

A Glivenko-Cantelli theorem for the empirical distribution function of uniform m-step spacings

Einmahl, J.H.J.

Published in:

Publications de l'Institut de Statistique de l'Universite de Paris

Publication date:

1987

[Link to publication](#)

Citation for published version (APA):

Einmahl, J. H. J. (1987). A Glivenko-Cantelli theorem for the empirical distribution function of uniform m-step spacings. *Publications de l'Institut de Statistique de l'Universite de Paris*, 32(4), 23-27.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Take down policy

If you believe that this document breaches copyright, please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Pub. Inst. Stat. Univ.
XXXII, fasc. 4, 1987, 23 à 27

A GLIVENKO-CANTELLI THEOREM FOR THE EMPIRICAL
DISTRIBUTION FUNCTION OF UNIFORM m -STEP SPACINGS

John H.J. Einmahl
University of Limburg, Maastricht

SUMMARY

In this note a Glivenko-Cantelli theorem for the empirical distribution function of uniform m -step spacings is derived, where m is allowed to tend to infinity with n at a certain rate. The short proof of this theorem is based on a powerful exponential probability bound.

AMS 1980 subject classification: 60F15.

Key words and phrases: empirical distribution function, Glivenko-Cantelli theorem, m -step uniform spacings.

INTRODUCTION AND NOTATION

Let U_1, U_2, \dots be a sequence of independent uniform $(0,1)$ random variables and let

$$0 := U_{0:n-1} \leq U_{1:n-1} \leq \dots \leq U_{n-1:n-1} \leq U_{n:n-1} := 1$$

be the order statistics at stage $n-1$. Moreover write

$$U_{i:n-1} = 1 + U_{i-n:n-1} \quad \text{for } i > n.$$

For $m < n$ the normalized uniform (overlapping) m -step spacings are defined by

$$D_{i,n}^{(m)} = n(U_{i+m-1:n-1} - U_{i-1:n-1}), \quad i = 1, 2, \dots, n,$$

and the empirical distribution function based on these m -step spacings will be written as

$$\hat{F}_n^{(m)}(t) = n^{-1} \sum_{i=1}^n 1_{[0,t]}(D_{i,n}^{(m)}), \quad 0 \leq t < \infty.$$

The common distribution function of the $D_{i,n}^{(m)}$ is given by

$$F_n^{(m)}(t) = P(nU_{m:n-1} \leq t) = \int_0^t f_n^{(m)}(u) du, \quad 0 \leq t \leq n,$$

with

$$f_n^{(m)}(u) = \frac{\Gamma(n)}{n^m \Gamma(m) \Gamma(n-m)} u^{m-1} (1-u/n)^{n-m-1}, \quad 0 \leq u \leq n.$$

As $n \rightarrow \infty$, $F_n^{(m)}$ converges to $F^{(m)}$, the distribution function of a Gamma $(m,1)$ random variable, i.e.

$$F^{(m)}(t) = \int_0^t (\Gamma(m))^{-1} u^{m-1} e^{-u} du, \quad 0 \leq t < \infty.$$

In Einmahl (1985) a Dvoretzky-Kiefer-Wolfowitz type exponential bound is derived for the Kolmogorov-Smirnov statistic of a certain class of dependent random variables. In this note that bound will be applied to give a short proof of a Glivenko-Cantelli theorem for $\hat{F}_n^{(m)} - F^{(m)}$, where m is allowed to tend to infinity with n at a certain rate. The thus obtained theorem is an improvement of Theorem 1.1 in Beirlant and van Zuijlen (1985) w.r.t. the rate at which m may tend to infinity, cf. also Beirlant (1984, page 72).

THE THEOREM

Theorem. Let $\alpha \geq 0$ and $0 \leq \beta < \frac{1}{2}$. If $m = o((n^{1-2\beta}/\log n)^{1/(1+2\alpha)})$, as $n \rightarrow \infty$, then

$$(1) \quad \limsup_{n \rightarrow \infty} \sup_{0 \leq t < \infty} m^{\alpha\beta} |\hat{F}_n^{(m)}(t) - F^{(m)}(t)| = 0 \quad \text{a.s.}$$

Corollary. If $m = o(n/\log n)$ as $n \rightarrow \infty$, then

$$\limsup_{n \rightarrow \infty} \sup_{0 \leq t < \infty} |\hat{F}_n^{(m)}(t) - F^{(m)}(t)| = 0 \quad \text{a.s.}$$

For the proof of the Theorem we shall need the following

Inequality. There exists a universal constant $c \in (0, \infty)$ such that for all $n \in \mathbb{N}$ and $\lambda \geq 0$

$$(2) \quad P(\sup_{0 \leq t < \infty} |\hat{F}_n^{(m)}(t) - F^{(m)}(t)| \geq \lambda) \leq C m \exp(-\frac{n\lambda^2}{2m}).$$

Proof. It is well known (see e.g. Beirlant and van Zuijlen (1985, page 303)) that $\hat{F}_n^{(m)} - F^{(m)}$ can be written as a sum of empirical processes based on *non-overlapping* m -step spacings and a remainder term. More precisely, we have

$$(3) \quad \hat{F}_n^{(m)}(t) - F^{(m)}(t) = \frac{\nu}{n} \sum_{j=1}^m ({}^{(j)}\hat{F}_\nu(t) - F^{(m)}(t)) + \frac{1}{n} \sum_{i=m\nu+1}^n (1_{[0,t]}(D_{i,n}^{(m)}) - F^{(m)}(t)),$$

where $\nu = [n/m]$ and

$${}^{(j)}\hat{F}_\nu(t) = \nu^{-1} \sum_{i=1}^{\nu} 1_{[0,t]}(D_{j+(i-1)m,n}^{(m)}), \quad j = 1, 2, \dots, m.$$

From the Theorem in Einmahl (1985) and its proof we immediately see that there exists a universal constant C such that for all $n \in \mathbb{N}$ and $\lambda \geq 0$

$$(4) \quad P(\sup_{0 \leq t < \infty} |\nu^{\frac{1}{2}} ({}^{(1)}\hat{F}_\nu(t) - F^{(m)}(t))| \geq \lambda) \leq C \exp(-2\lambda^2)$$

and

$$(5) \quad P(\sup_{0 \leq t < \infty} |1_{[0,t]}(D_{n,n}^{(m)}) - F^{(m)}(t)| \geq \lambda) \leq C \exp(-2\lambda^2).$$

Combination of (3), (4) and (5) yields

$$\begin{aligned}
 & P(\sup_{0 \leq t < \infty} |\tilde{F}_n^{(m)}(t) - F^{(m)}(t)| \geq \lambda) \\
 & \leq P(\sup_{0 \leq t < \infty} |\frac{\nu}{n} \sum_{j=1}^m (\tilde{F}_{\nu}^{(j)}(t) - F^{(m)}(t)) + \frac{1}{n} \sum_{i=m\nu+1}^n (1_{[0,t]}(D_{i,n}^{(m)}) - F^{(m)}(t))| \geq \lambda) \\
 & \leq mP(\sup_{0 \leq t < \infty} |\frac{\nu}{n} (\tilde{F}_{\nu}^{(1)}(t) - F^{(m)}(t))| \geq \frac{\lambda}{2m}) \\
 & \quad + mP(\sup_{0 \leq t < \infty} |\frac{1}{n} (1_{[0,t]}(D_{n,n}^{(m)}) - F^{(m)}(t))| \geq \frac{\lambda}{2m}) \\
 & \leq Cm \exp\left(\frac{-2\lambda^2 n^2}{4m^2 \nu}\right) + Cm \exp\left(\frac{-2\lambda^2 n^2}{4m^2}\right) \leq 2Cm \exp\left(\frac{-n\lambda^2}{2m}\right).
 \end{aligned}$$

Relabeling $2C$ by C completes the proof of the Inequality. □

Proof of the Theorem. For a proof of (1) it suffices to show that

$\sum_{n=1}^{\infty} PA_n(\epsilon) < \infty$, for every $\epsilon > 0$, where

$$A_n(\epsilon) = \left\{ \sup_{0 \leq t < \infty} |\tilde{F}_n^{(m)}(t) - F^{(m)}(t)| \geq \epsilon m^{-\alpha} n^{-\beta} \right\}.$$

Applying (2) gives

$$(6) \quad PA_n(\epsilon) \leq Cm \exp\left(-\frac{1}{2} \epsilon^2 n^{1-2\beta} m^{-1-2\alpha}\right).$$

Using the condition on m we see that for any $K > 0$, the right side of (6) is bounded from above by

$$Cn \exp(-\frac{1}{2} \epsilon^2 K \log n),$$

for large n . For sufficiently large K this last expression is summable in n . □

REFERENCES

- [1] BEIRLANT, J. (1984). Eigenschappen van de empirische verdelingsfunctie en het empirisch proces van toevalsafstanden en asymptotische theorie van hierop gebaseerde statistische functionalen. Ph.D. Dissertation, Catholic University, Leuven.

- [2] BEIRLANT, J. and VAN ZUIJLEN, M.C.A. (1985). The empirical distribution function and strong laws for functions of order statistics of uniform spacings. *J. Multivariate Anal.* 16, 300-317.
- [3] EINMAHL, J.H.J. (1985). On the Kolmogorov-Smirnov statistic of certain dependent random variables. *Publ. Inst. Statist. Univ. Paris.* 30, 53-57.

Reçu en Juin 1987

JOHN H.J. EINMAHL
Dept. of Med. Informatics and Statistics
University of Limburg
P.O. Box 616
6200 MD Maastricht
The Netherlands