Estimating a changed segment in a sample*

Alfredas Račkauskas
Vilnius University and Institute of Mathematics and Informatics
Department of Mathematics, Vilnius University
Naugarduko 24, Lt-2006 Vilnius, Lithuania
and
Charles Suquet
Laboratoire P. Painlevé (UMR 8524 CNRS)
Bât. M2 U.F.R. de Mathématiques
Université des Sciences et Technologies de Lille
F-59655 Villeneuve d'Ascq Cedex France

Abstract

In the paper we consider a changed segment model for sample distributions. We generalize Dümbgen's [6] change point estimator and obtain optimal rates of convergence of estimators of the begining and the length of the changed segment.

Keywords : changed segment, changed segment location, empirical process, epidemic model, probability metrics, reproducing kernel.

Mathematics Subject Classifications (2000): 62G05

1 Introduction

A general changed segment (called also epidemic) model can be described as follows. For $n = 3, 4, \dots$, let P_n and Q_n be two probability distributions on a measurable space E and let $X_{n,1}, X_{n,2}, \dots, X_{n,n}$ be a triangular array of independent random elements in E. There exist s_n^* and t_n^* such that for $1 \le i \le ns_n^*$ or $nt_n^* < i \le n$, the $X_{i,n}$'s have distribution P_n, while for $ns_n^* < i \le nt_n^*$, they have distribution Q_n. We refer to e.g., Avery and Handerson [1], Commenges et al. [3], Račkauskas and Suquet [12], [13],

*This work was supported by a cooperation agreement Lille-Vilnius EGIDE Gillibert.

Yao [18] for a comprehensive review. Our aim to this model is to estimate the pair (s_n^*, h_n^*), where $h_n^* = t_n^* - s_n^*$ measures the length of the changed segment.

This is the most simple type of the multiple change models which have attracted big attention and is well studied in the literature, e.g. Yao [17] (estimates a number of jumps in the mean), Schechtman and Wolfe [14] (propose sequential algorithm for estimating the number and the location of change points), Lavielle and Moulines [9] (estimates unknown number of shifts in time series) Lavielle [8] (derives asymptotic results for location and the number of changed segments), Lee [11] (gives asymptotic results for the location of segments, following Dümbgen [6])) to name a few. We also refer to Brodsky and Darkhovsky [2] and Csörgő and Horváth [4] for state-of-the art of change point problems.

Concerning asymptotic results in multiple change models it is usually assumed that the length of each changed segment tends to infinity at the same rate as the total number of observations, see, e.g., above mentioned paper by Lavielle [8] and references therein. In this paper we consider the changed segment tending to infinity possibly at much slower rate. For example, our results apply for segment growth at the rates of type $\log n \log\log n$. We also discuss very general examples of possible estimators.

2 Estimator and its consistency

The parameter we are estimating is $\theta_n = (s_n^*, h_n^*)$ where $h_n^* = t_n^* - s_n^*$ measures the length of the changed segment in the model described above. This unknown parameter θ_n belongs to the set

$$\Theta_n := \big\{(s,h) \in T_n^2 : \ s + h < 1\big\}$$

where

$$T_n := \{1/n, 2/n, \ldots, (n-1)/n\}.$$

For notational convenience we extend the sample $X_{n,1}, \ldots, X_{n,n}$ periodically by putting $X_{n,j} = X_{n,j-n}$ for $n < j \le 2n$. Now for $0 \le k \le m \le 2n$ introduce the empirical measure

$$P_n^{k,m} := \frac{1}{m-k} \sum_{j=k+1}^{m} \delta_{X_{n,j}}.$$

Clearly, $P_n^{0,n}$ is the empirical measure based on the sample $X_{n,1}, \ldots, X_{n,n}$.

For $s, h \in T_n$ denote

$$I_{s,h} = \{ns+1, \dots, ns+nh\}, \quad \text{and set} \quad I_* = I_{s_n^*, h_n^*}.$$

We denote by $|A|$ the number of elements of any finite set A. Introducing $I_{s,h}^c := \{1, \dots, n\} \setminus I_{s,h}$, we note that

$$|I_{s,h}| = nh, \quad |I_{s,h}^c| = n(1-h).$$

With the weight function

$$w(u) := u^{1/2}(1-u)^{1/2}, \quad u \in [0,1],$$

introduce the signed measure

$$D_n^{s,h} := w(h)\big(P_n^{ns,n(s+h)} - P_n^{n(s+h),n+ns}\big), \quad s, h \in T_n. \tag{1}$$

First we check that this measure can be represented as

$$D_n^{s,h} = \frac{1}{nw(h)} \sum_{i \in I_{s,h}} \big(\delta_{X_{n,i}} - P_n^{0,n}\big), \quad s, h \in T_n. \tag{2}$$

This results from the following elementary computation.

$$\begin{aligned}
D_n^{s,h} =& w(h)\left(\frac{1}{nh}\sum_{i \in I_{s,h}} \delta_{X_{n,i}} - \frac{1}{n(1-h)} \sum_{i \in I_{s,h}^c} \delta_{X_{n,i}}\right) \\
=& w(h)\left(\frac{1}{nh}\sum_{i \in I_{s,h}} \delta_{X_{n,i}} - \frac{1}{n(1-h)} \sum_{i=1}^n \delta_{X_{n,i}} + \frac{1}{n(1-h)} \sum_{i \in I_{s,h}} \delta_{X_{n,i}}\right) \\
=& \frac{w(h)}{n}\Big(\frac{1}{h} + \frac{1}{1-h}\Big) \sum_{i \in I_{s,h}} \delta_{X_{n,i}} - \frac{nhw(h)}{nh(1-h)} \frac{1}{n} \sum_{i=1}^n \delta_{X_{n,i}} \\
=& \frac{1}{nw(h)} \sum_{i \in I_{s,h}} \delta_{X_{n,i}} - \frac{1}{nw(h)}\big(nhP_n^{0,n}\big) \\
=& \frac{1}{nw(h)} \sum_{i \in I_{s,h}} \big(\delta_{X_{n,j}} - P_n^{0,n}\big), \quad \text{because } |I_{s,h}| = nh.
\end{aligned}$$

Next we check that the mean of $D_n^{s,h}$ is

$$\Delta_n^{s,h} := \mathbf{E}\, D_n^{s,h} = r_n(s,h)(Q_n - P_n), \tag{3}$$

where $r_n(s,h)$ admits the both representations

$$r_n(s,h) = \frac{1}{nw(h)}\Big(|I_{s,h} \cap I_*| - nhh_n^*\Big), \tag{4}$$

$$r_n(s,h) = \frac{1}{nw(h)}\Big(|I_{s,h}^c \cap I_*^c| - n(1-h)(1-h_n^*)\Big). \tag{5}$$

Indeed

$$\begin{aligned}
\mathbf{E}\, D_n^{s,h} &= \frac{1}{nw(h)}\Big(\sum_{i \in I_{s,h} \cap I_*} Q_n + \sum_{i \in I_{s,h} \cap I_*^c} P_n - hnh_n^* Q_n - hn(1-h_n^*)P_n\Big) \\
&= \frac{1}{nw(h)}\Big(|I_{s,h} \cap I_*| Q_n + |I_{s,h} \cap I_*^c| P_n - hnh_n^* Q_n - hn(1-h_n^*)P_n\Big) \\
&= \frac{1}{nw(h)}\Big(|I_{s,h} \cap I_*|(Q_n - P_n) + |I_{s,h}| P_n - hnh_n^* Q_n - hn(1-h_n^*)P_n\Big) \\
&= \frac{1}{nw(h)}\Big(|I_{s,h} \cap I_*| - nhh_n^*\Big)(Q_n - P_n).
\end{aligned}$$

It is worth noticing here that

$$r_n^* := r_n(s_n^*, h_n^*) = \sqrt{h_n^*(1-h_n^*)}. \tag{6}$$

Now we choose a seminorm N_n on the space $\mathcal{M}$ of all finite signed measures on E and note that $N_n(\Delta_n^{s,h}) = |r_n(s,h)| N_n(Q_n - P_n)$. As $|r_n|$ has a unique maximum on Θ_n reached at $(s,h) = (s_n^*, h_n^*)$, see Lemma 5 in Section 4, this leads to the estimator

$$(\widehat{s}_n^*, \widehat{h}_n^*) := \arg\max \big\{ N_n(D_n^{s,h}) :\ (s,h) \in T_n^2 \big\}, \tag{7}$$

which is a generalization of Dümbgen's estimator [6] in the setting of changed segment model. To incorporate the case where the length of changed segment h_n^* is close to either zero or to one, we shall assume that with a sequence (τ_n) of positive numbers which can tend to zero as n increases,

$$\frac{h_n^*(1-h_n^*)}{\tau_n} \xrightarrow[n\to\infty]{} a \in (0,1). \tag{8}$$

With this setup we define the estimator

$$(\widehat{s}_n^*, \widehat{h}_n^*) := \arg\max \big\{ N_n(D_n^{s,h}) :\ (s,h) \in T_n^2,\ h(1-h) \ge \tau_n \ \big\}. \tag{9}$$

Next we need some assumptions on the seminorms N_n which as in Dümbgen [6] are allowed to be random. Let the underlying probability space be $(\Omega, \mathcal{S}, \Pr)$. By $\Pr^*$ we denote then the outer probability. First we define an admissible class of measurable functions.

Definition 1. A seminorm $||\cdot||$ on $\mathcal{M}$ will be called admissible, if

> there are two constants $c_1, c_2 > 0$ such that for all $m \in \mathbb{N}$ and for arbitrary independent identically distributed random elements $X_1, \dots, X_m$ in E:
>
> $$\Pr{}^*\Big(\Big\|m^{-1/2}\sum_{i=1}^{m}\big(\delta_{X_i} - \mathbf{E}\,\delta_{X_i}\big)\Big\| > \lambda\Big) \le c_1 \exp(-c_2\lambda^2)$$
>
> for all $\lambda > 0$.

Now we introduce the following assumptions on N_n and on a distance between distributions P_n and Q_n.
Assumption (A): there is an admissible seminorm $||\cdot||$ on $\mathcal{M}$ such that

$$N_n(\nu) \le ||\nu|| \quad \text{for each} \quad \nu \in \mathcal{M}. \tag{10}$$

Assumption (B): there is a sequence of positive numbers (γ_n) possibly increasing such that

$$\Pr\Big(N_n(Q_n - P_n) \ge \frac{1}{\gamma_n}\Big) \xrightarrow[n\to\infty]{} 1. \tag{11}$$

In the case where N_n is non random, (11) reduces to $N_n(Q_n - P_n) \ge \gamma_n^{-1}$ for n large enough (one can put $\gamma_n^{-1} = N_n(Q_n - P_n)$). If moreover $P_n = P$ and $Q_n = Q$ do not depend on n, this in turn, reduces to $N_n(Q - P) \ge c_0$ with a constant $c_0 > 0$.

Theorem 2. *Assume that the seminorms N_n and the distributions P_n and Q_n satisfy the assumptions (A) and (B). Under (8) suppose that*

$$\frac{\gamma_n \log^{1/2}(1/\tau_n)}{n^{1/2}\tau_n^{1/2}} \xrightarrow[n\to\infty]{} 0. \tag{12}$$

Then

$$|\widehat{s}_n - s_n^*| + |\widehat{h}_n - h_n^*| = O_{\Pr}(\gamma_n^2 n^{-1}). \tag{13}$$

If $\gamma_n = const.$ does not depend on n, then the convergence rate provided by Theorem 2 is $O_{\Pr}(n^{-1})$ which is optimal since the number of observations equals to n.

In the case where the seminorm N_n is nonrandom (12) reads

$$\frac{\log^{1/2}(1/\tau_n)}{n^{1/2}\tau_n^{1/2}N_n(P_n - Q_n)} \xrightarrow[n\to\infty]{} 0. \tag{14}$$

In terms of the length $\ell_n^* = nt_n^* - ns_n^*$ of changed segment, this condition becomes, let us say when $\ell_n^* \le n/2$,

$$\frac{\log n}{\ell_n^* N_n^2(P_n - Q_n)} \xrightarrow[n\to\infty]{} 0. \tag{15}$$

3 Proof of Theorem 2

To simplify notation set, recalling (1) and (3)

$$D_n^* = D_n^{s_n^*,h_n^*}, \quad \widehat{D}_n^* = D_n^{\widehat{s}_n^*,\widehat{h}_n^*}, \quad \Delta_n^* = \Delta_n^{s_n^*,h_n^*}, \quad \widehat{\Delta}_n^* = \Delta_n^{\widehat{s}_n^*,\widehat{h}_n^*} \tag{16}$$

The idea of the proof of theorem 2 is the following. We analyze the difference $N_n(D_n^{s,h}) - N_n(D_n^*)$ and show that for each pair (s,h) which is at a certain distance b_n far-off the point (s_n^*, h_n^*) this quantity is negative with probability approaching one as $b_n \to \infty$. Since the estimator $(\widehat{s}_n^*, \widehat{h}_n^*)$ is the point of maxima of $N_n(D_n^{s,h})$, the difference $N_n(\widehat{D}_n^*) - N_n(D_n^*)$ is always nonnegative. These arguments give the rate b_n for the convergence $(\widehat{s}_n^*, \widehat{h}^*) \to (s_n^*, h_n^*)$ in probability.

Set

$$d_n^{s,h} = D_n^{s,h} - \Delta_n^{s,h}, \quad d_n^* = D_n^* - \Delta_n^*. \tag{17}$$

The main tool for the proof is the following.

Claim 3. *For n large enough*

$$N_n(D_n^{s,h}) - N_n(D_n^*) \le \|d_n^{s,h} - d_n^*\| - \big(r_n^* - r_n(s,h)\big)\gamma_n^{-1}\big(1 - o_{\Pr}(1)\big), \tag{18}$$

where the "$o_{\Pr}(1)$" does not depend on (s,h).

Proof of Claim 3. To deduce this estimate first we check that

$$\big|N_n(D_n^*) - r_n^* N_n(Q_n - P_n)\big| \le \|d_n^*\| = O_{\Pr}(n^{-1/2}). \tag{19}$$

Indeed, recalling (3), we have $N_n(\Delta_n^*) = |r_n^*| N_n(Q_n - P_n)$. As N_n is a seminorm, we get by assumption (A)

$$\big|N_n(D_n^*) - N_n(\Delta_n^*)\big| \le N_n(D_n^* - \Delta_n^*) = N_n(d_n^*) \le \|d_n^*\|.$$

Then we can express $d_n^* = D_n^* - \Delta_n^*$ as follows.

$$\begin{aligned} d_n^* &= w(h_n^*)\Big(\frac{1}{nh_n^*}\sum_{i\in I_*}\delta_{X_{n,i}} - \frac{1}{n(1-h_n^*)}\sum_{i\in I_*^c}\delta_{X_{n,i}}\Big) - w(h_n^*)(Q_n - P_n) \\ &= w(h_n^*)\Big(\frac{1}{nh_n^*}\sum_{i\in I_*}\delta_{X_{n,i}} - Q_n - \frac{1}{n(1-h_n^*)}\sum_{i\in I_*^c}\delta_{X_{n,i}} + P_n\Big) \\ &= \frac{w(h_n^*)}{nh_n^*}\sum_{i\in I_*}\big(\delta_{X_{n,i}} - \mathbf{E}\,\delta_{X_{n,i}}\big) - \frac{w(h_n^*)}{n(1-h_n^*)}\sum_{i\in I_*^c}\big(\delta_{X_{n,i}} - \mathbf{E}\,\delta_{X_{n,i}}\big). \end{aligned}$$

By admissibility of the seminorm $||\cdot||$ we have for any $\lambda > 0$,

$$\Pr\left(\frac{w(h_n^*)}{nh_n^*}\Big\|\sum_{i\in I_*}(\delta_{X_{n,i}} - Q_n)\Big\| > \lambda n^{-1/2}\right) \le c_1 \exp\Big(\frac{-c_2\lambda^2}{1-h_n^*}\Big) \le c_1\exp(-c_2\lambda^2),$$

whence

$$\frac{w(h_n^*)}{nh_n^*}\Big\|\sum_{i\in I_*}(\delta_{X_{n,i}} - Q_n)\Big\| = O_{\Pr}(n^{-1/2}).$$

Similarly

$$\frac{w(h_n^*)}{n(1-h_n^*)}\Big\|\sum_{i\in I_*^c}(\delta_{X_{n,i}} - P_n)\Big\| = O_{\Pr}(n^{-1/2}).$$

Hence $\|d_n^*\| = O_{\Pr}(n^{-1/2})$ and (19) is established.

From (3) we have

$$\Delta_n^{s,h} = \frac{r_n(s,h)}{r_n^*}\Delta_n^*.$$

Hence

$$\begin{aligned} &N_n(D_n^{s,h}) - N_n(D_n^*) = N_n(d_n^{s,h} - d_n^* + d_n^* + \Delta_n^{s,h}) - N_n(D_n^*) = \\ &N_n\big(d_n^{s,h} - d_n^* + d_n^* + (r_n(s,h)/r_n^*)\Delta_n^*\big) - N_n(D_n^*) = \\ &N_n\big(d_n^{s,h} - d_n^* + (1 - r_n(s,h)/r_n^*)d_n^* + (r_n(s,h)/r_n^*)D_n^*\big) - N_n(D_n^*) \le \\ &N_n(d_n^{s,h} - d_n^*) - \Big(1 - \frac{r_n(s,h)}{r_n^*}\Big)\big(N_n(D_n^*) - N_n(d_n^*)\big). \end{aligned} \tag{20}$$

By the inequality in (19), condition (A) and noting that $1 - r_n(s,h)/r_n^* \ge 0$, cf. Lemma 5, we see that the right hand side in (20) is bounded by the quantity

$$N_n(d_n^{s,h} - d_n^*) - \Big(1 - \frac{r_n(s,h)}{r_n^*}\Big)\big(r_n^* N_n(P_n - Q_n) - 2||d_n^*||\big),$$

which on the set $\{N_n(P_n - Q_n) \geq \gamma_n^{-1}\}$ does not exceed

$$N_n(d_n^{s,h} - d_n^*) - \Big(1 - \frac{r_n(s,h)}{r_n^*}\Big)(r_n^*\gamma_n^{-1} - 2||d_n^*||) =$$
$$N_n(d_n^{s,h} - d_n^*) - \big(r_n^* - r_n(s,h)\big)\gamma_n^{-1}\big(1 - 2r_n^{*-1}\gamma_n||d_n^*||\big);$$

Finally, applying condition (A) to the signed measure $d_n^{s,h} - d_n^*$, recalling that $||d_n^*|| = O_{\mathrm{Pr}}(n^{-1/2})$ and noting that by (8), (12), $r_n^{*-1}\gamma_n n^{-1/2}$ tends to 0, we complete the proof of Claim 3. □

The rest of proof of Theorem 2 is divided in two steps. In the first one we check, that under the conditions stated in the theorem, for each $\varepsilon \in (0,1)$,

$$\mathrm{Pr}\left(\frac{|h_n^* - \widehat{h}_n^*|}{h_n^*(1-h_n^*)} > \varepsilon\right) \xrightarrow[n\to\infty]{} 0. \tag{21}$$

In the next step we prove, that for any $\varepsilon \in (0,1)$

$$\mathrm{Pr}\big(|\widehat{s}_n^* - s_n^*| + |\widehat{h}_n^* - h_n^*| > bn^{-1}\gamma_n^2,\ \Omega_n(\varepsilon)\big) \to 0, \tag{22}$$

as $n \to \infty$ and $b \to \infty$, where

$$\Omega_n(\varepsilon) := \left\{\frac{|h_n^* - \widehat{h}_n^*|}{h_n^*(1-h_n^*)} \leq \varepsilon\right\}.$$

Proof of (21). Introduce the sets

$$T^\tau := \big\{(s,h) \in \Theta_n :\ h(1-h) \geq \tau_n\big\}$$

and for $\varepsilon \in (0,1), b \geq 1$,

$$T_\varepsilon = \big\{(s,h) \in T^\tau :\ |h - h_n^*| > \varepsilon h_n^*(1-h_n^*)\big\},$$
$$T(b) = \big\{(s,h) \in T^\tau :\ \max\{|s - s_n^*|, |h - h_n^*|\} > b\gamma_n^2 n^{-1}\big\}.$$

Lemma 7 in Section 4 provides

$$\max_{(s,h)\in T_n^\tau} ||d_n^{s,h}|| = O_{\mathrm{Pr}}(n^{-1/2}\kappa_n) \quad \text{where } \kappa_n = |\log\tau_n|^{1/2}. \tag{23}$$

Reporting this estimate in (18), we obtain

$$N_n(D_n^{s,h}) - N_n(D_n^*) \leq O_{\mathrm{Pr}}(n^{-1/2}\kappa_n) - (r_n^* - r_n(s,h))\gamma_n^{-1}\big(1 - o_{\mathrm{Pr}}(1)\big), \tag{24}$$

uniformly in $(s,h) \in T_n^\tau$.

Let us check now that there is some constant c_ε such that

$$\forall (s,h) \in T_\varepsilon, \quad r_n^* - r_n(s,h) \geq c_\varepsilon \tau_n^{1/2}. \tag{25}$$

Consider first the case where $h_n^* \leq 1/2$. Lemma 5 provides the lower bound

$$r_n^* - r_n(s,h) \geq w(h_n^*) \frac{|h - h_n^*|}{2\max(h, h_n^*)}. \tag{26}$$

As $1 - h_n^* \geq 1/2$, the definition of T_ε gives $|h - h_n^*| > h_n^* \varepsilon/2$. Note also that on T_ε, $h \neq h_n^*$.

1. If $h > h_n^*$, the lower bound (26) writes $\frac{1}{2} w(h_n^*)(1 - h_n^*/h)$ and as $h_n^*/h < (1+\varepsilon/2)^{-1}$, we obtain

$$r_n^* - r_n(s,h) \geq \frac{\varepsilon}{4+2\varepsilon} w(h_n^*), \quad (h_n^* \leq 1/2,\ h > h_n^*). \tag{27}$$

2. If $h < h_n^*$, the lower bound (26) writes $\frac{1}{2} w(h_n^*)(h_n^* - h)/h_n^*$ whence

$$r_n^* - r_n(s,h) \geq \frac{\varepsilon}{4} w(h_n^*), \quad (h_n^* \leq 1/2,\ h < h_n^*). \tag{28}$$

In the case where $h_n^* > 1/2$, we keep from Lemma 5 the lower bound

$$r_n^* - r_n(s,h) \geq w(h_n^*) \frac{|h - h_n^*|}{2\max(1-h, 1-h_n^*)} = w(h_n^*) \frac{|(1-h) - (1-h_n^*)|}{2\max(1-h, 1-h_n^*)}$$

and from the definition of T_ε we get $|(1-h) - (1-h_n^*|) > (1-h_n^*)\varepsilon/2$. This leads clearly to the lower bounds (27) when $h < h_n^*$ and (28) when $h > h_n^*$. In view of (8), these lower bounds give (25).

Now,(24), (25) and (12) yield

$$N_n(D_n^{s,h}) - N_n(D_n^*) \leq -c_\varepsilon \tau_n^{1/2} \gamma_n^{-1}(1 - o_{\mathrm{Pr}}(1)), \quad \text{uniformly in } (s,h) \in T_\varepsilon.$$

In other words for each $\varepsilon > 0$ there is a constant $c_\varepsilon > 0$ such that

$$\Pr\bigl(N_n(D_n^{s,h}) - N_n(D_n^*) \leq -c_\varepsilon \tau_n^{1/2} \gamma_n^{-1} \text{ for all } (s,h) \in T_\varepsilon\bigr) \xrightarrow[n\to\infty]{} 1. \tag{29}$$

As $(\widehat{s}_n^*, \widehat{h}_n^*) := \arg\max N_n(D_n^{s,h})$, the difference $N_n(\widehat{D}_n^*) - N_n(D_n^*)$ is always nonnegative. Hence (29) implies that $\Pr\bigl((\widehat{s}_n^*, \widehat{h}_n^*) \in T_\varepsilon\bigr)$ tends to 0. Since by (9), $\widehat{h}_n^*(1 - \widehat{h}_n^*) \geq \tau_n$, this leads to (21). □

Proof of (22). We investigate in details the difference $d_n^{s,h} - d_n^*$ for (s,h) in the set

$$T_0(b,\varepsilon) := (T^\tau \setminus T_\varepsilon) \cap T(b) = T_\varepsilon^c \cap T(b). \tag{30}$$

Putting for notational simplicity

$$\eta_j := \delta_{X_{n,j}} - \mathbf{E}\, \delta_{X_{n,j}}, \quad j = 1, \ldots, n,$$

we can write

$$d_n^{s,h} - d_n^* = I_1(s,h) + I_2(s,h) + I_3(s,h), \quad (s,h) \in T,$$

where

$$I_1(s,h) := \frac{1}{nw(h)}\Big(\sum_{j\in I_{s,h}} \eta_j - \sum_{j\in I_*} \eta_j\Big),$$

$$I_2(s,h) := \Big(\frac{1}{nw(h)} - \frac{1}{nw(h_n^*)}\Big)\sum_{j\in I_*} \eta_j,$$

$$I_3(s,h) := \Big(\frac{h}{nw(h)} - \frac{h_n^*}{nw(h_n^*)}\Big)\sum_{j=1}^n \eta_j.$$

By admissibility of the seminorm $||\cdot||$ we have

$$||I_2(s,h)|| = \Big\|\Big(1 - \frac{w(h_n^*)}{w(h)}\Big)\frac{1}{nw(h_n^*)}\sum_{j\in I_*} \eta_j\Big\| = \Big|1 - \frac{w(h_n^*)}{w(h)}\Big| O_{\mathrm{Pr}}(n^{-1/2})$$

and

$$||I_3(s,h)|| = \Big|h - \frac{h_n^* w(h)}{w(h_n^*)}\Big| O_{\mathrm{Pr}}(n^{-1/2}\tau_n^{-1/2}).$$

It is worth observing here that if $(s,h) \in T_\varepsilon^c$, then

$$\Big|\frac{h}{h_n^*} - 1\Big| \le (1 - h_n^*)\varepsilon, \qquad \Big|\frac{1-h}{1-h_n^*} - 1\Big| \le h_n^*\varepsilon,$$

whence

$$(1-\varepsilon)h_n^* \le h \le (1+\varepsilon)h_n^* \tag{31}$$

$$(1-\varepsilon)(1-h_n^*) \le 1 - h \le (1+\varepsilon)(1-h_n^*) \tag{32}$$

$$(1-\varepsilon)w(h_n^*) \le w(h) \le (1+\varepsilon)w(h_n^*). \tag{33}$$

Note also the elementary inequality

$$\forall t \in [0,1], \quad \min(t, 1-t) \le 2t(1-t) = 2w(t)^2. \tag{34}$$

Lemma 5 provides the lower bounds

$$r_n^* - r_n(s,h) \geq \begin{cases} \dfrac{|h-h_n^*|}{2\min(h_n^*, 1-h)} & \text{if } h < h_n^*, \\ \dfrac{|h-h_n^*|}{2\min(h, 1-h_n^*)} & \text{if } h \geq h_n^*. \end{cases}$$

If $h < h_n^*$, then by (32), (34) and as $0 < w(h_n^*)^2 < w(h_n^*) < 1$, we get

$$r_n^* - r_n(s,h) \geq \frac{|h-h_n^*|}{2(1+\varepsilon)\min(h_n^*, 1-h_n^*)} \geq \frac{|h-h_n^*|}{4(1+\varepsilon)w(h_n^*)^2} > \frac{|h-h_n^*|}{4(1+\varepsilon)w(h_n^*)}$$

We clearly obtain the same lower bound in the case $h \geq h_n^*$, so let us retain that with $c_\varepsilon = 1/(4+4\varepsilon)$,

$$\forall (s,h) \in \Theta_n \setminus T_\varepsilon, \qquad r_n^* - r_n(s,h) \geq c_\varepsilon \frac{|h-h^*|}{w(h_n^*)}. \tag{35}$$

Next we note here that

$$w(h) - w(h_n^*) = \sqrt{h(1-h)} - \sqrt{h_n^*(1-h_n^*)} = \frac{h(1-h) - h_n^*(1-h_n^*)}{\sqrt{h(1-h)} + \sqrt{h_n^*(1-h_n^*)}},$$

whence

$$|w(h) - w(h_n^*)| \leq \frac{|h-h_n^*||1-(h+h_n^*)|}{\sqrt{h_n^*(1-h_n^*)}} \leq \frac{|h-h_n^*|}{w(h_n^*)}. \tag{36}$$

Using (36) and the rough estimate $w(h) < 1 < w(h_n^*)^{-1}$ in the inequality

$$|hw(h) - h_n^* w(h_n^*)| \leq |h-h_n^*| w(h) + h_n^* |w(h) - w(h_n^*)|,$$

we obtain also

$$|hw(h) - h_n^* w(h_n^*)| \leq 2\frac{|h-h_n^*|}{w(h_n^*)}. \tag{37}$$

Recalling that for $(s,h) \in T_\varepsilon^c = T^\tau \setminus T_\varepsilon$, $w(h) \geq \tau_n^{1/2}$, we deduce from (36) and (37) that

$$\left|1 - \frac{w(h_n^*)}{w(h)}\right| \leq \tau_n^{-1/2} \frac{|h-h^*|}{w(h_n^*)}, \qquad \left|h - \frac{h_n^* w(h_n^*)}{w(h)}\right| \leq 2\tau_n^{-1/2} \frac{|h-h^*|}{w(h_n^*)}. \tag{38}$$

Hence combining the inequalities (38) with (35) we see that for $i = 2, 3$,

$$||I_i(s,h)|| = \tau_n^{-1/2}(r_n^* - r_n(s,h)) n^{-1/2} O_{\Pr}(1) = \gamma_n^{-1}(r_n^* - r_n(s,h)) o_{\Pr}(1)$$

and from (18) we deduce

$$N_n(D_n^{s,h}) - N_n(D_n^*) \leq ||I_1(s,h)|| - c_\varepsilon (r_n^* - r_n(s,h)) \gamma_n^{-1}(1 - o_{\Pr}(1)). \tag{39}$$

Next we use the following

Claim 4. *For each constant $c > 0$ and each $\varepsilon \in (0,1)$*

$$\limsup_{n\to\infty} \Pr\Big(\max_{(s,h)\in T_0(b,\varepsilon)} \frac{||I_1(s,h)||}{(r_n^* - r_n(s,h))} \geq c\gamma_n^{-1} \Big) \xrightarrow[b\to\infty]{} 0. \tag{40}$$

It follows from (40) that for any $\varepsilon \in (0,1)$, the quantity

$$\liminf_{n\to\infty} \Pr\big(N_n(D_n^{s,h}) - N_n(D_n^*) \leq -\frac{c_\varepsilon}{2}(r_n^* - r_n(s,h))\gamma_n^{-1},\ \forall (s,h) \in T_0(b,\varepsilon) \big)$$

tends to 1 as b tends to infinity. Since $N_n(D_n^{\widehat{s_n^*},\widehat{h_n^*}}) - N_n(D_n^*)$ is always non negative and recalling the definition of $T_0(b,\varepsilon)$, see (30), this gives (22) and the proof of the theorem is now reduced to that of the Claim 4. □

Proof of Claim 4. Decompose the set $T_0(b,\varepsilon)$ defined in (30) as

$$T_0(b,\varepsilon) = \bigcup_{j=1}^{5} T_{0j}(b,\varepsilon),$$

where

$$\begin{aligned}
T_{01}(b,\varepsilon) &= \{(s,h) \in T_0(b,\varepsilon) : I_{s,h} \cap I_* = \emptyset\},\\
T_{02}(b,\varepsilon) &= \{(s,h) \in T_0(b,\varepsilon) : I_{s,h} \supset I_*\},\\
T_{03}(b,\varepsilon) &= \{(s,h) \in T_0(b,\varepsilon) : I_{s,h} \subset I_*\},\\
T_{04}(b,\varepsilon) &= \{(s,h) \in T_0(b,\varepsilon) : s < s_n^* < s+h < s_n^* + h_n^*\},\\
T_{05}(b,\varepsilon) &= \{(s,h) \in T_0(b,\varepsilon) : s_n^* < s < s_n^* + h_n^* < s+h\}.
\end{aligned}$$

Now (40) follows if we show for any $c > 0$, $\varepsilon \in (0,1)$ and each $j = 1,\ldots,5$

$$\limsup_{n} \Pr\Big(\max_{(s,h)\in T_{0j}(b,\varepsilon)} \frac{||I_1(s,h)||}{(r_n^* - r_n(s,h))} \geq c\gamma_n^{-1} \Big) \xrightarrow[b\to\infty]{} 0. \tag{41}$$

If $(s,h) \in T_{01}(b,\varepsilon)$, then $r_n(s,h) = -hh_n^*/w(h) < 0$, so by (8) we have $r_n^* - r_n(s,h) > w(h_n^*) > \frac{a^{1/2}}{2}\tau_n^{1/2}$ for n large enough and applying Lemma 7, more precisely the upper bound for the sum indexed by $I_{s,h}$ obtained in its proof, we get

$$\sup_{(s,h)\in T_{01}(b,\varepsilon)} \frac{||I_1(s,h)||}{r_n^* - r_n(s,h)} = O_{\Pr}\Big(\frac{|\ln \tau_n|^{1/2}}{\sqrt{n\tau_n}} \Big).$$

Hence, by condition (12)

$$\Pr\Big(\max_{(s,h)\in T_{01}(b,\varepsilon)} \frac{||I_1(s,h)||}{(r_n^* - r_n(s,h))} \geq c\gamma_n^{-1} \Big) \xrightarrow[n\to\infty]{} 0.$$

So (41) is valid for $j=1$.

If $(s,h)\in T_{02}(b,\varepsilon)$ then combining (35) with (33) we get with $c'_\varepsilon := \frac{4(1+\varepsilon)}{1-\varepsilon}$

$$(r_n^* - r_n(s,h))^{-1}||I_1(s,h)|| \le \frac{c'_\varepsilon}{n(h-h^*)}\Big\| \sum_{j=ns+1}^{ns_n^*} \eta_j + \sum_{j=ns_n^*+nh_n^*+1}^{ns+nh} \eta_j \Big\|.$$

In this case $s_n^* - s \le h - h_n^*$. Therefore, $h - h_n^* \ge b\gamma_n^2 n^{-1}$ and writing $c' = cc'_\varepsilon$, we obtain

$$\begin{aligned}
&\Pr\Big(\max_{(s,h)\in T_{02}(b,\varepsilon)} \frac{||I_1(s,h)||}{(r_n^* - r_n(s,h))} \ge c'\gamma_n^{-1}\Big) \le \\
&\Pr\Big(\max_{(s,h)\in T_{02}(b,\varepsilon)} \Big\|\frac{1}{nh - nh_n^*}\Big(\sum_{j=ns+1}^{ns_n^*} \eta_j + \sum_{j=ns_n^*+nh_n^*+1}^{ns+nh} \eta_j\Big)\Big\| \ge c'\gamma_n^{-1}\Big) \le \\
&\Pr\Big(\max_{m\ge b\gamma_n^2} \Big\|\frac{1}{m}\sum_{j=1}^m \xi_j\Big\| \ge c'\gamma_n^{-1}\Big),
\end{aligned}$$

where $(\xi_j, j\ge 1)$ are independent copies of any η_i, $i\notin I_*$.

Now we apply the same arguments as Dümbgen, see the end of the proof of Proposition 1 in [6]. Namely, since $(m^{-1}||\sum_{j=1}^m \xi_j||,\ \ m\ge 1)$ is a reverse submartingale, by Chow's inequality

$$\Pr\Big(\max_{m\ge b\gamma_n^2} \Big\|\frac{1}{m}\sum_{j=1}^m \xi_j\Big\| \ge c'\gamma_n^{-1}\Big) \le \frac{\gamma_n}{c'}\mathbf{E}\,(m_0^{-1}||\zeta_{m_0}||),$$

where $m_0 = \min\{m\in\mathbb{N} : m\ge b\gamma_n^2\}$ and $\zeta_m = \sum_{j=1}^m \xi_j$, $m\ge 1$. Since $||\cdot||$ is admissible, we have

$$\begin{aligned}
\mathbf{E}\,||\zeta_{m_0}|| &= \int_0^\infty \Pr(||\zeta_{m_0}|| \ge t)\,\mathrm{d}t \\
&\le \int_0^\infty c_1 \exp\{-c_2 t^2 m_0^{-1}\}\,\mathrm{d}t = c_3 m_0^{1/2}.
\end{aligned}$$

Hence

$$\Pr\Big(\max_{m\ge b\gamma_n^2} \Big\|\frac{1}{m}\sum_{j=1}^m \xi_j\Big\| \ge c'\gamma_n^{-1}\Big) \le \frac{c_3\gamma_n}{c' m_0^{1/2}} \le \frac{c_3\gamma_n}{c'(b\gamma_n^2)^{1/2}} = \frac{c_3}{c' b^{1/2}}. \tag{42}$$

This completes the proof of (41) with $j=2$. For $j=3$ the proof of (41) is similar therefore we omit it.

Let $(s,h) \in T_{04}(b,\varepsilon)$. In this case $s_n^* - s \geq h - h_n^*$, so $s_n^* - s \geq b\gamma_n^2 n^{-1}$ and we have

$$||I_1(s,h)|| = \frac{1}{nw(h)}\Big\|\Big(\sum_{j=ns+1}^{ns_n^*} \eta_j + \sum_{j=ns+nh+1}^{ns_n^*+nh_n^*} \eta_j'\Big)\Big\|.$$

The number of summands in $I_1(s,h)$ equals to $n[2(s_n^* - s) + h_n^* - h]$ and is not less than $b\gamma_n^2$. For this configuration of changed segment we need to more carefully control the difference $r_n^* - r_n(s,h)$ in order to prove that

$$\forall n \geq 3,\ \forall (s,h) \in T_{04}(b,\varepsilon), \quad \frac{2(s_n^* - s) + h_n^* - h}{w(h)(r_n^* - r_n(s,h))} \leq K_\varepsilon, \tag{43}$$

for some constant K_ε.

If $h \geq h_n^*$ then we write

$$r_n^* - r_n(s,h) = r_n^* - \frac{s + h - s_n^* - hh_n^*}{w(h)} = \frac{s_n^* + h_n^* - s - h}{w(h)} + r_n^* - \frac{h_n^*(1-h)}{w(h)}.$$

Now if moreover $h \leq 1/2$, whence $1 - h_n^* \geq 1 - h \geq 1/2$, we get

$$\begin{aligned}
r_n^* - \frac{h_n^*(1-h)}{w(h)} &= \frac{\sqrt{hh_n^*(1-h)(1-h_n^*)} - h_n^*(1-h)}{w(h)} \\
&\geq (1-h)\sqrt{h_n^*}\frac{\sqrt{h} - \sqrt{h_n^*}}{w(h)} = (1-h)\frac{\sqrt{h_n^*}}{\sqrt{h} + \sqrt{h_n^*}}\frac{h - h_n^*}{w(h)} \\
&\geq \frac{1}{2(1+\varepsilon)^{1/2} + 2}\frac{h - h_n^*}{w(h)},
\end{aligned}$$

using (31). If $h > 1/2$, with still $h \geq h_n^*$, we have similarly

$$\begin{aligned}
r_n^* - \frac{h_n^*(1-h)}{w(h)} &\geq h_n^*\sqrt{1-h}\frac{\sqrt{1-h_n^*} - \sqrt{1-h}}{w(h)} \\
&= h_n^*\frac{\sqrt{(1-h)}}{\sqrt{1-h} + \sqrt{1-h_n^*}}\frac{h - h_n^*}{w(h)} \\
&\geq \frac{1}{2(1+\varepsilon)}\frac{1}{1 + (1-\varepsilon)^{-1/2}}\frac{h - h_n^*}{w(h)},
\end{aligned}$$

using (31), $h > 1/2$ and (32). So let us retain from the case $h \geq h_n^*$ that there is a constant $K_{1,\varepsilon} \in (0,1)$ such that

$$r_n^* - r_n(s,h) \geq \frac{s_n^* + h_n^* - s - h}{w(h)} + K_{1,\varepsilon}\frac{h - h_n^*}{w(h)}. \tag{44}$$

If $h < h_n^*$, writing

$$r_n^* - r_n(s,h) = r_n^* - \frac{s+h-s_n^*-hh_n^*}{w(h)} = \frac{s_n^*-s}{w(h)} + r_n^* - \frac{h(1-h_n^*)}{w(h)}$$

we obtain in the same way with a constant $K_{2,\varepsilon} \in (0,1)$,

$$r_n^* - r_n(s,h) \geq \frac{s_n^*-s}{w(h)} + K_{2,\varepsilon}\frac{h_n^*-h}{w(h)}. \tag{45}$$

Now we are in a position to check (43). First if $h \geq h_n^*$, then using (44) and recalling that here $h - h_n^* \leq s_n^* - s$ and $0 < K_{1,\varepsilon} < 1$, we get

$$\frac{2(s_n^*-s)+h_n^*-h}{w(h)(r_n^*-r_n(s,h))} \leq \frac{2(s_n^*-s)+h_n^*-h}{(s_n^*-s)+(K_{1,\varepsilon}-1)(h-h_n^*)} \leq \frac{2}{K_{1,\varepsilon}}.$$

If $h < h_n^*$, using (45), we get

$$\frac{2(s_n^*-s)+h_n^*-h}{w(h)(r_n^*-r_n(s,h))} \leq \frac{2(s_n^*-s)+h_n^*-h}{(s_n^*-s)+K_{2,\varepsilon}(h_n^*-h)} \leq 2 + \frac{1}{K_{2,\varepsilon}}.$$

Going back to the uniform control of $||I_1(s,h)||$ for $(s,h) \in T_{04}(b,\varepsilon)$, we deduce from (43) that with $C_\varepsilon = c/K_\varepsilon$,

$$\Pr\Big(\max_{(s,h)\in T_{04}(b,\varepsilon)} \frac{||I_1(s,h)||}{(r_n^*-r_n(s,h))} \geq c\gamma_n^{-1}\Big) \leq$$

$$\Pr\Big(\max_{(s,h)\in T_{04}(b,\varepsilon)} \Big\|\frac{1}{n(2(s_n^*-s)+h_n^*-h)}\Big(\sum_{j=ns+1}^{ns_n^*}\eta_j + \sum_{j=ns+nh+1}^{ns_n^*+nh_n^*}\eta_j'\Big)\Big\| \geq C_\varepsilon\gamma_n^{-1}\Big) \leq$$

$$\Pr\Big(\max_{m=m'+m''\geq b\gamma_n^2} \Big\|\frac{1}{m}\Big(\sum_{j=1}^{m'}\xi_j' + \sum_{j=1}^{m''}\xi_j\Big)\Big\| \geq C_\varepsilon\gamma_n^{-1}\Big).$$

We shall complete the treatment of $T_{04}(b,\varepsilon)$ together with $T_{05}(b,\varepsilon)$ since if $(s,h) \in T_{05}(b,\varepsilon)$ we have similarly

$$\Pr\Big(\max_{(s,h)\in T_{05}(b,\varepsilon)} \frac{||I_1(s,h)||}{(r_n^*-r_n(s,h))nw(h)} \geq c\gamma_n^{-1}\Big) \leq$$

$$\Pr\Big(\max_{m=m'+m''\geq b\gamma_n^2} \Big\|\frac{1}{m}\Big(\sum_{j=1}^{m'}\xi_j' + \sum_{j=1}^{m''}\xi_j\Big)\Big\| \geq c_\varepsilon\gamma_n^{-1}\Big).$$

The indexation set of the maximum inside the above probability is precisely

$$A := \big\{(m',m'') \in \{1,\dots,n\}^2 :\ m = m'+m'' \geq b\gamma_n^2\big\}.$$

In view of the inclusion $A \subset A_1 \cup A_2 \cup A_3$, where

$$A_1 := \{(m', m'') \in A : \ m' \le b\gamma_n^2/2, \ m'' \ge b\gamma_n^2/2\}$$
$$A_2 := \{(m', m'') \in A : \ m' \ge b\gamma_n^2/2, \ m'' \le b\gamma_n^2/2\}$$
$$A_3 := \{(m', m'') \in A : \ m' \ge b\gamma_n^2/2, \ m'' \ge b\gamma_n^2/2\}.$$

we have to estimate the probabilities

$$p_i := \Pr\Big(\max_{(m',m'')\in A_i} \Big\| \frac{1}{m}\Big(\sum_{j=1}^{m'} \xi'_j + \sum_{j=1}^{m''} \xi_j\Big)\Big\| \ge c\gamma_n^{-1}\Big), \quad i = 1, 2, 3.$$

We have $p_1 \le p'_1 + p''_1$ where

$$p'_1 := \Pr\Big(\max_{m' \le b\gamma_n^2/2} \frac{1}{b\gamma_n^2}\Big\| \sum_{j=1}^{m'} \xi'_j \Big\| \ge \frac{c}{2\gamma_n}\Big)$$
$$p''_1 := \Pr\Big(\max_{m'' \ge b\gamma_n^2/2} \frac{1}{m''}\Big\| \sum_{j=1}^{m''} \xi_j \Big\| \ge \frac{c}{2\gamma_n}\Big).$$

For the second probability, $\limsup_{n\to\infty} p''_1$ tends to zero when b tends to infinity by (42). As for the first probability, we have by Doob's inequality

$$p'_1 \le \frac{4}{b^2c^2\gamma_n^2}\mathbf{E} \max_{m' \le b\gamma_n^2/2} \Big\| \sum_{j=1}^{m'} \xi'_j \Big\|^2 \le \frac{16}{b^2c^2\gamma_n^2}\mathbf{E} \Big\| \sum_{j=1}^{b\gamma_n^2/2} \xi'_j \Big\|^2. \tag{46}$$

By admissibility of the seminorm $|| \cdot ||$ we have for any integer $m \ge 1$

$$\mathbf{E}\Big\| m^{-1/2} \sum_{i=1}^{m} \xi'_i \Big\|^2 \le \int_0^\infty 2\lambda c_1 \exp(-c_2\lambda^2)\, \mathrm{d}\lambda =: c_0 < \infty,$$

whence

$$\mathbf{E}\Big\| \sum_{j=1}^{b\gamma_n^2/2} \xi'_j \Big\|^2 \le \frac{c_0 b\gamma_n^2}{2},$$

which combined with (46) gives $p'_1 \le (8c_0c^{-2})b^{-1}$.

Clearly we can apply the same arguments to estimate p_2 and p_3. Hence the proof of the claim 4 is complete and this ends the proof of Theorem 2. □

4 Auxiliary results

We shall need the following two properties of the function r_n.

Lemma 5. *The function r_n defined by (4) satisfies*

(i) $|r_n|$ *has a unique maximum on* Θ_n*, reached at the point* $(s,h) = (s_n^*, h_n^*)$ *and* $|r_n(s_n^*, h_n^*)| = r_n^* := \sqrt{h_n^*(1-h_n^*)}$.

(ii) For every $(s,h) \in \Theta_n$,

$$1 - \frac{r_n(s,h)}{r_n^*} \geq \frac{|h_n^* - h|}{2\min\big(\max(h, h_n^*); \max(1-h, 1-h_n^*)\big)}. \tag{47}$$

Proof of (i). We separate the cases $s = s_n^*$ and $s \neq s_n^*$. In the first case we note that

$$r_n(s_n^*, h) = \frac{1}{nw(h)}\big(\min(nh, nh_n^*) - nhh_n^*\big) = \frac{1}{w(h)}\big(\min(h, h_n^*) - hh_n^*\big),$$

whence more explicitly

$$r_n(s_n^*, h) = \begin{cases} \frac{h(1-h_n^*)}{\sqrt{h(1-h)}} = (1-h_n^*)\sqrt{\frac{h}{1-h}} & \text{if } h \leq h_n^*, \\ \frac{h_n^*(1-h)}{\sqrt{h(1-h)}} = h_n^*\sqrt{\frac{1-h}{h}} & \text{if } h \geq h_n^*. \end{cases}$$

Observing that $h \mapsto h^{1/2}(1-h)^{-1/2}$ increases on $[0,1)$ and $h \mapsto (1-h)^{1/2}h^{-1/2}$ decreases on $(0,1]$, it is now clear that the non negative partial map $h \mapsto r_n(s_n^*, h)$ reaches its maximum at the point h_n^* and $r_n(s_n^*, h_n^*) = r_n^* := \sqrt{h_n^*(1-h_n^*)}$.

Next, considering different configurations of $(s,h) \in \Theta_n$ with $s \neq s_n^*$, we will check that

$$|r_n(s,h)| < r_n^* \quad \text{if} \quad (s,h) \neq (s_n^*, h_n^*). \tag{48}$$

1. If $I_{s,h} \cap I_* = \emptyset$ then $h < 1 - h_n^*$ and by increasingness on $[0,1)$ of $h \mapsto h/w(h) = h^{1/2}(1-h)^{-1/2}$ and from (4) we obtain

$$|r_n(s,h)| = \frac{h}{w(h)}h_n^* < \frac{1-h_n^*}{w(1-h_n^*)}h_n^* = r_n^*.$$

2. If $I_{s,h} \subset I_*$ and $s \neq s_n^*$, then necessarily $h < h_n^*$. In this configuration,

$$|r_n(s,h)| = \frac{1}{nw(h)}|nh - nhh_n^*| = \frac{h(1-h_n^*)}{w(h)} = (1-h_n^*)\sqrt{\frac{h}{1-h}}.$$

From increasingness of $h \mapsto h^{1/2}(1-h)^{-1/2}$ on $[0,1)$ and the fact that in this configuration, $h < h_n^*$, we obtain

$$|r_n(s,h)| < (1-h_n^*)\sqrt{\frac{h_n^*}{1-h_n^*}} = r_n^*.$$

3. If $I_* \subset I_{s,h}$ and $s \neq s_n^*$, then necessarily $h_n^* < h$. In this configuration,

$$|r_n(s,h)| = \frac{1}{nw(h)}|nh_n^* - nhh_n^*| = h_n^* \frac{1-h}{w(h)} = h_n^* \sqrt{\frac{1-h}{h}}.$$

From decreasingness on $(0,1]$ of $h \mapsto (1-h)^{1/2}h^{-1/2}$ and the fact that in this configuration, $h_n^* < h$, we obtain

$$|r_n(s,h)| < h_n^* \sqrt{\frac{1-h_n^*}{h_n^*}} = r_n^*.$$

4. If $s < s_n^* < s+h \leq s_n^* + h_n^*$, then

$$|r_n(s,h)| = \frac{1}{w(h)}|s+h-s_n^* - hh_n^*|.$$

4.1. If $s+h-s_n^*-hh_n^* \geq 0$, then $|r_n(s,h)| = w(h)^{-1}(s+h-s_n^*-hh_n^*)$. Note that $I_{s_n^*, s+h-s_n^*} \subsetneq I_{s,h}$, whence $s+h-s_n^* < h$. If moreover $s+h < s_n^* + h_n^*$, then $s+h-s_n^* < \min(h, h_n^*)$ and

$$|r_n(s,h)| < \frac{1}{w(h)}(\min(h,h_n^*) - hh_n^*) = r_n(s_n^*, h) \leq r_n^*,$$

so (48) holds. If $s+h = s_n^* + h_n^*$, then as $s < s_n^*$, necessarily $h > h_n^*$, so $r_n(s_n^*, h) < r_n(s_n^*, h_n^*)$ and

$$|r_n(s,h)| \leq \frac{1}{w(h)}(\min(h,h_n^*) - hh_n^*) = r_n(s_n^*, h) < r_n^*,$$

so (48) still holds.

4.2. If $s+h-s_n^*-hh_n^* < 0$, then $|r_n(s,h)| = w(h)^{-1}(s_n^*-s-h+hh_n^*)$. Recall that $s \geq 1/n > 0$ and $s_n^* + h_n^* \leq 1 - 1/n < 1$.

a) If $h + h_n^* \geq 1$, using the fact that $s_n^* - s < 1 - h_n^*$, we get

$$|r_n(s,h)| < \frac{1}{w(h)}\big(1-h_n^* - h(1-h_n^*)\big) = \sqrt{\frac{1-h}{h}}(1-h_n^*).$$

As $h \geq 1-h_n^*$, the decreasingness of $h \mapsto (1-h)^{1/2}h^{-1/2}$ on $(0,1]$ implies that this bound is maximal for $h = 1-h_n^*$, which gives $|r_n(s,h)| < r_n^*$.

b) If $h + h_n^* < 1$, noting that $s_n^* - s < h$, we obtain

$$|r_n(s,h)| < \frac{hh_n^*}{w(h)} = \sqrt{\frac{h}{1-h}} h_n^* < \sqrt{\frac{1-h_n^*}{1-(1-h_n^*)}} h_n^* = r_n^*,$$

since $h < 1 - h_n^*$ and $h \mapsto h^{1/2}(1-h)^{-1/2}$ increases on $[0,1)$.

5. If $s_n^* < s \leq s_n^* + h_n^* < s + h$, then

$$|r_n(s,h)| = \frac{1}{w(h)} |s_n^* + h_n^* - s - hh_n^*|.$$

5.1. If $s_n^* + h_n^* - s - hh_n^* \geq 0$, then $|r_n(s,h)| = w(h)^{-1}(s_n^* + h_n^* - s - hh_n^*)$. Noting that $s_n^* + h_n^* - s < h_n^*$ and $s_n^* + h_n^* - s < h$ we have

$$|r_n(s,h)| < \frac{1}{w(h)} \big(\min(h, h_n^*) - hh_n^* \big) = r_n(s_n^*, h) \leq r_n^*,$$

so (48) holds.

5.2. If $s_n^* + h_n^* - s - hh_n^* < 0$, then $|r_n(s,h)| = w(h)^{-1}(s - s_n^* - h_n^* + hh_n^*)$.

a) If $h + h_n^* \geq 1$, noting that $s - s_n^* < 1 - h$, we have

$$|r_n(s,h)| < \frac{1 - h - h_n^* + hh_n^*}{w(h)} = \frac{(1-h)(1-h_n^*)}{w(h)}.$$

The maximal value of this bound is r_n^* obtained for $h = 1 - h_n^*$ by decreasingness of $h \mapsto (1-h)^{1/2} h^{-1/2}$ on $(0,1]$. Hence (48) holds.

b) If $h + h_n^* < 1$, we simply note that $s - s_n^* - h_n^* \leq 0$, whence

$$|r_n(s,h)| \leq \frac{hh_n^*}{w(h)} = h_n^* \sqrt{\frac{h}{1-h}} < h_n^* \sqrt{\frac{1-h_n^*}{1-(1-h_n^*)}} = r_n^*,$$

by increasingness of $h \mapsto h^{1/2}(1-h)^{-1/2}$ on $[0,1)$.

The verification of (i) is now complete. □

Proof of (ii). From $|I_{s,h} \cap I_*| \leq n \min(h, h_n^*)$, we deduce

$$r_n^* - r(s,h) \geq \sqrt{h_n^*(1-h_n^*)} - \frac{\min(h, h_n^*) - hh_n^*}{\sqrt{h(1-h)}}.$$

If $h \le h_n^*$,

$$1 - \frac{r_n(s,h)}{r_n^*} \ge 1 - \frac{h(1-h_n^*)}{\sqrt{h_n^*(1-h_n^*)}\sqrt{h(1-h)}} = 1 - \frac{\sqrt{h}\sqrt{1-h_n^*}}{\sqrt{h_n^*}\sqrt{1-h}}$$
$$\ge 1 - \frac{\sqrt{h}}{\sqrt{h_n^*}} = \frac{h_n^* - h}{\sqrt{h_n^*}\big(\sqrt{h_n^*}+\sqrt{h}\big)} \ge \frac{h_n^* - h}{2h_n^*}.$$

Symmetrically, if $h \ge h_n^*$ we have

$$1 - \frac{r_n(s,h)}{r_n^*} \ge \frac{h - h_n^*}{2h}.$$

Hence we can summarize the two cases by writing

$$\forall (s,h) \in \Theta_n, \quad 1 - \frac{r_n(s,h)}{r_n^*} \ge \frac{|h - h_n^*|}{2\max(h, h_n^*)}. \tag{49}$$

Now using the alternative expression (5) of r_n and estimating $|I_{s,s+h}^c \cap I_*^c| \le n\min(1-h, 1-h_n^*)$ we similarly prove

$$\forall (s,h) \in \Theta_n, \quad 1 - \frac{r_n(s,h)}{r_n^*} \ge \frac{|h - h_n^*|}{2\max(1-h, 1-h_n^*)}. \tag{50}$$

Clearly (47) follows from (49) and (50). □

Lemma 6. *Let X_n, $n \ge 1$ be independent random elements in a measurable space E. Put $S_0 := 0$, $S_n := (\delta_{X_1} - \mathbf{E}\,\delta_{X_1}) + \cdots + (\delta_{X_n} - \mathbf{E}\,\delta_{X_n})$, $n \ge 1$. Assume that the seminorm $||\cdot||$ is admissible. Define*

$$R_n := n^{-1/2} \max_{0 \le i < j \le n} \frac{||S_j - S_i||}{\rho((j-i)/n)}, \quad n \ge 1 \tag{51}$$

where

$$\rho(h) = \sqrt{h(1-h)\log(\mathrm{e}/h(1-h))}, \quad 0 < h < 1.$$

Then the sequence $(R_n)_{n\ge1}$ is stochastically bounded.

Proof. It is easy to reduce the problem to proving stochastic boundedness of the sequence $(\widetilde{R}_n)_{n\ge1}$, where

$$\widetilde{R}_n = n^{-1/2} \max_{1\le \ell \le n} \frac{1}{\varrho(\ell/n)} \max_{0 \le k \le n-\ell} \|S_{k+\ell} - S_k\|,$$

with $\varrho(h) := \sqrt{h\log(\mathrm{e}/h)}$. We shall use dyadic splitting of the ℓ's and k's indexation ranges. Defining the integer J_n by

$$2^{J_n} \le n < 2^{J_n+1},$$

we get

$$\begin{aligned}
n^{1/2}\widetilde{R}_n &= \max_{1\le j\le J_n+1} \max_{n2^{-j}<\ell\le n2^{-j+1}} \frac{1}{\varrho(\ell/n)} \max_{1\le k\le n-\ell} \|S_{k+\ell}-S_k\| \\
&\le \max_{1\le j\le J_n+1} \max_{n2^{-j}<\ell\le n2^{-j+1}} \frac{1}{\varrho(2^{-j})} \max_{0\le k<n-n2^{-j}} \|S_{k+\ell}-S_k\| \\
&\le \max_{1\le j\le J_n+1} \max_{n2^{-j}<\ell\le n2^{-j+1}} \frac{1}{\varrho(2^{-j})} \max_{1\le i<2^j} \max_{(i-1)n2^{-j}\le k<in2^{-j}} \|S_{k+\ell}-S_k\|.
\end{aligned}$$

For $n2^{-j} < \ell \le n2^{-(j-1)}$ and $(i-1)n2^{-j} \le k < in2^{-j}$ we have

$$\begin{aligned}
\|S_{k+\ell}-S_k\| &\le \|S_{k+\ell}-S_{[in2^{-j}]}\| + \|S_{[in2^{-j}]}-S_k\| \\
&\le \max_{in2^{-j}<u<(i+2)n2^{-j}} \|S_u - S_{[in2^{-j}]}\| \\
&\quad + \max_{(i-1)n2^{-j}\le k<in2^{-j}} \|S_{[in2^{-j}]}-S_k\|,
\end{aligned}$$

where $[t]$ denotes the integer part of the real number t. Therefore

$$\widetilde{R}_n \le R'_n + R''_n,$$

where

$$R'_n = n^{-1/2} \max_{1\le j\le J_n+1} \frac{1}{\varrho(2^{-j})} \max_{1\le i<2^j} \max_{in2^{-j}<u<(i+2)n2^{-j}} \|S_u - S_{[in2^{-j}]}\|$$

$$R''_n = n^{-1/2} \max_{1\le j\le J_n+1} \frac{1}{\varrho(2^{-j})} \max_{1\le i<2^j} \max_{(i-1)n2^{-j}\le k<in2^{-j}} \|S_{[in2^{-j}]}-S_k\|.$$

Consider the probability $P_1(\lambda) = \Pr^*\{R'_n > \lambda\}$, $\lambda > 0$. We have

$$\begin{aligned}
P_1(\lambda) &\le \sum_{j=1}^{J_n+1} \Pr^*\Big\{ \max_{1\le i<2^j} \max_{in2^{-j}<u<(i+2)n2^{-j}} \|S_u - S_{[in2^{-j}]}\| > \lambda n^{1/2}\varrho(2^{-j})\Big\} \\
&\le \sum_{j=1}^{J_n+1} \sum_{1\le i<2^j} P^*\Big\{ \max_{in2^{-j}<u<(i+2)n2^{-j}} \|S_u - S_{[in2^{-j}]}\| > \lambda n^{1/2}\varrho(2^{-j})\Big\} \\
&= \sum_{j=1}^{J_n+1} \sum_{1\le i<2^j} P^*_{ij}(\lambda).
\end{aligned}$$

Applying Ottaviani's inequality (Ledoux and Talagrand (1991), Lemma 6.2) and admissibility of $\mathcal{F}$ we obtain

$$\Pr^*(R'_n > \lambda) \leq \sum_{j=1}^{J_n+1} 2^j \frac{c_1 \exp(-c_2\lambda^2 2^{j-3}\varrho^2(2^{-j}))}{1 - c_1 \exp(-c_2\lambda^2 2^{j-3}\varrho^2(2^{-j}))}$$

provided that the denominator above be positive for each $j \geq 1$. This condition is clearly satisfied for λ large enough (independently of n). Stochastic boundedness of $(R'_n)_{n\geq 1}$ is obtained now via the dominated convergence theorem for the series. The proof of stochastic boundedness of $(R''_n)_{n\geq 1}$ is clearly similar. □

Lemma 7. *Let the class $\mathcal{F}$ be admissible in the sense of Definition 1. For $d_n^{s,h}$ defined by* (17) *and for any sequence* $(\tau_n) \subset (0,1)$*,*

$$\sup_{h(1-h)\geq \tau_n} \|d_n^{s,h}\| = O_{\Pr}(\kappa_n n^{-1/2}).$$

where $\kappa_n = |\log \tau_n|^{1/2}$.

Proof. Set $\xi_j = \delta_{X_{n,j}} - \mathbf{E}\,\delta_{X_{n,j}}$ if $j \in I_*$ and $\xi'_j = \delta_{X_{n,j}} - \mathbf{E}\,\delta_{X_{n,j}}$ if $j \in I_*^c$. For $h \leq 1/2$

$$d_n^{s,h} = \frac{1}{nw(h)} \sum_{j\in I_{s,h}} (\alpha_j\xi_j + \beta_j\xi'_j) - \frac{h}{w(h)} \sum_{j=1}^{n} (\alpha_j\xi_j + \beta_j\xi'_j),$$

where $\alpha_j = 1-\beta_j = 1$ for $j \in I_*$ and $\alpha_j = 1-\beta_j = 0$ for $j \in I_*^c$. As $h \leq 1/2$, $h/w(h) \leq 1$, hence by admissibility of $\mathcal{F}$ we have

$$\frac{h}{w(h)} \Big\| \sum_{j=1}^{n} (\alpha_j\xi_j + \beta_j\xi'_j) \Big\| = O_{\Pr}(n^{-1/2}).$$

By Lemma 6

$$\sup_{h(1-h)\geq\tau_n} \frac{1}{nw(h)} \Big\| \sum_{j\in I_{s,h}} (\alpha_j\xi_j + \beta_j\xi'_j) \Big\| \leq$$
$$\log^{1/2}(\mathrm{e}/\tau_n) \sup_{h(1-h)\geq\tau_n} \frac{1}{n\rho(h)} \Big\| \sum_{j\in I_{s,h}} (\alpha_j\xi_j + \beta_j\xi'_j) \Big\| =$$
$$\log^{1/2}(\mathrm{e}/\tau_n) O_{\Pr}(n^{-1/2}).$$

If $h > 1/2$ we start with

$$d_n^{s,h} = \frac{1}{nw(h)} \sum_{j \in I_{s,h}^c} (\alpha_j \xi_j + \beta_j \xi_j') - \frac{1-h}{w(h)} \sum_{j=1}^{n} (\alpha_j \xi_j + \beta_j \xi_j')$$

and use the same arguments. □

5 Examples

In this section we discuss some examples of seminorms admissible in a sence of Definition 1. First two examples are taken from Dümbgen [6]. In what follows, for a measure ν on $\mathbb{R}$, $\nu(x) = \nu(-\infty, x]$.

Example 1. $E = \mathbb{R}^d$; $\mathcal{D}$ is a Vapnik-Červonenkis class of measurable subsets of $\mathbb{R}^d$,

$$N_n(\nu) = ||\nu||_{\mathcal{D}} = \sup_{A \in \mathcal{D}} |\nu(A)|.$$

The seminorm $||\nu||_{\mathcal{D}}$ is admissible. Particularly,

$$N_n(\nu) = ||\nu||_\infty = \sup_{x \in \mathbb{R}} |\nu(x)|.$$

Example 2. Consider $E = \mathbb{R}$. For $p \geq 1$, let

$$N_n(\nu) = \Big(\int_{\mathbb{R}} |\nu(x)|^p \, \mathrm{d}P_n^{0,n}(x) \Big)^{1/p}.$$

Evidently, $N_n(\nu) \leq ||\nu||_\infty$. To verify condition (11) with a given sequence (γ_n) it is sufficient to show that

$$\liminf_{n \to \infty} \gamma_n \int_{\mathbb{R}} |(Q_n - P_n)(x)| R_n(\mathrm{d}x) > 0,$$

where $R_n = h_n^* P_n + (1 - h_n^*) Q_n$. If for example $P_n = P$, $Q_n = Q$ and $h_n^* \to a \in (0,1)$, then this condition is valid, since $\int_{\mathbb{R}} |(P-Q)(x)|(P+Q)(\mathrm{d}x) > 0$, if $P \neq Q$.

Example 3 (p-variation norm)**.** Assume that the observations $X_{n,i}$ range in (a, b), $-\infty < a < b < \infty$.

In practice, the sup norm for empirical process is often not strong enough, see e.g., Dudley and Norvaiša [5] for examples. Instead p-variation norm is considered. For a real-valued function f on an interval J and $0 < p < \infty$, its p-variation on J is

$$v_p(f,J) := \sup\{\sum_{k=1}^{m} |f(t_i) - f(t_{i-1})|^p : t_0 \in J, t_0 < t_1 < \cdots < t_m \in J, m \geq 1\}.$$

Let f be such that $v_p(f) < \infty$. For $1 \leq p < \infty$ the p-variation seminorm is defined by $||f||_{(p)} := v_p^{1/p}(f)$ and the p-variation norm is then $||f||_{[p]} := ||f||_{(p)} + ||f||_\infty$.

If $p > 2$ then the p-variation norm is admissible in the sense of Definition 1. This easily follows from Huang and Dudley [7].

Consider now $N_n(\nu) = ||\nu||_{[p]}$. Then the condition (11) reduces to

$$\liminf_{n\to\infty} \gamma_n ||Q_n - P_n||_{[p]} > 0.$$

This condition is evidently satisfied if $P_n = P$, $Q_n = Q$ and $P \neq Q$.

May be the p-variation norm is too big to compute the quantities $||D_n^{s,h}||_{[p]}$ exactly. So, some weaker seminorms can be useful. For example one can consider variation of functions build on dyadic partition of the interval J. For simplicity, let $J = [0,1]$. Define

$$v_p^{\text{dyad}}(f) = \sup_{j\geq 0}\{\sum_{v\in V_j} |f(v^+) - f(v)|^p\}.$$

Here V_j is the set of dyadic numbers of the level j and $v^+ = v + 2^{-j}$ for $v \in V_j$. Define $||f||_{[p]}^{\text{dyad}} := ||f||_\infty + (v_p^{\text{dyad}}(f))^{1/p}$.

Example 4 (Reproducing kernel seminorms)**.** Let E be a metric space and let $\mathcal{M}(E)$ denote the space of signed measure on the Borel σ-field of E. As in [16], we consider the class of reproducing kernels $K : E \times E \to \mathbb{R}$ having the following representation

$$K(x,y) = \int_U r(x,u)\overline{r(y,u)}\rho(\mathrm{d}u), \quad x,y \in E, \tag{52}$$

where ρ is a positive measure on some measurable space $(U,\mathcal{U})$ and the function $r : E \times U \to \mathbb{C}$ satisfies

$$\sup_{x\in E} ||r(x,\cdot)||_{L_2(\rho)} < \infty. \tag{53}$$

We consider for $\nu \in \mathcal{M}(E)$

$$||\nu||_K = \Big(\int_{E^2} K(x,y)\nu\otimes\nu(\mathrm{d}x,\mathrm{d}y)\Big)^{1/2}.$$

Proposition 8. *With any reproducing kernel K defined by* (52) *and satisfying* (53)*, the seminorm $||\cdot||_K$ is admissible in the sense of Definition* 1*.*

The proof easily follows from well known exponential inequalities for sums of bounded Hilbert space random variables (see, e.g., [10])

Here are some examples of most interesting kernels.

1. Take for ρ the counting measure on $U = \mathbb{N}$ and define r by $r(x,i) = f_i(x)$, $x \in E$, where the sequence of functions $f_i : E \to \mathbb{R}$ separates the measures, i.e. the only $\nu \in \mathcal{M}(E)$ such that $\nu f_i := \int f_i \,\mathrm{d}\nu = 0$ for all $i \in \mathbb{N}$ is the null measure. Assume also $\sum_{i\in\mathbb{N}} ||f_i||_\infty^2 < \infty$. Then consider
$$K(x,y) = \sum_{i\in\mathbb{N}} f_i(x)f_i(y), \quad x, y \in E \times E.$$

2. Take $E = U = [0,1]$, $\rho = \lambda + \delta_1$, where λ is the Lebesgue measure and δ_1 is the Dirac mass at the point 1. With $r(x,u) = \mathbf{1}_{[x,1]}(u)$ we obtain $K(x,y) = 2 - \max\{x,y\}$, $x, y \in [0,1]$.

3. Let $E = U = \mathbb{R}^d$, $r(x,u) = \exp\{i\langle x,u\rangle\}$, $x, u \in \mathbb{R}^d$ and ρ a bounded positive measure on R^d. This gives the kernel
$$K(x,y) = \int_{\mathbb{R}^d} \exp\{i\langle x-y,u\rangle\}\rho(\mathrm{d}u), \quad x, y \in \mathbb{R}^d.$$
This example gives estimators based on empirical characteristic functions. The condition (11) becomes
$$\liminf_{n\to\infty} \gamma_n^2 \int_{\mathbb{R}^d} |\chi_{P_n}(u) - \chi_{Q_n}(u)|^2 \rho(\mathrm{d}u) > 0,$$
where χ_P denotes the characteristic function of the probability measure P.

References

[1] Avery, P.J., Henderson, D.A. (1999). Detecting a changed segment in DNA sequences. *J. of Roy. Statist. Soc.: Series C Appl. Stat.*, **48**, 489–503.

[2] Brodsky, B., Darkhovsky, B., (1993), *Nonparametric Methods in Change-Point Problems.* Kluwer Academic Publishers, the Netherland.

[3] Commenges, D., Seal, J., Pinatel, F. (1986). Inference about a Change Point in Experimental Neurophysiology. *Math. Biosc.*, **80**, 81–108.

[4] Csörgő, M., Horváth, L. (1997). *Limit Theorems in Change-Point Analysis.* John Wiley & Sons, New York.

[5] Dudley, R.M., Norvaiša, R., (1999), *Differentiability of Six Operators on Nonsmooth Functions and p-Variation*, Lecture Notes in Math., **1703**, Springer, New York.

[6] Dümbgen, L. (1991). The asymptotic behavior of some nonparametric change-point estimators, *Ann. Stat.*, **19**, No 3, 1471–1495.

[7] Huang, Yen-Chin, Dudley, R.M., (2001), Speed of convergence of classical empirical processes in p-variation norm, *Ann. Probab.* **29**, 1625–1636.

[8] Lavielle, M., (1999), Detection of multiple changes in a sequence of dependent variables, *Stoch. Proc. Applic.* **83**, 79–102.

[9] Lavielle, M., Moulines, E., (1999), Least squares estimation of an unknown number of shifts in a time series, *J. Time-Ser. Anal.* , ????.

[10] Ledoux, M., Talagrand, M. (1991). *Probability in Banach Spaces.* Springer Verlag, Berlin, Heidelberg.

[11] Chung-Bow Lee, (1996), Nonparametric multiple change-point estimators, *Stat. Probab. Lett.* **27**, 295–304.

[12] Račkauskas, A., Suquet, Ch. (2003a). Hölder norm test statistics for epidemic change. *Journal of Statistical Planning and Inference* 2004, vol. **126**, Issue 2, 495–520.

[13] Račkauskas, A., Suquet, Ch. (2006). Testing epidemic change of infinite dimensional parameters. *Stat. Inf. for Stoch. Proc.* **9**, 111–134.

[14] Schechtman, E., Wolfe, D., (1985), Multiple change points problem - nonparametric procedures for estimation of the points of change, *Comun.Statist.-Simul. Comput.* **14**, 615–631.

[15] Shorack G.R. and Wellner, J.A. (1986). *Empirical processes with applications to statistics.* John Wiley & Sons, New York.

[16] Suquet, Ch. (1995), Distances euclidiennes sur les mesures signées et application à des théorèmes de Berry-Esséen. *Bull. Belg. Math. Soc.* **2**, 161–181.

[17] Yao, Y.C., (1988), Estimating the number of the change-points by Schwarz's criterion, *Statist. Probab. Lett.* **6**, 181–187.

[18] Yao, Q. (1993). Tests for change-points with epidemic alternatives. *Biometrika*, **80**, 179–191.