TECHNOMETRICS VOL. 15, NO. 4 NOVEMBER 1973

Discrimination Between the Log-Normal and the Weibull Distributions

ROBERT DUMONCEAUX

St. Johns University
Collegeville, Minnesota

and

CHARLES E. ANTLE

The Pennsylvania State University
University Park, Pennsylvania

The log-normal and the Weibull are often considered for situations in which a skewed distribution for a non-negative random variable is needed. The ratio of maximized likelihoods provides a good test for selecting one of these. A table of the necessary critical values is given. The table may also be used for discriminating between the normal and the type 1 extreme value distributions.

KEY WORDS

Model Discrimination
Weibull
Log-normal
Ratio of Maximized Likelihoods

1. INTRODUCTION AND NOTATION

Suppose an experimenter has observed $x_1, x_2, \cdots, x_n$ and on the basis of these observations he wishes to choose either the log-normal with density function

$$f_{ln}(x; \mu, \sigma) = \exp\{-.5(ln x - \mu)^2/\sigma^2\}/(x\sqrt{2\pi\sigma^2}) \quad \text{for } x > 0, \qquad (1.1$$

or the two parameter Weibull with density function

$$f_w(x; b, c) = c(x/b)^{c-1}b^{-1}\exp\{-(x/b)^c\} \quad \text{for } x > 0, \qquad (1.2)$$

as a model. We assume the parameters would be unknown to him.

If the experimenter transforms these data by taking logarithms, say $z_i = ln_e x_i$, then he would be choosing between the normal with parameters μ and σ or the type 1 extreme value distribution for minimums given by

$$F_{1\,\min}(z; A, B) = 1 - \exp(-\exp((z - A)/B)).$$

Thus his problem becomes that of choosing between two location and scale parameter distributions with unknown parameters. It is noted in the paper by Dumonceaux, Antle and Haas (1973) that whenever the two models are location and scale parameter distributions the distribution of the ratio of maximized likelihoods (hereafter called the RML) does not depend upon the values of the unknown parameters and it is therefore reasonable to obtain critical values for

Received May 1970; revised March 1971

this test by simulation whenever they cannot be calculated analytically. We have done this and the resulting tables are given in Section 2.

It can easily be seen that the RML will be the same value when calculated with the z's, the normal and the type 1 distribution for minimums, or with the x's, the log-normal and the Weibull. Hence, the RML may be calculated in either setting. We shall think in terms of the log-normal vs. Weibull setting because this is the problem of particular interest to us.

We would also note that these tables may be used directly for the purpose of discriminating between the normal and the type 1 for maximums distribution given by

$$F_{1\ max}(y; A, B) = \exp(-\exp(-(y - A)/B)).$$

(This is a result of the fact that the negative of a type 1 for maximum random variable has a type 1 for minimums distribution.) An example for this test is given in Section 3.

2. The RML Test for Discrimination Between the Log-Normal and the Weibull Distributions

2.1 *Critical Values and Power of the* RML *Test with* H_0 : *Log-Normal and* H_1 : *Weibull*

Let $f_{ln}(x; \mu, \sigma)$ be the log-normal density given by 1.1 and let $f_W(x; b, c)$ be the Weibull density given by 1.2. The maximum likelihood estimates of b and c must be obtained by iterative methods. However, this is a simple task, requiring the solution of one nonlinear equation which always has a unique real root. Table 1 gives critical values for the test statistic, $(\mathrm{RML})_c^{1/n}$, and the power of the test for this problem. For this problem it is easily seen that

$$(\mathrm{RML})^{1/n} = (2\pi e\hat{\sigma}^2)^{\frac{1}{2}}\left[\prod_i^n x_i f_W(x_i ; \hat{b}, \hat{c})\right]^{1/n}$$

where

$$\hat{\sigma}^2 = \sum (\ln x_i - \hat{\mu})^2/n \quad \text{and} \quad \hat{\mu} = \frac{1}{n}\sum \ln x_i ,$$

and we reject the log-normal in favor of the Weibull distribution whenever $(\mathrm{RML})^{1/n} \geq (\mathrm{RML})_c^{1/n}$.

The values in Table 1 were the result of simulations in which 15,000 samples were used for each sample size.

Table 1

Critical Values of $(\mathrm{RML})^{1/n}$ *and Power of the Test for* H_0: *Log-normal and* H_1: *Weibull*

n	$\alpha = .20$		$\alpha = .10$		$\alpha = .05$		$\alpha = .01$	
	$(\mathrm{RML})_c^{1/n}$	Power	$(\mathrm{RML})_c^{1/n}$	Power	$(\mathrm{RML})_c^{1/n}$	Power	$(\mathrm{RML})_c^{1/n}$	Power
20	1.015	.75	1.038	.61	1.082	.48	1.144	.22
30	.993	.86	1.020	.75	1.044	.63	1.095	.39
40	.984	.93	1.007	.85	1.028	.76	1.070	.53
50	.976	.96	.998	.91	1.014	.83	1.054	.63

To illustrate the use of the RML test for this problem, suppose the following observations, as given by Lieblein and Zelen (1956) for the lifetime in millions of revolutions of 23 ball bearings, are used to test H_0 : Log-normal vs. H_1 : Weibull;

17.88	28.92	33.00	41.52	42.12	45.60
48.48	51.84	51.96	54.12	55.56	67.80
68.64	68.64	68.88	84.12	93.12	98.64
105.12	105.84	127.92	128.04	173.40	

For these data we find for the maximum likelihood estimates of the parameters in the Weibull model, $\hat{b} = 81.88$ and $\hat{c} = 2.102$. (See Harter and Moore (1965, 1967), Thoman, Bain and Antle (1969) for information on maximum likelihood estimation of the parameters in the Weibull model.) We also find $\hat{\sigma} = .522$ which results in a value of .976 for $(\mathrm{RML})^{1/n}$. Consequently we see from Table 1 that we cannot reject the log-normal model in favor of the Weibull.

As we are considering the choice of a model as a test of hypothesis, it is important to allow either of the models to be the null hypothesis. We suppose the experimenter would assign to the null hypothesis the model he prefers to use, unless there is convincing evidence that he should use the other. In order to allow him this choice, we next present a table of critical values with the Weibull as the null hypothesis.

2.2 Critical Values and Power of the RML *Test with H_0 : Weibull and H_1 : Log-normal*

If the Weibull model is to be the favored model, then the test statistic becomes

$$(\mathrm{RML})^{1/n} = (2\pi e\sigma^2)^{-\frac{1}{2}}\left[\prod_{i}^{n} x_i f_W(x_i\ ;\ \hat{b}, \hat{c})\right]^{-1/n}$$

and we reject the Weibull in favor of the log-normal whenever $(\mathrm{RML})^{1/n} \geq (\mathrm{RML})_c^{1/n}$. The critical values and powers for this test are given in Table 2.

The values in Table 2 were the result of simulations in which 15,000 samples were used for each sample size.

With the ball bearing data given in Section 2.1 and considering the Weibull as the null hypothesis, we find $(\mathrm{RML})^{1/n} = 1.025$. From Table 2 it is clear that we could not reject the Weibull model in favor of the log-normal at the .10 level, but it does appear that the log-normal may deserve some consideration for such data. We note, however, from the powers in either Table 1 or Table 2 that the ability to select between the Weibull and log-normal with sample size 20 is not good.

TABLE 2

Critical Values of $(\mathrm{RML})^{1/n}$ *and Power of the Test for* H_0*: Weibull and* H_1*: Log-normal*

n	$\alpha = .20$		$\alpha = .10$		$\alpha = .05$		$\alpha = .01$	
	$(\mathrm{RML})_c^{1/n}$	Power	$(\mathrm{RML})_c^{1/n}$	Power	$(\mathrm{RML})_c^{1/n}$	Power	$(\mathrm{RML})_c^{1/n}$	Power
20	1.008	.73	1.041	.57	1.067	.43	1.120	.20
30	.991	.84	1.019	.74	1.041	.62	1.088	.34
40	.980	.93	1.005	.85	1.026	.75	1.063	.51
50	.974	.96	.995	.91	1.016	.82	1.045	.66

3. An Example for Testing Normality vs. Type 1 Extreme Value for Maximums

I am indebted to Professor Byron Reich of our Civil Engineering Department for the data in this example. The data are the maximum flood level (in millions of cubic feet per second) for the Susquehanna River at Harrisburg, Pennsylvania. Each number is the maximum flood level for a four year period, the first, .654, being for the period 1890–1893, and the last, .265, being for the period 1966–1969. The data are

.654	.613	.315	.449	.297	.402	.379	.423	.379	.3235
.269	.740	.418	.412	.494	.416	.338	.392	.484	.265

With these data the maximum likelihood estimates for the parameters in the Type 1 extreme value for maximums distribution are $\hat{A} = .3688$, $\hat{B} = .089775$. The maximum likelihood estimate for σ in the normal model is .1221. To test H_0 : normal vs. H_1 : Type 1 extreme value for maximums we find $(\text{RML})^{1/n}$ to be 1.131. From Table 1 we see that the critical value for this test at the $\alpha = .05$ level is 1.082 while at the $\alpha = .01$ level it is 1.14. Hence normality may be rejected in favor of the Type 1 extreme value for maximums distribution at something between the .01 and the .05 level.

With these data and assuming a Type 1 extreme value for maximums distribution, one might have been interested in an upper .95 confidence limit on the maximum flood in 100 years, beginning in 1970. This is easily found to be .949 (See Antle and Rademaker (1972)). This appears to be quite extreme; however, something worse happened in Harrisburg in the summer of 1972. The flood level reached .960 on June 24, 1972.

A program for the calculation of $(\text{RML})^{1/n}$ is available at no cost from the authors.

References

[1] Antle, Charles E. and Rademaker, Fred, (1972). "An upper confidence limit on the maximum of m future observations from a type 1 extreme value distribution," *Biometrika*, *59*, 475–477.

[2] Dumonceaux, Robert, Antle, Charles E., and Haas, Gerald, (1973). "Likelihood ratio test for discrimination between two models with unknown location and scale parameters," *Technometrics*, *15*, 19–27.

[3] Harter, H. L. and Moore, A. H., (1965). "Maximum likelihood estimation of the parameters of gamma and Weibull populations from complete and censored samples," *Technometrics*, *7*, 639–643.

[4] Harter, H. L. and Moore, A. H., (1967). "A note on estimation from a type 1 extreme-value distribution," *Technometrics*, *9*, 325–331.

[5] Lieblein, J. and Zelen, M., (1956). "Statistical investigation of the fatigue life of deep-groove ball bearing," *J. Research National Bureau of Standards*, *57*, 273–316.

[6] Thoman, D. R., Bain, L. J., and Antle, C. E., (1969). "Inferences on the parameters of the Weibull distribution," *Technometrics*, *11*, 445–460.