Weierstraß-Institut
für Angewandte Analysis und Stochastik
Leibniz-Institut im Forschungsverbund Berlin e. V.

Preprint ISSN 2198-5855

Rates of convergence for extremes of geometric random variables and marked point processes

Alessandra Cipriani[1], Anne Feidt[2]

submitted: April 15, 2015

[1] Weierstrass Institute
Mohrenstr. 39
10117 Berlin
Germany
E-Mail: Alessandra.Cipriani@wias-berlin.de

[2] Institut für Mathematik
Universität Zürich
Winterthurerstr. 190
8057 Zurich
Switzerland
E-Mail: Anne.Feidt@gmx.net

No. 2099
Berlin 2015

2010 *Mathematics Subject Classification.* 60F99, 62E20.

Key words and phrases. Stein-Chen method, maxima of geometric random variables, Marshall-Olkin geometric distribution, Poisson approximation, marked point process of extremes.

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Leibniz-Institut im Forschungsverbund Berlin e. V.
Mohrenstraße 39
10117 Berlin
Germany

Fax: +49 30 20372-303
E-Mail: preprint@wias-berlin.de
World Wide Web: http://www.wias-berlin.de/

Abstract

We use the Stein-Chen method to study the extremal behaviour of univariate and bivariate geometric laws. We obtain a rate for the convergence, to the Gumbel distribution, of the law of the maximum of i.i.d. geometric random variables, and show that convergence is faster when approximating by a discretised Gumbel. We similarly find a rate of convergence for the law of maxima of bivariate Marshall-Olkin geometric random pairs when approximating by a discrete limit law. We introduce marked point processes of exceedances (MPPEs), both with univariate and bivariate Marshall-Olkin geometric variables as marks and determine bounds on the error of the approximation, in an appropriate probability metric, of the law of the MPPE by that of a Poisson process with same mean measure. We then approximate by another Poisson process with an easier-to-use mean measure and estimate the error of this additional approximation. This work contains and extends results contained in the second author's PhD thesis under the supervision of Andrew D. Barbour. The thesis is available at `http://arxiv.org/abs/1310.2564`.

1 Introduction

The problem of determining the behavior of extremes of random variables is a mathematically intriguing problem. Already the case of the maximum or minimum of a sample of n independent and identically distributed (i.i.d.) random variables $X_1, \ldots, X_n$ presents interesting aspects, starting from the existence of a domain of attraction under suitable conditions (Fisher and Tippett (1928)). In this work it is also highlighted that for discrete distributions for which the jump heights continue to be too large, no non-degenerate limit distribution may be found for $X_{(n)}$, the maximum of the sample (Leadbetter et al. (1983), Theorem 1.7.13). Well-known examples are the geometric and Poisson distributions (Leadbetter et al., 1983, Examples 1.7.14, 1.7.15). There is, however, a way to partially remedy this. By allowing the distributional parameter (or one of them) to vary with the sample size n at a suitable rate, Anderson et al. (1997) and Mitov and Nadarajah (2002) determined extremal limit laws for the Poisson and geometric distributions, respectively. In this paper, we will concentrate on the geometric distribution. We will consider both univariate and bivariate geometric random variables, as well as point processes with geometric marks, again both univariate and bivariate.

Using a result from the Stein-Chen method for Poisson approximation by Barbour and Hall (1984), contained in our work in Theorem 2.6, we determine bounds on the error, in the Kolmogorov distance, of the approximation of the maximum, under different normalisations, of i.i.d. geometric random variables by the Gumbel distribution. Our results show that convergence is faster and requires no conditions on the distributional parameter when approximating by a discretised Gumbel distribution. We similarly determine an error bound for the approximation of the joint law of maxima of random pairs, following the bivariate Marshall-Olkin geometric distribution, by an appropriate discrete limit law.

We further use the Stein-Chen method for Poisson process approximation to determine bounds on the error, in a suitable probability metric, of the law of a *marked point process of exceedances* (MPPE), defined by

$$\Xi_{u,n} := \sum_{i=1}^{n} I_{\{X_i>u\}}\delta_{X_i}, \tag{1.0.1}$$

by that of a Poisson process whose mean measure equals that of the MPPE, where the X_i are i.i.d. geometric random variables and u denotes a threshold. Though the MPPE does not mark the points that exceed u in the way that a *point process of exceedances* (PPE) of the form $\sum_{i=1}^{n} I_{\{X_i>u\}}\delta_{in^{-1}}$ does, it contains more information relevant to the study of extreme values than a *marked point process* (MPP) of the form $\sum_{i=1}^{n}\delta_{X_i}$, as it is not only a random configuration of points in space, but specifically a random configuration of points exceeding a threshold. For more details on the study of PPEs and MPPs in Extreme Value Theory, we refer to Leadbetter et al. (1983) and Resnick (1987), respectively. In addition to an MPPE with univariate geometric marks, we consider an MPPE with bivariate marks that follow the Marshall-Olkin geometric distribution, for which the MPPE indicates whether they lie in a subset A of extreme values of the marks' state space. For both cases, the estimate for the actual "Poisson approximation" comes easily. The reason for this is that we use i.i.d. samples $X_1, \ldots, X_n$ and i.i.d. indicators $I_{\{X_1\in A\}}, \ldots I_{\{X_n\in A\}}$. This allows us to apply Theorem 2.6, which reduces the problem to the approximation of a binomial by a Poisson distribution. However, as the marks have geometric, and thereby discrete margins, the mean measure will live on a lattice and be rather tedious to work with in practial applications. We would therefore prefer to approximate by a further Poisson process with a continuous mean measure. Since the total variation distance is too strong for this kind of approximation, we use the weaker d_2-distance instead, which is not as sensitive towards small changes in the positions of the points of the point processes. Our main effort thus lies in determining error bounds on the approximation of a Poisson process by another Poisson process. As the error given by Theorem 2.6 is only $P(X \in A)$, the error obtained by further approximating by a different Poisson process is typically the bigger of the two, both in the univariate and bivariate case.

For the MPPE with univariate geometric marks, we specifically add, to the first approximation by a Poisson process with mean measure equal to that of the MPPE, a further approximation by a Poisson process with continuous intensity function. For the MPPE with bivariate marks, we proceed by "spreading out" the point probabilities of the Marshall-Olkin distribution over the entire space. As the intensity function obtained through this depends on n, we make some additional assumptions on the parameters of the Marshall-Olkin geometric distribution and further approximate by a Poisson process with a "continuous" mean measure independent of n. By adding up the d_2-error bounds arising from each step, we give the total error bound for the approximation of the MPPE by the final Poisson process.

The structure of the paper is as follows: in Section 2 we recall necessary basic definitions as well as results from the Stein-Chen method. In Section 3 we treat maxima of univariate geometric random variables, as well as joint maxima of random pairs that follow the bivariate Marshall-Olkin geometric distribution. In Section 4 we first study the MPPE with univariate geometric marks, and then the various steps involved in approximating the MPPE with bivariate geometric marks by a Poisson process with "continuous" mean measure independent of n.

2 Background

Throughout, let E be a locally compact separable metric space. In later applications, we simply use $E \subseteq \mathbb{R}^d$, $d \geq 1$. Let E be equipped with its Borel σ-algebra $\mathcal{E} := \mathcal{B}(E)$. δ_z denotes the *Dirac measure* on $\mathcal{E}$ for a point $z \in E$. Suppose that ξ is an integer-valued Radon measure on $\mathcal{E}$. Denote by $\overline{M}_p(E)$ the space of all such point measures ξ on E and equip $\overline{M}_p(E)$ with the σ-algebra $\overline{\mathcal{M}}_p(E)$ that is the smallest σ-algebra making the evaluation maps $\xi \to \xi(B)$ from $M_p(E)$ to $[0, \infty]$ measurable for any set $B \in \mathcal{E}$. Similarly, denote by $M_p(E) \subset \overline{M}_p(E)$ the space of all *finite* point measures and equip $M_p(E)$ with the σ-algebra $\mathcal{M}_p(E)$.

Let $(\Omega, \mathcal{F}, P)$ be a probability space. We define a *point process* on E by $\Xi : (\Omega, \mathcal{F}, P) \to (\overline{M}_p(E), \overline{\mathcal{M}}_p(E))$, $\omega \mapsto \Xi(\omega) = \xi$ and denote its *intensity measure* or *mean measure* on $\mathcal{E}$ by $\boldsymbol{\lambda}$. Furthermore, we denote the law of *Poisson point processes* Ξ on E with mean measure $\boldsymbol{\lambda}$ by $\mathrm{PRM}(\boldsymbol{\lambda})$. For a general review of point processes, we refer to Resnick (1987).

2.1 Distance between measures

We recall here some known facts on a probability metric, the d_2-metric, that is weaker than the total variation metric, defined for example in (Barbour and Brown, 1992, Section 3). Let d_0 be a metric on E that is bounded by 1. We now define metrics on both the space $M_p(E)$ of finite point measures over E and on the set of probability measures over $M_p(E)$. Let $\mathcal{K}$ denote the set of functions $\kappa : E \to \mathbb{R}$ such that

$$s_1(\kappa) = \sup_{z_1 \neq z_2 \in E} \frac{|\kappa(z_1) - \kappa(z_2)|}{d_0(z_1, z_2)} < \infty,$$

which implies that for all $z_1 \neq z_2 \in E$, $|\kappa(z_1) - \kappa(z_2)| \leq s_1(\kappa) d_0(z_1, z_2)$. Thus each function $\kappa \in \mathcal{K}$ is Lipschitz continuous with constant $s_1(\kappa)$. Define a distance d_1 between two finite measures $\boldsymbol{\rho}$ and $\boldsymbol{\sigma}$ over E by

$$d_1(\boldsymbol{\rho}, \boldsymbol{\sigma}) = \begin{cases} 1, & \text{if } \boldsymbol{\rho}(E) \neq \boldsymbol{\sigma}(E), \\ \dfrac{1}{\boldsymbol{\rho}(E)} \sup\limits_{\kappa \in \mathcal{K}} \dfrac{\left| \int_E \kappa d\boldsymbol{\rho} - \int_E \kappa d\boldsymbol{\sigma} \right|}{s_1(\kappa)}, & \text{if } \boldsymbol{\rho}(E) = \boldsymbol{\sigma}(E). \end{cases} \tag{2.1.1}$$

Note that d_1 is bounded by 1. We can use d_1 as distance between point measures in $M_p(E)$. The d_1-distance is then a Wasserstein metric induced by d_0 over point measures on E.

We next construct a metric d_2 that is a Wasserstein metric induced by d_1 over probability measures on $M_p(E)$. Let $\mathcal{H}$ denote the set of functions $h : M_p(E) \to \mathbb{R}$ such that

$$s_2(h) = \sup_{\xi_1 \neq \xi_2 \in M_p(E)} \frac{|h(\xi_1) - h(\xi_2)|}{d_1(\xi_1, \xi_2)} < \infty, \tag{2.1.2}$$

i.e. each function $h \in \mathcal{H}$ is Lipschitz continuous with constant $s_2(h)$. We define a distance d_2 between probability measures μ and ν over $M_p(E)$ by

$$d_2(\mu, \nu) = \frac{1}{s_2(h)} \sup_{h \in \mathcal{H}} \left| \int_{M_p(E)} h d\mu - \int_{M_p(E)} h d\nu \right|. \tag{2.1.3}$$

Note that d_2 is bounded by 1. It can readily be shown that $d_2(\mu, \nu) \leq d_{TV}(\mu, \nu)$.

2.2 The Stein-Chen method

We briefly recall the theory of the Stein-Chen method, which was first worked out by Chen (1975). Let Z be a random variable with law μ. A characterising operator for μ is an operator A_μ on some class of functions $\mathcal{F}$ such that, for any random variable X,

$$\mathbb{E}\left[A_\mu f(X)\right] = 0 \quad \forall f \in \mathcal{F} \quad \text{iff } X \sim \mu.$$

The idea of the method is that $\mathbb{E}\left[A_\mu f(X)\right]$ determines the "distance" between Z and X. To develop this idea, given a function $h \in \mathcal{H}$ (for example, indicator functions of intervals), we need to find a function f_h such that $A_\mu f_h(x) = h(x) - \mathbb{E}\left[h(Z)\right]$. Hence this yields $\left|\mathbb{E}\left[A_\mu f_h(X)\right]\right| = \left|\mathbb{E}\left[h(X)\right] - \mathbb{E}\left[h(Z)\right]\right|$. If the left-hand side is small, then $\left|\mathbb{E}\left[h(X)\right] - \mathbb{E}\left[h(Z)\right]\right|$ is also small. A detailed description of the Stein-Chen method can be found, other than in the original article, in Barbour (1997). We recall also the following results to keep the paper self-contained: let $Z = \{Z_t, t \in \mathbb{R}_+\}$ be an immigration-death process on E with immigration intensity $\boldsymbol{\lambda}$[1] and unit per-capita death rate. For any bounded $h \in \mathcal{H}$, let the function $\gamma : M_p(E) \to \mathbb{R}$ be given by

$$\gamma(\xi) = -\int_0^\infty \left\{ \mathbb{E}^\xi h(Z_t) - \mathrm{PRM}(\boldsymbol{\lambda})(h) \right\} \mathrm{d}t.$$

Then

Proposition 2.1. *$\gamma(\xi)$ is well defined, and $\sup_{\xi:\, \xi(E)=k} |\gamma(\xi)| < \infty$ for each $k \in \mathbb{Z}_+$.*

Lemma 2.2. *For any $\xi \in M_p(E)$,*

$$\Delta_1 \gamma \leq s_2(h) \min\left\{1, \frac{1.65}{\sqrt{\lambda}}\right\}.$$

Proposition 2.3. *The function γ satisfies the Stein equation*

$$(\mathcal{A}\gamma)(\xi) = h(\xi) - \mathrm{PRM}(\boldsymbol{\lambda})(h),$$

for all $\xi \in M_p(E)$.

Lemma 2.4. *Let $\boldsymbol{\lambda}$ and $\tilde{\boldsymbol{\lambda}}$ be two finite measures over E such that $\boldsymbol{\lambda}(E) = \tilde{\boldsymbol{\lambda}}(E) = \lambda$. Then, for any $\xi \in M_p(E)$,*

$$\left| \int_E \left[\gamma(\xi + \delta_z) - \gamma(\xi)\right] \left(\boldsymbol{\lambda}(\mathrm{d}z) - \tilde{\boldsymbol{\lambda}}(\mathrm{d}z)\right) \right| \leq s_2(h)(1 - \mathrm{e}^{-\lambda}) \left(1 + \frac{\lambda}{|\xi| + 1}\right) d_1(\boldsymbol{\lambda}, \tilde{\boldsymbol{\lambda}}).$$

[1] $\boldsymbol{\lambda}$ is a finite measure over E with $\boldsymbol{\lambda}(E) = \lambda$.

Proofs of the above can be found in (Barbour et al., 1992, Prop. 10.1.1), (Barbour et al., 1992, Lemma 10.2.3), (Barbour et al., 1992, Prop. 10.1.2) and (Barbour et al., 1992, Lemma 10.2.2). We also remark that, by Proposition 2.3,

$$|\mathbb{E}(\mathcal{A}\gamma)(\Xi)| = |\mathbb{E}h(\Xi) - \mathrm{PRM}(\boldsymbol{\lambda})(h)| = \left| \int_{M_p(E)} h \mathrm{d}\mathcal{L}(\Xi) - \int_{M_p(E)} h \mathrm{dPRM}(\boldsymbol{\lambda}) \right|, \tag{2.2.1}$$

2.3 General results on approximation by Poisson processes

We use an argument first made by Michel (1988) to establish a bound on the total variation distance between a point process and an "easier" Poisson random measure as follows:

Theorem 2.5. *For each integer $n \geq 1$, let $I_1, \dots, I_n$ be Bernoulli random variables with probability of success $P(I_i = 1) = p_i \in (0,1)$. Let E be a locally compact separable metric space and let $X_1, \dots, X_n$ be i.i.d. E-valued random variables, independent of the I_i's. Moreover, let $\Xi = \sum_{i=1}^n I_i \delta_{X_i}$ and let $W = \sum_{i=1}^n I_i$. Then,*

$$d_{TV}(\mathcal{L}(\Xi), \mathrm{PRM}(\mathbb{E}\Xi)) \leq d_{TV}(\mathcal{L}(W), \mathrm{Poi}(\mathbb{E}W)).$$

Proof. Let $Z_1, \dots, Z_n$ be i.i.d. random variables with distribution $\mathcal{L}(X_1)$, and let them be independent of W. Then the process $\sum_{j=1}^W \delta_{Z_j}$ has the same distribution as the process of interest Ξ. Furthermore, note that a $\mathrm{PRM}(\mathbb{E}\Xi)$ can be realised as $\sum_{j=1}^{W^\star} \delta_{Z_j}$, where $W^\star \sim \mathrm{Poi}(\mathbb{E}W)$ is independent of the Z_j's. Then,

$$\begin{aligned}
&d_{TV}(\mathcal{L}(\Xi), \mathrm{PRM}(\mathbb{E}\Xi)) \\
&= \sup_R \left\{ P\left(\sum_{j=1}^W \delta_{Z_j} \in R \right) - P\left(\sum_{j=1}^{W^\star} \delta_{Z_j} \in R \right) \right\} \\
&= \sup_R \left\{ \sum_{l=0}^n P\left(\sum_{j=1}^l \delta_{Z_j} \in R \right) P(W = l) - \sum_{l=0}^\infty P\left(\sum_{j=1}^l \delta_{Z_j} \in R \right) P(W^\star = l) \right\} \\
&\leq \sup_R \sum_{l=0}^n P\left(\sum_{j=1}^l \delta_{Z_j} \in R \right) \{P(W = l) - P(W^\star = l)\}_+ \leq \sum_{l=0}^n \{P(W = l) - P(W^\star = l)\}_+,
\end{aligned}$$

where $\{\,.\,\}_+ = \max(\,.\,, 0)$. Now define $B_0 = \{l \in \{1, \dots, n\} : P(W = l) > P(W^\star = l)\}$. Then

$$\begin{aligned}
&\sum_{l=0}^n \{P(W = l) - P(W^\star = l)\}_+ = \sum_{l \in B_0} \{P(W = l) - P(W^\star = l)\} \\
&= P(W \in B_0) - P(W^\star \in B_0) \leq \sup_{B \subseteq \mathbb{Z}_+} |P(W \in B) - P(W^\star \in B)| = d_{TV}(\mathcal{L}(W), \mathrm{Poi}(\mathbb{E}W)).
\end{aligned}$$

□

With this theorem at hand, we are able to show

Theorem 2.6. *For each integer $n \geq 1$, let $\mathbf{X}_1, \ldots, \mathbf{X}_n$ be i.i.d. copies of a d-dimensional random vector $\mathbf{X}$ with state space $E \subseteq \mathbb{R}^d$, where $d \geq 1$. For a fixed set $A \in \mathcal{E}$, define $\Xi_A := \sum_{i=1}^n I_{\{\mathbf{X}_i \in A\}} \delta_{\mathbf{X}_i}$ and $W_A := \sum_{i=1}^n I_{\{\mathbf{X}_i \in A\}}$. Then,*

$$d_{TV}(\mathcal{L}(\Xi_A), \mathrm{PRM}(\mathbb{E}\Xi_A)) \leq d_{TV}(\mathcal{L}(W_A), \mathrm{Poi}(\mathbb{E}W_A)) \leq P(\mathbf{X} \in A).$$

Proof. Let $P_A = \mathcal{L}(\mathbf{X}|\mathbf{X} \in A)$ and define an i.i.d. random sample $\mathbf{X}'_1, \ldots, \mathbf{X}'_n$ with common distribution P_A that is independent of the sample $\mathbf{X}_1, \ldots, \mathbf{X}_n$. Then the process $\sum_{i=1}^n I_{\{\mathbf{X}_i \in A\}} \delta_{\mathbf{X}'_i}$ has the same distribution as the process of interest Ξ_A. Note that due to the independence of the samples $\mathbf{X}_1, \ldots, \mathbf{X}_n$ and $\mathbf{X}'_1, \ldots, \mathbf{X}'_n$, the process $\sum_{i=1}^n I_{\{\mathbf{X}_i \in A\}} \delta_{\mathbf{X}'_i}$ is distributed as $\sum_{j=1}^{W_A} \delta_{\mathbf{Z}'_j}$, where the $\mathbf{Z}'_j$ are independent, have common distribution P_A, and are independent of W_A. Furthermore, note that a $\mathrm{PRM}(\mathbb{E}\Xi_A)$ can be realised as $\sum_{j=1}^{W^\star} \delta_{\mathbf{Z}'_j}$, where $W^\star \sim \mathrm{Poi}(\mathbb{E}W_A)$ is independent of $\mathbf{Z}'_j$. It then follows from the proof of Theorem 2.5 that $d_{TV}(\mathcal{L}(\Xi_A), \mathrm{PRM}(\mathbb{E}\Xi_A)) \leq d_{TV}(\mathcal{L}(W_A), \mathrm{Poi}(\mathbb{E}W_A))$. Finally, an application of Theorem 1 by Barbour and Hall (1984) yields

$$d_{TV}(\mathcal{L}(W_A), \mathrm{Poi}(\mathbb{E}W_A)) \leq \frac{1 - e^{-\mathbb{E}W_A}}{\mathbb{E}W_A} \sum_{i=1}^n P(X_i \in A)^2 \leq P(X \in A).$$

□

Finally, we recall the following useful results:

Proposition 2.7. *Barbour et al. (1992, p.235) Let $\boldsymbol{\lambda}$ and $\tilde{\boldsymbol{\lambda}}$ be two finite measures on E. Then*

$$d_{TV}\left(\mathrm{PRM}(\boldsymbol{\lambda}), \mathrm{PRM}(\tilde{\boldsymbol{\lambda}})\right) \leq \int_E |\boldsymbol{\lambda} - \tilde{\boldsymbol{\lambda}}|(dz).$$

Proposition 2.8. *Barbour and Brown (1992, Theorem 3.6) Let $\boldsymbol{\lambda}$ and $\tilde{\boldsymbol{\lambda}}$ be two finite measures over E such that $\boldsymbol{\lambda}(E) = \tilde{\boldsymbol{\lambda}}(E) = \lambda$. Then*

$$d_2\left(\mathrm{PRM}(\boldsymbol{\lambda}), \mathrm{PRM}(\tilde{\boldsymbol{\lambda}})\right) \leq (1 - \mathrm{e}^{-\lambda})(2 - \mathrm{e}^{-\lambda}) d_1(\boldsymbol{\lambda}, \tilde{\boldsymbol{\lambda}}).$$

2.4 The bivariate Marshall-Olkin geometric distribution

The bivariate Marshall-Olkin geometric distribution arises as a natural generalisation of the geometric distribution to two dimensions. It was first introduced by Hawkes (1972) and later studied by Marshall and Olkin (1985) as the discrete counterpart to their bivariate exponential distribution, first derived by them in Marshall and Olkin (1967) using shock models. Limit distributions for maxima of i.i.d. Marshall-Olkin geometric random pairs were established in Mitov and Nadarajah (2005) and Feidt et al. (2010).

Underlying the Marshall-Olkin geometric distribution are Bernoulli trials. Suppose S and T are two Bernoulli random variables with joint probability mass function $P(S = i, T = j) = p_{ij}$, for all $i, j = 0, 1$, and let $S_1, S_2, \ldots$ and $T_1, T_2, \ldots$ be i.i.d. copies of S and T, respectively. Let X_1 and X_2 denote the numbers of 0s before the first 1 in the sequences $S_1, S_2, \ldots$ and $T_1, T_2, \ldots$, respectively. Obviously, X_1 and X_2 follow geometric distributions with failure probabilities $q_1 :=$

$P(S=0)=p_{00}+p_{01}$ and $q_2 := P(T=0)=p_{00}+p_{10}$, respectively. Their joint probability mass function is given by

$$P(X_1=k, X_2=l)=\begin{cases} p_{00}^k q_2^{l-k}(1-p_{00}/q_2-q_2+p_{00}) & \text{for } k<l, \\ p_{00}^k(1-q_1-q_2+p_{00}) & \text{for } k=l, \\ p_{00}^l q_1^{k-l}(1-q_1-p_{00}/q_1+p_{00}) & \text{for } k>l, \end{cases} \tag{2.4.1}$$

for any $(k,l)\in\mathbb{Z}_+^2$. The distribution of $\mathbf{X}=(X_1,X_2)$ thus depends on three parameters: the two marginal failure probabilities q_1 and q_2, as well as $p_{00}=P(S=0,T=0)$, the probability of joint failure. We assume that $p_{00}\geq q_1q_2$. We have

$$P(X_1\geq k, X_2\geq l)=\begin{cases} p_{00}^k q_2^{l-k} & \text{for } k<l, \\ p_{00}^k & \text{for } k=l, \\ p_{00}^l q_1^{k-l} & \text{for } k>l. \end{cases} \tag{2.4.2}$$

Note that if $p_{00}=q_1q_2$, the Marshall-Olkin geometric distribution corresponds to a bivariate distribution with independent geometric margins.

3 Rates of convergence for maxima of geometric random variables

In this section we determine bounds on the error of the approximation, in the Kolmogorov distance, of the laws of maxima of univariate geometric random variables and bivariate Marshall-Olkin geometric random pairs by appropriate limit distributions. The one-dimensional case is treated for example by Mitov and Nadarajah (2002) where they see that, for $X_1,\ldots,X_n$ i.i.d. geometric random variables with probability of success $p=p_n\in(0,1)$, if $p_n\to 0$ as $n\to\infty$ then, for all $x\in\mathbb{R}$,

$$\lim_{n\to+\infty} P\left(X_{(n)}\leq \frac{\log n + x}{p_n}\right)=\mathrm{e}^{-\mathrm{e}^{-x}}. \tag{3.0.3}$$

The following proposition investigates the rate of convergence of this limit result and suggests two improvements. One way to reduce the error is to approximate by a discretised version of the Gumbel distribution, the other is to use different normalising constants.

Proposition 3.1. *For each integer $n\geq 1$, let $X_1,\ldots,X_n$ be i.i.d. geometric random variables with success probability $p_n\in(0,1)$, failure probability $q_n=1-p_n$, probability mass function $P(X_1=k)=p_nq_n^k$ and survival function $\overline{F}(k)=q_n^{k+1}$, for any $k\in\mathbb{Z}_+$. Then:*

(a) (Approximation by a discretised Gumbel distribution) For all $k\in\mathbb{Z}_+$ and for all $k^\star\in\mathbb{R}$ defined by $k^\star=-\log n+k\log(1/q_n)$,

$$\left|P\left(X_{(n)}<\frac{\log n+k^\star}{\log(1/q_n)}\right)-\mathrm{e}^{-\mathrm{e}^{-k^\star}}\right|\leq\frac{\log n}{q_n n}+\frac{1}{n}=:\delta_{\mathrm{PoiAppr}}. \tag{3.0.4}$$

(b) (Approximation by a Gumbel distribution) For all $x\in\mathbb{R}$,

$$\left|P\left(X_{(n)}<\frac{\log n+x}{\log(1/q_n)}\right)-\mathrm{e}^{-\mathrm{e}^{-x}}\right|\leq\delta_{\mathrm{PoiAppr}}+\mathrm{e}^{-1}\log(1/q_n)=:\delta_{\mathrm{Cont}}.$$

(c) *(Using the normalising constants from Mitov and Nadarajah (2002))*

$$\left| P\left(X_{(n)} < \frac{\log n + x}{1 - q_n} \right) - \mathrm{e}^{-\mathrm{e}^{-x}} \right| \leq \delta_{\mathrm{Cont}} + \frac{1 - q_n}{2 q_n} \left(\log^2 n + \mathrm{e}^{-1} \right).$$

Note that the failure probability q_n need not vary with the sample size n for approximation by a discretised Gumbel distribution. The error bound is sharp for any constant $q_n \equiv q \in (0,1)$, showing clearly that it makes more sense to approximate a discrete distribution by another discrete distribution than by a continuous one, as there is no need to add an extra error as in (b).

The extra error in (b), $\mathrm{e}^{-1}\log(1/q_n)$, is the discretisation error that arises when going from the Gumbel concentrated on the lattice of points $k^\star$ to the continuous Gumbel distribution over $\mathbb{R}$. It dominates the overall error in (b) unless q_n tends to 1 fast enough as $n \to \infty$, that is, unless $1 - q_n = O(\log(n)/n)$, in which case the discretisation error is of the same order as the first error term from (a).

Part (c) shows that the choice of normalising constants, more precisely, of the scaling by p_n in (3.0.3), is far from optimal. In order for the approximation in (c) to be good we require $p_n = o(1/\log^2 n)$. Its being a stronger condition than the one for the asymptotic result from (3.0.3) is justified by (c) also being a stronger result in the sense that it gives a uniform bound. The error in (c) is of the same order as the error in (a) only if $1 - q_n = O(1/(n \log n))$.

Proof. For ease of notation we omit the subscript n.

(a) Let $A = [y, \infty)$ for any choice of $y \geq 0$. Then $P(X_1 \in A) = q^{\lceil y \rceil}$, and, setting $k := \lceil y \rceil \in \mathbb{Z}_+$, Theorem 2.6 gives

$$\left| P\left(X_{(n)} < k \right) - \mathrm{e}^{-nq^k} \right| \leq q^k. \tag{3.0.5}$$

With $k^\star \in \mathbb{R}$ chosen such that $k = (\log n + k^\star)/\log(1/q)$, we then have

$$\left| P\left(X_{(n)} < \frac{\log n + k^\star}{\log(1/q)} \right) - \mathrm{e}^{-\mathrm{e}^{-k^\star}} \right| \leq \frac{\mathrm{e}^{-k^\star}}{n}. \tag{3.0.6}$$

In order to find a uniform bound for all $k \in \mathbb{Z}_+$, choose $x_0 := x_{0n} := -\log\log n$. Then, for all k such that $k^\star \geq x_0$, we have $\exp(-k^\star)/n \leq \exp(-x_0)/n = \log(n)/n$, whereas for k such that $k^\star \leq x_0$, we may bound the error in (3.0.6) by further adding the Gumbel distribution to the error at $m^\star$, where $m := \lfloor y_0 \rfloor := \lfloor (\log n + x_0)/\log(1/q) \rfloor$, i.e.

$$\begin{aligned} \left| P\left(X_{(n)} < \frac{\log n + k^\star}{\log(1/q)} \right) - \mathrm{e}^{-\mathrm{e}^{-k^\star}} \right| &\leq \frac{\mathrm{e}^{-m^\star}}{n} + \mathrm{e}^{-\mathrm{e}^{-m^\star}} = \mathrm{e}^{-m\log(1/q)} + \mathrm{e}^{-\mathrm{e}^{-m^\star}} \\ &\leq \frac{\mathrm{e}^{-x_0}}{qn} + \mathrm{e}^{-\mathrm{e}^{-x_0}} \leq \frac{\log n}{qn} + \frac{1}{n}, \end{aligned} \tag{3.0.7}$$

where we used $m \geq y_0 - 1$ in the second inequality. See Figure 1 for a sketch.

(3.0.7) provides a bound for all $k^\star$, and thus

$$\left| P\left(X_{(n)} < \frac{\log n + k^\star}{\log(1/q)} \right) - \mathrm{e}^{-\mathrm{e}^{-k^\star}} \right| \leq \frac{\log n}{qn} + \frac{1}{n}.$$

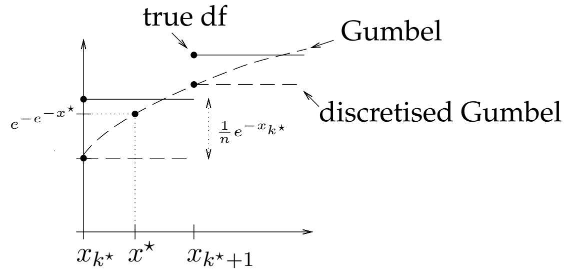

Figure 1: At all values $k^\star \le m^\star \le x_0$, the error cannot exceed the sum of the bound $\mathrm{e}^{-m^\star}/n$ on the difference between the two distributions functions and the height $\exp\{-\mathrm{e}^{-x_0}\}$ of the Gumbel distribution.

(b) Let $x = \log(1/q)y - \log n$. By adding and subtracting $\exp\{-\mathrm{e}^{-x}\}$ into (3.0.4), and noting that, since $y \le \lceil y \rceil = k$, we have $x \le k^\star$ and $\exp\{-\mathrm{e}^{-k^\star}\} - \exp\{-\mathrm{e}^{-x}\} \ge 0$. We thus obtain

$$\left| P\left(X_{(n)} < \frac{\log n + x}{\log(1/q)}\right) - \mathrm{e}^{-\mathrm{e}^{-x}} \right| \le \delta_{\mathrm{PoiAppr}} + \mathrm{e}^{-\mathrm{e}^{-k^\star}} - \mathrm{e}^{-\mathrm{e}^{-x}},$$

where

$$\mathrm{e}^{-\mathrm{e}^{-k^\star}} - \mathrm{e}^{-\mathrm{e}^{-x}} \le \int_x^{k^\star} \mathrm{e}^{-t}\mathrm{e}^{-\mathrm{e}^{-t}}\mathrm{d}t \le \mathrm{e}^{-1}(k^\star - x) = \mathrm{e}^{-1}\log(1/q)(\lceil y \rceil - y) \le \mathrm{e}^{-1}\log(1/q).$$

(c) From (a) and (b) we have

$$\left| P\left(X_{(n)} < \frac{\log n + x}{\log(1/q)}\right) - \mathrm{e}^{-\mathrm{e}^{-x}} \right| \le \frac{\mathrm{e}^{-x}}{n} + \mathrm{e}^{-1}\log(1/q). \tag{3.0.8}$$

Choose $x' \ge -\log n$ such that

$$y = \frac{\log n + x}{\log(1/q)} = \frac{\log n + x'}{1-q}.$$

By adding and subtracting $\exp\left\{-\mathrm{e}^{-x'}\right\}$ into (3.0.8) and observing that $x > x'$ since $\log(1/q)$ $> 1 - q$, we then obtain

$$\left| P\left(X_{(n)} < \frac{\log n + x'}{1-q}\right) - \mathrm{e}^{-\mathrm{e}^{-x'}} \right| \le \frac{\mathrm{e}^{-x'}}{n} + \mathrm{e}^{-1}\log(1/q) + \mathrm{e}^{-\mathrm{e}^{-x}} - \mathrm{e}^{-\mathrm{e}^{-x'}}.$$

For the latter error term we find

$$\begin{aligned} \mathrm{e}^{-\mathrm{e}^{-x}} - \mathrm{e}^{-\mathrm{e}^{-x'}} &= \mathrm{e}^{-\mathrm{e}^{-x}}\left[1 - \mathrm{e}^{-(\mathrm{e}^{-x'} - \mathrm{e}^{-x})}\right] \le \mathrm{e}^{-x'}\left[1 - \mathrm{e}^{-(x-x')}\right] \\ &\le \mathrm{e}^{-x'}(x - x') = \mathrm{e}^{-x'}\left[\log(1/q) - (1-q)\right] y, \end{aligned}$$

where we used $\exp\{-\mathrm{e}^{-x}\} \le 1$ in the first inequality and $1 - \mathrm{e}^{-z} \le z$ for $z \ge 0$ in both inequalities. Note that use of the definition of the logarithm and the geometric series give

$$\begin{aligned} \log(1/q) - (1-q) &= \sum_{j=2}^{\infty} \frac{(1-q)^j}{j} \le \sum_{j=2}^{\infty} \frac{(1-q)^j}{2} = \frac{1}{2}\left[\sum_{j=0}^{\infty}(1-q)^j - 1 - (1-q)\right] \\ &= \frac{1}{2}\left[\frac{1}{q} - 2 + q\right] = \frac{(1-q)^2}{2q}. \end{aligned}$$

Then,

$$\mathrm{e}^{-\mathrm{e}^{-x}} - \mathrm{e}^{-\mathrm{e}^{-x'}} \le \frac{(1-q)^2 y}{2q} \mathrm{e}^{-x'} = \frac{1-q}{2q}(\log n + x')\mathrm{e}^{-x'} \le \frac{1-q}{2q}(\mathrm{e}^{-x'} \log n + \mathrm{e}^{-1}).$$

Thus,

$$\left| P\left(X_{(n)} < \frac{\log n + x'}{1-q} \right) - \mathrm{e}^{-\mathrm{e}^{-x'}} \right| \le \frac{\mathrm{e}^{-x'}}{n} + \mathrm{e}^{-1}\log(1/q) + \frac{1-q}{2q}\left(\mathrm{e}^{-x'}\log n + \mathrm{e}^{-1} \right).$$

For a uniform bound over all x', we choose x_0 as before in (a) and (b), and obtain, with an analogous argument, the overall error bound

$$\frac{\log n}{n} + \mathrm{e}^{-1}\log(1/q) + \frac{1-q}{2q}\left(\log^2 n + \mathrm{e}^{-1}\right) + \frac{1}{n}.$$

□

We now pass to the bivariate case where we assume that the random pair $\mathbf{Z} = (X, Y)$ follows the Marshall-Olkin geometric distribution as defined in Section 2.4, taking values $(k, \ell) \in \mathbb{Z}_+^2$. We use the following normalisation:

$$\mathbf{Z}^\star = (X^\star, Y^\star) = \left(\log\left(\frac{1}{p_{00}}\right) X - \log n \,,\, \log\left(\frac{1}{p_{00}}\right) Y - \log n \right), \tag{3.0.9}$$

taking values $(k^\star, \ell^\star) \in [-\log n, \infty)^2$. Similarly to (a) in Proposition 3.1 above, the following proposition provides bounds on the error of the approximation, in the Kolmogorov distance, of the distribution of maxima of Marshall-Olkin geometric pairs by a discrete limit distribution.

Proposition 3.2. *For each integer $n \ge 3$, let $\mathbf{Z}_1^\star, \dots, \mathbf{Z}_n^\star$ be i.i.d. copies of the random pair $\mathbf{Z}^\star = (X^\star, Y^\star)$ as defined in (3.0.9). Moreover let $X_{(n)}^\star = \max_{1\le i\le n} X_i^\star$ and similarly for $Y_{(n)}^\star$. Then, for all $(k^\star, \ell^\star) \in [-\log n, \infty)^2$,*

$$\left| P\left(X_{(n)}^\star < k^\star,\, Y_{(n)}^\star < \ell^\star \right) - H(k^\star, \ell^\star) \right| \le \left(\frac{1}{n}\right)^{\frac{\log \min\{q_1, q_2\}}{\log p_{00}}} \cdot \frac{\log n}{p_{00}} + \mathrm{e}^{-\sqrt{\log n}},$$

where

$$H(x, y) = \begin{cases} \mathrm{e}^{-\mathrm{e}^{-\frac{\log(p_{00}/q_2)}{\log p_{00}} x - \frac{\log q_2}{\log p_{00}} y}} & x < y, \\ \mathrm{e}^{-\mathrm{e}^{-x}} & x = y, \\ \mathrm{e}^{-\mathrm{e}^{-\frac{\log(p_{00}/q_1)}{\log p_{00}} y - \frac{\log q_1}{\log p_{00}} x}} & x > y, \end{cases}$$

We remark here that the limit distribution H is different from the ones obtained in Mitov and Nadarajah (2005), who used different normalisations. The idea underlying the proof of Proposition 3.2 is to apply Theorem 2.6 to $A = [k, \infty) \times [\ell, \infty)$ and $W_A = \sum_{i=1}^n I_{\{\mathbf{Z}_i \in A\}}$ as follows:

$$\left| P(W_A = 0) - \mathrm{e}^{-nP(\mathbf{Z}\in A)} \right| \le P(\mathbf{Z} \in A).$$

The limit distribution H defined in Proposition 3.2 thus corresponds to $\mathrm{e}^{-nP(X^\star \ge k^\star, Y^\star \ge \ell^\star)}$.

Proof. 1. If $k^\star = \ell^\star$, we choose the auxiliary threshold $x_0 := -\log\log n$. For all k such that $k^\star \geq x_0$, Theorem 2.6 gives

$$\left| P\left(X^\star_{(n)} < k^\star,\, Y^\star_{(n)} < k^\star\right) - \mathrm{e}^{-\mathrm{e}^{-k^\star}} \right| \leq \frac{\mathrm{e}^{-k^\star}}{n} \leq \frac{\mathrm{e}^{-x_0}}{n} \leq \frac{\log n}{n}. \tag{3.0.10}$$

For k such that $k^\star \leq x_0$, we may bound the error in (3.0.10) (i.e. the absolute value in (3.0.10)) by further adding the limit distribution to the error bound at $m^\star$, where $m := \lfloor y_0 \rfloor$ and $y_0 := (\log n + x_0)/\log(1/p_{00})$, i.e.

$$\left| P\left(X^\star_{(n)} < k^\star,\, Y^\star_{(n)} < k^\star\right) - \mathrm{e}^{-\mathrm{e}^{-k^\star}} \right| \leq \frac{\mathrm{e}^{-m^\star}}{n} + \mathrm{e}^{-\mathrm{e}^{-m^\star}} \leq \frac{\mathrm{e}^{-x_0}}{n\,p_{00}} + \mathrm{e}^{-\mathrm{e}^{-x_0}} \leq \frac{\log n}{n\,p_{00}} + \frac{1}{n},$$

where we used $x_0 - \log(1/p_{00}) \leq m^\star \leq x_0$ in the second inequality.

2. Suppose instead $k^\star < \ell^\star$ (the case $k^\star > \ell^\star$ is symmetric). By Theorem 2.6 we have

$$\left| P\left(X^\star_{(n)} < k^\star,\, Y^\star_{(n)} < \ell^\star\right) - \mathrm{e}^{-\mathrm{e}^{-\frac{\log(p_{00}/q_2)}{\log p_{00}}k^\star - \frac{\log q_2}{\log p_{00}}\ell^\star}} \right| \leq \frac{1}{n}\,\mathrm{e}^{-\frac{\log(p_{00}/q_2)}{\log p_{00}}k^\star - \frac{\log q_2}{\log p_{00}}\ell^\star}. \tag{3.0.11}$$

We first assume

$$p_{00} \geq q_2^2. \tag{3.0.12}$$

Choose the following auxiliary points:

$$x_0 := -\frac{\log p_{00}}{2\log(p_{00}/q_2)}\log\log n, \quad y_0 := -\frac{\log p_{00}}{2\log q_2}\log\log n.$$

Note that, under (3.0.12), $x_0 \leq y_0$. Hence:

(a) If $k^\star \geq x_0$, $\ell^\star \geq y_0$, then

$$\frac{1}{n}\,\mathrm{e}^{-\frac{\log(p_{00}/q_2)}{\log p_{00}}k^\star - \frac{\log q_2}{\log p_{00}}\ell^\star} \leq \frac{\log n}{n}.$$

(b) If $k^\star \leq x_0$, $\ell^\star \leq y_0$, then the error in (3.0.11) may be bounded by further adding the limit distribution to the error bound at $(\kappa^\star,\, \lambda^\star)$, with

$$\kappa := \left\lfloor \frac{x_0 + \log n}{\log(1/p_{00})} \right\rfloor, \quad \lambda := \left\lfloor \frac{y_0 + \log n}{\log(1/p_{00})} \right\rfloor,$$

i.e. the error in (3.0.11) is bounded by

$$\begin{aligned}
&\frac{1}{n}\,\mathrm{e}^{-\frac{\log(p_{00}/q_2)}{\log p_{00}}\kappa^\star - \frac{\log q_2}{\log p_{00}}\lambda^\star} + \mathrm{e}^{-\mathrm{e}^{-\frac{\log(p_{00}/q_2)}{\log p_{00}}\kappa^\star - \frac{\log q_2}{\log p_{00}}\lambda^\star}} \\
&\quad \leq \frac{1}{n}\,\mathrm{e}^{-\frac{\log(p_{00}/q_2)}{\log p_{00}}x_0 - \frac{\log q_2}{\log p_{00}}y_0 - \log p_{00}} + \mathrm{e}^{-\mathrm{e}^{-\frac{\log(p_{00}/q_2)}{\log p_{00}}x_0 - \frac{\log q_2}{\log p_{00}}y_0}} = \frac{\log n}{np_{00}} + \frac{1}{n},
\end{aligned}$$

where we used $x_0 - \log(1/p_{00}) \leq \kappa^\star \leq x_0$ and $y_0 - \log(1/p_{00}) \leq \lambda^\star \leq y_0$ in the inequality.

(c) If $k^\star \geq x_0$, $\ell^\star \leq y_0$, then we may bound the error in (3.0.11) by further adding the limit distribution to the error bound at the point $(\lambda^\star,\ \lambda^\star)$. This yields an error bound of

$$\begin{aligned}
&\frac{1}{n}\,\mathrm{e}^{-\frac{\log(p_{00}/q_2)}{\log p_{00}}\lambda^\star-\frac{\log q_2}{\log p_{00}}\lambda^\star}+\mathrm{e}^{-\mathrm{e}^{-\frac{\log(p_{00}/q_2)}{\log p_{00}}\lambda^\star-\frac{\log q_2}{\log p_{00}}\lambda^\star}}=\frac{1}{n}\,\mathrm{e}^{-\lambda^\star}+\mathrm{e}^{-\mathrm{e}^{-\lambda^\star}}\\
&\leq \frac{1}{n}\,\mathrm{e}^{-y_0+\log(1/p_{00})}+\mathrm{e}^{-\mathrm{e}^{-y_0}}=\frac{(\log n)^{\frac{\log p_{00}}{2\log q_2}}}{np_{00}}+\mathrm{e}^{-(\log n)^{\frac{\log p_{00}}{2\log q_2}}}\leq \frac{\log n}{np_{00}}+\mathrm{e}^{-\sqrt{\log n}},
\end{aligned}$$

where we used $y_0 - \log(1/p_{00}) \leq \lambda^\star \leq y_0$, and $\frac{1}{2} \leq (\log p_{00})\,/\,(2\log q_2) \leq 1$ since $q_2^2 \leq p_{00} \leq q_2$.

(d) If $k^\star \leq x_0$, $\ell^\star \geq y_0$, the error bound is given by

$$\frac{1}{n}\,\mathrm{e}^{-\frac{\log(p_{00}/q_2)}{\log p_{00}}k^\star-\frac{\log q_2}{\log p_{00}}\ell^\star}\leq \frac{1}{n}\,\mathrm{e}^{\frac{\log(p_{00}/q_2)}{\log p_{00}}\log n-\frac{\log q_2}{\log p_{00}}y_0}=\frac{\sqrt{\log n}}{n^{\frac{\log q_2}{\log p_{00}}}}\leq\sqrt{\frac{\log n}{n}}\,,$$

since $k^\star \geq -\log n$ and $\frac{\log q_2}{\log p_{00}} \geq \frac{1}{2}$ by (3.0.12).

Therefore, if (3.0.12) holds, an overall bound on the error in (3.0.11) is given by $\frac{\log n}{np_{00}} + e^{-\sqrt{\log n}}$. If instead $p_{00} \leq q_2^2$ (hence $x_0 \geq y_0$), we consider only the cases (a), (b), (d) as above (since we must have $k^* < \ell^\star$). Hence

i. If $k^\star \geq x_0$, $\ell^\star \geq x_0$, then the error bound is given by

$$\frac{1}{n}\,\mathrm{e}^{-\frac{\log(p_{00}/q_2)}{\log p_{00}}k^\star-\frac{\log q_2}{\log p_{00}}\ell^\star}\leq\frac{1}{n}\,\mathrm{e}^{-\frac{\log(p_{00}/q_2)}{\log p_{00}}x_0-\frac{\log q_2}{\log p_{00}}x_0}=\frac{\mathrm{e}^{-x_0}}{n}\leq\frac{\log n}{n}\,.$$

Here we have used the fact that $p_{00} \leq q_2^2$ implies $\frac{\log p_{00}}{2\log(p_{00}/q_2)} \leq 1$.

ii. If $k^\star \leq x_0$, $\ell^\star \leq x_0$, then, choosing $\kappa := \lambda := \left\lfloor \frac{x_0+\log n}{\log(1/p_{00})} \right\rfloor$, the error in (3.0.11) may be bounded by

$$\begin{aligned}
&\frac{1}{n}\,\mathrm{e}^{-\frac{\log(p_{00}/q_2)}{\log p_{00}}\kappa^\star-\frac{\log q_2}{\log p_{00}}\kappa^\star}+\mathrm{e}^{-\mathrm{e}^{-\frac{\log(p_{00}/q_2)}{\log p_{00}}\kappa^\star-\frac{\log q_2}{\log p_{00}}\kappa^\star}}=\frac{\mathrm{e}^{-\kappa^\star}}{n}+\mathrm{e}^{-\mathrm{e}^{-\kappa^\star}}\\
&\leq\frac{\mathrm{e}^{-x_0}}{np_{00}}+\mathrm{e}^{-\mathrm{e}^{-x_0}}=\frac{(\log n)^{\frac{\log p_{00}}{2\log(p_{00}/q_2)}}}{np_{00}}+\mathrm{e}^{-\frac{\log p_{00}}{2\log(p_{00}/q_2)}}\leq\frac{\log n}{np_{00}}+\mathrm{e}^{-\sqrt{\log n}},
\end{aligned}$$

where we used $x_0 - \log(1/p_{00}) \leq \kappa^\star \leq x_0$ in the first inequality and $q_2^2 \geq p_{00}$ in the second.

iii. If $k^\star \leq x_0$, $\ell^\star \geq x_0$, as in (d) we obtain

$$\frac{1}{n}\,\mathrm{e}^{-\frac{\log(p_{00}/q_2)}{\log p_{00}}k^\star-\frac{\log q_2}{\log p_{00}}\ell^\star}\leq\frac{1}{n}\,\mathrm{e}^{\frac{\log(p_{00}/q_2)}{\log p_{00}}\log n-\frac{\log q_2}{\log p_{00}}x_0}\leq\frac{\mathrm{e}^{-\frac{\log q_2}{\log p_{00}}x_0}}{n^{\frac{\log q_2}{\log p_{00}}}}=\frac{\sqrt{\log n}}{n^{\frac{\log q_2}{\log p_{00}}}}\,.$$

Thus, if $p_{00} \leq q_2^2$, an overall bound on the error in (3.0.11) is given by $\frac{\log n}{n^{\frac{\log q_2}{\log p_{00}}}p_{00}} + \mathrm{e}^{-\sqrt{\log n}}$.

$\square$

4 Rates of convergence for MPPEs with geometric marks

4.1 Univariate geometric marks

In Proposition 3.1 we demonstrated that for maxima of geometric random variables the approximation by a discretised Gumbel distribution living on lattice points $k^\star$ gives a smaller error than the approximation by a continuous Gumbel distribution on $\mathbb{R}$. For the latter approximation to be sharp, we need the condition that the failure probability q_n depends on n in such a way that $1 - q_n = o(1/\log n)$ for $n \to \infty$. We encounter a similar behaviour when approximating an MPPE with geometric marks, defined by $\Xi_A := \sum_{i=1}^{n} I_{\{X_i \in A\}} \delta_{X_i}$, by a Poisson process. The set $A \in \mathcal{B}([0,\infty)^2)$ will, in all further applications, be chosen such that points falling into A can be considered "extreme". We consider here the MPPE

$$\Xi^\star_{A^\star} := \sum_{i=1}^{n} I_{\{X_i^\star \in A^\star\}} \delta_{X_i^\star}, \tag{4.1.1}$$

the normalised version of (4.1.1) where the marks are subject to the normalisation used in Proposition 3.1 (a). Proposition 4.1 below gives the error in total variation of the approximation of the law of $\Xi^\star_{A^\star}$ by a Poisson process with mean measure living on the lattice $E^\star$ of normalised points $k^\star$. On the other hand, Proposition 4.2 determines the error of the approximation by a Poisson process with an easy-to-use continuous mean measure, and uses the d_2-metric to achieve this.

Proposition 4.1. *For each integer $n \geq 1$, let $X_1, \ldots, X_n$ be i.i.d. geometric random variables with failure probability $q \in (0,1)$ and $P(X_1 \geq y) = q^{\lceil y \rceil}$, for any $y \geq 0$. Define the normalised random variables $X_i^\star = \log(1/q)X_i - \log n$, $i = 1, \ldots, n$, taking values in $E^\star = \log(1/q)\mathbb{Z}_+ - \log n$. Let $A^\star = [u^\star, \infty)$ for any choice of $u^\star \in [-\log n, \infty)$, and let $\Xi^\star_{A^\star}$ be defined as in (4.1.1). Then the mean measure of $\Xi^\star_{A^\star}$ is given by*

$$\boldsymbol{\pi}^\star(B^\star) = \sum_{k^\star \in A^\star \cap E^\star \cap B^\star} (1-q)\mathrm{e}^{-k^\star}, \quad \text{for any } B^\star \in \mathcal{B}([-\log n, \infty)), \tag{4.1.2}$$

and

$$d_{TV}\left(\mathcal{L}(\Xi^\star_{A^\star}), \mathrm{PRM}(\boldsymbol{\pi}^\star)\right) \leq \frac{\mathrm{e}^{-u^\star}}{n}.$$

Proof. For all $k \in \mathbb{Z}_+$, we use the normalisation $k = (k^\star + \log n)/\log(1/q)$, where $k^\star \in E^\star = \{-\log n, \log(1/q) - \log n, 2\log(1/q) - \log n, \ldots\}$. We then have $P(X_1 = k) = (1-q)q^k = (1-q)\,\frac{\mathrm{e}^{-k^\star}}{n} = P(X_1^\star = k^\star)$, and, for any $B^\star \in \mathcal{B}([-\log n, \infty))$,

$$\boldsymbol{\pi}^\star(B^\star) = nP(X^\star \in A^\star \cap B^\star) = \sum_{k \in A^\star \cap E^\star \cap B^\star} nP(X_1^\star = k^\star) = \sum_{k \in A^\star \cap E^\star \cap B^\star} (1-q)\mathrm{e}^{-k^\star}. \tag{4.1.3}$$

Using Theorem 2.6, we obtain

$$d_{TV}(\mathcal{L}(\Xi^\star_{A^\star}), \mathrm{PRM}(\boldsymbol{\pi}^\star)) \leq P(X_1^\star \geq u^\star) = P\left(X_1 \geq \frac{u^\star + \log n}{\log(1/q)}\right) = q^{\left\lceil \frac{u^\star + \log n}{\log(1/q)} \right\rceil} \leq \frac{\mathrm{e}^{-u^\star}}{n}.$$

□

The following proposition now uses the d_2-metric to approximate the MPPE with geometric marks by a Poisson process with continuous intensity, as the total variation metric is too strong to achieve this. The continuous intensity measure we aim for is the same as that of an MPPE with exponential marks. The result is achieved in two steps: we first estimate the error in the d_2-distance of the approximation by a Poisson process with mean measure given by (4.1.2), and then compare this Poisson process by another one with the desired continuous mean measure, again in the d_2-distance, by making use of Proposition 2.8. We assume here that d_0 is the Euclidean distance on $\mathbb{R}$ bounded by 1, i.e. $d_0(z_1, z_2) = \min(|z_1 - z_2|, 1)$ for any $z_1, z_2 \in \mathbb{R}$, and define the d_1- and d_2-distances as in (2.1.1) and (2.1.3), respectively, in Subsection 2.1.

Proposition 4.2. *For each integer $n \geq 1$, let X_i, $X_i^\star$, $i = 1, \ldots, n$, and $E^\star$ be defined as in Proposition 4.1. Let $A^\star = [u^\star, \infty)$ for any choice of $u^\star \in E^\star$, let $\Xi^\star_{A^\star}$ be defined as in (4.1.1) with mean measure $\boldsymbol{\pi}^\star$ as in (4.1.2), and define the continuous measure $\boldsymbol{\lambda}^\star(B^\star) = \int_{A^\star \cap B^\star} \mathrm{e}^{-x} \mathrm{d}x$, for any $B^\star \in \mathcal{B}([-\log n, \infty))$. Then*

$$d_2\left(\mathcal{L}(\Xi^\star_{A^\star}), \mathrm{PRM}(\boldsymbol{\lambda}^\star)\right) \leq \frac{\mathrm{e}^{-u^\star}}{n} + 2\min\left\{\log\left(1/q\right),\, 1\right\}.$$

Proof. We have

$$d_2\left(\mathcal{L}(\Xi^\star_{A^\star}), \mathrm{PRM}(\boldsymbol{\lambda}^\star)\right) \leq d_2\left(\mathcal{L}(\Xi^\star_{A^\star}), \mathrm{PRM}(\boldsymbol{\pi}^\star)\right) + d_2\left(\mathrm{PRM}(\boldsymbol{\pi}^\star), \mathrm{PRM}(\boldsymbol{\lambda}^\star)\right),$$

where, by Proposition 4.1,

$$d_2\left(\mathcal{L}(\Xi^\star_{A^\star}), \mathrm{PRM}(\boldsymbol{\pi}^\star\right) \leq d_{TV}\left(\mathcal{L}(\Xi^\star_{A^\star}), \mathrm{PRM}(\boldsymbol{\pi}^\star)\right) \leq \frac{\mathrm{e}^{-u^\star}}{n}.$$

It thus remains to determine an estimate of $d_2\left(\mathrm{PRM}(\boldsymbol{\pi}^\star), \mathrm{PRM}(\boldsymbol{\lambda}^\star)\right)$. Since $\boldsymbol{\lambda}^\star(A^\star) = \int_{u^\star}^\infty \mathrm{e}^{-x}\mathrm{d}x = \mathrm{e}^{-u^\star} = \boldsymbol{\pi}^\star(A^\star)$, Proposition 2.8 gives

$$d_2\left(\mathrm{PRM}(\boldsymbol{\pi}^\star), \mathrm{PRM}(\boldsymbol{\lambda}^\star)\right) \leq \left(1 - \mathrm{e}^{-\mathrm{e}^{-u^\star}}\right)\left(2 - \mathrm{e}^{-\mathrm{e}^{-u^\star}}\right) d_1(\boldsymbol{\pi}^\star, \boldsymbol{\lambda}^\star) \leq 2 d_1(\boldsymbol{\pi}^\star, \boldsymbol{\lambda}^\star). \tag{4.1.4}$$

By Definition (2.1.1) of the d_1-distance,

$$d_1(\boldsymbol{\pi}^\star, \boldsymbol{\lambda}^\star) = \mathrm{e}^{u^\star} \sup_{\kappa \in \mathcal{K}} \frac{1}{s_1(\kappa)} \left| \int_{-\log n}^\infty \kappa(x)\boldsymbol{\pi}^\star(\mathrm{d}x) - \int_{-\log n}^\infty \kappa(x)\boldsymbol{\lambda}^\star(\mathrm{d}x)\right|. \tag{4.1.5}$$

We may write the two integrals in the above expression as a sum of integrals over the "normalised unit intervals" $[k^\star, (k+1)^\star) = [k^\star, k^\star + \log(1/q))$, for all $k^\star \in E^\star \cap [u^\star, \infty)$. The modulus then equals

$$\left| \sum_{k^\star \geq u^\star} \left\{ \int_{k^\star}^{k^\star + \log(1/q)} \kappa(x)\boldsymbol{\pi}^\star(\mathrm{d}x) - \int_{k^\star}^{k^\star + \log(1/q)} \kappa(x)\boldsymbol{\lambda}^\star(\mathrm{d}x)\right\}\right|. \tag{4.1.6}$$

Since $\boldsymbol{\pi}^\star$ is concentrated on the lattice points $k^\star \in E^\star \cap [u^\star, \infty)$, we have

$$\int_{k^\star}^{k^\star + \log(1/q)} \kappa(x)\boldsymbol{\pi}^\star(\mathrm{d}x) = \kappa(k^\star)\boldsymbol{\pi}^\star(\{k^\star\}) = \kappa(k^\star)(1-q)\mathrm{e}^{-k^\star}.$$

Note that we obtain the same result by computing

$$\int_{k^\star}^{k^\star+\log(1/q)} \kappa(k^\star)\boldsymbol{\lambda}^\star(\mathrm{d}x) = \kappa(k^\star)\int_{k^\star}^{k^\star+\log(1/q)} \mathrm{e}^{-x}\mathrm{d}x = \kappa(k^\star)(1-q)\mathrm{e}^{-k^\star}.$$

We may thus express (4.1.6) as follows:

$$\left|\sum_{k^\star\geq u^\star}\int_{k^\star}^{k^\star+\log(1/q)} \{\kappa(k^\star)-\kappa(x)\}\,\boldsymbol{\lambda}^\star(\mathrm{d}x)\right| \leq \sum_{k^\star\geq u^\star}\int_{k^\star}^{k^\star+\log(1/q)} |\kappa(k^\star)-\kappa(x)|\,\boldsymbol{\lambda}^\star(\mathrm{d}x).$$

From Lipschitz continuity of κ, we know that $|\kappa(k^\star)-\kappa(x)| \leq s_1(\kappa)d_0(k^\star,x)$ for any $x \in [k^\star, k^\star + \log(1/q))$, where $k^\star \in E^\star \cap [u^\star,\infty)$. The maximum Euclidean distance between $k^\star$ and any point in $[k^\star, k^\star+\log(1/q))$ is of course given by $\log(1/q)$. Since we bound d_0 by 1, we have

$$|\kappa(k^\star)-\kappa(x)| \leq s_1(\kappa)\min\{\log(1/q),\,1\}\,.$$

For the d_1-distance in (4.1.5) we now find, using $\boldsymbol{\lambda}^\star([u^\star,\infty)) = \mathrm{e}^{-u^\star}$,

$$d_1(\boldsymbol{\pi}^\star,\boldsymbol{\lambda}^\star) \leq \mathrm{e}^{u^\star}\sum_{k^\star\geq u^\star}\int_{k^\star}^{k^\star+\log(1/q)} \min\{\log(1/q),\,1\}\,\boldsymbol{\lambda}^\star(\mathrm{d}y) = \min\{\log(1/q),\,1\}\,,$$

which we plug into (4.1.4) to obtain an estimate for $d_2(\mathrm{PRM}(\boldsymbol{\pi}^\star),\mathrm{PRM}(\boldsymbol{\lambda}^\star))$. □

The approximation of $\mathcal{L}(\Xi^\star_{A^\star})$ by $\mathrm{PRM}(\boldsymbol{\lambda}^\star)$, whose continuous intensity function e^{-x} corresponds to that of MPPEs with exponential marks, gives rise to an additional error term which depends only on the failure probability of the geometric distribution. The error will become small only if we allow the failure probability $q = q_n$ to tend to 1 as $n \to \infty$. Since $\log(1/q_n)$ is the length of the normalised unit intervals, this condition causes the lattice structure to melt into the whole real subset $[-\log n, \infty)$ as $n \to \infty$. Note that Proposition 4.2 does not require q_n to vary at a particular rate. The reason for that is that we chose the threshold $u_n^\star$ as element of the lattice $E^\star$. If we had not done so, we would have obtained an additional error term of size $\log(1/q_n)\mathrm{e}^{-u_n^\star}$. In this case, q_n would have needed to vary at a fast enough rate to guarantee a small error despite the factor $\mathrm{e}^{-u_n^\star}$, which roughly corresponds to the expected number of exceedances and should thus be greater than 1. We refer to Section 4.2.4, where we established the error estimate in full detail for MPPEs with bivariate geometric marks.

4.2 Bivariate Marshall-Olkin geometric marks

4.2.1 Approximation in d_{TV} by a Poisson process on a lattice

For any integer $n \geq 1$, let $\mathbf{X}_1,\ldots,\mathbf{X}_n$ be i.i.d. copies of the random pair $\mathbf{X} = (X_1, X_2)$, which follows the Marshall-Olkin geometric distribution from Section 2.4. We use the following normalisation for studying joint threshold exceedances, introduced in Section 3 above:

$$(k^\star, l^\star) = (k\log(1/p_{00}) - \log n\,,\, l\log(1/p_{00}) - \log n)\,,\quad \text{for any } (k,l)\in\mathbb{Z}_+^2, \tag{4.2.1}$$

and denote by $E^\star$ the lattice of normalised points $(k^\star, l^\star)$. The following proposition gives straightforward error estimates for the approximation of the law of $\Xi^\star_{A^\star}$ by that of a Poisson process with mean measure $\mathbb{E}\Xi^\star_{A^\star}$, both for general sets $A^\star$, and for the particular choice $A^\star = [u^\star_n, \infty)^2$ (for which $\Xi^\star_{A^\star}$ captures joint threshold exceedances of the components of the normalised random pairs $\mathbf{X}^\star_1, \ldots, \mathbf{X}^\star_n$).

Proposition 4.3. *Suppose $\mathbf{X} = (X_1, X_2)$ follows the Marshall-Olkin geometric distribution with parameters $q_1, q_2, p_{00} \in (0,1)$. For each integer $n \geq 1$, let $\mathbf{X}^\star_1, \ldots, \mathbf{X}^\star_n$ be i.i.d. copies of the normalised random pair $\mathbf{X}^\star = (X^\star_1, X^\star_2)$ with state space $E^\star$, where $X^\star_j = \log(1/p_{00})X_j - \log n$, for $j = 1, 2$. Let $A^\star \in \mathcal{B}([0,\infty)^2)$ and let $\Xi^\star_{A^\star}$ and $W^\star_{A^\star}$ be defined as follows:*

$$\Xi^\star_{A^\star} = \sum_{i=1}^n I_{\{\mathbf{X}^\star_i \in A^\star\}} \delta_{\mathbf{X}^\star_i}, \quad \textit{and} \quad W^\star_{A^\star} = \sum_{i=1}^n I_{\{\mathbf{X}^\star_i \in A^\star\}}.$$

Then the mean measure of $\Xi^\star_{A^\star}$ is given by

$$\boldsymbol{\pi}^\star(B^\star) := \boldsymbol{\pi}^\star_{A^\star}(B^\star) := \mathbb{E}\Xi^\star_{A^\star}(B^\star) = \sum_{(k^\star, l^\star) \in A^\star \cap E^\star \cap B^\star} nP(X^\star_1 = k^\star, X^\star_2 = l^\star),$$

for any $B^\star \in \mathcal{B}([-\log n, \infty)^2)$, where, for any $(k^\star, l^\star) \in E^\star$,

$$\begin{aligned} &P\left(X^\star_1 = k^\star, X^\star_2 = l^\star\right) \\ &= \begin{cases} \frac{1}{n}\left(1 - \frac{p_{00}}{q_2} - q_2 + p_{00}\right) \mathrm{e}^{-\frac{\log(p_{00}/q_2)}{\log p_{00}} k^\star} \mathrm{e}^{-\frac{\log q_2}{\log p_{00}} l^\star} & \textit{for } k^\star < l^\star, \\ \frac{1}{n}\left(1 - q_1 - q_2 + p_{00}\right) \mathrm{e}^{-k^\star} & \textit{for } k^\star = l^\star, \\ \frac{1}{n}\left(1 - q_1 - \frac{p_{00}}{q_1} + p_{00}\right) \mathrm{e}^{-\frac{\log q_1}{\log p_{00}} k^\star} \mathrm{e}^{-\frac{\log(p_{00}/q_1)}{\log p_{00}} l^\star} & \textit{for } k^\star > l^\star, \end{cases} \end{aligned} \tag{4.2.2}$$

and $d_{TV}\left(\mathcal{L}\left(\Xi^\star_{A^\star}\right), \mathrm{PRM}(\boldsymbol{\pi}^\star)\right) \leq P(\mathbf{X}^\star \in A^\star)$. With $A^\star = A^\star_n = [u^\star_n, \infty)^2$ for any choice of $u^\star_n \geq -\log n$, we obtain

$$d_{TV}\left(\mathcal{L}\left(\Xi^\star_{A^\star}\right), \mathrm{PRM}(\boldsymbol{\pi}^\star)\right) \leq \frac{\mathrm{e}^{-u^\star_n}}{n}. \tag{4.2.3}$$

Proof. With (2.4.1) and (4.2.1), we obtain (4.2.2). For any $B^\star \in \mathcal{B}([-\log n, \infty)^2)$, the mean measure of $\Xi^\star_{A^\star}$ applied to $B^\star$ is given by

$$nP(\mathbf{X}^\star \in A^\star \cap B^\star) = \sum_{(k^\star, l^\star) \in A^\star \cap E^\star \cap B^\star} nP(X^\star_1 = k^\star, X^\star_2 = l^\star).$$

By Theorem 2.6, $d_{TV}\left(\mathcal{L}\left(\Xi^\star_{A^\star}\right), \mathrm{PRM}(\boldsymbol{\pi}^\star)\right) \leq P(\mathbf{X}^\star \in A^\star)$, where, using (2.4.2), we find

$$P\left(\mathbf{X}^\star \in A^\star\right) = P\left(X_1 \geq \left\lceil \frac{u^\star_n + \log n}{\log(1/p_{00})} \right\rceil, X_2 \geq \left\lceil \frac{u^\star_n + \log n}{\log(1/p_{00})} \right\rceil\right) = p_{00}^{\left\lceil \frac{u^\star_n + \log n}{\log(1/p_{00})} \right\rceil} \leq \frac{\mathrm{e}^{-u^\star_n}}{n}. \tag{4.2.4}$$

□

4.2.2 Construction of a "continuous" intensity function

Proposition 4.3 gives an error bound for the approximation of the MPPE $\Xi^\star_{A^\star}$ by a Poisson process whose mean measure $\mathbb{E}\Xi^\star_{A^\star}$ lives on the lattice of normalised points $(k^\star, l^\star)$, i.e. on

$$E^\star = (\log(1/p_{00})\mathbb{Z}_+ - \log n)^2 \subset [-\log n, \infty)^2.$$

We would however prefer to approximate the law of the MPPE by that of a Poisson process with an easier-to-use and more flexible *continuous* intensity measure $\boldsymbol{\lambda}^\star = \boldsymbol{\lambda}^\star_{A^\star}$ living on $A^\star \cap [-\log n, \infty)^2$.

As the survival copula of the Marshall-Olkin geometric distribution is a Marshall-Olkin copula, and thereby consists of both an absolutely continuous part off the diagonal in $[-\log n, \infty)^2$ and a singular part on the diagonal (refer to Section 3.1.1 in Nelsen (2006)), the "continuous" intensity measure $\boldsymbol{\lambda}^\star$ will have to be of the form

$$\boldsymbol{\lambda}^\star(B^\star) = \int_{A^\star \cap B^\star} \lambda^\star(s,t)\mathrm{d}s\mathrm{d}t + \int_{A^\star \cap B^\star \cap \{(s,t):\, s=t\}} \acute{\lambda}^\star(s)\mathrm{dd}ts, \tag{4.2.5}$$

for any $B^\star \in \mathcal{B}([-\log n, \infty)^2)$, for "continuous" intensity functions $\lambda^\star$ and $\acute{\lambda}^\star$ that, if integrated over the entire space, will give n.

Remark 4.4. Note that for simplicity of language we here (and later on) somewhat abuse terminology when speaking of a "continuous" intensity function $\lambda^\star$ or a "continuous" intensity measure $\boldsymbol{\lambda}^\star$. The bivariate intensity function $\lambda^\star$ is not continuous, but piecewise continuous, having a jump along the diagonal. The measure $\boldsymbol{\lambda}^\star$ is continuous only in the sense that it has an intensity with respect to Lebesgue measure (2-dimensional on the off-diagonal and 1-dimensional on the diagonal) and not with respect to a point measure.

The idea is to spread the point mass sitting on each of the off-diagonal lattice points $(k^\star, l^\star) \in E^\star$, $k^\star \neq l^\star$, uniformly over each of their corresponding coordinate rectangles (or rather, coordinate squares)

$$R^\star_{k^\star,l^\star} = \left[k^\star, k^\star + \log\left(\frac{1}{p_{00}}\right)\right) \times \left[l^\star, l^\star + \log\left(\frac{1}{p_{00}}\right)\right), \quad k^\star \neq l^\star,$$

and to also spread the point probabilities of the diagonal points $(k^\star, k^\star)$ over the diagonal line $s = t$, where $s, t \geq -\log n$. We achieve this in the following three steps, illustrated in Figure 2.

Step 1. Consider only the off-diagonal lattice points. We of course have

$$P\left(\mathbf{X}^\star \in R^\star_{k^\star,l^\star}\right) = P\left(X_1^\star = k^\star, X_2^\star = l^\star\right),$$

which is given by (4.2.2), and we may express the mean $nP(\mathbf{X}^\star \in A^\star)$ as

$$\sum_{(k^\star,l^\star)\in A^\star, k^\star \neq l^\star} n \int\int_{R^\star_{k^\star,l^\star}} \frac{P(X_1^\star = k^\star, X_2^\star = l^\star)}{\log^2(1/p_{00})} \mathrm{d}s\mathrm{d}t + \sum_{(k^\star,k^\star)\in A^\star} nP\left(X_1^\star = k^\star, X_2^\star = k^\star\right), \tag{4.2.6}$$

where $\log^2(1/p_{00})$ is the surface area of $R^\star_{k^\star,l^\star}$. As we aim to find a continuous intensity function

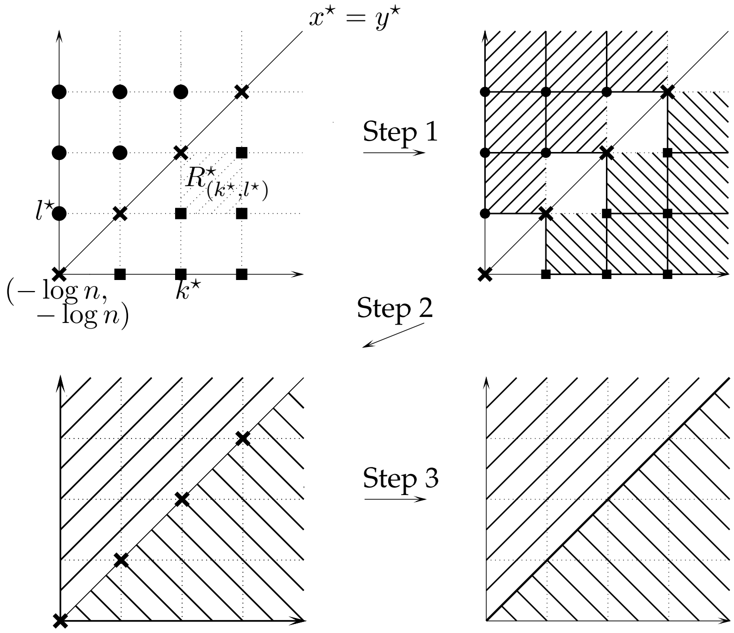

Figure 2: Three steps to construct a "continuous" intensity function.

over the entire space $[-\log n, \infty)^2$, we exchange $k^\star$ and $l^\star$ in the expression of the point probability $P(X_1^\star = k^\star, X_2^\star = l^\star)$ from (4.2.2) by s and t, respectively. E.g., suppose that $k^\star < l^\star$. Then we replace the integral in (4.2.6) by

$$\int\int_{R^\star_{k^\star,l^\star}} \frac{1 - p_{00}/q_2 - q_2 + p_{00}}{\log^2(1/p_{00})} \mathrm{e}^{-\frac{\log(p_{00}/q_2)}{\log p_{00}} s} \mathrm{e}^{-\frac{\log q_2}{\log p_{00}} t} \mathrm{d}s\mathrm{d}t = \frac{1 - p_{00}/q_2 - q_2 + p_{00}}{\log(p_{00}/q_2)\log q_2} P(X_1^\star = k^\star, X_2^\star = l^\star).$$

The switch to variable s and t thus results only in the multiplication of the original point probability by a factor. The goal, however, is to integrate a function in s and t over $R^\star_{k^\star,l^\star}$ and obtain the original point probability. This may be achieved by simply dividing the integrand by the multiplying factor . Hence, we rewrite the mean as follows

$$\sum_{(k^\star,l^\star)\in A^\star, k^\star\neq l^\star} \int\int_{R^\star_{k^\star,l^\star}} \lambda^\star(s,t)\mathrm{d}s\mathrm{d}t + n \sum_{(k^\star,k^\star)\in A^\star} P\left(X_1^\star = k^\star, X_2^\star = k^\star\right),$$

where

$$\lambda^\star(s,t) = \frac{\log(p_{00}/q_2)\log q_2}{\log^2(1/p_{00})} \mathrm{e}^{-\frac{\log(p_{00}/q_2)}{\log p_{00}} s} \mathrm{e}^{-\frac{\log q_2}{\log p_{00}} t}, \ \forall (s,t) \in R^\star_{k^\star,l^\star} \text{ with } k^\star < l^\star. \tag{4.2.7}$$

Analogously, we find

$$\lambda^\star(s,t) = \frac{\log(p_{00}/q_1)\log q_1}{\log^2(1/p_{00})} \mathrm{e}^{-\frac{\log q_1}{\log p_{00}} s} \mathrm{e}^{-\frac{\log(p_{00}/q_1)}{\log p_{00}} t}, \ \forall (s,t) \in R^\star_{k^\star,l^\star} \text{ with } k^\star > l^\star. \tag{4.2.8}$$

(4.2.7) and (4.2.8) supply suitable choices for the intensity function on coordinate rectangles lying above and below the diagonal, respectively.

Step 2. We expand $\lambda^\star(s,t)$ from (4.2.7) and (4.2.8) to the entire space (without the diagonal), i.e. we define

$$\lambda^\star(s,t) := \begin{cases} \frac{\log(p_{00}/q_2)\log q_2}{\log^2(1/p_{00})}\, \mathrm{e}^{-\frac{\log(p_{00}/q_2)}{\log p_{00}}s} \mathrm{e}^{-\frac{\log q_2}{\log p_{00}}t} & \text{for } s<t, \\ \frac{\log(p_{00}/q_1)\log q_1}{\log^2(1/p_{00})}\, \mathrm{e}^{-\frac{\log q_1}{\log p_{00}}s} \mathrm{e}^{-\frac{\log(p_{00}/q_1)}{\log p_{00}}t} & \text{for } s>t, \end{cases}$$

for all $(s,t) \in [-\log n, \infty)^2$. However, this adds surplus mass on the diagonal rectangles $R^\star_{k^\star,k^\star}$.

Step 3. We adjust for the surplus mass on the diagonal rectangles by subtracting it from the point probabilities of the diagonal lattice points $(k^\star,k^\star)$, and accordingly rewrite the mean as follows:

$$\int\int_{A^\star} \lambda^\star_n(s,t)\mathrm{d}s\mathrm{d}t + n \sum_{(k^\star,k^\star)\in A^\star} \left\{ P(X_1^\star = k^\star, X_2^\star = k^\star) - \frac{1}{n}\int_{R^\star_{k^\star,k^\star}} \lambda^\star_n(s,t)\mathrm{d}s\mathrm{d}t \right\}. \tag{4.2.9}$$

Computation of the term in curly brackets shows that the new mass that we put on the diagonal segments of each diagonal rectangle $R^\star_{k^\star,k^\star}$ is given by

$$\frac{\mathrm{e}^{-k^\star}}{n}(1-p_{00})\left[\frac{\log(1/q_1q_2)}{\log(1/p_{00})} - 1\right]. \tag{4.2.10}$$

Note that this equals

$$\int_{k^\star}^{k^\star+\log(1/p_{00})} \frac{\mathrm{e}^{-s}}{n}\left[\frac{\log(1/q_1q_2)}{\log(1/p_{00})} - 1\right]\mathrm{d}s,$$

for each $k^\star \in E^\star$, where we have parameterised the intensity function on the diagonal as projection along the s-axis. We thus define:

$$\begin{aligned} \lambda^\star(s,t) &= \begin{cases} \frac{\log(p_{00}/q_2)\log q_2}{\log^2(1/p_{00})}\, \mathrm{e}^{-\frac{\log(p_{00}/q_2)}{\log p_{00}}s} \mathrm{e}^{-\frac{\log q_2}{\log p_{00}}t} & \text{for} \quad s<t, \\ \frac{\log(p_{00}/q_1)\log q_1}{\log^2(1/p_{00})}\, \mathrm{e}^{-\frac{\log q_1}{\log p_{00}}s} \mathrm{e}^{-\frac{\log(p_{00}/q_1)}{\log p_{00}}t} & \text{for} \quad s>t, \end{cases} \\ \acute{\lambda}^\star(s) &= \frac{\log(p_{00}/q_1q_2)}{\log(1/p_{00})}\,\mathrm{e}^{-s} \quad \text{for } s=t. \end{aligned} \tag{4.2.11}$$

The above construction guarantees the following:

Proposition 4.5. *Let $\boldsymbol{\lambda}^\star$, $\lambda^\star$ and $\acute{\lambda}^\star$ be defined by (4.2.5) and (4.2.11). Then,*
(i) $\boldsymbol{\lambda}^\star\left(R^\star_{k^\star,l^\star}\right) = \boldsymbol{\pi}^\star\left(R^\star_{k^\star,l^\star}\right)$, for any $(k^\star,l^\star)\in E^\star$,
(ii) $\int_{[-\log n,\infty)^2} \lambda^\star(s,t)\mathrm{d}s\mathrm{d}t + \int_{-\log n}^{\infty} \acute{\lambda}^\star(s)\mathrm{d}s = n.$

□

Remark 4.6. It can readily be shown that the new intensity functions $\lambda^\star$ and $\acute{\lambda}^\star$ may be expressed in the original coordinate system by

$$\begin{aligned} \lambda(x,y) = \lambda_n(x,y) &= \begin{cases} n\log(q_2)\log\left(\frac{p_{00}}{q_2}\right) p_{00}^x q_2^{y-x} & \text{for } x<y, \\ n\log(q_1)\log\left(\frac{p_{00}}{q_1}\right) q_1^{x-y} p_{00}^y & \text{for } x>y, \end{cases} \\ \acute{\lambda}(x) = \acute{\lambda}_n(x) &= n\log\left(\frac{p_{00}}{q_1q_2}\right)p_{00}^x \quad \text{for } x=y, \end{aligned}$$

for any $(x,y)\in[0,\infty)^2$. We recognise a weighted and continuous version of $P(X_1 \geq k, X_2 \geq l)$ from (2.4.2).

4.2.3 Assumptions on the distributional parameters

The continuous intensity measure $\boldsymbol{\lambda}^\star$ defined by (4.2.5) and (4.2.11) depends on the parameters q_1, q_2 and p_{00} of the Marshall-Olkin geometric distribution. Our aim is to determine a bound on the error for the approximation of the Poisson process with mean measure $\mathbb{E}\Xi^\star_{A^\star}$, living on the lattice $E^\star$, by a Poisson process with mean measure $\boldsymbol{\lambda}^\star$. As Section 4.2.4 will show, the probability of simultaneous success, p_{11}, for the Marshall-Olkin geometric distribution, will have to tend to 0 as $n \to \infty$. Since $p_{00} + p_{01} + p_{10} + p_{11} = 1$, this of course influences the distributional parameters p_{00}, q_1 and q_2 in that it also makes them dependent on n. The continuous intensity functions $\lambda^\star$ and $\acute{\lambda}^\star$ thus have the drawback that, through their dependence on the parameters p_{00}, q_1 and q_2, they are also dependent on n. We thus try to find other suitable continuous intensity functions that no longer vary with the sample size.

For simplicity, we make the assumption that p_{10} and p_{01} vary at the same rate as $p_{11} = p_{11n}$; more precisely, assume $p_{10} = p_{10n} = \gamma p_{11n}$ and $p_{01} = p_{01n} = \delta p_{11n}$, where γ and δ are strictly positive real numbers, bounded such that p_{10} and p_{01} are smaller than 1. We assume that p_{11n} tends to 0 as $n \to \infty$ at a rate that will be determined later, and express the distributional parameters as functions of it:

$$\begin{aligned} q_{1n} &= 1 - (1+\gamma)p_{11n}, \\ q_{2n} &= 1 - (1+\delta)p_{11n}, \\ p_{00n} &= 1 - (1+\gamma+\delta)p_{11n}. \end{aligned} \tag{4.2.12}$$

Plugging into (4.2.11) and using the relation $\log(1-z) \sim -z$ for $|z| < 1$ and $z \to 0$, we find that $\lambda^\star(s,t)$ and $\acute{\lambda}^\star(s)$ are, for $p_{11n} \to 0$ as $n \to \infty$, asymptotically equal to

$$\begin{aligned} \lambda^\star_{\gamma,\delta}(s,t) &:= \begin{cases} \frac{\gamma(1+\delta)}{(1+\gamma+\delta)^2}\, \mathrm{e}^{-\frac{\gamma}{1+\gamma+\delta}\, s} \mathrm{e}^{-\frac{1+\delta}{1+\gamma+\delta}\, t} & \text{for} \quad s < t, \\ \frac{\delta(1+\gamma)}{(1+\gamma+\delta)^2}\, \mathrm{e}^{-\frac{1+\gamma}{1+\gamma+\delta}\, s} \mathrm{e}^{-\frac{\delta}{1+\gamma+\delta}\, t} & \text{for} \quad s > t, \end{cases} \\ \text{and} \quad \acute{\lambda}^\star_{\gamma,\delta}(s) &:= \frac{1}{1+\gamma+\delta}\, \mathrm{e}^{-s} \qquad \text{for } s = t, \end{aligned} \tag{4.2.13}$$

respectively, for all $(s,t) \in [-\log n, \infty)^2$. At first glance $\lambda^\star_{\gamma,\delta}$ and $\acute{\lambda}^\star_{\gamma,\delta}$ seem to be valid choices for continuous intensity functions independent of n. We will investigate in Section 4.2.5 whether a Poisson process with mean measure $\boldsymbol{\lambda}^\star$ on $A^\star$ may indeed be approximated by a Poisson process with mean measure

$$\boldsymbol{\lambda}^\star_{\gamma,\delta}(B^\star) := \int\int_{A^\star \cap B^\star} \lambda^\star_{\gamma,\delta}(s,t)\mathrm{d}s\mathrm{d}t + \int_{A^\star\cap B^\star \cap \{(s,t):\, s=t\}} \acute{\lambda}^\star_{\gamma,\delta}(s)\mathrm{d}s, \tag{4.2.14}$$

for all $B^\star \in \mathcal{B}([-\log n, \infty)^2)$. To do the corresponding error calculations for a fixed sample size n we first need to examine in further detail the differences between the exponent terms in $\lambda^\star(s,t)$ and $\lambda^\star_{\gamma,\delta}(s,t)$:

Lemma 4.7. *For each integer $n \geq 1$, let $p_{11n} \in (0,1)$ and let q_{1n}, q_{2n}, $p_{00n} \in (0,1)$ be defined by (4.2.12). Then,*

$$(i) \qquad 0 \leq \frac{1+\delta}{1+\gamma+\delta} - \frac{\log q_{2n}}{\log p_{00n}} \leq \frac{\gamma p_{11n}}{1-(1+\gamma+\delta)p_{11n}},$$
$$(ii) \qquad 0 \leq \frac{1+\gamma}{1+\gamma+\delta} - \frac{\log q_{1n}}{\log p_{00n}} \leq \frac{\delta p_{11n}}{1-(1+\gamma+\delta)p_{11n}}.$$

Moreover,

$$(iii) \qquad 0 \leq \log\left(\frac{q_{2n}}{p_{00n}}\right) \leq \frac{\gamma p_{11n}}{1-(1+\gamma+\delta)p_{11n}},$$
$$(iv) \qquad 0 \leq \log\left(\frac{q_{1n}}{p_{00n}}\right) \leq \frac{\delta p_{11n}}{1-(1+\gamma+\delta)p_{11n}},$$

and

$$(v) \qquad \log\left(\frac{1}{p_{00n}}\right)\log\left(\frac{p_{00n}}{q_{1n}q_{2n}}\right) \leq \frac{(1+\gamma+\delta)p_{11n}}{\{1-(1+\gamma+\delta)p_{11n}\}^2}.$$

Proof. (i) For ease of notation we omit the subscript n. Since, for all $|z| < 1$, $-\log(1-z)/z$ is increasing and $-(1-z)\log(1-z)/z$ is decreasing, we obtain the following lower and upper bound, respectively, for $-(\log q_2)/(\log p_{00})$, where $q_2 < p_{00}$:

$$-\frac{1+\delta}{1+\gamma+\delta} \leq -\frac{\log q_2}{\log p_{00}} \leq -\frac{(1+\delta)\cdot[1-(1+\gamma+\delta)p_{11}]}{[1-(1+\delta)p_{11}]\cdot(1+\gamma+\delta)}.$$

Therefore,

$$0 \leq \frac{1+\delta}{1+\gamma+\delta} - \frac{\log q_2}{\log p_{00}} \leq \frac{1+\delta}{1+\gamma+\delta}\left\{1 - \frac{1-(1+\gamma+\delta)p_{11}}{1-(1+\delta)p_{11}}\right\}$$
$$= \frac{(1+\delta)\gamma p_{11}}{(1+\gamma+\delta)[1-(1+\delta)p_{11}]} \leq \frac{\gamma p_{11}}{1-(1+\gamma+\delta)p_{11}}.$$

(iii) Moreover, since $q_2 = p_{00} + p_{10}$, we have $\log(q_2/p_{00}) \geq 0$. Using $\log(1+z) \leq z$ for positive z, we obtain

$$\log(q_2/p_{00}) = \log\left(\frac{p_{00}+p_{10}}{p_{00}}\right) \leq \frac{p_{10}}{p_{00}} = \frac{\gamma p_{11}}{1-(1+\gamma+\delta)p_{11}}.$$

(ii) and (iv) can be shown analogously to (i) and (iii), respectively.

(v) We have

$$\log\left(\frac{1}{p_{00n}}\right)\log\left(\frac{p_{00n}}{q_{1n}q_{2n}}\right) = (-\log p_{00})\left\{-\log(p_{00}+p_{01}) - \log(p_{00}+p_{10}) + \log p_{00}\right\}$$
$$\leq (-\log p_{00})\left\{-\log p_{00} - \log p_{00} + \log p_{00}\right\} = (-\log p_{00})^2 \leq \frac{(1-p_{00})^2}{p_{00}^2} \leq \frac{1-p_{00}}{p_{00}^2}$$
$$= \frac{(1+\gamma+\delta)p_{11}}{\{1-(1+\gamma+\delta)p_{11}\}^2}.$$

□

We will use Lemma 4.7 to determine error estimates in Sections 4.2.4 and 4.2.5.

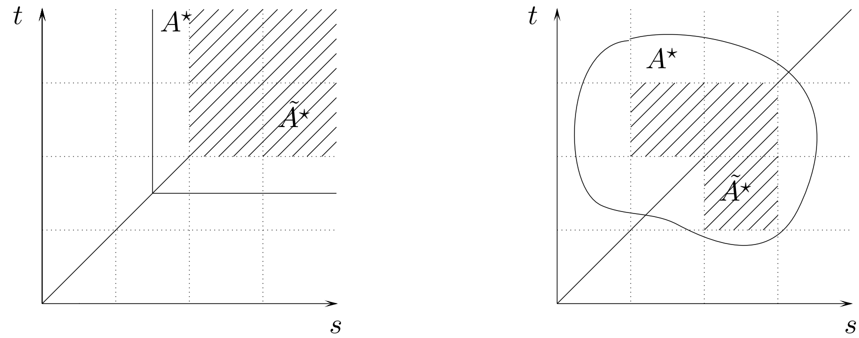

Figure 3: Examples of sets $\tilde{A}^\star$.

Remark 4.8. We suppose here that γ and δ do not vary with n. However, the asymptotic equivalence of (4.2.11) and (4.2.13), and later results (i.e. Propositions 4.10 and 4.12, as well as Corollary 4.13) also hold for the case $\gamma = \gamma_n$ and $\delta = \delta_n$. These results are thus actually stronger than we make them out to be.

4.2.4 Approximation in d_2 by a Poisson process with continuous intensity

We now determine the error of the approximation of the Poisson process with mean measure $\boldsymbol{\pi}^\star$, living on lattice points $(k^\star, l^\star) \in A^\star \cap E^\star$, and the Poisson process with continuous mean measure $\boldsymbol{\lambda}^\star$, living on $A^\star \cap [-\log n, \infty)^2$. Note that any not too small set $A^\star \in \mathcal{B}([-\log n, \infty)^2)$ contains subsets that are unions of coordinate rectangles $R^\star_{k^\star,l^\star}$, i.e. of the form

$$\bigcup_{(k^\star,l^\star)\in M^\star} R^\star_{k^\star,l^\star} \subseteq A^\star, \tag{4.2.15}$$

where $M^\star$ is a countable subset of $E^\star$. Let $\tilde{A}^\star$ denote the biggest subset of $A^\star$ of the form (4.2.15). In order to prove Theorem 4.9, we distinguish between the errors on $\tilde{A}^\star$ and $A^\star \setminus \tilde{A}^\star$. Even though $\boldsymbol{\pi}^\star(A^\star) = \boldsymbol{\lambda}^\star(A^\star)$ is not necessarily satisfied, Proposition 4.5 ensures that at least $\boldsymbol{\pi}^\star(\tilde{A}^\star) = \boldsymbol{\lambda}^\star(\tilde{A}^\star)$. We may therefore use Lemma 2.4 to bound the error on $\tilde{A}^\star$ by way of the d_1-distance between $\boldsymbol{\pi}^\star$ and $\boldsymbol{\lambda}^\star$ on $\tilde{A}^\star$. The size of the d_1-distance depends on the choice of the d_0-distance. We choose the Euclidean distance bounded by 1. For the remaining error, we rely on the "small" size of $A^\star \setminus \tilde{A}^\star$ and use Lemma 2.2 for an upper bound on $\Delta_1\gamma$, where γ is the solution to an appropriate Stein equation. More precisely, let γ be as defined in Proposition 2.1 for $h \in \mathcal{H}$.

Theorem 4.9. *With the notations from Sections 2.4-4.2.3, we obtain, for a set $A^\star \in \mathcal{B}([-\log n, \infty)^2)$,*

$$d_2(\mathrm{PRM}(\boldsymbol{\pi}^\star), \mathrm{PRM}(\boldsymbol{\lambda}^\star)) \le 2\sqrt{2}\log(1/p_{00}) + \min\left\{1, \frac{1.65}{\sqrt{\boldsymbol{\lambda}^\star(A^\star)}}\right\} \boldsymbol{\lambda}^\star(A^\star \setminus \tilde{A}^\star), \tag{4.2.16}$$

where $\tilde{A}^\star$ denotes the biggest subset of $A^\star$ of the form (4.2.15) *with $M^\star$ the biggest subset of $E^\star$ such that $\tilde{A}^\star \subseteq A^\star$.*

Proof. Let $\Xi_{\boldsymbol{\pi}^\star} \sim \mathrm{PRM}(\boldsymbol{\pi}^\star)$ and $\Xi_{\boldsymbol{\lambda}^\star} \sim \mathrm{PRM}(\boldsymbol{\lambda}^\star)$. Suppose that $Z = \{Z_t, t \in \mathbb{R}_+\}$ is an immigration-death process on $A^\star$ with immigration intensity $\boldsymbol{\lambda}^\star$, unit per-capita death rate, equilibrium distribution $\mathcal{L}(\Xi_{\boldsymbol{\lambda}^\star})$, and generator $\mathcal{A}$. Furthermore, let $\mathcal{H}$ denote the set of functions $h : M_p(A^\star) \to \mathbb{R}$ such that (2.1.2) is satisfied and let $\gamma : M_p(A^\star) \to \mathbb{R}$ be defined by $\gamma(\xi) = -\int_0^\infty \{\mathbb{E}^\xi h(Z_t) - \mathrm{PRM}(\boldsymbol{\lambda}^\star)\}\mathrm{d}t$, for any $\xi \in M_p(A^\star)$. By Proposition 2.1, γ is well-defined, and by (2.2.1), $|\mathrm{PRM}(\boldsymbol{\pi}^\star)(h) - \mathrm{PRM}(\boldsymbol{\lambda}^\star)(h)|$ equals $|\mathbb{E}(\mathcal{A}\gamma)(\Xi_{\boldsymbol{\pi}^\star})|$. Proceeding as in the proof of Proposition 2.8, we find that

$$\mathbb{E}(\mathcal{A}\gamma)(\Xi_{\boldsymbol{\pi}^\star}) = \mathbb{E}\int_{A^\star} \left[\gamma(\Xi_{\boldsymbol{\pi}^\star} + \delta_{\mathbf{z}}) - \gamma(\Xi_{\boldsymbol{\pi}^\star})\right] \left(\boldsymbol{\lambda}^\star(d\mathbf{z}) - \boldsymbol{\pi}^\star(d\mathbf{z})\right),$$

and thus

$$\frac{|\mathbb{E}(\mathcal{A}\gamma)(\Xi_{\boldsymbol{\pi}^\star})|}{s_2(h)} \le \frac{1}{s_2(h)}\,\mathbb{E}\left|\int_{A^\star} \left[\gamma(\Xi_{\boldsymbol{\pi}^\star} + \delta_{\mathbf{z}}) - \gamma(\Xi_{\boldsymbol{\pi}^\star})\right] \left(\boldsymbol{\lambda}^\star(d\mathbf{z}) - \boldsymbol{\pi}^\star(d\mathbf{z})\right)\right|,$$

where, for any $\xi \in M_p(A^\star)$,

$$\begin{aligned}&\left|\int_{A^\star} [\gamma(\xi + \delta_{\mathbf{z}}) - \gamma(\xi)](\boldsymbol{\lambda}^\star(d\mathbf{z}) - \boldsymbol{\pi}^\star(d\mathbf{z}))\right| \\ &\quad\le \left|\int_{\tilde{A}^\star} [\gamma(\xi + \delta_{\mathbf{z}}) - \gamma(\xi)](\boldsymbol{\lambda}^\star(d\mathbf{z}) - \boldsymbol{\pi}^\star(d\mathbf{z}))\right| + \left|\int_{A^\star\setminus\tilde{A}^\star} [\gamma(\xi + \delta_{\mathbf{z}}) - \gamma(\xi)](\boldsymbol{\lambda}^\star(d\mathbf{z}) - \boldsymbol{\pi}^\star(d\mathbf{z}))\right|.\end{aligned} \tag{4.2.17}$$

The second summand may be bounded by

$$\int_{A^\star\setminus\tilde{A}^\star} |\gamma(\xi + \delta_{\mathbf{z}}) - \gamma(\xi)| \cdot |\boldsymbol{\lambda}^\star(d\mathbf{z}) - \boldsymbol{\pi}^\star(d\mathbf{z})| \le \Delta_1\gamma \int_{A^\star\setminus\tilde{A}^\star} |\boldsymbol{\lambda}^\star(d\mathbf{z}) - \boldsymbol{\pi}^\star(d\mathbf{z})|. \tag{4.2.18}$$

Note that

$$\int_{A^\star\setminus\tilde{A}^\star} |\boldsymbol{\lambda}^\star(d\mathbf{z}) - \boldsymbol{\pi}^\star(d\mathbf{z})| \le \boldsymbol{\lambda}^\star(A^\star \setminus \tilde{A}^\star),$$

and that Lemma 2.2 gives

$$\Delta_1\gamma \le s_2(h)\min\left\{1,\ \frac{1.65}{\sqrt{\boldsymbol{\lambda}^\star(A^\star)}}\right\}. \tag{4.2.19}$$

By Proposition 4.5, $\boldsymbol{\lambda}^\star(\tilde{A}^\star) = \boldsymbol{\pi}^\star(\tilde{A}^\star) = nP(\mathbf{X}^\star \in \tilde{A}^\star) < \infty$. We may therefore use Lemma 2.4 to bound the first summand by

$$s_2(h)\left(1 - \mathrm{e}^{-\boldsymbol{\lambda}^\star(\tilde{A}^\star)}\right)\left(1 + \frac{\boldsymbol{\lambda}^\star(\tilde{A}^\star)}{|\xi| + 1}\right) d_1(\boldsymbol{\pi}^\star, \boldsymbol{\lambda}^\star)|_{\tilde{A}^\star}, \tag{4.2.20}$$

where $d_1(.\,,.)|_{\tilde{A}^\star}$ denotes the d_1-distance on $\tilde{A}^\star$ (instead of on $A^\star$). We have

$$\mathbb{E}\left(\frac{1}{|\Xi_{\boldsymbol{\pi}^\star}| + 1}\right) = \frac{1 - \mathrm{e}^{-\boldsymbol{\pi}^\star(A^\star)}}{\boldsymbol{\pi}^\star(A^\star)}, \tag{4.2.21}$$

since $|\Xi_{\boldsymbol{\pi}^\star}| \sim \mathrm{Poi}(\boldsymbol{\pi}^\star(A^\star))$. Taking expectations in (4.2.17) and using (4.2.18) - (4.2.21), we obtain

$$\begin{aligned}&|\mathbb{E}(\mathcal{A}\gamma)(\Xi_{\boldsymbol{\pi}^\star})|\,/s_2(h) \\ &\le \left(1 - \mathrm{e}^{-\boldsymbol{\lambda}^\star(\tilde{A}^\star)}\right)\left\{1 + \frac{\boldsymbol{\lambda}^\star(\tilde{A}^\star)}{\boldsymbol{\pi}^\star(A^\star)}\left(1 - \mathrm{e}^{-\boldsymbol{\pi}^\star(A^\star)}\right)\right\} d_1(\boldsymbol{\pi}^\star, \boldsymbol{\lambda}^\star)|_{\tilde{A}^\star} + \min\left\{1,\ \frac{1.65}{\sqrt{\boldsymbol{\lambda}^\star(A^\star)}}\right\}\boldsymbol{\lambda}^\star(A^\star\setminus\tilde{A}^\star).\end{aligned} \tag{4.2.22}$$

We may further simplify by bounding $1 - \mathrm{e}^{-\boldsymbol{\lambda}^\star(\tilde{A}^\star)}$ and $1 - \mathrm{e}^{-\boldsymbol{\pi}^\star(\tilde{A}^\star)}$ by 1 and noting that, since $\boldsymbol{\pi}^\star(A^\star) = \boldsymbol{\pi}^\star(A^\star \setminus \tilde{A}^\star) + \boldsymbol{\lambda}^\star(\tilde{A}^\star) \geq \boldsymbol{\lambda}^\star(\tilde{A}^\star)$, we have

$$1 + \frac{\boldsymbol{\lambda}^\star(\tilde{A}^\star)}{\boldsymbol{\pi}^\star(A^\star)} \leq 2. \tag{4.2.23}$$

With the definitions of $\mathcal{K}$ and $s_1(\kappa)$ from Barbour and Brown (1992) and Barbour et al. (1992), the d_1-distance between $\boldsymbol{\lambda}^\star$ and $\boldsymbol{\pi}^\star$ on $\tilde{A}^\star$ is given by

$$d_1(\boldsymbol{\pi}^\star, \boldsymbol{\lambda}^\star)|_{\tilde{A}^\star} = \frac{1}{\boldsymbol{\lambda}^\star(\tilde{A}^\star)} \sup_{\kappa \in \mathcal{K}} \frac{\left| \int_{\tilde{A}^\star} \kappa d\boldsymbol{\pi}^\star - \int_{\tilde{A}^\star} \kappa d\boldsymbol{\lambda}^\star \right|}{s_1(\kappa)}.$$

As $\tilde{A}^\star$ is a union of coordinate rectangles $R^\star_{k^\star,l^\star}$, the term $\int_{\tilde{A}^\star} \kappa d\boldsymbol{\pi}^\star - \int_{\tilde{A}^\star} \kappa d\boldsymbol{\lambda}^\star$ may be expressed as

$$\sum_{(k^\star,l^\star)\in\tilde{A}^\star} \left\{ \int_{R^\star_{k^\star,l^\star}} \kappa(\mathbf{z})\boldsymbol{\pi}^\star(d\mathbf{z}) - \int_{R^\star_{k^\star,l^\star}} \kappa(\mathbf{z})\boldsymbol{\lambda}^\star(d\mathbf{z}) \right\}. \tag{4.2.24}$$

Furthermore, again by Proposition 4.5,

$$\int_{R^\star_{k^\star,l^\star}} \kappa(\mathbf{z})\boldsymbol{\pi}^\star(d\mathbf{z}) = \kappa((k^\star,l^\star))\boldsymbol{\pi}^\star(R^\star_{k^\star,l^\star}) = \kappa((k^\star,l^\star))\boldsymbol{\lambda}^\star(R^\star_{k^\star,l^\star}).$$

Hence, we find the following upper bound for (4.2.24):

$$\sum_{(k^\star,l^\star)\in\tilde{A}^\star} \int_{R^\star_{k^\star,l^\star}} \left|\kappa((k^\star,l^\star)) - \kappa(\mathbf{z})\right| \boldsymbol{\lambda}^\star(d\mathbf{z}),$$

which, by definition of the Lipschitz constant $s_1(k)$, is smaller than

$$s_1(\kappa)d_0((k^\star,l^\star),\mathbf{z})\boldsymbol{\lambda}^\star(\tilde{A}^\star).$$

The biggest possible Euclidean distance between the lower left corner point $(k^\star, l^\star)$ and any other point $\mathbf{z}$ in the rectangle $R^\star_{k^\star,l^\star}$ is given by the length $\sqrt{2}\log(1/p_{00})$ of its diagonal. Thus,

$$d_1(\boldsymbol{\pi}^\star, \boldsymbol{\lambda}^\star)|_{\tilde{A}^\star} \leq \sqrt{2}\log(1/p_{00}). \tag{4.2.25}$$

(4.2.23) and (4.2.25) give the upper bound $2\sqrt{2}\log(1/p_{00})$ for the first summand of the error term in (4.2.22). This completes the proof. □

Theorem 4.9 gives sharp results only if the probability of simultaneous failure, $p_{00} = p_{00n}$ tends to 1 as $n \to \infty$. This makes sense since $\log(1/p_{00})$, introduced as scaling factor of the original marginal geometric random variables, provides the side lengths of the rescaled lattice squares. The condition $p_{00n} \uparrow 1$ makes the side lengths of the coordinate squares tend to 0 and thus causes the "disappearance" of the lattice into the whole real subset $[-\log n, \infty)^2$. The same holds for the area $A^\star \setminus \tilde{A}^\star$, thereby also causing the disappearance of the second error term as $n \to \infty$.

For sets $A^\star$ that are unions of coordinate rectangles, there is no left-over area $A^\star \setminus \tilde{A}^\star$, and by consequence no second error term. We now apply Theorem 4.9 to the case where $A^\star = A^\star_n = [u^\star_n, \infty)^2$ and express the error estimate in terms of the threshold $u^\star_n$ and the probability of simultaneous success p_{11n}. To achieve this we assume that the distributional parameters p_{00}, q_1 and q_2 are defined as in Section 4.2.3.

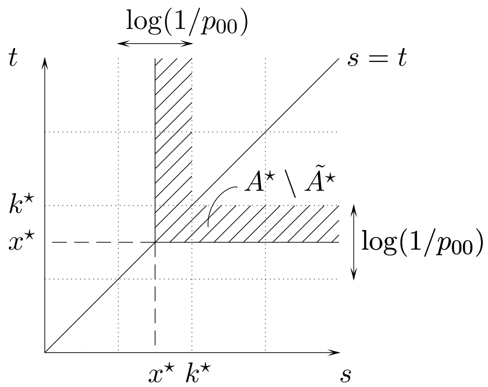

Figure 4: The set $A^\star \setminus \tilde{A}^\star$.

Proposition 4.10. *Let $p_{11n} \in (0,1)$ and assume that q_{1n}, q_{2n} and p_{00n} satisfy (4.2.12). For any choice of $u_n^\star \geq -\log n$, define $A^\star = [u_n^\star, \infty)^2$. With the notations from Theorem 4.9,*

$$d_2(\mathrm{PRM}(\boldsymbol{\pi}^\star), \mathrm{PRM}(\boldsymbol{\lambda}^\star)) \leq \frac{(1+\gamma+\delta)p_{11n}}{[1-(1+\gamma+\delta)p_{11n}]^2} \left\{2\sqrt{2} + 3\min\left\{\mathrm{e}^{-u_n^\star},\, 1.65\mathrm{e}^{-u_n^\star/2}\right\}\right\}.$$

Proof. For ease of notation we omit the subscript n. We apply result (4.2.16) from Theorem 4.9 to the special case $A^\star = [u^\star, \infty)^2$. Due to (4.2.12) and $-\log(1-z) \leq z/(1-z)$ for $|z| < 1$, we may bound the first of the two error terms in (4.2.16) as follows:

$$2\sqrt{2}\log(1/p_{00}) \leq \frac{2\sqrt{2}(1+\gamma+\delta)p_{11}}{1-(1+\gamma+\delta)p_{11}} \leq \frac{2\sqrt{2}(1+\gamma+\delta)p_{11}}{[1-(1+\gamma+\delta)p_{11}]^2}. \tag{4.2.26}$$

Direct computation yields $\boldsymbol{\lambda}^\star(A^\star) = \mathrm{e}^{-u^\star}$. As illustrated by Figure 4, $\boldsymbol{\lambda}^\star(A^\star \setminus \tilde{A}^\star)$ may be bounded by

$$\int_{u^\star}^{\infty} \log(1/p_{00}) \sup_{s\in[u^\star,k^\star]} \lambda^\star(s,t)\mathrm{d}t + \int_{u^\star}^{\infty} \log(1/p_{00}) \sup_{t\in[u^\star,k^\star]} \lambda^\star(s,t)\mathrm{d}s + \int_{u^\star}^{k^\star} \sqrt{2}\log(1/p_{00}) \sup_{s\in[u^\star,k^\star]} \acute{\lambda}^\star(s)\mathrm{d}s.$$

Note that

$$\sup_{s\in[u^\star,k^\star]} \exp\left\{-\frac{\log(p_{00}/q_2)}{\log p_{00}}\, s\right\} \leq \exp\left\{-\frac{\log(p_{00}/q_2)}{\log p_{00}}\, u^\star\right\},$$
$$\sup_{t\in[u^\star,k^\star]} \exp\left\{-\frac{\log(p_{00}/q_1)}{\log p_{00}}\, t\right\} \leq \exp\left\{-\frac{\log(p_{00}/q_1)}{\log p_{00}}\, u^\star\right\},$$

and $\sup_{s\in[u^\star,k^\star]} \mathrm{e}^{-s} \leq \mathrm{e}^{-u^\star}$. Thus, by definition (4.2.11) of $\lambda^\star(s,t)$,

$$\int_{u^\star}^{\infty} \log(1/p_{00}) \sup_{s\in[u^\star,k^\star]} \lambda^\star(s,t)\mathrm{d}t \leq \frac{\log(p_{00}/q_2)\log q_2}{\log(1/p_{00})}\, \mathrm{e}^{-\frac{\log(p_{00}/q_2)}{\log p_{00}}\, u^\star} \int_{u^\star}^{\infty} \mathrm{e}^{-\frac{\log q_2}{\log p_{00}}\, t}\mathrm{d}t,$$

which equals $\log(q_2/p_{00})\mathrm{e}^{-u^\star}$. Analogously,

$$\int_{u^\star}^{\infty} \log(1/p_{00}) \sup_{t\in[u^\star,k^\star]} \lambda^\star(s,t)\mathrm{d}s \le \log(q_1/p_{00})\mathrm{e}^{-u^\star},$$

whereas

$$\int_{u^\star}^{k^\star} \sqrt{2}\log(1/p_{00}) \sup_{s\in[u^\star,k^\star]} \acute{\lambda}^\star(s)\mathrm{d}s \le 2\log^2(1/p_{00}) \frac{\log(p_{00}/q_1q_2)}{\log(1/p_{00})}\,\mathrm{e}^{-u^\star},$$

since $k^\star - u^\star \le \sqrt{2}\log(1/p_{00})$. We obtain

$$\boldsymbol{\lambda}^\star(A^\star \setminus \tilde{A}^\star) = \mathrm{e}^{-u^\star}\left\{\log\left(\frac{q_2}{p_{00}}\right) + \log\left(\frac{q_1}{p_{00}}\right) + 2\log\left(\frac{1}{p_{00}}\right)\log\left(\frac{p_{00}}{q_1q_2}\right)\right\}.$$

By Lemma 4.7 (iii)-(v), the term in curly brackets may be bounded by

$$\frac{(\gamma+\delta)p_{11}}{1-(1+\gamma+\delta)p_{11}} + \frac{2(1+\gamma+\delta)p_{11}}{\{1-(1+\gamma+\delta)p_{11}\}^2} \le \frac{3(1+\gamma+\delta)p_{11}}{\{1-(1+\gamma+\delta)p_{11}\}^2}.$$

An upper bound for the second error term in (4.2.16) is thus given by

$$\min\left\{\mathrm{e}^{-u_n^\star},\ 1.65\mathrm{e}^{-u_n^\star/2}\right\} \frac{3(1+\gamma+\delta)p_{11n}}{[1-(1+\gamma+\delta)p_{11n}]^2}.$$

By adding this to the bound in (4.2.26) we obtain the result. □

The first of the error terms given by Proposition 4.10, i.e. $2\sqrt{2}(1+\gamma+\delta)p_{11n}\ /\ [1-(1+\gamma+\delta)p_{11n}]^2$, is a bound on the error $2\sqrt{2}\log(1/p_{00n})$ from Theorem 4.9, where we used the assumption from Section 4.2.3 that $p_{00n} = 1-(1+\gamma+\delta)p_{11n}$. This error term thus becomes small only if the probability of simultaneous success, p_{11n}, tends to 0 as n increases. The second error term, i.e.

$$\left\{\mathrm{e}^{-u_n^\star},\ 1.65\mathrm{e}^{-u_n^\star/2}\right\} \frac{3(1+\gamma+\delta)p_{11n}}{[1-(1+\gamma+\delta)p_{11n}]^2},$$

is the bigger of the two, and determines the rate at which p_{11n} must converge to 0. The reason for that is that p_{11n} must converge fast enough in order to offset the effect of the factor $\mathrm{e}^{-u_n^\star}$ which we will want to be increasing with increasing n, since $\mathrm{e}^{-u_n^\star} = \boldsymbol{\lambda}^\star(A_n^\star)$ is the expected number of points in $A_n^\star$ of the approximating Poisson process, as well as more or less the expected number of threshold exceedances of the MPPE, for which we have $\mathrm{e}^{-u_n^\star}/p_{00n} \le \boldsymbol{\pi}^\star(A_n^\star) \le \mathrm{e}^{-u_n^\star}$. For instance, for a threshold $u_n^\star$ of size $-\log\log n$, the expected number of points in $A^\star$ of the two Poisson processes is $\log n$, the MPPE captures roughly the biggest $\log n$ points of its sample, and we need $p_{11n} = o(\log^{-1} n)$ for a sharp error bound. Suppose, for example, that $p_{11n} = n^{-1}$. Then, by (4.2.12), the marginal probabilities of failure of $\mathbf{X}_n^\star$, q_{1n} and q_{2n}, as well as the probability of simultaneous failure, p_{00n}, tend to 1 very fast.

The mean measure $\boldsymbol{\lambda}^\star$ is by definition dependent on the values of the distributional parameters. Since these need to vary with the sample size n in order to obtain a small error for the approximation of $\mathrm{PRM}(\boldsymbol{\pi}^\star)$ by $\mathrm{PRM}(\boldsymbol{\lambda}^\star)$, it follows that $\boldsymbol{\lambda}^\star = \boldsymbol{\lambda}_n^\star$ (and of course also $\boldsymbol{\pi}^\star = \boldsymbol{\pi}_n^\star$). Though we have now achieved the goal of successfully approximating by a Poisson process with a continuous intensity, the conditions needed to accomplish this imply that we are not satisfied with our results yet, since we prefer to approximate by a Poisson process with continuous intensity that does not vary with n. As the next section will demonstrate, a suitable candidate is given by the Poisson process with intensity measure $\boldsymbol{\lambda}^\star_{\gamma,\delta}$ defined in (4.2.14).

4.2.5 Approximation in d_2 and d_{TV} by a Poisson process independent of n

We determine an error estimate for the approximation of the Poisson process with intensity measure $\boldsymbol{\lambda}^\star = \boldsymbol{\lambda}^\star_n$ by the Poisson process with intensity measure $\boldsymbol{\lambda}^\star_{\gamma,\delta}$, defined in (4.2.14), that does not depend on the sample size n. Since both intensities are continuous, there is no special need to use the d_2-distance. We give the error in both the total variation and the d_2 distances. For the error in total variation we may straightforwardly use Proposition 2.7 for the approximation of two Poisson processes. For the d_2-error, which will be smaller than the d_{TV}, we may additionally use Lemma 2.2 for an upper bound on $\Delta_1\gamma$, where γ is the solution of an adequate Stein equation. This bound, containing the factor $\boldsymbol{\lambda}^\star(A^\star)^{-1/2}$ (or $\boldsymbol{\lambda}^\star_{\gamma,\delta}(A^\star)^{-1/2}$), serves in reducing the d_2-error.

Theorem 4.11. *With the notations from Sections 2.4 and 4.2.1-4.2.3, we obtain, for any set $A^\star \in \mathcal{B}([-\log n, \infty)^2)$,*

$$(i) \qquad d_{TV}\left(\mathrm{PRM}(\boldsymbol{\lambda}^\star), \mathrm{PRM}(\boldsymbol{\lambda}^\star_{\gamma,\delta})\right) \le \int_{A^\star} |\boldsymbol{\lambda}^\star(d\mathbf{z}) - \boldsymbol{\lambda}^\star_{\gamma,\delta}(d\mathbf{z})|,$$

$$\begin{aligned}(ii) \qquad & d_2\left(\mathrm{PRM}(\boldsymbol{\lambda}^\star), \mathrm{PRM}(\boldsymbol{\lambda}^\star_{\gamma,\delta})\right) \\ & \le \min\left\{1,\, 1.65 \min\left\{\boldsymbol{\lambda}^\star(A^\star)^{-1/2},\, \boldsymbol{\lambda}^\star_{\gamma,\delta}(A^\star)^{-1/2}\right\}\right\} \int_{A^\star} |\boldsymbol{\lambda}^\star(d\mathbf{z}) - \boldsymbol{\lambda}^\star_{\gamma,\delta}(d\mathbf{z})|.\end{aligned}$$

Proof. (i) By Proposition 4.5 (ii), $\boldsymbol{\lambda}^\star$ is finite. Moreover, $\boldsymbol{\lambda}^\star_{\gamma,\delta}$ is finite since integration of $\lambda^\star_{\gamma,\delta}$ and $\acute{\lambda}^\star_{\gamma,\delta}$ over $[u^\star, \infty)^2$ gives $\mathrm{e}^{-u^\star}$ which equals n for $u^\star = -\log n$. Proposition 2.7 then immediately gives the result.
(ii) Using the same immigration-death process Z and arguments as in the proof of Theorem 4.9, we can show that for $\Xi^\star_{\gamma,\delta} \sim \mathrm{PRM}(\boldsymbol{\lambda}^\star_{\gamma,\delta})$,

$$\mathbb{E}h(\Xi^\star_{\gamma,\delta}) - \mathrm{PRM}(\boldsymbol{\lambda}^\star)(h) = \mathbb{E}\left\{\int_{A^\star} [\gamma(\Xi^\star_{\gamma,\delta} + \delta_{\mathbf{z}}) - \gamma(\Xi^\star_{\gamma,\delta})](\boldsymbol{\lambda}^\star(d\mathbf{z}) - \boldsymbol{\lambda}^\star_{\gamma,\delta}(d\mathbf{z}))\right\}.$$

Analogously to (4.2.18) and (4.2.19), the integrand may be bounded by

$$\Delta_1\gamma \int_{A^\star} |\boldsymbol{\lambda}^\star(d\mathbf{z}) - \boldsymbol{\lambda}^\star_{\gamma,\delta}(d\mathbf{z})| \le s_2(h) \min\left\{1,\, \frac{1.65}{\sqrt{\boldsymbol{\lambda}^\star(A^\star)}}\right\} \int_{A^\star} |\boldsymbol{\lambda}^\star(d\mathbf{z}) - \boldsymbol{\lambda}^\star_{\gamma,\delta}(d\mathbf{z})|.$$

Here, $1.65(\boldsymbol{\lambda}^\star(A^\star))^{-1/2}$ may be replaced by $1.65(\boldsymbol{\lambda}^\star_{\gamma,\delta}(A^\star))^{-1/2}$ by going through the same arguments as before, but instead starting with an immigration-death process over $A^\star$ with immigration intensity $\boldsymbol{\lambda}^\star_{\gamma,\delta}$, unit per-capita death rate, and equilibrium distribution $\mathrm{PRM}(\boldsymbol{\lambda}^\star_{\gamma,\delta})$. □

We now again assume that the distributional parameters p_{00n}, q_{1n} and q_{2n} satisfy (4.2.12) and apply Theorem 4.11 to the case where $A^\star = A^\star_n = [u^\star_n, \infty)^2$. We express the error bounds in terms of the threshold $u^\star_n$ and of the probability of simultaneous success p_{11n}.

Proposition 4.12. *Let $p_{11n} \in (0,1)$ and assume that q_{1n}, q_{2n} and p_{00n} satisfy (4.2.12). For any choice of $u_n^\star \geq -\log n$, define $A^\star = [u_n^\star, \infty)^2$. With the notations from Sections 2.4-4.2.3,*

$$(i) \qquad d_{TV}\left(\mathrm{PRM}(\boldsymbol{\lambda}^\star), \mathrm{PRM}(\boldsymbol{\lambda}^\star_{\gamma,\delta})\right) \leq \frac{4(1+\gamma+\delta)^2 p_{11n}}{[1-(1+\gamma+\delta)p_{11n}]^3}\, \mathrm{e}^{-u_n^\star},$$

$$(ii) \qquad d_2\left(\mathrm{PRM}(\boldsymbol{\lambda}^\star), \mathrm{PRM}(\boldsymbol{\lambda}^\star_{\gamma,\delta})\right) \leq \min\left\{\mathrm{e}^{-u_n^\star},\, 1.65\mathrm{e}^{-u_n^\star/2}\right\} \frac{4(1+\gamma+\delta)^2 p_{11n}}{[1-(1+\gamma+\delta)p_{11n}]^3}.$$

Proof. For ease of notation we again omit the subscript n.
(i) By Theorem 4.11,

$$d_{TV}\left(\mathrm{PRM}(\boldsymbol{\lambda}^\star), \mathrm{PRM}(\boldsymbol{\lambda}^\star_{\gamma,\delta})\right) \leq \int_{u^\star}^\infty \int_{u^\star}^t \left|\lambda^\star(s,t) - \lambda^\star_{\gamma,\delta}(s,t)\right| \mathrm{d}s\mathrm{d}t$$
$$+ \int_{u^\star}^\infty \int_{u^\star}^s \left|\lambda^\star(s,t) - \lambda^\star_{\gamma,\delta}(s,t)\right| \mathrm{d}t\mathrm{d}s + \int_{u^\star}^\infty \left|\acute{\lambda}^\star(s) - \acute{\lambda}^\star_{\gamma,\delta}(s)\right| \mathrm{d}s.$$

Define

$$h := h(p_{11}) := \frac{1+\delta}{1+\gamma+\delta} - \frac{\log q_2}{\log p_{00}} \quad \text{and} \quad g := g(p_{11}) := \frac{1+\gamma}{1+\gamma+\delta} - \frac{\log q_1}{\log p_{00}}.$$

We first consider the case $s = t$. Note that, with definitions (4.2.11) and (4.2.13),

$$\acute{\lambda}^\star(s) = \frac{\log(p_{00}/q_1 q_2)}{\log(1/p_{00})}\, \mathrm{e}^{-s} = \left[\frac{\log q_1 + \log q_2}{\log p_{00}} - \frac{2+\gamma+\delta}{1+\gamma+\delta} + \frac{1}{1+\gamma+\delta}\right] \mathrm{e}^{-s}$$
$$= \left[\frac{1}{1+\gamma+\delta} - h(p_{11}) - g(p_{11})\right] \mathrm{e}^{-s},$$

and that, since $h, g \geq 0$ by Lemma 4.7 (i) and (ii), we thus have $\acute{\lambda}^\star(s) \leq \acute{\lambda}^\star_{\gamma,\delta}(s)$. Hence,

$$\int_{u^\star}^\infty \left|\acute{\lambda}^\star(s) - \acute{\lambda}^\star_{\gamma,\delta}(s)\right| \mathrm{d}s = \int_{u^\star}^\infty (h+g)\, \mathrm{e}^{-s}\mathrm{d}s \leq \frac{(\gamma+\delta)p_{11}}{1-(1+\gamma+\delta)p_{11}}\, \mathrm{e}^{-u^\star},$$

again by Lemma 4.7 (i) and (ii). For $s < t$, note that

$$\lambda^\star(s,t) = \left[\frac{\gamma}{1+\gamma+\delta} + h\right]\left[\frac{1+\delta}{1+\gamma+\delta} - h\right] \mathrm{e}^{-\frac{\gamma}{1+\gamma+\delta}\, s}\mathrm{e}^{-\frac{1+\delta}{1+\gamma+\delta}\, t}\mathrm{e}^{h(t-s)}$$
$$= \lambda^\star_{\gamma,\delta}(s,t)\mathrm{e}^{h(t-s)} + \left(\frac{1+\delta-\gamma}{1+\gamma+\delta}\, h - h^2\right) \mathrm{e}^{-\frac{\gamma}{1+\gamma+\delta}\, s}\mathrm{e}^{-\frac{1+\delta}{1+\gamma+\delta}\, t}\mathrm{e}^{h(t-s)},$$

where $\lambda^\star(s,t)$ and $\lambda^\star_{\gamma,\delta}(s,t)$ are defined by (4.2.11) and (4.2.13), respectively. Thereby,

$$\left|\lambda^\star(s,t) - \lambda^\star_{\gamma,\delta}(s,t)\mathrm{e}^{h(t-s)}\right| = \left|\frac{1+\delta-\gamma}{1+\gamma+\delta}\, h - h^2\right| \mathrm{e}^{-\frac{\gamma}{1+\gamma+\delta}\, s}\mathrm{e}^{-\frac{1+\delta}{1+\gamma+\delta}\, t}\mathrm{e}^{h(t-s)},$$

where $\left|\frac{1+\delta-\gamma}{1+\gamma+\delta}\, h - h^2\right| \leq h + h^2$. Note that we have

$$\left|\lambda^\star(s,t) - \lambda^\star_{\gamma,\delta}(s,t)\right| \leq \left|\lambda^\star(s,t) - \lambda^\star_{\gamma,\delta}(s,t)\mathrm{e}^{h(t-s)}\right| + \lambda^\star_{\gamma,\delta}(s,t)\left|\mathrm{e}^{h(t-s)} - 1\right|. \tag{4.2.27}$$

We first compute the following integral:

$$\int_{u^\star}^{\infty}\int_{u^\star}^{t} \mathrm{e}^{-\frac{\gamma}{1+\gamma+\delta}\,s-\frac{1+\delta}{1+\gamma+\delta}\,t+h(t-s)}\mathrm{d}s\mathrm{d}t = \int_{u^\star}^{\infty}\int_{u^\star}^{t} \mathrm{e}^{-\frac{\log(p_{00}/q_2)}{\log p_{00}}\,s-\frac{\log q_2}{\log p_{00}}\,t}\mathrm{d}s\mathrm{d}t = \frac{\log p_{00}}{\log q_2}\,\mathrm{e}^{-u^\star}, \quad (4.2.28)$$

where, using $z \le -\log(1-z) \le \frac{z}{1-z}$ for all $|z| \le 1$, and (4.2.12),

$$\begin{aligned} \frac{\log p_{00}}{\log q_2} &= \frac{-\log[1-(1+\gamma+\delta)p_{11}]}{-\log[1-(1+\delta)p_{11}]} \le \frac{1+\gamma+\delta}{(1+\delta)[1-(1+\gamma+\delta)p_{11}]} \\ &\le \frac{1+\gamma+\delta}{1+\delta}\left\{1+\frac{(1+\gamma+\delta)p_{11}}{[1-(1+\gamma+\delta)p_{11}]^2}\right\}. \end{aligned} \quad (4.2.29)$$

Moreover, by Lemma 4.7 (i),

$$h + h^2 \le \frac{\gamma p_{11}}{1-(1+\gamma+\delta)p_{11}} + \left[\frac{\gamma p_{11}}{1-(1+\gamma+\delta)p_{11}}\right]^2 \le \frac{2\gamma p_{11}}{[1-(1+\gamma+\delta)p_{11}]^2},$$

since $\gamma p_{11} = p_{10} < 1$, and therefore $(\gamma p_{11})^2 \le \gamma p_{11}$. Then,

$$\int_{u^\star}^{\infty}\int_{u^\star}^{t} \left|\lambda^\star(s,t) - \lambda^\star_{\gamma,\delta}(s,t)\mathrm{e}^{h(t-s)}\right| \mathrm{d}s\mathrm{d}t \le \frac{2\gamma(1+\gamma+\delta)p_{11}}{(1+\delta)[1-(1+\gamma+\delta)p_{11}]^3}\,\mathrm{e}^{-u^\star}, \quad (4.2.30)$$

which gives a bound for the integral of the first error term in (4.2.27). For the second error term in (4.2.27), note first that $|\mathrm{e}^{h(t-s)}-1| = \mathrm{e}^{h(t-s)}-1$, since $h \ge 0$ and $t > s$. By (4.2.28) and (4.2.29), and with definition (4.2.13) of $\lambda^\star_{\gamma,\delta}(s,t)$, we obtain

$$\begin{aligned} \int_{u^\star}^{\infty}\int_{u^\star}^{t} \lambda^\star_{\gamma,\delta}(s,t)\mathrm{e}^{h(t-s)} &= \frac{\gamma(1+\delta)\log p_{00}}{(1+\gamma+\delta)^2\log q_2}\,\mathrm{e}^{-u^\star} \\ &\le \frac{\gamma}{1+\gamma+\delta}\left\{1+\frac{(1+\gamma+\delta)p_{11}}{[1-(1+\gamma+\delta)p_{11}]^2}\right\}\mathrm{e}^{-u^\star}, \end{aligned} \quad (4.2.31)$$

whereas

$$\int_{u^\star}^{\infty}\int_{u^\star}^{t} \lambda^\star_{\gamma,\delta}(s,t)\mathrm{d}s\mathrm{d}t = \frac{\gamma}{1+\gamma+\delta}\,\mathrm{e}^{-u^\star}. \quad (4.2.32)$$

By (4.2.31) and (4.2.32), we may thus bound the integral of the second error term in (4.2.27) as follows:

$$\int_{u^\star}^{\infty}\int_{u^\star}^{t} \lambda^\star_{\gamma,\delta}(s,t)\left|\mathrm{e}^{h(t-s)}-1\right| \mathrm{d}s\mathrm{d}t \le \frac{\gamma p_{11}}{[1-(1+\gamma+\delta)p_{11}]^2}\,\mathrm{e}^{-u^\star}. \quad (4.2.33)$$

Hence, for $s < t$, (4.2.27), (4.2.30) and (4.2.33) give

$$\int_{u^\star}^{\infty}\int_{u^\star}^{t} \left|\lambda^\star(s,t) - \lambda^\star_{\gamma,\delta}(s,t)\right| \mathrm{d}s\mathrm{d}t \le \frac{\gamma p_{11}}{[1-(1+\gamma+\delta)p_{11}]^3}\left\{\frac{2(1+\gamma+\delta)}{1+\delta}+1\right\}\mathrm{e}^{-u^\star}.$$

By proceeding analogously for $s > t$, we obtain

$$\int_{u^\star}^{\infty}\int_{u^\star}^{s} \left|\lambda^\star(s,t) - \lambda^\star_{\gamma,\delta}(s,t)\right| \mathrm{d}t\mathrm{d}s \le \frac{\delta p_{11}}{[1-(1+\gamma+\delta)p_{11}]^3}\left\{\frac{2(1+\gamma+\delta)}{1+\gamma}+1\right\}\mathrm{e}^{-u^\star}.$$

The sum of the bounds for the three cases $s = t$, $s < t$ and $s > t$ yields

$$\int_{A^\star} \left|\boldsymbol{\lambda}^\star(d\mathbf{z}) - \boldsymbol{\lambda}^\star_{\gamma,\delta}(d\mathbf{z})\right| \\ \leq \frac{2p_{11}}{[1-(1+\gamma+\delta)p_{11}]^3}\left\{\frac{\gamma(1+\gamma+\delta)}{1+\delta} + \frac{\delta(1+\gamma+\delta)}{1+\gamma} + \gamma + \delta\right\}\mathrm{e}^{-u^\star} \leq \frac{4(1+\gamma+\delta)^2 p_{11}}{[1-(1+\gamma+\delta)p_{11}]^3}\,\mathrm{e}^{-u^\star},$$

where we used $(1+\gamma)^{-1}, (1+\delta)^{-1} < 1$, and $\gamma + \delta \leq 1 + \gamma + \delta \leq (1+\gamma+\delta)^2$ for the second inequality.
(ii) Direct computations give $\boldsymbol{\lambda}^\star(A^\star) = \mathrm{e}^{-u^\star} = \boldsymbol{\lambda}^\star_{\gamma,\delta}(A^\star)$. Theorem 4.11 (ii), together with the bound from (i), then immediately gives the result. □

The error bounds established in Proposition 4.12 are similar to the error bound from Proposition 4.10. As before, p_{11n} needs to converge to 0 fast enough to make up for the factor $\mathrm{e}^{-u_n^\star}$ which increases the size of the error as soon as $u_n^\star < 0$. And since $u_n^\star \geq 0$ gives 1 or no points in $A^\star$, the mean number of points in $A^\star$ being given by $\mathrm{e}^{-u_n^\star}$ for either process, we would want the threshold $u_n^\star$ to be negative.

The biggest difference between the d_2-bounds from Propositions 4.10 and 4.12 is that the former contains the multiplicative factor $[1-(1+\gamma+\delta)p_{11n}]^{-2}$ and the latter the bigger factor $[1-(1+\gamma+\delta)p_{11n}]^{-3}$. However, since we need $p_{11n} \to 0$ as $n \to \infty$, we will have $(1+\gamma+\delta)p_{11n} \leq 1/2$ for all n large enough. Then $[1-(1+\gamma+\delta)p_{11n}]^{-3} \leq 2[1-(1+\gamma+\delta)p_{11n}]^{-2}$ so that both error bounds will be of the same rate. Hence, for large enough n, the approximation by a further Poisson process does not add an error of a bigger size than the one that arises from the approximation by only $\mathrm{PRM}(\boldsymbol{\lambda}^\star)$.

4.2.6 Final bound in the d_2-distance

The following corollary summarises the results from Sections 4.2.1, 4.2.4 and 4.2.5. It gives an estimate for the error in the d_2-distance of the approximation of the law of an MPPE $\Xi^\star_{A^\star}$ with i.i.d. Marshall-Olkin geometric marks, living on a lattice of points contained in $A^\star \cap [-\log n, \infty)^2$, by the law of a Poisson process with a continuous intensity measure $\boldsymbol{\lambda}^\star_{\gamma,\delta}$ over $A^\star \cap [-\log n, \infty)^2$, where $A^\star = [u^\star, \infty)^2$ for some choice of threshold $u^\star \geq -\log n$.

Corollary 4.13. *Let $p_{11n} \in (0,1)$ and assume that q_{1n}, q_{2n} and p_{00n} satisfy (4.2.12). For any choice of $u_n^\star \geq -\log n$, define $A^\star = [u_n^\star, \infty)^2$. With the notations from Sections 2.4 and 4.2.1-4.2.3,*

$$d_2\left(\mathcal{L}(\Xi^\star_{A^\star}), \mathrm{PRM}(\boldsymbol{\lambda}^\star_{\gamma,\delta})\right) \leq \frac{\mathrm{e}^{-u_n^\star}}{n} + \frac{(1+\gamma+\delta)^2 p_{11n}}{[1-(1+\gamma+\delta)p_{11n}]^3}\left\{2\sqrt{2} + 7\min\left\{\mathrm{e}^{-u_n^\star},\ 1.65\mathrm{e}^{-u_n^\star/2}\right\}\right\}.$$

Proof. We have

$$d_2\left(\mathcal{L}(\Xi^\star_{A^\star}), \mathrm{PRM}(\boldsymbol{\lambda}^\star_{\gamma,\delta})\right) \\ \leq d_2\left(\mathcal{L}(\Xi^\star_{A^\star}), \mathrm{PRM}(\boldsymbol{\pi}^\star)\right) + d_2\left(\mathrm{PRM}(\boldsymbol{\pi}^\star), \mathrm{PRM}(\boldsymbol{\lambda}^\star)\right) + d_2\left(\mathrm{PRM}(\boldsymbol{\lambda}^\star), \mathrm{PRM}(\boldsymbol{\lambda}^\star_{\gamma,\delta})\right).$$

By Theorem 4.3,

$$d_2\left(\mathcal{L}(\Xi^\star_{A^\star}), \mathrm{PRM}(\boldsymbol{\pi}^\star)\right) \leq d_{TV}\left(\mathcal{L}(\Xi^\star_{A^\star}), \mathrm{PRM}(\boldsymbol{\pi}^\star)\right) \leq \frac{\mathrm{e}^{-u_n^\star}}{n}.$$

Furthermore, with the results from Propositions 4.10 and 4.12, and using $(1+\gamma+\delta) \leq (1+\gamma+\delta)^2$ and $[1-(1+\gamma+\delta)p_{11}]^{-2} \leq [1-(1+\gamma+\delta)p_{11}]^{-3}$, we obtain

$$\begin{aligned} &d_2\left(\mathrm{PRM}(\boldsymbol{\pi}^\star), \mathrm{PRM}(\boldsymbol{\lambda}^\star)\right) + d_2\left(\mathrm{PRM}(\boldsymbol{\lambda}^\star), \mathrm{PRM}(\boldsymbol{\lambda}^\star_{\gamma,\delta})\right) \\ &\qquad \leq \frac{(1+\gamma+\delta)^2 p_{11n}}{[1-(1+\gamma+\delta)p_{11n}]^3} \left\{2\sqrt{2} + 7\min\left\{\mathrm{e}^{-u_n^\star},\ 1.65\mathrm{e}^{-u_n^\star/2}\right\}\right\}. \end{aligned}$$

□

By far the smallest component of the error estimate from Corollary 4.13 is given by $\mathrm{e}^{-u_n^\star}/n$, the error arising from approximating $\mathcal{L}(\Xi^\star_{A^\star})$ by $\mathrm{PRM}(\mathbb{E}\Xi^\star_{A^\star})$, which lives on the lattice $A^\star \cap E^\star$ just as $\Xi^\star_{A^\star}$. A far bigger error emerges for the MPPE with Marshall-Olkin geometric marks when going from the Poisson process on the lattice to a Poisson process on $A^\star \cap [-\log n, \infty)^2$ with continuous intensity. This error can only be small if the probability of simultaneous success of the Marshall-Olkin geometric distribution, p_{11}, and thereby also the marginal success probabilities $1-q_{1n}$ and $1-q_{2n}$, tend to zero as $n \to \infty$ at a rate fast enough to compensate for the factor $\mathrm{e}^{-u_n^\star}$, the (rough) number of points expected in $A_n^\star$ for each of the processes. For instance, for $A_n^\star = [-\log\log n, \infty)^2$ and $p_{11n} = 1/n$, we expect $\log n$ joint threshold exceedances, and obtain

$$\begin{aligned} &d_2\left(\mathcal{L}(\Xi^\star_{A^\star}), \mathrm{PRM}(\boldsymbol{\lambda}^\star_{\gamma,\delta})\right) \\ &\leq \frac{\log n}{n} + \frac{(1+\gamma+\delta)^2}{n[1-(1+\gamma+\delta)/n]^3}\left\{2\sqrt{2} + 7\min\left\{\log n,\ 1.65\sqrt{\log n}\right\}\right\} \leq \frac{C\log n}{n}, \end{aligned}$$

where C is some constant. With the (very strong) condition $p_{11n} = 1/n$, we thus obtain an error of the same size as the error that we obtain when approximating $\mathcal{L}(\Xi^\star_{A^\star})$ only by $\mathrm{PRM}(\mathbb{E}\Xi^\star_{A^\star})$.

References

Anderson, C. W., Coles, S. G. and Hüsler, J. (1997). Maxima of Poisson-like variables and related triangular arrays, *Ann. Appl. Probab.* **7**: 953–971.

Barbour, A. D. (1997). Stein's method, *Encyclopedia of statistical sciences, Update Volume 1*, Wiley, New York, pp. 513–521.

Barbour, A. D. and Brown, T. C. (1992). Stein's method and point process approximation, *Stochastic Process. Appl.* **43**(1): 9–31.

Barbour, A. D. and Hall, P. (1984). On the rate of Poisson convergence, *Math. Proc. Cambridge Philos. Soc.* **95**(3): 473–480.
URL: *http://dx.doi.org/10.1017/S0305004100061806*

Barbour, A. D., Holst, L. and Janson, S. (1992). *Poisson approximation*, The Clarendon Press Oxford University Press.

Chen, L. H. Y. (1975). An approximation theorem for sums of certain randomly selected indicators, *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **33**(1): 69–74.

Feidt, A., Genest, C. and Nešlehová, J. (2010). Asymptotics of joint maxima for discontinuous random variables, *Extremes* **13**: 35–53.

Fisher, R. A. and Tippett, L. H. C. (1928). Limiting forms of the frequency distributions of the largest or smallest member of a sample, *Proceedings of the Cambridge Philosophical Society* **24**: 180–190.

Hawkes, A. G. (1972). A bivariate exponential distribution with applications to reliability, *Journal of the Royal Statistical Society. Series B (Methodological)* **34**(1): pp. 129–131.
URL: *http://www.jstor.org/stable/2985057*

Leadbetter, M. R., Lindgren, G. and Rootzén, H. (1983). *Extremes and Related Properties of Random Sequences and Processes*, Springer, New York.

Marshall, A. W. and Olkin, I. (1967). A multivariate exponential distribution, *Journal of the American Statistical Association* **62**(317): pp. 30–44.

Marshall, A. W. and Olkin, I. (1985). A family of bivariate distributions generated by the bivariate Bernoulli distribution, *J. Amer. Statist. Assoc.* **80**: 332–338.

Michel, R. (1988). An improved error bound for the compound poisson approximation of a nearly homogeneous portfolio, *Astin Bull.* **17**: 165–169.

Mitov, K. and Nadarajah, S. (2002). Asymptotics of maxima of discrete random variables, *Extremes* **5**: 287–294.

Mitov, K. and Nadarajah, S. (2005). Limit distributions for the bivariate geometric maxima, *Extremes* **8**: 357–370.

Nelsen, R. B. (2006). *An Introduction to Copulas,* Second Edition, Springer, New York.

Resnick, S. I. (1987). *Extreme Values, Regular Variation, and Point Processes*, Springer, New York.