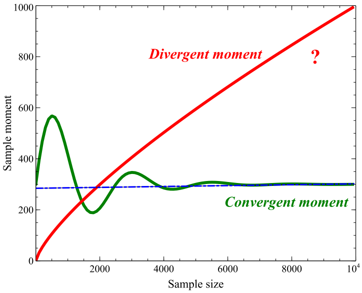
有限矩发散条件下的统计分析
A Statistical Analysis of Diverging Moments

郭方健
richardkwo@gmail.com
指导教师: 周涛

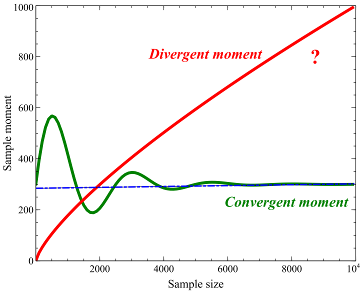
英才实验学院 2009 级 1 班

2013 年 6 月 7 日

Overview


Divergent moment
?
Convergent moment
Sample moment
Sample size
0
200
400
600
800
1000
2000
4000
6000
8000
10^4

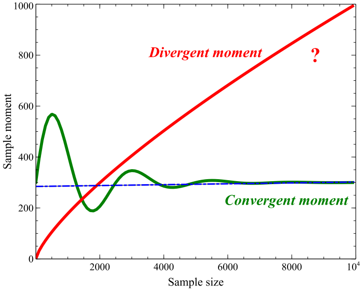
Overview



Question

- How do diverging moments grow with sample size?

Overview



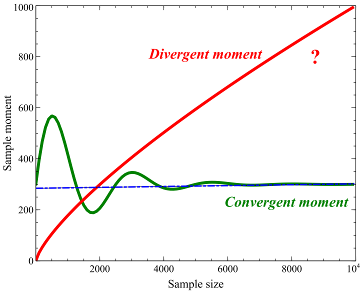
Question

- How do diverging moments grow with sample size?

Method

- EPM (Equiprobable Partition Method) $\rightarrow$ *Asymptotics*
- The precision of EPM moment estimators

Overview



Question

- How do diverging moments grow with sample size?

Method

- EPM (Equiprobable Partition Method) $\rightarrow$ *Asymptotics*
- The precision of EPM moment estimators

Application

- Theoretical approximation for statistics with several moments
- Memory constraints for power-law series

Moments

Moments are statistics that characterize

- the **shape** of **distributions** (e.g. *mean*, *variance*, *skewness*)
- the **interdependence** among variables (e.g. *Pearson's correlation coefficient*)

Moments

Moments are statistics that characterize

- the **shape** of **distributions** (e.g. *mean*, *variance*, *skewness*)
- the **interdependence** among variables (e.g. *Pearson's correlation coefficient*)

A moment corresponds to a **sample moment** and a **population moment**.

k-th moment

Sample moment *(algebraic)*

$$m_k(c) = \frac{1}{n}\sum_{i=1}^{n}(x_i - c)^k$$

Population moment *(integral)*

$$\mu_k(c) = \mathrm{E}[(X-c)^k] = \int_{-\infty}^{+\infty}(x-c)^k dF(x)$$

Moments

Moments are statistics that characterize

- the **shape** of **distributions** (e.g. *mean*, *variance*, *skewness*)
- the **interdependence** among variables (e.g. *Pearson's correlation coefficient*)

A moment corresponds to a **sample moment** and a **population moment**.

k-th moment

Sample moment *(algebraic)*

$$m_k(c) = \frac{1}{n}\sum_{i=1}^{n}(x_i - c)^k$$

Population moment *(integral)*

$$\mu_k(c) = \mathrm{E}[(X-c)^k] = \int_{-\infty}^{+\infty}(x-c)^k dF(x)$$

- $c = 0$: **raw moment**
- $c =$ mean: **central moment**

Yingcai Experimental School

Relation between two moments: convergent case

For i.i.d. $\{x_1, x_2, \cdots, x_n\}$, when the population moment (*Riemann-Stieltjes integral*) **converges**, we know that the corresponding *sample moment* is an **unbiased estimators** for the *population moment*, i.e.

$$\mathrm{E}[m_k(c)] = \mu_k(c).$$

Relation between two moments: convergent case

For i.i.d. $\{x_1, x_2, \cdots, x_n\}$, when the population moment (*Riemann-Stieltjes integral*) **converges**, we know that the corresponding *sample moment* is an **unbiased estimators** for the *population moment*, i.e.

$$\mathrm{E}[m_k(c)] = \mu_k(c).$$

Then, by *the Weak Law of Large Numbers*, we have

Theorem

A sample moment $m_k(c)$ ***converges in probability towards*** *the corresponding population moment $\mu_k(c)$ when the sample size $n \to \infty$, namely*

$$\lim_{n \to \infty} P(|m_k(c) - \mu_k(c)| > \epsilon) = 0, \ \forall \epsilon > 0.$$

given that μ'_k exists (is convergent).

In other words, the estimator is ***more accurate with more data***.

英才實驗學院 Yingcai Experimental School

Yet, what if population moments diverge?

But a population moment may ***diverge*** for a **heavy-tailed distribution** (e.g. *power-law, log-normal, Weibull*).

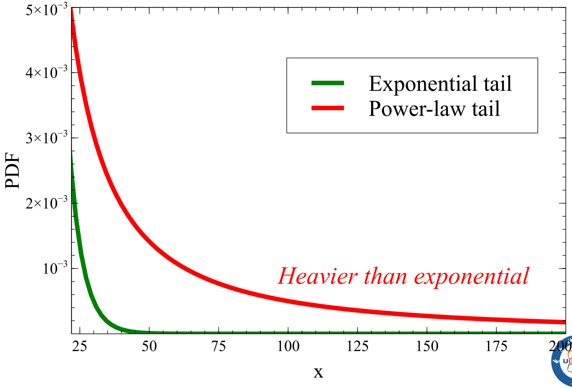
$$\int_{-\infty}^{+\infty} x^k f(x)\, dx = \infty, \text{ for } k > k_0.$$

Yingcai Experimental School

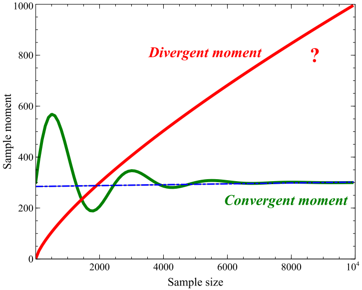
Yet, what if population moments diverge?

But a population moment may ***diverge*** for a **heavy-tailed distribution** (e.g. *power-law*, *log-normal*, *Weibull*).

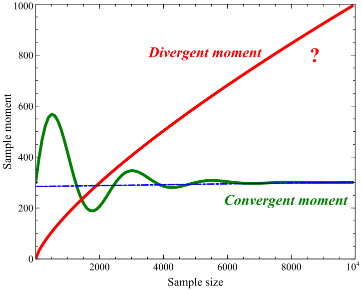
$$\int_{-\infty}^{+\infty} x^k f(x)\,dx = \infty, \text{ for } k > k_0.$$



Yet, what if population moments diverge?



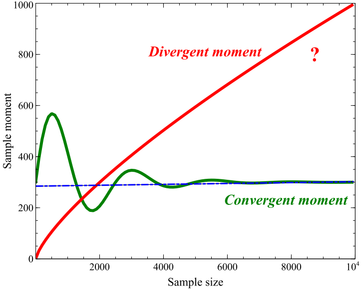
Yet, what if population moments diverge?



Objective

- Find the asymptotics for diverging moments in the form of $n^\gamma g(n)$.

Yet, what if population moments diverge?



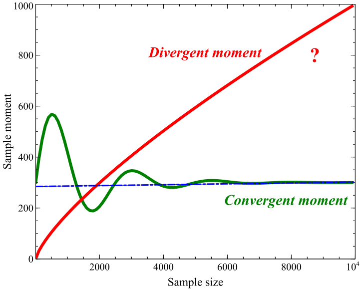
Objective

- Find the asymptotics for diverging moments in the form of $n^{\gamma} g(n)$.

Application

- A theoretical reference for diverging moments.

Yet, what if population moments diverge?



Objective

- Find the asymptotics for diverging moments in the form of $n^\gamma g(n)$.

Application

- A theoretical reference for diverging moments.
- Approximation for statistics involving several moments, e.g.

$$A = \frac{m_{k_1}}{m_{k_2}} \approx \frac{g_1(n)n^{\gamma_1}}{g_2(n)n^{\gamma_2}} \ (n \to \infty).$$

Equiprobable Partition Method (EPM)

Equiprobable Partition Method (EPM)

Equiprobable Partition Method is presented as a tool for constructing the **asymptotics of diverging sample moments**, i.e., answering — "How do diverging sample moments **grow** with n ?"

Equiprobable Partition Method (EPM)

Equiprobable Partition Method is presented as a tool for constructing the **asymptotics of diverging sample moments**, i.e., answering — "How do diverging sample moments **grow** with n ?"

EPM consists of 3 steps:

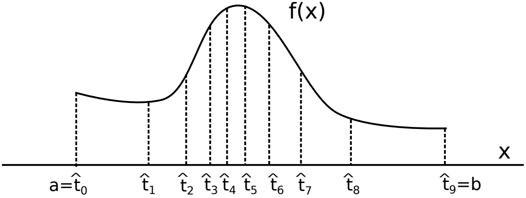
1. Constructing equiprobable partitions
2. Forming moment estimators
(substituting samples with representative points)
3. Rewriting into the asymptotic form

Step 1: Equiprobable partitions

Definition

Let X be a continuous random variable defined on the interval $[a, b] \in \mathbb{R}$ with the PDF $f(x)$ $(a \leq x \leq b)$, then its n-**separated equiprobable partition** $\mathcal{P}_n(a, b)$ is defined as $\mathcal{P}_n(a, b) = (\hat{t}_0, \hat{t}_1, \cdots, \hat{t}_n)$, where $\hat{t}_0 = a$, $\hat{t}_n = b$ and it satisfies that

$$\int_{\hat{t}_i}^{\hat{t}_{i+1}} f(x)\,dx = \frac{1}{n} \quad (i = 1, 2, \cdots, n-1).$$

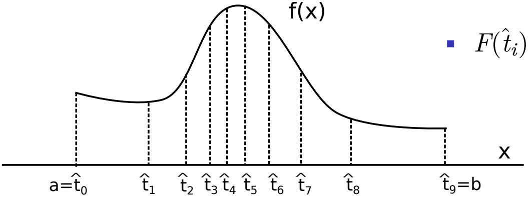


Step 1: Equiprobable partitions

Definition

Let X be a continuous random variable defined on the interval $[a, b] \in \mathbb{R}$ with the PDF $f(x)$ $(a \leq x \leq b)$, then its n-**separated equiprobable partition** $\mathcal{P}_n(a, b)$ is defined as $\mathcal{P}_n(a, b) = (\hat{t}_0, \hat{t}_1, \cdots, \hat{t}_n)$, where $\hat{t}_0 = a$, $\hat{t}_n = b$ and it satisfies that

$$\int_{\hat{t}_i}^{\hat{t}_{i+1}} f(x)\,dx = \frac{1}{n} \quad (i = 1, 2, \cdots, n-1).$$



- $F(\hat{t}_i) = \frac{i}{n} \quad (i = 0, 1, \cdots, n)$

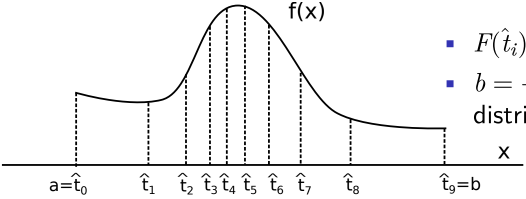
英才實驗學院
Yingcai Experimental School

Step 1: Equiprobable partitions

Definition

Let X be a continuous random variable defined on the interval $[a, b] \in \mathbb{R}$ with the PDF $f(x)$ $(a \leq x \leq b)$, then its n-**separated equiprobable partition** $\mathcal{P}_n(a, b)$ is defined as $\mathcal{P}_n(a, b) = (\hat{t}_0, \hat{t}_1, \cdots, \hat{t}_n)$, where $\hat{t}_0 = a$, $\hat{t}_n = b$ and it satisfies that

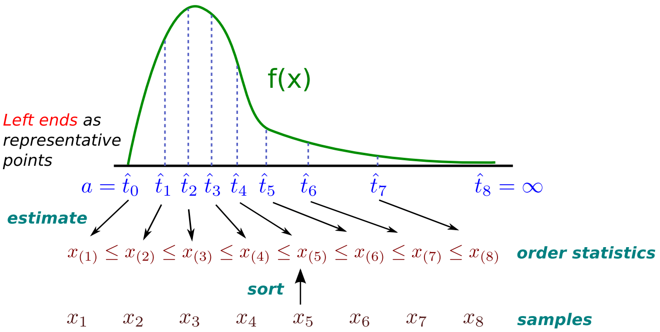
$$\int_{\hat{t}_i}^{\hat{t}_{i+1}} f(x)\, dx = \frac{1}{n} \quad (i = 1, 2, \cdots, n-1).$$



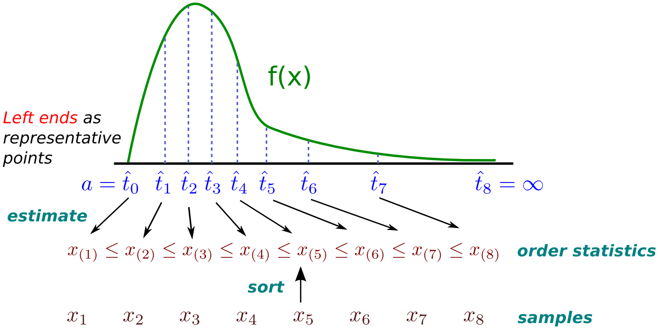
- $F(\hat{t}_i) = \frac{i}{n} \quad (i = 0, 1, \cdots, n)$
- $b = +\infty$ for a right-heavy-tailed distribution

Yingcai Experimental School

Step 2: Moment estimators via sample substitution



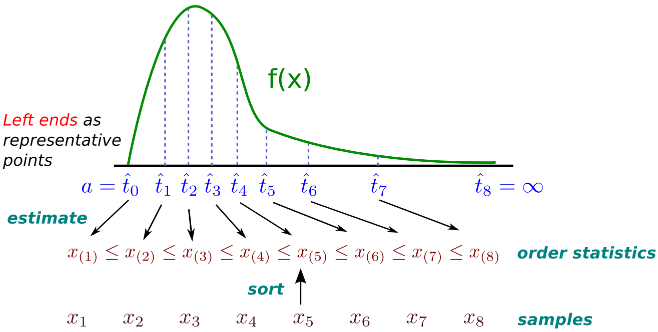
Step 2: Moment estimators via sample substitution



Sample moment

$$m_k(c) = \frac{1}{n}\sum_{i=1}^{n}(x_i - c)^k$$

Step 2: Moment estimators via sample substitution



Sample moment

$$m_k(c) = \frac{1}{n}\sum_{i=1}^{n}(x_i - c)^k$$

Substituting $x_{(i)}$ with $\hat{t}_{i-1}$,
EPM moment estimator

$$\hat{m}_k(n; c) = \frac{1}{n}\sum_{i=1}^{n}(\hat{t}_{i-1} - c)^k$$

Step 3: Rewriting into the asymptotic form

Then, we rewrite $\hat{m}_k(n)$ (c is left out for simplicity) into the form of

$$\hat{m}_k(n) = n^{\gamma} g(n),$$

Step 3: Rewriting into the asymptotic form

Then, we rewrite $\hat{m}_k(n)$ (c is left out for simplicity) into the form of

$$\hat{m}_k(n) = n^\gamma g(n),$$

where

- n^γ $(\gamma > 0)$ is the **leading order** that characterizes **the speed of convergence**.

Yingcai Experimental School

Step 3: Rewriting into the asymptotic form

Then, we rewrite $\hat{m}_k(n)$ (c is left out for simplicity) into the form of

$$\hat{m}_k(n) = n^\gamma g(n),$$

where

- n^γ ($\gamma > 0$) is the **leading order** that characterizes **the speed of convergence**.
- $g(n)$ is a function that satisfies

  $$0 < \lim_{n\to\infty} |g(n)| < \infty,$$

  which is the "**convergent remainder**" left by taking off the effect of divergence from the sample moment.

Yingcai Experimental School

Why equiprobable?

Rationales

Why equiprobable?

Rationales

1. EPM moments estimators **coincide** with population moments in the **convergent** case.

Theorem

Let X *be a random variable defined on the interval* $[a, b]$ *with its PDF* $f(x)$ $(a \leq x \leq b)$*. Supposing there exists some* M *that it holds that* $\forall x \in [a, b]$, $f(x) < M$*, we have*

$$\lim_{n \to \infty} \hat{m}'_k(n) = \mu'_k.$$

Why equiprobable?

Rationales

1. EPM moments estimators **coincide** with population moments in the **convergent** case.

Theorem

Let X be a random variable defined on the interval $[a, b]$ with its PDF $f(x)$ $(a \leq x \leq b)$. Supposing there exists some M that it holds that $\forall x \in [a, b]$, $f(x) < M$, we have

$$\lim_{n\to\infty} \hat{m}'_k(n) = \mu'_k.$$

2. If viewing *partition* as **discretizing** a continuous distribution, then EPM maximizes the entropy $H = -\sum_{i=1}^{n} p_i \log(p_i)$.

Why equiprobable?

Rationales

1. EPM moments estimators **coincide** with population moments in the **convergent** case.

Theorem

Let X *be a random variable defined on the interval* $[a, b]$ *with its PDF* $f(x)$ $(a \leq x \leq b)$. *Supposing there exists some* M *that it holds that* $\forall x \in [a, b]$, $f(x) < M$, *we have*

$$\lim_{n \to \infty} \hat{m}'_k(n) = \mu'_k.$$

2. If viewing *partition* as **discretizing** a continuous distribution, then EPM maximizes the entropy $H = -\sum_{i=1}^{n} p_i \log(p_i)$.
3. Getting $\hat{t}_i$ is easy with $F^{-1}(x)$.

Yingcai Experimental School

Example: Asymptotics for moments of power-law

EPM asymptotics for diverging moments of the power-law distribution.

Example: Asymptotics for moments of power-law

EPM asymptotics for diverging moments of the power-law distribution.
Power-law with $x_{\min} = 1$ and $\alpha > 1$,

$$\text{PDF } f(x) = (\alpha - 1)x^{-\alpha}, \quad \text{CDF } F(x) = 1 - x^{1-\alpha}.$$

Example: Asymptotics for moments of power-law

EPM asymptotics for diverging moments of the power-law distribution. Power-law with $x_{\min} = 1$ and $\alpha > 1$,

$$\text{PDF } f(x) = (\alpha - 1)x^{-\alpha}, \quad \text{CDF } F(x) = 1 - x^{1-\alpha}.$$

1. n-separated equiprobable partition $\mathcal{P}_n(1, +\infty) = \{\hat{t}_1, \hat{t}_2, \cdots, \hat{t}_n\}$. From

$$F(\hat{t}_i) = 1 - (\hat{t}_i)^{1-\alpha} = \frac{i}{n},$$

we have

$$\hat{t}_i = (1 - \frac{i}{n})^{-c} \quad (i = 0, 1, \cdots, n-1),$$

where $c = \frac{1}{\alpha - 1} > \frac{1}{2}$ is a shorthand.

Example: Asymptotics for moments of power-law

2. The EPM estimator for m'_k is

$$\hat{m}'_k(n) = \frac{1}{n}\sum_{i=0}^{n-1}(\hat{t}_i)^k = \frac{1}{n}\sum_{i=0}^{n-1}(1-\frac{i}{n})^{-ck} \quad (c = \frac{1}{\alpha-1}).$$

Example: Asymptotics for moments of power-law

2. The EPM estimator for m'_k is

$$\hat{m}'_k(n) = \frac{1}{n}\sum_{i=0}^{n-1}(\hat{t}_i)^k = \frac{1}{n}\sum_{i=0}^{n-1}(1-\frac{i}{n})^{-ck} \quad (c = \frac{1}{\alpha-1}).$$

3. Rewriting it into

$$\hat{m}'_k(n) = \frac{1}{n}\sum_{i=0}^{n-1} n^{ck}(n-i)^{-ck} = n^{ck-1}\sum_{i=1}^{n}\frac{1}{i^{ck}},$$

Example: Asymptotics for moments of power-law

2. The EPM estimator for m'_k is

$$\hat{m}'_k(n) = \frac{1}{n}\sum_{i=0}^{n-1}(\hat{t}_i)^k = \frac{1}{n}\sum_{i=0}^{n-1}(1-\frac{i}{n})^{-ck} \quad (c = \frac{1}{\alpha-1}).$$

3. Rewriting it into

$$\hat{m}'_k(n) = \frac{1}{n}\sum_{i=0}^{n-1} n^{ck}(n-i)^{-ck} = n^{ck-1}\sum_{i=1}^{n}\frac{1}{i^{ck}},$$

where

- n^{ck-1} is the leading order,

Example: Asymptotics for moments of power-law

2. The EPM estimator for m'_k is

$$\hat{m}'_k(n) = \frac{1}{n}\sum_{i=0}^{n-1}(\hat{t}_i)^k = \frac{1}{n}\sum_{i=0}^{n-1}(1-\frac{i}{n})^{-ck} \quad (c = \frac{1}{\alpha-1}).$$

3. Rewriting it into

$$\hat{m}'_k(n) = \frac{1}{n}\sum_{i=0}^{n-1} n^{ck}(n-i)^{-ck} = n^{ck-1}\sum_{i=1}^{n}\frac{1}{i^{ck}},$$

where

- n^{ck-1} is the leading order,
- $\sum_{i=1}^{n}\frac{1}{i^{ck}} \to$ *Riemann* $\zeta(ck)$ $(n \to \infty)$ is the remaining "convergent term".

Example: Asymptotics for moments of power-law

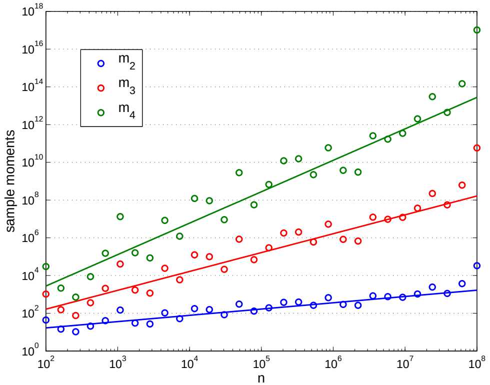


Figure : The comparison between numerical sample moments m_2, m_3, m_4 (circle) and their EPM asymptotics (line) for power-law with $\alpha = 2.5$.

英才實驗學院
Yingcai Experimental School

Time series and their distributions

Time series and their distributions

Time series — an ordered sequence of random variables

$$(t_1, t_2, \cdots, t_n).$$

Time series and their distributions

Time series — an ordered sequence of random variables

$$(t_1, t_2, \cdots, t_n).$$

Distribution charaterization

1. **Marginal** distribution, often assumed to be **stable**

$$f_{t_1}(t) = f_{t_2}(t) = \cdots = f_{t_n}(t).$$

Yingcai Experimental School

Time series and their distributions

Time series — an ordered sequence of random variables

$$(t_1, t_2, \cdots, t_n).$$

Distribution charaterization

1. **Marginal** distribution, often assumed to be **stable**

$$f_{t_1}(t) = f_{t_2}(t) = \cdots = f_{t_n}(t).$$

2. **Interdependence** among elements.
Memory (*1st-order autocorrelation*) is defined as the Pearson's correlation coefficient between $\{t_2, t_3, \cdots, t_n\}$ and its lag-1 counterpart $\{t_1, t_2, \cdots, t_{n-1}\}$, i.e.

$$M = \frac{1}{n-1}\sum_{i=1}^{n-1}\frac{(t_i - m_1)(t_{i+1} - m_2)}{\sigma_1\sigma_2} \quad (-1 \leq M \leq 1).$$

英才实验学院
Yingcai Experimental School

Correlated series with a specified marginal

Correlated series with a specified marginal

Modeling correlated series with marginal distribution $f(t)$.

Correlated series with a specified marginal

Modeling correlated series with marginal distribution $f(t)$.

1. Sampling i.i.d. $(t_1, t_2, \cdots, t_n)$ from the same marginal $f(t)$.

Correlated series with a specified marginal

Modeling correlated series with marginal distribution $f(t)$.

1. Sampling i.i.d. $(t_1, t_2, \cdots, t_n)$ from the same marginal $f(t)$.
2. Sorting samples to $t_{(1)} \leq t_{(2)} \leq \cdots \leq t_{(n)}$.

英才實驗學院 Yingcai Experimental School

Correlated series with a specified marginal

Modeling correlated series with marginal distribution $f(t)$.

1. Sampling i.i.d. $(t_1, t_2, \cdots, t_n)$ from the same marginal $f(t)$.
2. Sorting samples to $t_{(1)} \leq t_{(2)} \leq \cdots \leq t_{(n)}$.
3. Modifying interdependence with a **permutation** $\theta : \{1, \cdots, n\} \leftrightarrow \{1, \cdots, n\}$, so we have $(t_{(\theta_1)}, t_{(\theta_2)}, \cdots, t_{(\theta_n)})$.

[1] M. Hallin, G. Melard, X. Milhaud. Permutational extreme values of autocorrelation coefficients and a Pitman test against serial dependence. *Annals of Statistics*, 1992.

Correlated series with a specified marginal

Modeling correlated series with marginal distribution $f(t)$.

1. Sampling i.i.d. $(t_1, t_2, \cdots, t_n)$ from the same marginal $f(t)$.
2. Sorting samples to $t_{(1)} \leq t_{(2)} \leq \cdots \leq t_{(n)}$.
3. Modifying interdependence with a **permutation** $\theta : \{1, \cdots, n\} \leftrightarrow \{1, \cdots, n\}$, so we have $(t_{(\theta_1)}, t_{(\theta_2)}, \cdots, t_{(\theta_n)})$.

Interestingly, there are **fixed** $\theta_{\max}$ and $\theta_{\min}$ that maximizes and minimizes M for the permuted series [1].

[1] M. Hallin, G. Melard, X. Milhaud. Permutational extreme values of autocorrelation coefficients and a Pitman test against serial dependence. *Annals of Statistics*, 1992.

Correlated series with a specified marginal

Modeling correlated series with marginal distribution $f(t)$.

1. Sampling i.i.d. $(t_1, t_2, \cdots, t_n)$ from the same marginal $f(t)$.
2. Sorting samples to $t_{(1)} \leq t_{(2)} \leq \cdots \leq t_{(n)}$.
3. Modifying interdependence with a **permutation** $\theta : \{1, \cdots, n\} \leftrightarrow \{1, \cdots, n\}$, so we have $(t_{(\theta_1)}, t_{(\theta_2)}, \cdots, t_{(\theta_n)})$.

Interestingly, there are **fixed** $\theta_{\max}$ and $\theta_{\min}$ that maximizes and minimizes M for the permuted series [1].

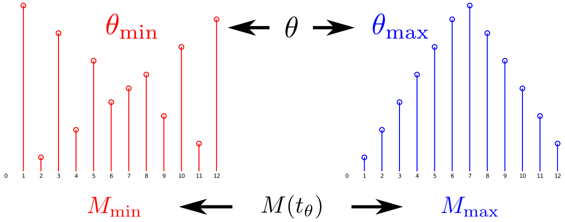
$$\theta_{\max}: \; t_{(1)}, t_{(3)}, \cdots, t_{(2l-1)}, t_{(2l)}, t_{(2l-2)}, \cdots, t_{(4)}, t_{(2)} \quad (n = 2l),$$

$$\theta_{\min}: \; t_{(2l)}, t_{(1)}, t_{(2l-2)}, \cdots, t_{(2l-3)}, t_{(2)}, t_{(2l-1)} \quad (n = 2l).$$

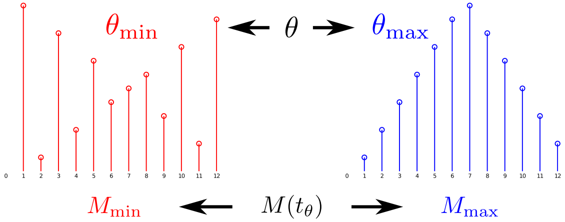
[1] M. Hallin, G. Melard, X. Milhaud. Permutational extreme values of autocorrelation coefficients and a Pitman test against serial dependence. *Annals of Statistics*, 1992.

Yingcai Experimental School

Correlated series with a specified marginal



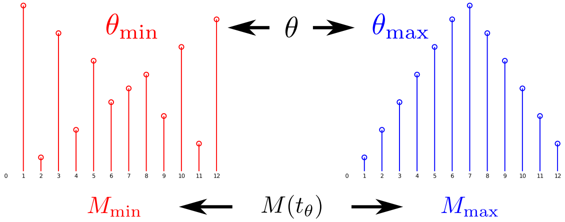
Correlated series with a specified marginal



Therefore, we define the **(tight) bounds** for memory as

$$M_{\max} = \lim_{n\to\infty} \mathrm{E}\left[M(t_{(\theta_{\max})})\right] \text{ and } M_{\min} = \lim_{n\to\infty} \mathrm{E}\left[M(t_{(\theta_{\min})})\right].$$

Correlated series with a specified marginal



Therefore, we define the **(tight) bounds** for memory as

$$M_{\max} = \lim_{n\to\infty} \mathrm{E}\left[M(t_{(\theta_{\max})})\right] \text{ and } M_{\min} = \lim_{n\to\infty} \mathrm{E}\left[M(t_{(\theta_{\min})})\right].$$

For **Gaussian** and **uniform** distributions, we have

$$M_{\max} = 1 \text{ and } M_{\min} = -1 \quad (\text{trivial, as } |M| \le 1).$$

Probabilistic method for the $\alpha > 3$ case

$\alpha > 3$ is necessary for the population variance $\sigma(\alpha)^2$ to converge.

$$\mathrm{E}[M(t_{(\theta)})] = \frac{1}{\sigma(\alpha)^2}\left(\frac{1}{n-1}\sum_{i=1}^{n-1}\mathrm{E}[t_{(\theta_i)}t_{(\theta_{i+1})}] - m(\alpha)^2\right),$$

英才實驗學院 Yingcai Experimental School

Probabilistic method for the $\alpha > 3$ case

$\alpha > 3$ is necessary for the population variance $\sigma(\alpha)^2$ to converge.

$$\mathrm{E}[M(t_{(\theta)})] = \frac{1}{\sigma(\alpha)^2}\left(\frac{1}{n-1}\sum_{i=1}^{n-1}\mathrm{E}[t_{(\theta_i)}t_{(\theta_{i+1})}] - m(\alpha)^2\right),$$

where

$$m(\alpha) = \int_1^{+\infty} xf(x)\,dx = \frac{\alpha-1}{\alpha-2},$$

$$\sigma(\alpha)^2 = \int_1^{+\infty} x^2 f(x)\,dx - m(\alpha)^2 = \frac{\alpha-1}{\alpha-3} - m(\alpha)^2.$$

Probabilistic method for the $\alpha > 3$ case

$\alpha > 3$ is necessary for the population variance $\sigma(\alpha)^2$ to converge.

$$\mathrm{E}[M(t_{(\theta)})] = \frac{1}{\sigma(\alpha)^2}\left(\frac{1}{n-1}\sum_{i=1}^{n-1}\mathrm{E}[t_{(\theta_i)}t_{(\theta_{i+1})}] - m(\alpha)^2\right),$$

where

$$m(\alpha) = \int_1^{+\infty} xf(x)\,dx = \frac{\alpha-1}{\alpha-2},$$

$$\sigma(\alpha)^2 = \int_1^{+\infty} x^2 f(x)\,dx - m(\alpha)^2 = \frac{\alpha-1}{\alpha-3} - m(\alpha)^2.$$

We can use the **distribution for order statistics** to derive

$$M_{\max} = 1,$$

$$M_{\min} = \frac{1}{\sigma(\alpha)^2}\left[2B\left(\frac{1}{2}; \frac{1}{m(\alpha)}, \frac{1}{m(\alpha)}\right) - m(\alpha)^2\right] \quad (\alpha > 3).$$

EPM for the $\alpha \leq 3$ case

When $\alpha \leq 3$, m and σ^2 may diverge, where we can use EPM estimators to tackle the diverging moments in the limit of large n.

英才實驗學院
Yingcai Experimental School

EPM for the $\alpha \leq 3$ case

When $\alpha \leq 3$, m and σ^2 may diverge, where we can use EPM estimators to tackle the diverging moments in the limit of large n.
Rewriting M to

$$M = \frac{s - m^2}{\overline{t^2} - m^2},$$

EPM for the $\alpha \leq 3$ case

When $\alpha \leq 3$, m and σ^2 may diverge, where we can use EPM estimators to tackle the diverging moments in the limit of large n.
Rewriting M to

$$M = \frac{s - m^2}{\overline{t^2} - m^2},$$

where

- $s = \frac{1}{n-1} S = \frac{1}{n-1} \sum_{i=1}^{n-1} t_i t_{i+1}$,
- $\overline{t^2} = \frac{1}{n} \sum_{i=1}^{n} t_i^2$,
- and $m^2 = (\frac{1}{n} \sum_{i=1}^{n} t_i)^2$

are the moments in concern.

Yingcai Experimental School

EPM for the $\alpha \leq 3$ case

$$M = \frac{s - m^2}{\overline{t^2} - m^2}.$$

EPM for the $\alpha \leq 3$ case

$$M = \frac{s - m^2}{\overline{t^2} - m^2}.$$

For $\theta_{\max}$, we have EPM estimators

$$\hat{s}_{\theta_{\max}} = n^{2c-1}\Big[\sum_{k=2}^{n-1}(k^2-1)^{-c} + 2^{-c}\Big] \quad (c = \frac{1}{\alpha - 1}),$$

$$\hat{\overline{t^2}}_{\theta_{\max}} = n^{2c-1}\sum_{k=0}^{n-1}(k+1)^{-2c},$$

Yingcai Experimental School

EPM for the $\alpha \leq 3$ case

$$M = \frac{s - m^2}{\overline{t^2} - m^2}.$$

For $\theta_{\max}$, we have EPM estimators

$$\hat{s}_{\theta_{\max}} = n^{2c-1}\Big[\sum_{k=2}^{n-1}(k^2-1)^{-c} + 2^{-c}\Big] \quad (c = \frac{1}{\alpha - 1}),$$

$$\hat{\overline{t^2}}_{\theta_{\max}} = n^{2c-1}\sum_{k=0}^{n-1}(k+1)^{-2c},$$

while

$$\hat{m}^2 \sim n^{2c-2}.$$

$\hat{m}^2$ can be thrown away.

EPM for the $\alpha \leq 3$ case

Therefore, we have

$$M_{\max} \approx \frac{\lim_{n\to\infty} \sum_{k=2}^{n-1} (k^2 - 1)^{-c} + 2^{-c}}{\zeta(2c)} \quad \left(c = \frac{1}{\alpha - 1},\ 1 < \alpha \leq 3\right).$$

Yingcai Experimental School

EPM for the $\alpha \leq 3$ case

Therefore, we have

$$M_{\max} \approx \frac{\lim_{n\to\infty} \sum_{k=2}^{n-1} (k^2-1)^{-c} + 2^{-c}}{\zeta(2c)} \quad \left(c = \frac{1}{\alpha - 1},\ 1 < \alpha \leq 3\right).$$

For $\theta_{\min}$, in $M = \frac{s-m^2}{\overline{t^2}-m^2}$, $\overline{\hat{t}^2}_{\theta_{\min}} \sim n^{2c-1}$ is the term diverging with the **biggest** order of n.

Yingcai Experimental School

EPM for the $\alpha \leq 3$ case

Therefore, we have

$$M_{\max} \approx \frac{\lim_{n\to\infty} \sum_{k=2}^{n-1} (k^2-1)^{-c} + 2^{-c}}{\zeta(2c)} \quad \left(c = \frac{1}{\alpha - 1},\ 1 < \alpha \leq 3\right).$$

For $\theta_{\min}$, in $M = \frac{s-m^2}{\overline{t^2}-m^2}$, $\overline{\hat{t}^2}_{\theta_{\min}} \sim n^{2c-1}$ is the term diverging with the **biggest** order of n.

Because the term with the biggest order only appears in the **denominator**, we have

$$M_{\min} \approx \lim_{n\to\infty} \hat{M}_{\theta_{\min}} = 0 \quad (1 < \alpha \leq 3).$$

Numerical simulation vs. theory

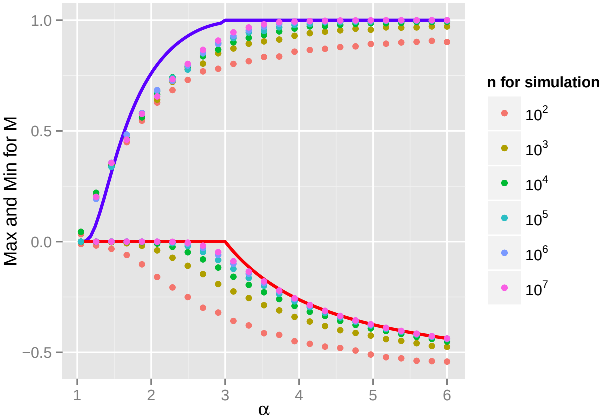


Figure : Theoretical bounds for $M_{\min}$ and $M_{\max}$ and numerical simulations with different series lengths. Each point in simulation is produced by averaging 1000 independent runs.

Empirical results

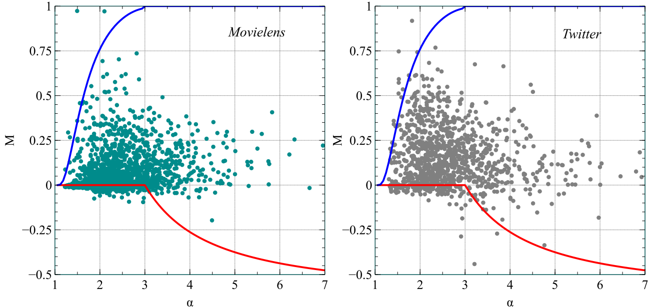


Figure : Memory for power-law distributed **inter-event time series** from empirical inter-event time series, where each series is represented by a point and theoretical bounds are drawn as solid curves. Left: *MovieLens* dataset for online movie rating. Right: *Twitter* dataset for sending tweets. Power-law is examined with KS-test and α is fitted with MLE [3].

---

[3] A. Clauset, C. Shalizi, M. Newman. Power-Law Distributions in Empirical Data. *SIAM Review*, 2009.

Summary

Summary

Contributions

- EPM asymptotics for diverging moments.

Summary

Contributions

- EPM asymptotics for diverging moments.
- Non-trivial bounds for the memory of power-law series (paper in preparation).

Summary

Contributions

- EPM asymptotics for diverging moments.
- Non-trivial bounds for the memory of power-law series (paper in preparation).

Future work

- Proving the *convergence in probability* of the EPM asymptotics, i.e.

$$\lim_{n\to\infty} P\left(\left|\frac{1}{n^{\gamma+1}}\sum_{i=1}^{n} x_i^k - g(n)\right| \geq \epsilon\right) = 0,\ \forall\epsilon > 0.$$

Yingcai Experimental School

Summary

Contributions

- EPM asymptotics for diverging moments.
- Non-trivial bounds for the memory of power-law series (paper in preparation).

Future work

- Proving the *convergence in probability* of the EPM asymptotics, i.e.

$$\lim_{n\to\infty} P\left(\left|\frac{1}{n^{\gamma+1}}\sum_{i=1}^{n} x_i^k - g(n)\right| \geq \epsilon\right) = 0,\ \ \forall\epsilon > 0.$$

Thanks!

Yingcai Experimental School