Taylor & Francis
Taylor & Francis Group

Minimum Kolmogorov–Smirnov test statistic parameter estimates

MICHAEL D. WEBER, LAWRENCE M. LEEMIS* and REX K. KINCAID

Department of Mathematics, The College of William and Mary, Williamsburg, VA 23187-8795, USA

We present and implement an algorithm for computing the parameter estimates in a univariate probability model for a continuous random variable that minimizes the Kolmogorov–Smirnov test statistic. The algorithm uses an evolutionary optimization technique to solve for the estimates. Several simulation experiments demonstrate the effectiveness of this approach.

Keywords: Algorithm; Evolutionary algorithm; Continuous probability distributions; Goodness-of-fit tests; Input modeling

1.   Introduction

When a reliability analysis, survival analysis, or discrete-event simulation input modeling analysis is performed, it is typically the case that a modeler collects data associated with some system under consideration. Although this raw data provides an initial insight into the stochastic elements of the system, often it is desirable to construct a model of the system which is tractable for mathematical analysis, while remaining consistent with the data. Typically, some common parametric model is chosen, such as the exponential distribution, then a point estimation technique, such as maximum likelihood, is used to determine reasonable parameter estimates for the model. Finally, some goodness-of-fit test is performed to validate the model.

However, this approach to modeling has two drawbacks. First, on the basis of experience and an initial examination of the raw data, the modeler chooses which model to use *a priori*, and then proceeds to choose reasonable parameters on the basis of data (for example $\hat{\lambda} = 3.2$, if an exponential (λ) model is chosen). The model is then assessed via a goodness-of-fit test. An obvious difficulty is that some other parametric model may have yielded an even better fit.

A second and related problem has to do with the computation involved in finding parameter estimates. Analytical techniques for optimizing distribution parameters involve calculus. Likelihood functions for distributions with multiple parameters can be cumbersome to analyze. Differentiating them and solving for the estimators requires a symbolic language or a well designed algorithm to ensure correctness and efficiency. Further, as stated above, we know

---

*Corresponding author. Email leemis@math.wm.edu

that the modeler must repeat this process for several common models so that the best model is not overlooked. Thus, finding a good model, along with good parameters, is a tedious task.

Popular parameter estimation techniques include maximum likelihood estimation (MLE), method of moments (MOM), and least squares (LS). Often statistical properties of estimators drive the choice of the parameter estimation technique selected. The exponential distribution typically relies on MLE in modeling lifetime data, for example, because the point estimates are closed form and the interval estimates are exact for complete and Type II censored data sets. There are occasions, however, when only fit considerations drive the choice of the estimation technique, as might be the case in the selection of an input model for a discrete-event simulation.

The purpose of this article is to consider an estimation technique which we refer to as minimum Kolmogorov–Smirnov estimation (MKSE), whose purpose is to minimize the value of the KS statistic. We use a heuristic optimization algorithm developed by Sobieszczanski–Sobieski *et al.* [1] to estimate the parameters.

In spite of the benefits to MKSE, there are three significant drawbacks to our approach:

- The method provides point estimators, but no interval estimators. Thus, MKSE is more closely aligned with the performance of MOM estimators, rather than MLE estimators, where interval estimators can be constructed based on the likelihood ratio statistic.
- The method requires a number of arbitrary parameters to be prescribed in order for the algorithm to converge to the MKSE. These parameters are described in section 3.
- As with all estimation procedures, sampling variability makes identifying the correct population distribution difficult, as shown in section 4.

The software tool MKSFitter described in this article is capable of finding these estimates. It is a tool that fits common models to data sets using a general-purpose Bell-Curve Based evolutionary algorithm (BCB). The tool maintains a collection of common models that may be easily extended by a user familiar with the C programming language. It takes as input an arbitrary data set, and outputs the model which best fits the data, along with the parameters that minimize the KS test statistic. Scott [2] contains related work.

Let $x_1, x_2, \ldots, x_n$ denote an independent and identically distributed sample drawn from a population with unknown parameters $\theta_1, \theta_2, \ldots, \theta_k$. The Kolmogorov–Smirnov (KS) test statistic D_n, is defined by

$$D_n = \sup_x |\hat{F}(x) - F_n(x)|,$$

where n is the sample size, $\hat{F}(x)$ is a fitted CDF, and $F_n(x)$ is the empirical CDF, a step-function that increases by $1/n$ at each data value. The KS statistic has been used for goodness-of-fit testing for continuous populations for decades, although other tests have made slight improvements in terms of power. The KS test's appeal includes the straightforward computation of the test statistic and the distribution-free characteristic of D_n. Its drawback is that its cumulative distribution function of D_n under the null hypothesis (*i.e.*, the data was drawn from a population with CDF $F(x)$) is difficult to determine, leaving one to calculate critical values with various approximation methods. An algorithm for computing the distribution of D_n for small to moderate values of n is given by Drew *et al.* [3]. As the supremum must be achieved at a data value, the computational formulas for computing D_n are:

$$D_n^+ = \max_{i=1,2,\ldots,n} \left(\frac{i}{n} - \hat{F}(x_{(i)}) \right), \quad D_n^- = \max_{i=1,2,\ldots,n} \left(\hat{F}(x_{(i)}) - \frac{i-1}{n} \right), \quad \text{and}$$

$$D_n = \max\{D_n^+, D_n^-\},$$

where $x_{(1)}, x_{(2)}, \ldots, x_{(n)}$ are the order statistics. The maximum positive difference, D_n^+, detects the largest vertical deviation between the two CDFs where the fitted CDF is below the empirical

CDF. Likewise, the maximum negative difference, D_n^-, detects the largest vertical deviation between the two CDFs where the fitted CDF is above the empirical CDF. The smallest value of D_n that can be achieved is $1/2n$, which corresponds to the CDF of the fitted distribution bisecting all of the risers of the steps associated with the empirical CDF.

We begin with three examples to motivate the technique. The first two involve fitting data sets to the exponential distribution with a mean of θ. When $x_1, x_2, \ldots, x_n$ denote an independent and identically distributed sample drawn from an exponential population with mean θ, Hogg and Craig [4] give the unbiased MLE as $\hat{\theta} = \bar{x}$. This estimator will be compared with the MKSE.

Example 1: Consider the case of sampling a single ($n = 1$) value, x_1, from an exponential population with unknown mean θ. In this case, the MLE is simply $\hat{\theta} = x_1$. Thus, the fitted CDF is

$$\hat{F}(x) = 1 - e^{-(x/x_1)} \quad x > 0.$$

This fitted CDF strikes the riser associated with x_1 of the empirical CDF at $1 - e^{-(x_1/x_1)} = 1 - 1/e \cong 0.6321$, regardless of the value of x_1. Thus, the distribution of the KS statistic when the MLE is used is degenerate at $1 - 1/e$. The MLE point estimator $\hat{\theta} = x_1$, although unbiased, does not minimize the KS test statistic. In order to find the MKSE which will bisect the riser, it is necessary to solve

$$\frac{1}{2} = 1 - e^{-(x_1/\hat{\theta})}$$

for $\hat{\theta}$, yielding $\hat{\theta} = x_1/\log 2$. Thus, the distribution of the KS statistic when the MKSE is used is degenerate at $1/2$. A plot of the MLE and MKSE fitted distributions associated with a single data value of $x_1 = 3$ is shown in figure 1. Figure 2 is the associated plot of D_1 *vs.* θ with the MLE and MKSE labeled.
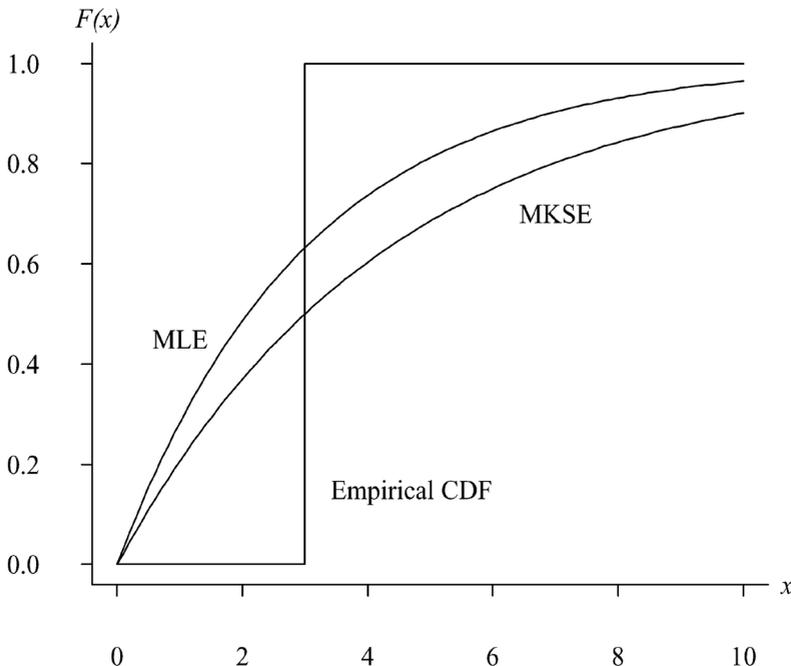

Figure 1. The empirical CDF, the fitted exponential MLE CDF, and the fitted exponential MKSE CDF for a single data value $x_1 = 3$.
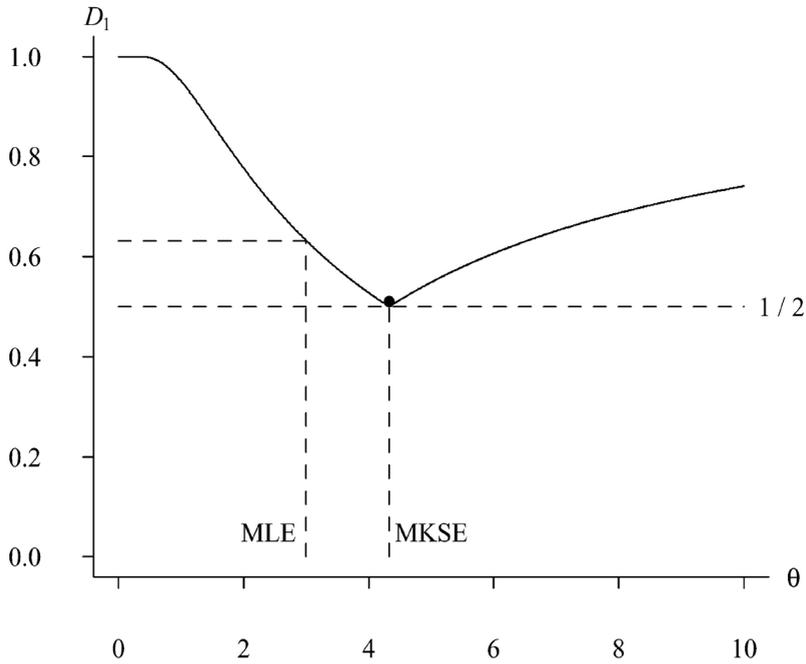
Figure 2.   The value of the KS test statistic for various values of θ associated with an exponential model and a single data value $x_1 = 3$.

The theoretical minimum that the KS statistic can assume, $1/2n = 1/2$ is shown by a horizontal dashed line. In this case, the KS statistic is able to achieve this theoretical minimum using the MKSE.

The next example illustrates the difference between MLE and MKSE on a non-trivial sample size where the MKSE is not able to achieve the theoretical minimum. This is almost always the case in practice.

Example 2:   Consider the oft-studied dataset of $n = 23$ ball bearing failure times in millions of revolutions [5]:

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

The exponential distribution can also be fitted to this data set, although the fit is poor because the ball bearings are wearing out. Figure 3 contains a plot of the empirical CDF, along with the fitted CDFs for the MLE and MKSE estimators. The MLE is $\hat{\theta} = \bar{x} = 1661.16/23 \cong 72.22$. The value of D_{23} associated with the MLE is shown with a vertical dashed line just to the left of $x_{(4)} = 41.52$. The MKSE is $\hat{\theta} \cong 96.10$. In this case (since the exponential distribution has a single parameter), the MKSE can be determined by solving the equation

$$1 - e^{-41.52/\theta} - \frac{3}{23} = \frac{22}{23} - (1 - e^{-128.04/\theta})$$
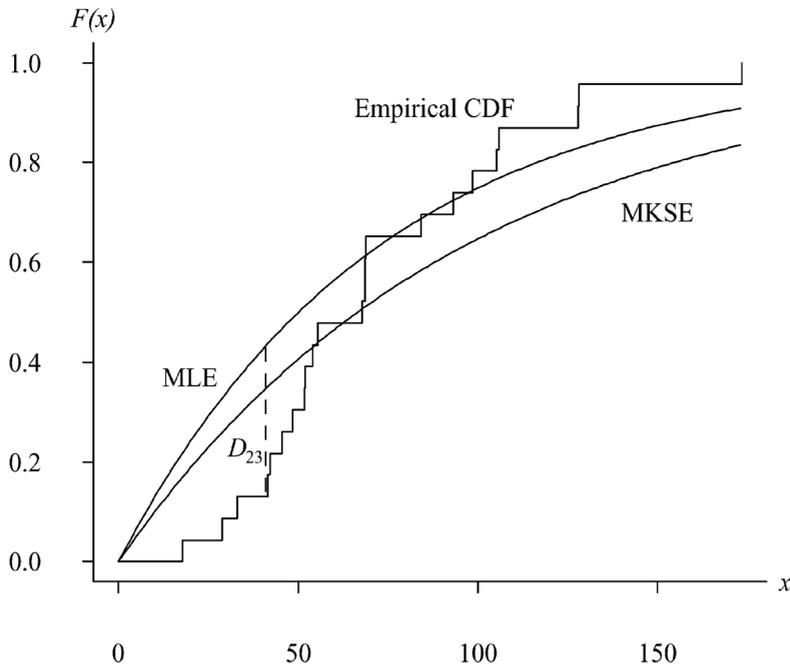
Figure 3. The empirical CDF, the fitted exponential MLE CDF, and the fitted exponential MKSE CDF for the ball bearing data set.

for θ, which corresponds to the point where the KS test statistics switches from occurring just to the left of $x_{(4)} = 41.52$ (detected by D_{23}^+) to occurring just to the right of $x_{(22)} = 128.04$ (detected by D_{23}^-). Figure 4 shows the value of D_{23} for various values of θ. The indices of the order statistics where the maximum vertical difference between the CDFs is achieved are shown above the curve, along with a $+$ or $-$ indicating whether the maximum difference was detected by D_{23}^+ or D_{23}^-. The dots that are placed along the curve denote the points where the order statistic associated with D_{23} switches from one data value to another. For the ball bearing data, the order statistics where the maximum differences occur are $x_{(1)} \rightarrow x_{(2)} \rightarrow x_{(4)} \rightarrow x_{(22)} \rightarrow x_{(23)}$. Note that the switch from $x_{(22)}$ to $x_{(23)}$ occurs off the horizontal scale of figure 4 at $\theta \cong 879$. (All data sets begin this curve with $1-$ for small θ and end with $n+$ for large θ.) The theoretical minimum, $1/2n = 1/46$ is indicated by a dashed horizontal line. The KS test statistic value is ~ 0.3068 at the MLE and is ~ 0.2204 at the MKSE.

The third and final example considers a two-parameter model.

Example 3: Consider fitting the ball bearing failures to the Weibull distribution with CDF

$$F(x) = 1 - e^{-(\lambda x)^{\kappa}} \quad x > 0,$$

where λ and κ are positive parameters. Using standard techniques [6, p. 217], the MLEs are $\hat{\lambda} \cong 0.0122, \hat{\kappa} \cong 2.10$. Using software to be described later in this article, the MKSEs are $\hat{\lambda} \cong 0.0132, \hat{\kappa} \cong 2.23$. The MLEs correspond to a KS statistic of $D_{23} \cong 0.151$ and the MKSEs correspond to a KS statistic of $D_{23} \cong 0.099$. A plot of the CDFs of these two fitted distributions is given in figure 5.
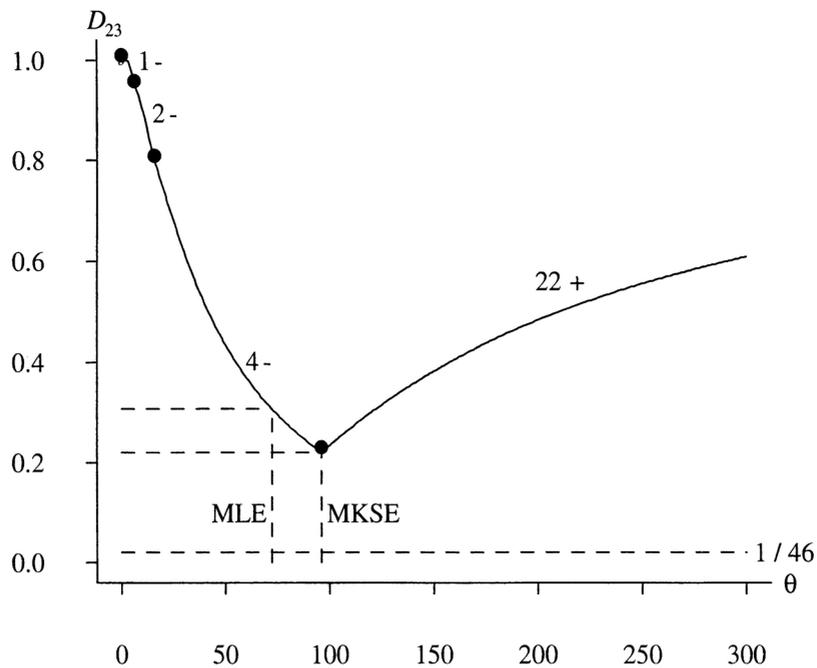
Figure 4.   The value of the KS test statistic for various values of θ associated with an exponential model and the ball bearing data set.
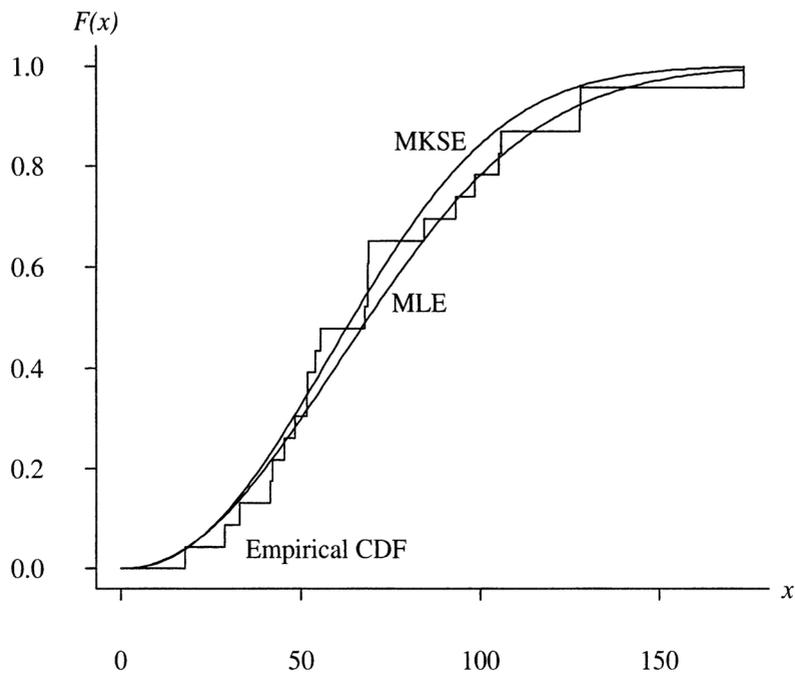

Figure 5.   The empirical CDF, the fitted Weibull MLE CDF, and the fitted Weibull MKSE CDF for the ball bearing data set.

2. Literature review

The literature available on the KS statistic is extensive. Stephens' article [7, chapter 4] contains comprehensive coverage on the use of the KS statistic, as well as other statistics based on the empirical distribution function. To the best of our knowledge, there is no literature on computational experiments in which the KS test statistic is minimized so as to compute the parameter estimates in a univariate probability model for a continuous random variable. Gyorfi *et al.* [8, 9], however, provide a theoretical treatment of minimum Kolmogorov distance estimates for both univariate [8] and multivariate [9] distributions. They give conditions for the underlying metric space under which parameter estimates are consistent. In particular, they show that all location and scale models, with parent distributions different from Dirac, and all standard exponential models satisfy these conditions. If inconsistent estimators arise, Donoho and Liu [10] indicate that consistency can be achieved using the Cramer–VonMises statistic rather than the KS statistic, which is a relatively minor modification to the methods presented in this article.

3. Problem formulation

The underlying algorithm of the MKSFitter is a general-purpose optimizer named BCB [1]. It is a population-based heuristic, similar in spirit to evolutionary strategies (ESs) and evolutionary programs (EPs) but has fewer parameters to adjust. See Back [11] for detailed information concerning ESs and EPs. A new generation in BCB is selected exactly the same as a $(\mu + \lambda) -$ ES with $\lambda = \mu$. That is, the best μ individuals out of μ parents plus λ children are selected for the next generation and fit individuals may continue from one generation to the next. The recombination and mutation mechanisms of BCB are illustrated in figure 6 for the case of $k = 3$ unknown parameters. Consider the line through two k-dimensional parent vectors $\vec{P}_1$ and $\vec{P}_2$ selected for mating. First, determine the weighted mean $\vec{M}$ of these two vectors where the weights are given by the fitness (KS statistic) of each parent. Next, generate a $N(0, \sigma_m^2)$ variate Z. The resulting point $\vec{B} = \vec{M} + |\vec{P}_2 - \vec{P}_1| \cdot Z$ is the child, prior to mutation. Note that $\vec{B}$ is *not* restricted to lie on the line segment $\overline{P_1 P_2}$ because the support of Z is $(-\infty, \infty)$. Mutation ensues by first generating a radius r for an $k - 1$ dimensional hypersphere. The radius is a $N(0, \sigma_r^2)$ variate. Typically $\sigma_r \gg \sigma_m$. Finally, the mutated child $\vec{C}$ is selected by sampling uniformly on the surface of the $k - 1$ dimensional hypersphere. Ideally, when the algorithm terminates after several generations, the population has converged to some optimal region of the search space. The specifics of BCB can be found by Kincaid *et al.* [12] and Sobieszczanski-Sobieski *et al.* [1].

We use BCB as an alternative to analytical optimization for choosing near-optimal parameters for some given model. As BCB is a general-purpose optimizer, we simply need to define the objective function, the decision variables and set the BCB parameter values. The four parameters and their default values are: population size (50), number of generations (100), σ_m (1.0), and σ_r (4.0).

The MKSFitter tool maintains a set of continuous, univariate probability distributions $\mathcal{F}$. Each model $F \in \mathcal{F}$, for example the Weibull or normal distributions, has associated parameters [(λ, κ) for the Weibull model, and (μ, σ) for the normal model]. For each model, the parameter(s) serve as decision variable(s).

We define our objective function to be the minimization of the KS statistic. Thus, for each parametric model F with parameters $(\theta_1, \theta_2, \ldots, \theta_k) \in \Omega$ that we consider, we use BCB to
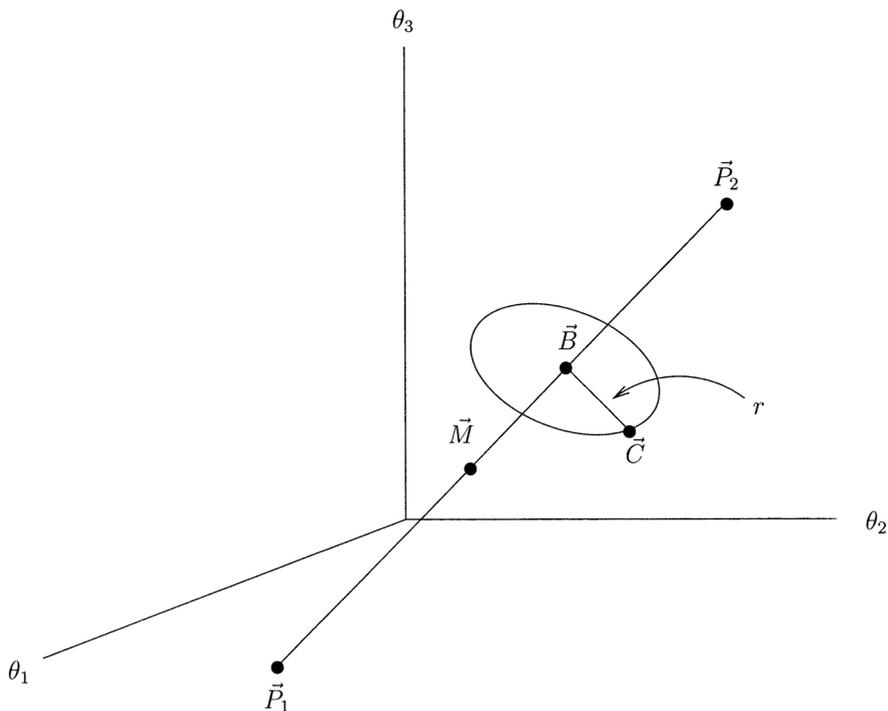
Figure 6.   BCB geometrical construct in 3D space.

heuristically solve

$$\min_{\theta_1, \theta_2, ..., \theta_\kappa} \sup_x |\hat{F}(x) - F_n(x)|.$$

Finally, we do one such optimization for each of several common models maintained in the software, so the overall optimization using one application of BCB for each model $F \in \mathcal{F}$, is

$$\min_{F \in \mathcal{F}} \left(\min_{\theta_1, \theta_2, ..., \theta_\kappa} \sup_x |\hat{F}(x) - F_n(x)| \right).$$

In the one-dimensional examples in the Introduction, there was a single valley, and a single minimum. This unimodality may or may not extend to several dimensions when multiple parameter values are to be calculated. However, unimodality is not required for BCB to perform well. In fact, it has been demonstrated to be extremely effective in solving multidimensional problems with very large numbers of local optima in Sobieszczanski–Sobieski *et al.* [1].

4.   Experimentation

To verify that MKSFitter chooses the correct population a high percentage of the time and that the MKSE performs comparable to MLE, we performed two sets of experiments.

In our first set of experiments, we constructed several random data sets and used MKS-Fitter to analyze them. We generated an $n = 10, 30$, and 100 point data set from each of four models [normal(30, 3), Weibull(1, 1/2), exponential power (1, 1/2), and log normal (2.0, 0.5)], as parameterized by Leemis [6]. These four distributions correspond to the IFR,

Table 1. Best distributions identified: $n = 10$ point data sets.

	Normal(30, 9)		Weibull(1, 0.5)		ExpPow(1, 0.5)		Log norm(2.0, 0.5)	
	Frequency	KS	Frequency	KS	Frequency	KS	Frequency	KS
Exponential								
Normal	0.53	0.141	0.03	0.136	0.13	0.155	0.15	0.124
Weibull			0.22	0.122	0.16	0.121	0.11	0.131
Log normal	0.47	0.139					0.41	0.139
Log logistic			0.34	0.126	0.28	0.136	0.25	0.134
Gompertz			0.05	0.159	0.04	0.128		
Gamma			0.11	0.129	0.11	0.132	0.08	0.113
Exponential power			0.25	0.128	0.28	0.136		

DFR, BT, and UBT survivor distribution classes, respectively. For 100 replications, we recorded the proportion of times that MKSFitter identified the various distributions in $\mathcal{F}$.

Tables 1–3 report the fraction of time that the various distributions are identified by MKS-Fitter as the parent population. Also reported are the mean KS statistic values over the data sets that the respective distributions were identified as the parent population.

We may make several observations about the results. First, we note that MKSE improves with its ability to detect the correct distribution as n increases. For example, for the exponential power data sets of size 10, 30, and 100, the proportion of the time that this distribution was identified as the best fit was 0.28, 0.47, and 0.69, respectively. Second, we note that, as expected, the mean KS values decrease as the sample size increases. For the three exponential power data sets, the mean KS value decreased from 0.136 to 0.080 to 0.047. These results give us confidence that MKSE is able to consistently identify good parameter estimates, and

Table 2. Best distributions identified: $n = 30$ point data sets.

	Normal(30, 9)		Weibull(1, 0.5)		ExpPow(1, 0.5)		Log norm(2.0, 0.5)	
	Frequency	KS	Frequency	KS	Frequency	KS	Frequency	KS
Exponential								
Normal	0.54	0.084			0.01	0.121	0.06	0.076
Weibull			0.37	0.075	0.25	0.073	0.10	0.079
Log normal	0.46	0.083					0.43	0.079
Log logistic			0.27	0.080	0.09	0.078	0.28	0.079
Gompertz					0.03	0.075		
Gamma			0.14	0.085	0.15	0.079	0.13	0.076
Exponential power			0.22	0.075	0.47	0.080		

Table 3. Best distributions identified: $n = 100$ point data sets.

	Normal(30, 9)		Weibull(1, 0.5)		ExpPow(1, 0.5)		Log norm(2.0, 0.5)	
	Frequency	KS	Frequency	KS	Frequency	KS	Frequency	KS
Exponential								
Normal	0.72	0.045						
Weibull			0.69	0.044	0.12	0.046	0.02	0.042
Log normal	0.28	0.044					0.39	0.043
Log logistic			0.11	0.047	0.01	0.057	0.39	0.046
Gompertz								
Gamma			0.08	0.044	0.18	0.044	0.20	0.042
Exponential power			0.12	0.041	0.69	0.047		

Table 4. Distances from parameter estimates to $(\lambda, \kappa) = (1, 1/2)$.

	Mean	Std. Dev.	Min.	Max.
MKSE	0.21	0.16	0.02	0.83
MLE	0.20	0.17	0.01	0.99

this ability improves with more input data. The log normal population was the most difficult to identify. This is partly due to the fact that the log logistic distribution is also in the UBT class, and that the logarithms of random variables having these distributions have symmetric probability density functions.

In our second set of experiments, we compare MKSE and MLE. We generate 100 data sets of 100 Weibull(1, 1/2) variates. To each of these data sets, we apply MKSE and MLE, which each yield a pair of estimates $(\hat{\lambda}, \hat{\kappa})$. For each such pair, we compute its distance from (1, 1/2). The goal is to determine which method yields the most accurate estimates. The distance results are given in table 4. A plot of the estimates is given in figure 7. We note that both methods yield estimates with similar accuracy. While the mean and minimum distance of MLE
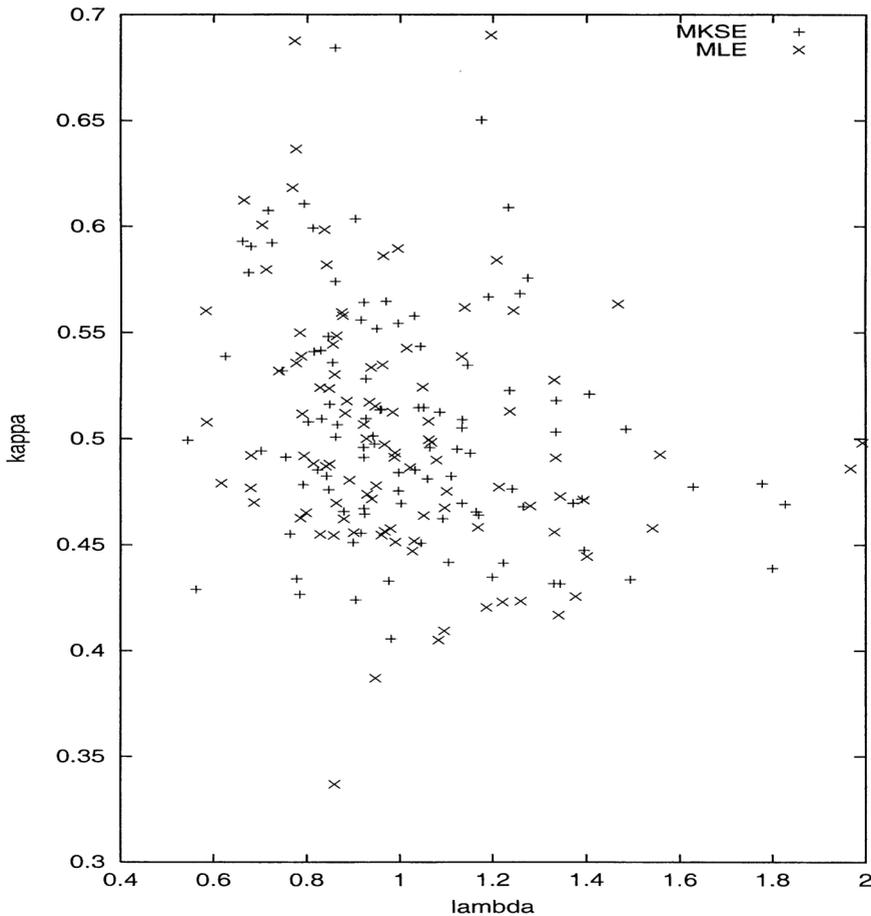

Figure 7. Weibull(1, 1/2) parameter estimates from MKSE and MLE.

is slightly smaller, MLE has a slightly greater variability and a larger maximum distance. With our sample size chosen, these differences are most likely owing to sampling variability. These results give us confidence that MKSE is on par with the more established method of MLE, with respect to accuracy of parameter estimates.

5. Using MKSfitter

5.1 *Common usage*

MKSFitter may be used as a black-box tool. That is, if a user has a sorted data set in a file named, for example, `ballbearings.dat`, the user may type at the command line:

```
mksfitter < ballbearings.dat
```

and expect to see the following output for the ball bearing data set:

```
Model          K-S          Scale Param  Shape Param
-----          ---          -----------  -----------
Log logistic  0.08079062    0.01589375    3.17549147
Log normal    0.08128213    4.14000931    0.51698360
Gamma         0.08953485   17.16327119    4.03181847
Weibull       0.09868065    0.01318428    2.23422230
Normal        0.10664124   65.11006950   32.95900275
Exponential   0.22038238   96.10201571   -----------
Gompertz      0.42456225    0.01000000    1.02967888
```

5.2 *Possible modifications*

There are a few cases in which the user may need to make very minor changes to the inner workings of MKSFitter. These involve large parameter values, 'hard' data sets, and extending the current list of models.

Like most optimization algorithms, BCB must know what the allowable ranges for decision variable values are. The default configuration allows distribution parameters to lie in [0.001, 100.0]. This information is contained in the input file `bcbparams.dat`. If the user's data is such that this range needs to be extended, the user may simply edit this file and change the value(s). No recompilation is needed.

Some optimization problems are harder than others for population-based heuristics. That is, some problems require a larger population size or more generations. Increasing either of these will allow the optimizer to examine more solutions before it terminates, thereby likely improving upon the quality of solutions found. These two parameters to BCB are also found in `bcbparams.dat`. The default population size is 50, and the default number of generations is 100. As above, these values may be changed by the user by simply editing this file. Again, no recompilation is needed. The user may want to increase these values if all models yield a poor KS, but it is believed that the data should fit some model. (Note that for some data sets, no model currently available in MKSFitter may be appropriate.)

Finally, the user may wish to incorporate additional models into MKSFitter. This requires knowledge of C programming, editing two C source files, and recompiling the software. The user should examine the function `evalobj()` in the file `bcb.c`, as well as the function `print_results()` in `runbcb.c`. These contain similar case statements which loop over the available models. It should be clear how to extend these to consider another

model. The software has been placed on the web and the source code is available at `www.math.wm.edu/~leemis/bcb.html`.

6. Conclusions and extensions

We have demonstrated the need for a data fitting tool which reduces the guesswork and computational cost of fitting failure time data to common distributions. We have shown that MKSFitter fulfills this need, in a heuristic sense, as demonstrated by our experimental results. Finally, MKSFitter is a tool that is simple to use as a black-box tool, while also being easy for advanced users to extend to include any number of arbitrary failure time models.

For modelers interested in adapting our general framework to their own setting, there are two extensions that are easily be made to the software. First, the software could be rewritten to accommodate right-censored observations, which are prevalent in reliability and biostatistical applications. In this case, the software would minimize the vertical distance between the Kaplan–Meier product-limit estimator and the fitted CDF up to the last observed failure. Second, the objective function can be changed from the KS statistic to any other goodness-of-fit statistic. The Cramer–VonMises or Anderson–Darling statistics [13] may be used in place of the KS statistic.

Acknowledgement

References

[1] Sobieszczanski-Sobieski, J., Laba, K. and Kincaid, R.K., 1998, *Bell-curve based evolutionary optimization algorithm*. Proceedings of the 7th AIAA Symposium on Multidisciplinary Analysis and Optimization, St. Louis, MO, 2–4 September, AIAA Paper 98-4971.
[2] Scott, D.W., 2001, Parametric statistical modeling by minimum integrated square error. *Technometrics*, **43**, 274–285.
[3] Drew, J., Glen, A. and Leemis, L., 2000, Computing the cumulative distribution function of the Kolmogorov–Smirnov statistic. *Computational Statistics and Data Analysis*, **34**(1), 1–15.
[4] Hogg, R.V., McKean, J.W. and Craig, A.T., 1995, *Introduction to Mathematical Statistics* (6th edn) (Englewood Cliffs, NJ: Prentice-Hall).
[5] Lieblein, J. and Zelen, M., 1956, Statistical investigation of the fatigue Life of deepgroove ball bearings. *Journal of Research of the National Bureau of Standards*, **57**(5), 273–316.
[6] Leemis, L., 1995, *Reliability: Probabilistic Models and Statistical Methods* (Englewood Cliffs, NJ: Prentice-Hall).
[7] D'Agostino, R. and Stephens, M., 1986, *Goodness-of-Fit Techniques* (Marcel Dekker).
[8] Gyorfi, L., Vajda, I. and Van Der Meulen, E., 1996a, Minimum Kolmogorov distance estimates of parameters and parametrized distributions. *Metrika*, **43**, 237–255.
[9] Gyorfi, L., Vajda, I. and Van Der Meulen, E., 1996b, Minimum Kolmogorov distance estimates for multivariate parametrized families. *American Journal of Mathematical and Management Sciences*, **16**, 167–191.
[10] Donoho, D. and Liu, R., 1988, Pathologies of some minimum distance estimators. *The Annals of Statistics*, **16**(2), 587–608.
[11] Back, T., 1996, *Evolutionary Algorithms in Theory and Practice* (Oxford: Oxford University Press).
[12] Kincaid R.K., Weber, M. and Sobieszczanski-Sobieski, J., 2000, Performance of a bell-curve based evolutionary optimization algorithm. Proceedings of the 41st AIAA Structures, Structural Dynamics, and Materials Conference, Atlanta, GA, 3–6 April, AIAA Paper 2000-1388.
[13] Lawless, J.F., 2003, *Statistical Models and Methods for Lifetime Data* (2nd edn) (New York: John Wiley & Sons, Inc.).