Chapter 2
The Stable Approach to the Risk Assessment: Estimation of Quantiles of Maximum Event

Abstract The "peak-over-threshold" method is suggested for determination of the limit distribution of the maximum event in a future time period τ, Generalized Extreme Value Distribution (GEV). This approach is based on the Extreme Value Theory, EVT. The method of Maximum Likelihood Estimation (MLE) of unknown parameters is exposed. Several statistical models of the non-stationarity of point process are studied. A modification of the suggested method for the aggregated annual data is given.

Keywords Extreme value theory · Generalized Pareto distribution, GPD · "Peak-over-threshold" method · Intensity of point process · Non-stationarity of point process · Annual data

2.1 The Method

The problem of statistical characterization of extreme, rare events is reduced to estimation of quantiles of high level (close to unity) with the estimates based on a finite sample. We recall that the quantile Q_q of level q, $0 < q < 1$ of a continuous, monotone distribution function $F(x)$ is defined as the root of the equation

$$F(x) = q.$$

Thus, the quantile q is inverse function with respect to the distribution function $F(x)$.

The problem of statistical estimation of quantiles of high level is extremely important for practice. Many applied problems boil down, in fact, to the estimation of such quantiles. If $x_1 < x_2 < \ldots < x_n$ is an ordered sample of iid (independent identically distributed) random variables with a continuous distribution function (DF) then the quantile Q_q of a fixed level q can be estimated by the kth term of the

V. F. Pisarenko and M. V. Rodkin, *Statistical Analysis of Natural Disasters and Related Losses*, SpringerBriefs in Earth Sciences,
DOI: 10.1007/978-3-319-01454-8_2, © The Author(s) 2014

ordered sample x_k ($k =$ entire part of $q \bullet n$). This estimate is known to be consistent for fixed q as $n \to \infty$. The limit distribution of normalized sample quantile

$$n^{1/2}\left(x_k - Q_q\right)/\sigma$$

is the standard Gauss distribution; here $\sigma^2 = F\left(Q_q\right) \cdot \left(1 - F\left(Q_q\right)\right)/f^2\left(Q_q\right)$; $f(x) = F'(x)$—probability density. We stress that the sample quantile tends to its theoretical analog for any distribution with continuous density whatever heavy tail is. In particular, the sample median is consistent estimate of the middle point of a symmetrical distribution (theoretical median) for any tail, whereas the sample mean (arithmetic mean of the sample values) is consistent only for distributions with a light tail. It is true that the efficiency of the sample mean is sometimes higher (e.g. for the Gauss distribution), but not much. For example, the limit standard deviation of sample mean of the standard Gauss distribution is $1/\sqrt{n}$, whereas sample median has limit standard deviation $1.25/\sqrt{n}$. The gain is not big, but the median is guaranteed against possible presence of a heavy tail component in the sample. However, if we try to estimate quantiles of higher levels, $q > 1 - 1/n$, the consistency of sample quantiles disappear. But just such levels are of the most practical interest. For example, suppose that a sample size equal to $n = 500$, so $1 - 1/n = 0.998$, whereas we need the quantile of level $q = 0.9999$ The estimation of quantiles that are "out of sample range", i.e., for $q > 1 - 1/n$, can be effectuated only under some extra assumptions about the distribution in question. There is no magical technique that would yield reliable results for free. Rephrasing a financial truth one can say:

There is no free lunch, when it comes to high quantile estimation!

We shall use for this purpose the Limit Theorem of the extreme value theory (*EVT*) *assuming its validity* (*see, e.g. Embrechts et al.* 1997, *Theorem 3.4.13*). The conditions guaranteeing the validity of this Limit Theorem include the regularity of the original distributions of event sizes in extreme range and boil down to the existence of a *non-degenerate* limit distribution of $\mu_n = \max(x_1, x_2, \ldots, x_n)$ after a proper centering and normalization.

If the Limit Theorem of EVT is valid, then observations exceeding a threshold h, tends (as both h and sample size n tend to infinity) to the Generalized Pareto Distribution (GPD). This approach is called sometimes the "peak over threshold" method. The GDP depends on two unknown parameters (ξ, s):

$$\begin{aligned} GPD_h(x|\xi, s) &= 1 - [1 + (\xi/s) \times (x - h)]^{-1/\xi}, \quad \xi \neq 0; \\ GPD_h(x|0, s) &= 1 - \exp(-(x - h)/s), \quad \xi = 0; \end{aligned} \tag{2.1}$$

here, ξ is the form parameter $(-\infty < \xi < \infty)$, s is the scale parameter $(s > 0)$. The domain of definition depends on parameter values:

$$\text{if } \xi \geq 0, \text{then } x \geq h;$$

$$\text{if } \xi < 0, \text{ then } h \leq x \leq h - s/\xi.$$

We see that for negative ξ the domain of definition of GPD is limited within a finite interval. Because of evident finiteness of any possible physical event (e.g., energy of earthquake) and of the loss values (e.g., number of fatalities) this case is mostly expected in an analysis of empirical data on parameters and losses from natural hazards, but nevertheless, sometimes the unbounded distributions can model empirical data better.

Suppose, the sample $(x_1, x_2, \ldots, x_n)$ is result of observations (peaks over a threshold h) occurred at moments $t_1, \ldots, t_n$ that represent stationary Poisson process with intensity λ. The sample is observed on time interval $[-T;\ 0]$, thus the intensity can be estimated as $\lambda = n/T$. We assume that the threshold h is high enough, so that the conditions of the Limit Theorem of EVT are fulfilled, and, consequently, observations $(x_1, x_2, \ldots, x_n)$ have GPD distribution (2.18) with some parameters $(\xi\,, s)$.

We put down the formulae of the Maximum Likelihood Estimates of GPD-parameters. The GPD-density has form:

$$f(x) = \frac{1}{s}\left(1 + \frac{\xi}{s}(x - h)\right)^{-1/\xi - 1}.$$

Thus, the log-likelihood function equals:

$$L = -n \cdot \log(s) - \left(\frac{1}{\xi} + 1\right)\sum_{k=1}^{n} \log\left(1 + \frac{\xi}{s}(x_k - h)\right).$$

Now one can find numerically the ML-estimates $\hat{\xi}$, $\hat{s}$ providing maximum to L. These estimates are proved to be consistent at least for $\xi > -1/2$ and in that case the limit distribution of normalized quantities

$$\sqrt{n}\left(\hat{\xi} - \xi\right)\Big/|1 + \xi|; \quad \sqrt{n}(\hat{s}/s - 1)\sqrt{2|1 + \xi|}; \quad \xi > -0.5$$

is the standard Gauss distribution. These relations give possibility to construct confidence intervals for parameters ξ, s.

It can be proved (Embrechts et al. 1997, Theorem 3.4.13) that if times of occurrence form a stationary Poisson process and individual sizes are GPD-distributed then the distribution of maximum $M_\tau = \max(x(t_1), \ldots, x(t_m))$ observed on interval $[0; \tau]$, $0 \le t_1, \ldots, t_m \le \tau$, has DF

$$\Phi_\tau(x) = \exp\left(-\lambda\tau[1 + (\xi/s) \cdot (x - h)]^{-1/\xi}\right),$$

apart from terms of the order $\exp(-\lambda\tau)$ which we assume to be negligible.

Our statistical problem consists in estimating quantiles $Q_q(\tau)$ of maximum M_τ in a future time interval τ that we propose as stable robust characteristics of the tail distribution. The quantiles $Q_q(\tau)$ are the roots of the following equation:

$$\Phi_\tau(x) = \exp\left(-\lambda\tau[1 + (\xi/s) \cdot (x - h)]^{-1/\xi}\right) = q. \tag{2.2}$$

Inverting (2.2) as a function of x depending on parameters q, τ we get:

$$Q_q(\tau) = h + (s/\xi) \cdot \left[a \cdot (\lambda\tau)^{\xi} - 1\right],$$

where $a = [\log(1/q)]^{-\xi}$

The GPD-distribution includes all types of tails: power-like ($\xi > 0$), exponential ($\xi = 0$), finite boundary ($\xi < 0$). If $\xi < 0$, then the rightmost boundary of GPD-distribution designed as $M_{\max}$ equals to:

$$M_{\max} = h - s/\xi \tag{2.3}$$

On Fig. 2.1 we show a set of GPD-tails for negative form parameters $\xi = -0.8$; -0.35; -0.1; -0.01; corresponding scale parameters $s = 8.0; 3.5; 1.0; 0.1$ and threshold h = 0; in all cases $M_{\max} = 10$. We see that the more absolute value $|\xi|$ is the more the tail curvature becomes and the steeper the extreme part of the tail decreases. The last curve ($\xi = -0.01$) practically coincides with the tail graph of the exponential distribution.

On Fig. 2.2 we show two GPD-densities, corresponding to negative form parameters $\xi = -0.05$ and $\xi = -0.20$. It is seen from Eq. (2.3) that the more the ratio $-s/\xi$ is, the further to right side $M_{\max}$ is shifted. The density looks like a duck beak. This explains instability of the parameter $M_{\max}$: small variation of ξ-estimate can lead to large excursion of $M_{\max}$.

The main difficulty in parameter estimation of GPD consists in the choice of a proper threshold h. How to cut off the utmost part of the tail for further analysis and statistical inference about asymptotical tail behavior? We use for this purpose the Kolmogorov test with corrections for estimated parameters. The Kolmogorov

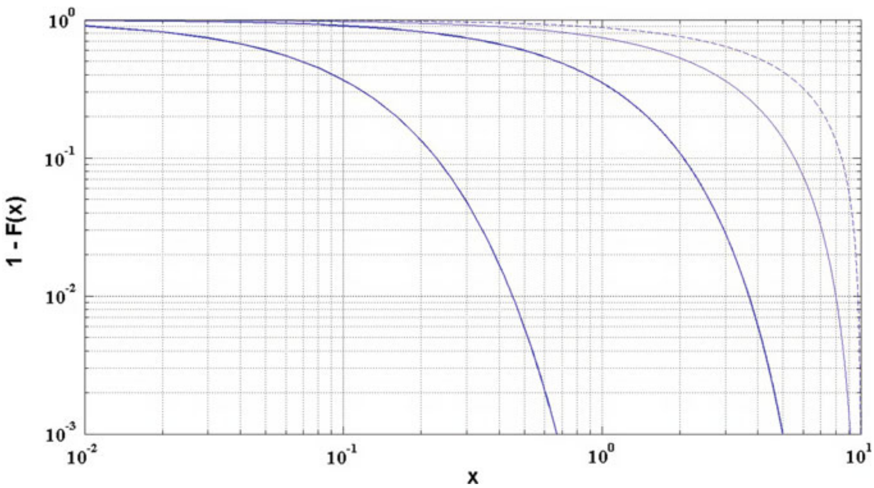

Fig. 2.1 GPD-tails. *Thick line* $\xi = -0.01; s = 0.1$; *intermediate line* $\xi = -0.10; s = 1.0$; *thin line* $\xi = -0.35; s = 3.5$ *dotted line* $\xi = -0.80; s = 8.0$; threshold $h = 0$; in all cases $M_{\max} = 10$

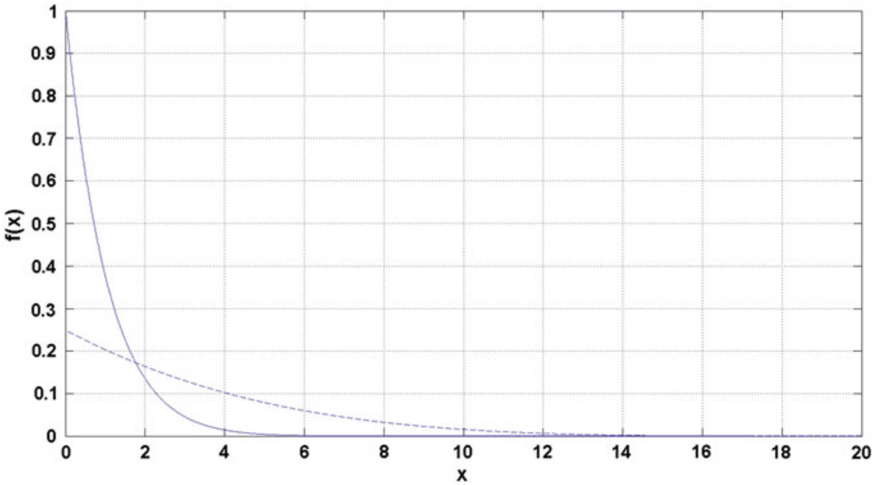

Fig. 2.2 GPD-densities. *Line* $\xi = -0.05,\ s = 1$. *Dotted line* $\xi = -0.20,\ s = 4$

test is a powerful statistical tool, but it needs a representative sample of sufficient size for fully reliable inference. This condition is not always fulfilled in practice, as we shall see below, since the limit theorem of EVT demands a sufficiently high threshold for its validity, and it is left less and less "peaks over threshold" for higher thresholds. This contradiction needs some compromise, sometimes resulting in small size of sample left for GPD-fitting.

The choice of threshold h should satisfy the following restrictions:

1. The threshold h should be high enough, so that the Limit Theorem of EVT mentioned above can be applied. The Kolmogorov distance *KD* between a fitted GPD and the actual sample DF should be small enough. The estimation of significance level for *KD* should take into account the fact that the two fitted GPD-parameters decrease the quantiles of the Kolmogorov distribution (see Pisarenko et al. 2008).
2. Sample size n of observations exceeding h should be large enough to provide reliable estimates of ξ, s and applicability of the Kolmogorov test. Our numerical experiments showed that it is necessary to have $n > 50 \div 80$ (however, in hard situations of deficit of data we were forced to use samples of size $n \cong 30$; of course, in such cases the reliability of statistical estimates is lowered).
3. If ξ is negative (positive), then parameter s (as it was shown in Pisarenko and Rodkin 2010) decreases (increases) with threshold h. Thus, it is reasonable to use thresholds providing decreasing (increasing) s-estimates. If $\xi = 0$ s-estimates should not significantly vary with h.

We use below the restrictions 1–3 to determine the range of h-values suitable for a proper estimation of parameters of the GPD-distribution fitting the tail of the studied sample.

2.2 Non-Stationarity of Natural Processes, the "Operational Time" Method

In this section we consider the problem of non-stationarity of catalogs of natural disasters and related losses. This aspect is very important for application of statistical methods to practical problems. Usually, catalogs consist of pairs (t, X), where t is time of an event (natural catastrophe like earthquake, tornado etc.) and X is information data about this event: coordinates, size, other characteristics. We are going to study events relating to a particular region, so coordinates just belong to this region. Thus, we can consider our catalogs consisting of pairs $(t_i,\ x_i)$, where t_i is time and x_i is size of the i-th event (earthquake magnitude, ground acceleration, number of fatalities, economic loss, etc.). The time occurrences t_i are modeled by a random point process. We shall use as a model the Poisson process with intensity λ (see Embrechts et al. 1997). As to the sizes x_i we assume, that they are random variables (independent of the Poisson process governing occurrences t_i), obeying to a certain (unknown) probability distribution with distribution function (DF) $F(x)$. The non-stationarity of the catalog can be caused by non-stationarity of the Poisson process (in this case the intensity is not constant, but

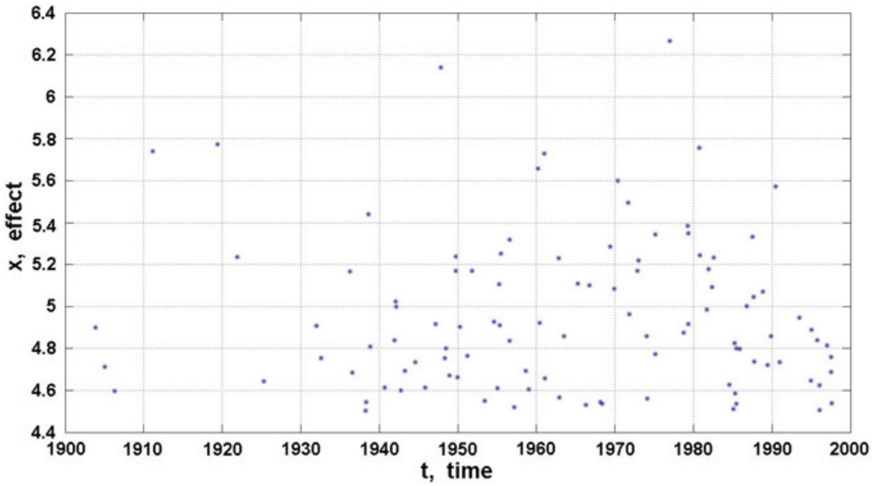

Fig. 2.3 Artificial sample ($n = 105$) of non-stationary Poisson process with the intensity $\lambda(t) = 0.375$, $1900 \leq t \leq 1940$; $\lambda(t) = 1.5$, $1940 < t \leq 2000$. Exponential distribution of effect sizes: $F(x) = 1 - \exp[-2.0 \cdot (x - 4.5)], \quad x \geq 4.5$

depends on t: $\lambda(t)$) and by non-stationarity of the distribution $F(x)$, which results in dependence of distribution function $F(x)$ on time: $F(x, t)$. Of course, both these types of non-stationarity can occur simultaneously.

Figure 2.3 shows an artificial sample $(n = 105)$ of non-stationary Poisson process with the intensity $\lambda(t)$:

$$\lambda(t) = \begin{matrix} 0.375\frac{1}{year}; & 1900 \le t \le 1940 \\ 1.5\frac{1}{year}; & 1940 < t \le 2000 \end{matrix}. \tag{2.4}$$

$$F(x) = 1 - \exp[-2.0 \cdot (x - 4.5)], \quad x \ge 4.5\,(\text{exponential distribution}).$$

Figure 2.4 shows a sample (n = 300) of non-stationary process with an intensity growing with time:

$$\lambda(t) = (t - 1900)/25 + 1; \quad 1900 \le t \le 2000 \tag{2.5}$$

Both figures give a general idea about behavior of intensity, but of course more rigorous statistical tools are needed for accurate estimation of $\lambda(t)$.

In order to test whether the occurrence times are generated by a stationary Poisson process (with constant $\lambda(t)$) we can use following known property of such processes: the conditional distribution of n time occurrences over time interval (0; T) under fixed n coincides with uniform distribution of n points on the interval (0; T). The uniformity of distribution can be tested by the standard Kolmogorov test. We take time occurrences on Fig. 2.3 and calculate the Kolmogorov distance $DK = \sqrt{n}\max|F_n(t) - t/T|$. Here $F_n(t)$ is sample DF of occurrences$F_n(t) = \#(t_i \le t)/n$, and the maximum is taken over all values of t. We

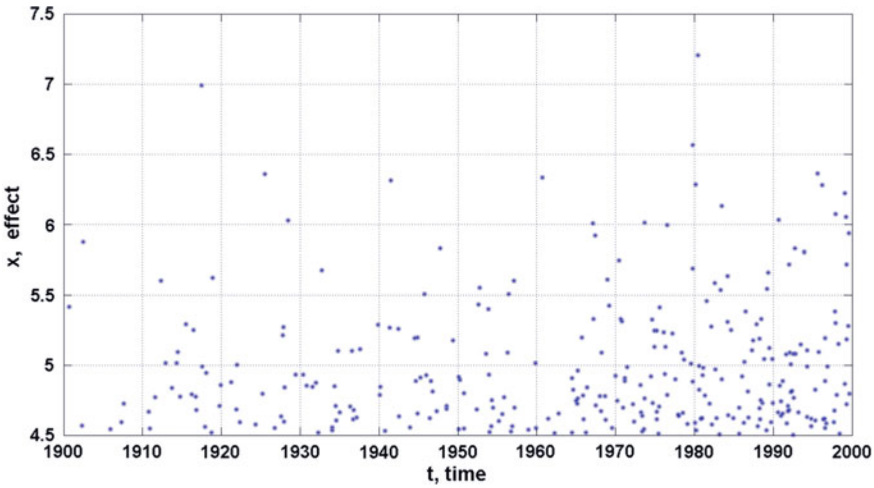

Fig. 2.4 Artificial sample ($n = 300$) of non-stationary process with an intensity growing with time: $\lambda(t) = (t - 1900)/25 + 1$; $1900 \le t \le 2000$. Exponential distribution of effect sizes: $F(x) = 1 - \exp[-2.0 \times (x - 4.5)]$, $x \ge 4.5$

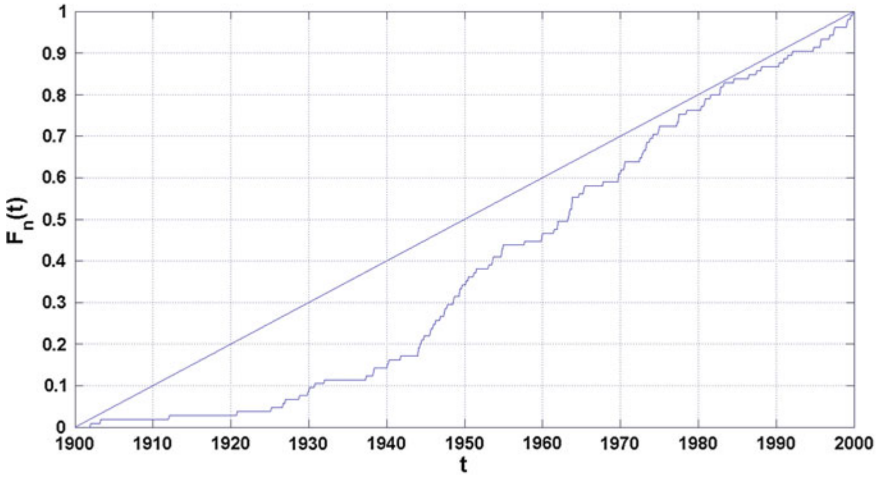

Fig. 2.5 Sample $F_n(t)$ of occurrence times t_j (Fig. 2.3) compared with the uniform DF (*diagonal line*)

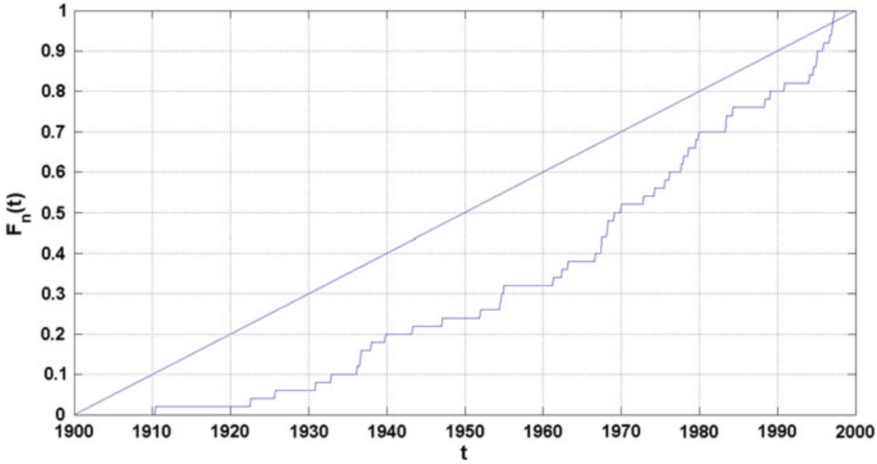

Fig. 2.6 Sample $F_n(t)$ of occurrence times t_j (Fig. 2.4) compared with the uniform DF (*diagonal line*)

get for Fig. 2.3 $DK = 2.79$ and for Fig. 2.4 $DK = 2.11$. On Figs. 2.5 and 2.6 we show corresponding sample $F_n(t)$ compared with the uniform DF. These figures correspond to small p-values: $p = 0.00083$ (Fig. 2.5) and $p = 0.023$ (Fig. 2.4). Thus, hypotheses of stationary should be rejected with very high confidence level.

On Fig. 2.5 one sees a more or less noticeable change of average slope somewhere near 1940 caused by jump of the intensity at $t = 1940$. In order to detect such intensity changes we suggest following procedure. We divide our

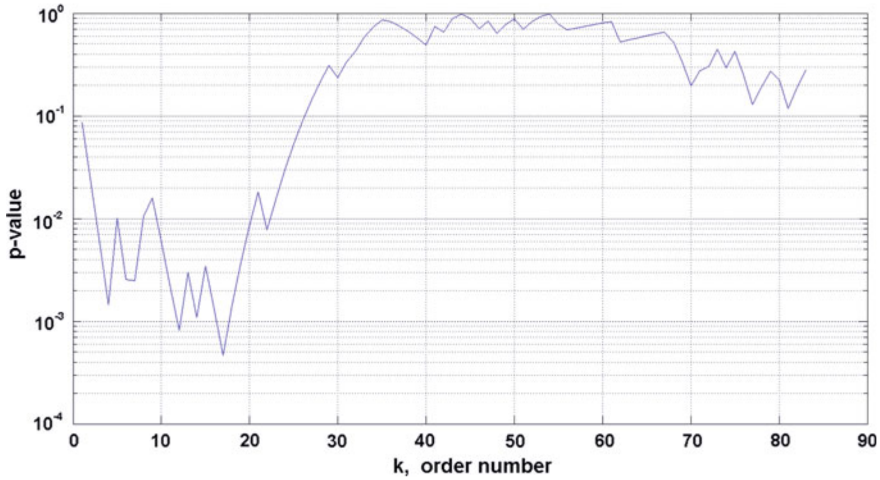

Fig. 2.7 p-values of K–S test: sample $(1:k)$ versus sample $(k+1:n)$ as function of k. k is the order number of the kth event. Minimum point ($k = 18$) corresponds to the date 1940.3

catalog in two parts: sample S_t^1, containing occurrences $t_i \le t$, and sample S_t^2, containing occurrences $t_i > t$. Now, we test the hypothesis H_0 that both samples are generated by one DF using the standard Kolmogorov–Smirnov test. Varying t we find t_0 giving the least p value. It corresponds to the most distinguishable samples $S_{t_0}^1$ and $S_{t_0}^2$. So, one can take point t_0 as an appropriate estimate of a sharp change of intensity. Figure 2.7 shows p-value as function of t. Minimum point t_0 corresponds to the date 1940.3.

An alternative way to judge about stationarity is allowed by looking at cumulative sums of time intervals between adjacent occurrences. If there is a tendency to increase (or decrease) intervals it often can be observed more or less clearly on cumulative curves. On Fig. 2.8 we see the cumulative sum of successive ordered time intervals between occurrences. A visible change of slope (marked by star) is observed at point 18, corresponding to the same time 1940.3 as minimum point at Fig. 2.7.

We have considered two examples of violation of stationarity represented by intensities (2.4)–(2.5). Of course, there are a lot of alternative ways of the violation. E.g. the intensity $\lambda(t)$ can be a polynomial of the 2nd or higher degree. Each practical situation needs its concrete study and appropriate statistical modeling. In our applications to real catalogs we restrict ourselves by intensity models (2.4)–(2.5) taking into account that non-stationarities in these catalogs can be satisfactorily modeled by (2.4)–(2.5), and the use of more sophisticated models are practically excluded by very limited size of available catalogs.

The non-stationarity of catalogs can be connected often with a range of event sizes. Say, seismic catalogs are less representative in lower ranges at the first half of the twentieth century. This fact can be explained by a low level of registration

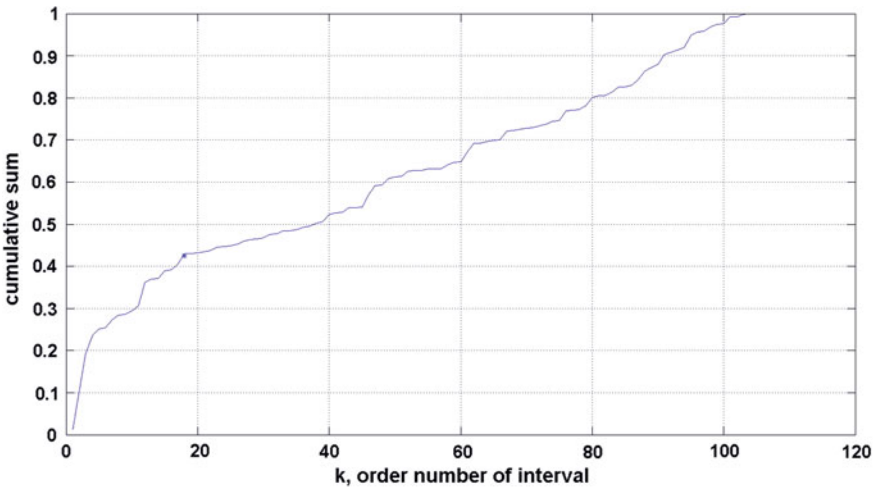

Fig. 2.8 Cumulative sum of successive time intervals between adjacent occurrences in sample Fig. 2.3. Point of slope break is marked by *star*

of small size eartquakes by existed at that time seismic networks. On the other hand, an evolution in preventive services measures can cause the essential decrease in fatalities and in loss values from the natural disasters. The latter tendency competes with a tendency of an loss increase due to the Earth's population growth and increasing value of the technosphere. The use of anti-seismic construction in some countries (e.g. in Japan) gives

an example of decrease of damages caused by earthquakes. Thus, it would be reasonable to look at intensity of events exceeding some lower threshold.

Let us consider an artificial example of seismic catalog. Suppose, seismic flow in the magnitude range $M \geq 4.5$ is formed by stationary Poisson time occurrences t with constant intensity $\lambda_0 = 5$ events per year. But the network registers earthquakes with random probability of registration $p(M,\ t)$:

$$p(M,\ t) = \begin{matrix} 0.4 \cdot M - 1.4; & 1900 \leq t < 1960; \\ 1; & 1960 \leq t \leq 2000. \end{matrix} \quad (2.6)$$

Thus, the intensity of evens registered in our catalog is $\lambda_0 \cdot p(M,\ t)$, and the stationarity is violated. Suppose further, that magnitude distribution is given by the Gutenberg-Richter law:

$$F(M) = 1 - \exp[-2.0 \cdot (M - 4.5)], \quad M \geq 4.5 \quad (2.7)$$

Thus, the stationarity is guaranteed only for events registered after $t = \mathit{1960}$. Figure 2.9 shows the magnitude-time diagram of our schematic example ($n = 299$). Figure 2.10 shows the intensity of three sub-catalogs corresponding to three lower magnitude thresholds: $h_1 = 5.0$, $n = 140$; $h_2 = 5.3$, $n = 87$; $h_3 = 6.0$, $n = 29$; intensities were smoothed by moving 40-year time interval.

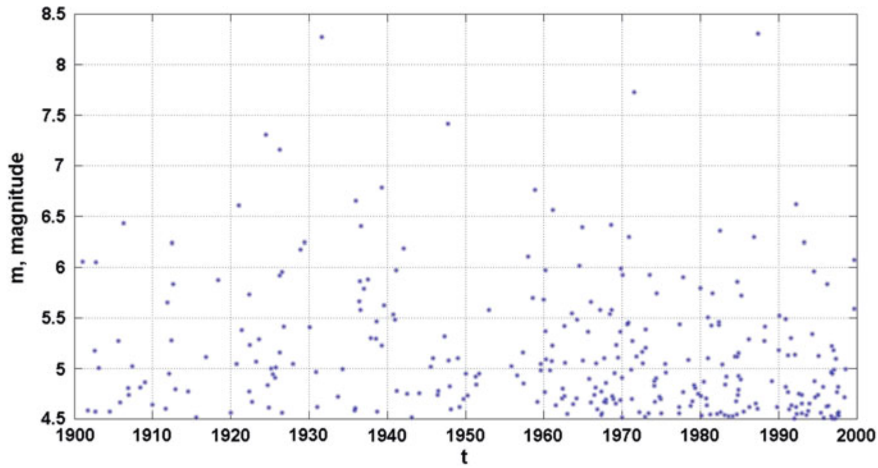

Fig. 2.9 Magnitude-time diagram of artificial example described by Eqs. (2.6) and (2.7), $n = 299$

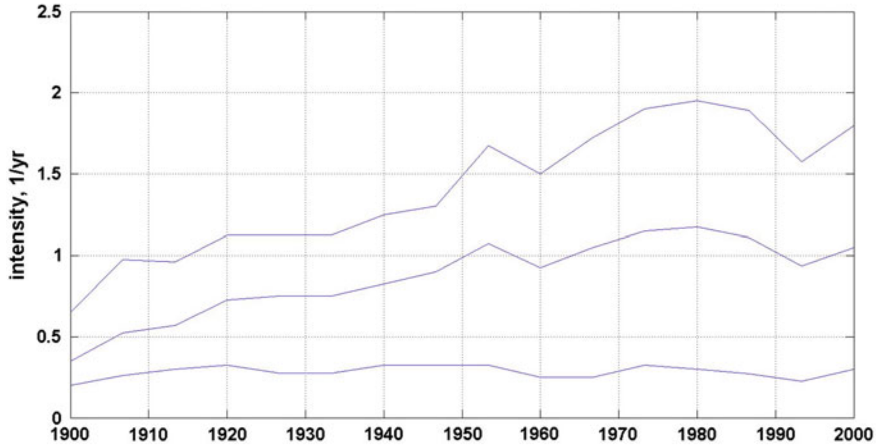

Fig. 2.10 Intensities of three sub-catalogs corresponding to three lower magnitude thresholds: $h_1 = 5.0,\ n = 140$ (*upper curve*); $h_2 = 5,\ n = 87$ (*middle curve*); $h_3 = 6.0,\ n = 29$ (*lower curve*), smoothed by 40-year window

We see that for thresholds $h_1 = 5.0,\ h_2 = 5.3$ intensities vary significantly, whereas for $h_3 = 6.0$ the intensity looks stable.

Suppose, the intensity λ of the point process $\xi(t)$ is not constant but vary with time as arbitrary positive function $\lambda(t)$. If we know $\lambda(t)$, we can transform time scale and pass to a new time $\tau(t)$ (sometimes, it is called "operational time"), so that the process $\xi(t(t))$ is stationary. We can apply to process $\xi(t(t))$ our statistical methods assuming stationarity (in particular, we can estimate desirable quantiles

$Q_q(t)$ for a future time interval τ, and then to return to the original time t. Important condition: the size distribution of sizes should not depend on time (i.e. stationarity of the distribution $F(x)$ is assumed). The needed time transformation is determined as follows. We take inverse function $G(.)$ with respect to $g(t)$

$$g(t) = \int_0^t \lambda(s)\, ds \tag{2.8}$$

Then the direct verification shows that point process $\xi[G(t)]$ is stationary with intensity one. Let us consider an example.

Suppose,

$$\lambda(s) = as + b; \quad a > 0 \tag{2.9}$$

Then

$$g(t) = at^2/2 + bt, \tag{2.10}$$

$$G(t) = \left[(b/a)^2 + 2t/a\right]^{1/2} - b/a \tag{2.11}$$

We affirm that the process $\eta(t) = \xi[\left((b/a)^2 + 2t/a\right)^{1/2} - b/a]$ is stationary with intensity one. Indeed, the mean number of occurrences of the process $\eta(t)$ on interval $[0, T]$ equals to mean number of occurrences of the process $\xi(t)$ on interval $[0, T_1]$, $T_1 = \sqrt{(b/a)^2 + 2T/a} - b/a$ which equals to the integral

$$\int_0^{T_1} \lambda(s) ds = \int_0^{T_1} (as + b) ds = \frac{aT_1^2}{2} + bT_1 = T.$$

The last identity just means that intensity of the process $\eta(t)$ equals unity.

Though the case examined above when the intensity of events' flow changes whereas the DF is stationary appears to be not quite natural it is a reasonable approach to the historical catalogs of strong earthquakes. The completeness of such catalogs depends frequently mainly on casual safety of information about ancient events whereas the DF is rather stationary for a long time because the slow change in technosphere of the ancient society.

Application of the operational time method to study of seismic regime can be found in Ogata (1989, 1993).

2.3 Parametrical Estimation of the Intensity $\lambda(t)$

The log-likelihood of a realization of event occurrences $(t_1, \ldots, t_n)$ in the time interval [0, T] is given by Vere-Jones (1995) and Ogata (1993):

$$\log L = \sum_{k=1}^{n} \log \lambda(t_k) - \int_0^T \lambda(s) ds \tag{2.12}$$

Suppose, that the dependence of the intensity $\lambda(t)$ on time can be modeled by some parametric function. For simplicity of exposition let us take linear function:

$$\lambda(t) = at + b. \tag{2.13}$$

Inserting (2.13) into (2.12) we get:

$$\log L(a, b) = \sum_{k=1}^{n} \log(at_k + b) - \frac{aT^2}{2} - bT \tag{2.14}$$

Now we can apply the full machinery of the likelihood methods for to derive estimators of parameters (a, b). The likelihood equations determining the estimators of maximum likelihood (MLE) are:

$$\frac{\partial L}{\partial a} = -\frac{T^2}{2} + \sum_{k=1}^{n} \frac{t_k}{at_k + b} = 0 \tag{2.15}$$

$$\frac{\partial L}{\partial b} = -T + \sum_{k=1}^{n} \frac{1}{at_k + b} = 0 \tag{2.16}$$

Using these equations we can express parameter a through b:

$$a = \frac{2(n - bT)}{T^2}. \tag{2.17}$$

Inserting (2.17) into (2.16), we get one equation with one unknown parameter b that should be solved numerically.

Using standard technique of the maximum likelihood estimation (Embrechts et al. 1997) one can derive standard deviations of MLE $\hat{a}$, $\hat{b}$:

$$std(\hat{a}) = \left[\frac{a}{T^2} \cdot \frac{w \log(1 + w)}{(1 + w/2)\log(1 + w) - w}\right]^{1/2}, \tag{2.18}$$

$$std(\hat{b}) = \left[\frac{b}{T} \cdot \frac{w(w/2 - 1) + \log(1 + w)}{(1 + w/2)\log(1 + w) - w}\right]^{1/2}, \tag{2.19}$$

where $w = aT/b$. In our example shown on Fig. 2.4 where $\lambda(1900 + t) = 0.04 \cdot t + 1,\ 0 \le t \le 100$, we got $\hat{a} = 0.034 \pm 0.0056$; $\hat{b} = 1.14 \pm 0.26$. Similarly one could use more sophisticated models for $\lambda(t)$.

Finishing discussion on the non-stationarity we consider a generalization of our problem, where the intensity depends both on t and on size x in a very general form given by function $\lambda(t,\ x)$ In this case the log-likelihood is:

$$\log L((t_1, x_1), \ldots (t_n, x_n)) = \sum_{k=1}^{n} \log \lambda(t_k, x_k) - \int_0^T ds \int_m^M \lambda(s, y) dy \tag{2.20}$$

and the second integral represents the intensity of events at time s in the range $(m,\ M)$, where registration was effectuated. It should be remarked, that for fixed t the density $\lambda(t,\ x)$ is proportional to PDF $f(t, x) = \frac{\partial F(x,t)}{\partial x}$. Application of the exposed likelihood technique to numerous problems connected with seismic regime can be found in Vere-Jones (1995), Ogata (1993) and Lyubushin and Pisarenko (1994, 1998).

2.4 Annual Data

Sometimes statistics of the natural catastrophes are published in form of annual data. Now we are going to modify the exposed above method for such data.

Suppose, we have a list of sizes (economic losses, fatalities, etc.) representing yearly figures for N sequential years: $x_1, \ldots, x_N$ Our problem consists in statistical estimation of quantiles $Q_q(\tau)$ of maximum size in question for future τ years (now in contrast to the considered above situation τ is an entire number of years). We assume that the limit theorem (Theorem 3.4.13, Embrechts et al. 1997) is applicable to our data and there exists such sufficiently high threshold h that distribution of peaks over this threshold is well approximated by GPD. We denote parameters of this GPD by $\xi,\ s,$ number of peaks by m (m depends on h), and the sample of peaks by

$$z_1, \ldots, z_m. \tag{2.38}$$

We shall use ratio m/N as natural estimator of probability p exceeding h in future observations: $p \cong m/N$ According to our assumption the conditional distribution of x above h is:

$$GPD_h(x|\xi,\ s) = 1 - [1 + (\xi/s) \cdot (x - h)]^{-1/\xi}, \quad x \ge h. \tag{2.39}$$

Parameters $\xi,\ s$ are estimated by the maximum likelihood method with use of sample (2.38), and the goodness of fit is tested by the Kolmogorov test taking into account the presence of two estimated parameters in the distribution function, as it was remarked earlier. Let us denote τ future sizes as $X_1, \ldots, X_\tau$. Here, the sizes are annual losses. We are interested in distribution of

$$M_\tau = \max(X_1, \ldots, X_\tau) \tag{2.40}$$

The random value X_1 (the annual loss) is less than h with probability $(1-p)$. Under condition that $X_1 > h$ (which happens with probability p) its distribution is governed by GPD (2.39). Thus, for $x \geq h$ the distribution of X_1 equals to

$$(1-p) + p \cdot GPD_h(x|\xi, s) \tag{2.41}$$

We do not study the distribution of X for range $X < h$ since this is of no (or very small) importance for distribution of M_τ. So, if $X < h$ we take $X \approx h$. Then distribution function of X_1, denoted as $F_1(x)$, is zero for $x < h$ and equals to (2.41) for arguments exceeding h. The distribution function $F_M(x)$ of the maximum M_τ equals to the τth degree of $F_1(x)$.

$$F_M(x) = \begin{matrix} 0; & x<h; \\ [(1-p) + p \cdot GPD_h(x|\xi, s)]^\tau; & x > h. \end{matrix} \tag{2.42}$$

We can find the q-level quantile $Q_q(\tau)$ of M_τ from equation:

$$F_M(x) = q \tag{2.43}$$

It follows from (2.42), that if $q > (1-p)^\tau$, then

$$Q_q(\tau) = h + (s/\xi) \cdot \left[\left(\frac{1 - q^{1/\tau}}{p} \right)^{-\xi} - 1 \right] \tag{2.44}$$

For $q < (1-p)^\tau$ the quantile $Q_q(t)$ is not defined. We could put it conditionally equal to h (it can be remarked that usually only quantiles of high level are of interest for risk problems).

If we take instead of $F_1(x)$ the exponential distribution with DF $1 - \exp(-\alpha \cdot (x-h)), \quad x \geq h$ (we recall that the exponential distribution is the limit of GPD as $\xi \to 0$), then we get:

$$Q_q(\tau) = h - (1/\alpha) \cdot \log\left(\frac{1 - q^{1/\tau}}{p} \right) \tag{2.45}$$

References

Embrechts P, Klueppelberg C, Mikosch T (1997) Modelling extremal events. Springer, Berlin

Lyubushin AA, Pisarenko VF (1994) Research on seismic regime using linear model of intensity interaction point processes. Phys Solid Earth 29:1108–1113

Lyubushin AA, Pisarenko VF (1998) A new method for identifying seismicity periodicities. Volcanol Seismolog 20:73–89

Ogata Y (1989) Statistical model for standard seismicity and detection of anomalies by residual analysis. Tectonophysics 169:159–174

Ogata Y (1993) Fast likelihood computation of epidemic type aftershock-sequence model. Geophys Res Lett 20:2143–2146

Pisarenko VF, Rodkin MV (2010) Heavy-tailed distributions in disaster analysis. Springer, Dordrecht-Heidelberg-London-New York

Pisarenko VF, Sornette A, Sornette D, Rodkin MV (2008) New approach to the characterization of M_{max} and of the tail of the distribution of earthquake magnitudes. Pure Appl Geophys 65:847–888

Vere-Jones D (1995) Forecasting earthquakes and earthquake risk. Int J Forecast 11:503–538

http://www.springer.com/978-3-319-01453-1

Statistical Analysis of Natural Disasters and Related Losses
Pisarenko, V.F.; Rodkin, M.V.
2014, XI, 81 p. 84 illus., Softcover
ISBN: 978-3-319-01453-1