



HAL
open science

A correction on approximation of smoothing probabilities for hidden Markov models

Jüri Lember

► **To cite this version:**

Jüri Lember. A correction on approximation of smoothing probabilities for hidden Markov models. Statistics and Probability Letters, 2011, 81 (9), pp.1463. 10.1016/j.spl.2011.04.009 . hal-00762903

HAL Id: hal-00762903

<https://hal.science/hal-00762903>

Submitted on 9 Dec 2012

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.

Accepted Manuscript

A correction on approximation of smoothing probabilities for hidden Markov models

Jüri Lember

PII: S0167-7152(11)00148-9
DOI: [10.1016/j.spl.2011.04.009](https://doi.org/10.1016/j.spl.2011.04.009)
Reference: STAPRO 5984

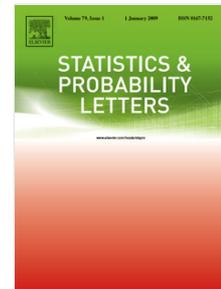
To appear in: *Statistics and Probability Letters*

Received date: 13 April 2011

Accepted date: 15 April 2011

Please cite this article as: Lember, J., A correction on approximation of smoothing probabilities for hidden Markov models. *Statistics and Probability Letters* (2011), doi:[10.1016/j.spl.2011.04.009](https://doi.org/10.1016/j.spl.2011.04.009)

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.



A correction on approximation of smoothing probabilities for hidden Markov models

Jüri Lember¹

Tartu University, J. Liivi 2 - 507, Tartu 50408, Estonia

Abstract

In this note, we correct a mistake concerning Theorem 2.1 in (Lember 2011a)

Keywords: Hidden Markov models, smoothing, segmentation

1. Introduction

Let $Y = \{Y_k\}_{-\infty}^{\infty}$ be a double-sided stationary MC with states $S = \{1, \dots, K\}$ and irreducible aperiodic transition matrix $(P(i, j))$. Let $X = \{X_k\}_{-\infty}^{\infty}$ be the (double-sided) process such that: 1) given $\{Y_k\}$ the random variables $\{X_k\}$ are conditionally independent; 2) the distribution of X_n depends on $\{Y_k\}$ only through Y_n . The process X is sometimes called the *hidden Markov process* (HMP) and the pair (Y, X) is referred to as *hidden Markov model* (HMM).

In (Lember 2011a), the approximation of smoothing probabilities $P(Y_t \in \cdot | X_1, \dots, X_n) =: P(Y_t \in \cdot | X^n)$ is considered. The only assumption on HMM is the so-called cluster assumption **A** that relaxes the ones usually made in the literature like the strong mixing assumption. Unfortunately, a mistake occurred in one of the main results in (Lember 2011a), namely in Theorem 2.1. In this note, we present the corrected version of Theorem 2.1 and briefly discuss the validity of other main results. The full corrected paper with complete proofs can be found in arXiv (Lember 2011b).

2. Main results

The first main result of (Lember 2011a) is Corollary 2.1 that states that under **A**, there exists a finite random variable C_1 such that for every z, t, n such that $z \leq 1 \leq t \leq n$,

$$\|P(Y_t \in \cdot | X_z^n) - P(Y_t \in \cdot | X_1^n)\| \leq C_1 \rho_1^{t-1}, \quad \text{a.s.} \quad (1)$$

In a sense, this is the most important result, generalizing the previous results of the kind (actually, it should be called as a "Theorem", whereas Theorem 2.1 is essentially a corollary). This result as well as preceding results Proposition 2.1 and Lemma 2.1 holds true; moreover (1) holds true when "1" is replaced by more general $z_1 \leq 1$. With this little but important generalization, Corollary 2.1 is as follows.

Corollary 2.1. *Assume **A**. Then, there exists a non-negative finite random variable C_1 as well as constant $\rho_1 \in (0, 1)$ such that for every $z_2 \leq z_1 \leq 1 \leq t \leq n$,*

$$\|P(Y_t \in \cdot | X_{z_1}^n) - P(Y_t \in \cdot | X_{z_2}^n)\| \leq C_1 \rho_1^{t-1}, \quad \text{a.s.} \quad (2)$$

For detailed proof and explicit definition of C_1 , see (Lember 2011b).

The mistake in (Lember 2011a) is due to misinterpretation of the stationarity of the hidden Markov process. Indeed, by the stationarity, the probabilistic behavior is invariant under the shift so that shifting everything in time does not change the law of the bounds C_1 but changes them as the random variables. To make this clear and to simplify the proof of Theorem 2.1, we state the following corollary.

Email address: jyri1@ut.ee (Jüri Lember)

¹Estonian science foundation grant no 7553

Corollary 2.2. *There exist a constant $\rho_1 \in (0, 1)$ and an ergodic process $\{C_z\}_{z=-\infty}^{\infty}$ so that for any $z_2 \leq z_1 \leq z \leq t \leq n$*

$$\|P(Y_t \in \cdot | X_{z_1}^n) - P(Y_t \in \cdot | X_{z_2}^n)\| \leq C_z \rho_1^{t-z}, \quad \text{a.s.}$$

The proof of the corollary consists of one line, see (Lember 2011b). Of course, without the ergodicity, the existence of random variables C_z would be trivial.

In the original version, the second main result of (Lember 2011a) states that under **A**, there exists a finite random variable C and $\rho_o \in (0, 1)$ such that for every z, t, n satisfying $z \leq t \leq n$

$$\|P(Y_t \in \cdot | X_z^n) - P(Y_t \in \cdot | X_{-\infty}^{\infty})\| \leq C(\rho_o^{t-z} + \rho_o^{n-t}) \quad \text{a.s.} \quad (3)$$

Unfortunately in such a form, the statement does not hold: the random variable C in (3) would in general depend on z and n (unless, obviously, it is bounded). Of course, if C is allowed to depend on z and n , the existence of it would be trivial. Hence, we restate (3) as follows.

Theorem 2.1. *Assume **A**. Then there exist a constant $\rho_o \in (0, 1)$ and a finite ergodic process $\{C'_z\}_{z=-\infty}^{\infty}$ so that for every $z \leq 1 \leq t \leq k \leq n$*

$$\|P(Y_t \in \cdot | X_z^n) - P(Y_t \in \cdot | X_{-\infty}^{\infty})\| \leq C_1 \rho_o^{t-1} + C'_k \rho_o^{k-t} \quad \text{a.s.}, \quad (4)$$

where C_1 is a finite random variable as in Corollary 2.1.

For the proof, see (Lember 2011b). Now, the theorem is more informative, because the both side bounds are separated, n is allowed to be greater than k (important in applications, see the proof of Theorem 3.1) and z is allowed to be smaller than 1. Again, the existence of C'_k is trivial, but the importance lies in the ergodicity of $\{C'_k\}$.

3. Convergence of risks

As an application, in (Lember 2011a), the convergence of risks $R_1(X^n)$ is considered. The third main result of (Lember 2011a) is the following.

Theorem 3.1. *Suppose **A** holds. Then there exists a constant R such that $R(X^n) \rightarrow R$ a.s. and in L_1 .*

That theorem holds true. The proof would remain unchanged, if $E(C'_n) < \infty$, since then $\frac{C'_n}{n} \rightarrow 0$, a.s. Unfortunately, it has not been proven that $E(C'_n) < \infty$, so we use another slightly different argument that relies on the ergodicity of $\{C'_n\}$. The new proof, given in (Lember 2011b) is a couple of lines longer but is based on the same ideas.

References

- Jüri Lember. On approximation of smoothing probabilities for hidden Markov models. *Statistics & Probability Letters*. **81** (2011a), 310 – 316.
 Jüri Lember. On approximation of smoothing probabilities for hidden Markov models. arXiv:0910.4636v2, (2011b).