Asymptotics of Joint Maxima of Discrete Random Variables

Anne Feidt
University of Zurich
with Christian Genest and Johanna Nešlehová

Disentis, 21st July 2008

Introduction

Notation

- $X_1, \ldots, X_n$ i.i.d. with cdf F
- Maximum $X_{(n)} := \max_{1 \leq i \leq n} X_i$
- $x_F = \sup_{x \in \mathbb{R}} \{F(x) < 1\}$ right endpoint of F

Question

Under which conditions on F do there exist $a_n, b_n \in \mathbb{R}$, $a_n > 0$, and a non-degenerate df F^* such that

$$\lim_{n \to \infty} P\left(\frac{X_{(n)} - b_n}{a_n} \leq x\right) = F^*(x),$$

i.e. such that F is in the maximum domain of attraction of F^*, $F \in \mathrm{MDA}(F^*)$?

Answer

F has to satisfy (Leadbetter et al., 1983)

$$\lim_{x \to x_F} \frac{1 - F(x)}{1 - F(x-)} = 1 \qquad (1)$$

Fisher-Tippett Theorem

If (1) is fulfilled, there exist only 3 possible limit laws for the normalized maximum $(X_{(n)} - a_n)/b_n$:

- Fréchet: $\Phi_\alpha(x) = \begin{cases} 0 & , x \leq 0 \\ \exp\{-x^{-\alpha}\} & , x > 0 \end{cases} \quad , \alpha > 0$
- Weibull: $\Psi_\alpha(x) = \begin{cases} \exp\{-(-x)^{-\alpha}\} & , x \leq 0 \\ 1 & , x > 0 \end{cases} \quad , \alpha > 0$
- Gumbel: $\Lambda(x) = \exp\{-e^{-x}\}, x \in \mathbb{R}.$

(the extreme-value distributions F^*)

Univariate discrete random variables

Problem

(1) is not satisfied for discrete distributions such as the Binomial, Poisson, Geometric, Negative Binomial $\Rightarrow$ no limit law for maxima!

Univariate discrete random variables

Problem

(1) is not satisfied for discrete distributions such as the Binomial, Poisson, Geometric, Negative Binomial $\Rightarrow$ no limit law for maxima!

Remedy

Let a distribution parameter vary with the sample size n at a suitable rate. Then

- Poisson in MDA(Gumbel) (Anderson et al., 1997)
- Binomial, Geometric, Negative Binomial in MDA(Gumbel) (Nadarajah and Mitov, 2002)

Example: Geometric

- $X_1, \dots, X_n$ i.i.d. $\sim \mathrm{Geo}(p)$, $0 < p < 1$, q=1-p
- $W^{(x)} := \sum_{i=1}^{n} \mathbb{1}_{\{X_i \geq x\}} = \#$ exceedances of level x
- $\left\{W^{(x)} = 0\right\} = \{\max_{1\leq i\leq n} X_i < \lfloor x \rfloor\}$

Approximate $W^{(x)}$ by a $\mathrm{Poi}(nq^{\lfloor x \rfloor})$ distribution:

$$\left| P\left(\max_{1\leq i\leq n} X_i < \lfloor x \rfloor\right) - e^{-nq^{\lfloor x \rfloor}} \right| \leq q^{\lfloor x \rfloor} \quad \text{(Stein-Chen method)}$$

Choose $p = p_n \overset{n\to\infty}{\longrightarrow} 0$ and $a_n = 1/p_n$, $b_n = \log n / p_n$. Then

$$\left| P\left(\max_{1\leq i\leq n} X_i \leq a_n x + b_n\right) - \exp\{-e^{-x}\} \right| \leq q_n^{a_n x + b_n} = O\left(\frac{1}{n}\right).$$

But, there exist discrete distributions such that (1) holds!

Example Let

- $X \geq 0$ absolutely continuous rv
- $x_F = \infty$
- hazard rate $f(x)/(1 - F(x)) \to 0$ as $x \to \infty$.
- e.g. Pareto distribution
- $\lceil x \rceil := \min\{n \in \mathbb{N} : n \geq x\}$

Then we discretize X to obtain $\lceil X \rceil$ with df

$$\lceil F \rceil(x) = P(\lceil X \rceil \leq x) = P(\lceil X \rceil \leq \lfloor x \rfloor) = P(X \leq \lfloor x \rfloor) = F(\lfloor x \rfloor)$$

$\to$ Can show that (1) holds for $\lceil X \rceil$ and

$$\lceil F \rceil \in \mathrm{MDA}(F^*) \Leftrightarrow F \in \mathrm{MDA}(F^*)$$

In higher dimensions?

Notation (d=2)

- $(X_1, Y_1), \dots, (X_n, Y_n)$ i.i.d. with joint df H and margins F, G
- componentwise maxima $X_{(n)}$, $Y_{(n)}$

Question

When do there exist a_n, b_n, c_n and $d_n \in \mathbb{R}$, $b_n, d_n > 0$, and a non-degenerate df H^* such that

$$\lim_{n\to\infty} P\left(\frac{X_{(n)} - a_n}{b_n} \leq x, \frac{Y_{(n)} - c_n}{d_n} \leq y\right) = H^*(x,y),$$

i.e. when is $H \in \mathrm{MDA}(H^*)$?

In higher dimensions?

Notation (d=2)

- $(X_1, Y_1), \dots, (X_n, Y_n)$ i.i.d. with joint df H and margins F, G
- componentwise maxima $X_{(n)}$, $Y_{(n)}$

Question

When do there exist a_n, b_n, c_n and $d_n \in \mathbb{R}$, $b_n, d_n > 0$, and a non-degenerate df H^* such that

$$\lim_{n\to\infty} P\left(\frac{X_{(n)} - a_n}{b_n} \le x, \frac{Y_{(n)} - c_n}{d_n} \le y\right) = H^*(x, y),$$

i.e. when is $H \in \mathrm{MDA}(H^*)$?

Answer for continuous margins

Galambos' Thm (1978). Uses copulas for modelling joint dfs.

What is a copula?

What is a copula?

Definition

A bivariate copula $C : [0,1]^2 \to [0,1]$ is a joint distribution function with standard uniform margins.

What is a copula?

Definition

A bivariate copula $C : [0,1]^2 \to [0,1]$ is a joint distribution function with standard uniform margins.

Idea

$$
\begin{aligned}
H(x,y) &= P\left(X \le x, Y \le y\right) \\
&= P\left[F(X) \le F(x), G(Y) \le G(y)\right] \\
&= P\left[U \le F(x), V \le G(y)\right], \text{ with } U, V \sim \mathcal{U}[0,1] \\
&= C\left(F(x), G(y)\right)
\end{aligned}
$$

(for F, G continuous)

Sklar's Theorem

(i) If H is a joint df with margins F and G, then $\exists$ a copula C s.t.

$$H(x, y) = C\left(F(x), G(y)\right) \quad \forall x, y \in [-\infty, \infty] \tag{2}$$

If F, G are continuous, then C is unique. If F, G are discrete, then C is uniquely determined on $\mathrm{Ran}(F) \times \mathrm{Ran}(G)$.

Sklar's Theorem

(i) If H is a joint df with margins F and G, then $\exists$ a copula C s.t.

$$H(x, y) = C\left(F(x), G(y)\right) \quad \forall x, y \in [-\infty, \infty] \tag{2}$$

If F, G are continuous, then C is unique. If F, G are discrete, then C is uniquely determined on $\mathrm{Ran}(F) \times \mathrm{Ran}(G)$.

(ii) If C is a copula and F and G are dfs, then H defined by (2) is a joint df with margins F and G.

Sklar's Theorem

(i) If H is a joint df with margins F and G, then $\exists$ a copula C s.t.

$$H(x, y) = C(F(x), G(y)) \quad \forall x, y \in [-\infty, \infty] \tag{2}$$

If F, G are continuous, then C is unique. If F, G are discrete, then C is uniquely determined on $\mathrm{Ran}(F) \times \mathrm{Ran}(G)$.

(ii) If C is a copula and F and G are dfs, then H defined by (2) is a joint df with margins F and G.

If (2) holds, say $C \in \mathcal{C}(H)$, the class of copulas compatible with H.

Galambos' Theorem

For continuous margins

Let H and H^* be joint dfs such that $H(x, y) = C(F(x), G(y))$ with F and G continuous, and $H^*(x, y) = C^*(F^*(x), G^*(y))$.
Then, with $u, v \in [0, 1]$,

$$H \in \mathrm{MDA}(H^*) \quad \Leftrightarrow \quad \begin{cases} \text{(i) } F \in \mathrm{MDA}(F^*) \text{ and } G \in \mathrm{MDA}(G^*) \\ \text{(ii) } \lim_{t\to\infty} C^t\left(u^{1/t}, v^{1/t}\right) = C^*(u, v) \end{cases},$$

i.e. the extremal behaviour of H is determined by the extremal behaviour of its margins and its underlying copula.

What if the margins are discrete?

Problem

C is not unique, $|\mathcal{C}(H)| = \infty$ (Genest and Nešlehová, 2007).

What if the margins are discrete?

Problem

C is not unique, $|\mathcal{C}(H)| = \infty$ (Genest and Nešlehová, 2007).

$\longrightarrow$ Can apply the following weak convergence result to prove Galambos' theorem for the discrete case.

Proposition 1

Let (X_1, Y_1), $(X_2, Y_2), \ldots$ be mutually independent random pairs such that (X_n, Y_n) has joint df H_n and margins F_n, G_n.
Let (X, Y) be a random pair with joint df H and margins F, G.
Then, the following are equivalent:

(a) $(X_n, Y_n) \xrightarrow{w} (X, Y)$, as $n \longrightarrow \infty$.

Proposition 1

Let $(X_1, Y_1), (X_2, Y_2), \ldots$ be mutually independent random pairs such that (X_n, Y_n) has joint df H_n and margins F_n, G_n.
Let (X, Y) be a random pair with joint df H and margins F, G.
Then, the following are equivalent:

(a) $(X_n, Y_n) \xrightarrow{w} (X, Y)$, as $n \to \infty$.

(b) $X_n \xrightarrow{w} X$ and $Y_n \xrightarrow{w} Y$, as $n \to \infty$,
and $\exists$ $C \in \mathcal{C}(H)$ and $\exists$ a sequence (C_n) with $C_n \in \mathcal{C}(H_n)$ such that $C_n \to C$ on $\mathrm{Ran}(F) \times \mathrm{Ran}(G)$.

Proposition 1

Let $(X_1, Y_1), (X_2, Y_2), \ldots$ be mutually independent random pairs such that (X_n, Y_n) has joint df H_n and margins F_n, G_n.
Let (X, Y) be a random pair with joint df H and margins F, G.
Then, the following are equivalent:

(a) $(X_n, Y_n) \xrightarrow{w} (X, Y)$, as $n \to \infty$.

(b) $X_n \xrightarrow{w} X$ and $Y_n \xrightarrow{w} Y$, as $n \to \infty$,
and $\exists\ C \in \mathcal{C}(H)$ and $\exists$ a sequence (C_n) with $C_n \in \mathcal{C}(H_n)$
such that $C_n \to C$ on $\mathrm{Ran}(F) \times \mathrm{Ran}(G)$.

(c) $X_n \xrightarrow{w} X$ and $Y_n \xrightarrow{w} Y$ as $n \to \infty$,
and $\forall\ C_n \in \mathcal{C}(H_n)$ and $\forall\ C \in \mathcal{C}(H)$,
we have $C_n \to C$ uniformly on $\overline{\mathrm{Ran}(F)} \times \overline{\mathrm{Ran}(G)}$.

Proof: use triangle inequality, Lipschitz-property of copulas, Continuous Mapping thm, Arzelá-Ascoli thm

General Galambos (for i.i.d. pairs)

Apply Proposition 1 to normalized maxima:

Proposition 2

Let $(X_1, Y_1), (X_2, Y_2), \dots$ be mutually independent random pairs with common joint df H and margins F, G.
Let H^* be a joint df with margins F^*, G^* and copula C^*.
Then, the following are equivalent:

(a) $H \in \mathrm{MDA}(H^*)$

(b) $F \in \mathrm{MDA}(F^*)$ and $G \in \mathrm{MDA}(G^*)$ and $\exists\ C \in \mathcal{C}(H)$ such that $\lim_{t\to\infty} C^t\left(u^{1/t}, v^{1/t}\right) = C^*(u, v)$ for all $(u, v) \in [0, 1]^2$.

(c) $F \in \mathrm{MDA}(F^*)$ and $G \in \mathrm{MDA}(G^*)$ and $\forall\ C \in \mathcal{C}(H)$, $\lim_{t\to\infty} C^t\left(u^{1/t}, v^{1/t}\right) = C^*(u, v)$ holds uniformly on $[0, 1]^2$.

General Galambos (for triangular arrays)

If margins are Bin, Poi, Geo, NB, ... $\Rightarrow$ let parameter vary with n

Proposition 2

Let $(X_1, Y_1), (X_2, Y_2), \ldots$ be mutually independent random pairs with common joint df H and margins F, G.
Let H^* be a joint df with margins F^*, G^* and copula C^*.
Then, the following are equivalent:

(a) $H \in \mathrm{MDA}(H^*)$

(b) $F \in \mathrm{MDA}(F^*)$ and $G \in \mathrm{MDA}(G^*)$ and $\exists\ C \in \mathcal{C}(H)$ such that $\lim_{t\to\infty} C^t\left(u^{1/t}, v^{1/t}\right) = C^*(u,v)$ for all $(u,v) \in [0,1]^2$.

(c) $F \in \mathrm{MDA}(F^*)$ and $G \in \mathrm{MDA}(G^*)$ and $\forall\ C \in \mathcal{C}(H)$, $\lim_{t\to\infty} C^t\left(u^{1/t}, v^{1/t}\right) = C^*(u,v)$ holds uniformly on $[0,1]^2$.

General Galambos (for triangular arrays)

If margins are Bin, Poi, Geo, NB, $\ldots \Rightarrow$ let parameter vary with n

Proposition 2'

Let (X_{n1}, Y_{n1}), $(X_{n2}, Y_{n2}), \ldots$ be mutually independent random pairs with common joint df H_n and margins F_n, G_n.
Let H^* be a joint df with margins F^*, G^* and copula C^*.
Then, the following are equivalent:

(a) $(H_n) \in \mathrm{MDA}(H^*)$

(b) $(F_n) \in \mathrm{MDA}(F^*)$ and $(G_n) \in \mathrm{MDA}(G^*)$ and $\exists\ (C_n) \in \mathcal{C}(H_n)$ such that $\lim_{n\to\infty} C_n^n \left(u^{1/n}, v^{1/n}\right) = C^*(u, v)$ for all $(u, v) \in [0, 1]^2$.

(c) $(F_n) \in \mathrm{MDA}(F^*)$ and $(G_n) \in \mathrm{MDA}(G^*)$ and $\forall\ (C_n) \in \mathcal{C}(H_n)$, $\lim_{n\to\infty} C_n^n \left(u^{1/n}, v^{1/n}\right) = C^*(u, v)$ holds uniformly on $[0, 1]^2$.

Idea why Prop. 1 $\Rightarrow$ Prop. 2, 2'

- $\widetilde{H}_n(x, y) := P\left(X_{(n)} \le a_n x + b_n, Y_{(n)} \le c_n y + d_n\right)$
- $\widetilde{F}_n(x) := P\left(X_{(n)} \le a_n x + b_n\right) = F_n^n(a_n x + b_n)$
- $\widetilde{G}_n(y) := P\left(Y_{(n)} \le c_n y + d_n\right) = G_n^n(c_n y + d_n)$

$$\begin{aligned}
\widetilde{H}_n(x, y) &= H_n^n(a_n x + b_n, c_n y + d_n) \\
&= C_n^n(F_n(a_n x + b_n), G_n(c_n y + d_n)), \text{ for } C_n \in \mathcal{C}(H_n) \\
&= C_n^n(\widetilde{F}_n^{1/n}(x), \widetilde{G}_n^{1/n}(y)) \\
&= D_n(\widetilde{F}_n(x), \widetilde{G}_n(y)),
\end{aligned}$$

where $D_n(u, v) := C_n^n\left(u^{1/n}, v^{1/n}\right)$ is a copula $\Rightarrow D_n \in \mathcal{C}(\widetilde{H}_n)$.
Therefore,

$$C_n^n(u^{1/n}, v^{1/n}) \to C^*(u, v) \iff D_n \to C^*$$

Examples

Proposition 2 (i.i.d. pairs)

- Pareto distribution of the first kind (Kotz et al., 2000) with discretized margins
- Marshall-Olkin exponential distribution (Nelsen, 2006) with discretized margins

Proposition 2' (triangular arrays)

- Marshall-Olkin geometric distribution (Marshall and Olkin, 1985)
- Poisson (Coles and Pauli, 2001), copula not tractable?

Thanks for listening

Enjoy dinner