Discrimination between close hypotheses on tail distributions belonging to the Gumbel maximum domain of attraction.

Igor Rodionov
(Moscow State University)

02.09.13
Ulm

Presentation plan

1. Introduction

The problem of discrimination between distributions with similar tails appears in many statistical applications of extreme value theory.

Herewith it's often convenient to model distributions of medium values by standard distributions, that differs from the asymptotical distribution of tails, but not always.

Fisher-Tippet-Gnedenko theorem (Fisher, Tippet(1928) and Gnedenko(1943)) lies in the foundation of extreme value theory.

Let $X_1, \dots, X_n$ be independent and identically distributed (i.i.d.) random variables with the distribution function $F(x)$. Denote $M_n = \max(X_1, \dots, X_n)$. If there exists a sequence of constants $a_n > 0$ and b_n (n = 1, 2, . . .), such that $\frac{M_n - b_n}{a_n}$ has a nondegenerate limit distribution as $n \to \infty$, i.e.

$$\lim_{n\to\infty} F^n(a_n x + b_n) = G(x), \tag{1}$$

then there exists constants a and b, such that $G(x) = G_\gamma(ax + b)$, where the distribution function G_γ belongs to one of three classes of distributions:

1)Fréchet class of distributions $(\gamma > 0)$:

$$G_\gamma(x) = \begin{cases} 0, & x \leq 0, \\ \exp(-x^{-1/\gamma}), & x > 0, \end{cases} .$$

2)Gumbel distribution $(\gamma = 0)$:

$$G_0(x) = \exp(-e^{-x}), \quad x \in \mathbf{R}.$$

3)reverse-Weibull class of distributions $(\gamma < 0)$:

$$G_\gamma(x) = \begin{cases} \exp(-(-x)^{1/\gamma}), & x < 0, \\ 1, & x \geq 0, \end{cases} .$$

The class of distributions F satisfying (1) is called the *maximum domain of attraction* of G.

The methods of the ratio of likelihoods and the ratio of maximal likelihoods (RML-test) are well-known and often used for discriminating between close types of distributions. These methods were used for discrimination between distributions belonging to the Gumbel maximum domain of attraction in the following works:

Gupta, Kundu and Manglick (2001) (between Weibull and generalized exponential distributions)

Gupta and Kundu (2003) (between Weibull and generalized exponential distributions)

Kundu and Raqab (2007) (between the generalized Rayleigh and log-normal distributions)

Dey and Kundu (2012) (between log-normal and Weibull distributions) etc.

Seem, that contiguity theory of probability measures (see Roussas's monograph) is important instrument for discriminating between families of distributions with close tails and estimating the power of different criteria of discriminating. The key result of this theory is following:

Let $X_1, \ldots, X_n$ be i.i.d. random variables with density $f(x, \gamma_0)$ belonging to family of densities $F = \{f(x, \gamma), \gamma \in \Theta\}$. Let $L(x_1, \ldots, x_n; \gamma)$ be the likelihood function. If family of densities F satisfies some regularity conditions (see Le Cam (1970)), then

$$\ln \frac{L\left(X_1, \ldots, X_n; \gamma_0 + \frac{h}{\sqrt{n}}\right)}{L(X_1, \ldots, X_n; \gamma_0)} \xrightarrow{d} N\left(-\frac{h^2}{2}, h^2\right), \quad n \to +\infty,$$

where $\xrightarrow{d}$ means convergence in distribution, $N(0,1)$ is standard normal random variable and h is the arbitrary constant.

The method of the ratio of likelihoods may be also used for discriminating between close types of distributions by the first order statistics. Consider the ratio of likelihoods

$$R_n(u) = \frac{L(X_{(n)}, \ldots, X_{(n-k_n+1)}; \gamma + t(k_n, u))}{L(X_{(n)}, \ldots, X_{(n-k_n+1)}; \gamma)}$$

as $n \to \infty, k_n \to \infty, \frac{n}{k_n} \to \infty$ ($X_{(1)} \leq \ldots \leq X_{(n)}$ are the n-th order statistics) for family of densities:

$$f(x, \gamma) = C(x, \gamma) \exp(-V(x, \gamma)),$$

where $C(x, \gamma) = C_1(\gamma) + C_2(\gamma)x^{-\beta} + o(x^{-\beta})$, $\beta > 0$, as $x \to \infty$, the function $S(x, \gamma) = V(x, \gamma) - \ln C(x, \gamma)$ is monotone and four times continuous differentiable as $x > x_0(\gamma) > 0$.

2. Regularity conditions and examples

Discuss the regularity conditions, that is imposed on the function $S(x, \gamma)$. The first variant of conditions is following:

1. There exists $\delta = \delta(\gamma) > 0$, such that $\lim\limits_{x \to +\infty} \frac{S(x,\gamma)}{x^{1+\delta}} = +\infty$ for all γ.
2. All of the partial derivatives of $S(x, \gamma)$ up to order 4 either aren't equal to 0 or are equal to 0 identically as $x > x_0(\gamma)$ for all γ. Furthermore, all of the partial derivatives of $S(x, \gamma)$ up to order 3 have a finite or infinite limit as $x \to +\infty$.
3. There exists a finite limit of the expressions like $\frac{\ln |F(x,\gamma)|}{\ln S(x,\gamma)}$ as $x \to +\infty$ for all γ, where $F(x, \gamma)$ is any partial derivative of $S(x, \gamma)$ up to order 3.
   Furthermore, $\lim\limits_{x \to +\infty} \frac{\ln \left| \frac{\partial^k S(x,\gamma)}{\partial \gamma^k} \right|}{\ln S(x,\gamma)} = 1$ for $k = 1, 2, 3$, if these partial derivatives aren't equal to 0 identically.
4. There exists a limit of expressions like $\frac{\frac{\partial \ln |F(x,\gamma)|}{\partial x}}{\frac{\partial \ln S(x,\gamma)}{\partial x}}$ as $x \to +\infty$ for all γ, where $F(x, \gamma)$ is any partial derivative of $S(x, \gamma)$ up to order 3, if this partial derivative isn't equal to 0 identically.

Despite of complicated regularity conditions, the class of distributions satisfying the first variant of conditions is the quite extensive. Let's give some examples of families of densities satisfying these conditions.

1. Weibull-type family of densities:

$$f_W(x,\gamma) = C(\gamma)x^{a(\gamma)}\exp(-x^{b(\gamma)}),\ x \geq 0,$$

   where functions $C(\gamma)$, $a(\gamma)$ and $b(\gamma) > 1$ are four times continuous differentiable.

2. Family of normal densities:

$$f_N(x,\gamma) = \frac{1}{\sqrt{2\pi\gamma}}e^{-\frac{x^2}{2\gamma}},\ \gamma > 0.$$

3. Gumbel-type family of densities

$$f_G(x,\gamma) = \gamma e^{\gamma x}e^{-e^{\gamma x}},\ x \geq 0,\ \gamma > 0.$$

The second variant of conditions is following:

1. There exists $\delta = \delta(\gamma) > 0$, such that $\lim\limits_{x \to +\infty} \frac{S(x,\gamma)}{(\ln x)^{1+\delta}} = +\infty$ for all γ.
2. There exists $\kappa = \kappa(\gamma)$, $0 \le \kappa < 1$, such that $\lim\limits_{x \to +\infty} \frac{\ln \frac{\partial S(x,\gamma)}{\partial x}}{(S(x,\gamma))^{1-\kappa}} = 0$ for all γ.
3. All of the partial derivatives of $S(x,\gamma)$ up to order 4 either aren't equal to 0 or are equal to 0 identically as $x > x_0(\gamma)$ for all γ. Furthermore, all of the partial derivatives of $S(x,\gamma)$ up to order 3 have a finite or infinite limit as $x \to +\infty$.
4. There exists a finite limit of the expressions like $\frac{\ln |F(x,\gamma)|}{\ln \frac{\partial S(x,\gamma)}{\partial x}}$ as $x \to +\infty$ and $\lim\limits_{x \to +\infty} \frac{\ln |F(x,\gamma)|}{S(x,\gamma)} = 0$ for all γ, where $F(x,\gamma)$ is any partial derivative of $S(x,\gamma)$ up to order 3. Furthermore, $\lim\limits_{x \to +\infty} \frac{\ln \left| \frac{\partial^k S(x,\gamma)}{\partial \gamma^k} \right|}{\ln S(x,\gamma)} = 1$ for $k = 1, 2, 3$, if these partial derivatives aren't equal to 0 identically.
5. There exists a limit of expressions like

$$\frac{\frac{\partial \ln |F(x,\gamma)|}{\partial x}}{\frac{\partial \ln S(x,\gamma)}{\partial x}}, \quad \frac{\frac{\partial \ln |F(x,\gamma)|}{\partial x}}{\frac{\partial S(x,\gamma)}{\partial x}} \quad \text{and} \quad \frac{\frac{\partial \ln \left| \frac{\partial^k S(x,\gamma)}{\partial \gamma^k} \right|}{\partial x}}{\frac{\partial \ln S(x,\gamma)}{\partial x}}$$

as $x \to +\infty$ for all γ, where $F(x,\gamma)$ is any partial derivative of $S(x,\gamma)$ up to order 3, if this partial derivative isn't equal to 0 identically.

Let's give some examples of families of densities satisfying the second variant of regularity conditions.

1. log-Weibull-type family of densities:

$$f_{LW}(x,\gamma) = C(\gamma)x^{a(\gamma)}\exp(-(\ln x)^{b(\gamma)}),\ x \geq 0,$$

   where functions $C(\gamma)$, $a(\gamma)$ and $b(\gamma) > 1$ are four times continuous differentiable.

2. Weibull-type family of densities:

$$f_{W}(x,\gamma) = C(\gamma)x^{a(\gamma)}\exp(-x^{b(\gamma)}),\ x \geq 0,$$

   where functions $C(\gamma)$, $a(\gamma)$ and $b(\gamma) \neq 1$ are four times continuous differentiable.

3. Family of log-normal densities:

$$f_{LN}(x,\gamma) = \frac{1}{\sqrt{2\pi\gamma}x}e^{-\frac{(\ln x)^2}{2\gamma}},\ \gamma > 0, x \geq 0.$$

3. Main results

We'll write $S_{x\gamma}$ instead of $\frac{\partial^2 S}{\partial x \partial \gamma}$ and so on. Denote

$$a_{n/k_n} = \overline{F}^{\leftarrow}\left(\frac{k_n}{n}, \gamma\right),$$

where $F(x,\gamma)$ is the distribution function with the density $f(x,\gamma)$, $\overline{F}(x,\gamma) = 1 - F(x,\gamma)$ and $\overline{F}^{\leftarrow}(x,\gamma) = \inf\{t : \overline{F}(x,\gamma) = t\}$.

Denote also

$$t(k_n, u) = \frac{u}{\sqrt{k_n}} \frac{S_x(a_{n/k_n}, \gamma)}{S_{x\gamma}(a_{n/k_n}, \gamma)}.$$

Theorem 1. Suppose, that family of densities $f(x, \gamma)$ satisfies to the first type of regularity conditions, $n \to \infty$, $k_n \to \infty$ and $n/k_n \to \infty$, and there exists ε, $0 < \varepsilon < 2$, such that

$$\lim_{n\to\infty} \frac{k_n}{(\ln \frac{n}{k_n})^{\varepsilon}} = 1.$$

Then

$$\ln R_n(u) \xrightarrow{d} N\left(-\frac{u^2}{2}, u^2\right).$$

Denote

$$H(x)=\sqrt{k_n}t(k_n)\left(\frac{\int\limits_x^{\infty} S_\gamma(y,\gamma)\exp(-S(y,\gamma))dx}{\int\limits_x^{\infty}\exp(-S(y,\gamma))dx}-S_\gamma(x,\gamma)-\right.$$

$$\left.-\frac{S_{x\gamma}(x,\gamma)}{S_x(x,\gamma)}-\frac{S_{xx\gamma}(x,\gamma)}{(S_x(x,\gamma))^2}+2\frac{S_{xx}(x,\gamma)S_{x\gamma}(x,\gamma)}{(S_x(x,\gamma))^3}\right).$$

Note, that $H(x)=-\frac{S_{xx}(x,\gamma)}{S_x^2(x,\gamma)}(1+o(1))$ as $x\to+\infty$.

Theorem 2. Suppose, that family of densities $f(x, \gamma)$ satisfies to the second type of regularity conditions, $n \to \infty$, $k_n \to \infty$ and $n/k_n \to \infty$, and there exists ε, $0 < \varepsilon < 1$, such that

$$\lim_{n\to\infty} \frac{k_n}{n^{\varepsilon}} = 1.$$

Then

$$\ln R_n(u) - \sqrt{k_n} H(a_{n/k_n}) \xrightarrow{d} N\left(-\frac{u^2}{2}, u^2\right).$$

4. References

1. Gupta, R.D., Kundu, D., 2003. Discriminating between Weibull and generalized exponential distributions. *Computational Statistics and Data Analysis* 43 (2), 179 196.
2. Kundu, D., Raqab, M. Z., 2007. Discriminating Between the Generalized Rayleigh and Log-Normal Distribution. *Statistics* 41 (6), 505-515.
3. Roussas G.G.. Contiguity of probability measures: some applications in statistics. *Cambridge Univ. Press*, 1972.
4. Fereira A., Haan L. de.. Extreme value theory. An introduction. *Springer Series in Operations Research and Financial Engineering.* N. Y.: Springer, 2006.
5. Gupta, R.D., Kundu, D., Manglick, A., 2001. Probability of correct selection of Gamma versus GE or Weibull versus GE based on likelihood ratio test. *Technical Report, The University of New Brunswick, Saint John.*
6. Dey, A. K., Kundu, D., 2012. Discriminating between the Weibull and log-normal distributions for Type-II censored data. *Statistics*, 46 (4), 197-214.

Thanks Vladimir I. Piterbarg for scientific advising on my work.