Risk and Environmental Sustainability

Linda Mhalla

3 Poisson Process **87**
Improved inferences 88
Eskdalemuir rainfall 89

3.1 Point Processes **93**

3.2 Poisson Processes **99**

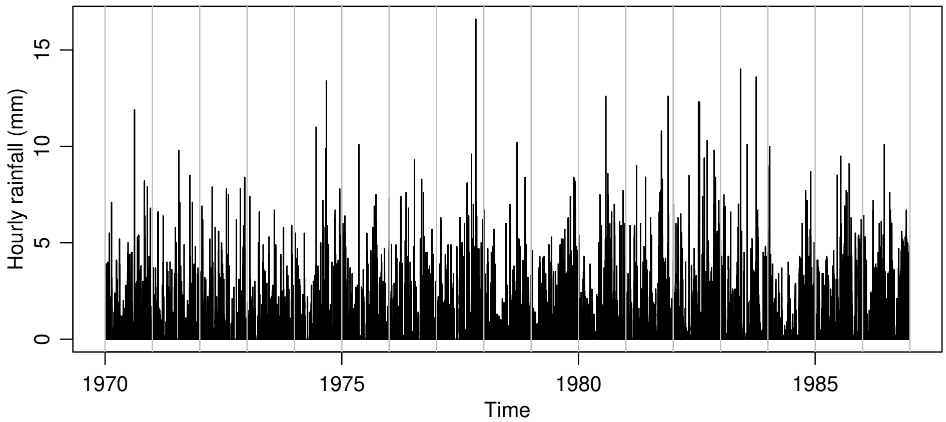
3.3 Poisson Modelling of Extremes **106**

Improved inferences

- The block maximum method can be inefficient if other data are available. Alternative methods include:
  - peaks over thresholds,
  - r-largest order statistics.
- Both are special cases of a **point process** representation, under which we use a Poisson process to approximate the occurrence of those values that exceed a (high) threshold.
- The Poisson process representation is also very powerful in more general settings, so we discuss it before returning to data analysis.

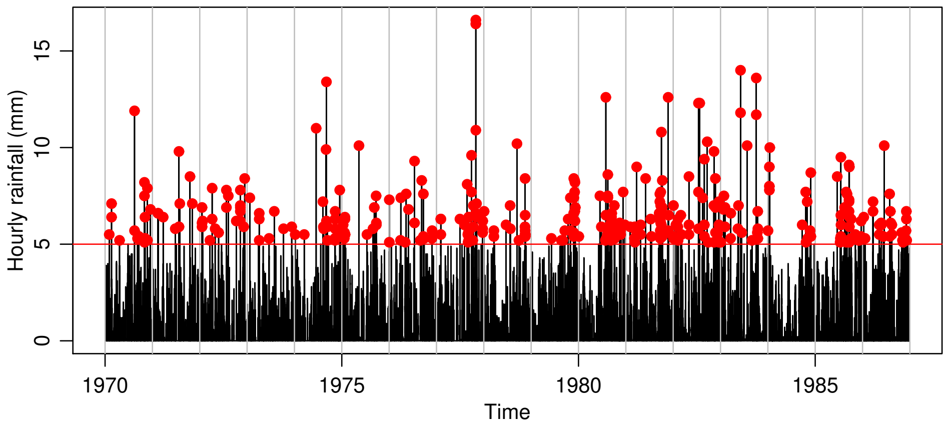
slide 88

Example: Hourly rainfall at Eskdalemuir, 1970–86



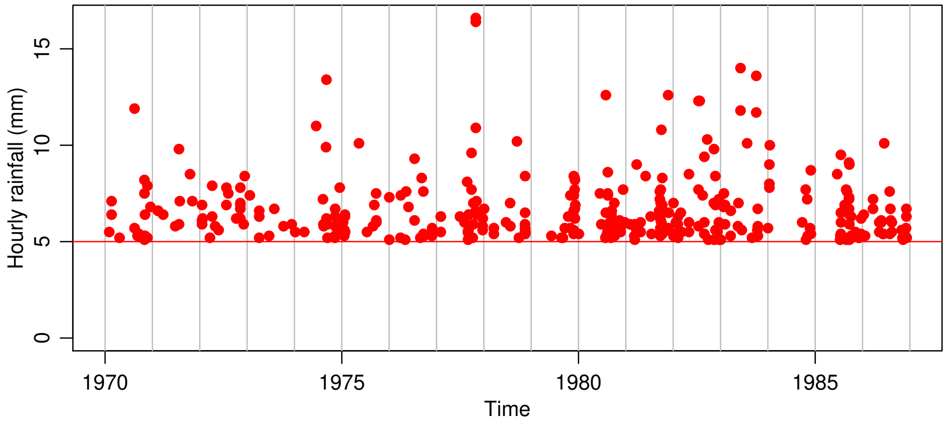
slide 89

Example: Eskdalemuir rainfall



slide 90

Example: Eskdalemuir rainfall



slide 91

Comments

- ☐ The fixed number of annual maxima has been replaced by a random number of exceedances over the threshold.
- ☐ We now retain more observations in the tail of the distribution.
- ☐ Dependence in the underlying series means that exceedances occur in clusters, which we may need to model.
- ☐ For now we suppose that the underlying series comprises independent identically distributed observations, whose maxima have a non-degenerate limit, after renormalisation.

slide 92

3.1 Point Processes

slide 93

Point process

- ☐ A **point process** is a stochastic model for a **point pattern** $\mathcal{P} = \{x_1, x_2, \ldots\}$ lying in a **state space** $\mathcal{E}$. We also call a point an **event**.
- ☐ We visualise $\mathcal{E} \subset \mathbb{R}^2$, but $\mathcal{E}$ might be more complex, e.g., $\mathcal{E} = \mathbb{R} \times \mathcal{C}$, where $\mathcal{C}$ is a space of functions—then a 'point' would be $x = (u, f) \in \mathcal{E}$, with $u \in \mathbb{R}$ and $f \in \mathcal{C}$.
- ☐ The set $\mathcal{E}$ must allow us to count how many points of $\mathcal{P}$ lie in any suitable subset $\mathcal{A} \subset \mathcal{E}$, giving

$$N(\mathcal{A}) = |\mathcal{P} \cap \mathcal{A}| = \sum_x I(x \in \mathcal{P} \cap \mathcal{A}), \quad \mathcal{A} \subset \mathcal{E},$$

  where $I(\cdot)$ is an indicator function.
- ☐ Two points cannot exactly coincide: $\mathcal{P}$ must be **simple** (or **orderly**) — otherwise we would not know how many points there are.
- ☐ If you know about measures . . . the function $N(\mathcal{A})$ is
  - a **counting measure** on $\mathcal{E}$, since it counts the number of elements of $\mathcal{P}$ in any (measurable) set $\mathcal{A}$,
  - a **Radon measure** if $N(\mathcal{A}) < \infty$ for any $\mathcal{A}$ compact (in a suitable topology on $\mathcal{E}$),
  - a **random measure** if the points $\mathcal{P}$ arise at random, since then $N(\mathcal{A})$ is a random variable computed from the (random) $\mathcal{P}$.

Laplace transform

- If it exists, the **Laplace transform** of a scalar random variable X is defined as

$$\mathrm{E}\{\exp(-tX)\} = M_X(-t),$$

  where M_X is the **moment-generating function (MGF)**. This is useful because
  - there is a bijection between distributions and MGFs, i.e., if we recognise M_X, then we know the corresponding distribution;
  - the **continuity theorem** tells us that if $\{X_n\}$, X have CDFs $\{F_n\}$, F for which the MGFs $M_n(t)$, $M(t)$ exist and there exists $a > 0$ such that

$$\lim_{n\to\infty} M_n(t) = M(t), \quad 0 \le |t| < a,$$

    then $X_n \xrightarrow{D} X$, i.e., X_n converges in distribution (weakly, in law) to X.
  - Hence for large enough n we can approximate the distribution of X_n by that of X.
- On the next slide we will extend this to point processes, but first, a simple example:

Theorem 7 (Law of small numbers) *If $X_n \sim B(n, p_n)$ and $np_n \to \lambda > 0$ when $n \to \infty$, then the limiting distribution of X_n is $\mathrm{Pois}(\lambda)$, i.e., $X_n \xrightarrow{D} X$, where $X \sim \mathrm{Pois}(\lambda)$.*

Note to Theorem 7

- The MGF of X is

$$M(t) = \mathrm{E}(e^{tX}) = \sum_{x=0}^{\infty} e^{tx}\lambda^x e^{-\lambda}/x! = e^{-\lambda}\sum_{x=0}^{\infty}(\lambda e^t)^x/x! = \exp\left\{\lambda(e^t - 1)\right\}, \quad t \in \mathbb{R}.$$

- The MGF of X_n is

$$M_n(t) = \mathrm{E}(e^{tX_n}) = \sum_{x=0}^{n} e^{tx}\binom{n}{x} p_n^x (1-p_n)^{n-x} = \left(1 - p_n + p_n e^t\right)^n, \quad t \in \mathbb{R}.$$

  Let $p_n = \lambda_n/n$, where $\lambda_n \to \lambda$, and note that as $n \to \infty$ and for any real t,

$$\left(1 - p_n + p_n e^t\right)^n = \left(1 + \frac{\lambda_n(e^t - 1)}{n}\right)^n \to \exp\left\{\lambda(e^t - 1)\right\}.$$

- As $M_n(t) \to M(t)$ for all real t, the continuity theorem implies that $X_n \xrightarrow{D} X$.

Laplace functional

- ☐ We specify properties of $\mathcal{P}$ through the finite-dimensional distributions of $N(\cdot)$, i.e.,

$$\mathrm{P}\{N(\mathcal{A}_1) = n_1, \ldots, N(\mathcal{A}_k) = n_k\}, \quad n_1, \ldots, n_k \in \{0, 1, 2, \ldots, \},$$

  for all possible choices of sets $\mathcal{A}_1, \ldots, \mathcal{A}_k$, and all $k = 0, 1, 2, \ldots$.
- ☐ An efficient way to do this is through the **Laplace functional**,

$$\mathcal{L}_{\mathcal{P}}(f) = \mathrm{E}\left\{\exp\left(-\int f \mathrm{d}\mathcal{P}\right)\right\}, \quad \text{where} \quad \int f \mathrm{d}\mathcal{P} = \int f(x)\mathcal{P}(\mathrm{d}x) = \sum_{x \in \mathcal{P}} f(x),$$

  for functions $f \geq 0$ that are positive only on a bounded set. If $f(x) = \sum_r t_r I(x \in \mathcal{A}_r)$, then $\mathcal{L}_{\mathcal{P}}(f)$ is the joint MGF for the $N(\mathcal{A}_r)$.
- ☐ Under mild conditions, there is
  - a bijection between point processes and Laplace functionals; and
  - the continuity theorem can be generalised.

slide 96

Convergence of point processes

Definition 8 *A sequence of random variables $\{X_n\}$ with corresponding CDFs $\{F_n\}$* **converges weakly (or in distribution)** *to a random variable X with CDF F if*

$$\lim_{n \to \infty} F_n(x) = F(x) \text{ at every } x \text{ where } F \text{ is continuous.}$$

Definition 9 *A sequence of point processes $\{\mathcal{P}_n\}$ with corresponding counts $\{N_n(\cdot)\}$ on $\mathcal{E}$* **converges weakly (or in distribution)** *to a point process $\mathcal{P}$ with count $N(\cdot)$, written $\mathcal{P}_n \xrightarrow{D} \mathcal{P}$, if for all choices of k and all compact sets $\mathcal{A}_1, \ldots, \mathcal{A}_k \subset \mathcal{E}$ such that*

$$\mathrm{P}\{N(\partial\mathcal{A}_j) = 0\} = 1, \quad j = 1, \ldots, k,$$

where $\partial\mathcal{A}_j$ is the boundary of $\mathcal{A}_j$,

$$\{N_n(\mathcal{A}_1), \ldots, N_n(\mathcal{A}_k)\} \xrightarrow{D} \{N(\mathcal{A}_1), \ldots, N(\mathcal{A}_k)\}, \quad n \to \infty.$$

Theorem 10 (No proof) *The point processes $\mathcal{P}_1, \mathcal{P}_2, \ldots$ converge weakly to the point process $\mathcal{P}$ on $\mathcal{E}$ if and only if the corresponding Laplace functionals converge for every continuous non-negative function f on $\mathcal{E}$ with compact support, i.e., as $n \to \infty$,*

$$\mathcal{L}_{\mathcal{P}_n}(f) = \mathrm{E}\left\{\exp\left(-\int f \mathrm{d}\mathcal{P}_n\right)\right\} \to \mathcal{L}_{\mathcal{P}}(f) = \mathrm{E}\left\{\exp\left(-\int f \mathrm{d}\mathcal{P}\right)\right\}.$$

slide 97

Kallenberg's theorem

- **Kallenberg's theorem** gives another way to establish the weak convergence of $\{\mathcal{P}_n\}$ to a simple process $\mathcal{P}$ when $\mathcal{E} \subset \mathbb{R}$.
- For any $\mathcal{A} \subset \mathcal{E}$, let $N_n(\mathcal{A}) = |\mathcal{P}_n \cap \mathcal{A}|$. Then if
  - $\mathcal{B} \subset \mathcal{E}$ is any interval,
  - $\mathcal{C}$ is any finite union of disjoint sub-intervals of $\mathcal{E}$,

  and if

$$\mathrm{E}\{N_n(\mathcal{B})\} \to \mathrm{E}\{N(\mathcal{B})\}, \quad \mathrm{P}\{N_n(\mathcal{C}) = 0\} \to \mathrm{P}\{N(\mathcal{C}) = 0\}, \quad n \to \infty, \tag{8}$$

  then $\mathcal{P}_n$ converges weakly to $\mathcal{P}$.
- When $\mathcal{E} \subset \mathbb{R}^D$, the same result holds if intervals are replaced by **rectangles**,

$$(a, b] = \{x = (x_1, \ldots, x_D) : a_d < x_d \leq b_d, d = 1, \ldots, D\} \subset \mathcal{E},$$

  where $a_d < b_d$ for each d.
- Thus weak convergence of point processes to a simple limiting process in $\mathbb{R}^D$ entails establishing convergence of expected counts for rectangles and of the **void probabilities** of finite unions of rectangles.
- See Kingman (1993) *Poisson Processes* and Daley and Vere-Jones (2002, 2008), *An Introduction to the Theory of Point Processes*.

slide 98

3.2 Poisson Processes

slide 99

Poisson process

Definition 11 *A* **Poisson process** *is a random countable subset $\mathcal{P}$ of a state space $\mathcal{E}$ such that*

- *the random variables $N(\mathcal{A}_1), \ldots, N(\mathcal{A}_k)$ corresponding to any collection of disjoint subsets $\mathcal{A}_1, \ldots, \mathcal{A}_k$ of $\mathcal{E}$ are independent; and*
- *for any $\mathcal{A} \subset \mathcal{E}$, $N(\mathcal{A})$ has the Poisson distribution with mean $\mu(\mathcal{A})$, where $0 \leq \mu(\mathcal{A}) \leq \infty$, and $\mu(\mathcal{A}) < \infty$ for compact $\mathcal{A}$.*

Comments:

- if $\mathcal{A} = \bigcup_j \mathcal{A}_j$ is a countable union of disjoint sets, then $N(\mathcal{A}) = \sum_j N(\mathcal{A}_j)$, so $\mu(\mathcal{A}) = \sum_j \mu(\mathcal{A}_j)$, and μ is a measure; called the **mean measure** of $\mathcal{P}$;
- μ must be **diffuse**, i.e., $\mu(\{x\}) = 0$ for every $x \in \mathcal{E}$;
- if $\mathcal{E} \subset \mathbb{R}^D$, $\mathcal{A} = [a_1, x_1] \times \cdots \times [a_D, x_D]$, and if

$$\dot{\mu}(x_1, \ldots, x_D) = \frac{\partial^D \mu(\mathcal{A})}{\partial x_1 \cdots \partial x_D}$$

  exists and is finite, then $\dot{\mu}$ is called the **intensity function** of $\mathcal{P}$;
- if $\dot{\mu}(x) \equiv \dot{\mu}$, then $\mathcal{P}$ is called **homogeneous**. Otherwise it is **inhomogeneous**.
- We simplify notation by replacing $\mu\{(a, b]\}$ by $\mu(a, b]$, etc.

slide 100

Conditioning

Theorem 12 (Conditioning) *Let $\mathcal{P}$ be a Poisson process with mean measure μ, and suppose that $\mathcal{A} \subset \mathcal{E}$ is such that $0 < \mu(\mathcal{A}) < \infty$. Conditional on the event $N(\mathcal{A}) = n$, the n points of $\mathcal{P} \cap \mathcal{A}$ have the same distribution as n points generated independently at random in $\mathcal{A}$ with measure $\mu_{\mathcal{A}}(\mathcal{B}) = \mu(\mathcal{B})/\mu(\mathcal{A})$, for $\mathcal{B} \subset \mathcal{A}$.*

- ☐ If μ has intensity $\dot{\mu}(x)$, then we can generate points of $\mathcal{P}$ in $\mathcal{A}$ by
  - generating a value n of $N(\mathcal{A}) \sim \text{Poiss}\{\mu(\mathcal{A})\}$;
  - then generating $X_1, \ldots, X_n \overset{\text{iid}}{\sim} \dot{\mu}(x)/\mu(\mathcal{A})$ for $x \in \mathcal{A}$.
- ☐ The process generated at the second step is a **binomial process**.

Lemma 13 *The* **Laplace functional** *of a Poisson process $\mathcal{P}$ on $\mathcal{E}$ with mean measure μ is*

$$\mathcal{L}_{\mathcal{P}}(f) = \exp\left[-\int_{\mathcal{E}} \left\{1 - e^{-f(x)}\right\} \mu(\mathrm{d}x)\right].$$

slide 101

Note to Theorem 12

If we observe a Poisson process with intensity $\dot{\mu}(x)$ on the set $\mathcal{A}$, and there are points at $\{x_1, \ldots, x_n\}$, then the corresponding probability element is

$$\exp\{-\mu(\mathcal{A})\} \times \prod_{j=1}^{n} \dot{\mu}(x_j), \quad \{x_1, \ldots, x_n\} \subset \mathcal{A}.$$

Properties of the Poisson process imply that $N(\mathcal{A})$ has a Poisson distribution with mean $\mu(\mathcal{A})$, so the conditional density of the n points in $\mathcal{A}$, given that $N(\mathcal{A}) = n$, is the ratio

$$\frac{\exp\{-\mu(\mathcal{A})\} \times \prod_{j=1}^{n} \dot{\mu}(x_j)}{\mu(\mathcal{A})^n \exp\{-\mu(\mathcal{A})\}/n!} = n! \prod_{j=1}^{n} \left\{\frac{\dot{\mu}(x_j)}{\mu(\mathcal{A})}\right\}, \quad \{x_1, \ldots, x_n\} \subset \mathcal{A}.$$

Now consider the measure $\mu_{\mathcal{A}}(\mathcal{B}) = \mu(\mathcal{B})/\mu(\mathcal{A})$, for $\mathcal{B} \subset \mathcal{A}$, which is a probability measure on subsets of $\mathcal{A}$, because it is non-negative and $\mu_{\mathcal{A}}(\mathcal{A}) = 1$. The corresponding probability density is $\dot{\mu}(x)/\mu(\mathcal{A})$ ($x \in \mathcal{A}$), so the joint density for independent identically distributed variables $X_1, \ldots, X_n$ with distribution $\mu_{\mathcal{A}}$ is $\prod_{j=1}^{n}\{\dot{\mu}(x_j)/\mu(\mathcal{A})\}$, which is almost the conditional probability above. The additional factor $n!$ arises because the point process is unlabelled: the same density would arise for any of the $n!$ permutations of $X_1, \ldots, X_n$ that gave the outcome $\{x_1, \ldots, x_n\}$.

note 1 of slide 101

Note to Lemma 13

Let $f \geq 0$ have support only on a compact $\mathcal{A}$, so $\mu(\mathcal{A}) < \infty$. Conditional on $N(\mathcal{A}) = n$, $\int f(x)\mathcal{P}(\mathrm{d}x) = \sum_{j=1}^{n} f(X_j)$, where $\{X_1, \ldots, X_n\} \subset \mathcal{A}$ are independent with density $\dot{\mu}(x)/\mu(\mathcal{A})$. Thus

$$\begin{aligned}
\mathrm{E}\left[\exp\left\{-\int f(x)\mathcal{P}(\mathrm{d}x)\right\} \,\middle|\, N(\mathcal{A}) = n\right] &= \mathrm{E}\left[\exp\left\{-\sum_{j=1}^{n} f(X_j)\right\} \,\middle|\, N(\mathcal{A}) = n\right] \\
&= \left\{\int_{\mathcal{A}} e^{-f(x)}\mu(\mathrm{d}x)/\mu(\mathcal{A})\right\}^n .
\end{aligned}$$

Hence

$$\begin{aligned}
\mathrm{E}\left[\exp\left\{-\int f(x)\mathcal{P}(\mathrm{d}x)\right\}\right] &= \sum_{n=0}^{\infty} \left\{\int_{\mathcal{A}} e^{-f(x)}\mu(\mathrm{d}x)/\mu(\mathcal{A})\right\}^n \frac{\mu(\mathcal{A})^n}{n!} e^{-\mu(\mathcal{A})} \\
&= \exp\left[\int_{\mathcal{A}} e^{-f(x)}\mu(\mathrm{d}x) - \mu(\mathcal{A})\right] \\
&= \exp\left[-\int_{\mathcal{A}} \left\{1 - e^{-f(x)}\right\} \mu(\mathrm{d}x)\right] \\
&= \exp\left[-\int_{\mathcal{E}} \left\{1 - e^{-f(x)}\right\} \mu(\mathrm{d}x)\right],
\end{aligned}$$

as required, since $1 - \exp\{-f(x)\} \equiv 0$ outside $\mathcal{A}$.

note 2 of slide 101

Superposition and colouring

Theorem 14 (Superposition) *If $\mathcal{P}_1$, $\mathcal{P}_2$ are independent Poisson processes on $\mathbb{R}^D$ with mean measures μ_1, μ_2, then their union $\mathcal{P}_1 \cup \mathcal{P}_2$ is a Poisson process with mean measure $\mu_1 + \mu_2$.*

Theorem 14 extends to a countable number of Poisson processes.

Theorem 15 (Colouring) *Let $\mathcal{P}$ be a Poisson process with intensity $\dot{\mu}(x)$. Colour a point of $\mathcal{P}$ at x red with probability $\gamma(x)$; otherwise colour it green. Then the red and green sets of points $\mathcal{P}_{\mathrm{red}}$ and $\mathcal{P}_{\mathrm{green}}$ are independent Poisson processes with intensity functions*

$$\dot{\mu}_{\mathrm{red}}(x) = \dot{\mu}(x)\gamma(x), \quad \dot{\mu}_{\mathrm{green}}(x) = \dot{\mu}(x)\{1 - \gamma(x)\}.$$

The colouring theorem is in some sense the inverse of the superposition theorem, and it too applies with a countable number of colours.

slide 102

Note to Theorem 14

☐ This looks easy using the Laplace functional for $\mathcal{P}_1 \cup \mathcal{P}_2$, which is

$$\mathcal{L}_{\mathcal{P}_1\cup\mathcal{P}_2}(f) = \mathrm{E}\left[\exp\left\{-\int f(x)(\mathcal{P}_1 \cup \mathcal{P}_2)(\mathrm{d}x)\right\}\right].$$

Now

$$\int f(x)(\mathcal{P}_1 \cup \mathcal{P}_2)(\mathrm{d}x) = \int f(x)\mathcal{P}_1(\mathrm{d}x) + \int f(x)\mathcal{P}_2(\mathrm{d}x),$$

and the two processes are independent, so

$$\begin{aligned}
\mathrm{E}\left[\exp\left\{-\int f\mathrm{d}(\mathcal{P}_1 \cup \mathcal{P}_2)\right\}\right] &= \mathrm{E}\left\{\exp\left(-\int f\mathrm{d}\mathcal{P}_1\right)\right\} \times \mathrm{E}\left\{\exp\left(-\int f\mathrm{d}\mathcal{P}_2\right)\right\} \\
&= \exp\left\{-\int_{\mathcal{E}}\left(1-e^{-f}\right)\mathrm{d}\mu_1\right\} \times \exp\left\{-\int_{\mathcal{E}}\left(1-e^{-f}\right)\mathrm{d}\mu_2\right\} \\
&= \exp\left\{-\int_{\mathcal{E}}\left(1-e^{-f}\right)\mathrm{d}(\mu_1+\mu_2)\right\},
\end{aligned}$$

which is the Laplace functional of a Poisson process with mean measure $\mu_1 + \mu_2$.

☐ The catch with the argument above is the assumption that points of $\mathcal{P}_1$ and $\mathcal{P}_2$ do not coincide, so that

$$\mathrm{P}(\mathcal{P}_1 \cap \mathcal{P}_2 \cap \mathcal{A} = \emptyset) = 1$$

for any $\mathcal{A}$ for which $\mu_1(\mathcal{A})$, $\mu_2(\mathcal{A})$ are both finite. This is intuitively obvious but takes a bit of measure-theoretic work to prove.

note 1 of slide 102

Mapping

Theorem 16 (Mapping) *Let $\mathcal{P}$ be a Poisson process on $\mathcal{E}$ with mean measure μ, and suppose that the function $g : \mathcal{E} \to \mathcal{E}^*$ maps $\mathcal{E}$ into $\mathcal{E}^*$. Define*

$$\mu^*(\mathcal{A}^*) = \mu\{g^{-1}(\mathcal{A}^*)\}, \quad \mathcal{A}^* \subset \mathcal{E}^*.$$

If

(i) $\mu^*(\{x^*\}) = \mu^*(x^*) = 0$ for every $x^* \in \mathcal{E}^*$, and

(ii) $\mu^*(\mathcal{A}^*) < \infty$ for any compact $\mathcal{A}^*$,

then $\mathcal{P}^* = g(\mathcal{P})$ is a Poisson process on $\mathcal{E}^*$ with mean measure μ^*.

Here

☐ (i) implies that g does not create atoms in $\mathcal{E}^*$,

☐ (ii) implies that no compact set $\mathcal{A}^* \subset \mathcal{E}^*$ has infinite measure,

which are both needed for $\mathcal{P}^*$ to be Poisson.

Example 17 *If $\mathcal{P}$ is a homogeneous Poisson process of unit rate on $(0,\infty)$, and $g(x) = 1/x$, show that $g(\mathcal{P})$ is a Poisson process and find its intensity function. What if $g(x) = \lceil x \rceil$ or $g(x) = |\sin x|$?*

slide 103

Note to Example 17

☐ The mean measure of $\mathcal{P}$ is given by $\mu[a,b] = (b-a)$, for $0 < a < b < \infty$.
(i) The function g maps $(0,\infty)$ to $(0,\infty)$, and $g \equiv g^{-1}$, so $g^{-1}(x^*) = 1/x^*$ satisfies $\mu[1/x^*, 1/x^*] = (1/x^* - 1/x^*) = 0$ for any $0 < x^* < \infty$.
(ii) Any compact set $\mathcal{A}$ of $(0,\infty)$ is a subset of a set $[a,b]$ for some $0 < a < b < \infty$, so

$$\mu^*(\mathcal{A}) = \mu(g^{-1}\mathcal{A}) = \int_{g^{-1}\mathcal{A}} \dot{\mu}(x)\,dx \le \int_{g^{-1}[a,b]} 1\,dx = \int_{1/b}^{1/a} dx = (1/a - 1/b) < \infty.$$

Hence $g(\mathcal{P})$ is indeed a Poisson process, and since $\mu[a,b] = (1/a - 1/b)$, its intensity function is $d\mu[a,b]/db = 1/b^2$, for $b > 0$.
A sketch shows what happens to the intensities of $\mathcal{P}$ and $g(\mathcal{P})$.

☐ With $g(x) = \lceil x \rceil$, where $\lceil x \rceil$ is the smallest integer greater than or equal to x, then condition (i) fails whenever $x^* \in \mathbb{N}$, so the resulting process is not Poisson, as points of $\mathcal{P}^*$ could be superposed on the positive integers and thus $N^*(\cdot)$ is not well-defined. Equivalently,

$$\mu^*(\{n\}) = \mu[g^{-1}(\{n\})] = \mu\{(n-1,n]\} = 1, \quad n \in \mathbb{N},$$

so μ^* has atoms on every positive integer and thus is not diffuse.

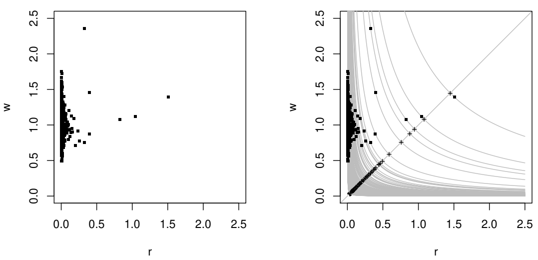
☐ With $g(x) = |\sin(x)|$ we have $\mathcal{E}^* = [0,1]$, and it is easy to check that while condition (i) is satisfied, $\mu^*([a,b]) = \infty$ for any $0 < a < b < 1$, so condition (ii) fails; $\mathcal{P}^*$ has an infinite number of points in any interval.

Example

Example 18 *Let $\mathcal{P}$ be a Poisson process with $\mathcal{E} = \mathbb{R}_+^2$ with $x = (r,w)$ generated by*

$$\mu\{(r,\infty) \times (w,\infty)\} = \frac{1}{r} \times \{1 - F(w)\}, \quad r, w > 0,$$

where F is the CDF of a positive continuous random variable W with unit expectation. Show that $q = rw$ defines a Poisson process and find its intensity.



Left panel: first 1000 points (r,w) of a Poisson process sequentially generated on $\mathbb{R}_+^2$. Right panel: mapping of the points shown in the left panel to $q = rw$, shown as $+$ on the diagonal, with the mapping function shown by the curved grey lines.

Note to Example 18

☐ In the picture $\mathcal{P} = \{(R_i, W_i) : i = 1, 2, \ldots\}$, where $R_1 > R_2 > \cdots > 0$ are generated sequentially by setting $R_i = (E_1 + \cdots + E_i)^{-1}$, with $E_i \stackrel{\text{iid}}{\sim} \exp(1)$, and $W_i = \exp(\sigma\varepsilon_i - \sigma^2/2)$, where $\varepsilon_i \stackrel{\text{iid}}{\sim} N(0, 1)$, independent of the E_i; note that $\mathrm{E}(W_i) = 1$. The first 1000 points of a realisation of such a process are shown in the left-hand panel of the figure; the full realisation would have an infinity of points at the left-hand edge of the panel, because $\mu\{(r, \infty) \times (0, \infty)\} = 1/r \to \infty$ as $r \to 0$.

☐ In the general case the mean measure has an *intensity function* $\dot{\mu}$ given by its derivative at the upper right corner of a rectangle $(r', r) \times (w', w)$, i.e.,

$$\begin{aligned} \dot{\mu}(r, w) &= \frac{\partial^2 \mu\{(r', r) \times (w', w)\}}{\partial r \partial w} \\ &= \frac{\partial^2}{\partial r \partial w} \{\mu(r', w') - \mu(r, w') - \mu(r', w) + \mu(r, w)\} \qquad (9) \\ &= \frac{1}{r^2} \times f(w), \quad r, w > 0, \qquad (10) \end{aligned}$$

where we have written $\mu(r, w) = \mu\{(r, \infty) \times (w, \infty)\}$ and so forth, and f denotes the density function corresponding to F.

☐ Let $g(r, w) = rw$, corresponding to setting $Q_i = R_i W_i$, which amounts to collapsing the points shown in the left-hand panel onto the diagonal line shown in the right-hand panel. For any $q > 0$, $\mu^*(q) = \mu[\{(r, q/r) : r > 0\}] = 0$ because μ has a density with respect to Lebesgue measure and the set $\{(r, q/r) : r > 0\}$ has Lebesgue measure zero, so this transformation does not create atoms. We can check the second property of μ^* once it is calculated. Note that $Q = RW > q$ if and only if $R > q/W$, and that $\mathcal{A}_q = \{(r, w) : rw > q\}$ has measure

$$\begin{aligned} \mu^*(q) = \mu(\mathcal{A}_q) &= \int_0^\infty f(w) \int_{r=q/w}^\infty \frac{1}{r^2} \,\mathrm{d}r \,\mathrm{d}w \\ &= \int_0^\infty f(w) \left[-\frac{1}{r}\right]_{q/w}^\infty \mathrm{d}w \\ &= \int_0^\infty f(w) \frac{1}{q/w} \,\mathrm{d}w \\ &= \frac{1}{q}\mathrm{E}(W) = \frac{1}{q}, \quad q > 0. \qquad (11) \end{aligned}$$

Hence $Q_i = R_i W_i$ is also Poisson, with the same mean measure as the R_i. This implies that the second property is also satisfied: any compact set $\mathcal{A}^*$ is a subset of (q_1, q_2) for some $q_2 > q_1$, so

$$\mu^*(\mathcal{A}^*) \le \mu^*(q_1, \infty) = \mu^*(q_2, \infty) = q_1^{-1} - q_2^{-1} < \infty.$$

☐ The restriction of $\mathcal{P}$ to a subset $\mathcal{E}'$ of $\mathcal{E}$ clearly also follows a Poisson process, with mean measure $\mu'(\mathcal{A}) = \mu(\mathcal{E}' \cap \mathcal{A})$. For example, if we let $\mathcal{E} = (0, \infty)$, consider $R_1, R_2, \ldots$ and let $\mathcal{E}' = (z', \infty)$ for some $z' > 0$, then we retain only those points R_i exceeding z'. As $\mu(\mathcal{E}') = 1/z'$ is finite, these R_i can be generated by first simulating a Poisson variable N' with mean $1/z'$, and if $N' = n$, simulating n independent variables on the interval (z', ∞) with survivor function z'/z; these Pareto variables have probability density function z'/z^2 $(z > z')$.

note 1 of slide 104

Marking

Theorem 19 (Marking) *Let $\mathcal{P}$ be a Poisson process on $\mathcal{E}$ with mean measure μ. Attach a random variable y_x, called the mark, to each point x of $\mathcal{P}$; the distribution of $y_x \in \mathcal{Y}$ may depend on x but not on any other point of $\mathcal{P}$. Then the points (x, y_x) form a Poisson process $\mathcal{P}^*$ in the product space $\mathcal{E} \times \mathcal{Y}$ with mean measure*

$$\mu(\mathcal{C}) = \iint_{(x,y)\in\mathcal{C}} \nu_x(\mathrm{d}y)\mu(\mathrm{d}x), \quad \mathcal{C} \subset \mathcal{E} \times \mathcal{Y},$$

where $\nu_x(\cdot)$ is the conditional probability measure of y_x given x.

- ☐ This provides an approach to making new Poisson processes, by attaching random variables to existing processes, and (perhaps) then applying the mapping theorem.
- ☐ If y_x takes a countable number of values ($\equiv$ colours), then the colouring theorem shows that the corresponding subsets of $\mathcal{P}$ are independent Poisson processes.

slide 105

Note to Theorem 19

The Laplace functional of $\mathcal{P}^*$ is

$$\mathrm{E}\left\{\exp\left(-\int f \mathrm{d}\mathcal{P}^*\right)\right\} = \mathrm{E}_{\mathcal{P}}\left[\mathrm{E}\left\{\exp\left(-\int f \mathrm{d}\mathcal{P}^*\right) \,\middle|\, \mathcal{P}\right\}\right]$$

and the inner expectation on the right-hand side is

$$\prod_{x\in\mathcal{P}} \int_{\mathcal{Y}} e^{-f(x,y)}\, \nu_x(\mathrm{d}y) = \exp\left(-\int f^*\, \mathrm{d}\mathcal{P}\right),$$

say, where

$$f^*(x) = -\log \int_{\mathcal{Y}} e^{-f(x,y)}\, \nu_x(\mathrm{d}y).$$

Thus the Laplace functional of $\mathcal{P}^*$ is that of the Poisson process $\mathcal{P}$ with f replaced by f^*. But since

$$\begin{aligned}
\int_{\mathcal{E}} \left\{1 - e^{-f^*(x)}\right\} \mu(\mathrm{d}x) &= \int_{\mathcal{E}} \left\{1 - \int_{\mathcal{Y}} e^{-f(x,y)}\, \nu_x(\mathrm{d}y)\right\} \mu(\mathrm{d}x) \\
&= \int_{\mathcal{E}} \int_{\mathcal{Y}} \left\{1 - e^{-f(x,y)}\right\} \nu_x(\mathrm{d}y)\mu(\mathrm{d}x),
\end{aligned}$$

the result is established.

note 1 of slide 105

3.3 Poisson Modelling of Extremes

slide 106

Basic result

Theorem 20 *Let* $X_1, \ldots, X_{nt_0} \overset{\text{iid}}{\sim} F$ *form* t_0 *blocks each of* n *observations, and suppose that sequences* $\{a_n\} > 0$ *and* $\{b_n\}$ *exist such that*

$$\mathrm{P}\left[\{\max(X_1, \ldots, X_n) - b_n\}/a_n \le x\right] \to G(x), \quad n \to \infty,$$

where G *is non-degenerate. Then as* $n \to \infty$ *the point processes*

$$\mathcal{P}_n = \{(j/(n+1), (X_j - b_n)/a_n) : j = 1, \ldots, nt_0\}$$

on $\mathcal{E} = [0, t_0] \times \mathcal{E}_x$ *converge in distribution to a Poisson process* $\mathcal{P}$ *with mean measure*

$$\mu\{(t', t) \times [x, \infty)\} = (t - t')\Lambda(x), \quad 0 \le t' < t \le t_0, \quad x \in \mathcal{E}_x = \{x' \in \mathbb{R} : \Lambda(x') < \infty\}, \qquad (12)$$

where

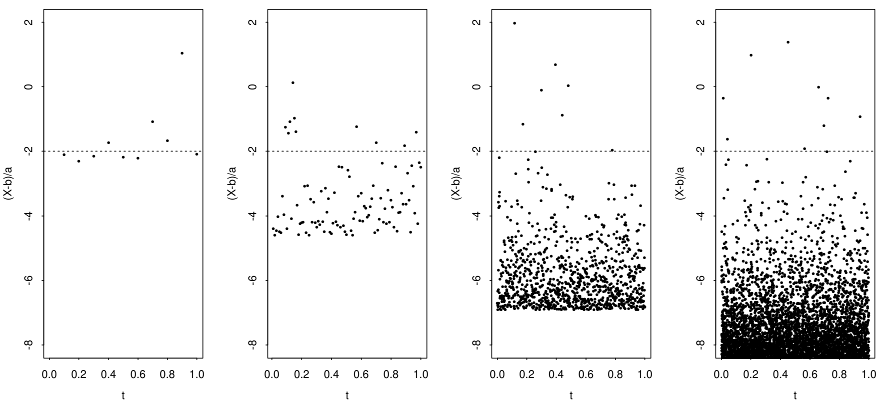
$$\Lambda(x) = \left(1 + \xi \frac{x - \eta}{\tau}\right)_+^{-1/\xi}$$

depends on parameters $\eta, \xi \in \mathbb{R}$ *and* $\tau > 0$ *and* $a_+ = \max(a, 0)$ *for real* a*. The corresponding intensity function is*

$$-\dot{\Lambda}(x) = \tau^{-1}\left(1 + \xi \frac{x - \eta}{\tau}\right)_+^{-1/\xi - 1} \ge 0.$$

slide 107

Point process limit



□ Here $\mathcal{E} \subset \mathbb{R}^D$, so we only need Kallenberg's theorem: for $\mathcal{A} \subset \mathcal{E}$, let $N_n(\mathcal{A}) = |\mathcal{P}_n \cap \mathcal{A}|$. Then if $\mathcal{B} \subset \mathcal{E}$ is any rectangle, and $\mathcal{C}$ is any finite union of disjoint rectangles of $\mathcal{E}$, and if

$$\mathrm{E}\{N_n(\mathcal{B})\} \to \mathrm{E}\{N(\mathcal{B})\}, \quad \mathrm{P}\{N_n(\mathcal{C}) = 0\} \to \mathrm{P}\{N(\mathcal{C}) = 0\}, \quad n \to \infty, \qquad (13)$$

then $\mathcal{P}_n \xrightarrow{D} \mathcal{P}$ as $n \to \infty$.

slide 108

Forms of $\Lambda(x)$

- $\Lambda(x)$ is decreasing, but has three distinct forms:
  - when $\xi > 0$,

$$\Lambda(x) = \begin{cases} +\infty, & x \le \eta - \tau/\xi, \\ \left(1 + \xi\frac{x-\eta}{\tau}\right)_+^{-1/\xi}, & x > \eta - \tau/\xi, \end{cases}$$

    which is finite only for $x > \eta - \tau/\xi$, so $\Lambda(\mathcal{A}) = +\infty$, giving infinite counts, for any set $\mathcal{A}$ that goes below $\eta - \tau/\xi$;
  - for $\xi = 0$ we take the limit when $\xi \to 0$, giving

$$\Lambda(x) = \exp\{-(x-\eta)/\tau\}, \quad x \in \mathbb{R},$$

    which is finite for all x;
  - when $\xi < 0$,

$$\Lambda(x) = \begin{cases} \left(1 + \xi\frac{x-\eta}{\tau}\right)_+^{-1/\xi}, & x < \eta - \tau/\xi, \\ 0, & x \ge \eta - \tau/\xi, \end{cases}$$

    which is finite for all x.
- When $\xi \le 0$ the limiting mass at $-\infty$ is infinite, so any compact set $\mathcal{A}$ considered must have a finite lower bound.

slide 109

Implications: Maxima

- A rescaled block maximum $Y_n = \{\max(X_1, \ldots, X_n) - b_n\}/a_n$ satisfies

$$\begin{aligned} \mathrm{P}(Y_n \le y) &= \mathrm{P}[N_n\{(0,1) \times [y,\infty)\} = 0] \\ &\to \mathrm{P}[N\{(0,1) \times [y,\infty)\} = 0] \quad n \to \infty, \\ &= \exp[-\mu\{(0,1) \times [y,\infty)\}], \\ &= \exp\{-\Lambda(y)\}, \quad y \in \mathbb{R}, \end{aligned}$$

  so a block maximum has a limiting **generalized extreme-value (GEV)** distribution,

$$G(y) = \exp\left\{-\left(1 + \xi\frac{y-\eta}{\tau}\right)_+^{-1/\xi}\right\}, \quad y \in \mathbb{R}.$$

slide 110

Implications: Threshold exceedances I

☐ Consider the 'forgetting' mapping $g : \mathbb{R}^2 \to \mathbb{R}$ with $g(t, x) = x$, giving the process of event sizes $\mathcal{P}^* = g(P)$ without their times. The mapping theorem (Theorem 16) implies that $\mathcal{P}^*$ is Poisson with mean measure

$$\mu^*\{[x,\infty)\} = \mu[g^{-1}\{[x,\infty)\}] = \mu\{[0,t_0] \times [x,\infty)\} = t_0\{1 + \xi(x-\eta)/\tau\}_+^{-1/\xi}.$$

☐ The conditional property (Theorem 12) implies that conditional on $N^*([u,\infty)) = n$, these n threshold exceedances have the same distribution as n points generated independently on $\mathcal{A}_u$ with measure

$$\frac{\mu^*(\mathcal{A}_{u+x})}{\mu^*(\mathcal{A}_u)} = \frac{t_0\{1+\xi(x-u-\eta)/\tau\}_+^{-1/\xi}}{t_0\{1+\xi(u-\eta)/\tau\}_+^{-1/\xi}} = \left(1 + \xi\frac{x}{\sigma_u}\right)_+^{-1/\xi}, \quad x > 0,$$

where $\sigma_u = \tau + \xi(u - \eta)$. This corresponds to the **generalized Pareto distribution (GPD)**.

☐ The mapping $g_1(t, x) = t$ giving the rescaled times of those events that exceed u clearly satisfies the conditions of the mapping theorem and has measure

$$\mu_1^*\{(s,t]\} = \mu[g_1^{-1}\{(s,t]\}] = \mu\{(s,t] \times (u,\infty]\} = (t-s)\Lambda(u),$$

and intensity function $\Lambda(u)$. Thus the rescaled times of those events whose sizes exceed u form a homogeneous Poisson process on $(0, t_0)$ with rate $\Lambda(u)$.

slide 111

Implications: Threshold exceedances II

This yields two fitting approaches:

☐ estimate η, τ and ξ directly by fitting the Poisson process likelihood

$$\exp\{-\mu(\mathcal{A}_u)\} \times \prod_{j=1}^{n} \dot\mu(t_j, x_j),$$

for the region $\mathcal{A}_u = [0, t_0] \times [u, \infty)$ containing $\{x_1, \ldots, x_n\}$ and where the intensity function for threshold exceedances is

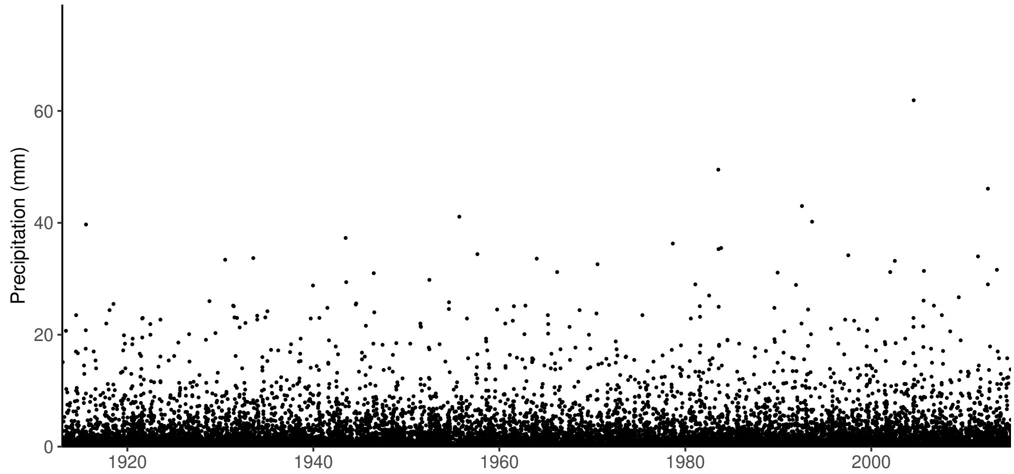
$$\begin{aligned}\dot\mu(t,x) &= \frac{\partial^2 \mu\{[s,t] \times [u,x]\}}{\partial t \partial x} \\ &= \frac{\partial^2}{\partial t \partial x}\left((t-s)\left[\{1+\xi(u-\eta)/\tau\}_+^{-1/\xi} - \{1+\xi(x-\eta)/\tau\}_+^{-1/\xi}\right]\right) \\ &= \frac{1}{\tau}\{1+\xi(x-\eta)/\tau\}_+^{-1/\xi-1}.\end{aligned}$$

☐ estimate σ_u and ξ from the exceedances and p_u from the number of exceedances, n_u.

slide 112

Abisko daily rainfall data

- Daily precipitation in Abisko, in northern Sweden, 1913–2014. The largest value is 61.9 mm, but many values are zero and most of the positive values are quite small.



slide 113

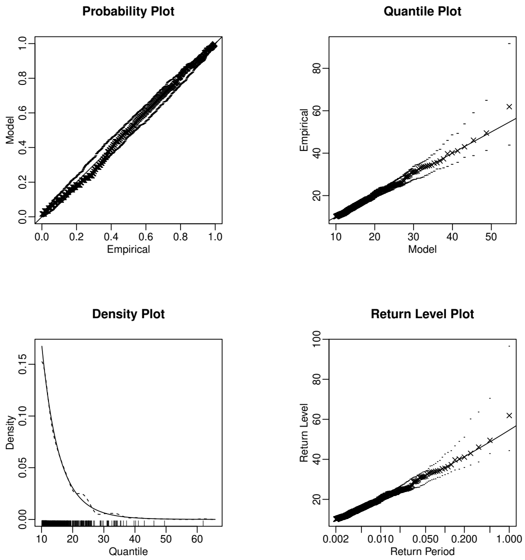
Abisko Poisson process fit

```
(fit.pp <- fpot(abisko$precip, threshold=10, model="pp", npp=365.25,
   start=list(loc=20,scale=6.5,shape=0.01)))
# needs initial values and number of points/block

Deviance: 2241.606

Threshold: 10
Number Above: 499
Proportion Above: 0.0134

Estimates
     loc     scale     shape
19.79658   6.52110   0.07026

Standard Errors
    loc    scale    shape
0.55597  0.37895  0.05088

Optimization Information
  Convergence: successful
  Function Evaluations: 20  ...  Gradient Evaluations: 8
```

slide 114

Abisko Poisson process fit

☐ Let's check the fit using `plot(fit.pp)`:



slide 115

Note 1 to Theorem 20

☐ If a limiting distribution G for rescaled maxima exists, then

$$\begin{aligned} \mathrm{P}[\{\max(X_1,\ldots,X_n)-b_n\}/a_n \le y] &= \mathrm{P}\{\max(X_1,\ldots,X_n) \le b_n + a_n y\} \\ &= F^n(b_n + a_n y) \\ &= \left[1 - \frac{n\{1-F(b_n+a_n y)\}}{n}\right]^n. \end{aligned}$$

Hence a limiting function $\Lambda(y)$ must exist such that

$$\Lambda_n(y) = n\{1 - F(b_n + a_n y)\} \to \Lambda(y), \quad n \to \infty.$$

☐ Let $H(x) = -\log\{1 - F(x)\}$ denote the *cumulative hazard function* corresponding to F, and choose $b_n = b_n^*$ such that $H(b_n^*) = -\log n$, so that

$$\log \Lambda_n(y) = H(b_n + a_n y) - H(b_n).$$

☐ We suppose that F is continuous, places probability in an interval $[x_*, x^*]$, where either or both limits might be infinite, F is not defective (so there is no mass at x^*), that H is twice continuously differentiable with reciprocal hazard function $r(x) = 1/H'(x)$, and that $\lim_{x \to x^*} r'(x) = \xi$ is real and finite. These are sometimes called the *von Mises conditions*.

☐ Then

$$H(b_n + a_n y) - H(b_n) = a_n \int_0^y \frac{1}{r(b_n + a_n x)}\, \mathrm{d}x = a_n \int_0^y \frac{1}{r(b_n) + a_n x r'\{b_n + s_n(x)\}}\, \mathrm{d}x,$$

where $s_n(x)$ lies between zero and x. If we now choose $a_n = a_n^* = r(b_n^*)$, which is positive because $r(x) = \{1 - F(x)\}/f(x)$, we have

$$H(b_n^* + a_n^* y) - H(b_n^*) = \int_0^y \frac{1}{1 + x r'\{b_n^* + s_n(x)\}}\, \mathrm{d}x = \int_0^y \frac{1}{1 + \xi_n x} g_n(x)\, \mathrm{d}x,$$

where $\xi_n = r'(b_n^*)$ and $g_n(x) = (1 + \xi_n x)/\{1 + x r'\{b_n^* + s_n(x)\}$.

☐ The implicit function theorem implies that $s_n(x)$ is continuous in x and so is r', so $g_n(x)$ is continuous in x, and one can check that $g_n(x) \to 1$ as $n \to \infty$. Hence in the interval where $1 + \xi_n x$ does not change sign, we can use a mean value theorem for integrals and choose y^* such that

$$H(b_n^* + a_n^* y) - H(b_n^*) = g_n(y^*) \int_0^y \frac{1}{1 + \xi_n x}\, \mathrm{d}x = g_n(y^*) \times \xi_n^{-1} \log(1 + \xi_n x)_+,$$

where we add the $(\cdot)_+$ to remind us that the term in brackets must be positive. Now $\xi_n = r'(b_n^*) \to \xi$ and $0 < y^* < y$ as $n \to \infty$, so

$$\lim_{n \to \infty} H(b_n^* + a_n^* y) - H(b_n^*) = \xi^{-1} \log(1 + \xi y)_+ = \log \Lambda(y),$$

as required. This establishes sufficient conditions under which a maximum has limiting distribution $\exp\{-\Lambda(y)\}$, with $\eta = 0$ and $\tau = 1$. We need the more general case to allow for the fact that b_n and a_n are unknown in applications (because F is unknown).

note 1 of slide 115

Note 2 to Theorem 20

- ☐ To establish the Poisson convergence, define the binomial processes

$$\mathcal{P}_n = \{(j/(n+1), (X_j - b_n)/a_n) : j = 1, \ldots, nt_0\}, \quad n = 1, 2, \ldots,$$

and the corresponding count process $N_n(\cdot)$ on $\mathcal{E} = [0, t_0] \times \mathcal{E}_x$.

- ☐ Let $0 < t_1 < t_2 \leq t_0$ and $x_1 < x_2$ determine the rectangle $\mathcal{A} = (t_1, t_2] \times (x_1, x_2]$, let

$$\mu(\mathcal{A}) = (t_2 - t_1)\{\Lambda(x_1) - \Lambda(x_2)\}, \quad \mathcal{A} \subset \mathcal{E},$$

and let $\mathcal{P}$ denote a Poisson process on $\mathcal{E}$ with mean measure μ.

- ☐ We now check Kallenberg's conditions. If $\lfloor x \rfloor$ is the integer part of x, then

$$\begin{aligned}
\mathrm{E}\{N_n(\mathcal{A})\} &= \lfloor (n+1)t_2 - (n+1)t_1 \rfloor \times \mathrm{P}\{x_1 < (X_j - b_n)/a_n \leq x_2\} \\
&= \frac{\lfloor (n+1)(t_2 - t_1) \rfloor}{n} \times \Lambda_n(x_1, x_2] \\
&\to (t_2 - t_1)\Lambda(x_1, x_2] = \mu(\mathcal{A}), \quad n \to \infty,
\end{aligned}$$

which verifies the first condition.

- ☐ For the second condition, let $\mathcal{C}$ be a union of a finite number of disjoint rectangles of $\mathcal{E}$, and note that we can write $\mathcal{C} = \bigcup_{i=1}^{k} \mathcal{T}_i \times \bigcup_{l=1}^{L_i} \mathcal{X}_{i,l}$, where the $\mathcal{T}_i \subset [0, t_0]$ are disjoint intervals, and the intervals $\mathcal{X}_{i,l} \subset \mathbb{R}$ are disjoint for each i. Let $\mathcal{T}_1 = (t_1, t_2]$, let $\mathcal{X}_i = \bigcup_{l=1}^{L_i} \mathcal{X}_{i,l}$ and $\mathcal{B}_i = \mathcal{T}_i \times \mathcal{X}_i$, and note that independence and identical distribution of the X_j gives

$$\begin{aligned}
\mathrm{P}\{N_n(\mathcal{B}_1) = 0\} &= \mathrm{P}\{(X_1 - b_n)/a_n \notin \mathcal{X}_1\}^{\lfloor (n+1)(t_2 - t_1) \rfloor} \\
&= \left[\left\{1 - \frac{\Lambda_n(\mathcal{X}_1)}{n}\right\}^n\right]^{\lfloor (n+1)(t_2 - t_1) \rfloor / n} \\
&\to \exp\{-|\mathcal{T}_1|\Lambda(\mathcal{X}_1)\}, \quad n \to \infty, \\
&= \exp\{-\mu(\mathcal{T}_1 \times \mathcal{X}_1)\}.
\end{aligned}$$

This applies for each $\mathcal{T}_i$, and the corresponding variables X_j are independent, so

$$\begin{aligned}
\mathrm{P}\{N_n(\mathcal{C}) = 0\} &= \prod_{i=1}^{k} \mathrm{P}\{N_n(\mathcal{B}_i) = 0\} \\
&\to \prod_{i=1}^{k} \exp\{-\mu(\mathcal{T}_i \times \mathcal{X}_i)\}, \quad n \to \infty, \\
&= \exp\left\{-\sum_{i=1}^{k} \mu(\mathcal{B}_i)\right\} \\
&= \exp\{-\mu(\mathcal{C})\},
\end{aligned}$$

which establishes the second condition. Thus $\mathcal{P}_n \xrightarrow{D} \mathcal{P}$.