On the mixing advantage

Kais Hamza

Monash University

ANZAPW, July 2013, University of Queensland

Joint work with Aidan Sudbury, Peter Jagers & Daniel Tokarev

Introduction

Introduction

- X_i^j, X_i are the lifetime of an individual/unit and, $\max(X_i^1, \ldots, X_i^n)$ and $\max(X_1, X_2, \ldots, X_n)$ represent the lifetime of population/system.
  All random variables are assumed to be non-negative.
  $X_i, X_i^1, \ldots, X_i^n$ are identically distributed.

Introduction

- X_i^j, X_i are the lifetime of an individual/unit and, $\max(X_i^1, \ldots, X_i^n)$ and $\max(X_1, X_2, \ldots, X_n)$ represent the lifetime of population/system.
  All random variables are assumed to be non-negative.
  $X_i, X_i^1, \ldots, X_i^n$ are identically distributed.
- $X_1 \quad X_1^1 \quad X_1^2 \quad \ldots \quad X_1^n \quad \rightarrow \quad M_1 = \mathbb{E}[\max(X_1^1, \ldots, X_1^n)]$

Introduction

- X_i^j, X_i are the lifetime of an individual/unit and, $\max(X_i^1, \ldots, X_i^n)$ and $\max(X_1, X_2, \ldots, X_n)$ represent the lifetime of population/system.
  All random variables are assumed to be non-negative.
  $X_i, X_i^1, \ldots, X_i^n$ are identically distributed.
- $$\begin{array}{cccccccc} X_1 & X_1^1 & X_1^2 & \ldots & X_1^n & \to & M_1 = \mathbb{E}[\max(X_1^1, \ldots, X_1^n)] \\ X_2 & X_2^1 & X_2^2 & \ldots & X_2^n & \to & M_2 = \mathbb{E}[\max(X_2^1, \ldots, X_2^n)] \end{array}$$

Introduction

- X_i^j, X_i are the lifetime of an individual/unit and, $\max(X_i^1, \ldots, X_i^n)$ and $\max(X_1, X_2, \ldots, X_n)$ represent the lifetime of population/system.
  All random variables are assumed to be non-negative.
  $X_i, X_i^1, \ldots, X_i^n$ are identically distributed.
- $$\begin{array}{cccccccl} X_1 & X_1^1 & X_1^2 & \ldots & X_1^n & \to & M_1 = \mathbb{E}[\max(X_1^1, \ldots, X_1^n)] \\ X_2 & X_2^1 & X_2^2 & \ldots & X_2^n & \to & M_2 = \mathbb{E}[\max(X_2^1, \ldots, X_2^n)] \\ \vdots & & & & & & \end{array}$$

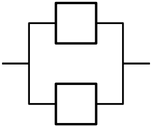
Introduction

- X_i^j, X_i are the lifetime of an individual/unit and, $\max(X_i^1, \ldots, X_i^n)$ and $\max(X_1, X_2, \ldots, X_n)$ represent the lifetime of population/system.
  All random variables are assumed to be non-negative.
  $X_i, X_i^1, \ldots, X_i^n$ are identically distributed.
- $$\begin{array}{llllllll} X_1 & X_1^1 & X_1^2 & \ldots & X_1^n & \rightarrow & M_1 = \mathbb{E}[\max(X_1^1, \ldots, X_1^n)] \\ X_2 & X_2^1 & X_2^2 & \ldots & X_2^n & \rightarrow & M_2 = \mathbb{E}[\max(X_2^1, \ldots, X_2^n)] \\ \vdots & & & & & & \\ X_n & X_n^1 & X_n^2 & \ldots & X_n^n & \rightarrow & M_n = \mathbb{E}[\max(X_n^1, \ldots, X_n^n)] \end{array}$$

Introduction

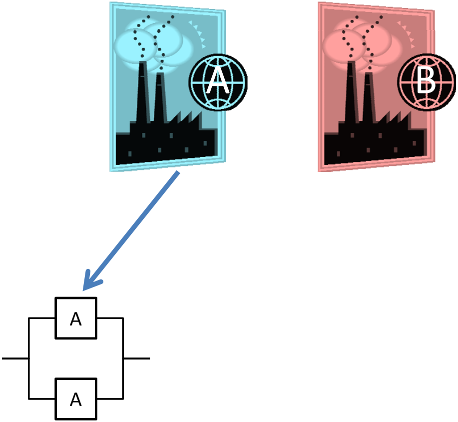
- X_i^j, X_i are the lifetime of an individual/unit and, $\max(X_i^1, \ldots, X_i^n)$ and $\max(X_1, X_2, \ldots, X_n)$ represent the lifetime of population/system.
  All random variables are assumed to be non-negative.
  $X_i, X_i^1, \ldots, X_i^n$ are identically distributed.
- $$\begin{array}{llllllll} X_1 & X_1^1 & X_1^2 & \ldots & X_1^n & \rightarrow & M_1 = \mathbb{E}[\max(X_1^1, \ldots, X_1^n)] \\ X_2 & X_2^1 & X_2^2 & \ldots & X_2^n & \rightarrow & M_2 = \mathbb{E}[\max(X_2^1, \ldots, X_2^n)] \\ \vdots & & & & & & \\ X_n & X_n^1 & X_n^2 & \ldots & X_n^n & \rightarrow & M_n = \mathbb{E}[\max(X_n^1, \ldots, X_n^n)] \end{array}$$
- We wish to compare $\mathbb{E}[\max(X_1, X_2, \ldots, X_n)]$ to $M_i = \mathbb{E}[\max(X_i^1, \ldots, X_i^n)]$, $i = 1, \ldots, n$.

Introduction



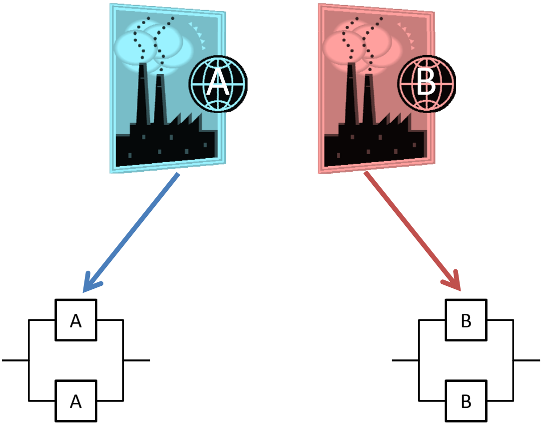
Reliability – Warm Duplication Method

Introduction



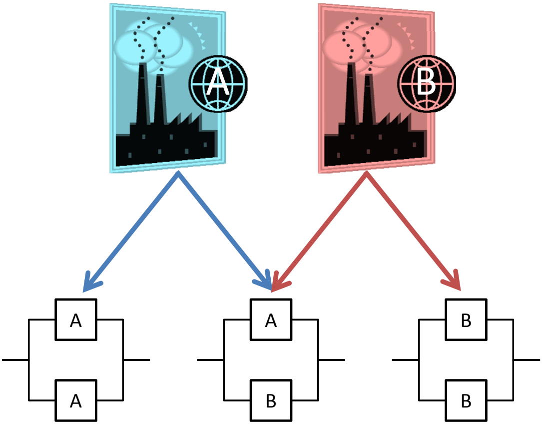
Reliability – Warm Duplication Method

Introduction



Reliability – Warm Duplication Method

Introduction



Reliability – Warm Duplication Method

Introduction

Introduction

- Question: Is it better to mix or go with a single type?

Introduction

- Question: Is it better to mix or go with a single type?
- Obviously, if one type dominates all others, then choosing that type only is optimum.

Introduction

- Question: Is it better to mix or go with a single type?
- Obviously, if one type dominates all others, then choosing that type only is optimum.
- Question: What if all types are similar (no dominant type); i.e.

$$\mathbb{E}[\max(X_1^1, \ldots, X_1^n)] = \ldots = \mathbb{E}[\max(X_n^1, \ldots, X_n^n)]?$$

Introduction

Introduction

Assume all random variables are independent.

Introduction

Assume all random variables are independent.

- It is easy to show (direct consequence of the arithmetic-geometric mean inequality) that

$$\mathbb{E}[\max(X_1, \dots, X_n)] \geq \mathbb{E}[\max(X_i^1, \dots, X_i^n)].$$

In fact, the same arithmetic-geometric mean inequality shows that

$$\mathbb{E}[\max(X_1, \dots, X_n)] \geq \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}[\max(X_i^1, \dots, X_i^n)].$$

In other words, mixing is advantageous.

Introduction

Assume all random variables are independent.

- It is easy to show (direct consequence of the arithmetic-geometric mean inequality) that

$$\mathbb{E}[\max(X_1, \dots, X_n)] \geq \mathbb{E}[\max(X_i^1, \dots, X_i^n)].$$

  In fact, the same arithmetic-geometric mean inequality shows that

$$\mathbb{E}[\max(X_1, \dots, X_n)] \geq \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}[\max(X_i^1, \dots, X_i^n)].$$

  In other words, mixing is advantageous.

- If $M_i = \mathbb{E}[\max(X_i^1, \dots, X_i^n)]$, $i = 1, \dots, n$, we call mixing factor

$$\theta = \frac{\mathbb{E}[\max(X_1, \dots, X_n)]}{\max(M_1, \dots, M_n)}.$$

  We show that when $M_i = M$, $\theta \leq 2 - 1/n < 2$.

Existing literature

Existing literature

- An extensive literature exists on $\mathbb{E}[\max(X_1, \ldots, X_n)]$ in the iid case – see David and Nagaraja (2003). However, very little work exists for the non-identically distributed case.

Existing literature

- An extensive literature exists on $\mathbb{E}[\max(X_1, \ldots, X_n)]$ in the iid case – see David and Nagaraja (2003). However, very little work exists for the non-identically distributed case.
- Arnold and Groeneveld (1979) obtain upper and lower bounds on $\mathbb{E}[\max(X_1, \ldots, X_n)]$ even when $X_1, \ldots, X_n$ are not independent and not identically distributed, but in terms of $\mathbb{E}[X_1]$ and $\text{var}(X_i)$, not $M_1, \ldots, M_n$.
  This generalises Hartley and David (1954) and Gumbel (1954) who deal with the iid case.

Existing literature

Existing literature

- Sen (1970) shows that $\max(X_1, \ldots, X_n)$ stochastically dominates $\max(Y^1, \ldots, Y^n)$, where $Y^1, \ldots, Y^n$ are iid equally-weighted probability mixtures of $X_1, \ldots, X_n$:

$$\mathbb{P}(\max(X_1, \ldots, X_n) \leq z) \leq \mathbb{P}(\max(Y^1, \ldots, Y^n) \leq z).$$

In particular

$$\begin{aligned} &\frac{1}{n}\sum_{i=1}^{n} \mathbb{E}[\max(X_i^1, \ldots, X_i^n)] \\ &\leq \; \mathbb{E}[\max(Y^1, \ldots, Y^n)] \; \leq \; \mathbb{E}[\max(X_1, \ldots, X_n)]. \end{aligned}$$

However, $\mathbb{E}[\max(Y^1, \ldots, Y^n)]$ cannot be expressed in terms of $M_1, \ldots, M_n$.

Unbounded independent case

Unbounded independent case

Theorem (H., Jagers, Sudbury & Tokarev, 2009)

If $X_1, \ldots, X_n$ are independent random variables with the property that $\mathbb{E}[\max(X_i^1, \ldots, X_i^n)] = M_i$, $i = 1, 2, ..., n$, then

$$\begin{aligned} \frac{1}{n}\sum_{i=1}^{n} M_i &\leq \mathbb{E}[\max(X_1, \ldots, X_n)] \\ &\leq \frac{1}{n}\sum_{i=1}^{n} M_i + \frac{n-1}{n}\max(M_1, \ldots, M_n). \end{aligned}$$

In particular, if $M_i = M$, $i = 1, ..., n$,

$$M \leq \mathbb{E}[\max(X_1, \ldots, X_n)] \leq (2 - 1/n)M.$$

Unbounded independent case

Theorem (H., Jagers, Sudbury & Tokarev, 2009)

If $X_1, \ldots, X_n$ are independent random variables with the property that $\mathbb{E}[\max(X_i^1, \ldots, X_i^n)] = M_i$, $i = 1, 2, ..., n$, then

$$\begin{aligned} \frac{1}{n}\sum_{i=1}^{n} M_i &\leq \mathbb{E}[\max(X_1, \ldots, X_n)] \\ &\leq \frac{1}{n}\sum_{i=1}^{n} M_i + \frac{n-1}{n}\max(M_1, \ldots, M_n). \end{aligned}$$

In particular, if $M_i = M$, $i = 1, ..., n$,

$$M \leq \mathbb{E}[\max(X_1, \ldots, X_n)] \leq (2 - 1/n)M.$$

The upper bound is obtained by letting some of the random variables be concentrated on 0 and x and letting $x \to \infty$.

Bounded independent case

Bounded independent case

Theorem (H. & Sudbury, 2011)

If a set of random variables $X_1, \dots, X_n$ are independent, concentrated on $[0, b]$ and s.t.

$$\mathbb{E}[\max(X_i^1, \dots, X_i^n)] = M_i, i = 1, \dots, n,$$

then, putting $M_n = \max(M_1, \dots, M_n)$,

$$\begin{aligned} b - \prod_{i=1}^{n}(b - M_i)^{1/n} &\leq \mathbb{E}[\max(X_1, \dots, X_n)] \\ &\leq b - (b - M_n)\prod_{i=1}^{n-1}(1 - M_i/b)^{1/n}. \end{aligned}$$

Bounded independent case

Bounded independent case

Corollary

In the case $M_i = M, i = 1, \dots, n$ we have

$$M \leq \mathbb{E}[\max(X_1, \dots, X_n)] \leq b - b(1 - M/b)^{2-1/n}$$

where the latter expression approaches $(2-1/n)M$ as $b \to +\infty$ and $M(2 - M/b)$ as $n \to +\infty$.

Bounded independent case

Bounded independent case

Changing X into $b - X$ transforms maxima into minima immediately yielding the following result.

Bounded independent case

Changing X into $b - X$ transforms maxima into minima immediately yielding the following result.

Corollary

The equivalent result for the minima, with $m_1 = \min(m_1, \ldots, m_n)$, is

$$m_1 \prod_{i=2}^{n} (m_i/b)^{1/n} \leq \mathbb{E}[\min(X_1, \ldots, X_n)] \leq \prod_{i=1}^{n} m_i^{1/n}.$$

Dependent case

Dependent case

- What if the random variables are NOT independent.

Dependent case

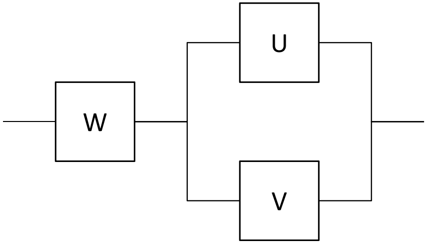
- What if the random variables are NOT independent.
- U, V and W are independent continuous random variables. Let $X = U \wedge W$ and $Y = V \wedge W$ ($a \wedge b = \min(a, b)$).

Dependent case

- What if the random variables are NOT independent.
- U, V and W are independent continuous random variables. Let $X = U \wedge W$ and $Y = V \wedge W$ ($a \wedge b = \min(a, b)$).

Dependent case

- What if the random variables are NOT independent.
- U, V and W are independent continuous random variables. Let $X = U \wedge W$ and $Y = V \wedge W$ ($a \wedge b = \min(a, b)$).



Dependent case – Copulas

Dependent case – Copulas

- If X has marginal F, Y has marginal G and they assume a copula C, then (X, Y) has joint distribution

$$H(x, y) = C(F(x), G(y)).$$

Dependent case – Copulas

- If X has marginal F, Y has marginal G and they assume a copula C, then (X, Y) has joint distribution

$$H(x, y) = C(F(x), G(y)).$$

- Recall that a copula is defined as satisfying:

Dependent case – Copulas

- If X has marginal F, Y has marginal G and they assume a copula C, then (X, Y) has joint distribution

$$H(x, y) = C(F(x), G(y)).$$

- Recall that a copula is defined as satisfying:
  - C is defined on $[0, 1] \times [0, 1]$;

Dependent case – Copulas

- If X has marginal F, Y has marginal G and they assume a copula C, then (X, Y) has joint distribution

$$H(x, y) = C(F(x), G(y)).$$

- Recall that a copula is defined as satisfying:
  - C is defined on $[0, 1] \times [0, 1]$;
  - $C(s, 0) = C(0, t) = 0$;

Dependent case – Copulas

- If X has marginal F, Y has marginal G and they assume a copula C, then (X, Y) has joint distribution

$$H(x, y) = C(F(x), G(y)).$$

- Recall that a copula is defined as satisfying:
  - C is defined on $[0, 1] \times [0, 1]$;
  - $C(s, 0) = C(0, t) = 0$;
  - $C(s, 1) = s$ and $C(1, t) = t$;

Dependent case – Copulas

- If X has marginal F, Y has marginal G and they assume a copula C, then (X, Y) has joint distribution
$$H(x, y) = C(F(x), G(y)).$$
- Recall that a copula is defined as satisfying:
  - C is defined on $[0, 1] \times [0, 1]$;
  - $C(s, 0) = C(0, t) = 0$;
  - $C(s, 1) = s$ and $C(1, t) = t$;
  - $C(s_2, t_2) - C(s_2, t_1) - C(s_1, t_2) + C(s_1, t_1) \geq 0$.

Dependent case – Copulas

- If X has marginal F, Y has marginal G and they assume a copula C, then (X, Y) has joint distribution

$$H(x, y) = C(F(x), G(y)).$$

- Recall that a copula is defined as satisfying:
  - C is defined on $[0, 1] \times [0, 1]$;
  - $C(s, 0) = C(0, t) = 0$;
  - $C(s, 1) = s$ and $C(1, t) = t$;
  - $C(s_2, t_2) - C(s_2, t_1) - C(s_1, t_2) + C(s_1, t_1) \geq 0$.
- Three examples

Dependent case – Copulas

- If X has marginal F, Y has marginal G and they assume a copula C, then (X, Y) has joint distribution

$$H(x, y) = C(F(x), G(y)).$$

- Recall that a copula is defined as satisfying:
  - C is defined on $[0, 1] \times [0, 1]$;
  - $C(s, 0) = C(0, t) = 0$;
  - $C(s, 1) = s$ and $C(1, t) = t$;
  - $C(s_2, t_2) - C(s_2, t_1) - C(s_1, t_2) + C(s_1, t_1) \geq 0$.
- Three examples
  - $\Pi(s, t) = st$ – independent case;

Dependent case – Copulas

- If X has marginal F, Y has marginal G and they assume a copula C, then (X, Y) has joint distribution

$$H(x,y) = C(F(x), G(y)).$$

- Recall that a copula is defined as satisfying:
  - C is defined on $[0,1] \times [0,1]$;
  - $C(s,0) = C(0,t) = 0$;
  - $C(s,1) = s$ and $C(1,t) = t$;
  - $C(s_2,t_2) - C(s_2,t_1) - C(s_1,t_2) + C(s_1,t_1) \geq 0$.
- Three examples
  - $\Pi(s,t) = st$ – independent case;
  - $M(s,t) = s \wedge t$ – perfectly positively related case;

Dependent case – Copulas

- If X has marginal F, Y has marginal G and they assume a copula C, then (X, Y) has joint distribution
$$H(x, y) = C(F(x), G(y)).$$
- Recall that a copula is defined as satisfying:
  - C is defined on $[0, 1] \times [0, 1]$;
  - $C(s, 0) = C(0, t) = 0$;
  - $C(s, 1) = s$ and $C(1, t) = t$;
  - $C(s_2, t_2) - C(s_2, t_1) - C(s_1, t_2) + C(s_1, t_1) \geq 0.$
- Three examples
  - $\Pi(s, t) = st$ – independent case;
  - $M(s, t) = s \wedge t$ – perfectly positively related case;
  - $W(s, t) = (s + t - 1)^+$ – perfectly negatively related case;

Dependent case – Copulas

- If X has marginal F, Y has marginal G and they assume a copula C, then (X, Y) has joint distribution

$$H(x, y) = C(F(x), G(y)).$$

- Recall that a copula is defined as satisfying:
  - C is defined on $[0, 1] \times [0, 1]$;
  - $C(s, 0) = C(0, t) = 0$;
  - $C(s, 1) = s$ and $C(1, t) = t$;
  - $C(s_2, t_2) - C(s_2, t_1) - C(s_1, t_2) + C(s_1, t_1) \geq 0$.
- Three examples
  - $\Pi(s, t) = st$ – independent case;
  - $M(s, t) = s \wedge t$ – perfectly positively related case;
  - $W(s, t) = (s + t - 1)^+$ – perfectly negatively related case;
  - $K(s, t) = s \wedge t - \psi(s \wedge t) + (s \vee t)\psi(s \wedge t)$ – $(U \wedge W, V \wedge W)$.

Dependent case – A toy example

Dependent case – A toy example

- $n = 2$.

Dependent case – A toy example

- $n = 2$.
- X_1, X_2 take at most 2 values and assume a copula C.

Dependent case – A toy example

- $n = 2$.
- X_1, X_2 take at most 2 values and assume a copula C.
- $p_i = \mathbb{P}(X_i = a_i)$, $\mathbb{P}(X_i = x_i) = 1 - p_i$, $a_i \leq x_i$.

Dependent case – A toy example

- $n = 2$.
- X_1, X_2 take at most 2 values and assume a copula C.
- $p_i = \mathbb{P}(X_i = a_i)$, $\mathbb{P}(X_i = x_i) = 1 - p_i$, $a_i \leq x_i$.
- X_i^1 and X_i^2 inherit the copula of X_1 and X_2, C:

$$\begin{aligned}
\mathbb{P}(X_i^1 = a_i, X_i^2 = a_i) &= C(p_i, p_i) \\
\mathbb{P}(X_i^1 = a_i, X_i^2 = x_i) &= p_i - C(p_i, p_i) \\
\mathbb{P}(X_i^1 = x_i, X_i^2 = a_i) &= p_i - C(p_i, p_i) \\
\mathbb{P}(X_i^1 = x_i, X_i^2 = x_i) &= 1 - 2p_i + C(p_i, p_i)
\end{aligned}$$

Dependent case – A toy example

- $n = 2$.
- X_1, X_2 take at most 2 values and assume a copula C.
- $p_i = \mathbb{P}(X_i = a_i)$, $\mathbb{P}(X_i = x_i) = 1 - p_i$, $a_i \leq x_i$.
- X_i^1 and X_i^2 inherit the copula of X_1 and X_2, C:

$$\begin{aligned}
\mathbb{P}(X_i^1 = a_i, X_i^2 = a_i) &= C(p_i, p_i) \\
\mathbb{P}(X_i^1 = a_i, X_i^2 = x_i) &= p_i - C(p_i, p_i) \\
\mathbb{P}(X_i^1 = x_i, X_i^2 = a_i) &= p_i - C(p_i, p_i) \\
\mathbb{P}(X_i^1 = x_i, X_i^2 = x_i) &= 1 - 2p_i + C(p_i, p_i)
\end{aligned}$$

- $M_i = \mathbb{E}[\max(X_i^1, X_i^2)] = C(p_i, p_i)a_i + (1 - C(p_i, p_i))x_i$, $i = 1, 2$.

Dependent case – A toy example

Dependent case – A toy example

Assumption

We assume that for any (s,t),

$$C(s,t) - sC(t,t) \geq 0 \text{ and } C(s,t) - tC(s,s) \geq 0. \qquad (\star)$$

- Π, M and K satisfy this condition; W does not.

Dependent case – A toy example

Assumption

We assume that for any (s, t),

$$C(s,t) - sC(t,t) \geq 0 \text{ and } C(s,t) - tC(s,s) \geq 0. \qquad (\star)$$

- Π, M and K satisfy this condition; W does not.
- If (U, V) are uniform $(0, 1)$ and have copula C, then $(\star)$ translates to

$$\mathbb{P}(U \leq s | \max(U,V) \leq t) \geq \mathbb{P}(U \leq s), \quad s < t.$$

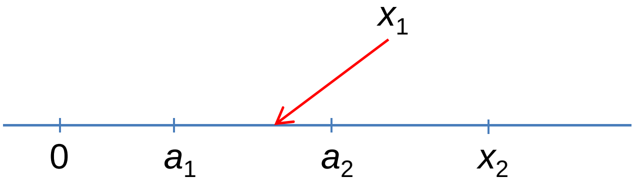
Dependent case – A toy example

Dependent case – A toy example

Assume WLOG that $a_1 \leq a_2$. We need to consider 3 cases:

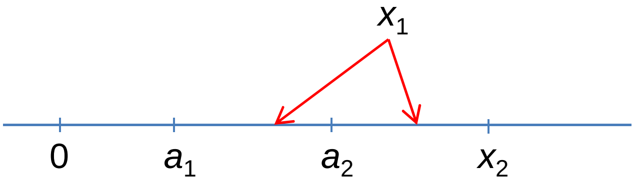
Dependent case – A toy example

Assume WLOG that $a_1 \leq a_2$. We need to consider 3 cases:



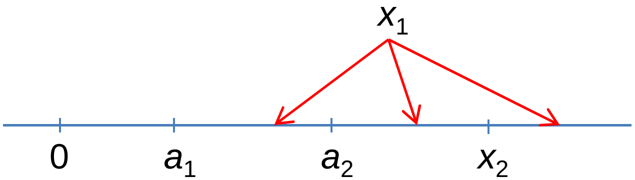
Dependent case – A toy example

Assume WLOG that $a_1 \leq a_2$. We need to consider 3 cases:



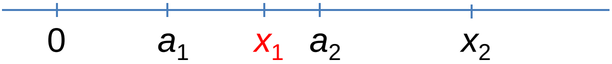
Dependent case – A toy example

Assume WLOG that $a_1 \leq a_2$. We need to consider 3 cases:



Dependent case – A toy example

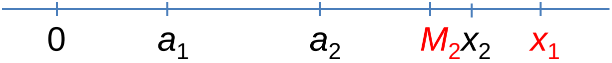
Dependent case – A toy example



The case $a_1 \leq x_1 \leq a_2 \leq x_2$ is trivial since in this case X_2 dominates X_1.

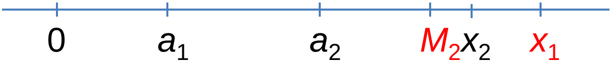
Dependent case – A toy example

Dependent case – A toy example



Assume $a_1 \leq a_2 \leq x_2 \leq x_1$.

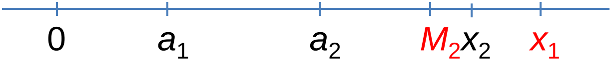
Dependent case – A toy example



Assume $a_1 \leq a_2 \leq x_2 \leq x_1$.

$$\begin{aligned}
&\mathbb{E}[\max(X_1, M_2)] - \mathbb{E}[\max(X_1, X_2)] \\
&\quad = \quad p_1 M_2 - C(p_1, p_2) a_2 - (p_1 - C(p_1, p_2)) x_2 \\
&\quad = \quad \Big(C(p_1, p_2) - p_1 C(p_2, p_2) \Big) (x_2 - a_2) \ \geq \ 0.
\end{aligned}$$

Dependent case – A toy example



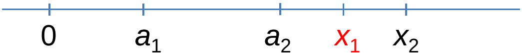
Assume $a_1 \leq a_2 \leq x_2 \leq x_1$.

$$
\begin{aligned}
&\mathbb{E}[\max(X_1, M_2)] - \mathbb{E}[\max(X_1, X_2)] \\
&\quad = \ p_1 M_2 - C(p_1, p_2)a_2 - (p_1 - C(p_1, p_2))x_2 \\
&\quad = \ \Big(C(p_1, p_2) - p_1 C(p_2, p_2)\Big)(x_2 - a_2) \ \geq \ 0.
\end{aligned}
$$

Therefore we may replace X_2 with M_2.

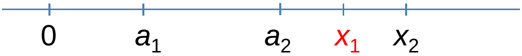
Dependent case – A toy example

Dependent case – A toy example



Assume $a_1 \leq a_2 \leq x_1 \leq x_2$.

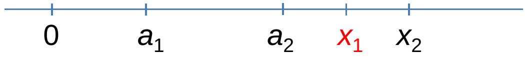
Dependent case – A toy example



Assume $a_1 \leq a_2 \leq x_1 \leq x_2$.
We vary a_2 and x_2 keeping p_2 (and a_1, x_1, p_1) constant:

$$\mathbb{E}[\max(X_2^1, X_2^2)] = C(p_2, p_2)a_2 + (1 - C(p_2, p_2))x_2 = M_2.$$

Dependent case – A toy example



Assume $a_1 \leq a_2 \leq x_1 \leq x_2$.
We vary a_2 and x_2 keeping p_2 (and a_1, x_1, p_1) constant:

$$\mathbb{E}[\max(X_2^1, X_2^2)] = C(p_2, p_2)a_2 + (1 - C(p_2, p_2))x_2 = M_2.$$

Then, the linear function

$$\mathbb{E}[\max(X_1, X_2)] = C(p_1, p_2)a_2 + (1 - p_2)x_2 + (p_2 - C(p_1, p_2))x_1$$

is maximum at one of the 3 boundary points.

Dependent case – A toy example

Dependent case – A toy example

Assume $a_1 \leq a_2 \leq x_1 \leq x_2$.
We vary a_2 and x_2 keeping p_2 (and a_1, x_1, p_1) constant:

$$\mathbb{E}[\max(X_2^1, X_2^2)] = C(p_2, p_2)a_2 + (1 - C(p_2, p_2))x_2 = M_2.$$

Then, the linear function

$$\mathbb{E}[\max(X_1, X_2)] = C(p_1, p_2)a_2 + (1 - p_2)x_2 + (p_2 - C(p_1, p_2))x_1$$

is maximum at one of the 3 boundary points.

Dependent case – A toy example

- $a_2 = x_1$. In this case X_2 dominates X_1.

Dependent case – A toy example

- $a_2 = x_1$. In this case X_2 dominates X_1.
- $a_2 = a_1$. In this case we may collapse X_1 into M_1.

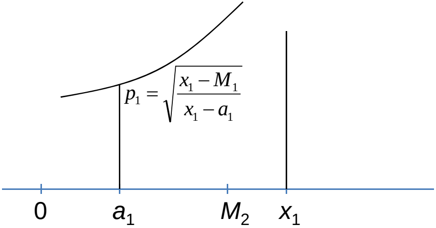
Dependent case – A toy example

- $a_2 = x_1$. In this case X_2 dominates X_1.
- $a_2 = a_1$. In this case we may collapse X_1 into M_1.
- $x_2 = x_1$. In this case we may collapse X_2 into M_2.

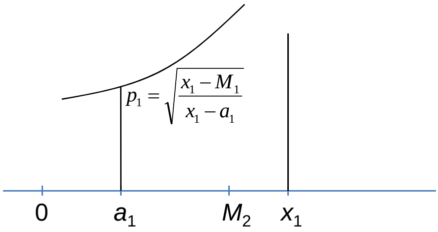
Dependent case – A toy example

- $a_2 = x_1$. In this case X_2 dominates X_1.
- $a_2 = a_1$. In this case we may collapse X_1 into M_1.
- $x_2 = x_1$. In this case we may collapse X_2 into M_2.
- In any case, we may assume that $X_2 = M_2$ and $a_1 \leq M_2 \leq x_1$.

Dependent case – A toy example



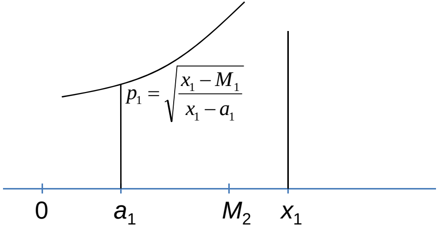
Dependent case – A toy example



We vary a_1 and p_1 keeping x_1 constant:

$$\mathbb{E}[\max(X_1^1, X_1^2)] = C(p_1, p_1)a_1 + (1 - C(p_1, p_1))x_1 = M_1.$$

Dependent case – A toy example



We vary a_1 and p_1 keeping x_1 constant:

$$\mathbb{E}[\max(X_1^1, X_1^2)] = C(p_1, p_1)a_1 + (1 - C(p_1, p_1))x_1 = M_1.$$

Then,

$$\mathbb{E}[\max(X_1, M_2)] = p_1 M_2 + (1 - p_1)x_1 = x_1 - (x_1 - M_2)p_1$$

is maximum for p_1 minimum i.e. $a_1 = 0$.

Dependent case – A toy example

Therefore we may assume that $X_2 = M_2$, $a_1 = 0$ and $0 \leq M_2 \leq x_1$.

Dependent case – A toy example

Therefore we may assume that $X_2 = M_2$, $a_1 = 0$ and $0 \leq M_2 \leq x_1$. In this case

$$\mathbb{E}[\max(X_1, X_2)] = p_1 M_2 + (1 - p_1) \frac{M_1}{1 - C(p_1, p_1)}$$

Theorem

$$\mathbb{E}[\max(X_1, X_2)] \leq \sup_{0 \leq r < 1} \left(M_2 r + M_1 \frac{1 - r}{1 - C(r, r)} \right).$$

The upper bound for Π, M and K

The upper bound for Π, M and K

Let $\gamma(r) = C(r, r)$ and assume that $M_1 \leq M_2$.

- $C = \Pi$. In this case $\gamma(r) = r^2$, $\gamma'(1) = 2$ and

$$\mathbb{E}[\max(X_1, X_2)] \leq \frac{1}{2} M_1 + M_2.$$

The upper bound for Π, M and K

Let $\gamma(r) = C(r, r)$ and assume that $M_1 \leq M_2$.

- $C = \Pi$. In this case $\gamma(r) = r^2$, $\gamma'(1) = 2$ and

$$\mathbb{E}[\max(X_1, X_2)] \leq \frac{1}{2} M_1 + M_2.$$

- $C = M$. In this case $\gamma(r) = r$, $\gamma'(1) = 1$ and

$$\mathbb{E}[\max(X_1, X_2)] \leq M_1 + M_2.$$

The upper bound for Π, M and K

Let $\gamma(r) = C(r,r)$ and assume that $M_1 \leq M_2$.

- $C = \Pi$. In this case $\gamma(r) = r^2$, $\gamma'(1) = 2$ and

$$\mathbb{E}[\max(X_1, X_2)] \leq \frac{1}{2}M_1 + M_2.$$

- $C = M$. In this case $\gamma(r) = r$, $\gamma'(1) = 1$ and

$$\mathbb{E}[\max(X_1, X_2)] \leq M_1 + M_2.$$

- $C = K$. In this case $\gamma(r) = r - \psi(r) + r\psi(r)$, $\gamma'(1) = 2$ and

$$\mathbb{E}[\max(X_1, X_2)] \leq \frac{1}{2}M_1 + M_2.$$

The upper bound for Π, M and K

Let $\gamma(r) = C(r, r)$ and assume that $M_1 \leq M_2$.

- $C = \Pi$. In this case $\gamma(r) = r^2$, $\gamma'(1) = 2$ and

$$\mathbb{E}[\max(X_1, X_2)] \leq \frac{1}{2} M_1 + M_2.$$

- $C = M$. In this case $\gamma(r) = r$, $\gamma'(1) = 1$ and

$$\mathbb{E}[\max(X_1, X_2)] \leq M_1 + M_2.$$

- $C = K$. In this case $\gamma(r) = r - \psi(r) + r\psi(r)$, $\gamma'(1) = 2$ and

$$\mathbb{E}[\max(X_1, X_2)] \leq \frac{1}{2} M_1 + M_2.$$

- Note that the definition of M_i depends on C and the three bounds cannot be compared.

References

- Arnold B.C. and Groeneveld R.A. (1979), Bounds on Expectations of Linear Systematic Statistics Based on Dependent Samples, *Ann. Stat.* **7**, pp 220–223..
- Balakrishnan N. and Balasubramanian K. (2008), Revisiting Sen's inequalities on order statistics, *Statist. Probab. Letters* **78**, pp 616–621.
- David H.A. and Nagaraja H.N. (2003), *Order Statistics*, John Wiley & Sons, Hoboken, New Jersey, 3rd ed.
- Hamza K., Jagers P., Sudbury A. and Tokarev D. (2009), The mixing advantage is less than 2, *Extremes* **12**, no. 1, pp 19–31.
- Hamza K. and Sudbury A. (2011), The mixing advantage for bounded random variables, *Statist. Probab. Letters* **81**, no. 8, pp 1190–1195.

Refrences

- Hartley H.O. and David H.A. (1954), Universal Bounds for Mean Range and Extreme Observation, *Ann. Math. Stat.* **25**, pp 85–89..
- Gumbel E.J. (1954), The Maxima of the Mean Largest Value and of the Range, *Ann. Math. Stat.* **25**, pp 76–84..
- Sen P.K. (1970), A Note on Order Statistics for Heterogeneous Distributions, *Ann. Math. Stat.* **41**, pp 2137–2139..
- Tokarev D. (2007), Galton-Watson processes and extinction in population systems, PhD Thesis, Monash University.