A study on the extreme value distribution of the minimum tensile strength of bolt

Geping Bi[1,*], *Jinzhe* Chen[2], and *Wenhua* Tai[1]

[1]China National Quality Supervision and Inspection Center for Knife, Scissors and Domestic Metal Tool, Yangjiang, Guangdong, 529539, China
[2]China National Accreditation Service for Conformity Assessment, Beijing, 100062, China

Abstract. Taking M12.5 × 1.25 × 65-10.9 bolt as an example, this paper studies the extreme distribution of the minimum value of bolt tensile strength in order to evaluate the reliability and stability of bolt product quality and to verify the production process. Through data collection, parameter estimation, distribution test and extreme prediction, it is concluded that: 1) the distribution of the minimum value of bolt tensile strength conforms to Gumbel extreme distribution; 2) when the return period is 10000, the predicted minimum tensile strength is 1071.7 MPa ± 3.2 MPa, $k = 2$. It is higher than the minimum value of 1040 MPa required by ISO 898-1 standard.

1 Introduction

Fasteners are the most widely used standard parts in the mechanical industry. According to ISO 898-1:2013 standard [1], the tensile strength of fasteners cannot be lower than the specified value. When the chemical composition of the raw material is determined, the tensile strength of the fastener depends on the standardized production process, and the process evaluation is required before the standardized production process is determined.

The study of extreme value has a long history, the earliest research literature is 1824. At present, Gumbel maximum distribution has been widely used in finance, hydrology, architecture, bridge, metallurgy and other fields [2-9]. However, Gumbel minimum application research is less [2, 10-11].

The general requirements for nondestructive testing and metallographic structure defects of metal materials shall be less than or equal to the specified value, i.e. the maximum value shall not be exceeded, which belongs to the application research of maximum value distribution. The mechanical properties of metal materials, such as strength, plasticity, hardness, impact absorbed energy, fatigue and so on, are generally required to be greater than or equal to the specified value, i.e. not less than the minimum value, which belongs to the application research of minimum value distribution. The application research of minimum value has important guiding significance in process design, process evaluation, process control, product quality evaluation and so on.

In this paper, samples were randomly collected from M12.5 × 1.25 × 65-10.9 bolts made of 35CrMo by a specific process. The distribution of the minimum tensile strength of the bolts was studied. Predict the minimum tensile strength that may occur for the production department to evaluate the reliability and stability of bolt product quality and to verify the specific production process.

2 Data collection

Twenty four groups of testing data of M12.5 × 1.25 × 65-10.9 bolts were collected from the testing history samples. Five tensile strength bolts were tested in each group. The testing was conducted according to Clause 9.2 of ISO 898-1:2013 standard [1]. See Table 1 for the data of tensile strength samples.

The minimum values of 24 groups of tensile strength in Table 1 are arranged in non descending order, which are $x_{(1)}$~$x_{(24)}$, and $x_{(1)} \leq x_{(2)} \leq \ldots \leq x_{(24)}$, see the "Data collection" column of Table 2.

The sample average value (Avg) of the minimum value is 1130.3 MPa, the sample standard deviation(S) is 8.42 MPa, the sample skewness coefficient (b_s) is - 0.4 ≤ 0, and the sample kurtosis coefficient (b_k) is 2.0 ≤ 3. See the last row of Table 2. From the skewness coefficient and kurtosis coefficient of the sample, we can see that: 1) the distribution of the minimum tensile strength at the low value range tends to deviate from the center more than that at the high value range; 2) it may have a few small extreme values; 3) the minimum value is not concentrated near the average value; 4) it has a long left tail [11-12].

For the minimum value distribution in Gumbel extreme value distribution, when $x \to -\infty$, the density function of the minimum value distribution decreases exponentially [2]. In this paper, it is assumed that the minimum value distribution of bolt is consistent with

*Corresponding author: bigeping@sina.com

Gumbel extreme value distribution firstly.

Table 1. Sample data of bolt tensile strength / MPa

№	01	02	03	04	05	06	07	08	09	10	11	12
1	1152	1154	1140	1171	1157	1141	1143	1146	1139	1134	1144	1140
2	1151	1145	1149	1130	1121	1141	1133	1152	1152	1145	1147	1144
3	1139	1130	1136	1141	1151	1140	1147	1154	1136	1126	1138	1148
4	1134	1159	1137	1155	1130	1149	1130	1144	1129	1150	1139	1141
5	1142	1129	1138	1161	1133	1139	1145	1138	1159	1134	1151	1163
Min	1134	1129	1136	1130	1121	1139	1130	1138	1129	1126	1138	1140
№	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**
1	1141	1151	1160	1143	1140	1139	1143	1155	1132	1153	1112	1160
2	1138	1136	1140	1145	1138	1159	1131	1154	1137	1153	1126	1150
3	1142	1141	1153	1146	1147	1165	1137	1162	1155	1144	1125	1152
4	1121	1165	1151	1154	1148	1136	1156	1124	1123	1141	1132	1156
5	1122	1143	1118	1143	1135	1144	1154	1171	1120	1152	1143	1120
Min	1121	1136	1118	1143	1135	1136	1131	1124	1120	1141	1112	1120

3 Parameter estimation

The distribution function and probability density function of Gumbel minimum distribution are formula (1) and formula (2), respectively. The empirical guarantee function of Gumbel is formula (3), and the natural logarithm maximum likelihood function of Gumbel distribution samples is formula (4).

$$F(x) = 1 - e^{-e^{\frac{(x-\lambda)}{\sigma}}} \tag{1}$$

$$f(x) = \frac{1}{\sigma} e^{\frac{(x-\lambda)}{\sigma} - e^{\frac{(x-\lambda)}{\sigma}}} \tag{2}$$

$$F(x_i) = P\{X \le x_i\} = p_i = \frac{i}{n+1} \qquad i = 1, \cdots, n \tag{3}$$

$$LL(x,\lambda,\sigma) = \sum_{i=1}^{N} \left[Ln(\frac{1}{\sigma}) + \frac{(x-\lambda)}{\sigma} - e^{\frac{(x-\lambda)}{\sigma}} \right] \tag{4}$$

In formula (1), (2), (3) and (4), where $\infty < x < \infty$, $\infty < \lambda < \infty$, $\sigma$is the scale parameter, λ is the location parameter, X is the random variable, x is the value of X, n is the number of samples used for the evaluation of extreme distribution parameters, p_i is the empirical cumulative density [2, 11-13].

Set:

$$y = \frac{x-\lambda}{\sigma} \tag{5}$$

Then:

$$x = \sigma y + \lambda \tag{6}$$

Formula (1), (2) and (5) can be used to deduce:

$$F(x) = 1 - e^{-e^{y}} \tag{7}$$

$$f(x) = \frac{1}{\sigma} e^{y-e^{y}} \tag{8}$$

Formula (6) and (7) can further deduce:

$$y = Ln\{-Ln[1 - F(x)]\} \tag{9}$$

$$x = \sigma y + \lambda = \sigma Ln\{-Ln[1 - F(x)]\} + \lambda \tag{10}$$

$$x_p = \sigma y + \lambda = \sigma Ln[-Ln(1-p)] + \lambda \tag{11}$$

In formula (11), x_p is called the p-quantile of distribution function $F(x)$. p is the probability associated with the p-quantile of $F(x)$ [12].

If T is the return period, the minimum strength of the bolt predicted at the return period T is:

$$x(T) = \sigma Ln\left[-Ln(1-\frac{1}{T})\right] + \lambda \tag{12}$$

In formula (10-12), there are two unknown parameters, scale parameter σ and location parameter λ. Two unknown parameters were evaluated by the maximum likelihood method [2-11]. The specific methods are as follows:

1) The minimum value samples of bolts in Table 1 are arranged in non descending order.

2) The empirical cumulative density p_i of each sample is calculated by formula (3).

3) y_i is calculated from formulas (3) and (9).

4) Set the average value as the initial value of λ, and the standard deviation as the initial value of σ, that is, λ = 1130.3 MPa, σ = 8.42 MPa. From formula (4), the component $L(x_i)$ of the maximum likelihood function of the natural logarithm of the sample is calculated in turn.

5) The maximum likelihood function $LL(x, \lambda,\sigma)$ is calculated from formula (4).

6) Using the Excel's "planning solving" function, the σ and λ that make LL get the maximum value are calculated by Newton iterative method. From this, It can be concluded that: $LL(x, \lambda_{ML},\sigma_{ML})$ =-83.863, σ_{ML}=6.79, λ_{ML}=1134.2. σ_{ML} and λ_{ML} are scale and location parameters calculated by maximum likelihood method.

The above keyabbreviations, symbols and datas are in the "parameter estimation" column and the last row in Table 2.

Through the parameter evaluation, the extremum distribution function and probability density function of the minimum value of bolt tensile strength are formula (13) and formula (14), respectively.

$$F(x)=1-e^{-e^{\frac{(x-1134.2)}{6.79}}} \tag{13}$$

$$f(x)=\frac{1}{6.79}e^{\frac{(x-1134.2)}{6.79}-e^{\frac{(x-1134.2)}{6.79}}} \tag{14}$$

4 Distribution test

Table 2. Sample analysis and calculation of minimum tensile strength of bolt

Data collection			Parameter estimation			K-S test			$x_i\,(T)$
i	x_i	Min /MPa	p_i	y_i	$L(x_i)$	Fn(x_i)	Fo(x_i)	Dn(i)	
1	x_1	1112	0.04	-3.1985	-5.226	0.000	0.037	0.0373	1112.5
2	x_2	1118	0.08	-2.4843	-4.396	0.042	0.088	0.0462	1117.3
3	x$_3$	1120	0.12	-2.0570	-4.133	0.083	0.116	0.0329	1120.2
4	x_4	1120	0.16	-1.7467	-4.133	0.125	0.116	0.0505	1122.3
5	x_5	1121	0.20	-1.4999	-4.005	0.167	0.133	0.0750	1124.0
6	x_6	1121	0.24	-1.2930	-4.005	0.208	0.133	*0.1166*	1125.4
7	x_7	1124	0.28	-1.1132	-3.643	0.250	0.200	0.0921	1126.6
8	x_8	1126	0.32	-0.9528	-3.424	0.292	0.258	0.0750	1127.7
9	x_9	1129	0.36	-0.8068	-3.148	0.333	0.372	0.0385	1128.7
10	x_{10}	1129	0.40	-0.6717	-3.148	0.375	0.372	0.0448	1129.6
11	x_{11}	1130	0.44	-0.5450	-3.074	0.417	0.417	0.0418	1130.5
12	x_{12}	1130	0.48	-0.4248	-3.074	0.458	0.417	0.0835	1131.3
13	x_{13}	1131	0.52	-0.3093	-3.012	0.500	0.464	0.0774	1132.1
14	x_{14}	1134	0.56	-0.1973	-2.916	0.542	0.621	0.0796	1132.9
15	x_{15}	1135	0.60	-0.0874	-2.922	0.583	0.675	0.0920	1133.6
16	x_{16}	1136	0.64	0.0214	-2.953	0.625	0.728	0.1034	1134.3
17	x_{17}	1136	0.68	0.1305	-2.953	0.667	0.728	0.0618	1135.1
18	x_{18}	1136	0.72	0.2413	-2.953	0.708	0.728	0.0216	1135.8
19	x_{19}	1138	0.76	0.3557	-3.104	0.750	0.826	0.0762	1136.6
20	x_{20}	1138	0.80	0.4759	-3.104	0.792	0.826	0.0346	1137.4
21	x_{21}	1139	0.84	0.6057	-3.233	0.833	0.868	0.0350	1138.3
22	x_{22}	1140	0.88	0.7515	-3.407	0.875	0.905	0.0296	1139.3
23	x_{23}	1141	0.92	0.9265	-3.631	0.917	0.934	0.0241	1140.5
24	x_{24}	1143	0.96	1.1690	-4.266	0.958	0.974	0.0259	1142.1
Avg = 1130.3 MPa, S = 8.42 MPa, b_s = -0.4, b_k = 2.0; $LL(x,\lambda_{ML},\sigma_{ML})$ = -83.863, λ_{ML}=1134.2, σ_{ML} = 6.79; D_n= 0.1166 <$D_{24,0.05}$ = 0.2776.									

Whether the theoretical distribution function and the theoretical density function of the minimum tensile strength of bolts are consistent with the actual needs distribution test. In this paper, K-S test method [2-9] is used. See Table 2 for the process data of K-S test. The specific methods are as follows:

1) Calculate the empirical distribution function $F\text{n}(x_i)=(i-1)/\text{n}$;

2) Calculate the theoretical distribution function $F\text{o}(x_i)$ according to formula (13).

3) Calculation statistics Dn(i), $D\text{n(i)} = \max(\mid F\text{n}(x_i)-F\text{o}(x_i)\mid, \mid i/\text{n}-F\text{o}(x_i)\mid)$.

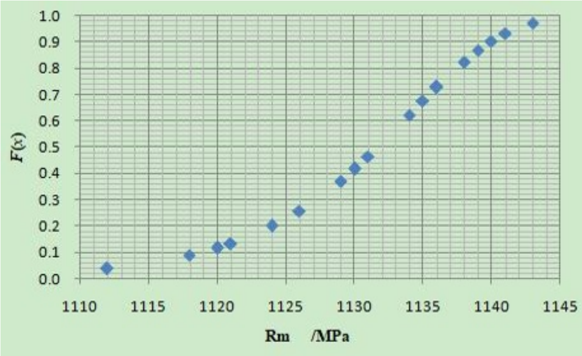
4) Let Dn be the maximum of Dn(i), with Dn= Dn(16)=0.1034.

5) In the case of n = 24 and significance level of 95%, the critical value of K-S test $D_{24,0.05}$=0.2776 can be obtained.

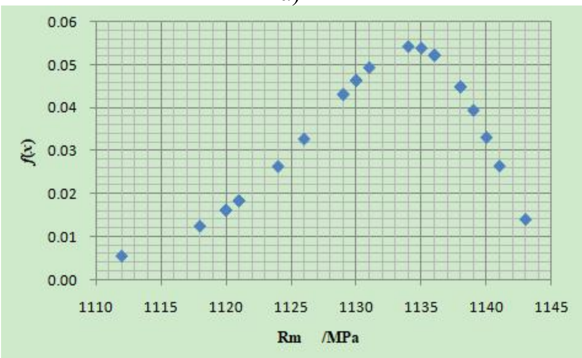
It can be seen that $D\text{n}=0.1166<D_{24,0.05}=0.2776$, so it can be determined that the theoretical distribution function and theoretical density function formula (13-14) of the minimum value of the bolt tensile strength are consistent with the actual extreme value distribution.

The above key abbreviations, symbols and datas are in the " K-S test " column and the last row in Table 2.

The distribution curve of the minimum value of the bolt tensile strength obtained according to formula (13-14) is shown in Figure 1.



a)



b)

Fig.1. Distribution of minimum tensile strength of bolt. a) Distribution curve of minimum; b) Probability density curve of minimum

5 Mimimum value prediction

According to the theoretical extreme value distribution function formula (13-14) and formula (12) of minimum value of bolt tensile strength, it can be concluded that the prediction function of minimum value of bolt strength with return period of T is formula (14):

$$x(T) = 6.79Ln\left[-Ln(1-\frac{1}{T})\right]+1134.2 \tag{14}$$

$$SE = \sqrt{\frac{\sum_{i=1}^{n}\left[x_i - x_i(T)\right]^2}{n-1}} \tag{15}$$

According to formula (14), the theoretical value corresponding to 24 samples can be calculated, as shown in $x_i(T)$ in Table 2. The fitting standard deviation between the theoretical value and the actual value of the sample is calculated according to formula (15) [14], then the expanded uncertainty (U) of the theoretical value is $\pm k\times SE$. After calculation, SE = 1.6Mpa, 95% confidence interval: U= ±3.2MPa， k=2.

Figure 2 is the quantile plot of the actual value and theoretical value of the sample, i.e. Q-Q Plot. The dotted line in the figure is the upper and lower limit of the expanded uncertainty.

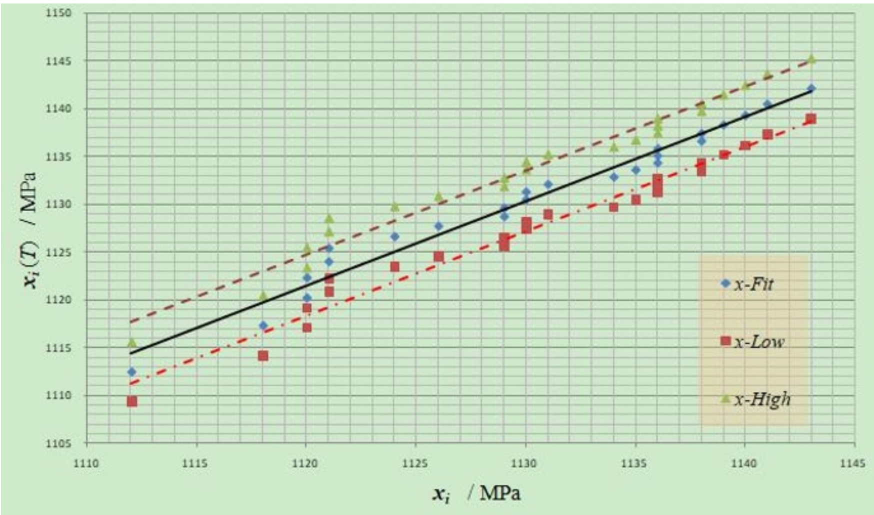


Fig.2. Q-Q Plot

Through the calculation and analysis of the extreme value of 24 groups of samples, it can be predicted that when the return period T = 10000, the minimum value of bolt tensile strength (Rm) is 1071.7 MPa ± 3.2 MPa, k = 2. That is to say, for M12.5 × 1.25 × 65-10.9 bolt made of 35CrMo, the technical requirement [1] is Rm ≥ 1040 MPa. When the return period is 10000, the minimum Rm is higher than the technical requirement. It can be determined that the quality of the bolt is stable and reliable and the specific production process is reliable. The prediction results can be explained as follows: For every 10000 samples, the probability of Rm ≤ 1071.7 MPa ± 3.2 MPa is 1 / 10000, i.e. 0.01%. Or, Rm ≤ 1071.7 MPa ± 3.2 MPa will appear in every 10000 samples; Or, in every 10000 samples, Rm of each sample will be less than 1071.7 MPa ± 3.2 MPa with a probability of 0.01%.

6 Conclusion

Through the research, analysis and calculation of the minimum value distribution of 24 groups of bolts, it can be concluded that:

1) the distribution of the minimum value of bolt tensile strength conforms to Gumbel extreme distribution;

2) when the return period is 10000, the predicted minimum tensile strength is 1071.7 MPa ± 3.2 MPa, k = 2.

It is higher than the minimum value of 1040 MPa required by ISO 898-1 standard and it can be determined that the quality of the bolt is stable and reliable and the specific production process is reliable.

References

[1] ISO 898-1:2013 Mechanical properties of fasteners made of carbon steel and alloy steel — Part 1: Bolts, screws and studs with specified property classes — Coarse thread and fine pitch thread.

[2] Shi Daoji, 2006, Practical statistical method of extreme value. (Tianjin: Tianjin Science and Technology Press). (In Chinese)

[3] Hou Shu-min, LI You-rong, LIU Guang-lin. 2005. Application of K-S Test and Wavelet Analysis in Rolling Bearing Fault Diagnosis. *Coal Mine Machinery*, 1, 129-130. (In Chinese)

[4] Nan Bo. 2012. Statistical analysis of snow load based on K-S inspection method. *Journal of Lanzhou University of technology*. **38,** 1, 115-119. (In Chinese)

[5] LIU Zhi-fei, YAN Pei-jun, XU Ren-ping, etc. 2013. Effect of Strain Age on Impact Toughness of ADB610 Steel. *Development and Application of Materials*. **28**, 3, 36-39. (In Chinese)

[6] GB 50199-2013 Unified standard for reliability design of hydraulic engineering structures. (In Chinese)

[7] DIAO Hongquan, LIN Jiahui, Yao Huiming. 2019. Statistical parameter estimation method and application of Gumbel distribution. *Science and Technology Innovation Herald*. 08, 130-132. (In Chinese)

[8] ZHANG Xinsheng, CAO Naining, LI Yayun. 2016. Residual Life Prediction of Buried Oil and Gas Pipelines Based on Gumbel Extreme Value Type I Distribution. *Journal of Chinese Society for Corrosion and Protection*. **36**, 4, 370-374. (In Chinese)

[9] ASTM E2283-08 (2019) Standard Practice for Extreme Value Analysis of Nonmetallic Inclusions in Steel and Other Microstructural Features.

[10] Hong Wei. 1988. The Limited Value Distribution and Application of Annual Lowest Climate Temperature . *Journal of Fujian College of Forestry*, **8**, 3, 229-234. (In Chinese)

[11] IEC/TR 60727-2:1993 Evaluation of electrical endurance of electrical insulation systems - Part 2: Evaluation procedures based on extreme-value distributions.

[12] ISO 3534-1:2006 Statistics — Vocabulary and symbols — Part 1: General statistical terms and terms used in probability.

[13] ISO 5479:1997 Statistical interpretation of data — Tests for departure from the normal distribution.

[14] Shi Daoji. 1988. An Analysis of Variance Test for the Extreme Value Distribution. *Journal of Tianjin University*, 2, 116-121. (In Chinese)