F}(u, v-)} \right) - \log \left(\frac{\tilde{F}(u, v)}{\tilde{F}(u, v-)} \right) - \log \left(\frac{\tilde{F}(u-, v-)}{\tilde{F}(u, v-)} \right) \right] \\
 &= \sum_{u>s} \sum_{v \leq t} I((u, v) \in E_4) \left[\log \left(1 - \frac{\Lambda_{01}(u, \Delta v)\Lambda_{10}(\Delta u, v-) - \Lambda_{11}(\Delta u, \Delta v)}{[1 - \Lambda_{10}(\Delta u, v-)][1 - \Lambda_{01}(u, \Delta v)]} \right) \right]
 \end{aligned}$$

$$\begin{aligned}
 A_1(s, t) &= \int_s^\infty \int_0^t I((u, v) \in E_1) \left[\frac{\tilde{F}(du, v-)\tilde{F}(u, dv)}{\tilde{F}(u, v-)\tilde{F}(u, v-)} - \frac{\tilde{F}(du, dv)}{\tilde{F}(u, v-)} \right] \\
 &= \int_s^\infty \int_0^t I((u, v) \in E_1) [\Lambda_{11}(du, dv) - \Lambda_{10}(du, v-)\Lambda_{01}(u, dv)].
 \end{aligned}$$

D'après la définition des ensembles E_i , on a $\Lambda_{10}(\Delta u, v-) = 0$ pour tout $(u, v) \in E_1 \cup E_3$, et $\Lambda_{01}(u, \Delta v) = 0$ pour tout $(u, v) \in E_1 \cup E_2$. Ainsi, les égalités précédentes permettent d'obtenir le résultat suivant.

Proposition 1 Pour tout (s, t) tel que $\tilde{F}(s, t) > 0$,

$$\tilde{F}(s, t) = \tilde{F}(\infty, t)\tilde{F}(s, 0) \prod_{i=1}^4 \exp\{A_i(s, t)\}$$

avec

$$\exp\{A_i(s, t)\} = \exp \left\{ - \int_s^\infty \int_0^t I((u, v) \in E_i) L(du, dv) \right\}, \text{ pour } i = 1, 2, 3,$$

$$\exp\{A_4(s, t)\} = \prod_{\substack{u>s \\ (u, v) \in E_4}} \prod_{v \leq t} (1 - L(\Delta u, \Delta v))$$

$$L(du, dv) = \frac{\Lambda_{10}(du, v-)\Lambda_{01}(u, dv) - \Lambda_{11}(du, dv)}{[1 - \Lambda_{10}(\Delta u, v-)][1 - \Lambda_{01}(u, \Delta v)]}$$

et

$$\begin{aligned}
 \tilde{F}(s, 0) &= \prod_{u>s} (1 - \Lambda_{10}(du, 0)) \\
 \tilde{F}(\infty, t) &= \prod_{v \leq t} (1 - \Lambda_{01}(\infty, dv)).
 \end{aligned}$$

Dans le paragraphe suivant, nous définissons un estimateur de \tilde{F} grâce à la proposition 1 et montrons sa consistance forte.

3 Estimation

On observe un échantillon i.i.d. de la forme

$$Y_i = (Y_{1,i}, Y_{2,i}) \text{ et } \delta_i = (\delta_{1,i}, \delta_{2,i}) \text{ pour } i = 1, \dots, n.$$

Définissons l'estimateur \widehat{F} à partir de la proposition 1. Les fonctions \widehat{H} , K_1 , K_2 et K_3 sont estimées empiriquement, on définit donc pour tout $(s, t) \in \mathbb{R}_+^2$:

$$\widehat{H}(s, t) = \frac{1}{n} \sum_{i=1}^n I(Y_{i,1} \leq s, Y_{2,i} > t),$$

$$\widehat{K}_1(s, t) = \frac{1}{n} \sum_{i=1}^n I(Y_{i,1} \leq s, Y_{2,i} > t, \delta_{1,i} = 1, \delta_{2,i} = 1),$$

$$\widehat{K}_2(s, t) = \frac{1}{n} \sum_{i=1}^n I(Y_{i,1} \leq s, Y_{2,i} > t, \delta_{1,i} = 1) \text{ et}$$

$$\widehat{K}_3(s, t) = \frac{1}{n} \sum_{i=1}^n I(Y_{i,1} \leq s, Y_{2,i} > t, \delta_{2,i} = 1).$$

On peut alors définir les estimateurs des fonctions de risque cumulé pour tout $(s, t) \in \mathbb{R}_+^2$ par :

$$\widehat{\Lambda}_{11}(s, t) = - \int_0^s \int_0^t \frac{\widehat{K}_1(du, dv) I(\widehat{H}(u, v-) > 0)}{\widehat{H}(u, v-)},$$

$$\widehat{\Lambda}_{10}(s, t) = \int_0^s \frac{\widehat{K}_2(du, t) I(\widehat{H}(u, t) > 0)}{\widehat{H}(u, t)} \text{ et}$$

$$\widehat{\Lambda}_{01}(s, t) = - \int_0^t \frac{\widehat{K}_3(s, dv) I(\widehat{H}(s, v-) > 0)}{\widehat{H}(s, v-)},$$

avec la convention $0/0 = 0$. Nous sommes alors en mesure de définir l'estimateur de la fonction L par

$$\widehat{L}(\Delta u, \Delta v) = \frac{\widehat{\Lambda}_{10}(\Delta u, v-) \widehat{\Lambda}_{01}(u, \Delta v) - \widehat{\Lambda}_{11}(\Delta u, \Delta v)}{[1 - \widehat{\Lambda}_{10}(\Delta u, v-)][1 - \widehat{\Lambda}_{01}(u, \Delta v)]}.$$

Comme la fonction \widehat{L} est purement discrète par construction, on obtient finalement l'estimateur \widehat{F} de la fonction \widehat{F} en posant :

$$\widehat{F}(s, t) = \widehat{F}(\infty, t) \widehat{F}(s, 0) \prod_{u>s, 0<v\leq t} (1 - \widehat{L}(du, dv)), \quad (3)$$

où $t \mapsto \widehat{F}(\infty, t)$ est l'estimateur de Kaplan-Meier associé à l'observation de $(Y_{2,i}, \delta_{2,i})$ pour $i = 1, \dots, n$, enfin $t \mapsto \widehat{F}(s, 0)$ est l'estimateur de Kaplan-Meier dans le cas de censure à gauche (cf. Gomez *et al.* [5]) associé à l'observation de $(Y_{1,i}, \delta_{1,i})$ pour $i = 1, \dots, n$.

Pour la preuve de la convergence de l'estimateur, l'équation 3 ne suffit pas car il faut prouver la convergence sur les quatre ensembles E_1, \dots, E_4 (même s'ils sont inconnus). On décompose donc le produit dans l'équation 3 et on obtient :

$$\widehat{F}(s, t) = \widehat{F}(\infty, t) \widehat{F}(s, 0) \prod_{i=1}^4 \exp\{\widehat{A}_i(s, t)\}$$

avec

$$\exp\{\widehat{A}_i(s, t)\} = \exp\left\{-\int_s^\infty \int_0^t I((u, v) \in E_i) \widehat{L}(du, dv)\right\}, \text{ pour } i = 1, 2, 3,$$

$$\exp\{\widehat{A}_4(s, t)\} = \prod_{\substack{u > s \\ (u, v) \in E_4}} \prod_{\substack{v \leq t \\ (u, v) \in E_4}} (1 - \widehat{L}(\Delta u, \Delta v)),$$

On peut alors établir la consistance forte de l'estimateur \widehat{F} . Pour $\tau = (\tau_1, \tau_2)$, notons $\|\cdot\|_\tau$ denote la norme sup sur $[\tau_1, \infty) \times [0, \tau_2]$.

Proposition 2 *Supposons que le vecteur de v.a. (L, C) est indépendant du vecteur (T_1, T_2) et que $\tau = (\tau_1, \tau_2)$ vérifie $\widehat{H}(\tau_1, \tau_2) > 0$, alors*

$$\|\widehat{F} - \tilde{F}\|_\tau \xrightarrow{p.s.} 0.$$

Cette proposition découle de la consistance forte de l'estimateur de Kaplan-Meier $t \mapsto \widehat{F}(\infty, t)$ sur l'intervalle $[0, \tau_2]$ (cf. Andersen *et al.* [1]), de celle de l'estimateur de Kaplan-Meier à gauche $t \mapsto \widehat{F}(s, 0)$ sur l'intervalle $[\tau_1, \infty[$ (cf. Gomez *et al.* [5]) et des lemmes suivants.

Lemme 1 *Sous les hypothèses de la Proposition 2,*

$$\|\widehat{\Lambda}_{10} - \Lambda_{10}\|_\tau \xrightarrow{p.s.} 0, \|\widehat{\Lambda}_{01} - \Lambda_{01}\|_\tau \xrightarrow{p.s.} 0 \text{ et } \|\widehat{\Lambda}_{11} - \Lambda_{11}\|_\tau \xrightarrow{p.s.} 0.$$

Ce lemme est une simple conséquence du théorème de Glivenko-Cantelli appliqué aux quantités empiriques \widehat{H} , \widehat{K}_1 , \widehat{K}_2 et \widehat{K}_3 . La consistance des estimateurs $\exp\{\widehat{A}_i\}$ pour $i = 1, 2, 3$ est donnée par le lemme suivant.

Lemme 2 *Sous les hypothèses de la Proposition 2, pour $i = 1, 2, 3$:*

$$\|\exp\{\widehat{A}_i\} - \exp\{A_i\}\|_\tau \xrightarrow{p.s.} 0$$

Ce résultat est obtenu en utilisant le lemme 1 et des arguments similaires à ceux de Drabrowska [4]. Enfin, on peut montrer le lemme ci-dessous, en suivant à nouveau Drabrowska [4].

Lemme 3 *Sous les hypothèses de la Proposition 2,*

$$\|\exp\{\widehat{A}_4\} - \exp\{A_4\}\|_\tau \xrightarrow{p.s.} 0.$$

Références

- [1] Per Kragh Andersen, Ørnulf Borgan, Richard D. Gill, and Niels Keiding. *Statistical models based on counting processes*. Springer Series in Statistics. Springer-Verlag, New York, 1993.
- [2] G. Campbell and A. Földes. Large-sample properties of nonparametric bivariate estimators with censored data. In *Nonparametric statistical inference, Vol. I, II (Budapest, 1980)*, volume 32 of *Colloq. Math. Soc. János Bolyai*, pages 103–121. North-Holland, Amsterdam, 1982.
- [3] D. R. Cox and D. Oakes. *Analysis of survival data*. Monographs on Statistics and Applied Probability. Chapman & Hall, London, 1984.
- [4] Dorota M. Dabrowska. Kaplan-Meier estimate on the plane. *Ann. Statist.*, 16(4) :1475–1489, 1988.
- [5] Guadalupe Gómez, Olga Julià, and Frederic Utzet. Survival analysis for left censored data. In *Survival analysis : state of the art (Columbus, OH, 1991)*, volume 211 of *NATO Adv. Sci. Inst. Ser. E Appl. Sci.*, pages 269–288. Kluwer Acad. Publ., Dordrecht, 1992. With a discussion by Melvin L. Moeschberger and a rejoinder by the authors.
- [6] T. H. Hildebrandt. *Introduction to the theory of integration*. Pure and Applied Mathematics, Vol. XIII. Academic Press, New York, 1963.
- [7] Philip Hougaard. *Analysis of multivariate survival data*. Statistics for Biology and Health. Springer-Verlag, New York, 2000.
- [8] Odile Pons and Élisabeth de Turckheim. Tests of independence for bivariate censored data based on the empirical joint hazard function. *Scand. J. Statist.*, 18(1) :21–37, 1991.