Robust estimation on Controlled Branching Processes: minimum disparity approach

M. González, C. Minuesa, I. del Puerto

III Workshop
On Branching Processes and their Applications
April, 7-10, 2015
Badajoz, Spain

Spanish Branching Processes Group

Department of Mathematics
Faculty of Sciences
University of Extremadura. Spain

Contents

Contents

Contents

Contents

Contents

Probability model

Definition (Yanev (1975))

Let $\{X_{ni} : n = 0, 1, \ldots; i = 1, 2, \ldots\}$ and $\{\phi_n(k) : n, k = 0, 1, \ldots\}$ be two independent families of non negative integer valued random variables which are defined on the same probability space, $(\Omega, \mathcal{A}, P)$.

(i) $\{X_{ni} : n = 0, 1, \ldots; i = 1, 2, \ldots\}$ are i.i.d. random variables whose distribution is denoted by $p = \{p_k\}_{k\geq 0}$, $p_k = P[X_{01} = k]$, $k \geq 0$.

(ii) For each $n = 0, 1, \ldots$, $\{\phi_n(k) : k = 0, 1, \ldots\}$ are independent stochastic processes with equal one-dimensional probability distributions, i.e., for each n, $p_j(k) = P[\phi_n(k) = j]$, $j, k \geq 0$.

The stochastic process $\{Z_n\}_{n\geq 0}$ defined as:

$$Z_0 = N \geq 0, \quad Z_{n+1} = \sum_{i=1}^{\phi_n(Z_n)} X_{ni}, \quad n = 0, 1, \quad \left(\sum_1^0 = 0\right),$$

is known as **Controlled Branching Process (CBP) with random control function**.

Probability model

Examples of CBPs

- CBP with determinist control (Sebast'yanov and Zubkov (1974)).
- Standard branching process (Bienaymè-Galton-Watson).
- Branching process with immigration (Sriram (1994)).
- Branching process with immigration at state zero (Bruss and Slavtchova-Bojkova (1999)).
- Branching process with random migration (Yanev and Yanev (1996))
- Branching process with bounded emigration (del Puerto and Yanev (2008)).
- Branching process with adaptive control (Bercu (1999)).
- Branching process with continuous state space (Rahimov and Al-Sabah (2007)).

Probability model

Main parameters of the model

- $p = \{p_k\}_{k\geq 0}$: **offspring distribution** or **reproduction law**.
- $m = E[X_{01}]$: **offspring mean**.
- $\sigma^2 = Var[X_{01}]$: **offspring variance**.

- $\{p_j(k)\}_{k,j\geq 0}$: **control law**.
- $\varepsilon(k) = E[\phi_0(k)], \quad k = 0, 1, \ldots$: **control mean**.
- $\sigma^2(k) = Var[\phi_0(k)], \quad k = 0, 1, \ldots$: **control variance**.

Probability model

Main parameters of the model

- $p = \{p_k\}_{k\geq 0}$: **offspring distribution** or **reproduction law**.
- $m = E[X_{01}]$: **offspring mean**.
- $\sigma^2 = Var[X_{01}]$: **offspring variance**.

- $\{p_j(k)\}_{k,j\geq 0}$: **control law**.
- $\varepsilon(k) = E[\phi_0(k)], \quad k = 0, 1, \ldots$: **control mean**.
- $\sigma^2(k) = Var[\phi_0(k)], \quad k = 0, 1, \ldots$: **control variance**.

Main investigated topics

Properties

- $\{Z_n\}_{n\geq 0}$ is an homogeneous Markov Chain with stationary transition probabilities.
- Duality Extinction-Explosion: $P(Z_n \to 0) + P(Z_n \to \infty) = 1$.

Main investigated topics

- Extinction Problem:
  - Sevastyanov and Zubkov (1974).
  - Zubkov (1974).
  - Yanev (1975).
  - González, Molina, and del Puerto (2002, 2005a).
- Asymptotic Behaviour. Growth rates:
  - Bagley (1986).
  - González, Molina, and del Puerto (2002, 2003, 2005a,b).

Main investigated topics

Properties

- $\{Z_n\}_{n\geq 0}$ is an homogeneous Markov Chain with stationary transition probabilities.
- Duality Extinction-Explosion: $P(Z_n \to 0) + P(Z_n \to \infty) = 1$.

Main investigated topics

- Extinction Problem:
  - Sevastyanov and Zubkov (1974).
  - Zubkov (1974).
  - Yanev (1975).
  - González, Molina, and del Puerto (2002, 2005a).
- Asymptotic Behaviour. Growth rates:
  - Bagley (1986).
  - González, Molina, and del Puerto (2002, 2003, 2005a,b).

Main investigated topics

Main investigated topics: Statistical Inference

- Bercu, B. (1999) Weighted estimation and tracking for Bienaymè-Galton-Watson processes with adaptive control. *Statistics & Probability Letters*, **42**, 415-421.
- Dion, J. P. and Essebbar, B. (1995). On the statistics of controlled branching processes. *Lecture Notes in Statistics*, **99**, 14-21.
- Mohan, K. (2000) Estimation of parameters in controlled branching process. *Stochastic Modelling and Applications*, **3**, 1-15.
- M. González, R. Martínez, I. and del Puerto (2004). Nonparametric estimation of the offspring distribution and the mean for a controlled branching process. *Test*, **13(2)**, 465-479.
- M. González, R. Martínez, I. and del Puerto (2005). Estimation of the variance for a controlled branching process. *Test*, **14(1)**, 199-213.
- T.N. Sriram, A. Bhattacharya, M. González, R. Martínez, and I. del Puerto (2007). Estimation of the offspring mean in a controlled branching process with a random control function. *Stochastic Processes and their Applications*, **117**, 928-946.
- R. Martínez, I. del Puerto and M. Mota (2009). On asymptotic posterior normality for controlled branching processes. *Statistics*, **43**, 367-378.
- M. González and I. del Puerto (2012). Diffusion Approximation of an Array of Controlled Branching Processes. *Methodology and Computing in Applied Probability*, **14**, 843-861.
- M. González, C. Gutiérrez, R. Martínez, and I. del Puerto (2013). Bayesian inference for controlled branching processes through MCMC and ABC methodologies. *Revista de la Real Academia de Ciencias Exactas, Físicas y Naturales. Serie A. Matemáticas*, **107(2)**, 459-473.
- González, M., M.C., del Puerto, I. (2015). Maximum likelihood estimation and Expectation - Maximization algorithm for controlled branching processes. Computational Statistics and Data Analysis, DOI: 10.1016/j.csda.2015.01.015

The problem

Assumption

The offspring distribution belongs to a parametric family

$$\mathcal{F}_\theta = \{p(\theta) : \theta \in \Theta\}, \qquad \Theta \subseteq \mathbb{R},$$

that is, $p = p(\theta_0)$, with $\theta_0 \in \Theta$.

Aim

To obtain a robust estimator of θ_0 and, in consequence, of $p(\theta_0)$, $m(\theta_0)$ and $\sigma^2(\theta_0)$.

The problem

Assumption

The offspring distribution belongs to a parametric family

$$\mathcal{F}_\theta = \{p(\theta) : \theta \in \Theta\}, \qquad \Theta \subseteq \mathbb{R},$$

that is, $p = p(\theta_0)$, with $\theta_0 \in \Theta$.

Aim

To obtain a robust estimator of θ_0 and, in consequence, of $p(\theta_0)$, $m(\theta_0)$ and $\sigma^2(\theta_0)$.

The problem

Maximum likelihood estimation in a nonparametric context

Assuming that we can observe the random variables

$$\mathcal{Z}_n^* = \left\{ Z_l(k) = \sum_{i=1}^{\phi_l(Z_l)} I_{\{X_{li}=k\}} : k \geq 0; l = 0, \dots, n-1 \right\},$$

it is proved that the maximum likelihood estimator (MLE) of p_k, $k \geq 0$, m, and σ^2 are:

$$\hat{p}_k = \frac{\sum_{l=0}^{n-1} Z_l(k)}{\sum_{l=0}^{n-1} \phi_l(Z_l)}, \quad \hat{m}_n = \frac{\sum_{l=0}^{n-1} Z_{l+1}}{\sum_{l=0}^{n-1} \phi_l(Z_l)}, \quad \hat{\sigma}_n^2 = \sum_{k=0}^{\infty} (k - \hat{m}_n)^2 \hat{p}_k.$$

- González, M., M.C., del Puerto, I. (2015). CSDA.

The problem

Maximum likelihood estimation in a parametric context

Assuming that $p \in \mathcal{F}_\Theta$ and that we observe $\mathcal{Z}_n^*$, one can obtain the maximum likelihood estimator of θ_0 by maximizing the log-likelihood

$$\ell(\theta \mid \mathcal{Z}_n^*) = \sum_{l=0}^{n-1} \log\left(\frac{\phi_l^*!}{\prod_{k=0}^{\infty} Z_l(k)!}\right) + \sum_{l=0}^{n-1} \log(p_{\phi_l^*}(Z_l)) + \sum_{l=0}^{n-1}\sum_{k=0}^{\infty} Z_l(k)\log(p_k(\theta)).$$

The problem

Maximum likelihood estimation in a parametric context

Assuming that $p \in \mathcal{F}_\Theta$ and that we observe $\mathcal{Z}_n^*$, one can obtain the maximum likelihood estimator of θ_0 by maximizing the log-likelihood

$$\ell(\theta \mid \mathcal{Z}_n^*) = \sum_{l=0}^{n-1} \log\left(\frac{\phi_l^*!}{\prod_{k=0}^{\infty} Z_l(k)!}\right) + \sum_{l=0}^{n-1} \log(p_{\phi_l^*}(Z_l)) + \sum_{l=0}^{n-1}\sum_{k=0}^{\infty} Z_l(k)\log(p_k(\theta)).$$

Thus,

$$\hat{\theta}_n = \arg\max_{\theta\in\Theta} \sum_{l=0}^{n-1}\sum_{k=0}^{\infty} Z_l(k)\log(p_k(\theta))$$

The problem

Maximum likelihood estimation in a parametric context

Assuming that $p \in \mathcal{F}_\Theta$ and that we observe $\mathcal{Z}_n^*$, one can obtain the maximum likelihood estimator of θ_0 by maximizing the log-likelihood

$$\ell(\theta \mid \mathcal{Z}_n^*) = \sum_{l=0}^{n-1} \log\left(\frac{\phi_l^*!}{\prod_{k=0}^{\infty} Z_l(k)!}\right) + \sum_{l=0}^{n-1} \log(p_{\phi_l^*}(Z_l)) + \sum_{l=0}^{n-1}\sum_{k=0}^{\infty} Z_l(k)\log(p_k(\theta)).$$

Thus,

$$\hat{\theta}_n = \arg\max_{\theta\in\Theta} \sum_{l=0}^{n-1}\sum_{k=0}^{\infty} Z_l(k)\log(p_k(\theta)) = \arg\max_{\theta\in\Theta} \sum_{k=0}^{\infty} \hat{p}_{n,k}\log(p_k(\theta))$$

$$\widehat{p(\theta_0)}_{n,k} = p_k(\hat{\theta}_n), \quad \widehat{m(\theta_0)}_n = m(\hat{\theta}_n), \quad \widehat{\sigma(\theta_0)}_n^2 = \sigma(\hat{\theta}_n)^2.$$

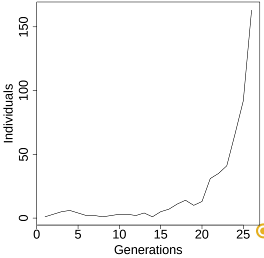
Simulated example

Behaviour of the MLEs against model perturbations

- **Parametric family**: $\mathcal{F}_\theta = \{\mathcal{P}(5,\theta) : \theta \in (0,\infty)\}$,
  where $\mathcal{P}(5,\theta)$ denotes a Poisson distribution with parameter θ truncated at 5.
- **Mixture model for gross errors**: $p(\theta,\alpha,L) = (1-\alpha)p(\theta) + \alpha\delta_L$,
  where $p(\theta)$ is the probability mass function of $\mathcal{P}(5, 1.2)$, $\alpha = 0.1$ and $L = 5$.

We have simulated the first 25 generations of a CBP which verifies:

- It starts with $Z_0 = 1$ individual.
- The distribution of the variables X_{ij} is $p(\theta,\alpha,L)$, for $i = 0,1,\ldots$, $j = 1,\ldots$.
- $\phi_n(k) \sim B(k,q)$, with $q = 0.9$, $k \geq 0$.



Simulated example

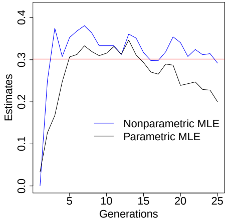


Fig: MLEs of p_0.

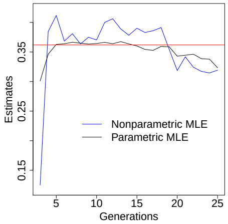


Fig: MLEs of p_1.

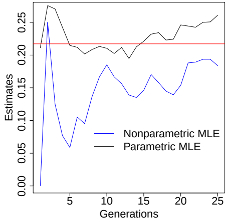


Fig: MLEs of p_2.

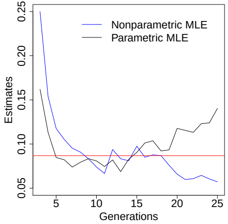


Fig: MLEs of p_3.

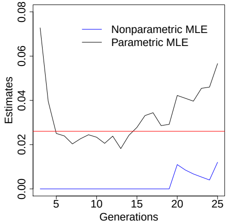


Fig: MLEs of p_4.

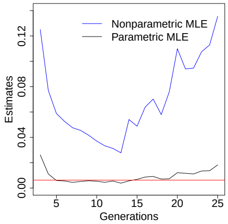


Fig: MLEs of p_5.

Minimum disparity estimation

Minimum Hellinger distance estimation

- Beran, R.J. (1977). Minimum Hellinger distance estimates for parametric models. *Annals of Statistics*, **5**, 445-463.
- Sriram, T.N, Vidyashankar, A.N. (2000). Minimum Hellinger distance estimation for supercritical Galton-Watson Process. *Statistics and Probability Letters*, **50**, 331-342.

Minimum disparity estimation

Minimum disparity estimation

- Lindsay, B. G. (1994). Efficiency versus robustness: The case for minimum Hellinger distance and related methods. *The Annals of Statistics*, **22**, 1081-1114.
- Basu, A., Sarkar, S., Vidyashankar, A. (1997). Minimum negative exponential disparity estimation in parametric models. *Journal of Statistical Planning and Inference*, **58**, 349-370.
- Park, C., Basu, A. (2004). Minimum disparity estimation: asymptotic normality and breakdown point results. *Bulletin of Informatics and Cybernetics*, **36**, 19-33.
- Hooker, G., Vidyashankar, A.N. (2014). Bayesian model robustness via disparities. *Test*, **23(3)**, 556-584.

Minimum disparity estimation

Disparity measure

Let Γ be the set of all probability mass functions defined on non-negative integers. A **disparity measure** between $q \in \Gamma$ and $p(\theta) \in \mathcal{F}_\theta$ is defined by:

$$\rho(q, \theta) = \sum_{k=0}^{\infty} G(\delta(q, \theta, k)) p_k(\theta),$$

with $G(\cdot)$ a three times differentiable and strictly convex function on $[-1, \infty)$ with $G(0) = 0$ and

$$\delta(q, \theta, k) = \frac{q_k}{p_k(\theta)} - 1 \quad (\textit{Pearson residual}).$$

Our aim:

To determine

$$\arg\min_{\theta \in \Theta} \rho(p, \theta).$$

Minimum disparity estimation

Disparity measure

Let Γ be the set of all probability mass functions defined on non-negative integers. A **disparity measure** between $q \in \Gamma$ and $p(\theta) \in \mathcal{F}_\theta$ is defined by:

$$\rho(q, \theta) = \sum_{k=0}^{\infty} G(\delta(q, \theta, k)) p_k(\theta),$$

with $G(\cdot)$ a three times differentiable and strictly convex function on $[-1, \infty)$ with $G(0) = 0$ and

$$\delta(q, \theta, k) = \frac{q_k}{p_k(\theta)} - 1 \quad (\textit{Pearson residual}).$$

Our aim: the minimum disparity estimator (MDE)

To determine

$$\tilde{\theta}_n^{\rho}(\tilde{p}_n) = \arg\min_{\theta \in \Theta} \rho(\tilde{p}_n, \theta).$$

Minimum disparity estimation

$$-\frac{\partial \rho}{\partial \theta}(\tilde{p}_n, \theta) = \sum_{k=0}^{\infty} p_k'(\theta) A(\delta(\tilde{p}_n, \theta, k)) = 0,$$

with

$$A(\delta) = (\delta + 1) G'(\delta) - G(\delta) \quad \textit{(RAF)}.$$

Examples of disparity measures

Disparity measure	MDE	Notation	$G(\delta)$	$A(\delta)$
Likelihood disparity	MLDE	$LD(\tilde{p}_n, \theta)$	$(\delta+1)\log(\delta+1)$	δ
Squared Hellinger distance	MHDE	$HD(\tilde{p}_n, \theta)$	$[(\delta+1)^{1/2}-1]^2$	$2[(\delta+1)^{1/2}-1]$
Negative exponential disparity	MNEDE	$NED(\tilde{p}_n, \theta)$	$\exp(-\delta)-1$	$1-(2+\delta)\exp(-\delta)$

Minimum disparity estimation

$$-\frac{\partial \rho}{\partial \theta}(\tilde{p}_n, \theta) = \sum_{k=0}^{\infty} p_k'(\theta) A(\delta(\tilde{p}_n, \theta, k)) = 0,$$

with

$$A(\delta) = (\delta + 1)G'(\delta) - G(\delta) \quad (RAF).$$

Examples of disparity measures

Disparity measure	MDE	Notation	$G(\delta)$	$A(\delta)$
Likelihood disparity	MLDE	$LD(\tilde{p}_n, \theta)$	$(\delta + 1)\log(\delta + 1)$	δ
Squared Hellinger distance	MHDE	$HD(\tilde{p}_n, \theta)$	$[(\delta + 1)^{1/2} - 1]^2$	$2[(\delta + 1)^{1/2} - 1]$
Negative exponential disparity	MNEDE	$NED(\tilde{p}_n, \theta)$	$\exp(-\delta) - 1$	$1 - (2 + \delta)\exp(-\delta)$

Minimum disparity estimation

Disparity functional

- The **disparity functional** is a functional $T^{\rho} : \Gamma \to \Theta$ satisfying the condition that for every $q \in \Gamma$

$$T^{\rho}(q) = \arg\min_{\theta \in \Theta} \rho(q, \theta)$$

  if $T^{\rho}(q)$ exists.

- $\tilde{\theta}_n^{\rho}(\tilde{p}_n) = T^{\rho}(\tilde{p}_n)$.

Minimum disparity estimation

Disparity functional

- The **disparity functional** is a functional $T^\rho : \Gamma \to \Theta$ satisfying the condition that for every $q \in \Gamma$

$$T^\rho(q) = \arg\min_{\theta\in\Theta} \rho(q,\theta)$$

  if $T^\rho(q)$ exists.

- $\tilde{\theta}_n^\rho(\tilde{p}_n) = T^\rho(\tilde{p}_n)$.

Minimum disparity estimation

- **Problem**: the existence of

$$min_{\theta\in\Theta}\rho(q,\theta).$$

- **Assumptions**: $\rho(q,\cdot)$ is continuous in Θ and Θ is a compact set.

Theorem 1: Existence and uniqueness of a disparity functional

Under the conditions $\rho(q,\theta)$ is continuous in θ, Θ is a compact set and the identifiability of $\mathcal{F}_\theta$, the existence and uniqueness the functional T^ρ are verified.

Theorem 2: Continuity of a disparity functional

Under the conditions of Theorem 1, if

$$\sup_{\theta\in\Theta}|\rho(q_n,\theta)-\rho(q,\theta)|\to 0,$$

as $q_n\to q$ in ℓ_1 then, $T^\rho(q_n)\to T^\rho(q)$, that is, the continuity of the functional T^ρ holds.

Minimum disparity estimation

- **Problem**: the existence of

$$min_{\theta \in \Theta} \rho(q, \theta).$$

- **Assumptions**: $\rho(q, \cdot)$ is continuous in Θ and Θ is a compact set.

Theorem 1: Existence and uniqueness of a disparity functional

Under the conditions $\rho(q, \theta)$ is continuous in θ, Θ is a compact set and the identifiability of $\mathcal{F}_\theta$, the existence and uniqueness the functional T^ρ are verified.

Theorem 2: Continuity of a disparity functional

Under the conditions of Theorem 1, if

$$\sup_{\theta \in \Theta} |\rho(q_n, \theta) - \rho(q, \theta)| \to 0,$$

as $q_n \to q$ in ℓ_1 then, $T^\rho(q_n) \to T^\rho(q)$, that is, the continuity of the functional T^ρ holds.

Minimum disparity estimation

- **Problem**: the existence of

$$min_{\theta \in \Theta} \rho(q, \theta).$$

- **Assumptions**: $\rho(q, \cdot)$ is continuous in Θ and Θ is a compact set.

Theorem 1: Existence and uniqueness of a disparity functional

Under the conditions $\rho(q, \theta)$ is continuous in θ, Θ is a compact set and the identifiability of $\mathcal{F}_\theta$, the existence and uniqueness the functional T^ρ are verified.

Theorem 2: Continuity of a disparity functional

Under the conditions of Theorem 1, if

$$\sup_{\theta \in \Theta} |\rho(q_n, \theta) - \rho(q, \theta)| \to 0,$$

as $q_n \to q$ in ℓ_1 then, $T^\rho(q_n) \to T^\rho(q)$, that is, the continuity of the functional T^ρ holds.

Minimum disparity estimation

- **Problem**: the existence of

$$min_{\theta \in \Theta} \rho(q, \theta).$$

- **Assumptions**: $\rho(q, \cdot)$ is continuous in Θ and Θ is a compact set.

Theorem 1: Existence and uniqueness of a disparity functional

Under the conditions $\rho(q, \theta)$ is continuous in θ, Θ is a compact set and the identifiability of $\mathcal{F}_\theta$, the existence and uniqueness the functional T^ρ are verified.

Theorem 2: Continuity of a disparity functional

Under the conditions of Theorem 1, if

$$\sup_{\theta \in \Theta} |\rho(q_n, \theta) - \rho(q, \theta)| \to 0,$$

as $q_n \to q$ in ℓ_1 then, $T^\rho(q_n) \to T^\rho(q)$, that is, the continuity of the functional T^ρ holds.

Minimum disparity estimation: sample $\mathcal{Z}_n^*$

Theorem 3: Consistency of the MDE

Suppose that $T^\rho(p)$ is unique, where $p = p(\theta_0)$ is the true reproduction law. Then, under conditions which guarantee $\tilde{p}_{n,k}$ is a consistent estimator of p_k and the assumptions of Theorems 1 and 2,

$$\tilde{\theta}_n^\rho(\tilde{p}_n) = T^\rho(\tilde{p}_n) \to T^\rho(p) = \theta_0 \quad \text{a.s.}$$

In particular, for $\hat{p}_n$,

$$\hat{\theta}_n^\rho(\hat{p}_n) \to \theta_0 \quad \text{a.s. on } \{Z_n \to \infty\}.$$

Minimum disparity estimation: sample $\mathcal{Z}_n^*$

Theorem 4: Asymptotic normality of the MDE

Let $p = p(\theta_0)$ be the true reproduction law. Under certain assumptions,

$$\left(\sum_{l=0}^{n-1} \phi_l(Z_l)\right)^{1/2} (\tilde{\theta}_n^{\rho}(\hat{p}_n) - \theta_0) \to N\left(0, I(\theta_0)^{-1}\right),$$

with respect to the distribution $P[\cdot|Z_n \to \infty]$, being

$$I(\theta_0) = \sum_{k=0}^{\infty} \left(\frac{p_k'(\theta_0)}{p_k(\theta_0)}\right)^2 p_k(\theta_0).$$

Minimum disparity estimation: sample $\overline{\mathcal{Z}}_n$

- **Sample**:

$$\overline{\mathcal{Z}}_n = \{Z_0, \ldots, Z_n, \phi_0(Z_0), \ldots, \phi_{n-1}(Z_{n-1})\}$$

- **Problem**: to determine a nonparametric estimator of p_k, $k \geq 0$, based on $\overline{\mathcal{Z}}_n$.

$$\ell(p|\overline{\mathcal{Z}}_n) = \sum_{l=0}^{n-1} \log\left(P[\phi(z_l) = \phi_l^*]\right) + \sum_{l=0}^{n-1} \log\left(\sum_{i_1+\ldots+i_{\phi_l^*}=z_{l+1}} p_{i_1} \cdots p_{i_{\phi_l^*}}\right)$$

- **Methodology**: Expectation-Maximization algorithm.
- **Nonparametric estimator of p based on $\overline{\mathcal{Z}}_n$**: $\widehat{p}_{n,k}^{EM}$.
  - González, M., M.C., del Puerto, I. (2015). CSDA.
- **Consistency**: $\hat{\theta}_n^{\rho}(\widehat{p}_n^{EM})$ is strongly consistent as $\widehat{p}_{n,k}^{EM}$ is for each $k \geq 0$.

Minimum disparity estimation: sample $\overline{\mathcal{Z}}_n$

- **Sample**:

$$\overline{\mathcal{Z}}_n = \{Z_0, \ldots, Z_n, \phi_0(Z_0), \ldots, \phi_{n-1}(Z_{n-1})\}$$

- **Problem**: to determine a nonparametric estimator of p_k, $k \geq 0$, based on $\overline{\mathcal{Z}}_n$.

$$\ell(p|\overline{\mathcal{Z}}_n) = \sum_{l=0}^{n-1} \log\left(P[\phi(z_l) = \phi_l^*]\right) + \sum_{l=0}^{n-1} \log\left(\sum_{i_1+\ldots+i_{\phi_l^*}=z_{l+1}} p_{i_1}\ldots p_{i_{\phi_l^*}}\right)$$

- **Methodology**: Expectation-Maximization algorithm.
- **Nonparametric estimator of p based on $\overline{\mathcal{Z}}_n$**: $\widehat{p}_{n,k}^{EM}$.
  - González, M., M.C., del Puerto, I. (2015). CSDA.

- **Consistency**: $\hat{\theta}_n^{\rho}(\widehat{p}_n^{EM})$ is strongly consistent as $\widehat{p}_{n,k}^{EM}$ is for each $k \geq 0$.

Minimum disparity estimation: sample $\overline{\mathcal{Z}}_n$

- **Sample**:

$$\overline{\mathcal{Z}}_n = \{Z_0, \ldots, Z_n, \phi_0(Z_0), \ldots, \phi_{n-1}(Z_{n-1})\}$$

- **Problem**: to determine a nonparametric estimator of p_k, $k \geq 0$, based on $\overline{\mathcal{Z}}_n$.

$$\ell(p|\overline{\mathcal{Z}}_n) = \sum_{l=0}^{n-1} \log\left(P[\phi(z_l) = \phi_l^*]\right) + \sum_{l=0}^{n-1} \log\left(\sum_{i_1+\ldots+i_{\phi_l^*}=z_{l+1}} p_{i_1}\ldots p_{i_{\phi_l^*}}\right)$$

- **Methodology**: Expectation-Maximization algorithm.
- **Nonparametric estimator of p based on $\overline{\mathcal{Z}}_n$**: $\widehat{p}_{n,k}^{EM}$.
  - González, M., M.C., del Puerto, I. (2015). CSDA.
- **Consistency**: $\hat{\theta}_n^{\rho}(\widehat{p}_n^{EM})$ is strongly consistent as $\widehat{p}_{n,k}^{EM}$ is for each $k \geq 0$.

Minimum disparity estimation: sample $\overline{\mathcal{Z}}_n$

- **Sample**:

$$\overline{\mathcal{Z}}_n = \{Z_0, \ldots, Z_n, \phi_0(Z_0), \ldots, \phi_{n-1}(Z_{n-1})\}$$

- **Problem**: to determine a nonparametric estimator of p_k, $k \geq 0$, based on $\overline{\mathcal{Z}}_n$.

$$\ell(p|\overline{\mathcal{Z}}_n) = \sum_{l=0}^{n-1} \log\left(P[\phi(z_l) = \phi_l^*]\right) + \sum_{l=0}^{n-1} \log\left(\sum_{i_1+\ldots+i_{\phi_l^*}=z_{l+1}} p_{i_1} \cdots p_{i_{\phi_l^*}} \right)$$

- **Methodology**: Expectation-Maximization algorithm.
- **Nonparametric estimator of p based on $\overline{\mathcal{Z}}_n$**: $\widehat{p}_{n,k}^{EM}$.
  - González, M., M.C., del Puerto, I. (2015). CSDA.

- **Consistency**: $\hat{\theta}_n^{\rho}(\widehat{p}_n^{EM})$ is strongly consistent as $\widehat{p}_{n,k}^{EM}$ is for each $k \geq 0$.

Minimum disparity estimation: sample $\mathcal{Z}_n$

- **Sample**:

$$\mathcal{Z}_n = \{Z_0, \dots, Z_n\}$$

- **Problem**: to determine the MLE of p_k, $k \geq 0$, based on $\mathcal{Z}_n$.

$$\ell(p|\mathcal{Z}_n) = \sum_{l=0}^{n-1} \log \left(\delta_0(z_{l+1}) P[\phi(z_l) = 0] + \sum_{j=1}^{\infty} P[\phi(z_l) = j] \sum_{i_1+\ldots+i_j = z_{l+1}} \prod_{k=1}^{j} p_{i_k} \right)$$

- **Methodology**: Expectation-Maximization algorithm.
- **Nonparametric estimator of p based on $\mathcal{Z}_n$**: $p_{n,k}^{EM}$.
  - González, M., M.C., del Puerto, I. (2015). CSDA.

- **Consistency**: $\hat{\theta}_n^{\rho}(p_n^{EM})$ is strongly consistent as $p_{n,k}^{EM}$ is for each $k \geq 0$.

Minimum disparity estimation

- **Mixture model for gross error**:

$$p(\theta, \alpha, L) = (1-\alpha)p(\theta) + \alpha\delta_L.$$

- **α-influence curves of T^ρ**:

$$L \in \mathbb{N}_0 \mapsto \alpha^{-1}(T^\rho(p(\theta, \alpha, L)) - \theta), \quad \alpha \in (0,1).$$

Theorem 5: Robustness of a disparity functional

Under conditions of Theorems 1 and 2, for every $\alpha \in (0,1)$, every $\theta \in \Theta$ and if $T^\rho(p(\theta, \alpha, L))$ is unique for all L we have the functional T^ρ is robust at $p(\theta)$ against $100\alpha\%$ contamination by gross errors at arbitrary integer L.

Minimum disparity estimation

- **Mixture model for gross error**:

$$p(\theta, \alpha, L) = (1 - \alpha)p(\theta) + \alpha\delta_L.$$

- **α-influence curves of T^ρ**:

$$L \in \mathbb{N}_0 \mapsto \alpha^{-1}(T^\rho(p(\theta, \alpha, L)) - \theta), \quad \alpha \in (0, 1).$$

Theorem 5: Robustness of a disparity functional

Under conditions of Theorems 1 and 2, for every $\alpha \in (0, 1)$, every $\theta \in \Theta$ and if $T^\rho(p(\theta, \alpha, L))$ is unique for all L we have the functional T^ρ is robust at $p(\theta)$ against $100\alpha\%$ contamination by gross errors at arbitrary integer L.

Minimum disparity estimation

- **Asymptotic breakdown point of a disparity functional T^ρ at $q \in \Gamma$:**

$$\alpha^*(T^\rho, q) = \inf \{\alpha \in (0,1) : b(\alpha; T^\rho, q) = \infty\},$$

with

$$b(\alpha; T^\rho, q) = \sup \{|T^\rho((1-\alpha)q + \alpha\overline{q}) - T^\rho(q)| : \overline{q} \in \Gamma\}.$$

Theorem 6: Asymptotic breakdown point

(i) Under certain conditions on the function G and the contaminant distributions,

$$\alpha^*(T^\rho, p) \geq \frac{1}{2}.$$

(ii) Under conditions of Theorems 1 and 2, if $\hat{\varrho} = \max_{t\in\Theta} \sum_{k=0}^{\infty} (p_k(\theta_0)p_k(t))^{1/2}$ and $\varrho^* = \lim_{M\to\infty} \sup_{|t|>M} \sum_{k=0}^{\infty} (p_k(\theta_0)p_k(t))^{1/2}$, then

$$\alpha^*(T^\rho, p) \geq \frac{(\hat{\varrho} - \varrho^*)^2}{[1 + (\hat{\varrho} - \varrho^*)^2]}.$$

Minimum disparity estimation

- **Asymptotic breakdown point of a disparity functional T^ρ at $q \in \Gamma$:**

$$\alpha^*(T^\rho, q) = \inf\left\{\alpha \in (0,1) : b(\alpha; T^\rho, q) = \infty\right\},$$

with

$$b(\alpha; T^\rho, q) = \sup\ \{|T^\rho((1-\alpha)q + \alpha\overline{q}) - T^\rho(q)| : \overline{q} \in \Gamma\}.$$

Theorem 6: Asymptotic breakdown point

(i) Under certain conditions on the function G and the contaminant distributions,

$$\alpha^*(T^\rho, p) \geq \frac{1}{2}.$$

(ii) Under conditions of Theorems 1 and 2, if $\hat{\varrho} = \max_{t\in\Theta} \sum_{k=0}^{\infty}(p_k(\theta_0)p_k(t))^{1/2}$ and $\varrho^* = \lim_{M\to\infty} \sup_{|t|>M} \sum_{k=0}^{\infty}(p_k(\theta_0)p_k(t))^{1/2}$, then

$$\alpha^*(T^\rho, p) \geq \frac{(\hat{\varrho} - \varrho^*)^2}{[1 + (\hat{\varrho} - \varrho^*)^2]}.$$

Simulated example 1: sample $\mathcal{Z}_n^*$

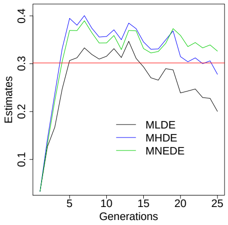


Fig: MDEs of p_0.

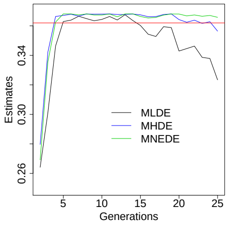


Fig: MDEs of p_1.

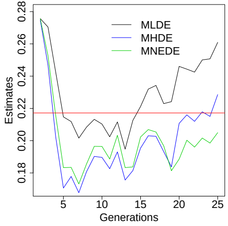


Fig: MDEs of p_2.

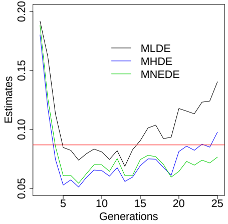


Fig: MDEs of p_3.

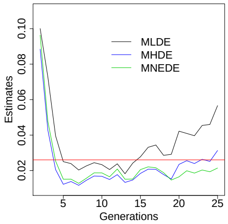


Fig: MDEs of p_4.

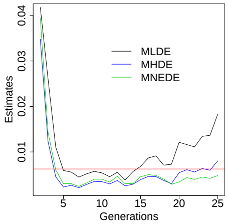


Fig: MDEs of p_5.

Simulated example 1: samples $\mathcal{Z}_n^*$, $\overline{\mathcal{Z}}_n$, $\mathcal{Z}_n$

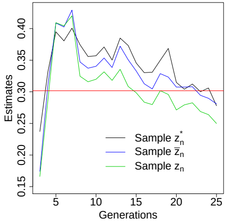


Fig: MHDEs of p_0.

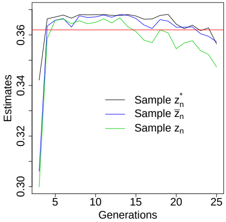


Fig: MHDEs of p_1.

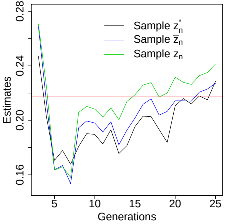


Fig: MHDEs of p_2.

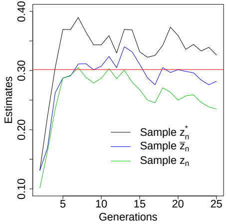


Fig: MNEDEs of p_0.

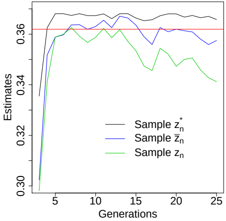


Fig: MNEDEs of p_1.

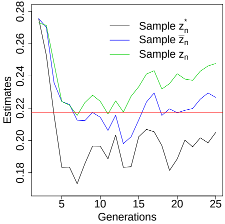


Fig: MNEDEs of p_2.

Simulated example 1: samples $\mathcal{Z}_n^*$, $\overline{\mathcal{Z}}_n$, $\mathcal{Z}_n$

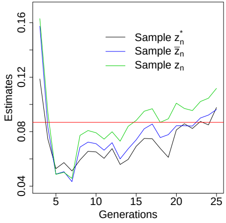


Fig: MHDEs of p_3.

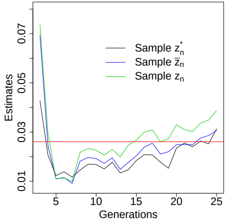


Fig: MHDEs of p_4.

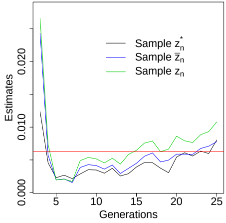


Fig: MHDEs of p_5.

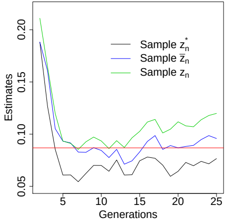


Fig: MNEDEs of p_3.

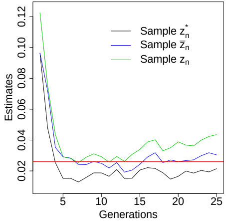


Fig: MNEDEs of p_4.

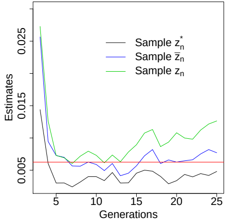


Fig: MNEDEs of p_5.

Simulated example 1: samples $\mathcal{Z}_n^*$, $\overline{\mathcal{Z}}_n$, $\mathcal{Z}_n$

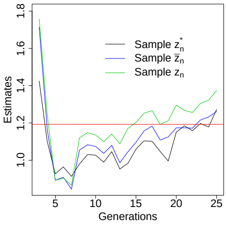


Fig: MHDEs of m.

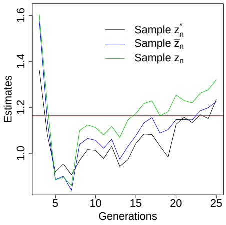


Fig: MHDEs of σ^2.

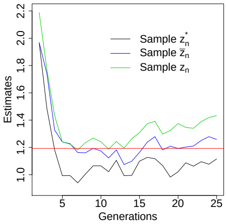


Fig: MNEDEs of m.

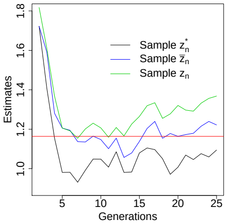


Fig: MNEDEs of σ^2.

Simulated example 2: uncontaminated model

We have simulated the first 50 generations of 100 CBPs verifying:

- They start with $Z_0 = 1$ individual.
- The distribution of the variables $X_{ij} \sim \mathcal{P}(\theta_0)$, with $\theta_0 = 4$ for $i = 0, 1, \ldots$, $j = 1, \ldots$
- $\phi_n(k) \sim \mathcal{P}(k\lambda)$, with $\lambda = 0.3$, $n \in \mathbb{N}$, $k \geq 0$.

- $\frac{1}{100}\sum_{i=1}^{100} \tilde{\theta}_i^{LD}$ (black line).
- $\frac{1}{100}\sum_{i=1}^{100} \tilde{\theta}_i^{HD}$ (blue line).
- $\frac{1}{100}\sum_{i=1}^{100} \tilde{\theta}_i^{NED}$ (green line).

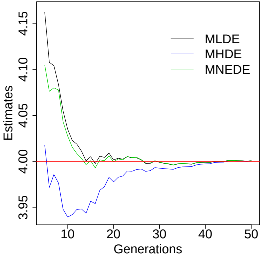


Fig: Evolution of the mean of the estimates of θ_0.

Simulated example 2: under mixture models for gross errors

We have simulated 100 CBPs following the previous model and with offspring distribution contaminated according to the mixture model for gross error:

- $L = 0, \ldots, 25$.
- $\alpha = 0.05, 0.1, 0.15, 0.2, 0.25, 0.3, 0.35, 0.4, 0.45, 0.5$.

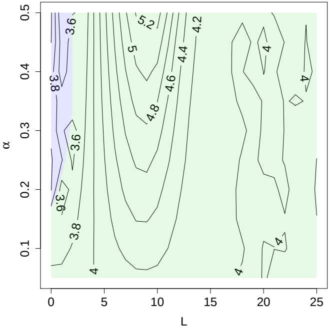


Fig: Means of the MHDEs of θ_0.

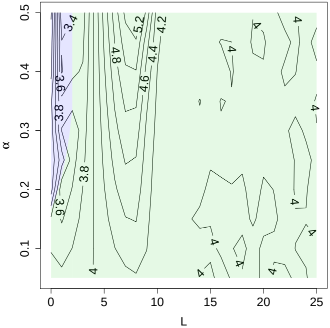


Fig: Means of the MNEDEs of θ_0.

Simulated example 2: under mixture models for gross errors

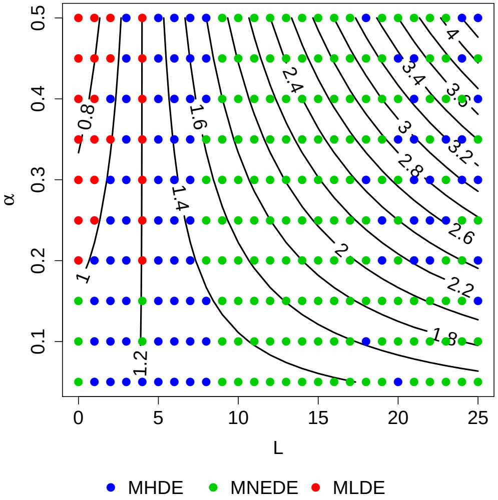


Fig: Disparity which provides the minimum of MSE of the estimates of θ_0

Concluding remarks

- For a CBP with offspring distribution belonging to a parametric family $\mathcal{F}_\theta$, we have deduced the **minimum disparity estimator of** θ_0**,** $p(\theta_0)$**,** $m(\theta_0)$ **and** $\sigma^2(\theta_0)$ based on the **whole family tree** and we have established the **consistency** and **asymptotic normality** of the minimum disparity estimator of θ_0.
- For a CBP with offspring distribution belonging to a parametric family $\mathcal{F}_\theta$, we have also determined the **minimum disparity estimators** of the main parameters of the model considering only the **total number of individuals and progenitors in each generation** or **only the population sizes**.
- We have studied **robustness against model perturbations and resistance to outliers of the minimum disparity estimator** for a certain family of disparities. These properties show that the related minimum disparity estimators are better choices than maximum likelihood estimators.
- We have **implemented** the **minimum disparity estimation** using statistical software and programming environment **R**.

Concluding remarks

- For a CBP with offspring distribution belonging to a parametric family $\mathcal{F}_\theta$, we have deduced the **minimum disparity estimator of θ_0, $p(\theta_0)$, $m(\theta_0)$ and $\sigma^2(\theta_0)$** based on the **whole family tree** and we have established the **consistency** and **asymptotic normality** of the minimum disparity estimator of θ_0.
- For a CBP with offspring distribution belonging to a parametric family $\mathcal{F}_\theta$, we have also determined the **minimum disparity estimators** of the main parameters of the model considering only the **total number of individuals and progenitors in each generation** or **only the population sizes**.
- We have studied **robustness against model perturbations and resistance to outliers of the minimum disparity estimator** for a certain family of disparities. These properties show that the related minimum disparity estimators are better choices than maximum likelihood estimators.
- We have **implemented** the **minimum disparity estimation** using statistical software and programming environment **R**.

Concluding remarks

- For a CBP with offspring distribution belonging to a parametric family $\mathcal{F}_\theta$, we have deduced the **minimum disparity estimator of θ_0, $p(\theta_0)$, $m(\theta_0)$ and $\sigma^2(\theta_0)$** based on the **whole family tree** and we have established the **consistency** and **asymptotic normality** of the minimum disparity estimator of θ_0.
- For a CBP with offspring distribution belonging to a parametric family $\mathcal{F}_\theta$, we have also determined the **minimum disparity estimators** of the main parameters of the model considering only the **total number of individuals and progenitors in each generation** or **only the population sizes**.
- We have studied **robustness against model perturbations and resistance to outliers of the minimum disparity estimator** for a certain family of disparities. These properties show that the related minimum disparity estimators are better choices than maximum likelihood estimators.
- We have **implemented** the **minimum disparity estimation** using statistical software and programming environment **R**.

Concluding remarks

- For a CBP with offspring distribution belonging to a parametric family $\mathcal{F}_\theta$, we have deduced the **minimum disparity estimator of θ_0, $p(\theta_0)$, $m(\theta_0)$ and $\sigma^2(\theta_0)$** based on the **whole family tree** and we have established the **consistency** and **asymptotic normality** of the minimum disparity estimator of θ_0.
- For a CBP with offspring distribution belonging to a parametric family $\mathcal{F}_\theta$, we have also determined the **minimum disparity estimators** of the main parameters of the model considering only the **total number of individuals and progenitors in each generation** or **only the population sizes**.
- We have studied **robustness against model perturbations and resistance to outliers of the minimum disparity estimator** for a certain family of disparities. These properties show that the related minimum disparity estimators are better choices than maximum likelihood estimators.
- We have **implemented** the **minimum disparity estimation** using statistical software and programming environment **R**.

References I

Bagley, J.H. (1986). On the almost sure convergence of controlled branching processes. *Journal of Applied Probability*, **23**, 827-831.

Basu, A., Sarkar, S., Vidyashankar, A. (1997). Minimum negative exponential disparity estimation in parametric models. *Journal of Statistical Planning and Inference*, **58**, 349-370.

Beran, R.J. (1977). Minimum Hellinger distance estimates for parametric models. *Annals of Statistics*, **5**, 445-463.

Bercu, B. (1999) Weighted estimation and tracking for Bienaymè Galton Watson processes with adaptive control. *Statistics & Probability Letters*, **42**, 415-421.

Dion, J. P. and Essebbar, B. (1995). On the statistics of controlled branching processes. *Lecture Notes in Statistics*, **99**, 14-21.

González, M., Gutiérrez, C., Martínez, R., del Puerto, I. (2013). Bayesian inference for controlled branching processes through MCMC and ABC methodologies. *Revista de la Real Academia de Ciencias Exactas, Físicas y Naturales. Serie A. Matemáticas*, **107(2)**, 459-473.

González, M., Martínez, R., del Puerto, I. (2004). Nonparametric estimation of the offspring distribution and the mean for a controlled branching process. *Test*, **13(2)**, 465-479.

González, M., M., C., del Puerto, I. (2015). Maximum likelihood estimation and Expectation-Maximization algorithm for controlled branching processes. Computational Statistics and Data Analysis, DOI: 10.1016/j.csda.2015.01.015

References II

González, M., Molina, M., del Puerto, I. (2002). On the class of controlled branching process with random control functions. *Journal of Applied Probability*, **39(4)**, 804-815.

González, M., Molina, M., del Puerto, I. (2003). On the geometric growth in controlled branching processes with random control function. *Journal of Applied Probability*, **40(4)**, 995-1006.

González, M., Molina, M., del Puerto, I. (2005a). Asymptotic behaviour of critical controlled branching process with random control function. *Journal of Applied Probability*, **42(2)**, 463-477.

González, M., Molina, M., del Puerto, I. (2005b). On the L^2-convergence of controlled branching processes with random control function. *Bernoulli*, **11(1)**, 37-46.

González, M. and del Puerto, I. (2012). Diffusion Approximation of an Array of Controlled Branching Processes. *Methodology and Computing in Applied Probability*, **14**, 843-861.

Hooker, G., Vidyashankar, A.N. (2014). Bayesian model robustness via disparities. *Test*, **23(3)**, 556-584.

Lindsay, B. G. (1994). Efficiency versus robustness: The case for minimum Hellinger distance and related methods. *The Annals of Statistics*, **22**, 1081-1114.

Martínez, R., del Puerto, I., Mota, M. (2009). On asymptotic posterior normality for controlled branching processes. *Statistics*, **43**, 367-378.

Mohan, K. (2000) Estimation of parameters in controlled branching process. *Stochastic Modelling and Applications*, **3**, 1-15.

References III

Park, C., Basu, A. (2004). Minimum disparity estimation: asymptotic normality and breakdown point results. *Bulletin of Informatics and Cybernetics*, **36**, 19-33.

Sevastyanov, B.A. and Zubkov, A. (1974). Controlled branching processes. *Theory of Probability and its Applications*, **19**, 14-24.

Sriram, T.N, Vidyashankar, A.N. (2000). Minimum Hellinger distance estimation for supercritical Galton-Watson Process. *Statistics and Probability Letters*, **50**, 331-342.

Sriram, T.N., Bhattacharya, A., González, M., Martínez, R., del Puerto, I. (2007). Estimation of the offspring mean in a controlled branching process with a random control function. *Stochastic Processes and their Applications*, **117**, 928-946.

Yanev, N.M. (1975). Conditions for degeneracy of φ-branching processes with random φ. *Theory of Probability and its Applications*, **20**, 421-428.

Zubkov, A.M. (1974). Analogies between Galton - Watson processes and ϕ-branching processes. *Theory of Probability and its Applications*, **19**, 309-331.

Thank you very much!

Acknowledgements: This research has been supported by the Ministerio de Educación, Cultura y Deporte (grant FPU13/03213), Ministerio de Economía y Competitividad (grant MTM2012-31235), the Gobierno de Extremadura (grant GR15105) and the FEDER.

GOBIERNO DE EXTREMADURA

GOBIERNO DE ESPAÑA MINISTERIO DE ECONOMÍA Y COMPETITIVIDAD

GOBIERNO DE ESPAÑA MINISTERIO DE EDUCACIÓN, CULTURA Y DEPORTE

Unión Europea Fondo Europeo de Desarrollo Regional "Una manera de hacer Europa"