# The Likelihood for a Two-Dimensional Poisson Exceedance Point Process Model

Seokhoon Yun<sup>1)</sup>

### Abstract

Extreme value inference deals with fitting the generalized extreme value distribution model and the generalized Pareto distribution model, which are recently combined to give a single model, namely a two-dimensional non-homogeneous Poisson exceedance point process model. In this paper, we extend the two-dimensional non-homogeneous Poisson process model to include non-stationary effect or dependence on covariates and then derive the likelihood for the extended model.

*Keywords:* Generalized extreme value distribution; generalized Pareto distribution; exceedance point process; non-homogeneous Poisson process.

## 1. Introduction

Let  $X_1, ..., X_n$  be i.i.d. (independent and identically distributed) random variables with common d.f. (distribution function) F. Suppose F belongs to the domain of attraction of an extreme value d.f.  $G_{\xi}$  for some  $\xi \in \mathbb{R}$   $[F \in \mathcal{D}(G_{\xi})]$ , where

$$G_{\xi}(x) := \exp\{-(1+\xi x)_{+}^{-1/\xi}\}, \ x \in \mathbb{R},$$

 $y_+$  being defined by  $\max\{y, 0\}$ . Throughout the paper the case  $\xi = 0$  is interpreted as the limit when  $\xi \to 0$ , so that  $G_0(x) = \exp(-e^{-x}), x \in \mathbb{R}$ .

Write  $M_n := \max\{X_1, ..., X_n\}$ . Then the assumption  $F \in \mathcal{D}(G_{\xi})$  means that there exist normalizing constants  $a_n > 0$  and  $b_n \in \mathbb{R}$  such that, as  $n \to \infty$ ,

$$P\{(M_n - b_n)/a_n \le x\} = P\{M_n \le a_n x + b_n\} = F^n(a_n x + b_n) \to G_{\xi}(x), \ x \in \mathbb{R},$$
(1.1)

from which by letting  $z = a_n x + b_n$ ,  $\mu = b_n$  and  $\sigma = a_n$  we have, for large n,

$$P\{M_n \le z\} = P\{M_n \le a_n x + b_n\} \approx G_{\xi}(x) = G_{\xi}((z - b_n)/a_n) = G_{\xi}((z - \mu)/\sigma)$$
$$= \exp\left[-\left\{1 + \xi\left(\frac{z - \mu}{\sigma}\right)\right\}_{+}^{-1/\xi}\right] =: H(z; \mu, \sigma, \xi), \ z \in \mathbb{R}$$
(1.2)

(cf. Fisher and Tippett (1928), von Mises (1936) and Gnedenko (1943)). Here,  $H(z; \mu, \sigma, \xi)$  is called the GEV (generalized extreme value d.f.) with  $\mu \in \mathbb{R}$ ,  $\sigma > 0$  and  $\xi \in \mathbb{R}$  as its location, scale and shape parameters, respectively. Based on the approximation

Associate Professor, Department of Applied Statistics, University of Suwon, Suwon 445-743, Korea. E-mail: syun@suwon.ac.kr

#### Seokhoon Yun

(1.2), one often fits the GEV to the annual maxima of a data series if one is interested in estimating high quantiles of F.

On the other hand, Pickands (1975) showed that if  $F \in \mathcal{D}(G_{\xi})$ , then for high u less than the right end-point of F,

$$P\{X_i - u \le y | X_i > u\} \approx 1 - \left(1 + \xi \frac{y}{\phi}\right)_+^{-1/\xi} =: G(y; \phi, \xi), \ y > 0$$
(1.3)

for some  $\phi > 0$ . Here,  $G(y; \phi, \xi)$  is called the GPD (generalized Pareto d.f.) with  $\phi > 0$ and  $\xi \in \mathbb{R}$  as its scale and shape parameters, respectively. Based on the approximation (1.3), one may fit the GPD to the exceedances over a high threshold u, instead of fitting the GEV to the annual maxima.

Though each of the two models, GEV and GPD, is useful in its own right, there is also an alternative technique combining them. In fact, Leadbetter, Lindgren and Rootzén (1983) showed that an exceedance point process for exceedances over a high threshold is asymptotically a Poisson process, which again can be extended to the case of a twodimensional non-homogeneous Poisson process (cf. Smith (1989)). The main attraction of this result is that it allows the GEV and GPD models to be combined into a single model. In this paper, we further extend the two-dimensional non-homogeneous Poisson process approximation to include non-stationary effect or dependence on covariates and then derive the likelihood for the extended model. This extended model is particularly useful in dealing with extremal behavior of environmental time series data which typically show non-stationarity such as intra-year seasonality, annual trend and etc.

## 2. The Two-Dimensional Poisson Exceedance Point Process Model and Its Likelihood

As before, let  $X_1, ..., X_n$  be i.i.d. random variables with common d.f. F and assume that  $F \in \mathcal{D}(G_{\xi})$  for some  $\xi \in \mathbb{R}$ . Then, from (1.1), we have, as  $n \to \infty$ ,

$$n(1 - F(a_n x + b_n)) \to -\log G_{\xi}(x) = (1 + \xi x)_+^{-1/\xi}, \ x \in \mathbb{R}.$$

Letting  $u = a_n x + b_n$ ,  $\mu = b_n$  and  $\sigma = a_n$ , we thus have, for large n and high u,

$$n(1 - F(u)) \approx \left\{ 1 + \xi \left( \frac{u - \mu}{\sigma} \right) \right\}_{+}^{-1/\xi}.$$

Define the exceedance point process  $N_n = \{N_n(B), B \in \mathcal{B}((0,1])\}$  by

$$N_n(B) = \#\{i : i/n \in B, X_i > u, i = 1, ..., n\},\$$

i.e.  $N_n(B)$  is the number of points  $X_i$ , i = 1, ..., n, with  $i/n \in B$ , which exceed the threshold u, where  $\mathcal{B}((0, 1])$  denotes the Borel  $\sigma$ -field in (0, 1]. Then it is well known that  $N_n$  is asymptotically a Poisson process on (0, 1] with intensity  $\{1 + \xi(u - \mu)/\sigma\}^{-1/\xi}_+$  for large n and high u (e.g. see Leadbetter, Lindgren and Rootzén (1983) and Resnick (1987)).

This result can be extended to a two-dimensional Poisson process on  $(0, 1] \times (u, \infty)$ (see Definition 2.1 for the definition of a *d*-dimensional Poisson process). From now on, let *n* and *u* be fixed large and high, respectively. In fact, from the point of view of statistical applications, one often assumes that the two-dimensional exceedance point process  $N = \{N(A), A \in \mathcal{B}(D)\}$  on  $D = (0, 1] \times (u, \infty)$  defined by

$$N(A) = \#\{i : (i/n, X_i) \in A, i = 1, ..., n\}$$

is approximately a non-homogeneous Poisson process on D with intensity function

$$\lambda(t,y) = \frac{1}{\sigma} \left\{ 1 + \xi \left( \frac{y-\mu}{\sigma} \right) \right\}_{+}^{-1/\xi-1}, \ (t,y) \in D$$

(cf. Smith (1989)). Then, for z > u, we have

$$P\{M_n \le z\} = P\{N(A) = 0\} \approx \exp\left[-\left\{1 + \xi\left(\frac{z-\mu}{\sigma}\right)\right\}_+^{-1/\xi}\right] = H(z;\mu,\sigma,\xi)$$

and

$$\begin{split} P\{X_i > z | X_i > u\} &= \frac{nP\{X_i > z\}}{nP\{X_i > u\}} = \frac{E(N(A))}{E(N(D))} \approx \frac{\{1 + \xi(z - \mu)/\sigma\}_+^{-1/\xi}}{\{1 + \xi(u - \mu)/\sigma\}_+^{-1/\xi}} \\ &= \left\{1 + \xi\left(\frac{z - u}{\phi}\right)\right\}_+^{-1/\xi} = 1 - G(z - u; \phi, \xi), \end{split}$$

where  $A = (0, 1] \times (z, \infty)$  and  $\phi = \sigma + \xi(u - \mu)$ , since N(A) is approximately a Poisson random variable with mean

$$\int_0^1 \int_z^\infty \lambda(t, y) \, dy dt = \left\{ 1 + \xi \left( \frac{z - \mu}{\sigma} \right) \right\}_+^{-1/\xi}.$$

This result well agrees with those of both (1.2) and (1.3) for z > u, which is of course an advantage in working directly with the non-homogeneous Poisson point process approximation.

**Definition 2.1** For a subset D of  $\mathbb{R}^d$   $(d \in \mathbb{N})$ , let  $\mathcal{B}(D)$  denote the Borel  $\sigma$ -field in Dand let N(A) denote the number of certain random points in  $A \in \mathcal{B}(D)$ . Then the point process  $N = \{N(A), A \in \mathcal{B}(D)\}$  is called a *d*-dimensional non-homogeneous Poisson process on D with intensity function  $\lambda(x) \geq 0, x \in D$  if

- (i) for any  $A \in \mathcal{B}(D)$ , N(A) is a Poisson random variable with mean  $\Lambda(A) = \int_A \lambda(x) dx$ ,
- (ii) for every  $k = 2, 3, ..., N(A_1), ..., N(A_k)$  are independent whenever  $A_1, ..., A_k$  are disjoint in  $\mathcal{B}(D)$ .

In particular, if  $\lambda(x) \equiv \lambda_0, x \in D$  for some constant  $\lambda_0 > 0$ , then N is called a homogeneous Poisson process on D with intensity  $\lambda_0$ .

The Poisson point process approximation to the two-dimensional exceednace point process may be further extended to include non-stationary effect or dependence on covariates by allowing the parameters  $\mu$ ,  $\sigma$  and  $\xi$  to be time-dependent. From now on, let  $X_1, \ldots, X_n$  be independent and let n and u be fixed large and high, respectively. For  $(t, y) \in D = (0, 1] \times (u, \infty)$ , define

$$\lambda(t,y) = \frac{1}{\sigma_t} \left\{ 1 + \xi_t \left( \frac{y - \mu_t}{\sigma_t} \right) \right\}_+^{-1/\xi_t - 1}$$
(2.1)

for some  $\mu_t \in \mathbb{R}$ ,  $\sigma_t > 0$  and  $\xi_t \in \mathbb{R}$ . Assume that the two-dimensional exceedance point process  $N = \{N(A), A \in \mathcal{B}(D)\}$  on D defined by

$$N(A) = \#\{i : (i/n, X_i) \in A, i = 1, ..., n\}$$

is (approximately) a non-homogeneous Poisson process on D with intensity function  $\lambda(t, y), (t, y) \in D$ . Then, for z > u, we have

$$P\{M_n \le z\} = P\{N(A) = 0\} = \exp\left[-\int_0^1 \left\{1 + \xi_t\left(\frac{z - \mu_t}{\sigma_t}\right)\right\}_+^{-1/\xi_t} dt\right],$$

where  $A = (0, 1] \times (z, \infty)$ , since N(A) is a Poisson random variable with mean

$$\int_0^1 \int_z^\infty \lambda(t, y) \, dy dt = \int_0^1 \left\{ 1 + \xi_t \left( \frac{z - \mu_t}{\sigma_t} \right) \right\}_+^{-1/\xi_t} dt$$

Take  $N_1 = N(D)$ , the number of exceedance points of u, and let  $(T_1, Z_1), ..., (T_{N_1}, Z_{N_1}) \in D$  denote the corresponding exceedance points. We then derive the likelihood based on  $N_1$  and  $(T_1, Z_1), ..., (T_{N_1}, Z_{N_1})$ . Write  $A_i = (t_i - \Delta t_i, t_i] \times (z_i - \Delta z_i, z_i]$  for small  $\Delta t_i, \Delta z_i > 0, i = 1, 2, ...,$  which are assumed to be disjoint subsets of D. Then  $N(A_1), N(A_2), ...$  are independent and each  $N(A_i)$  is a Poisson random variable satisfying that  $P\{N(A_i) = 1\} = \lambda(t_i, z_i) \Delta t_i \Delta z_i + o(\Delta t_i \Delta z_i)$  and  $P\{N(A_i) \geq 2\} = o(\Delta t_i \Delta z_i)$ . Thus,

$$P\{N_{1} = n, (T_{1}, Z_{1}) \in A_{1}, ..., (T_{N_{1}}, Z_{N_{1}}) \in A_{N_{1}}\}$$
  
=  $P\{N_{1} = n, (T_{1}, Z_{1}) \in A_{1}, ..., (T_{n}, Z_{n}) \in A_{n}\}$   
=  $P\{N(A_{1}) = \cdots = N(A_{n}) = 1, N(D \setminus \bigcup_{i=1}^{n} A_{i}) = 0\}$   
=  $P\{N(D \setminus \bigcup_{i=1}^{n} A_{i}) = 0\} \times \prod_{i=1}^{n} P\{N(A_{i}) = 1\}$   
=  $\left(e^{-\Lambda(D)} \prod_{i=1}^{n} \lambda(t_{i}, z_{i})\right) \prod_{i=1}^{n} (\triangle t_{i} \triangle z_{i}) + o\left(\prod_{i=1}^{n} (\triangle t_{i} \triangle z_{i})\right)$ 

where

$$\Lambda(D) = \int_0^1 \int_u^\infty \lambda(t, y) \, dy dt = \int_0^1 \left\{ 1 + \xi_t \left( \frac{u - \mu_t}{\sigma_t} \right) \right\}_+^{-1/\xi_t} dt,$$

since

$$P\{N(D \setminus \bigcup_{i=1}^{n} A_i) = 0\} = \frac{P\{N(D) = 0\}}{\prod_{i=1}^{n} P\{N(A_i) = 0\}}$$
$$= \frac{e^{-\Lambda(D)}}{\prod_{i=1}^{n} (1 - \lambda(t_i, z_i) \triangle t_i \triangle z_i + o(\triangle t_i \triangle z_i))}$$
$$= e^{-\Lambda(D)} \prod_{i=1}^{n} (1 + \lambda(t_i, z_i) \triangle t_i \triangle z_i + o(\triangle t_i \triangle z_i))$$
$$= e^{-\Lambda(D)} \left(1 + \sum_{i=1}^{n} \lambda(t_i, z_i) \triangle t_i \triangle z_i + \sum_{i=1}^{n} o(\triangle t_i \triangle z_i)\right).$$

Hence, the likelihood based on  $N_1$  and  $(T_1, Z_1), ..., (T_{N_1}, Z_{N_1})$  is given by

$$e^{-\Lambda(D)} \prod_{i=1}^{N_1} \lambda(T_i, Z_i) = \exp\left[-\int_0^1 \left\{1 + \xi_t \left(\frac{u - \mu_t}{\sigma_t}\right)\right\}_+^{-1/\xi_t} dt\right] \\ \times \prod_{i=1}^{N_1} \left[\frac{1}{\sigma_{T_i}} \left\{1 + \xi_{T_i} \left(\frac{Z_i - \mu_{T_i}}{\sigma_{T_i}}\right)\right\}_+^{-1/\xi_{T_i} - 1}\right].$$

## 3. Maximum Likelihood Estimation

In the previous section we derived the likelihood for a two-dimensional Poisson exceedance point process model with unit time of a typical one-year period. In practice we usually have a long time series of several-year data.

Now, suppose we have daily observations  $X_1, ..., X_n$  of m years, which are assumed to be independent. For a high threshold u, the two-dimensional exceedance point process  $N = \{N(A), A \in \mathcal{B}(D)\}$  on  $D = (0, m] \times (u, \infty)$  defined by

$$N(A) = \#\{i : (im/n, X_i) \in A, i = 1, ..., n\}$$

is assumed to be (approximately) a non-homogeneous Poisson process on D with intensity function  $\lambda(t, y)$ ,  $(t, y) \in D$ , given by (2.1), where  $\mu_t = \mu_t(\beta) \in \mathbb{R}$ ,  $\sigma_t = \sigma_t(\beta) > 0$ ,  $\xi_t = \xi_t(\beta) \in \mathbb{R}$ , and  $\beta$  is a vector of another parameters. For instance, if one wants to consider models of linear annual trends with the three extreme value parameters and intra-year seasonality with the location parameter, then one may choose  $\mu_t(\beta) = \beta_{10} + \beta_{11}t + \beta_{12}\cos(2\pi t) + \beta_{13}\sin(2\pi t), \sigma_t(\beta) = \beta_{20} + \beta_{21}t$  and  $\xi_t(\beta) = \beta_{30} + \beta_{31}t$ , where  $\beta = (\beta_{10}, \beta_{11}, \beta_{12}, \beta_{13}, \beta_{20}, \beta_{21}, \beta_{30}, \beta_{31}).$ 

If  $N_m$  (= N(D)) exceedance points of u are observed at  $(T_1, Z_1), ..., (T_{N_m}, Z_{N_m}) \in D$ , then the maximum likelihood estimator of  $\beta$  is the value of  $\beta$  which maximizes the likelihood  $L(\beta)$  given by

$$L(\beta) = \exp\left[-\int_0^m \left\{1 + \xi_t(\beta) \left(\frac{u - \mu_t(\beta)}{\sigma_t(\beta)}\right)\right\}_+^{-1/\xi_t(\beta)} dt\right]$$
$$\times \prod_{i=1}^{N_m} \left[\frac{1}{\sigma_{T_i}(\beta)} \left\{1 + \xi_{T_i}(\beta) \left(\frac{Z_i - \mu_{T_i}(\beta)}{\sigma_{T_i}(\beta)}\right)\right\}_+^{-1/\xi_{T_i}(\beta) - 1}\right]$$

Computing the maximum likelihood estimator requires a numerical method like the Newton-Raphson method.

## References

- Fisher, R. A. and Tippett, L. H. C. (1928). Limiting forms of the frequency distribution of the largest or smallest member of a sample, *Proceedings of the Cambridge Philosophical Society*, 24, 180-190.
- Gnedenko, B. V. (1943). Sur la distribution limite du terme maximum d'une série aléatoire, Annals of Mathematics, 44, 423-453.
- Leadbetter, M. R., Lindgren, G. and Rootzén, H. (1983). Extremes and Related Properties of Random Sequences and Processes, Springer, New York.
- Pickands, J. (1975). Statistical inference using extreme order statistics, Annals of Statistics, **3**, 119-131.
- Resnick, S. (1987). *Extreme Values, Point Processes, and Regular Variation*, Springer, New York.
- Smith, R. L. (1989). Extreme value analysis of environmental time series: an application to trend detection in ground-level ozone (with discussion), *Statistical Science*, 4, 367-393.
- von Mises, R. (1936). La distribution de la plus grande de valeurs, Reprinted in Selected Papers II, American Mathematical Society, Providence (1954), 271-294.

[Received August 2008, Accepted September 2008]