

Multiple monotonicity of discrete distributions: The case of the Poisson model

Fadoua Balabdaoui

*Seminar für Statistik, ETH Zürich,
8092, Zürich, Schweiz*
e-mail: fadoua.balabdaoui@stat.math.ethz.ch

Gabriella de Fournas-Labrosse

*Seminar für Statistik, ETH Zürich,
8092, Zürich, Schweiz*
e-mail: gdefourn@student.ethz.ch

and

Jade Giguelay

AgroParisTech, 16 rue Claude Bernard, F-75231, Paris Cedex 05
e-mail: jade.giguelay@agroparistech.fr

Abstract: Shape constrained estimation in discrete settings has received increasing attention in statistics. Among the most important shape constrained models is multiple monotonicity, including k -monotonicity, for a given integer $k \in [1, \infty)$, and complete monotonicity. Multiple monotonicity provides a nice generalization of monotonicity and convexity and has been successfully used in applications related to estimation of species richness. Although fully nonparametric, it is of great interest to determine some of the well-known parametric distributions which belong to this model. Among the most important examples are the family of Poisson distributions and mixtures thereof. In Giguelay (2017) k -monotonicity of Poisson distributions was connected to the roots of a certain polynomial, but a typographical error occurred while writing its expression. In this note, we correct that typographical error and give a detailed proof that a Poisson distribution with rate $\lambda \in [0, \infty)$ is k -monotone if and only if $\lambda \leq \lambda_k$, where λ_k is the smallest zero of the k -th degree Laguerre polynomial $L_k(x) = \sum_{j=0}^k (-1)^j \binom{k}{j} x^j / j!$, $x \geq 0$. This result yields the sufficient condition that a mixture of Poisson distributions is k -monotone if the support of the mixing distribution is included in $[0, \lambda_k]$. Furthermore, we show that the only complete monotone Poisson distribution is the Dirac distribution at 0.

Keywords and phrases: completely monotone, discrete distribution, k -monotone, Laguerre polynomials, Poisson model, species richness.

Received March 2019.

1. Background and motivation

Let p be a given function defined on \mathbb{N} . The r -th differential operator $\Delta^{(r)}$ is defined as follows: $\Delta^{(0)}p(i) = p(i)$ and $\Delta^{(r)}p(i) = \Delta\Delta^{(r-1)}p(i)$ for all $i \in \mathbb{N}$, with Δ the usual difference operator. Thus, $\Delta^{(r)}$ is recursively defined with $\Delta^{(0)}$ being the identity. For example, $\Delta^{(1)}p(i) = p(i+1) - p(i)$ gives the slopes of p , whereas $\Delta^{(2)}p(i) = p(i+2) + p(i) - 2p(i+1)$ measures its curvature at $i \in \mathbb{N}$. Furthermore, it can be easily shown that

$$(-1)^r \Delta^{(r)}p(i) = \sum_{j=0}^r (-1)^j \binom{r}{j} p(i+j)$$

for all $r, i \in \mathbb{N}$. For a given integer $k \geq 1$ p is said to be k -monotone if $(-1)^k \Delta^{(k)}p(i) \geq 0$ for all $i \in \mathbb{N}$. Therefore, k -monotonicity generalizes monotonicity ($k = 1$) and convexity ($k = 2$).

Research work on k -monotone sequences can be traced back to the paper of Knopp (1925). From the definition of k -monotonicity, it is clear that this property is preserved under an arbitrary linear combination with positive coefficients. In Lefèvre and Loisel (2013) it is shown that p is a k -monotone probability mass function (pmf) if and only if there exists a sequence of probabilities $(\pi_j^k)_{j \in \mathbb{N}}$ such that p can be written as

$$p(i) = \sum_{j \geq 0} \pi_j^k Q_j^k(i) \tag{1.1}$$

where Q_j^k is the finitely supported k -monotone spline given by

$$Q_j^k(i) = \frac{\binom{j-i+k-1}{k-1}}{\binom{j+k}{k}} \mathbb{I}_{0 \leq i \leq j}, \tag{1.2}$$

see the proof of their Proposition 2.3. Above, the notation \mathbb{I} is used for the indicator function. Thus, k -monotonicity defines a mixture model, and estimation of an unknown pmf p in this class is then equivalent to estimation of the mixing sequence $(\pi_j^k)_{j \in \mathbb{N}}$. Other interesting results including alternative stochastic representations can be found in Lefèvre and Loisel (2013). Given n independent random variables X_1, \dots, X_n assumed to be i.i.d. from an unknown k -monotone pmf p , Giguélay (2017) considered unconstrained and constrained least squares estimation. The resulting least square estimators (LSE) are found to have a convergence rate of order \sqrt{n} . Although no limit distribution was derived, this gives a nice generalization for the rate of convergence of the Gnenander estimator of Jankowski and Wellner (2009) and the convex least squares estimator of Durot et al. (2013). See also Balabdaoui et al. (2017). According to the numerical results obtained by Giguélay (2017), it is more advantageous to fit the k -monotone LSEs than the empirical estimator even for moderate sample sizes n , provided that the model is well-specified. Note that k -monotone pmfs with large values of k exhibit a deeper hollow than those with smaller values. For an

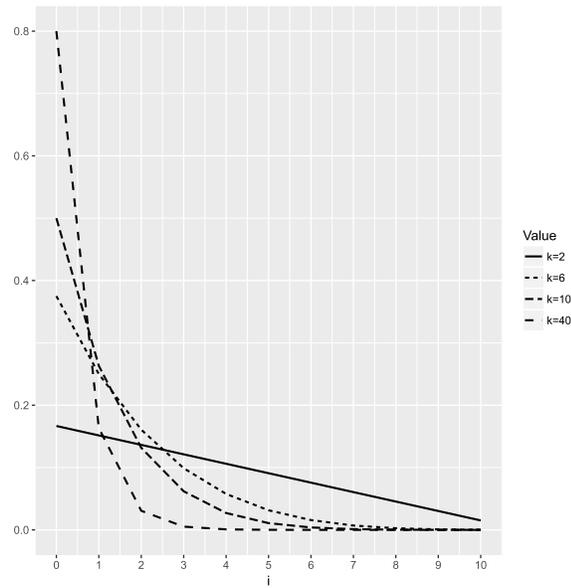


FIG 1. Plots of the splines Q_{10}^k for $k \in \{2, 6, 10, 40\}$.

illustration of this fact see Figure 1 where plots of the k -monotone spline Q_{10}^k are shown for different values of k . The same feature can be observed in the plots of the probability mass functions of the k -monotone Poisson distributions for $k \in \{2, 4, 6, 8, 10\}$; see Figure 3.

Extending the work of Durot et al. (2015), Giguelay and Huet (2018) showed that k -monotone estimation can be applied to estimating species richness; i.e., the total (unknown) number of species of a given population. If p is the pmf of abundances, then assuming that p is k -monotone comes with the interpretation that π_j^k in the mixture model (1.1) is the probability that a species belongs to the j -th sub-population whose abundance is distributed according to the pmf Q_j^k . Since $Q_0^k = \delta_0$, the Dirac at 0, we necessarily have $\pi_0^k = 0$ since π_0^k is then the probability that a species belongs to the sub-population whose individuals cannot be observed!

Chee and Wang (2016) also considered an estimation approach based on k -monotonicity. Their method differs however in that they considered mixtures of the k -monotone splines $s_\theta(i) = (\theta - i)_+^{k-1} / \sum_{j \leq \theta} j^{k-1}$, $i \in \mathbb{N}$, $\theta \in [1, \infty)$ instead of Q_j^k , $j \in \mathbb{N}$. Also, the authors considered maximum likelihood instead of least squares estimation. In applications related to species richness, the authors found that their method gives better results than those obtained with mixtures of Poisson distributions, a model that has become popular for estimation of species richness since the work of Norris and Pollock (1998). Interestingly, it can be shown that if a mixture of Poisson distributions has a mixing distribution with a restricted support then this mixture is k -monotone; see Corollary 2.7 below.

We would like to mention that for a density f with respect to Lebesgue measure which is defined on $[0, \infty)$, k -monotonicity can be also defined in terms of higher derivatives: $(-1)^{k-2} f^{(k-2)}$ has to be non-negative and convex on $[0, \infty)$. Characterization of k -monotonicity through scale mixtures of Beta(1, k) goes back to Williamson (1956). See also Fejér (1936), Feller (1939), Lefèvre and Loisel (2013), Balabdaoui and Wellner (2007) and Balabdaoui and Wellner (2010) where estimation via maximum likelihood and least squares methods was considered.

The other multiple monotonicity model is that of complete monotonicity. As $k \rightarrow \infty$ the classes of k -monotone pmfs converge, in the sense of intersection, to the class of completely monotone pmfs, formally, p defined on \mathbb{N} is completely monotone if it is k -monotone for all $k \geq 1$. It has been established in Steutel (1969) that p is a completely monotone pmf on \mathbb{N} if and only if it is a mixture of geometric distributions. In the recent work of Balabdaoui and de Fournas-Labrosse (2019) nonparametric least squares estimation in this class was considered. There, it was shown that the complete monotone LSE converges at the rate \sqrt{n} , to a limit distribution which is characterized as the unique solution of fully described limiting quadratic criterion.

When considering a nonparametric class of distributions such as the multiple monotone class, it is of much interest to identify some of the well-known parametric models which are elements thereof. This can be very helpful when looking for concrete examples in simulation settings. This also gives reference distributions representing the associated model. In the log-concave model, it is known that Binomial, Poisson, Skellam (convolution of two Poisson distributions) and geometric distributions are all log-concave on their respective supports, without any restriction on their parameters. As k gets larger the constraint of k -monotonicity becomes much more difficult to check. As opposed to log-concavity of a pmf p where only the sign of $\Delta^{(2)} \log p(i)$ for i in the support of p needs to be determined, it follows from Property 2 of Giguélay (2017) that k -monotonicity of p actually implies that the k conditions $(-1)^r \Delta^{(r)} p(i) \geq 0$, $r = 1, \dots, k$ should be satisfied for all $i \in \mathbb{N}$. Obviously, the splines Q_j^k , $j \in \mathbb{N}$ which form the basis of the class of k -monotone pmfs are themselves k -monotone. However, they are not as meaningful as Poisson distributions and their mixtures. In Giguélay and Huet (2018), the problem of testing k -monotonicity was considered with the goal of estimating the parameter k based on a random sample from a given multiple monotone pmf. To evaluate the power of these nested testing procedures a simulation study was designed using the Poisson distribution as the truth. This necessitates knowing the range of the rate λ of the Poisson distribution such that the latter is k -monotone but not $(k+1)$ -monotone. Indeed, if we write \mathcal{M}_k the class of k -monotone pmfs, then the procedures in Giguélay and Huet (2018) are designed for testing

$$H_0^k : p \in \mathcal{M}_{k+1} \quad \text{versus} \quad H_1^k : p \in \mathcal{M}_k \setminus \mathcal{M}_{k+1}.$$

It can be easily shown that a Poisson pmf is 1-monotone; i.e., monotone non-increasing, if and only if its rate parameter belongs to $[0, 1]$. With some addi-

tional effort, one can show that if the rate belongs to $[0, 2 - \sqrt{2}]$, then the corresponding Poisson distribution is 2-monotone; i.e., convex (and non-increasing). Our first goal in this paper is to give a precise characterization of k -monotonicity of such an important discrete distribution, thereby correcting the typographical error occurring in the condition given in Property 8 of Giguelay (2017). For the interested reader, we draw attention to the fact that values of the roots given by Giguelay (2017) in page 14 are correct, but they are actually not the roots of the polynomial considered in Property 8. Rather, they are the roots of certain Laguerre polynomials. The second main goal of this paper is to make the link between the roots of these polynomials and the range of the rate parameter of a k -monotone Poisson distribution.

The paper is organized as follows. In the next section we recall the k -monotonicity condition for the Poisson model and then make the connection between this property and the (α, k) -Laguerre polynomials where $\alpha \in \mathbb{N}$ and $k \in \mathbb{N} \setminus \{0\}$. Our main result given in Proposition 2.6 shows that a necessary and sufficient condition for k -monotonicity of a Poisson distribution with rate λ is that $\lambda \leq \lambda_k$ with λ_k the smallest zero of the $(0, k)$ -Laguerre polynomial. A byproduct of this characterization is a sufficient condition of k -monotonicity of a mixture of Poisson distributions. Using monotonicity of λ_k in k , we show that the result further implies that the sole complete monotone Poisson distribution is the Dirac at 0. In Section 3 we gather some concluding remarks and remaining questions. A short appendix is added for an intermediate result and a proof.

2. Characterization of k -monotonicity of Poisson distribution

Let $\lambda \in [0, \infty)$ and $\mathcal{P}(\lambda)$ denote the Poisson distribution with rate λ . For $\lambda = 0$, $\mathcal{P}(\lambda)$ coincides with δ_0 , the Dirac distribution at 0. Let p_λ denote the probability mass function of $\mathcal{P}(\lambda)$; i.e., for $i \in \mathbb{N}$

$$p_\lambda(i) = \frac{\lambda^i e^{-\lambda}}{i!}.$$

Thus, saying that p_λ is k -monotone is equivalent to the statement that $(-1)^k \Delta^{(k)} p_\lambda(i) \geq 0$ for all $i \in \mathbb{N}$, with $\Delta^{(k)}$ is the k -th differential operator already defined above. In other words, the parameter λ needs to satisfy

$$\sum_{j=0}^k (-1)^j \binom{k}{j} \frac{\lambda^{i+j}}{(i+j)!} \geq 0 \quad (2.3)$$

for all $i \in \mathbb{N}$.

2.1. Link to Laguerre polynomials

We now make the connection between k -monotonicity of $\mathcal{P}(\lambda)$ and certain Laguerre polynomials. Before doing so we would like to first recall their definition.

Given a real number $\alpha \in (-1, \infty)$ and an integer $k \geq 1$, consider the ordinary differential equation

$$xy''(x) + (\alpha + 1 - x)y'(x) + ky(x) = 0,$$

also referred to as the Laguerre's equation. Its polynomial solution is the well-known generalized (α, k) -Laguerre polynomial, $L_k^{(\alpha)}$, given by

$$L_k^{(\alpha)}(x) = \sum_{j=0}^k (-1)^j \binom{k+\alpha}{k-j} \frac{x^j}{j!}$$

where

$$\binom{k+\alpha}{k-j} = \frac{(k+\alpha)(k+\alpha-1)\dots(\alpha+j+1)}{(k-j)!}$$

is the generalized binomial coefficient. It is known that $L_k^{(\alpha)}$ admits k simple zeros, which all belong to the oscillatory region $(0, 4k + 2\alpha + 2)$, see for example Gatteschi (2002). Several properties of the zeros of $L_k^{(\alpha)}$ have been extensively studied in the literature. This includes lower and upper bounds, asymptotic behavior when $k \rightarrow \infty$ and α is fixed, monotonicity in certain functions of α and Stieltjes interlacing, to mention only a few. Investigation of these properties and related theorems can be found for example in Gatteschi (2002), Krasikov (2003), Natalini and Palumbo (2003), Dimitrov and Rafaeli (2009) and Driver et al. (2011). For our purpose, we will mainly need three results which we will present next. In the sequel, $\lambda_{k,1}^{(\alpha)} < \dots < \lambda_{k,k}^{(\alpha)}$ denote the k simple zeros of $L_k^{(\alpha)}$, rearranged in increasing order. The following theorem is due to Natalini and Palumbo (2003).

Theorem 2.1. *Natalini and Palumbo (2003)*

For any integer $k \geq 1$ and $r \in \{1, \dots, k\}$, the function

$$\alpha \mapsto \frac{\lambda_{k,r}^{(\alpha)}}{\sqrt{\alpha + 2k + 1}}$$

is increasing as α increases in $(-1, \infty)$.

As already noted by Natalini and Palumbo (2003) Theorem 2.1 implies that $\lambda_{k,r}^{(\alpha)}$ increases in α . This property will be used in the proof of Proposition 2.6 below. For illustration, we have plotted in Figure 2 the $(\alpha, 3)$ -Laguerre polynomials, $L_3^{(\alpha)}$, for $\alpha \in \{0, 2, 4, 6\}$. It can be seen from the plots that each of the three roots of $L_3^{(\alpha)}$ increases in α .

The second crucial result on a remarkable interlacing property of zeros of shifted Laguerre polynomials is due to Driver and Jordaan (2007). The complete version of their result can be found in their Theorem 2.3. For more results on interlacing properties of zeros of generalized Laguerre and other orthogonal polynomials we refer to Driver (2009), Driver and Jordaan (2011), Driver (2012) and the references therein.

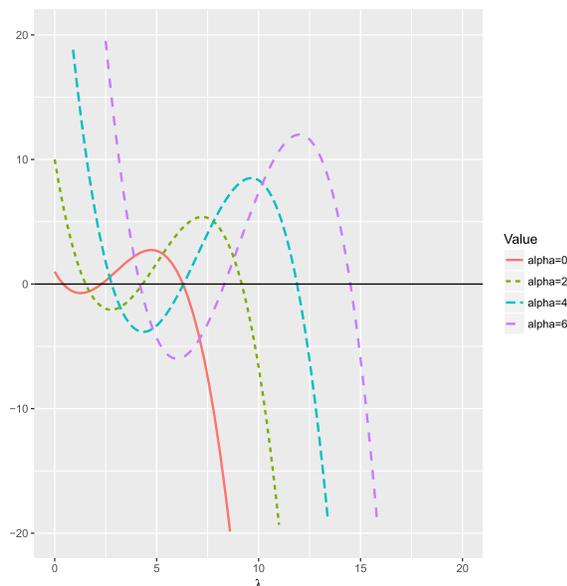


FIG 2. Plots of the $(\alpha, 3)$ -Laguerre polynomials for $\alpha \in \{0, 2, 4, 6\}$.

Theorem 2.2. *Driver and Jordaan (2007)*

Fix $\alpha > -1$ and $k \geq 1$ an integer. If $\lambda_{k,r}^{(\alpha)}$, $r = 1, \dots, k$ are the k zeros of the Laguerre polynomial $L_k^{(\alpha)}$ then

$$0 < \lambda_{k,1}^{(\alpha)} < \lambda_{k,1}^{(\alpha+1)} < \lambda_{k,2}^{(\alpha)} < \lambda_{k,2}^{(\alpha+1)} < \dots < \lambda_{k,k-1}^{(\alpha)} < \lambda_{k,k-1}^{(\alpha+1)} < \lambda_{k,k}^{(\alpha)} < \lambda_{k,k}^{(\alpha+1)}.$$

We illustrate this important result in Table 1 and 2 where it can be seen from the pattern exhibited by the boxes and circles how the zeros of $L_4^{(i)}$ and $L_5^{(i)}$ do interlace for $i \in \{0, 1, 2, 3, 4, 5, 6\}$.

The third result we need focuses on the special case $\alpha = 0$, and the smallest root of $L_k^{(0)}$.

Proposition 2.3. *The sequence $(\lambda_{k,1}^{(0)})_{k \geq 1}$ is decreasing.*

This property is well known and is a special case of a more general decreasing monotonicity property of the zeros of Laguerre polynomials $L_k^{(\alpha)}$ when k increases and α is held fixed. See for example the statement in Driver and Jordaan (2007) on pages 615-616 in the Introduction, where n can be taken here to be k and the polynomial p_n to be $L_k^{(0)}$, with the corresponding weight $w(x) = e^{-x}$. Showing this property is not difficult as it only uses induction and some well-known recursive relationships satisfied by Laguerre polynomials. For completeness we give a proof of Proposition 2.3 in the appendix. Now we shall make the link between k -monotonicity of $\mathcal{P}(\lambda)$ and the generalized Laguerre

TABLE 1

Table illustrating the interlacing property of the zeros of the Laguerre polynomials $L_k^{(i)}$ for $k = 4$ and $i \in \{0, 1, 2, 3, 4, 5, 6\}$. The zeros in bold are the smallest and largest in this table.

Along the boxes and circles, it can be seen that $\lambda_{4,r}^{(i)} > \lambda_{4,r-1}^{(i+1)}$ for all $r = 2, 3, 4$ and $i = 0, 1, 2, 3, 4, 5, 6$. Also, the boxes and circles interlace.

r	1	2	3	4
$i = 0$	0.322	1.745	4.536	9.395
$i = 1$	0.743	2.571	5.731	10.953
$i = 2$	1.226	3.412	6.902	12.458
$i = 3$	1.755	4.265	8.057	13.920
$i = 4$	2.319	5.128	9.200	15.351
$i = 5$	2.910	6.000	10.334	16.755
$i = 6$	3.525	6.878	11.459	18.136

polynomials more explicit. Recall that $\mathcal{P}(\lambda)$ is k -monotone if and only if the inequality in (2.3) is satisfied for all $i \in \mathbb{N}$. Since

$$\begin{aligned} \binom{k}{j} \frac{1}{(i+j)!} &= \frac{k!}{(k-j)!j!(i+j)!} \\ &= \frac{k!}{(k+i)!} \frac{(k+i)!}{(k-j)!(i+j)!} \frac{1}{j!} = \frac{k!}{(k+i)!} \binom{k+i}{k-j} \frac{1}{j!} \end{aligned}$$

we have the following proposition.

Proposition 2.4. For a given integer $k \geq 1$, $\mathcal{P}(\lambda)$ is k -monotone if and only if λ satisfies the inequalities

$$L_k^{(i)}(\lambda) \geq 0$$

for all $i \in \mathbb{N}$.

2.2. Characterization of k -monotonicity of Poisson distributions

The proof of Proposition 2.4 is immediate as the result follows directly from the definition of $L_k^{(\alpha)}$. Since $L_k^{(\alpha)}(0) > 0$ and all the zeros $\lambda_{k,r}^{(\alpha)}$, $r = 1, \dots, k$ are

TABLE 2

Table illustrating the interlacing property of the zeros of the Laguerre polynomials $L_k^{(i)}$ for $k = 5$ and $i \in \{0, 1, 2, 3, 4, 5, 6\}$. The zeros in bold are the smallest and largest in this table.

Along the boxes and circles, it can be seen that $\lambda_{5,r}^{(i)} > \lambda_{5,r-1}^{(i+1)}$ for all $r = 2, 3, 4, 5$ and $i = 0, 1, 2, 3, 4, 5, 6$. Also, the boxes and circles interlace.

r	1	2	3	4	5
$i = 0$	0.263	1.413	3.596	7.085	12.640
$i = 1$	0.617	2.112	4.610	8.399	14.260
$i = 2$	1.031	2.837	5.620	9.682	15.828
$i = 3$	1.490	3.581	6.626	10.944	17.356
$i = 4$	1.985	4.341	7.631	12.188	18.852
$i = 5$	2.510	5.115	8.635	13.417	20.320
$i = 6$	3.059	5.901	9.638	14.634	21.765

simple, $L_{k,r}^{(\alpha)}$ has to be non-negative on $[0, \lambda_{k,1}^{(\alpha)}]$, non-positive on $[\lambda_{k,1}^{(\alpha)}, \lambda_{k,2}^{(\alpha)}]$, and so on as the polynomial switches sign to the right of every zero. This yields the following corollary.

Corollary 2.5. For a given integer $k \geq 1$, $\mathcal{P}(\lambda)$ is k -monotone if and only if $\lambda \in \mathcal{R}^+$ with

$$\mathcal{R}^+ = \begin{cases} \bigcap_{i \in \mathbb{N}} \left\{ [0, \lambda_{k,1}^{(i)}] \cup \dots \cup [\lambda_{k,k}^{(i)}, \infty) \right\}, & \text{if } k \text{ is even,} \\ \bigcap_{i \in \mathbb{N}} \left\{ [0, \lambda_{k,1}^{(i)}] \cup \dots \cup [\lambda_{k,k-1}^{(i)}, \lambda_{k,k}^{(i)}] \right\}, & \text{if } k \text{ is odd.} \end{cases}$$

To guarantee a consistent notation in the corollary above we have used the convention $\lambda_{k,k-1}^{(i)} = 0$ if $k = 1$. In the sequel we will write λ_k instead of $\lambda_{k,1}^{(0)}$. The following proposition characterizes k -monotonicity of $\mathcal{P}(\lambda)$, which is one of our main goals in this work.

Proposition 2.6. For a given integer $k \geq 1$, $\mathcal{P}(\lambda)$ is k -monotone if and only if $\lambda \in [0, \lambda_k]$. Furthermore, $k = \max\{\ell \geq 1 : \mathcal{P}(\lambda) \text{ is } \ell\text{-monotone}\}$ if and only if $\lambda \in (\lambda_{k+1}, \lambda_k]$.

In Table 3 we give the values of λ_k for $k \in \{2, \dots, 10\}$. For $k > 3$, the value is preceded by “ \approx ” to indicate that it is only an approximation. The latter was found using the R-function “uniroot”. Figure 3 shows the plots of the probability

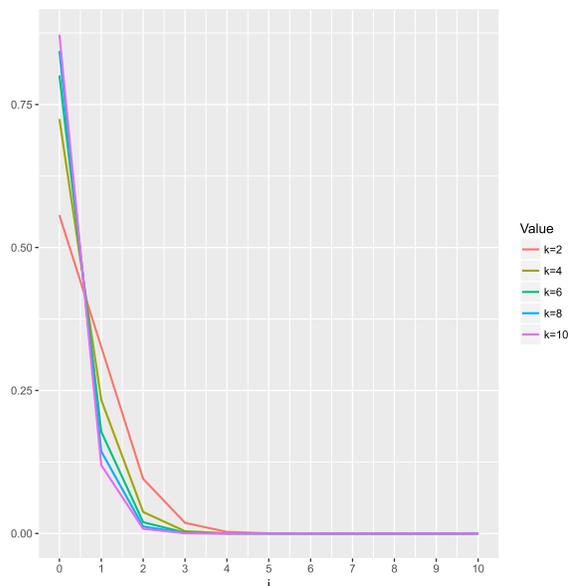


FIG 3. Plots of the probability mass functions of $\mathcal{P}(\lambda_k)$ for $k \in \{2, 4, 6, 8, 10\}$.

TABLE 3
The values of the first root λ_k for $k \in \{2, 3, 4, 5, 6, 7, 8, 9, 10\}$.

k	2	3	4	5	6	7	8	9	10
λ_k	$2 - \sqrt{2}$	≈ 0.415	≈ 0.322	≈ 0.264	≈ 0.222	0.193	0.170	0.152	0.137

mass functions of $\mathcal{P}(\lambda_k)$ for $k \in \{2, 4, 6, 8, 10\}$. It follows from Proposition 2.6 that $\mathcal{P}(\lambda_k)$ is k - but not $(k + 1)$ -monotone. The shapes of the pmfs of $\mathcal{P}(\lambda_k)$ exhibit deeper hollows as k grows. Note also that their value at 0, $e^{-\lambda_k}$, should increasingly converge to 1 the value taken by the Dirac at 0, the completely monotone limit of $\mathcal{P}(\lambda_k)$, as $k \rightarrow \infty$.

Proof of Proposition 2.6. Let k be an even positive integer. To have a more readable proof, let us write

$$I_{k,0}^{(i)} =: [0, \lambda_{k,1}^{(i)}], \quad I_{k,r}^{(i)} =: [\lambda_{k,2r}^{(i)}, \lambda_{k,2r+1}^{(i)}], \quad \text{for } r = 1, \dots, \frac{k}{2} - 1$$

and $I_{k,k/2}^{(i)} =: [\lambda_{k,k}^{(i)}, \infty)$.

Then, using the fact that intersection is distributive over union, the region \mathcal{R}^+ can also be re-written as

$$\mathcal{R}^+ = \bigcup_{\substack{r(i) \in \{0, \dots, k/2\} \\ i \in \mathbb{N}}} \left\{ \bigcap_{i \in \mathbb{N}} I_{k,r(i)}^{(i)} \right\}.$$

Since $\{0, \dots, k/2\}$ is finite, there must exist an $r \in \{0, \dots, k/2\}$ such that the set $S_r =: \{i \in \mathbb{N} : r(i) = r\}$ is infinite. Theorem 2.1 implies that the sequences $(\lambda_{k,1}^{(i)})_{i \in \mathbb{N}}, \dots, (\lambda_{k,k}^{(i)})_{i \in \mathbb{N}}$ are all increasing, which means that

$$\lim_{i \rightarrow \infty} \lambda_{k,1}^{(i)} = \lim_{i \rightarrow \infty} \lambda_{k,2}^{(i)} = \dots = \lim_{i \rightarrow \infty} \lambda_{k,k}^{(i)} = \infty.$$

Thus, it follows that

$$r \neq 0 \Rightarrow \bigcap_{i \in S_r} I_{k,r(i)}^{(i)} = \bigcap_{i \in S_r} I_{k,r}^{(i)} = \emptyset.$$

This allows us to conclude that $\bigcap_{i \in \mathbb{N}} I_{k,r(i)}^{(i)} = \emptyset$ whenever S_r is infinite for some $r > 0$. We consider now the case where $S_1, \dots, S_{k/2}$ are all finite, which implies that $S_0 =: \{i \in \mathbb{N} : r(i) = 0\}$ must be infinite. We will now show that

$$S_0 \neq \mathbb{N} \Rightarrow \left\{ \bigcap_{i \in S_0} I_{k,0}^{(i)} \right\} \cap \left\{ \bigcap_{\substack{j \in S_r \\ 1 \leq r \leq k/2}} I_{k,r}^{(j)} \right\} = \emptyset. \tag{2.4}$$

If $S_0 \neq \mathbb{N}$, then there exists necessarily integers $i \geq 1$, $0 \leq j < i$ and $r \in \{1, \dots, k/2\}$ such that both $I_{k,0}^{(i)}$ and $I_{k,r}^{(j)}$ are sets belonging to the intersection on the right side of (2.4). Now, from both Theorem 2.1 and Theorem 2.2 it follows that

$$\lambda_{k,2r}^{(j)} > \lambda_{k,2r-1}^{(j+1)} > \lambda_{k,2r-2}^{(j+2)} > \dots > \lambda_{k,1}^{(i)}. \tag{2.5}$$

Since $I_{k,0}^{(i)} \cap I_{k,r}^{(j)} = [0, \lambda_{k,1}^{(i)}] \cap [\lambda_{k,2r}^{(j)}, \lambda_{k,2r+1}^{(j)}]$ if $r \leq k/2 - 1$ and $[0, \lambda_{k,1}^{(i)}] \cap [\lambda_{k,k}^{(j)}, \infty)$ if $r = k/2$, the inequalities in (2.5) imply that $I_{k,0}^{(i)} \cap I_{k,r}^{(j)} = \emptyset$. Hence, the intersection on the right side of the display in (2.4) has to be empty. We conclude that the only element remaining in the union defining \mathcal{R}^+ is the one corresponding to the case $S_0 = \mathbb{N}$, that is,

$$\mathcal{R}^+ = \bigcap_{i \in \mathbb{N}} I_{k,0}^{(i)} = [0, \lambda_{k,1}^{(0)}] = [0, \lambda_k].$$

The same reasoning applies to the case when k is odd.

To show the second part of the proposition, we know that $\mathcal{P}(\lambda)$ is k -monotone but not $(k + 1)$ -monotone if and only if $\lambda \leq \lambda_k$ and $\lambda > \lambda_{k+1}$. Also, by Proposition 2.3, we know that $\lambda_{k+1} < \lambda_k$. Thus, the latter statement is equivalent to $\lambda \in (\lambda_{k+1}, \lambda_k]$, which completes the proof. \square

Proposition 2.6 implies the two following results.

Corollary 2.7. *Let F be some distribution function supported on $[0, \infty)$. Then, the pmf of the mixture of Poisson distributions with mixing distribution F given by*

$$p(i) = \int_0^\infty e^{-\lambda} \frac{\lambda^i}{i!} dF(\lambda), \quad i \in \mathbb{N}$$

is k -monotone if the support of F is included in $[0, \lambda_k]$.

The proof of this corollary is immediate, and hence is omitted. Note that the condition is only sufficient. More discussion about this result can be found in Section 3.

Corollary 2.8. *A Poisson distribution $\mathcal{P}(\lambda)$ is completely monotone if and only if it is equal to the Dirac distribution at 0; i.e., $\lambda = 0$.*

Proof. It follows from Theorem A.1 (see Appendix) that

$$\frac{j_{0,1}^2}{2(2k+1)} \leq \lambda_k \leq \frac{3}{2k+1}. \quad (2.6)$$

with $j_{0,1} \approx 2.40$ is the first zero of the Bessel function of the first kind, J_0 . Now, Proposition 2.6 and the inequality in (2.6) imply that $\mathcal{P}(\lambda)$ is completely monotone if $\lambda \leq 3/(2k+1)$ for all $k \in \mathbb{N}$, that is, if $\lambda = 0$. This finishes the proof. \square

3. Some conclusions

In this article we give a necessary and sufficient on λ for $\mathcal{P}(\lambda)$ to be k -monotone for a given integer $k \geq 1$ and a formal proof thereof. It turned out that k -monotonicity of this distribution holds if and only if λ is smaller or equal than λ_k , the first root of the Laguerre polynomial $L_k^{(0)}$. The values reported for $k \in \{2, 3, 4, 5\}$ in p.14 of Giguelay (2017) and $k \in \{2, 3, \dots, 10\}$ in Table 1 on p. 99 of Giguelay and Huet (2018) match exactly with the values of λ_k given in our Table 3. However, the theoretical statement of Property 8 has a typographical error because it involves the first root of a polynomial that is different from $L_k^{(0)}$. The main result of Proposition 2.6 shows also that $\lambda \in (\lambda_{k+1}, \lambda_k]$ is the characterizing condition for $\mathcal{P}(\lambda)$ to be exactly k -monotone; i.e., k - but not $(k+1)$ -monotone. This condition gives a full theoretical justification for the choice of λ in the simulation study in Giguelay and Huet (2018). We would like to note that the result of Proposition 2.6 is in a perfect agreement with the intuition: as k grows the value of the k -monotone pmf at 0 should increase, and hence the rate λ should decrease with k . It is also worthy to note that $\mathcal{P}(\lambda)$ is not k -monotone for any $k \geq 1$ if $\lambda \in (1, \infty)$. For that range, the Poisson model is log-concave with mode location different from 0. Thus, the adaptive shape of the Poisson distributions and related models can be used differently depending on the statistical application.

Chee and Wang (2016) and Giguelay and Huet (2018) both considered k -monotone pmfs to model the distribution of species abundances with the aim of estimating species richness of a given population. Their approaches to k -monotone estimation are different since Chee and Wang (2016) considered maximum likelihood whereas the method implemented by Giguelay (2017) was based on least squares estimation. Also, Chee and Wang (2016) considered the cone generated by the splines $s_\theta(i) = (\theta - i)_+^{k-1} / \sum_{j \leq \theta} j^{k-1}$ for $i \in \mathbb{N}$, $\theta \in [1, \infty)$, instead of the family of splines $(Q_j^k)_{j \geq 0}$ given in (1.2), which is known to generate

the whole k -monotone class of pmfs. If it is easy to show that s_θ is k -monotone for any $\theta \in [1, \infty)$, it is less clear whether these splines also generate all k -monotone pmfs.

Chee and Wang (2016) showed that their k -monotone maximum likelihood estimation improves upon the one based on mixture of Poisson distributions under a variety of scenarios; see their Table 3. According to Corollary 2.7 mixtures of Poisson distributions are actually part of the k -monotone model provided that the range of the rate is confined to the interval $[0, \lambda_k]$. This condition is only sufficient. In fact, in the simplest situation of a mixture of two Poisson distributions with probability mass function $\pi\mu_1^i e^{-\mu_1}/i! + (1-\pi)\mu_2^i e^{-\mu_2}/i!$, $i \in \mathbb{N}$, we can check that this pmf is monotone non-increasing (1-monotone) if $\pi = 1/2$, $\mu_1 = 0.5$ and $\mu_2 = 2$. In this example, the support of the mixing distribution is $\{0.5, 2\}$ and hence it is not included in $[0, \lambda_1] = [0, 1]$. Finding a characterization for k -monotonicity of mixtures of Poisson distributions, as done here for Poisson distributions, is far from being obvious and hence the question remains open.

Acknowledgments

The first author owes special thanks to professor Kathy Driver for her great help by pointing us to a reference for a proof of the interlacing property of the zeros of Laguerre polynomials stated in Theorem 2.2.

Appendix A: Appendix

Proof of Proposition 2.3. In the following we use again the notation λ_k for $\lambda_{k,1}^{(0)}$ and also write L_k for $L_k^{(0)}$. We will use induction to show that $\lambda_{k+1} < \lambda_k$ for all $k \geq 1$. We have $L_1(\lambda) = 1 - \lambda$ and $L_2(\lambda) = 1 - 2\lambda + \lambda^2/2$ with $\lambda_1 = 1 > \lambda_2 = 2 - \sqrt{2}$. Suppose now that $\lambda_{k-1} > \lambda_k$ for all $k \geq 3$. Recall the following recursive well-known identity

$$L_{k+1}(\lambda) = \frac{(2k+1-\lambda)L_k(\lambda) - kL_{k-1}(\lambda)}{k+1}$$

for all $\lambda \in [0, \infty)$. Thus, by definition of λ_k we have that

$$L_{k+1}(\lambda_k) = -\frac{k}{k+1}L_{k-1}(\lambda_k) < 0$$

where the negative sign is a consequence of the fact that it is assumed that $\lambda_{k-1} > \lambda_k$, which means that $L_{k-1} \geq 0$ on $[0, \lambda_k]$. As $L_{k+1} \geq 0$ on $[0, \lambda_{k+1}]$ the inequality $L_{k+1}(\lambda_k) < 0$ implies that we necessarily have $\lambda_k > \lambda_{k+1}$. This completes the proof. \square

The following theorem is due to Gatteschi (2002).

Theorem A.1. For any positive integer k and $\alpha \in (-1, \infty)$ it holds true that

$$\frac{(j_{\alpha,1})^2}{2(2k + \alpha + 1)} \leq \lambda_{k,1}^{(\alpha)} \leq \frac{(\alpha + 1)(\alpha + 3)}{2k + \alpha + 1}$$

where $j_{\alpha,1}$ is the first zero of the Bessel function of the first kind J_α .

References

- BALABDAOUI, F. and DE FOURNAS-LABROSSE, G. (2019). Least squares estimation of a completely monotone pmf: from analysis to statistics, in revision.
- BALABDAOUI, F., DUROT, C. and KOLADJO, F. (2017). On asymptotics of the discrete convex LSE of a p.m.f. *Bernoulli* **23** 1449–1480. [MR3624867](#)
URL <https://doi.org/10.3150/15-BEJ754>
- BALABDAOUI, F. and WELLNER, J. A. (2007). Estimation of a k -monotone density: limit distribution theory and the spline connection. *Ann. Statist.* **35** 2536–2564. [MR2382657](#)
- BALABDAOUI, F. and WELLNER, J. A. (2010). Estimation of a k -monotone density: characterizations, consistency and minimax lower bounds. *Stat. Neerl.* **64** 45–70. [MR2830965](#)
- CHEE, C.-S. and WANG, Y. (2016). Nonparametric estimation of species richness using discrete k -monotone distributions. *Comput. Statist. Data Anal.* **93** 107–118. [MR3406199](#)
- DIMITROV, D. K. and RAFAELI, F. R. (2009). Monotonicity of zeros of Laguerre polynomials. *J. Comput. Appl. Math.* **233** 699–702. [MR2583005](#)
- DRIVER, K. (2009). A note on the interlacing of zeros and orthogonality. *J. Approx. Theory* **161** 508–510. [MR2563067](#)
URL <https://doi.org/10.1016/j.jat.2008.11.008>
- DRIVER, K. (2012). Interlacing of zeros of Gegenbauer polynomials of non-consecutive degree from different sequences. *Numer. Math.* **120** 35–44. [MR2885596](#)
URL <https://doi.org/10.1007/s00211-011-0407-y>
- DRIVER, K., JOOSTE, A. and JORDAAN, K. (2011). Stieltjes interlacing of zeros of Jacobi polynomials from different sequences. *Electron. Trans. Numer. Anal.* **38** 317–326. [MR2871872](#)
- DRIVER, K. and JORDAAN, K. (2007). Interlacing of zeros of shifted sequences of one-parameter orthogonal polynomials. *Numer. Math.* **107** 615–624. [MR2342645](#)
URL <https://doi.org/10.1007/s00211-007-0100-3>
- DRIVER, K. and JORDAAN, K. (2011). Stieltjes interlacing of zeros of Laguerre polynomials from different sequences. *Indag. Math. (N.S.)* **21** 204–211. [MR2835928](#)
URL <https://doi.org/10.1016/j.indag.2011.04.003>
- DUROT, C., HUET, S., KOLADJO, F. and ROBIN, S. (2013). Least-squares estimation of a convex discrete distribution. *Comput. Statist. Data Anal.* **67** 282–298. [MR3079603](#)

- DUROT, C., HUET, S., KOLADJO, F. and ROBIN, S. (2015). Nonparametric species richness estimation under convexity constraint. *Environmetrics* **26** 502–513. [MR3415569](#)
- FEJÉR, L. (1936). Trigonometrische Reihen und Potenzreihen mit mehrfach monotoner Koeffizientenfolge. *Trans. Amer. Math. Soc.* **39** 18–59. [MR1501832](#)
URL <https://doi.org/10.2307/1989642>
- FELLER, W. (1939). Completely monotone functions and sequences. *Duke Math. J.* **5** 661–674. [MR0000315](#)
URL <http://projecteuclid.org/euclid.dmj/1077491420>
- GATTESCHI, L. (2002). Asymptotics and bounds for the zeros of Laguerre polynomials: a survey. *J. Comput. Appl. Math.* **144** 7–27. [MR1909981](#)
- GIGUELAY, J. (2017). Estimation of a discrete probability under constraint of κ -monotonicity. *Electronic journal of statistics* **11** 1–49. [MR3592697](#)
URL <https://hal.archives-ouvertes.fr/hal-01601603>
- GIGUELAY, J. and HUET, S. (2018). Testing k -monotonicity of a discrete distribution. Application to the estimation of the number of classes in a population. *Comput. Statist. Data Anal.* **127** 96–115. [MR3820312](#)
URL <https://doi.org/10.1016/j.csda.2018.02.006>
- JANKOWSKI, H. K. and WELLNER, J. A. (2009). Estimation of a discrete monotone distribution. *Electron. J. Stat.* **3** 1567–1605. [MR2578839](#)
- KNOPP, K. (1925). Mehrfach monotone Zahlenfolgen. *Math. Z.* **22** 75–85. [MR1544715](#)
- KRASIKOV, I. (2003). Bounds for zeros of the Laguerre polynomials. *J. Approx. Theory* **121** 287–291. [MR1971774](#)
- LEFÈVRE, C. and LOISEL, S. (2013). On multiply monotone distributions, continuous or discrete, with applications. *J. Appl. Probab.* **50** 827–847. [MR3102517](#)
URL <https://doi.org/10.1239/jap/1378401239>
- NATALINI, P. and PALUMBO, B. (2003). Some monotonicity results on the zeros of the generalized Laguerre polynomials. In *Proceedings of the Sixth International Symposium on Orthogonal Polynomials, Special Functions and their Applications (Rome, 2001)*, vol. 153. [MR1985706](#)
- NORRIS, J. and POLLOCK, K. (1998). Non-parametric mle for poisson species abundance models allowing for heterogeneity between species. *Environ. Ecol. Stat.* **5** 391–402.
- STEUTEL, F. W. (1969). Note on completely monotone densities. *Ann. Math. Statist.* **40** 1130–1131. [MR0254949](#)
- WILLIAMSON, R. E. (1956). Multiply monotone functions and their Laplace transforms. *Duke Math. J.* **23** 189–207. [MR0077581](#)
URL <http://projecteuclid.org/euclid.dmj/1077466811>