Extreme values, couplings and graphs

Jesse Goodman
Joint work with M. Eckhoff, R. van der Hofstad and F. Nardi

University of Auckland

5 December 2019
WWPMS 2019

Outline

1. The smallest values in a large sample
   - Poisson point process approximation – Uniform
   - An exact Poisson point process coupling – Exponential
   - An exact Poisson point process coupling – mapping to general random variables
   - The Extreme Value Theorem via this mapping

2. Random graph analogues
   - Poisson-weighted infinite tree
   - Coupling the complete graph to the PWIT
   - Example

Uniform random variables

Let $U_1, \ldots, U_n$ be i.i.d. $\mathrm{Uniform}[0,1]$ random variables. Randomly reorder them into the order statistics

$$U_{(1)} < U_{(2)} < \cdots < U_{(n)}.$$

Then:

- $nU_{(1)} \xrightarrow{d} \mathrm{Exponential}(1)$ as $n \to \infty$, and
- the points $(nU_{(1)}, nU_{(2)}, \ldots)$ can be approximated by a Poisson point process $(X_1, X_2, \ldots)$ of rate 1 on $(0, \infty)$, in the sense that for each fixed k,

$$(nU_{(1)}, \ldots, nU_{(k)}) \xrightarrow{d} (X_1, \ldots, X_k).$$

Uniform random variables

The points $(nU_{(1)}, nU_{(2)}, \dots)$ can be approximated by a Poisson point process $(X_1, X_2, \dots)$ of rate 1 on $(0, \infty)$, in the sense that for each fixed k,

$$(nU_{(1)}, \dots, nU_{(k)}) \xrightarrow{d} (X_1, \dots, X_k).$$

However, this result is not strong enough for all purposes:

Example

Let $0 < \beta < 1$. How big is

$$N = \min\left\{k \geq 2\colon U_{(k)} - U_{(k-1)} < n^{-1-\beta}\right\}?$$

Uniform random variables – stronger version

The points $(nU_{(1)}, nU_{(2)}, \dots)$ can be approximated by a Poisson point process $(X_1, X_2, \dots)$ of rate 1 on $(0, \infty)$, in the sense that for a fixed sequence $k = k_n$ with $k = o(\sqrt{n})$, there is a coupling of $(nU_{(1)}, nU_{(2)}, \dots)$ and $(X_1, X_2, \dots)$ such that

$$\mathbb{P}(nU_{(i)} = X_i \text{ for all } i = 1, \dots, k) = 1 - o(1).$$

Example

Let $0 < \beta < 1$. How big is

$$N = \min\left\{k \geq 2 \colon U_{(k)} - U_{(k-1)} < n^{-1-\beta}\right\}?$$

Heuristic: $nU_{(k)} - nU_{(k-1)} \approx X_k - X_{k-1}$, which are i.i.d. $\mathrm{Exponential}(1)$. Each one has probability $\approx n^{-\beta}$ to be smaller than $n^{-\beta}$, so N should be of order n^{β}.

If $1/2 < \beta < 1$, we cannot apply the coupling stated above.

Uniform random variables – stronger version

The points $(nU_{(1)}, nU_{(2)}, \dots)$ can be approximated by a Poisson point process $(X_1, X_2, \dots)$ of rate 1 on $(0, \infty)$, in the sense that for a fixed sequence $k = k_n$ with $k = o(\sqrt{n})$, there is a coupling of $(nU_{(1)}, nU_{(2)}, \dots)$ and $(X_1, X_2, \dots)$ such that

$$\mathbb{P}(nU_{(i)} = X_i \text{ for all } i = 1, \dots, k) = 1 - o(1).$$

Example

Let $0 < \beta < 1$. How big is

$$N = \min\left\{k \geq 2 \colon U_{(k)} - U_{(k-1)} < n^{-1-\beta}\right\}?$$

Heuristic: $nU_{(k)} - nU_{(k-1)} \approx X_k - X_{k-1}$, which are i.i.d. $\mathrm{Exponential}(1)$. Each one has probability $\approx n^{-\beta}$ to be smaller than $n^{-\beta}$, so N should be of order n^{β}.

If $1/2 < \beta < 1$, we cannot apply the coupling stated above.

Exponential random variables – exact version

Consider now $E_1, \ldots, E_n$ i.i.d. $\mathrm{Exponential}(1)$ random variables. We can produce an exact coupling for the whole sequence of order statistics by marking and thinning the Poisson point process $(X_k)_{k=1}^{\infty}$.

Definition

Associated with X_k, let M_k be a mark chosen i.i.d. and uniformly from $\{1, 2, \ldots, n\}$. Say that $k \in \mathbb{N}$ is thinned if $M_k = M_{k'}$ for some $k' < k$.

In other words, a point from the Poisson point process is thinned if it has the same mark as a smaller point.

Theorem

We can couple $E_1, \ldots, E_n$ and $(X_k)_{k=1}^{\infty}$ such that

$$(E_{(i)})_{i=1}^{n} = (\tfrac{1}{n} X_k)_{k \in \mathbb{N} \text{ is unthinned}},$$

i.e.,

$$\{E_i, i = 1, \ldots, n\} = \{\tfrac{1}{n} X_k \colon k \text{ is unthinned}\}$$

Exponential random variables – exact version

Theorem

We can couple $E_1, \ldots, E_n$ and $(X_k)_{k=1}^{\infty}$ such that

$$(E_{(i)})_{i=1}^{n} = (\tfrac{1}{n} X_k)_{k \in \mathbb{N} \text{ is unthinned}},$$

i.e.,

$$\{E_i, i = 1, \ldots, n\} = \{\tfrac{1}{n} X_k \colon k \text{ is unthinned}\}$$

Proof: The coupling is to define

$$E_i = \frac{1}{n} \min \{X_k \colon M_k = i\}.$$

By the Poisson splitting property, $\{X_k \colon M_k = i\}$ is a Poisson point process of rate $1/n$, independently for each i, so its first point is $\mathrm{Exponential}$(rate $1/n$). Thus the E_i's have the correct distribution. No extra work is needed to handle the order statistics because the X_k's are already naturally ordered.

Exponential random variables – example

Theorem

We can couple $E_1, \dots, E_n$ and $(X_k)_{k=1}^{\infty}$ such that

$$(E_{(i)})_{i=1}^{n} = (\tfrac{1}{n} X_k)_{k \in \mathbb{N} \text{ is unthinned}}$$

Example

Let $0 < \beta < 1$. How big is

$$N = \min \left\{ k \geq 2 \colon E_{(k)} - E_{(k-1)} < n^{-1-\beta} \right\}?$$

Now we can put the previous heuristic into effect:

- $N' = \min \left\{ k \geq 2 \colon X_k - X_{k-1} < n^{-\beta} \right\}$ is of order n^{β}
- $k = N'$ and $k = N' - 1$ are unthinned with high probability because the average number of smaller k' with the same marks will be $O(n^{\beta}/n) = o(1)$
- $N' - N$ will be the number of $k \leq N'$ that are thinned. This number will be $o(N')$ by the same reasoning.

General random variables

For a general random variable Y, we can always write

$$Y = g(E)$$

for $E \sim \mathrm{Exponential}(1)$ and g non-decreasing. Then

Theorem

We can couple $Y_1, \dots, Y_n$ and $(X_k)_{k=1}^{\infty}$ such that

$$(Y_{(i)})_{i=1}^{n} = (g(\tfrac{1}{n} X_k))_{k\in\mathbb{N}} \text{ is unthinned.}$$

We are applying an n-dependent function

$$f_n(x) = g(x/n)$$

to the randomness in $(X_k)_{k=1}^{\infty}$ which does not depend on n. This gives a non-standard way to couple samples for different values of n.

General random variables

For a general random variable Y, we can always write

$$Y = g(E)$$

for $E \sim \mathrm{Exponential}(1)$ and g non-decreasing. Then

Theorem

We can couple $Y_1, \ldots, Y_n$ and $(X_k)_{k=1}^{\infty}$ such that

$$(Y_{(i)})_{i=1}^{n} = (g(\tfrac{1}{n}X_k))_{k\in\mathbb{N}} \text{ is unthinned}.$$

We are applying an n-dependent function

$$f_n(x) = g(x/n)$$

to the randomness in $(X_k)_{k=1}^{\infty}$ which does not depend on n. This gives a non-standard way to couple samples for different values of n.

The Extreme Value Theorem via this mapping

From the above,

$$Y_{(1)} = \min\{Y_1, \ldots, Y_n\} = g(\tfrac{1}{n}X_1).$$

Extreme Value Theorem

If there are affine rescalings such that $a_n Y_{(1)} + b_n$ converges in distribution to a non-trivial limit, then, after a further affine rescaling, the limit distribution is either Gumbel, Weibull or Fréchet.

Extreme Value Theorem in terms of g

If g is such that $a_n g(x/n) + b_n$ converges to a non-constant limit $h(x)$, then, after a further affine rescaling, the limit function is either

$$h(x) = x^{\alpha}, \qquad h(x) = -x^{-\alpha} \quad \text{or} \quad h(x) = \log x.$$

Poisson-weighted infinite tree

The natural analogue to the Poisson point process for graphs is the Poisson-weighted infinite tree.

Definition

The Poisson-weighted infinite tree is the rooted tree such that every vertex v has infinitely many (ordered) children $v1, v2, \dots$, equipped with edge weights X_{vk} (joining vk to its parent v) such that

$$(X_{v1}, X_{v2}, \dots) \text{ forms a Poisson point process of rate 1 on } (0, \infty),$$

independently for each v.

Coupling the complete graph to the PWIT

Theorem, loosely paraphrased from EGvdHN

An exploration process on the complete graph with n vertices and edge weights $Y_e = g(E_e)$ can be coupled to an exploration process on the PWIT with edge weights $Y_v = g(\frac{1}{n}X_v)$.

This theorem is based on an explicit construction involving marks M_v and thinning for vertices v of the PWIT.

Under moderate additional assumptions, the coupling has the property that the complete graph exploration process is the image of the unthinned part of the PWIT exploration process under the mapping $v \mapsto M_v$.

Example – first passage percolation

Theorem (EGvdHN)

First passage percolation on the complete graph with edge weights $Y_e = g(E_e)$ is the image of a unthinned part of a continuous-time branching process in which each individual v has children at ages $g(\frac{1}{n}X_{v1}), g(\frac{1}{n}X_{v2}), \ldots$.

Special case:

$$g(x) = \begin{cases} 1 & \text{if } x \leq c, \\ \infty & \text{if } x > c, \end{cases}$$

gives the Erdős-Rényi random graph $G(n,p)$ with $p = 1 - e^{-c/n}$.

Example – first passage percolation

Theorem (EGvdHN)

First passage percolation on the complete graph with edge weights $Y_e = g(E_e)$ is the image of a unthinned part of a continuous-time branching process in which each individual v has children at ages $g(\frac{1}{n}X_{v1}), g(\frac{1}{n}X_{v2}), \dots$.

Special case:

$$g(x) = \begin{cases} 1 & \text{if } x \leq c, \\ \infty & \text{if } x > c, \end{cases}$$

gives the Erdős-Rényi random graph $G(n, p)$ with $p = 1 - e^{-c/n}$.

Thank you.