01 Jan 1992

Prediction Intervals, Based On Ranges And Waiting Times, For An Exponential Distribution

Laura Colangelo

Jagdish K. Patel
Missouri University of Science and Technology

Prediction Intervals, Based on Ranges and Waiting Times, for an Exponential Distribution

Laura Colangelo
University of Missouri, Rolla
Jagdish K. Patel
University of Missouri, Rolla

Key Words — **Life test, Range, Waiting time, Exponential distribution**

Reader Aids —
Purpose: Widen state of art
Special math needed for explanations: Statistics
Special math needed to use results: Statistics
Results useful to: Practicing statisticians

Abstract — **This article contains two prediction intervals applicable to a 2-parameter as well as a 1-parameter exponential distribution. One can be used to predict a future sample range on the basis of an observed sample range. Appropriate prediction factors are tabulated. The other can be used to predict a waiting time between two future successive failures on the basis of an observed waiting time between two previous successive failures.**

1. INTRODUCTION

In life testing and reliability studies prediction intervals, which use the results of a past sample, provide useful information about the realization of a random variable in a future sample from the same distribution. For example, one might wish to predict the number of product failures during some future time period using past data. A recent comprehensive review article by Patel [5] describes the availability of a large variety of prediction intervals for several life distributions.

The exponential distribution is a common probability model. One of the first papers on prediction intervals based on this model was by Lawless [4]. Since then, this model has been investigated extensively, and several prediction intervals covering different situations are available for it.

This paper adds two more prediction intervals that apply to a 2-parameter as well as a 1-parameter exponential distribution.

Let $R_0(m)$ be the sample range of the life times when m items are put on a life test (without replacement). Similarly, let $R_f(n)$ be the future s-independent sample range of the life times when n items are put on a similar life test. Then the first prediction interval that we have obtained could be used to predict $R_f(n)$ on the basis of the observed $R_0(m)$.

The prediction intervals are generally wider than s-confidence intervals because there are two sources of variation in their construction: 1) the presence of unknown population parameter(s) which may require some estimation, and 2) the variation due to the future sample [5, p 2398].

The exponential model is "conservative" because of its "no-memory" property. As a consequence, even its s-confidence intervals are wider unless the sample sizes are extremely large. The prediction interval for the duration of a life test [4] is quite wide as seen from the prediction factors computed for it by Bain & Patel [1].

Range is a quick rough measure of the variability of the sample data. Hahn [3] has obtained prediction intervals to contain future ranges from a s-normal population. The knowledge of the value of a sample range provides good information about the quality of the experimental units involved in the study. As pointed out [3], such prediction intervals could be used in the following situation. Suppose a random sample of size n is to be taken from a future shipment of a product and the shipment is accepted only if the range of the sample is below a specified limit. Using the information from m past units, such a limit is a 1-sided upper prediction limit. Acceptance of the shipment using this procedure would help maintain the quality of the product. Prediction intervals to contain future ranges for populations other than s-normal do not seem to be available at present. Section 2 covers the prediction interval for future ranges based on the exponential model. Table 1 provides the prediction factors to compute the prediction limits.

Let $W(i)$ be the waiting time between failures $(i-1)$ and i when m items are put on a life test (without replacement), $(i = 1,2,\ldots,m)$. We have obtained a prediction interval which can be used to predict the future waiting time $W(s)$ on the basis of the observed waiting time $W(r)$, $(1 \le r < s \le m)$. This information could be useful in planning a maintenance program. The necessary prediction factors can be obtained from known tables. Section 3 covers this result.

2. MODEL

The Cdf of the lifetimes of all items is a 2-parameter exponential:

$\text{expf}[(u-\beta)/\theta]$, for $u \ge \beta$.

If β is known, the distribution becomes, effectively, the 1-parameter exponential.

Notation

θ scale parameter, unknown
β location parameter, known or unknown
$X_{(i)}, Y_{(j)}$ ordered failure times from two s-independent samples $(i=1,2,\ldots m;\ j=1,2,\ldots,n)$.

TABLE 1
Lower Quantiles $k(\gamma;m,n)$

		γ									
n	m	.005	.01	.025	.05	.10	.90	.95	.975	.99	.995
5	5	.112	.141	.195	.256	.349	2.866	3.899	5.125	7.106	8.936
	10	.091	.113	.155	.201	.269	1.710	2.185	2.699	3.445	4.065
	15	.082	.102	.139	.179	.238	1.394	1.751	2.127	2.658	3.087
	20	.077	.095	.129	.166	.220	1.236	1.540	1.856	2.295	2.646
	25	.073	.090	.123	.158	.208	1.138	1.410	1.691	2.080	2.387
10	5	.246	.290	.371	.458	.585	3.717	4.966	6.441	8.817	11.009
	10	.208	.243	.304	.369	.460	2.172	2.711	3.287	4.119	4.808
	15	.191	.222	.276	.332	.411	1.757	2.151	2.563	3.139	3.602
	20	.181	.209	.259	.310	.382	1.552	1.882	2.223	2.693	3.067
	25	.173	.200	.247	.295	.362	1.424	1.718	2.018	2.431	2.756
15	5	.324	.376	.470	.571	.717	4.203	5.578	7.198	9.806	12.210
	10	.278	.319	.390	.465	.569	2.435	3.010	3.624	4.506	5.235
	15	.257	.293	.356	.420	.509	1.963	2.378	2.810	3.412	3.895
	20	.244	.277	.335	.394	.474	1.730	2.075	2.430	2.919	3.306
	25	.235	.266	.321	.376	.450	1.585	1.891	2.203	2.629	2.963
20	5	.378	.436	.539	.649	.809	4.544	6.010	7.734	10.508	13.063
	10	.326	.371	.450	.531	.645	2.621	3.222	3.862	4.780	5.538
	15	.303	.343	.412	.482	.578	2.108	2.539	2.985	3.606	4.103
	20	.288	.325	.388	.452	.539	1.855	2.211	2.576	3.078	3.474
	25	.277	.312	.372	.432	.512	1.699	2.013	2.332	2.768	3.110
25	5	.419	.481	.591	.709	.879	4.809	6.344	8.150	11.053	13.726
	10	.363	.411	.495	.582	.702	2.764	3.386	4.047	4.995	5.773
	15	.337	.380	.454	.529	.631	2.220	2.663	3.120	3.756	4.264
	20	.322	.361	.429	.497	.589	1.952	2.317	2.690	3.201	3.604
	25	.310	.348	.411	.474	.560	1.786	2.107	2.433	2.875	3.222

$R_0(m), R_f(n)$ sample ranges: $R_0(m) = X_{(m)} - X_{(1)}$, $R_f(n) = Y_{(n)} - Y_{(1)}$.

$W(i)$ waiting time between failures $(i-1)$ and i: $W(i) = X_{(i)} - X_{(i-1)}$, $(i=1,2,\ldots,m)$

$X^*_{(i)}, Y^*_{(j)}$ ordered standardized r.v.'s from two s-independent samples from expf($\cdot$): $X^*_{(i)} = (X_{(i)} - \beta)/\theta$, $Y^*_{(j)} = (Y_{(j)} - \beta)/\theta$, $(i=1,2,\ldots m;\ j=1,2,\ldots,n)$

$R^*_0(m), R^*_f(n)$ standardized sample ranges: $R^*_0(m) = X^*_{(m)} - X^*_{(1)}$, $R^*_f(n) = Y^*_{(n)} - Y^*_{(1)}$.

$W^*(i)$ standardized waiting time: $W^*(i) = X^*_{(i)} - X^*_{(i-1)}$, $(i=1,2,\ldots,m)$

$1-\alpha$ prediction probability

Other, standard notation is given in "Information for Readers & Authors" at the rear of each issue.

3. A PREDICTION INTERVAL BASED ON RANGES

Consider the ratio of two s-independent ranges:

$$K = R_f(n)/R_0(m) = R^*_f(n)/R^*_0(m).$$

The probability distribution of the r.v. K, which we obtained below, does not depend on any parameter(s). The Cdf $H_f(y)$ of the r.v. $R^*_f(n)$ is known [2, p 12]:

$$H_f(y) = \int_0^\infty n(e^{-u} - e^{-(y+u)})^{n-1} e^{-u} du$$

$$= \int_0^\infty n(1-e^{-y})^{n-1} e^{-nu}\, du$$

$$= (1-e^{-y})^{n-1}, \text{ for } y \geq 0.$$

Similarly, the Cdf $H_0(x)$ of the rv $R^*_0(m)$ is:

$$H_0(x) = (1-e^{-x})^{m-1}, \text{ for } x \geq 0.$$

Now, consider:

$$\Pr\{K \leq k\} = \Pr\{R^*_f(n) \leq k\, R^*_0(m)\}$$

$$= \int_0^\infty H_f(kx) dH_0(x) dx$$

$$= \int_0^\infty (m-1)e^{-x}(1-e^{-x})^{m-2}(1-e^{-kx})^{n-1} dx \tag{1}$$

$$= (m-1) \int_0^\infty \sum_{i=0}^{m-2} \sum_{j=0}^{n-1} \binom{m-2}{i} \binom{n-1}{j} (-1)^{i+j} \cdot e^{-x(1+i+kj)}\, dx$$

$$= (m-1) \sum_{i=0}^{m-2} \sum_{j=0}^{n-1} \binom{m-2}{i} \binom{n-1}{j} (-1)^{i+j} \cdot \frac{1}{(1+i+kj)}, \quad (2)$$

by using the binomial theorem to expand:

$$(1-e^{-x})^{m-2}(1-e^{-kx})^{n-1}$$

Let $k_\gamma \equiv k(\gamma;m,n)$ be the γ lower quantile for the Cdf$\{K\}$. Then

$$\Pr\left\{k_{\alpha_1} \le \frac{R_f(n)}{R_0(m)} \le k_{1-\alpha_2}\right\} = 1-\alpha$$

with $0 < \alpha_1 < 1$, $0 < \alpha_2 < 1$ and $\alpha_1 + \alpha_2 = \alpha$. The $(1-\alpha)$ 2-sided prediction interval of the future sample range $R_f(n)$ on the basis of the past sample range $R_0(m)$ is:

$$[k_{\alpha_1} R_0(m),\ k_{1-\alpha_2} R_0(m)]. \quad (3)$$

Similarly, $(1-\alpha)$ 1-sided lower and upper prediction limits for $R_f(n)$ are:

$$[k_\alpha R_0(m),\ \infty), \quad (4)$$

$$[0, k_{1-\alpha} R_0(m)] \quad (5)$$

We have computed prediction factors k_γ for $m=5(5)25$, $n=5(5)25$, $\gamma=0.005$, 0.01, 0.025, 0.05, 0.10, 0.90, 0.95, 0.975, 0.99, 0.995. These values are in table 1, and were computed with the Macintosh version 1.2.2 of 'Mathematica'. A numerical root-finding routine, "Find Root", was used. Table 1 gives these factors accurate to 3 decimal places. Table 2 illustrates the effect of the increase in sample sizes m and n on the ratio of the limits using (3).

Example 1

Consider the following failure times (in weeks) from an accelerated life test of 10 transistors having a 2-parameter exponential life distribution: 7,9,9,10,13,14,16,17,19,25. We want to predict the range of the failure times of a future such test of 15 transistors using a 90% prediction interval. Then, for $m = 10$, $R_0(10) = 25-7 = 18$, $n = 15$, and the equal-tail case of $\alpha_1 = \alpha_2 = 0.05$, we find $k_{0.05} = 0.464797$ and $k_{0.95} = 3.01029$ from table 1. This provides a 90% 2-sided prediction interval for $R_f(15)$ as:

TABLE 2
Ratio of the Upper Prediction Limit to the Lower Prediction Limit [with selected m, n, α]

n	m	$(k_{.95})/(k_{.05})$	$(k_{.975})/(k_{.025})$
5	5	15.205	26.264
	10	10.852	17.383
	15	9.767	15.312
	20	9.253	14.351
	25	8.945	13.784
10	5	10.852	17.383
	10	7.348	10.807
	15	6.477	9.287
	20	6.063	8.584
	25	5.816	8.168
15	5	9.767	15.312
	10	6.477	9.287
	15	5.657	7.896
	20	5.268	7.253
	25	5.035	6.873
20	5	9.253	14.332
	10	6.068	8.584
	15	5.268	7.253
	20	4.891	6.638
	25	4.664	6.274
25	5	8.945	13.784
	10	5.816	8.168
	15	5.035	6.873
	20	4.664	6.274
	25	4.442	5.920

$$[0.464797\cdot 18,\ 3.01029\cdot 18] = [8.37,\ 54.19].$$

Similarly, 90% 1-sided lower and upper prediction intervals are:

$$[0.569004\cdot 18,\ \infty) = [10.24,\ \infty)$$

$$[0, 2.43540\cdot 18] = [0, 43.84].$$

Side Remark

Table 1 can be used to obtain a certain type of *s*-confidence interval. Suppose X's have unknown parameters (θ_1,β), and Y's have unknown parameters (θ_2,β).

$$K = \frac{R_f^*(n)}{R_0^*(m)} = \frac{R_f(n)/\theta_2}{R_0(m)/\theta_1},$$

$$\Pr\{k \le K_\gamma\} = \Pr\left\{\frac{\theta_1}{\theta_2} \le \frac{R_0(m)}{R_f(n)} K_\gamma\right\} = \gamma,$$

giving a γ upper *s*-confidence limit for θ_1/θ_2. Similarly a lower limit can be obtained. This *s*-confidence interval is useful where extreme observations are available but the complete samples are not.

4. PREDICTION INTERVAL BASED ON WAITING TIMES

We consider the ratio,

$$F = \frac{(m-s+1)W(s)}{(m-r+1)W(r)} = \frac{2(m-s+1)W^*(s)}{2(m-r+1)W^*(r)},$$

$$(1 \le r < s \le n).$$

The $(m-s+1)W^*(s)$ is s-independent of $(m-r+1)W^*(r)$, and both have the 2-parameter exponential distribution [2, p 153]. The r.v. F has an F-distribution with 2 degrees of freedom in both numerator and denominator. Let $f_\gamma = f_{(\gamma;2,2)}$ be its γ lower quantile: $\Pr\{F \le f_\gamma\} = \gamma$. Since

$$\Pr\left\{f_{\alpha_1} \le \frac{(m-s+1)W(s)}{(m-r+1)W(r)} \le f_{1-\alpha_2}\right\} = 1-\alpha$$

with $\alpha_1 + \alpha_2 = \alpha$ defined as in section 3, we have a $(1-\alpha)$ 2-sided prediction interval of a future waiting time $W(s)$ on the basis of the past waiting time $W(r)$:

$$[cW(r)f_{\alpha_1},\ cW(r)f_{1-\alpha_2}] \quad (6)$$

$$c \equiv (m-r+1)/(m-s+1).$$

Similarly $(1-\alpha)$ 1-sided lower and upper prediction limits for $W(s)$ are:

$$[cW(r)f_\alpha, \infty) \quad (7)$$

$$[0,\ cW(r)f_{1-\alpha}] \quad (8)$$

Since $f_\gamma = f_{(\gamma;2,2)}$ can be found from the F-distribution tables, prediction factors $c \cdot f_{(\gamma;2,2)}$ do not require computing any new tables.

Since the distribution of the $rv\,(m-i+1)W(i)$ is exponential, it has the "no memory" property. Because of this, our prediction interval is based only on the present information given by $W(r)$ to predict the future value of $W(s)$.

Example 2

Let 15 transistors be put on an accelerated life test (without replacement) and let failure times have a 2-parameter exponential distribution. Let the failure times for transistors #4 & #5 be 10 and 13 weeks, respectively. Using this information, we want to find a 90% prediction interval for the waiting time between, future failures #9 & #10. Here $m=15$, $r=5$, $W(5)=13-10=3$, $s=10$; and $f_{0.90}=f(0.90;2,2)=9.0$. This provides a 90% 1-sided upper prediction limit of:

$$\frac{15-5+1}{15-10+1} \cdot 3 \cdot 9 = 49.5.$$

REFERENCES

[1] P. T. Bain, J. K. Patel, "Factors for calculating prediction intervals for samples from a one-parameter exponential distribution", *J. Quality Technology*, vol 23, 1991 Jan, pp 48-52.

[2] H. A. David, *Order Statistics*, (2^{nd} ed), 1981; John Wiley & Sons.

[3] G. J. Hahn, "Simultaneous prediction intervals to contain the standard deviations or ranges of future samples from a normal distribution", *J. American Statistical Assoc*, vol 67, 1972 Dec, pp 938-942.

[4] J. F. Lawless, "A prediction problem concerning samples from the exponential distribution, with applications in life testing", *Technometrics*, vol 13, 1971 Nov, pp 725-730.

[5] J. K. Patel, "Prediction intervals - A review", *Communications in Statistics - Theory & Methods*, vol 18, 1989, pp 2393-2465.

AUTHORS

Laura A. Colangelo; Department of Mathematics and Statistics; University of Missouri-Rolla; Rolla, Missouri 65401 USA.

Laura Colangelo received her BS in Mathematics from the University of Idaho in 1985 and her MS in Statistics from the University of Missouri-Rolla in 1990.

Dr. Jagdish K. Patel; Department of Mathematics and Statistics; University of Missouri-Rolla; Rolla, Missouri 65401 USA.

Jagdish K. Patel is a Professor of Statistics. He got his PhD from University of Minnesota in 1968. His research interests are in ranking & selection problems and statistical reliability. He is a co-author of *Handbook of Statistical Distributions* (1976) and *Handbook of the Normal Distribution* (1982).

Manuscript TR90-232 received 1990 December 15; revised 1991 June 11.

IEEE Log Number 06596 ◄TR►