

NONPARAMETRIC DENSITY ESTIMATION WITH REGION-CENSORED DATA

Youssef Bennani, Luc Pronzato & Maria João Rendas

Univ. Nice Sophia Antipolis, CNRS, I3S, UMR 7271, 06900 Sophia Antipolis, France

ABSTRACT

The paper proposes a new Maximum Entropy estimator for non-parametric density estimation from region censored observations in the context of population studies, where standard Maximum Likelihood is affected by over-fitting and non-uniqueness problems. The link between Maximum Entropy and Maximum Likelihood estimation for the exponential family has often been invoked in the literature. When, as it is the case for censored observations, the constraints on the Maximum Entropy estimator are derived from independent observations of a set of non-linear functions, this link is lost increasing the difference between the two criteria. By combining the two criteria we propose a novel density estimator that is able to overcome the singularities of the Maximum Likelihood estimator while maintaining a good fit to the observed data, and illustrate its behavior in real data (hyperbaric diving).

Index Terms— Censored observations, non-parametric maximum likelihood, constrained maxent, regularisation.

1. INTRODUCTION

1.1. Motivation

The density estimation problem addressed in this paper is motivated by a problem of population analysis: we are interested on the distribution π_θ of the biophysical parameters θ of a mathematical model [1] for the instantaneous volume of micro-bubbles flowing through the right ventricle of a diver's heart when executing a decompression profile P :

$$(\theta, \{P(u)\}_{u \leq t}) \rightarrow B(t, \theta, \{P(u)\}_{u \leq t}).$$

The problem appears in the context of prevention of decompression sickness (DCS) in deep sea diving: since DCS is known to be highly correlated with the presence of gas bubbles in the blood, ability to correctly predict the probability that this volume becomes exceedingly high can be used to establish safety rules that avoid profiles that will be dangerous for a non-negligible part of the population.

The instantaneous gas volume B is observed through periodic measurements of bubble grades G . Since it is usually accepted that DCS is related to the maximum observed

This work has been partially supported by DGA (France), through contract SAFEDIVE.

grade, only strongly quantified versions of the peak volume $b(\theta, P) = \max_t B(t, \theta, \{P(u)\}_{u \leq t})$ have been recorded:

$$G(\theta, P) = l \Leftrightarrow b(\theta, P) \in [\tau_l, \tau_{l+1}[, \quad l \in \{0, \dots, L\}. \quad (1)$$

In our case $L = 4$, and thresholds $\tau = \{\tau_l\}_{l=1}^L$, where $\tau_0 = 0 < \tau_1 < \dots < \tau_L < \tau_{L+1} = \infty$, are assumed known. Fig. 1 illustrates these definitions. A simplified model with $\theta \in \Theta \subset \mathbb{R}^2$ has been used, all other parameters of model [1] being held fixed. Note that all θ in region R_i

$$R_i \equiv \{\theta \in \Theta : b(\theta, P_i) \in [\tau_{G_i}, \tau_{G_i+1}[\} ,$$

yield the same observed grade G_i for profile P_i . In Fig. 1 (right) we draw the regions corresponding to the profile P shown on the left. The plot in the centre shows $B(t, \theta, \{P(u)\}_{u \leq t})$ (thresholds τ are indicated by the horizontal red lines). A value grade $G = 3$ corresponds to a value of θ inside the orange region.

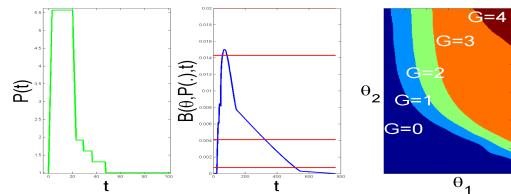


Fig. 1. Left: profile $P(t)$. Centre: instantaneous volume B (blue) and thresholds (red). Right: regions corresponding to the 5 possible grades.

The remark above shows that estimation of π_θ from observations $\{(G_i, P_i)\}_{i=1}^n$ is equivalent to the problem of estimating π_θ from observation of the set of regions $\{R_i\}_{i=1}^n$. In the absence of knowledge about the expected dispersion of the biological parameters, we estimate π_θ non-parametrically, imposing no constraints on its shape.

1.2. Background

Finding $\hat{\pi}_\theta$, the NPMLE (Non-parametric Maximum Likelihood Estimate) of π_θ from censored observations has been studied by many authors starting with the pioneering formulation of the Kaplan-Meier product-limit estimator [2]. Several types of censoring (one-sided, interval) have been considered since, first for scalar and more recently for multivariate

distributions. The problem assessed here departs from previous studies in that our (multi-dimensional) censoring regions can have arbitrary geometry. To emphasise this we speak of “region-censoring”, instead of “interval-censoring.”

Several facts are known about the NPMLE for interval-censored observations: (i) its support $\mathcal{S}_{\text{NPMLE}} = \{\theta : \hat{\pi}_\theta(\theta) > 0\}$ is confined to a finite number K of disjoint intervals (the so called “elementary regions”): $\mathcal{S}_{\text{NPMLE}} = \bigcup_{\ell=1}^K E_\ell, E_\ell \cap E_q = \emptyset, \ell \neq q$; (ii) all distributions that put the same probability mass $w_\ell \equiv \{\pi_\theta(E_\ell)\}, \ell = 1, \dots, K$ in these intervals have the same likelihood; (iii) there is in general no unique assignment of probabilities $\{\hat{w}_\ell\}_{\ell=1}^K$ that maximises the likelihood.

Turnbull [3] has first demonstrated (i), giving an algorithm to find the $\{E_\ell\}_{\ell=1}^K$ for the scalar case. Gentleman and Vandal [4] addressed the multivariate case, showing the E_ℓ 's are determined by the maximal cliques of the intersection graph of the set of observed regions $\{R_i\}_{i=1}^n$. We have shown elsewhere [5] that (i) also holds when the sets R_i have arbitrary geometry, but that some elementary regions are now associated to non-maximal cliques of the intersection graph, requiring a slightly more complex identification of the sets E_ℓ which we do not detail here. Fact (i) shows that the NPMLE problem can be studied in the K -dimensional probability simplex \mathbb{S}^K , since $\hat{\pi}_\theta(\cdot)$ is determined only up to the probability vector $\hat{\mathbf{w}} = \{\hat{w}_1, \dots, \hat{w}_K\}$. The two types of “non-uniqueness” of the NPMLE, (ii) and (iii), have been first pointed out by Turnbull [3]. More recently, they were studied in detail for the multi-variate case in [4], where the authors coined the terms *representational* – (ii) – and *mixture* – (iii) – non-uniqueness, further showing that the set of NPMLE's $\hat{\pi}_\theta$ is a polytope.

1.3. Contributions

Together, the facts presented above indicate that the NPMLE may frequently exhibit a singular behaviour, in the sense that its mass is concentrated in a subset of Θ of small Lebesgue measure, indicating “over-homogenous” population models. This may lead to dangerous biases in the context of risk assessment, by not taking into consideration the presence of individuals for which risk can be large.

We propose to remove ambiguity in the estimation of π_θ by relying in the principle of Maximum Entropy (*maxent*) [6], that finds the most un-informative density that can explain the observed data. While *maxent* has been frequently used for density estimation from joint observation of empirical moments of a set of features, its use for region-censored data arising from strongly quantified data, as we propose in this paper, is, as far as we know, novel. In particular, the equivalence between regularised *maxent* and penalised Maximum Likelihood in the exponential family [7, 8] is lost in our case. This leads us to formulate the density estimation problem as finding the most likely constrained *maxent* density. We com-

pare the proposed estimator to the NPMLE and to the best fitting *maxent* solutions, in real data from hyperbaric diving, showing that the resulting distribution is a better candidate than NPMLE or *maxent* alone for the distribution of biological parameters in a given population. The Rényi entropy [9] is a generalised entropy, obtained by considering an exponential mean in the postulates that lead to the usual Shannon entropy, which can also be shown to yield a measure of mutual information. By using the Rényi entropy of order two instead of the Shannon entropy, we are led to a quadratic optimisation problem with linear inequality constraints that has an efficient numerical solution.

The paper is organised as follows. Section 2 presents NPMLE estimation and shows (section 2.2) that even the most uncertain NPMLE will in general still present singularities that are unlikely to occur in a natural population. Section 3 considers Maximum entropy estimators, and presents (section 3.2) the main contribution of the paper, the most likely Rényi-*maxent* estimator, discussing relations to previous work and demonstrating that it leads to better estimates.

2. THE NPMLE

$\mathcal{Q}^{(j)}$: partition associated to profile $P^{(j)}$.

Let m denote the number of distinct profiles $\{P^{(j)}\}_{j=1}^m$ in the available dataset, and $\{R_i^{(j)}\}_{i=0,j=1}^{L,m}$ the regions

$$R_i^{(j)} = \left\{ \theta \in \Theta : b(\theta, P^{(j)}) \in [\tau_i, \tau_{i+1}] \right\}. \quad (2)$$

$\mathcal{Q}^{(j)} = \{R_i^{(j)}\}_{i=0}^L$ is a partition of Θ , see Fig. 1 (right).

Let \mathcal{Q} denote the smallest partition of Θ whose generated σ -algebra, $\sigma(\mathcal{Q})$, contains all partitions $\{\mathcal{Q}^{(j)}\}_{j=1}^m$. Denote by M its size. Let $\mathbf{E}_i^{(j)}$ be the set of elements of \mathcal{Q} that intersect $R_i^{(j)}$, and L_A be the list of regions $R_i^{(j)}$ that contain $A \in \mathcal{Q}$. Since the “elementary regions” $\{E_\ell\}_{\ell=1}^K$ are elements of \mathcal{Q} , notations $\mathbf{E}_i^{(j)}$ and L_{E_ℓ} are well defined.

2.1. Non-Parametric Maximum Likelihood Estimate

Let n be the total number of dives observed. In general, the same grade $G = i$ has been observed for the same profile $P = P^{(j)}$ more than once. Let n_j be the number of times $P^{(j)}$ has been executed, and $n_i^{(j)}$ the number of times grade i has been observed for $P^{(j)}$, such that $\sum_{i=0}^L n_i^{(j)} = n_j$ and $\sum_{j=1}^m n_j = n$. Denote by $\mathbf{f}^{(j)} = \{f_i^{(j)}\}_{i=0}^L$ the empirical distribution of the grades in $P^{(j)}$, $f_i^{(j)} = \frac{n_i^{(j)}}{n_j}$. Assuming that divers have been independently drawn in a population with probability distribution π_θ , the log-likelihood is

$$\mathcal{L}(\pi_\theta; \{n_i^{(j)}, R_i^{(j)}\}) = \sum_{j=1}^m \sum_{i=0}^L n_i^{(j)} \log \pi_\theta(R_i^{(j)}). \quad (3)$$

From property (i) of the NPMLE, $\hat{\pi}_\theta(R_i^{(j)} \cap \overline{S_{\text{NPMLE}}}) = 0$, with \bar{A} the complement of set A , and thus

$$\pi_\theta(R_i^{(j)}) = \sum_{E \in \mathbf{E}_i^{(j)}} \pi_\theta(E) = \mathbf{B}_{i.}^{(j)} \mathbf{w} , \quad (4)$$

where $\mathbf{B}_{i.}^{(j)}$ is the i -th row of $\mathbf{B}^{(j)}$, the $(L+1) \times K$ binary matrix, with $\mathbf{B}^{(j)}(i, \ell) = 1 \Leftrightarrow E_\ell \in \mathbf{E}_i^{(j)}$, and $\mathbf{w} \in \mathbb{S}^K$ the vector of probabilities of the E_ℓ 's: $\{w_\ell = \pi_\theta(E_\ell)\}_{\ell=1}^K$. Eqs. (3)–(4) show that all π_θ leading to the same \mathbf{w} have the same likelihood (property (ii)). In general, see [10], there is no single \mathbf{w} maximising (3): let $\hat{\mathbf{w}}$ be an NPMLE, then all elements of $\mathcal{P} = \{\mathbf{w}, \text{ s.t. } \forall j, \mathbf{B}^{(j)}\mathbf{w} = \mathbf{B}^{(j)}\hat{\mathbf{w}}\}$ are NPMLE's. We call \mathcal{P} the NPMLE polytope.

Several algorithms have been proposed to maximise (3), see e.g. [10]. It can be shown that the problem is equivalent to an optimal design problem, where \mathbf{w} plays the role of the design measure, enabling application of a vast collection of efficient algorithms originating from optimal design for NPMLE estimation. We rely on a multiplicative algorithm ([11, 12]):

$$w_\ell^{(t+1)} = \frac{1}{n} \left(\sum_{j=1}^m \sum_{i=0}^L n_i^{(j)} \frac{\mathbf{B}_{i\ell}^{(j)}}{\mathbf{B}_{i.}^{(j)} \mathbf{w}^{(t)}} \right) w_\ell^{(t)} , \quad (5)$$

initialized at some strictly positive $\mathbf{w}^{(0)}$ until

$$\max_{\ell \in \{1, \dots, K\}} \frac{1}{n} \left(\sum_{j=1}^m \sum_{i=0}^L n_i^{(j)} \frac{\mathbf{B}_{i\ell}^{(j)}}{\mathbf{B}_{i.}^{(j)} \mathbf{w}} - 1 \right) < \delta \ll 1.$$

The speed of the algorithm is improved by relying on the ability to detect the entries of \mathbf{w} that will converge to zero, as shown in [13]. We remark that this multiplicative algorithm can be applied to the complete partition \mathcal{Q} , and will automatically set to zero the entries of \mathbf{w} that do not correspond to “elementary sets” $\{E_\ell\}_{\ell=1}^K$. The computationally expensive analysis of the intersection graph described in [5] is traded by the need to solve a larger optimisation problem. For the dataset sizes of our study ($m = 19, M = 665$), we observed very fast convergence of (5) for the complete \mathcal{Q} .

The result for a total of $m = 19$ profiles, repeated a number of times ranging from 12 to 41 is shown on the left of Fig. 2 (see also top left of Fig. 3), clearly displaying the singularities that affect NPMLE's. The white regions have zero probability mass, and the estimated density has large peaks in a few small dispersed regions.

2.2. Rényi-maxent NPMLE

Motivated by the context of risk assessment, we rely on the notion of entropy to select the element of the NPMLE polytope that is the least informative, and that will thus better reflect the possible diversity of the population analysed.

All $\mathbf{w} \in \mathcal{P}$ define the same measures $\hat{f}_i^{(j)}$ over the partitions $\mathcal{Q}^{(j)}$ associated to the profiles $P^{(j)}$. Let $\tilde{\pi}_\theta \in \mathcal{P}$ be the

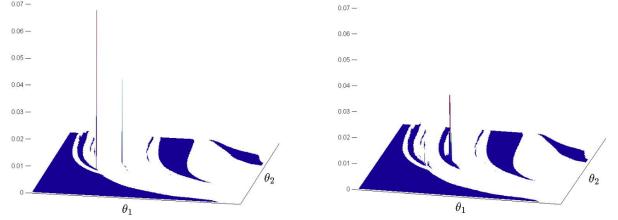


Fig. 2. Left: $\hat{\pi}_\theta$, one NPMLE solution found by (5). Right: $\tilde{\pi}_\theta$, the Rényi-maxent NPMLE. The white regions have zero probability mass.

distribution maximising the entropy while keeping unchanged the $\{\hat{f}_i^{(j)}\}_{i=0,j=1}^{L,m}$:

$$\tilde{\pi}_\theta(R_i^{(j)}) = \hat{f}_i^{(j)} = \hat{\pi}_\theta(R_i^{(j)}) . \quad (6)$$

For the Shannon entropy $H(\pi) = \text{E}_\pi[-\log(\pi(\cdot))]$ it is well known (Boltzman theorem) that the *maxent* estimate of π under a set of constraints of the form $\text{E}_\pi[f_i^{(j)}(\cdot)] = \gamma_i^{(j)}$ is the Gibbs distribution. In our problem, the $f_i^{(j)}(\cdot)$ – usually called “features” in the context of *maxent* – are the indicator functions $f_i^{(j)}(\theta) \equiv \mathbf{1}_{R_i^{(j)}}(\theta)$, showing that the Shannon-maxent solution is indeed constant inside the elements of \mathcal{Q} :

$$\tilde{\pi}_\theta^H(\theta) = \frac{1}{Z_\lambda} \prod_{j=1}^m \exp \left(\sum_{i=0}^L \lambda_i^{(j)} f_i^{(j)}(\theta) \right) ,$$

where Z_λ is a normalising constant (the partition function), and the $\{\lambda_i^{(j)}\}_{i=0,j=1}^{L,m}$ are determined such that (6) are satisfied.

Finding the $\tilde{\pi}_\theta$ for the Shannon entropy is a complex non-linear constrained optimisation problem. If instead we maximise the Rényi entropy of order 2

$$h_2(\pi_\theta) = -\log \int_{\Theta} \pi_\theta(\theta)^2 d\theta ,$$

also called *extension entropy* [9], it can be shown [14] that

$$\tilde{\pi}_{h_2}(\theta) = \left[-\frac{1}{2} \sum_{j=1}^m \sum_{i=0}^L \lambda_i^{(j)} f_i^{(j)}(\theta) \right]_+ ,$$

where $[.]_+ = \max(\cdot, 0)$ and the $\lambda_i^{(j)}$ are again such that (6) are satisfied, showing that the Rényi-maxent distribution is also constant inside the elements of \mathcal{Q} .

Let $\hat{\mathbf{w}}$ be a solution obtained at convergence of (5) and \mathbf{B} the matrix that stacks the $\mathbf{B}^{(j)}, j = 1, \dots, m$. The Rényi-maxent NPMLE probability vector $\tilde{\mathbf{w}}$ is the solution of the following quadratic program with linear equality constraints, for which efficient solutions exist

$$\tilde{\mathbf{w}} = \underset{\mathbf{w} \in \mathbb{S}^K}{\operatorname{argmin}} \quad \sum_{\ell=1}^K \frac{1}{\nu(E_\ell)} w_\ell^2 , \quad \text{s.t.} \quad \mathbf{B}\mathbf{w} = \mathbf{B}\hat{\mathbf{w}} .$$

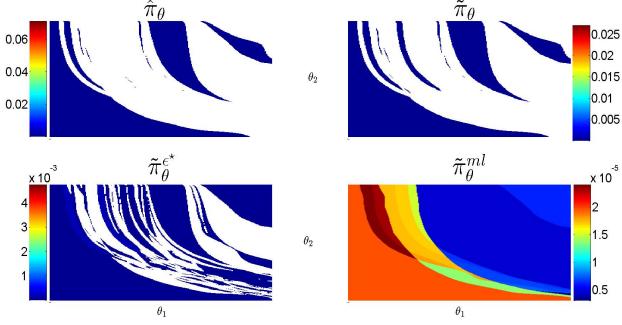


Fig. 3. Estimates of π_θ . Top left: one NPMLE solution found by (5). Top right: Rényi-maxent NPMLE. Bottom left: Rényi-maxent $\tilde{\pi}_\theta^{\epsilon^*}$. Bottom right: ML-Rényi-maxent $\tilde{\pi}_\theta^{ml}$. White regions have zero probability mass.

The Rényi-maxent NPMLE computed using the routine *quadprog* of Matlab is displayed in the right of Fig. 2 (3D plot) and in the top right of Fig. 3 (top view). We can see that the support of $\tilde{\pi}_\theta$ is larger than the support of $\hat{\pi}_\theta$, but that restriction of the solution to the NPMLE polytope keeps the density concentrated in a set of small disconnected regions, with large areas of zero measure (white areas).

3. REGULARIZED MAXENT

As we saw, even the *maxent* NPMLE still assigns negligible or zero mass to most regions of Θ . This is inherent to the likelihood criterion, that favours the most concentrated densities that are able to explain the observed data. It is easy to see that the support of an NPMLE density may suffer an important decrease when a single profile that is executed only once is added to the dataset, confirming that maximum likelihood estimation is ill-conditioned for censored observations. Suppose a new profile $P^{(m+1)}$ is added to the dataset, being executed only once. i.e., $n_{i^*}^{(m+1)} = n^{(m+1)} = 1$ and $n_i^{(m+1)} = 0, i \neq i^*$. Let \mathcal{Q} be the partition of Θ corresponding to the profiles with $j \leq m$ and \mathcal{Q}' the new partition, that also integrates $P^{(m+1)}$. If $R_{i^*}^{(m+1)}$ intersects an elementary set $E_\ell \in \mathcal{Q}$, such that

$$E'_\ell = E_\ell \cap R_{i^*}^{(m+1)} \in \mathcal{Q}', \text{ with } \nu(E'_\ell) \ll \nu(E_\ell),$$

then $E_\ell \setminus E'_\ell$ will no longer be an elementary set, shrinking the support of the NPMLEs.

3.1. Rényi-maxent

To avoid this singular behaviour, we must estimate π_θ with a criterion other than Maximum Likelihood. We propose to estimate π_θ as the Rényi-maxent distribution that best matches the *observed frequencies* for each profile, $\mathbf{f}^{(j)}$, which, as we

saw above, can be written as empirical averages of the indicator functions of the elements of the $\mathcal{Q}^{(j)}$. Note that while in the previous section the constraints were determined from the NPMLE polytope, here they are directly obtained from the data.

If there exists a π that can satisfy all constraints, the corresponding \mathbf{w} belongs to the NPMLE polytope \mathcal{P} . However, in general the m constraints will be inconsistent and, as in [8], we consider entropy maximisation under relaxed constraints. Let $\epsilon^* \geq 0$ be the smallest value of ϵ for which there exists a solution to the problem

$$\begin{aligned} \tilde{\pi}_\theta^\epsilon &= \operatorname{argmax}_\pi h_2(\pi) \\ \text{s.t. } & \left\| \Sigma^{(j)^{-1/2}} \left(\mathbf{E}_\pi[\check{\mathbf{f}}^{(j)}] - \check{\mathbf{f}}^{(j)} \right) \right\|_\infty \leq \epsilon, \quad \forall j, \end{aligned}$$

where $\Sigma^{(j)}$ is the $L \times L$ covariance matrix of the empirical distribution of the vector $\check{\mathbf{f}}^{(j)}$ obtained from $\mathbf{f}^{(j)}$ by removing the last entry. Note that the equivalence between relaxed *maxent* and penalised likelihood used in [8] does not hold in our case.

For our hyperbaric dataset, the Rényi-maxent solution corresponding to ϵ^* is shown in the bottom left of Fig. 3. As we can see, the support of this best fitting *maxent* density is still composed of a number of disjoint regions, and does not seem a plausible model of a biological population.

3.2. Most likely Rényi-maxent

In [7] duality arguments were used to show that, for the Shannon entropy, the solution to the problem above, when the constraints $f_i^{(j)}$ on the expected values are all obtained using the same empirical distribution – derived from an underlying data set $\{\theta_i\}_{i=1}^N$ – is the same as a L_1 -penalised maximum likelihood estimate of π from data $\{\theta_i\}_{i=1}^N$. For the censored data problem considered in this paper, the constraints are empirical averages, derived from independent datasets, and this equivalence is lost.

Moreover, since we rely on the L_∞ metric to evaluate deviation of the modelled distributions (by $\tilde{\pi}_\theta$) with respect to the empirical $\check{\mathbf{f}}^{(j)}$, and L_∞ is not equivalent to the (Riemannian) metric induced by Maximum Likelihood estimation for the exponential family, we cannot guarantee that $\mathcal{L}(\tilde{\pi}_\theta^\epsilon) < \mathcal{L}(\tilde{\pi}_\theta^{\epsilon^*})$, for $\epsilon > \epsilon^*$. In fact, as shown in Fig. 4 that plots the likelihood of $\tilde{\pi}_\theta^\epsilon$ as a function of ϵ/ϵ^* , this is not true.

We denote by $\tilde{\pi}_\theta^{ml}$ the most likely Rényi-maxent solution:

$$\tilde{\pi}_\theta^{ml} = \operatorname{argmax}_{\tilde{\pi}_\theta^\epsilon, \epsilon > \epsilon^*} \mathcal{L}(\tilde{\pi}_\theta^\epsilon; \{n^{(j)}, R^{(j)}\}). \quad (7)$$

This solution is displayed in the bottom left of Fig. 3. We can see that the support of $\tilde{\pi}_\theta^{ml}$ is the entire Θ , with a smoother distribution of the probability mass, being a more plausible characterisation of the natural variation within a biological population.

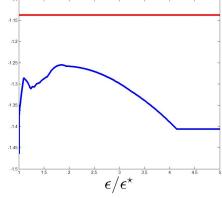


Fig. 4. Variation of $\mathcal{L}(\tilde{\pi}_\theta^\epsilon)$ with ϵ/ϵ^* . Red line: $\mathcal{L}(\hat{\pi}_\theta)$.

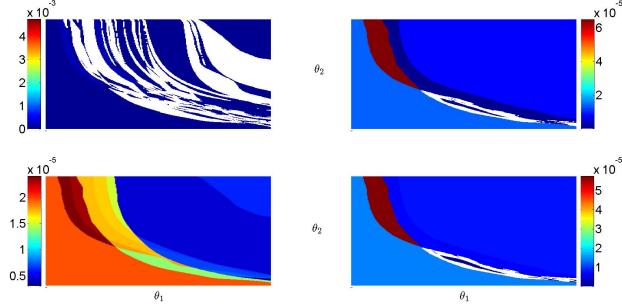


Fig. 5. Top: $\tilde{\pi}_\theta^{\epsilon^*}$; bottom: $\tilde{\pi}_\theta^{ml}$. Left: full covariance matrix; right: diagonal weighting.

Finally, we remark that it is important to take into account the correlation between the observed frequencies, opposed to what is done in [8]. Fig. 5 plots $\tilde{\pi}_\theta^{\epsilon^*}$ (top) and $\tilde{\pi}_\theta^{ml}$ (bottom) obtained using (7) with the full matrix (left) and using only the diagonal of $\Sigma^{(j)}$ (right). As we can see, use of an inappropriate metric directed the estimator towards sets of solutions that have no ability to reproduce the observed frequencies, and have thus lower likelihood.

4. SUMMARY

The paper studied identification of a probability density from region-censored observations, with application to modelling of decompression sickness during hyperbaric diving. We show that the NPMLE is intrinsically ill-posed, leading to unstable solutions which are biologically implausible. Expressing counts of the censored observations as empirical means of a set of features, we derive the *maxent* solution that best approximates the empirical distributions. The degree of fitting to the observed frequencies is chosen by selecting the *maxent* solution that has largest likelihood. The tests conducted show that the proposed most likely Rényi-*maxent* estimator has superior behaviour compared to the simpler relaxed-constraints *maxent*, being able to approximate the observed dataset while at the same time being compatible with description of a natural population. In particular, our numerical experiments show that our new algorithm leads to a distribution estimate that can predict observed grades well, and can thus be used to detect profiles with high risk of

decompression sickness.

Acknowledgements

We acknowledge the support of J. Hugon and A. Barbaud, from BF Systèmes (France), on biophysical modelling of hyperbaric diving, as well as providing the data used for the study.

References

- [1] J. Hugon, *Vers une modélisation biophysique de la décompression*, Ph.D. thesis, Aix Marseille 2, 2010.
 - [2] E. L. Kaplan and P. Meier, “Nonparametric estimation from incomplete observations,” *J. Amer. Statist. Assn.*, no. 53, pp. 457–481, 1958.
 - [3] B.W. Turnbull, “The empirical distribution function with arbitrarily grouped, censored and truncated data,” *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 290–295, 1976.
 - [4] R. Gentleman and A.C. Vandal, “Computational algorithms for censored-data problems using intersection graphs,” *Journal of Computational and Graphical Statistics*, vol. 10, no. 3, pp. 403–421, 2001.
 - [5] Y. Bennani, “Rapport de recherche i3s : Intersection graph for region-censored data,” 2013.
 - [6] E. T. Jaynes, “Information theory and statistical mechanics,” *Physics Reviews*, vol. 106, pp. 620–630, 1957.
 - [7] M. Dudik, *Maximum entropy density estimation and modeling geographic distributions of species*, Ph.D. thesis, Princeton University, Department of Computer Science, 2007.
 - [8] M. Dudik, M. Phillips, and R. Schapire, “Performance guarantees for regularised maximum entropy density estimation,” *Proc. 7th Annual Conf. on Comp. Learning Theory*, 2004.
 - [9] A. Renyi, “On measures of entropy and information,” *Fourth Berkeley Symposium on Mathematical Statistics and Probability*, pp. 547–561, 1961.
 - [10] X. Liu, *Nonparametric estimation with censored data: a discrete approach*, Ph.D. thesis, McGill University, Montreal, 2005.
 - [11] S.D. Silvey, D.H. Titterington, and B. Torsney, “An algorithm for optimal designs on a design space,” *Communications in Statistics-Theory and Methods*, vol. 7, no. 14, pp. 1379–1389, 1978.
 - [12] B. Torsney, “A moment inequality and monotonicity of an algorithm,” in *Semi-Infinite Programming and Applications*, pp. 249–260. Springer, 1983.
 - [13] R. Harman and L. Pronzato, “Improvements on removing nonoptimal support points in d-optimum design algorithms,” *Statistics & probability letters*, vol. 77, no. 1, pp. 90–94, 2007.
 - [14] B. Grechuk, A. Molyboha, and M. Zabarankin, “Maximum entropy principle with general deviation measures,” *Mathematics of Operations Research*, vol. 34, no. 2, pp. 445–467, 2009.