BULLETIN OF NATIONAL ACADEMY OF SCIENCES
OF THE REPUBLIC OF KAZAKHSTAN
ISSN 1991–3494
Volume 3, Number 385 (2020), 94 – 102 **https://doi.org/10.32014/2020.2518–1467.74**

Sanem ŞEHRİBANOĞLU

Van Yuzuncu Yil University Statistic Department, Van, Turkey.
E-mail: sanem@yyu.edu.tr

GEV DISTRIBUTION AND PARAMETER ESTIMATION FOR VAN EARTHQUAKE DATA

Abstract. Probabilistic extreme value theory is an interesting and fascinating theory with a great variety of applications. In probability theory and statistics, this distribution is used to model extreme (maximum or minimum) observations. Generalized Extreme Value (GEV)distribution is frequently applied to forecast natural events such as floods, air pollution, extreme sea levels, hydrology, meteorology, climatology, insurance, finance, geology and seismology. In this study general information about Maximum Likelihood Estimation and Bayesian Inference were investigated using the parameter estimation methods. The application of the study was completed using earthquake data from Van provincial center in Turkey from 1995 to 2017. This data used the Maximum Likelihood Estimation and Bayesian Estimation methods in an attempt to predict the severity of earthquakes expected to occur in the future.

Key words: Generalized Extreme Value, GEV, Earthquake Estimation.

Introduction. This theory is applied to model the maximum or minimum distribution of a series of observations. Extreme value theory deals with probabilistic and statistical problems involving the maximum or minimum values of random variables. The theoryattemptsto answer questions about extreme distribution; for example, the possibility of the occurrence of a windspeed in a given place during a certain year or the possibility of a river exceeding a certain height, etc [1]. The distinguishing property of extreme value analysis is measurement of the stochastic behavior of a process at unlikely large or small levels. Especially, it attempts to estimate the probability of things that exceed the monitored values [2]. This distribution was developed as the largest of a set of values. It is first thought to have been applied to the estimation of flood levels and it was also applied to the estimation of the magnitude of earthquakes [3]. In addition, the distribution is applied in forecasting natural events such as floods, earthquakes, volcanic eruptions [4] global warming problems, offshore modeling, rainfall modeling, and wind speed modeling [5]; in engineering such as breaking strength of materials [6]; and in insurance and finance [7].

[8] and [9] мstate that the three types of extreme value distributions can be combined to form a formula like 1.1 for a single parametric family (GEV)with parametersμ, σ, and ξ.

$$G(x) = exp\left[-\left\{1 + \xi\,\frac{(x-\mu)}{\sigma}\right\}_{+}^{1/\xi}\right] \tag{1.1}$$

The GEV family of distributions can be classified as Gumbel (type I), Fréchet (type II) and reverse-Weibull (type III) [10]. Gumbeldistribution is a special case within the GEV distributions (Fisher-Tippett distributions).In addition, the Gumbel distribution is also a member of the Gompertz–Verhulst (GV) family of distributions [11,12]. Although thesethree types of EV distributions are used for sample maxima, reverse-Gumbel, reverse-Fréchet, and Weibull distributions are used for modeling sample minima. Gumbel distribution is a transitional form between the Fréchet and the inverse-Weibull distributions. The GEV distribution is anirregular one which means that the distribution depends on the parameters [13].

The GEV distribution has three parameters; $-\infty < \mu < +\infty$ and $\sigma > 0$ are location andscale parameters, while ξ is a shape parameter $-\infty < \xi < +\infty$. The ξ value determines the type of GEV distribution. $\xi \to 0$ corresponds to Frechet distribution with $\alpha = 1/\xi$, $\xi < 0$ corresponds to Weibull distribution with $\alpha = -1/\xi$, and $\xi \to 0$ corresponds to Gumbel distribution [1,2,9,14].

Provided that for i=1,2,....n

$$z = 1 + \left(\frac{x-\mu}{\sigma}\right)\xi > 0 \tag{1.2}$$

the probability density function (pdf) is given by

$$g(x;\mu,\sigma,\xi) = \begin{cases} \sigma^{-1}\left(1+\xi\left(\frac{x-\mu}{\sigma}\right)\right)^{-\left(\frac{1}{\xi}+1\right)} exp\left[-\left(1+\xi\left(\frac{x-\mu}{\sigma}\right)\right)^{-\frac{1}{\xi}}\right] & ,\xi \neq 0 \\ \sigma^{-1} exp\left[-\left(\frac{x-\mu}{\sigma}\right) exp\left(-\left(\frac{x-\mu}{\sigma}\right)\right)\right] & ,\xi \to 0 \end{cases} \tag{1.3}$$

To model extremes in observational data, the ranges of observations of n length must form a block. Most of the time, these blocks are chosen as annual time periods. An attempt is made to predict the expected maximum numbers occurring after k years taking z_p (recurrence period) with the aid of equation (1.4) with $p = 1/k$ and $p \in [0,1]$ [1,2,15].

$$z_p = \begin{cases} \mu - \frac{\sigma}{\xi}\left[1 - \{-log(1-p)\}^{-\xi}\right] & ,\xi \neq 0 \\ \mu - \sigma log\{-log(1-p)\} & ,\xi = 0 \end{cases} \tag{1.4}$$

Earthquakes occurred with magnitude 7.0 in Van-Erciş county in September 1941, with magnitude 7.3 in Van-Muradiye county in November 1976 and with magnitude 7.2 in Van-Erciş county in October 2011 [16].These earthquakes occurred nearly every 30-35 years, and it is thought that generalized extreme value distribution may provide an answer to the question of what is the highest magnitude earthquake that will occur in Van province in future years.

Estimating the location and time of earthquakes with maximum magnitude is a topic that makes assessment of seismic risk parameters difficult. The extreme value statistics developed by Gumbel provide an appropriate method to estimate the frequency and recurrence intervals of naturally occurring events [17].

Generally, for earthquake prediction methods the $\log(N) = a - bM$ equation developed by Gutenberg-Richter in 1956 is used. This equation estimates using the total number of earthquakes occurring within one year. However, in this study instead of using Gutenberg-Richter's recommended equation, an attempt is made to predict the expected earthquake magnitude for future years by taking the highest earthquake magnitudes (xM) occurring in Van province in Turkey during 23 years. This is similar to studies by [2] about "annual maximum sea levels", [18] about national data buoy center (NDBC) using 21-year wave height (Hs), De Paola et al [19] about precipitation data from Dar Esamam and Addis Ababa, and Gilleland et al. [15] about excessive ozone amounts observed in 184 days from 72 stations in North Carolina.

Parameter estimation. GEV distribution is as follows: the random variable $\boldsymbol{x_i, i = 1,\dots,n}$, $\boldsymbol{F = \{f(x;\theta): \theta \in \Theta\}}$ The GEV distribution, with $\boldsymbol{\mu, \sigma}$ and $\boldsymbol{\xi}$ parameters, is denoted as in equation 1.1. A wide variety of parameter estimation techniques are available but, in this section, we chose to use MLE and Bayesian inference for estimation of GEV parameters.

A variety of techniques including graphical procedures, moment and probability-based estimators are recommended to predict extreme value models [20]. Each technique has its own positive and negative aspects. In this study, we chose to use MLE and Bayesian techniques.

1. Maximum Likelihood Estimation (MLE). MLE is a method that determines values for the parameters of a model. This method attempts to find the parameters reaching the highest levels of a probability function and is among the commonly used approaches. ML estimators are more effective when asymptotic. In some cases, MLE may remain irresolute (i.e., small sample estimators); in these cases, numerical methods like ML Newton-Raphson may be applied.

There are many reasons for using MLE for extreme value models. It is easy to numerically assess log-probability functions. Asymptotic theory ensures mere approaches for standard deviation and confidence intervals. Additionally, likelihood may be generalized to more complicated model structures [20].

Suppose $x_1, x_2 \dots, x_n$are i.i.d. observations with joint probability density function $f(x_1, x_2 \dots, x_n|\theta) = f(x_1|\theta).f(x_2|\theta) \dots f(x_n|\theta)$ is called the likelihood function, where $\theta = (\mu, \sigma, \xi)$ is a vector of unknown parameters. It is often more useful to work with the logarithm of the likelihood function, called the log-likelihood function:

$$L(\theta) = \mathrm{Ln}\left(f(x_1, x_2 \dots, x_n|\theta)\right) = \sum_{i=1}^{n} \mathrm{Ln}\,(f(x_i\,|\theta)).$$

The ML procedure can be used to estimate the GEV parameter θ withthe likelihood function;

$$L(\mu, \sigma, \xi) = \frac{1}{\sigma^n} \prod_i^n \left[\left[1 + \xi\left(\frac{x_i - \mu}{\sigma}\right)\right]^{-\left(\frac{1}{\xi}+1\right)} \exp\left[-\left[1 + \xi\left(\frac{x_i - \mu}{\sigma}\right)\right]^{\frac{-1}{\xi}}\right]\right] \tag{2.1}$$

when z (1.2) is violated, the likelihood is zero and the log-likelihood equals ∞. GEV log-likelihood function can be written as

$$Ln(L) = -n\log\sigma - \left(\frac{1}{\xi} + 1\right)\sum_{i=1}^{n} \log\left(1 + \xi\left(\frac{x_i - \mu}{\sigma}\right)\right) - \sum_{i=1}^{n}\left(1 + \xi\left(\frac{x_i - \mu}{\sigma}\right)\right)^{\frac{-1}{\xi}} \tag{2.2}$$

Then, by definition, MLE estimator $\hat{\theta} = (\hat{\mu}, \hat{\sigma}, \hat{\xi})$ for the unknown parameter θ, is obtained by bringing the first derivative to zero according to the LnL parameter. As used by Hosking et al. [21] andMartins and Stedinger [22] and Raynal-Villasenor [23] the MLE estimations for GEV are as follows:

$$\hat{\theta} = argmaxL\,(\theta, X)_{\theta\epsilon\Theta}$$

If we place this in Equation 1.2;

$$\frac{\partial LnL}{\partial \mu} = \frac{1}{\sigma}\sum_{i=1}^{n} \frac{\left(1 + \xi - z_i^{-1/\xi}\right)}{z_i} = 0$$

$$\frac{\partial LnL}{\partial \sigma} = \frac{1}{\sigma}\left[-n + \sum_{i=1}^{n} \frac{\left(1 + \xi - z_i^{-1/\xi}\right)}{z_i}\left(\left(\frac{x - \mu}{\sigma}\right)\right)\right] = 0$$

$$\frac{\partial LnL}{\partial \xi} = -\frac{1}{\xi^2}\sum_{i=1}^{n}\left[\ln(y_i)\left(1 + \xi - z_i^{-1/\xi}\right) - \frac{\left(1+\xi-z_i^{-1/\xi}\right)}{z_i}\xi\left(\frac{x-\mu}{\sigma}\right)\right] = 0 \tag{2.3}$$

Maximizing Eq. (2.3) according to the parameter $\hat{\theta}$ leads to the maximum likelihood estimate for the GEV distribution.Though this matrix can be analytically calculated, it is easier to use numerical differentiation techniques to complete secondary derivatives and inversion to assess standard numerical routines [2]. The Newton-Raphson algorithm is a powerful technique for solving equations numerically. It solves the likelihood equations $\partial L/\partial\theta$ by iteration. For detailed information see [21,24,25].

Taking $\theta_0$as the initial estimation value for the θ parameter vector, a Taylor series up to second order opens and if it corresponds to 0, the root of θ is obtained.

$$\theta = \theta_0 - \left[\frac{\partial^2 L(\theta)}{\partial\theta\partial\theta'}\right]^{-1} f'(\theta_0) \tag{2.4}$$

If the matrix of the second-degree partial derivative of the $H(\theta^m)$ function is taken as the Hessian matrix and if the vector of the first-degree partial derivative of the $\nabla L(\theta^m)$ functionis taken as the gradient vector, the estimation value for the parameter vector of the $m+1$ iteration of this root (2.4) is obtained as [26].

$$\theta^{m+1} = \theta^m - [H(\theta^m)]^{-1}\nabla L(\theta^m) \quad (2.5)$$

As the MLE method has asymptotic properties, it is a frequently chosen statistical method. MLE can handle cases like missing data, non-stationarity, temporal dependence and covariate effects. MLE may show better performance compared to other methods for small samples; however, the general problem with MLE is its lack of robustness [25].

2. Bayesian Inference. The aim of Bayesian approaches is to determine how the previously obtained data, used as prior knowledge, and the obtained posterior knowledge update the available information [27,28]. In the complex models of Bayesian techniques based on Bayes theorem, if θ is a high-dimensional vector of parameters, calculating the share of (Equation 2.6) may cause a problem even if complicated numerical integrations are used. Simulation-based techniques developed due to these difficulties such as Markov chain Monte Carlo (MCMC) have helped the use ofBayes techniques to become more widespread [2]. The advantage of MCMC is that it does not require asymptotic normality in samples and as a result provides reliable results for small samples [28]. MCMC includes all simulation techniques ensuring parameter estimations by pulling the sample to the simulation path from conditional distributions instead of by using complicated integration techniques. Basically, MCMC uses Markov chain and Monte Carlo iterations [29,30].

Bayes' theorem states that,

$$f(\theta|x) = \frac{f(\theta).f(x|\theta)}{\int_{\Theta} f(\theta).f(x|\theta)d\theta} \ \alpha\, f(\theta).L(\theta;x) \quad (2.6)$$

This distribution (2.6) does not entail a closed form, because of this, it cannot be used in the rest of the inference.

The joint prior density function can be written as $f(\mu,\sigma,\xi) = f(\mu)f(\sigma)f(\xi)$ [2].In situations where there is no information about these three parameters, the noninformative prior of improper prior $f(\mu,\sigma,\xi) \propto 1/\sigma$ may be used, as recommended in the study by Stephenson [31]. When using the MCMC method, the GEV scale parameter is commonly reorganized and

$\varphi = log\sigma$is employed to preserve the positivity of this parameter [2,32]. In this study, $f\ (\mu,\sigma,\xi) \propto 1$ was accepted.

In this situation, using likelihood function in Eq. (2.1), the joint posterior distribution of the parameters can be obtained as follows;

$$f(\mu,\sigma,\xi|x) \propto\ f(\mu,\sigma,\xi)L(\mu,\sigma,\xi|x)$$

$$f(\mu,\sigma,\xi|x) \propto \frac{1}{\sigma^n}\prod_{i}^{n}\left[z^{-\left(\frac{1}{\xi}+1\right)}\exp\left[-(z)^{-\frac{1}{\xi}}\right]\right], \mu,\sigma > 0$$

Similarly, when μ, $\xi$and$x$are given, the conditional posterior distribution of σ with given μ and$x$ is;

$$f(\sigma|\mu,\xi,x) \propto \frac{1}{\sigma^n} + \sum_{i=1}^{n}\left[(1-z)^{-\left(\frac{1}{\xi}+1\right)} - exp[1-z]^{\frac{-1}{\xi}}\right], \mu,\xi,\sigma > 0$$

the conditional posterior distribution of μgiven σ, ξ and $x$is obtained as below;

$$f(\mu|\sigma,\xi,x) \propto \sum_{i=1}^{n}\left[(1-z)^{-\left(\frac{1}{\xi}+1\right)} - exp[1-z]^{\frac{-1}{\xi}}\right], \mu,\xi,\sigma > 0$$

Similarly, the conditional posterior distribution of $\xi$with given μ,σ and x is;

$$f(\xi|\mu,\sigma,x) \propto \sum_{i=1}^{n}\left[(1-z)^{-\left(\frac{1}{\xi}+1\right)} - exp[1-z]^{\frac{-1}{\xi}}\right], \mu,\xi,\sigma > 0$$

For the remainder of the analysis, MCMC is used to obtain a random sample from this distribution. The aim of the MCMC simulation method is to create a random walk in $\theta_t = (\mu_t, \sigma_t, \xi_t)$ parameter space and converge to the final targeted distribution. The Markov chain is a stochastic process and $\theta_{t+1} = (\mu_{t+1}, \sigma_{t+1}, \xi_{t+1})$ value is linked to the previous value in the chain $\theta_t = (\mu_t, \sigma_t, \xi_t)$; however, it forms a stochastic process independent of the others and produces data. If this chain works long enough, it reaches the final distribution of interest [27].

Coles [2] collected reasons why Bayesian analysis needs to be used for extreme value data under two headings. The first is that due to the low number of data it allows the possibility of including other data sources via previous distributions. Secondly the posterior distribution provides better outcomes than ML analysis. However, Coles and Powell [33] emphasized that in cases where prior information is large, inquiries should be made about whether extreme behavior is formulated or not. They mention that when the Bayesian approach to inference is used there may be a contradiction. There are opposing opinions about determining prior knowledge. Especially when personal views about prior knowledge are considered, [34] and others Savage, 1972; Barnet, 1973; Wright and Ayton,1994are encountered [27]. However, during the continuation of studies by Coles and Powell [33] they state the benefits of using spatial prior knowledge in studies with very few data, that the inclusion of a previous distribution form in spatial knowledge stabilizes estimations without causing prejudice and are more variable than ML estimators. Ashour and El-Adl [35] used simple numerical techniques to obtain a range of simulated data. Studies comparing the Bayes estimator with the ML estimators emphasized that the Bayes estimator was more productive than the ML estimators. Kumar et al. [11] applied the quasi Newton-Raphson algorithm ML estimates and Bayesian estimations with the MCMC simulation method, respectively, and found uniform priors and gamma priors results were very close to each other in ML and Bayes estimation results for Gumbel. The study stated that the MCMC method was a substitute method for parameter estimations in the Gumbel model and was more responsive compared to the MLE method. Martins and Stedinger [22] used Monte Carlo simulations in studies and compared the GML, ML, MOM and LM quartile estimators for a GEV distribution. They determined that in small samples, MLE may cause unreasonable and low performance results for GEV distributions while these types of problems were resolved by using Bayesian prior distributions. The study by Gholami [18] concluded that Bayes results were better than ML. Studies De Paola et al. by [19] considered MLE and Bayes methods and stated that the μ parameter estimated with the Bayesian method was more sensitive. Coles et al. [36] stated that Bayes analysis did not provide a completely different interpretation of data; however, it presented a more appropriate and direct route to managing and expressing uncertainties.

The prior and posterior distributions used in the simulation section of this study are listed as follows, respectively.

Prior1:

Priors: $\mu \sim N(\mu_0, \sigma_0)$, $\sigma \sim Ga(\alpha_0, \lambda_0)$, $\xi \sim Ga(\alpha_1, \lambda_1)$

Posterior:

$$exp\left[\sum_{i}^{n} -(z)^{-\frac{1}{\xi}} - \frac{1}{2\sigma_0}(\mu-\mu_0)^2 - \lambda_0\sigma - \lambda_1\xi\right].\sum_{i}^{n}(z)^{-\left(\frac{1}{\xi}+1\right)}.[\sigma^{\alpha_0-n-1}].[\xi^{\alpha_1-1}]$$

Prior 2:

Priors: $\mu \sim GEV(\mu_0, \sigma_0, \xi_0)$, $\sigma \sim Ga(\alpha_0, \lambda_0)$, $\xi \sim Ga(\alpha_1, \lambda_1)$

Posterior:

$$exp\left[\sum_{i}^{n} -(z)^{-\frac{1}{\xi}} - \left(1+\xi_0\left(\frac{\mu-\mu_0}{\sigma_0}\right)\right)^{-\frac{1}{\xi}} - \lambda_0\sigma - \lambda_1\xi\right].\sum_{i}^{n}\left[(z)^{-\left(\frac{1}{\xi}+1\right)}\left(1+\xi_0\left(\frac{\mu-\mu_0}{\sigma_0}\right)\right)^{-\left(\frac{1}{\xi}+1\right)}\right].[\sigma^{\alpha_0-n-1}].[\xi^{\alpha_1-1}]$$

Prior 3:
Priors: $\mu\sim G(\mu_0,\sigma_0)$, $\sigma\sim Ga(\alpha_0,\lambda_0)$, $\xi\sim Ga(\alpha_1,\lambda_1)$
Posterior:

$$exp\left[\sum_{i}^{n} -(z)^{-\frac{1}{\xi}} - \frac{\mu-\mu_0}{\sigma_0} - exp\left(-\frac{\mu-\mu_0}{\sigma_0}\right) - \lambda_0\sigma - \lambda_1\xi\right].\sum_{i}^{n}(z)^{-\left(\frac{1}{\xi}+1\right)}.[\sigma^{\alpha_0-n-1}].[\xi^{\alpha_1-1}]$$

Simulation studies. With the aim of comparing the performance of ML and Bayesian methods, μ, σ and ξfor the GEV (10, 1, 1) were produced from datasets containing 10, 30, 50 and 100 numbers respectively. Openbugsand extremespackage in the R program was used for comparisons.

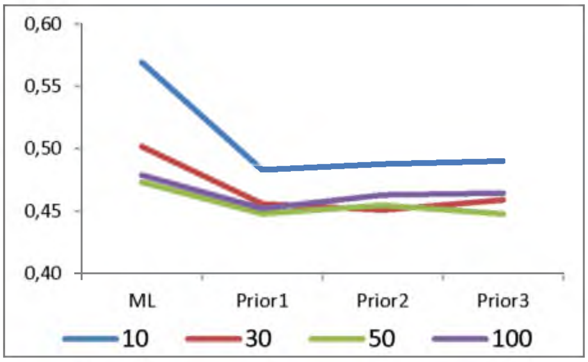
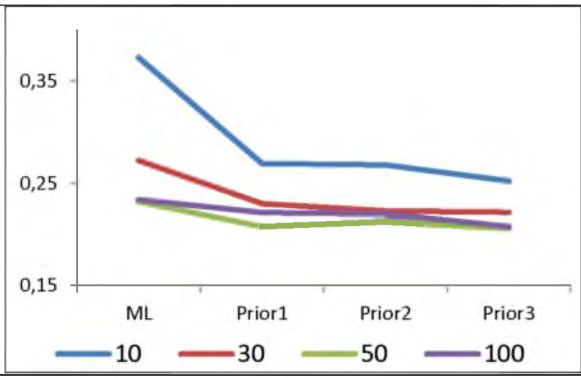
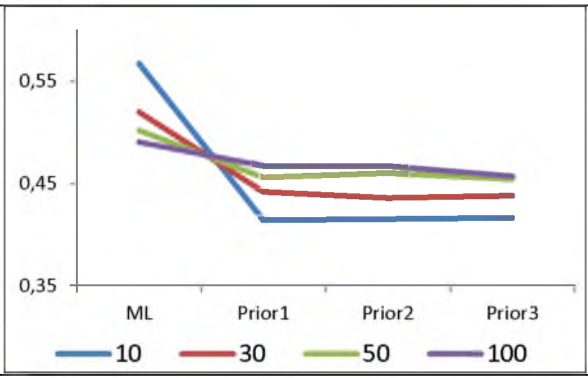
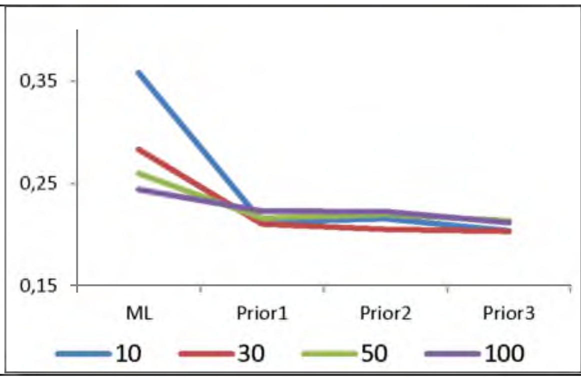
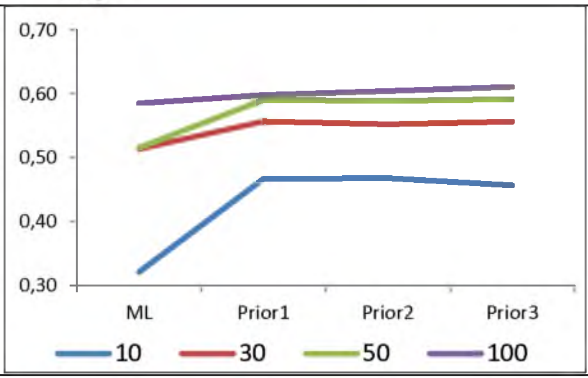
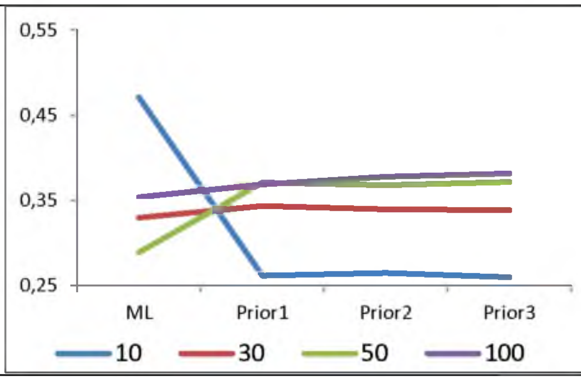
The bias informs how close the expected value of the estimator is to the actual value of the parameter, but not how far from the actual value. MSE is used to assess how close an estimator is to the actual value [37].

There were four applications of ML, Prior 1(μ Normal, σ,ξ gamma), Prior 2(μ Gumbel, σ,ξ gamma) and Prior 3 (μGEV, σ,ξ gamma) and 100 simulation studies were performed for each dataset. In order, Location ~ N (10,1), Location ~ Gumbel (10,1), and Location ~GEV (10,1,1) priors were taken for location, while for scale and shape Scale ~ Gamma (1,1) and Shape ~ Gamma (1,1) priors were used. The bias and MSE results for the simulations are given in table and figure.

Bias and MSE results

		Bias				MSE			
	n	ML	Prior1	Prior2	Prior3	ML	Prior1	Prior2	Prior3
Location	10	0.5691	0.4899	0.4873	0.4827	0.3729	0.2692	0.2679	0.2522
	30	0.5015	0.4588	0.4507	0.4557	0.2722	0.2299	0.2229	0.2216
	50	0.4730	0.4476	0.4545	0.4477	0.2319	0.2074	0.2122	0.2056
	100	0.4786	0.4643	0.4627	0.4519	0.2339	0.2214	0.2199	0.2076
Scale	10	0.5671	0.4139	0.4149	0.4163	0.3580	0.2122	0.2153	0.2036
	30	0.5197	0.4415	0.4356	0.4378	0.2830	0.2102	0.2049	0.2031
	50	0.5016	0.4557	0.4599	0.4539	0.2596	0.2159	0.2196	0.2132
	100	0.4903	0.4669	0.4668	0.4569	0.2439	0.2227	0.2223	0.2115
Shape	10	0.3206	0.4668	0.4678	0.4565	0.4708	0.2613	0.2645	0.2594
	30	0.5134	0.5567	0.5521	0.5565	0.3293	0.3432	0.3392	0.3382
	50	0.5154	0.5905	0.5889	0.5914	0.2885	0.3705	0.3675	0.3713
	100	0.5849	0.5980	0.6042	0.6109	0.3536	0.3684	0.3776	0.3813

When the simulation results are investigated, for the n=10 dataset, ML was seen to have high bias and MSE values (apart from bias results belonging to shape). The MLE simulation results belonging to shape were between -0.4 and 1.9. This causes the bias results for ML to fall. Again, in the same way the MSE values belonging to shape had higher bias values for Prior1, Prior2 and Prior3 compared to ML causing the MSE values to be larger than ML.

<table>
<tr><th>Bias</th><th>MSE</th></tr>
<tr><td>Location</td><td></td></tr>
<tr><td>
</td><td>
</td></tr>
<tr><td>Scale</td><td></td></tr>
<tr><td>
</td><td>
</td></tr>
<tr><td>Shape</td><td></td></tr>
<tr><td>
</td><td>
</td></tr>
</table>

Simulation results

With both methods (ML and Bayesian), they appeared to provide better results as sample numbers increased. As can be seen from studies with different priors, as the sample number increased Bayesian results became more consistent and additionally provided better results than ML in situations with n=10 (table, figure).

Conclusion. Currently it is still unknown when and with what severity earthquakes will occur at a certain point. When the historical process is examined, a large earthquake occurs in Van province nearly every 30-35 years. In this study, it was seen that earthquake data was appropriately modeled by GEV distribution. GEV used in situations showing extreme behavior with the ML and Bayesian approaches were used and the obtained results were compared. An attempt was made to predict earthquakes that will occur in future years. Simulation studies show that as the number of data increase, the ML and Bayesian results show similarities. Bayesian analysis outputs provided more complete inference than MLE. The simulation results applied to different datasets show us that the Bayesian approach is reliable, in addition to being an alternative statistical analysis. Simulation studies show that as the number of data increase the

Bayesian method provides better results than ML. In this study, four different applications were completed.

As stated by Coles and Tawn [39], it is necessary to take care with selection of prior information in the structure of asymptotic models for extreme values. Distributions from similar distribution families ensure we obtain close and consistent results. In simulation studies, Prior2 and Prior3 come from the same distribution family. When the obtained results are investigated, it appears we obtained similar results from them.

When earthquake data is investigated, the four simulation results are close to each other. For Bayesian approaches, the lowest DIC value belongs to Prior3. The results confirm there will be a large Mx earthquake again within the next 30 years.

As stated in the conclusion of the study by Pisarenko et al. [40], extreme value theory provides a good statistical approach to calculate the magnitudes that will occur in future time intervals. However, whether this is accurate or not should always be carefully researched.

Sanem Şehribanoğlu

Ван Юзунджу Иыл (Van Yüzüncü Yıl) университетінің статистика бөлімі, Ван, Түркия

ЖЭЖ (ЖАЛПЫЛАНҒАН ТӨТЕНШЕ ЖАҒДАЙ/МӘН) ТАРАЛУЫ МЕН ВАНДАҒЫ ЖЕР СІЛКІНІСІ ДЕРЕКТЕРІНІҢ ПАРАМЕТРЛЕРІН БАҒАЛАУ

Sanem Şehribanoğlu

Отдел статистики Университета Ван Юзунджу Иыл (Van Yüzüncü Yıl), Ван, Турция

РАСПРЕДЕЛЕНИЕ ОЭЗ И ОЦЕНКА ПАРАМЕТРОВ ДЛЯ ДАННЫХ О ЗЕМЛЕТРЯСЕНИИ В ВАНЕ

Аннотация. Вероятностная теория экстремальных значений – это интересная и увлекательная теория с большим разнообразием применений. В теории вероятностей и статистике это распределение используется для моделирования экстремальных (максимальных или минимальных) наблюдений. Распределение обобщенных экстремальных значений (ОЭЗ) часто применяется для прогнозирования природных явлений, таких как наводнения, загрязнение воздуха, экстремальные уровни моря, гидрология, метеорология, климатология, страхование, финансы, геология и сейсмология. В этом исследовании общая информация об оценке максимального правдоподобия и байесовском выводе была исследована с использованием методов оценки параметров. Применение исследования было завершено с использованием данных о землетрясениях из провинциального центра Ван в Турции с 1995 по 2017 год. В этих данных использовались методы оценки максимального правдоподобия и байесовской оценки в попытке предсказать силу землетрясений, которые могут произойти в будущем.

Ключевые слова: обобщенная экстремальная величина, ОЭЗ, оценка землетрясения.

Information about author:

Şehribanoğlu Sanem, PhD, Asts. Prof., Van YuzuncuYil University Statistic Department, Faculty of Economics and Administrative Sciences, Van, Turkey; sanem@yyu.edu.tr; https://orcid.org/0000-0002-3099-7599

REFERENCES

[1] Smith R.L. Statistics of extremes, with applications in environment, insurance and finance. In: Finkenstadt B, Rootzen H, editors. Extreme values in finance, telecommunication and the environment. Kondon: Chapman and Hall CRC Press; 2004. P. 1-78.

[2] Coles S. An Introduction to Statistical Modeling of Extreme Values. Verlag2001. (Springer).

[3] Forbes C., Evans M., Hastings N., et al. Statistical distributions. 4th ed. Canada: John Willey&Sons.Inc; 2011.

[4] Abbas K., Fu J., Tang Y. Bayesian Estimation of Gumbel Type-II Distribution. Data Science Journal. 2013; 12 (0): 33-46. doi: 10.2481/dsj.13-022.

[5] Tarvirdizade B. Estimation of Pr(X >Y) for Exponentiated Gumbel distribution based on lower record values. Journal of The Turkish Statistical Association. 2013; 6 (3): 103–109.

[6] Lye L.M., Hapuarachchi K.P., Ryan S. Bayes Estimation of the Extreme-Value Reliability Function. IEEE Transactions on Reliability. 1993; 42 (4): 641-644.

[7] Embrechts P., Klüppelberg C., Mikosch T. Modelling extremal events for insurance and finance. Berlin: Springer; 1997.

[8] Jenkinson A.F. The frequency distribution of the annual maximum (or minimum) values of meteorological elements. Quarterly Journal of the Royal Meteorological Society. 1955; 81: 145-158.

[9] Stephenson A., Tawn J. Bayesian Inference for Extremes: Accounting for the Three Extremal Types. Extremes. 2004; 7: 291–307.

[10] Persson K., Rydén J. Exponentiated Gumbel distribution for dstimationof return levels of significant wave height Journal of Environmental Statistics. 2010; 1 (3): 1-12.

[11] Kumar R., Srivastava A.K., Kumar V. Analysis of Gumbel Model for Software Reliability Using Bayesian Paradigm. International Journal of Advanced Research in Artificial Intelligence (IJARAI). 2012; 1 (9): 39-45.

[12] Demirhan H. A generalized Gumbel distribution and its parameter estimation. Communications in Statistics - Simulation and Computation. 2017; 0 (0): 1–20. doi: 10.1080/03610918.2017.1361976.

[13] Cooray K. Generalized Gumbel distribution. Journal of Applied Statistics. 2010; 37 (1): 171-179. doi: https://doi.org//10.1080/02664760802698995. PubMed PMID: WOS:000272848000012; English.

[14] Amin N., Adam M., Ibrahim N., editors. Prior selection for Gumbel distribution parameters using multiple-try metropolis algorithm for monthly maxima PM10 data. AIP Conference Proceedings; 2015.

[15] Gilleland E., Nychka D., Schneider U. Spatial models for the distribution of extremes. In: Clark J, Gelfand A, editors. Hierarchical modelling for the environmental sciences- statistical methods and applications. New York: Oxford University Press; 2006. P. 170-183.

[16] Aldemir M. 1941 depremini hatırlayan var mı? Dogruhaber2013 [cited 2018 22 May]. Available from: https://dogruhaber.com.tr/haber/97447-1941-depremini-hatirlayan-var-mi/

[17] Al-Abbasi J., Fahmi J. Estimating maximum magnitude earthquakes in Iraq using extreme value statistics. Geophys J R Astron Soc 1985; 82: 535-548.

[18] Gholami G. Bayesian estimation of parameters in the exponentiated Gumbel distribution. J Statist Res Iran. 2016; 13: 181–195. doi: 10.18869/acadpub.jsri.13.2.181.

[19] De Paola F., Giugni M., Pugliese F., et al. GEV Parameter Estimation and Stationary vs. Non-Stationary Analysis of Extreme Rainfall in African Test Cities. Hydrology-Basel. 2018 Jun; 5 (2). doi: https://doi.org//10.3390/hydrology5020028. PubMed PMID: WOS:000455156700007; English.

[20] Coles S. The use and misuse of extreme value models in practice. In: Finkenstadt B, Rootzen H, editors. Extreme values in finance, telecommunication and the environment. . London: Chapman and Hall CRC Press; 2004. P. 79-100.

[21] Hosking J.R.M., Wallis J.R., Wood E.F. Estimation of the generalized extreme-value distribution by the method of probability weighted moments. Technometrics. 1985; 27 (3): 251-261.

[22] Martins E.S., Stedinger J.R. Generalized maximum-likelihood generalized extreme-value quantile estimators for hydrologic data. Water Resources Research. 2000; 36 (3): 737-744.

[23] Raynal-Villasenor J.A. Maximum Likelihood Parameter Estimators for The Two Populations GEV Distribution. IJRRAS. 2012; 11 (3): 350-357.

[24] Prescott P., Walden A.T. Maximum-Likelihood estimation of the parame- ters of the three-parameter generalized extreme-value distribution from censored samples. Journal of Statistical Computation and Simulation. 1983; 16 (2): 241-250.

[25] Tsourti Z.A. Evaluation of semi-parametric estimation methods of the extreme-value index and robustifying modifications 2001 [cited 2018 06 June]. Available from: http://www2.stat-athens.aueb.gr/~jpan/diatrives/Tsourti/chapter3.pdf

[26] Searle S.R., Casella G., McCulloch C.E. Variance Components. New York: John Wiley & Sons; 1992.

[27] Gill J. Bayesian methods (A social and behavioral sciences approach). USA: Chapman & Hall/CRC Statistics in the Social and Behavioral Sciences; 2002.

[28] Congdon P. Bayesian statistical modelling. London,UK: John Wiley & Sons; 2003.

[29] Gilks W.R., Richardson S., Spiegelhalter D.J. Markov Chain Monte Carlo in Practice. London: Chapman and Hall; 1996.

[30] Gamerman D. Markov Chain Monte Carlo stochastic simulation for bayesian inference. London: Chapman and Hall; 1997.

[31] Stephenson A. Bayesian Inference for Extreme Value Modeling In: Dey DK, Yan J, editors. Extreme Value Modeling and Risk Analysis Methods and Applications. LLC: Taylor & Francis Group; 2016.

[32] Fawcett L. Short course: Modelling Environmental Extremes 2013 [cited 2018 06 June]. Available from: http://www.mas.ncl.ac.uk/~nlf8/shortcourse/

[33] Coles S.G., Powell E.A. Bayesian methods in extreme value modelling: A review and new developments. Int Stat Rev. 1996 Apr; 64 (1): 119-136. doi: Doi 10.2307/1403426. PubMed PMID: WOS:A1996UD21100006; English.

[34] Berger J.O. Statistical decision theory and bayesian analysis. 2nd ed. Nwe York: Springer; 1985.

[35] Ashour S.K., El-Ad Y.M. Bayesian estimation of the parameters of the extreme value distribution. Egyptian Statistical Journal. 1980; 24: 140-152.

[36] Coles S., Pericchi L.R., Sisson S. A fully probabilistic approach to extreme rainfall modelling. Journal of Hydrology. 2003; 273 (1-4): 35-50.

[37] Hoff P.D. A First Course in Bayesian Statistical Methods 1st ed.: (Springer Texts in Statistics); 2009.

[38] Koeri.boun.edu.tr. 1995-2017 Data 2018 [updated 12 May 2018]. Available from: http://www.koeri.boun.edu.tr

[39] Coles S.G., Tawn J.A. A Bayesian Analysis of Extreme Rainfall Data. Journal of the Royal Statistical Society Series C (Applied Statistics. 1996; 45 (4): 463-478.

[40] Pisarenko V.F., Sornette D., Rodkin M.V. Distribution of maximum earthquake magnitudes in future time intervals: application to the seismicity of Japan (1923–2007). Earth, Planets and Space. 2010; 62 (7): 567-578. doi: 10.5047/eps.2010.06.003.