Bayesian inference in based-kernel regression: comparison of count data of condition factor of fish in pond systems

Tristan Senga Kiessé, Etienne Rivot, Christophe Jaeger, Joël Aubin

Journal Of Applied Statistics
February 2022, Volume 49 Issue 3 Pages 676-693
https://doi.org/10.1080/02664763.2020.1830953
https://archimer.ifremer.fr/doc/00810/92245/

Bayesian inference in based-kernel regression: comparison of count data of condition factor of fish in pond systems

Senga Kiessé T. [1, *], Rivot Etienne [2], Jaeger Christophe [1], Aubin Joel [1]

[1] SAS, INRAE, Institut Agro, Rennes, France
[2] ESE, Ecology and Ecosystems Health, Institut Agro, INRAE, Rennes, France

* Corresponding author : T. Senga Kiessé, email address : tristan.senga-kiesse@inra.fr

Abstract :

The discrete kernel-based regression approach generally provides pointwise estimates of count data that do not account for uncertainty about both parameters and resulting estimates. This work aims to provide probabilistic kernel estimates of count regression function by using Bayesian approach and then allows for a readily quantification of uncertainty. Bayesian approach enables to incorporate prior knowledge of parameters used in discrete kernel-based regression. An application was proposed on count data of condition factor of fish (K) provided from an experimental project that analyzed various pond management strategies. The probabilistic distribution of estimates were contrasted by discrete kernels, as a support to theoretical results on the performance of kernels. More practically, Bayesian credibility intervals ofK-estimates were evaluated to compare pond management strategies. Thus, similarities were found between performances of semi-intensive and coupled fishponds, with formulated feed, in comparison with extensive fishponds, without formulated feed. In particular, the fish development was less predictable in extensive fishpond, dependent on natural resources, than in the two other fishponds, supplied in formulated feed.

Keywords : Aquaculture, discrete kernel, fishpond, prior and posterior distributions, uncertainty analysis

1. Introduction

Let $(X_i, Y_i)_{i=1,2,\ldots,n} \subseteq \mathbb{N}^d \times \mathbb{R}$ be a sequence of i.i.d. random variables related by the unknown count regression function (c.r.f.) $m(\cdot) = \mathbb{E}(Y|X = \cdot)$ on a discrete support $\mathcal{S}$, including in the set $\mathbb{Z}$ of integers. Many works in the literature abound for parametric and nonparametric modeling of the c.r.f. m. Nonparametric regression methods were developed with the advantage to avoid any assumption on the form of the count data distribution to estimate. As such, the kernel-based estimation is among the most popular smoothing methods since it is globally efficient and easy to implement [19]. Discrete kernel regression approaches were adapted from continuous kernel ones to primarily estimate the function m at each target point $x \in \mathcal{S}$ [3]. The accuracy of discrete kernel regression estimators depends on choices of both a discrete kernel and a smoothing bandwidth. Classical kernel-based estimation methods only provides pointwise values

of kernel estimates, without taking into account any type of uncertainty and variability in data and parameters. Considering the lack of information or wrong information (uncertainty) on data and parameters or their heterogeneous nature (variability) will enable to assess the reliability of pointwise kernel-based estimation. Only few approaches exist in the literature to quantify uncertainty about kernel-based estimates. For instance, the imprecise functional estimation using maxitive kernels provides interval range of continuous kernel-based estimates [13]. One well-recognized approach to conduct uncertainty analysis is the Bayesian one, which is used to update the prior knowledge of parameters into their posterior distributions by using the information from the data. Bayesian (nonparametric) hierarchical models are also widely developed to accomodate dependence in multivariate and longitudinal data collected in biomedical studies [6] and aquatic sciences [15]; see also [2], for more details on bayesian nonparametric models. To our knowledge, the development of Bayesian approach for kernel-based estimators is restricted to bandwidth selection to obtain a better accuracy of the resulting estimates than the cross-validation procedure, in density kernel estimation [4, 21] and nonparametric kernel regression [20].

This work is concerned with a larger investigation of the possibility that the Bayesian approach provides to incorporate (informative or non-informative) prior knowledge of parameters used in nonparametric discrete kernel-based estimation. The main objective of this work is to provide a probabilistic kernel estimation of c.r.f. m at each target point $x \in \mathcal{S}$, rather than just a point estimate. The Bayesian analysis thus enables to provide 95% Bayesian credibility interval of kernel estimates of data. This work points out also other practical advantages of Bayesian analysis in kernel-based estimation for practitioners, such as to compare the performance of different kernel functions, as a support to the theoretical results on their comparison.

We illustrate the method through an example in the context of an aquaculture project that aimed to mobilize ecosystem services towards aqua-ecosystem (combining productive system and its natural trophic web) to design a new aquaculture system and to monitor performances. The study focused both on fish production supported by formulated feed supplied and the ability of pond planted with macrophytes to improve water quality. Fishpond performances were evaluated by collecting count data of the evolution of fish in experimental ponds with some indicators of fish development. We illustrate how combining kernel-based and Bayesian approaches improves the reliability of fishpond performances by providing probabilistic distribution of estimates of count data of a given indicator of fish development (K) and by improving estimates of K to evaluate fish growth performances.

2. Discrete nonparametric kernel regression

Consider the sequence $(X_i, Y_i)_{i=1,2,\ldots,n} \subseteq \mathbb{N}^d \times \mathbb{R}$ of i.i.d. random variables (r.v.) such that

$$Y_i = m(X_i) + \epsilon_i, \tag{1}$$

where m is the unknown c.r.f. on a discrete support $\mathcal{S}$, defined as being the conditional expectation of Y on X, and the ϵ_i's are assumed to have zero mean and finite variance. For a fixed point $x \in \mathcal{S}$ and a bandwidth parameter $h > 0$, the discrete nonparametric

regression estimator $\widehat{m}$ of m is defined as

$$\widehat{m}(x) = \sum_{i=1}^{n} \frac{Y_i Kern_{x,h}(X_i)}{\sum_{j=1}^{n} Kern_{x,h}(X_j)} =: \widehat{m}_{Kern,h}(x), \tag{2}$$

where $Kern_{x,h}$ is the *discrete associated kernel* and $h = h(n) > 0$ an arbitrary sequence of smoothing parameters that fullfills $\lim_{n\to\infty} h(n) = 0$ [3]. The following section provides a brief recall of the kernel method for estimating discrete functions on support $\mathcal{S} \subseteq \mathbb{Z}$ [11].

2.1. Generalities on discrete kernels

Definition. The discrete associated kernel $Kern_{x,h}$ in equation (2) is a probability mass function associated with a r.v. $\mathcal{K}_{x,h}$, i.e.

$$0 \leq Kern_{x,h}(y) = \Pr(\mathcal{K}_{x,h} = y) \leq 1 \quad \text{and} \quad \sum_{y\in\mathcal{S}_x} Kern_{x,h}(y) = 1,$$

on support $\mathcal{S}_x$ having its probability at target point x such that

$$Kern_{x,h}(x) \to D_x(x) = 1 \text{ as } h \to 0, \tag{3}$$

where D_x is the Dirac type kernel. The idea is that the discrete associated kernel must attribute the probability mass closest to one at target $x \in \mathcal{S}$, while having a smoothing parameter $h > 0$ to take into account the probability mass at points $y \in \mathcal{S}_x \setminus \{x\}$ in the neighbourhood of x. The following expressions of $Kern_{x,h}$'s expectation and variance result from Equation (3):

$$(\mathrm{E}_1) : \mathrm{E}(\mathcal{K}_{x,h}) = x + a(x,h) \quad \text{and} \quad (\mathrm{E}_2) : \mathrm{Var}(\mathcal{K}_{x,h}) = b(x,h),$$

where both $a(x,h)$ and $b(x,h)$ tend to 0 as h goes to 0 [17]. The kernel function and the bandwidth parameter are the two most important issues of the discrete associated kernel procedure, since they affect the accuracy of estimates and the degree of smoothing.

Classes of kernels. Two classes of discrete associated kernels were proposed, depending on whether they satisfy Equation (3) or not (Figure 1). For x in $\mathcal{S}$, p in the set $\mathbb{N}$ of nonnegative integers and $h > 0$, the first class of kernels contains discrete symmetric triangular kernels $Kern_{p;x,h}$ on support $\mathcal{S}_{p;x} = \{x, x \pm 1, ..., x \pm p\}$ such that

$$Kern_{p;x,h}(y) = \frac{(p+1)^h - |y-x|^h}{(2p+1)(p+1)^h - 2\sum_{k=0}^{p} k^h} =: T_{p;x,h}(y).$$

The mean of $Kern_{p;x,h}$ associated with the r.v. $\mathcal{T}_{p;x,h}$ is equal to x and its modal probability and variance can be expressed as follows:

$$(\mathrm{E}_3) : \Pr(\mathcal{T}_{p;x,h} = x) = 1 - 2h\mathrm{A}(p) + O(h^2) \quad \text{and} \quad (\mathrm{E}_4) : \mathrm{Var}(\mathcal{T}_{p;x,h}) = 2h\mathrm{V}(p) + O(h^2),$$

with $\mathrm{A}(p) = p\log(p+1) - \sum_{k=1}^{p}\log(k)$ and $\mathrm{V}(p) = \{p(2p^2+3p+1)/6\}\log(p+1) - \sum_{k=1}^{p} k^2\log(k)$ [17].

The second class of kernels contains standard asymmetric kernels constructed from usually discrete probability distributions (Table 1). They are similar to continuous asymmetric kernels such as beta [5] and inverse and reciprocal inverse Gaussian kernels [9]. The behaviour of discrete standard kernels are generalized through expressions $(\mathrm{E}_5) - (\mathrm{E}_6)$ of both their probability mass at target x and variance (Table 1).

Table 1. Summary of properties of discrete asymmetric standard kernels [17]

Discrete kernel	**Distribution**	**Support $\mathbf{S}_x$**	$\mathbf{E}(\mathcal{K}_{x,h})$	$\mathbf{Var}(\mathcal{K}_{x,h})$
Poisson	$\mathcal{P}(x+h)$, with $h>0$	$\mathbb{N}$	$x+h$	$x+h$
Binomial	$\mathcal{B}\left(x+1, \frac{x+h}{x+1}\right)$, with $h\in(0,1]$	$\{0,1,\ldots,x+1\}$	$x+h$	$(x+h)\left(\frac{1-h}{x+1}\right)$
Negative binomial	$\mathcal{NB}\left(x+1, \frac{x+1}{2x+1+h}\right)$, with $h>0$	$\mathbb{N}$	$x+h$	$(x+h)\left(1+\frac{x+h}{x+1}\right)$
Generalized properties				
Modal probability	$(\mathrm{E}_5) : \Pr(\mathcal{K}_{x,h}=x) = (1-h^2)Kern_{x;0}(x) + O(h^2)^{(*)}$			
Variance	$(\mathrm{E}_6) : \mathrm{Var}(\mathcal{K}_{x,h}) = V_{\mathcal{K}}(x) + hU_{\mathcal{K}}(x) + O(h^2)^{(**)}$			

$^{(*)}$ where $\sum_{y\in\mathcal{S}_x\setminus\{x\}}\Pr(\mathcal{K}_{x,h}=y) = 1-(1-h^2)Kern_{x;0}(x)+O(h^2)$ and $Kern_{x;0}(x)$ being the probability at mode x when $h\to 0$
$^{(**)}$ the terms $V_{\mathcal{K}}$ and $U_{\mathcal{K}}$ depend on the discrete kernel

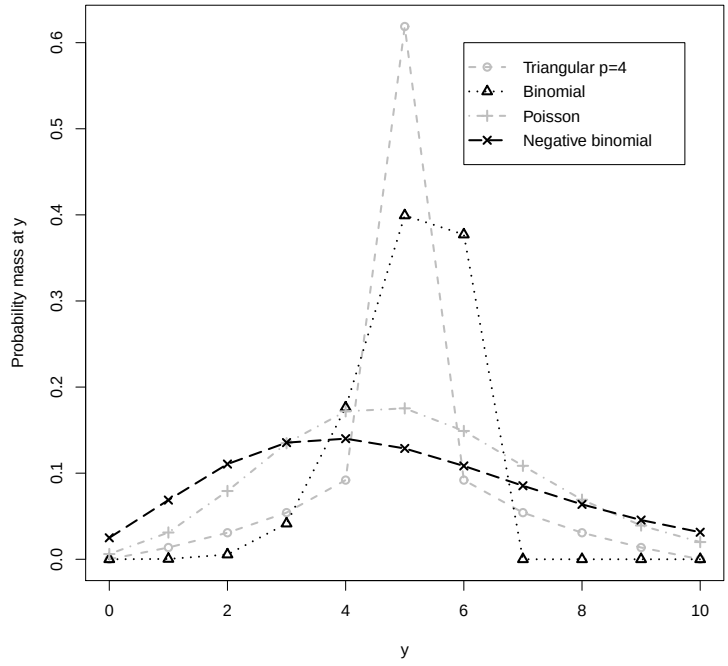

Figure 1. Discrete symmetric triangular and asymmetric standard kernels on support $S_x = \{1,2,\ldots,10\}$ with target point $x=5$ and bandwidth parameter $h=0.1$.

The main difference between those two classes of kernels is that the discrete standard kernels do not satisfy Equation (3), unlike the discrete symmetric triangular kernels; but, the discrete standard kernels may be more efficient for estimating small

or moderate sample sizes [11, 17]. Particularly, they can be less affected by boundary bias effect.

Remark 1. To find the optimal h-value, the cross-validation is one of the main method applied as an entirely data-driven procedure [14]. For discrete symmetric triangular kernel estimators, the cross-validation procedure minimizes h while $p \in \mathbb{N}$ remains fixed. In the literature, there is a gap in searching the potential pairs of parameters $(p, h) \in \mathbb{N} \setminus \{0\} \times (0, \infty)$, which provide good quality of adjustment for estimates.

2.2. Comparison of kernels

We now provide results on the comparison of discrete symmetric triangular kernels and standard asymmetric kernels, details not presented in existing references on discrete associated kernels.

Comparison of modal probabilities. First, we provide a result on the comparison of modal probability of discrete symmetric triangular kernels with respect to parameter $p \in \mathbb{N}$.

Proposition 2.1. *Consider any fixed $x \in \mathbb{N}$ and $h > 0$. Under the expression* (E_3)*, for $p \in \mathbb{N}$, the modal probability of discrete symmetric triangular satisfies:*

$$\Pr(\mathcal{T}_{p;x,h} = x) \geq \Pr(\mathcal{T}_{p+1;x,h} = x), \ \ as \ \ h \to 0.$$

Thus, more discrete symmetric triangular kernels have smaller p-values more they attribute an important probability mass at the target point as $h \to 0$.

Comparison of variances. Similar to the previous part on comparison of modal probabilities of discrete kernels, we provide a proposition on the variance of discrete symmetric triangular kernels with respect to $p \in \mathbb{N}$.

Proposition 2.2. *Consider any fixed $x \in \mathbb{N}$ and $h > 0$. Under the expression* (E_4)*, for $p \in \mathbb{N}$, the variance of discrete symmetric triangular satisfies:*

$$\mathrm{Var}(\mathcal{T}_{p;x,h}) \leq \mathrm{Var}(\mathcal{T}_{p+1;x,h}), \ \ as \ \ h \to 0.$$

Ultimately, next result compares discrete symmetric triangular kernels and standard asymmetric kernels via their modal probability and variance.

Proposition 2.3. *Consider $(x, p) \in \mathbb{N} \times \mathbb{N}$ and $h > 0$. As $h \to 0$, the modal probability and variance of discrete symmetric triangular ($T_{p;x,h}$) and standard kernels ($K_{x,h}$) satisfy:*

$$\Pr(\mathcal{T}_{p;x,h} = x) \geq \Pr(\mathcal{K}_{x,h} = x) \ \ and \ \ \mathrm{Var}(\mathcal{T}_{p;x,h}) \leq \mathrm{Var}(\mathcal{K}_{x,h}).$$

The ranking of modal probability and variance of discrete standard kernels were already provided in [17] such that, as $h \to 0$, $\Pr(\mathcal{B}_{x,h} = x) \geq \Pr(\mathcal{P}_{x,h} = x) \geq \Pr(\mathcal{NB}_{x,h} = x)$ and $\mathrm{Var}(\mathcal{B}_{x,h}) \leq \mathrm{Var}(\mathcal{P}_{x,h}) \leq \mathrm{Var}(\mathcal{NB}_{x,h})$.

Proofs of Propostitions 2.1 to 2.3 are postponed to Appendix.

2.3. Comparison of regression estimators

Previous results enable to contrast mean integrated squared error (MISE) of $\widehat{m}_{Kern,h}$ by the different discrete kernels, as $h \to 0$. To this end, we consider the following expressions of $\widehat{m}_{Kern,h}$'s bias and variance such that, as $h \to 0$,

$$\text{Bias}\{\widehat{m}_{Kern,h}(x)\} = \left\{ m^{(2)}(x) + 2m^{(1)}(x) \left(\frac{f^{(1)}}{f} \right)(x) \right\} \frac{Var(\mathcal{K}_{x,h})}{2} + O\left(\frac{1}{n}\right) + o(h),$$

and

$$\text{Var}\{\widehat{m}_{Kern,h}(x)\} = \frac{\text{Var}(Y|X=x)}{nf(x)} \{\Pr(\mathcal{K}_{x,h} = x)\}^2 + o\left(\frac{1}{n}\right),$$

where f is the probability mass function of r.v. X and, $f^{(1)}$, $m^{(1)}$ and $m^{(2)}$ are finite differences of f and m [3]. The expression of $\widehat{m}_{Kern,h}$'s variance traduces that as n increases, the variance term tends to 0 since it is penalized by the factor $1/n$. As $n \to \infty$ and $h \to 0$, the decrease in $\widehat{m}_{Kern,h}$'s variance term leads to considering mainly the influence of $\widehat{m}_{Kern,h}$'s bias term on approximate mean integrated squared error (AMISE) such that we obtain

$$\text{MISE}(\widehat{m}) = \sum_{x \in \mathbb{N}} \text{Bias}^2\{\widehat{m}(x)\} + \sum_{x \in \mathbb{N}} \text{Var}\{\widehat{m}(x)\} = \text{AMISE}(\widehat{m}) + o\left(h^2 + \frac{1}{n}\right).$$

It can be seen that the estimator $\widehat{m}$ using discrete symmetric triangular kernels is consistent, i.e $\text{AMISE}(\widehat{m}_{T,p,h})$ goes to 0 since $\text{Var}(\mathcal{T}_{p;x,h}) \to 0$ as $h \to 0$. At the opposite, the estimator $\widehat{m}$ using discrete standard kernels is non-consistent since $\text{Var}(\mathcal{K}_{x,h})$ does not tend to as $h \to 0$. However, this theoretical drawback of discrete standard kernel estimators is balanced by their interesting features in small/medium samples. The comparison of variances of discrete symmetric triangular (T) and standard kernels (K) in the Proposition 2.3 results in the following inequality about AMISE of regression estimators $\widehat{m}$, as $h \to 0$,

$$\text{AMISE}(\widehat{m}_{T,p,h}) \leq \text{AMISE}(\widehat{m}_{K,h}).$$

On the basis of Propositions 2.1 and 2.2, the function $p \mapsto \text{AMISE}(\widehat{m}_{T,p,h})$ in Equation (4) is increasing, for h fixed and $h \to 0$, such that discrete symmetric triangular kernels with $p = 1$ are practically recommended. The case $p = 0$ coincides with Dirac type kernel.

3. Bayesian inferences

Bayesian point of view. Let $(x_i, y_i), i = 1, 2, ..., n$ be i.i.d. bivariate observations and let ϵ_i be i.i.d. residuals following the Gaussian distribution $\mathcal{N}(0, \sigma^2)$ with mean zero and constant variance σ^2. Nonparametric kernel inferences on the c.r.f. m require an optimization step to find the point value of parameters involved in kernel estimation, such as h, p and σ^2. By contrast, optimization has nothing to do with inferences in the Bayesian context. The Bayesian treatment of a model only relies on the theory of probability. For instance, by simply considering the parameter h as random and

assigning a prior probability density function to it, the Bayesian approach enables us to express the posterior probability distribution of the parameters h and σ^2 given the observed data y_n.

Consider the following expression of the model in equation (1) given by:

$$Y_i - m(X_i) \sim \mathcal{N}(0, \sigma^2).$$

Let θ denote the vector of parameters involved in kernel-based estimation. The likelihood function for the set of observations $y_1, y_2, \ldots, y_n$ given parameter θ is

$$\mathrm{L}(y_1, y_2, \ldots, y_n|\theta) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\{y_i - \widehat{m}(x_i)\}^2\right]. \tag{4}$$

For the nonparametric estimator $\widehat{m}_{Kern,h}$ using binomial kernel with $h \in (0, 1]$ and Poisson kernel with $h > 0$, we assume a (non-informative) uniform prior on $(0, 1]$, which corresponds to a beta prior distribution $\mathcal{B}(\alpha, \beta)$ with positive parameters $\alpha = \beta = 1$.

For nonparametric estimator $\widehat{m}_{Kern,h}$ using discrete symmetric triangular kernels, continuous and discrete non-informative uniform priors are assumed for $h > 0$ and $p \in \mathbb{N} \setminus \{0\}$, respectively. Particularly, one objective is to investigate the variations of h according to p since until now these parameters are chosen independently from each other. Without loss of generality, Gamma(g_1, g_2) density, with $g_1 = g_2 = 0.001$, is considered for prior of precision $\tau = 1/\sigma^2$ since it is a natural conjugate distribution for the variance in a Gaussian likelihood model. The set of parameters considered is either the triplet (h, p, τ) or the pair (h, τ), when using either discrete symmetric triangular kernels or binomial and Poisson kernels, respectively.

The JAGS software. The estimation of the posterior p.d.f. is realized by means of the Markov Chain Monte Carlo (MCMC) algorithms [7, 8] using the JAGS software (http://mcmc-jags.sourceforge.net; release 3.4.0;[1]) through the rjags package (www.Rproject.org;[16]). JAGS only requires the declaration of the Bayesian model, which consists of the prior density of unknown variables, the sampling distributions relating the data and the variables and the deterministic equations linking the variables.

Details on expressions of prior distributions of parameters h and σ^2 and on MCMC sampling procedure within the framework of kernel estimation can be found in [18].

4. Application and results

4.1. Experimental design and datasets

Our case study is an experimental design to assess the impact of a formulated feed supplied on fishpond. We compared the fish growth performances between three different treatments.

Experimental data were collected in IMTA EFFECT Project (Integrated Multi Trophic Aquaculture for EFFiciency and Environmental ConservaTion). Different fish species were reared together in each fishpond: roach (*Rutilus rutilus*), common carp (*Cyprinus carpio*) and perch (*Perca fluviatilis*). The choice of these species was based on the complementarity of their trophic level and behavior in order to maximize use of

natural resources. Each experimental pond had an area of 500 m^2 and two replicates of each treatment were done (Figure 2):

- the first treatment represented extensive fishponds (N°508 and N°509) with a low fish density and fishes ate only natural biomass produced in the fishpond itself from nutrients available in the water.
- the second treatment represented semi-intensive fishponds (N°504 and N°505) with a fish density equivalent to the double of the previous treatment. Fish were fed with a commercial formulated feed on the basis of carp requirements (in quantity and quality). Formulated feed was issue from a unique batch to be sure that the same formula and the same raw material was used during all the experiment.
- the third treatment called "coupled" was composed of fishponds (N°502 and N°507), similar to the semi-intensive fishponds described above, associated to a pond of a similar size, planted with macrophytes. It was expected that plants play a role in purification of the water and as support for biodiversity. A water pump was set to circulate water between fishpond and planted pond (Figure 2).

Separated analyses were conducted in each pond, for the experimental design was not the same.

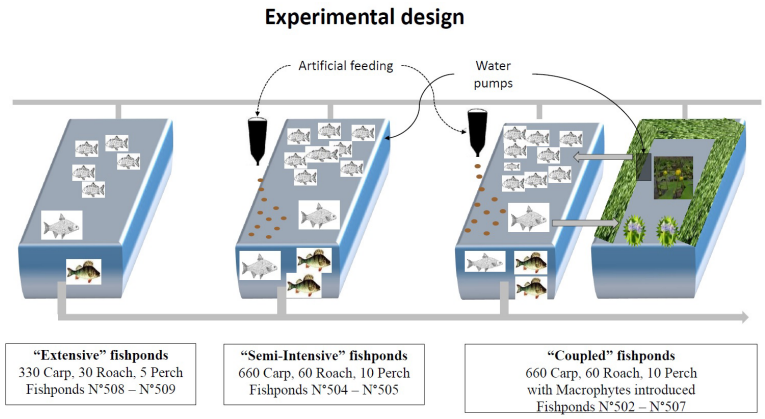

Figure 2. Experimental design of fishpond systems.

Throughout this study, we only consider carps. At the beginning of the experiment, the sample of carps stocked in each fishpond was assumed to be similar and representative from a total population of carps. The number of carps was observed and indicators of fish growth performances were measured, at the beginning and the end of the experiment (from March to December 2016). Fish weight (w in gram) and length to caudal fork (l in centimeter) were measured for each fish individually and the condition factor K was calculated as $K = w/l^3$ (without unit). The condition factor K is an indicator of body shape, and gives information on the nutritional status of the fish: higher the value of K is and more the fish has a round shape, meaning that resources in its environment were sufficient for its development. At the end of the experimental time, the count data distribution of the condition factor of carps was compared between two replicates within the same treatment. The values of K are classified in 4 categories ($\mathbf{1} : 1.5 < K \leq 2$, $\mathbf{2} : 2 < K \leq 2.5$, $\mathbf{3} : 2.5 < K \leq 3$, $\mathbf{4} : 3 < K \leq 3.5$). The K-classes were set to obtain an homogeneous repartition of number of fish similar to each fishpond. For each fishpond, the count dataset of carps (with sample size around

one hundred) was randomly selected at the end of the experiment (Table 2). Thus, we were concerned with the sequence of K-class (x_i) and observations of number of carps (y_i) in each class, $i = 1, \ldots, n_K$, $n_K = 4$, to estimate c.r.f. m.

Table 2. Samples of carps randomly selected in each experimental fishpond and replicate from the total population of carps, at the end of the experiment. The size N of each sample was around one hundred.

	Observed numbers of carps (y_i) per K-class in fishponds					
K-class (x_i)	Extensive		Semi-intensive		Coupled	
	N°508	N°509	N°504	N°505	N°502	N°507
1	6	2	12	0	1	1
2	68	54	21	16	31	20
3	31	45	71	89	72	83
4	0	1	18	9	3	9
Total N	105	102	122	114	107	113

4.2. Results

We used results obtained from the extensive fishpond N°509 to illustrate our approach. We presented descriptive statistics of estimated number of carp fish-species at harvesting. We provided descriptive statistics of discrete kernel estimates using Bayesian approach (Table 3). Performance of discrete kernel regression estimators $\widehat{m}_{Kern,h}$ of observations y_i of number of carp fish-species at harvesting was assessed using the root mean squared error (RMSE) such as

$$\text{RMSE} = \sqrt{\frac{1}{n_K}\sum_{i=1}^{n_K}\{\widehat{m}_{Kern,h}(x_i) - y_i)\},^2}$$

with x_i being the K-classes and $n_K = 4$ their total number.

4.2.1. Choices of parameters for discrete symmetric triangular kernels

The discrete kernel estimation using Bayesian approach provided joint and marginal posterior probability density functions (p.d.f.) of parameters (Figure 3). This was particularly informative when using discrete triangular symmetric kernels to point out the behavior of parameter $p \in \{1, 2, 3, 4, 5\}$ according to parameter $h_{Bayes} > 0$. Joint distributions showed that the parameter p was highly correlated to bandwidth parameter $h_{Bayes} > 0$ and precision parameter $\tau > 0$. The marginal posterior mode of p was at $p = 1$. A strong negative linear correlation was found between values of p and h (correlation coefficient $\rho = -0.99$). Hence, a similar quality of fit was obtained by different pairs of parameters values, e.g. the pairs ($p = 1$, h_{Bayes}-mean ≈ 0.55), ($p = 2$, h_{Bayes}-mean ≈ 0.45), ($p = 3$, h_{Bayes}-mean ≈ 0.40) (Figure 3). The correlations between parameters p and τ ($\rho = 0.59$), on one side, and h and τ ($\rho = -0.55$), on other side, appeared to be slightly weaker than between parameters p and h_{Bayes}.

In what follows, we applied the semiparametric estimation-procedure using discrete symmetric triangular kernel with $p = \{1, 2\}$ in comparison with asymmetric Poisson and binomial kernels.

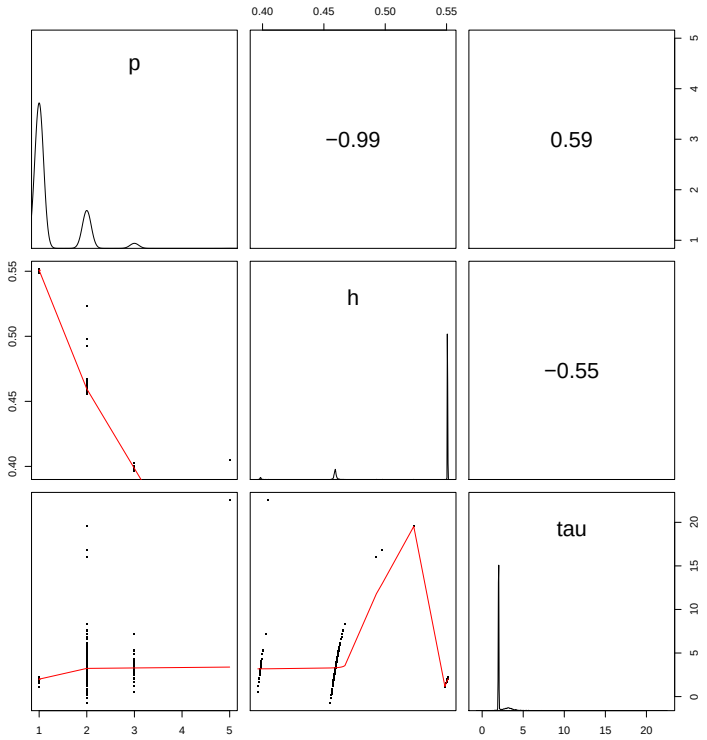

Figure 3. Joint and marginal posterior distributions of parameters $p \in \{1, 2, 3, 4, 5\}$, $h_{Bayes} > 0$ and τ (tau) > 0 involved in the discrete symmetric triangular kernel estimation using Bayesian approach, for count data distributions of carps per K-class in the extensive fishpond N°509.

4.2.2. Probabilistic kernel estimates of count data

As expected, incorporating Bayesian approach with kernel-based regression enabled to provide probabilistic kernel estimates of number of carps in each K-class, which is helpful to visualize the uncertainty and/or reliability linked to the use of a given discrete kernel (Figure 4). The p.d.f of Bayesian kernel-based estimates had a symmetric shape when using discrete triangular kernels (Figure 4). In addition, the p.d.f. of Bayesian kernel estimates had generally higher modal probabilities than when using the other discrete kernels, for each K-class. That indicated a high reliability of estimates with reduced count uncertainty. The discrete symmetric triangular kernel regression with $p = 1$ using Bayesian approach barely took into account any uncertainty, consequentially the means of the p.d.f. of kernel estimates was almost equal to pointwise observations for each K-class (Table 3). The discrete symmetric triangular kernel regression with $p = 2$ using Bayesian approach enabled to take into account more uncertainty than with $p = 1$. However, kernel-based estimates were affected by boundary bias, in particular with $p = 2$, in the sense of attributing probability mass at points outside the support $\mathcal{S} = \mathbb{N}$ of the count data to estimate. Thus, for K-class$= 1$ and $p = 2$, the discrete symmetric triangular kernel $Kern_{p=2;x=1,h}$ had support $\mathcal{S}_{p=2,x=1} = \{-1, 0, 1, 2, 3\} \nsubseteq \mathbb{N}$, which resulted in negative estimated frequencies when incorporating Bayesian approach (e.g., see the K-class$= 4$, Table 3). A solution was proposed through a modified version p_0 of the parameter $p \in \mathbb{N} \setminus \{0\}$ [12], but we did not illustrate this solution in this work. This did not change the sense of

our results.

The p.d.f. of Bayesian kernel estimates had an asymmetric shape when using Poisson and binomial kernels (Figure 4). By using these two discrete kernels, the obtained results obtained were generally less reliable (traduces by smaller values of probability density) than those obtained by using discrete symmetric triangular kernels. Poisson and binomial kernels were not affected by boundary bias effect, as it was defined above for discrete symmetric triangular kernels, since they have support $\mathcal{S}_x \subseteq \mathbb{N}$, for each K-class. A global bias reduction was proposed in the literature when using binomial kernel but we did not illustrate it in this study [10].

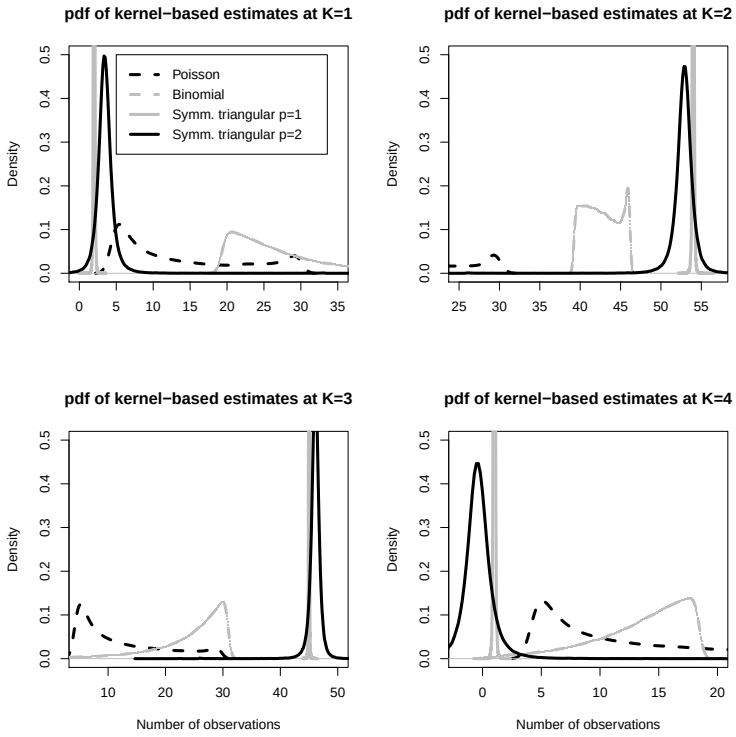

Figure 4. Marginal posterior probability density functions (p.d.f.) of discrete kernel-estimates of count data distributions of carps per K-class obtained in the extensive fishpond N°509

4.2.3. Bayesian credibility intervals

The RMSE criterion was calculated between observations and mean values of Bayesian kernel estimates for count data of fishpond N°509 (Table 3). According to the previous results, using discrete symmetric triangular kernels provided most accurate results in term of RMSE. Except for the discrete symmetric triangular kernel with $p = 1$, the calculated RMSE was smallest when using the discrete symmetric triangular kernel with $p = 2$. Then, the discrete kernel regression estimator incorportating Bayesian approach provided smaller RMSE when using the binomial kernel than the Poisson kernel.

The discrete kernel regression estimator incorportating Bayesian approach enabled to obtain variation ranges of the total number of observations $n = 102$, by summing up the upper and lower limits of the 95% Bayesian credibility interval $\mathrm{CI}_{95\%}$ of estimates. The estimated total number of observations $\widehat{n}$ stayed within the intervals

[70; 141] and [17; 115], using binomial and Poisson kernels, respectively. According to previous comments, the total number of observations barely had any uncertainty using discrete symmetric triangular kernel with $p = 1$ ($\widehat{n} \in [101; 102]$), while using discrete symmetric triangular kernel with $p = 2$ provided a variation range that was biased by negative estimates at K-class= 4 ($\widehat{n} \in [92; 112]$).

Table 3. Main statistics of Bayesian posterior estimates of count data of carps per K-class in the extensive fishpond N°509. Statistics of marginal posterior distribution of bandwidth parameter are in parenthesis.

		pond N°509			
K-class	No of	Poisson kern.		Symm. triang. kern. $p = 1$	
	obs.	mean ($h_{Bayes} = 19.5$)	$CI_{95\%}$ ($h_{Bayes} \in [0.50; 48.19]$)	mean ($h_{Bayes} = 0.5508$)	$CI_{95\%}$ ($h_{Bayes} \in [0.5504; 0.5511]$)
1	2	13.5	[4.5; 29.6]	2.0	[1.8; 2.1]
2	54	13.0	[4.5; 29.6]	54.0	[53.8; 54.1]
3	45	12.2	[4.4; 28.9]	45.0	[44.9; 45.1]
4	1	11.5	[4.3; 26.7]	1.0	[0.9; 1.1]
	RMSE	**27.4**		**0**	
K-class	No of	Binom. kern.		Symm. triang. kern. $p = 2$	
	obs.	mean ($h_{Bayes} = 0.29$)	$CI_{95\%}$ ($h_{Bayes} \in [0.01; 0.84]$)	mean ($h_{Bayes} = 0.459$)	$CI_{95\%}$ ($h_{Bayes} \in [0.457; 0.462]$)
1	2	27.3	[19.6; 46.3]	3.6	[1.2; 6.1]
2	54	42.6	[39.4; 46.0]	52.7	[50.1; 55.2]
3	45	24.4	[7.6; 30.7]	45.9	[44.1; 47.7]
4	1	13.5	[3.7; 18.2]	−0.2	[−2.8; 2.6]
	RMSE	**18.4**		**1.3**	

The discrete kernel regression estimator incorporating Bayesian approach was applied to the other fishponds. A similar behavior as in Figure 3 was generally observed for marginal distributions of parameters $p \in \{1, 2, 3, 4, 5\}$ and $h > 0$ involved in discrete symmetric triangular kernels. We were particularly interested in discrete symmetric triangular kernel with $p = 2$, since using Poisson and binomial kernels provided lower performances and discrete symmetric triangular kernel with $p = 1$ did not enable to take into account dispersion in observed counts. Table 4 presents means and 95th percentile of Bayesian posterior credibility intervals resulting from using discrete symmetric triangular kernel with $p = 2$. One could also observe that the boundary bias affected the estimated number of carp fish-species for K-class= 4 for fishponds N°508 - N°509 and for K-class= 1 for fishponds N°504 - N°505 and N°502 - N°507. The bias effect resulted in negative values of the lower limit of confidence intervals. Finally, the quality of adjustment varied according to observed count data of each fishpond: the smallest RMSE was obtained for fishpond N°509.

5. Interpretations

Incorporating uncertainty about parameters via Bayesian analysis enabled to evaluate the reliability linked to use of a given discrete kernel in nonparametric kernel regression. Discrete symmetric triangular kernel with $p = 2$ appeared as a compromise between discrete symmetric triangular kernel with $p = 1$ that did not allow to incorporate any uncertainty of parameter $h > 0$ and Poisson and binomial kernels that had a quality of adjustment not sufficiently reliable and accurate. By using discrete symmetric triangular kernel with $p = 2$, the 95% Bayesian CI generally contained observed count data of carps (Table 4). These CI were larger for the modal K-class: that concerned the K-class= 3 for pairs of fishponds N°502 - N°507 and N°504 - N°505, and the K-class= 2 for fishpond N°508. However, estimates of number of carps in K-class at boundary values of CI were affected by bias effect. The most accurate estimation were obtain for fishpond N°509, which also had the smallest RMSE in comparison with the other fishponds.

Table 4. Comparison of posterior means and 95% Bayesian credibility intervals (CI) of estimated count data distributions of carps per K-class in all fishponds, by aplying the discrete kernel regression estimator using symmetric triangular kernel with $p = 2$.

K-class	Extensive					
	pond N°508			**pond N°509**		
	Number of observations	Bayesian estimates		Number of observations	Bayesian estimates	
		mean	$CI_{95\%}$		mean	$CI_{95\%}$
1	6	14.1	[6.5; 34.5]	2	3.6	[1.2; 6.1]
2	68	60.0	[27.2; 73.8]	54	52.7	[50.1; 55.2]
3	31	34.4	[26.6; 37.6]	45	45.9	[44.1; 47.7]
4	0	−2.0	[−14.5; 31.8]	1	−0.2	[−2.8; 2.6]
RMSE		**6.0**			**1.3**	
	Semi-intensive					
	pond N°504			**pond N°505**		
	Number of observations	Bayesian estimates		Number of observations	Bayesian estimates	
		mean	$CI_{95\%}$		mean	$CI_{95\%}$
1	12	12.8	[−5.5; 40.6]	0	0.9	[−25.9; 43.7]
2	21	28.1	[25.5; 31.8]	16	25.8	[23.1; 30.1]
3	71	54.9	[23.7; 78.6]	89	66.1	[18.7; 100.3]
4	18	29.2	[22.2; 40.3]	9	25.4	[14.7; 43.7]
RMSE		**10.4**			**14.9**	
	Coupled					
	pond N°502			**pond N°507**		
	Number of observations	Bayesian estimates		Number of observations	Bayesian estimates	
		mean	$CI_{95\%}$		mean	$CI_{95\%}$
1	1	−0.9	[−15.3; 33.6]	1	1.3	[−21.9; 42.5]
2	31	35.2	[27.0; 39.0]	20	27.7	[27.0; 29.1]
3	72	62.1	[27.4; 78.6]	83	64.2	[19.7; 92.8]
4	3	12.4	[2.9; 34.9]	9	23.4	[13.1; 42.9]
RMSE		**7.2**			**12.4**	

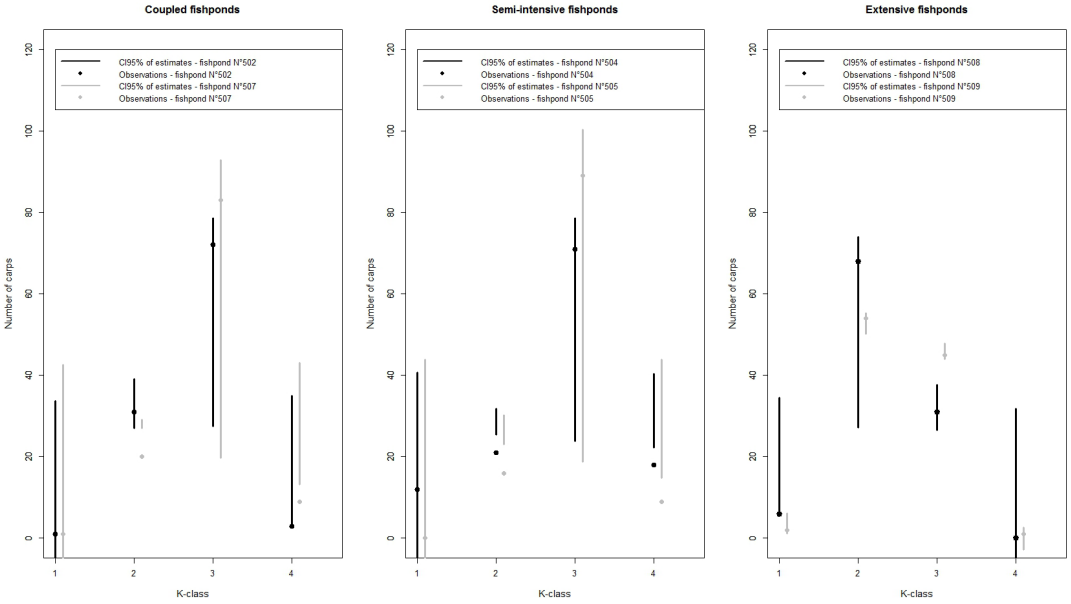

Figure 5. The 95% Bayesian credibility intervals ($CI_{95\%}$) of the kernel estimates of counts data (carps) per K-class in all fishponds. The discrete symmetric triangular kernel with $p = 2$ was used.

Similarities between replicates within the same fishpond (coupled and semi-intensive) were pointed out through overlapping of Bayesian CI of the kernel estimates of count data (Figure 5). For semi-intensive and coupled fishponds at harvesting, kernel-based estimates of the number of carp counted in each K-class were similar with a maximum amount of carps in the K-class= 3. Nevertheless, there was a significant amount of fish belonging to the three other K-classes, particularly for the fishpond N°504 belonging to the semi-intensive fishpond. Thus, fish fed with formulated feed had a relatively high value of K. This is in accordance with the fact that

K is a synthetic factor representing fish nutritional status.

For the extensive fishpond, at harvesting, count data distributions of carps for the two fishponds had a maximum amount of carps belonging to the K-class= 2. It is worth noting that no carp had its condition factor belonging to the K-class= 4 for fishpond N°508 and one carp had its condition factor belonging to this K-class for fishpond N°509. The fish development in extensive fishpond depends on natural resources available that makes it less predictable than the fish development in the two other fishponds supplied in formulated feed. That was traduced by differences between the count data of carps in K-classes of each replicate within the extensive fishponds. The 95% Bayesian CI of kernel-based estimates of count data in fishponds N°508 and N°509 did not overlap for K-class= $\{1, 3\}$ (Table 4 and Figure 5).

Note finally that the 95% Bayesian CI of kernel-based estimates of count data within extensive fishpond N°509 did not overlap with those obtained in semi-intensive and coupled fishponds, particularly for K-classes= $\{2, 4\}$ (Table 4).

6. Concluding remarks

Bayesian approach enables to incorporate prior knowledge of parameters used in nonparametric discrete kernel-based approach to provide posterior probability distribution of estimates instead of classical pointwise results. To use together Bayesian and kernel-based approaches is useful to assess the reliability of estimates by assessing how uncertainty in parameters estimates propagate to results. Moreover, the Bayesian approach would eventually allow to incorporate additional sources of information in the analysis through the use of informative priors. Depending on the research purpose, the judicious application of Bayesian approach offers a novel type of analysis in the portfolio of practitioners, that now has the opportunity to choose between the pure frequentist approach and the Bayesian approach for continuous and discrete kernel estimation.

Illustrative examples also contrasted the performance of different kernel functions, as a support to theoretical results on comparison of discrete kernels. Thus the discrete symmetric triangular kernel with parameter $p = 1$ provides an accurate estimation but barely integrates uncertainty, while the discrete symmetric triangular kernels with parameter $p = 2$ provides less accurate estimates but with the advantage to better integrate uncertainty. In addition, providing credibility intervals of pointwise kernel estimates enables to assess the performances of fishpond via hypotheses on the evolution of the number of fish from the beginning to the end of considered experiment. In our case study, the fish development was found to be less predictable in extensive fishpond, dependent on natural resources, than in the two other fishponds, supplied in formulated feed.

Research prospects would consist of performing a Bayesian nonparametrical hierarchical modeling with nonparametric kernel estimators, if data are collected under the same experimental design. A Bayesian risk analysis would be also conducted to investigate the probability of an undesirable event in discrete nonparametric kernel estimation of count data. For instance, one should be able to estimate the probability $P(X < x_{threshold})$ that the realization of a count variable X does not exceed a threshold value $x_{threshold}$.

Acknowledgments

The dataset was obtained in the IMTA-Effect project funded by Agence Nationale de la Recherche (ANR, France grant ANR-15-COFA-0001) in the framework of ERANET COFASP.

Appendix

Proof of Proposition 2.1

Let us express the difference

$$\Pr(\mathcal{T}_{p;x,h} = x) - \Pr(\mathcal{T}_{p+1;x,h} = x) = 2h\{A(p+1) - A(p)\}.$$

To establish this proof, we just show that the function $A(p)$ is increasing with respect to $p \in \mathbb{N}$. One has

$$A(p+1) - A(p) = (p+1)\log(p+2) - p\log(p+1) - \log(p+1) = (p+1)\log\big(\frac{p+2}{p+1}\big) > 0.$$

■

Proof of Proposition 2.2

Let us express the difference

$$\mathrm{Var}(\mathcal{T}_{p+1;x,h}) - \mathrm{Var}(\mathcal{T}_{p;x,h}) = 2h\{V(p+1) - V(p)\} + O(h^2).$$

To establish this proof, we show that the function $p \mapsto V(p)$ is increasing. One has

$$\begin{aligned} V(p+1) - V(p) &= \frac{(p+1)\{2(p+1)^2 + 3(p+1) + 1\}}{2}\log(p+2) \\ &\quad -\frac{p(2p^2+3p+1)}{2}\log(p+1) - 3(p+1)^2\log(p+1) \\ &= \frac{(p+1)(2p^2+7p+6)}{2}\log(p+2) - \frac{(p+1)(2p^2+7p+6)}{2}\log(p+1) \\ &= \frac{(p+1)(2p^2+7p+6)}{2}\log\left(\frac{p+2}{p+1}\right) > 0. \end{aligned}$$

■

Proof of Proposition 2.3

Comparison of modal probabilities We express the ratio of modal probabilities of kernels in (E_3) and (E_5), by using a Taylor expansion as $h \to 0$, such that:

$$r_1(x) = \frac{T_{p;x,h}(x)}{K_{x,h}(x)} \approx \frac{1-2hA(p)}{(1-h^2)K(x;0)(x)} = \frac{1-2hA(p)}{K(x;0)(x)} + O(h^2),$$

where $r_1(x) \to \{K(x;0)(x)\}^{-1}$ and, particularly, $r_1(0) \to 1$ as $h \to 0$. In addition, we show hereafter that the function $x \in \mathbb{N} \mapsto r_1(x)$ is increasing, for each discrete standard kernel.

For the Poisson kernel, by using Taylor expansion as $x \to \infty$, we get

$$\ln\left\{\frac{r_1(x+1)}{r_1(x)}\right\} \approx \ln\left\{\frac{K(x;0)(x)}{K(x+1;0)(x+1)}\right\} = 1-x\ln\left(1+\frac{1}{x}\right) = 1-x\left(\frac{1}{x}-\frac{1}{2x^2}\right) > 0.$$

Hence, we have both $r_1(x+1) \geq r_1(x)$ and $r_1(0) \to 1$ as $h \to 0$. Thus, we get $\Pr(\mathcal{T}_{p;x,h} = x) \geq \Pr(\mathcal{P}_{x,h} = x)$ as $h \to 0$, with $\mathcal{P}_{x,h}$ being the r.v. associated with the Poisson kernel.

For binomial kernel, we get

$$\begin{aligned}
\ln\left\{\frac{r_1(x+1)}{r_1(x)}\right\} &\approx \ln\left\{\left(\frac{x}{x+1}\right)^x\left(\frac{x+2}{x+1}\right)^{(x+1)}\right\} \\
&= (x+1)\left(\frac{1}{x+1}-\frac{1}{2(x+1)^2}\right) - x\left(\frac{1}{x}-\frac{1}{2x^2}\right) = \frac{1}{2x(x+1)} > 0.
\end{aligned}$$

That results in $\Pr(\mathcal{T}_{p;x,h} = x) \geq \Pr(\mathcal{B}_{x,h} = x)$ as $h \to 0$, with $\mathcal{B}_{x,h}$ being the r.v. associated with the binomial kernel.

Finally, denoting $\mathcal{NB}_{x,h}$ the r.v. associated with the negative binomial kernel, we also obtain $\Pr(\mathcal{T}_{p;x,h} = x) \geq \Pr(\mathcal{NB}_{x,h} = x)$ as $h \to 0$ since

$$\begin{aligned}
\ln\left\{\frac{r_1(x+1)}{r_1(x)}\right\} &\approx \ln\left\{\left(\frac{x+1}{2x+1}\right)^{x+1}\left(\frac{2x+3}{x+2}\right)^{x+2}\right\} \\
&= (x+2)\ln\left(1-\frac{1}{2(x+2)}\right) - (x+1)\ln\left(1-\frac{1}{2(x+1)}\right) + \ln(2) \\
&\approx \ln(2) > 0.
\end{aligned}$$

Comparison of variances We study the ratio

$$r_2(x) = \frac{\mathrm{Var}(\mathcal{T}_{p;x,h})}{\mathrm{Var}(\mathcal{K}_{x,h})}.$$

As the variance of discrete symmetric triangular kernel at the numerator of r_2 is independent on x, for each discrete standard asymmetric kernel, we get

$$\frac{r_2(x+1)}{r_2(x)} = \frac{\mathrm{Var}(\mathcal{K}_{x,h})}{\mathrm{Var}(\mathcal{K}_{x+1,h})}.$$

Hereafter, we show that the function $x \in \mathbb{N} \setminus \{0\} \mapsto r_2(x)$ is decreasing and less than 1 by using the expression of $T_{p;x,h}$'s variance in (E_4).

For the Poisson kernel, one has directly $r_2(x+1)/r_2(x) = (x+h)/(x+1+h) < 1$. Moreover, for $0 < h < 1$ and $h \to 0$, we get $r_2(1) = 2h/(1+h)V(p) + O(h^2) < V(p) + O(h^2) \leq 1$ since $V(1) = \log(2) < 1$, $p \in \mathbb{N}$. Thus, we have shown that $r_2(x) \leq 1$ and $\mathrm{Var}(\mathcal{T}_{p;x,h}) \leq \mathrm{Var}(\mathcal{P}_{x,h})$, as $h \to 0$.

For the binomial kernel, by using a Taylor expansion as $x \to \infty$, we successively obtain

$$\begin{aligned} \ln\left\{\frac{r_2(x+1)}{r_2(x)}\right\} = \ln\left(\frac{x+h}{x+1+h} \times \frac{x+2}{x+1}\right) &= \ln\left(1+\frac{1}{x+1}\right) - \ln\left(1+\frac{1}{x+h}\right) \\ &\approx \frac{1}{x+1} - \frac{1}{x+h} < 0, \end{aligned}$$

resulting in $r_2(x+1)/r_2(x) < 1$. For any $p \in \mathbb{N}$, we then get

$$r_2(1) = \frac{4h}{1-h^2}V(p) + O(h^2) < V(p) + O(h^2),\ 0 < h < \sqrt{5} - 2,$$

with $V(p) < 1$ at $p = 1$. Hence, we get $\mathrm{Var}(\mathcal{T}_{p;x,h}) \leq \mathrm{Var}(\mathcal{B}_{x,h})$, as $h \to 0$.

Similarly, for the negative binomial kernel, we show that $\mathrm{Var}(\mathcal{T}_{p;x,h}) \leq \mathrm{Var}(\mathcal{NB}_{x,h})$, as $h \to 0$. To this end, as $x \to \infty$, we show that

$$\begin{aligned} \ln\left\{\frac{r_2(x+1)}{r_2(x)}\right\} &= \ln\left\{\frac{(x+h)\left(1+\frac{x+h}{x+1}\right)}{(x+1+h)\left(1+\frac{x+1+h}{x+2}\right)}\right\} \\ &= \ln\left(1+\frac{1}{x+1}\right) - \ln\left(1+\frac{1}{x+h}\right) - \ln\left(1+\frac{1}{2x+1+h}\right) - \ln(2) \\ &\approx \frac{1}{x+1} - \frac{1}{x+h} - \frac{1}{2x+1+h} - \ln(2) \leq 0,\ \text{for } 0 < h \leq 1. \end{aligned}$$

Moreover, for any $p \in \mathbb{N}$, we obtain

$$r_2(1) = \frac{4h}{(1+h)(3+h)}V(p) + O(h^2) < 1 + O(h^2).$$

■

References

[1] *JAGS: a program for analysis of Bayesian graphical models using Gibbs sampling*, in *Proceedings of the 3rd International Workshop on Distributed Statistical Computing (DSC 2003)*, 20-22 march.

[2] *Bayesian Nonparametrics*, Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press, 2010.

[3] B. Abdous, C.C. Kokonendji, and T. Senga Kiessé, *On semiparametric regression for count explanatory variables*, Journal of Statistical Planning and Inference 142 (2012), pp. 1537 – 1548.

[4] M.J. Brewer, *A bayesian model for local smoothing in kernel density estimation*, Statistics and Computing 10 (2000), pp. 299–309.
[5] S.X. Chen, *A beta kernel estimator for density functions*, Computational Statistics and Data Analysis 31 (2000), pp. 131–145.
[6] D.B. Dunson, *Bayesian nonparametric hierarchical modeling*, Biometrical Journal 51 (2009), pp. 273–284.
[7] A. Gelman, J.B. Carlin, H. Stern, and D.B. Rubin, *Bayesian Data Analysis*, Chapman and Hall, 1995.
[8] W.R. Gilks, S. Richardson, and D. Spiegelhalter, *Markov Chain Monte Carlo in Practice*, Chapman and Hall, 1996.
[9] M. Hagmann and M. Scaillet, *Local multiplicative bias correction for asymmetric kernel density estimators*, Journal of Econometrics 141 (2007), pp. 213–249.
[10] L. Harfouche, S. Adjabi, N. Zougab, and B. Funke, *Multiplicative bias correction for discrete kernels*, Statistical Methods & Applications (2017). Available at https://doi.org/10.1007/s10260-017-0395-x.
[11] C.C. Kokonendji and T. Senga Kiessé, *Discrete associated kernel method and extensions*, Statistical Methodology 8 (2011), pp. 497–516.
[12] C.C. Kokonendji, T. Senga Kiessé, and S.S. Zocchi, *Discrete triangular distributions and non-parametric estimation for probability mass function*, Journal of Nonparametric Statistics 19 (2007), pp. 241–254.
[13] K. Loquin and O. Strauss, *Imprecise Functional Estimation: The Cumulative Distribution Case*, in *Soft Methods for Handling Variability and Imprecision*, D. Dubois, M.A. Lubiano, H. Prade, M.Á. Gil, P. Grzegorzewski, and O. Hryniewicz, eds., Springer Berlin Heidelberg, Berlin, Heidelberg (2008), pp. 175–182.
[14] J.S. Marron, *A comparison of cross-validation techniques in density estimation*, The Annals of Statistics 15 (1987), pp. 152–162.
[15] E. Parent and E. Rivot, *Introduction to Hierarchical Bayesian Modelling for Ecological Data*, Chapman/CRC, 2013.
[16] M. Plummer, *rjags: Bayesian graphical models using mcmc* (2016). Available at http://CRAN.R-project.org/package=rjags, R package version 4-6.
[17] T. Senga Kiessé, *On finite sample properties of nonparametric discrete asymmetric kernel estimators*, Statistics 51 (2017), pp. 1046–1060.
[18] T. Senga Kiessé, N. Zougab, and C.C. Kokonendji, *Bayesian estimation of bandwidth in semiparametric kernel estimation of unknown probability mass and regression functions of count data*, Computational Statistics 31 (2016), p. 189.
[19] A.B. Tsybakov, *Introduction à l'estimation non-paramétrique*, Springer Paris, 2004.
[20] X. Zhang, M.L. King, and H.L. Shang, *Bayesian bandwidth selection for a nonparametric regression model with mixed types of regressors*, Econometrics 4 (2016).
[21] N. Zougab, S. Adjabi, and C.C. Kokonendji, *Comparison study to bandwidth selection in binomial kernel estimation using bayesian approaches*, Journal of Statistical Theory and Practice 10 (2016), pp. 133–153.