The Distribution of the Kolmogorov–Smirnov, Cramer–von Mises, and Anderson–Darling Test Statistics for Exponential Populations with Estimated Parameters

Article *in* Communication in Statistics- Simulation and Computation · August 2008

DOI: 10.1080/03610910801983160 · Source: DBLP

Diane Lynn Evans
Rose-Hulman Institute of Technology

Communications in Statistics—Simulation and Computation®, 37: 1396–1421, 2008

ISSN: 0361-0918 print/1532-4141 online
DOI: 10.1080/03610910801983160

Taylor & Francis
Taylor & Francis Group

Goodness-of-Fit Test

The Distribution of the Kolmogorov–Smirnov, Cramer–von Mises, and Anderson–Darling Test Statistics for Exponential Populations with Estimated Parameters

DIANE L. EVANS[1], JOHN H. DREW[2], AND LAWRENCE M. LEEMIS[2]

[1]Department of Mathematics, Rose–Hulman Institute of Technology, Terre Haute, Indiana, USA
[2]Department of Mathematics, The College of William & Mary, Williamsburg, Virginia, USA

This article presents a derivation of the distribution of the Kolmogorov–Smirnov, Cramer–von Mises, and Anderson–Darling test statistics in the case of exponential sampling when the parameters are unknown and estimated from sample data for small sample sizes via maximum likelihood.

Keywords Distribution functions; Goodness-of-fit tests; Maximum likelihood estimation; Order statistics; Transformation technique.

Mathematics Subject Classification 62F03; 62E15.

1. The Kolmogorov–Smirnov Test Statistic

The Kolmogorov–Smirnov (K–S) goodness-of-fit test compares a hypothetical or fitted cumulative distribution function (cdf) $\widehat{F}(x)$ with an empirical cdf $F_n(x)$ in order to assess fit. The empirical cdf $F_n(x)$ is the proportion of the observations $X_1, X_2, \ldots, X_n$ that are less than or equal to x and is defined as:

$$F_n(x) = \frac{I(x)}{n},$$

where n is the size of the random sample and $I(x)$ is the number of X_i's less than or equal to x.

Received July 23, 2007; Accepted February 12, 2008
Address correspondence to Lawrence M. Leemis, Department of Mathematics, The College of William & Mary, Williamsburg, VA 23187, USA; E-mail: leemis@math.wm.edu

The K–S test statistic D_n is the largest vertical distance between $F_n(x)$ and $\widehat{F}(x)$ for all values of x, i.e.,

$$D_n = \sup_x \{|F_n(x) - \widehat{F}(x)|\}.$$

The statistic D_n can be computed by calculating (Law and Kelton, 2000, p. 364)

$$D_n^+ = \max_{i=1,2,\ldots,n} \left\{ \frac{i}{n} - \widehat{F}(X_{(i)}) \right\}, \quad D_n^- = \max_{i=1,2,\ldots,n} \left\{ \widehat{F}(X_{(i)}) - \frac{i-1}{n} \right\},$$

where $X_{(i)}$ is the ith order statistic, and letting

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

Although the test statistic D_n is easy to calculate, its distribution is mathematically intractable. Drew et al. (2000) provided an algorithm for calculating the cdf of D_n when all the parameters of the hypothetical cdf $\widehat{F}(x)$ are known (referred to as the all-parameters-known case). Assuming that $\widehat{F}(x)$ is continuous, the distribution of D_n, where $X_1, X_2, \ldots, X_n$ are independent and identically distributed (iid) observations from a population with cdf $F(x)$, is a function of n, but does not depend on $F(x)$. Marsaglia et al. (2003) provided a numerical algorithm for computing $\Pr(D_n \le d)$.

The more common and practical situation occurs when the parameters are unknown and are estimated from sample data, using an estimation technique such as maximum likelihood. In this case, the distribution of D_n depends upon both n and the particular distribution that is being fit to the data. Lilliefors (1969) provides a table (obtained via Monte Carlo simulation) of selected percentiles of the K–S test statistic D_n for testing whether a set of observations is from an exponential population with unknown mean. Durbin (1975) also provides a table (obtained by series expansions) of selected percentiles of the distribution of D_n. This article presents the derivation of the *distribution* of D_n in the case of exponential sampling for $n = 1$, $n = 2$, and $n = 3$. Additionally, the distribution of the Cramer–von Mises and Anderson–Darling test statistics for $n = 1$ and $n = 2$ are derived in Sec. 2. Two case studies that analyze real-world data sets (Space shuttle accidents and commercial nuclear power accidents), where $n = 2$ and the fit to an exponential distribution is important, are given in Sec. 3. Future work involves extending the formulas established for the exponential distribution with samples of size $n = 1, 2$, and 3 to additional distributions and larger samples. For a summary of the literature available on these test statistics and goodness-of-fit techniques (including tabled values, comparative merits, and examples), see D'Agostino and Stephens (1986).

We now define notation that will be used throughout the article. Let X be an exponential random variable with probability distribution function (pdf) $f(x) = \frac{1}{\theta} e^{-x/\theta}$ and cdf $F(x) = 1 - e^{-x/\theta}$ for $x > 0$ and fixed, unknown parameter $\theta > 0$. If $x_1, x_2, \ldots, x_n$ are the sample data values, then the maximum likelihood estimator (MLE) $\hat{\theta}$ is:

$$\hat{\theta} = \frac{1}{n} \sum_{i=1}^{n} x_i.$$

We test the null hypothesis H_0 that $X_1, X_2, \ldots, X_n$ are iid exponential(θ) random variables.

1.1. *Distribution of D_1 for Exponential Sampling*

If there is only $n = 1$ sample data value, which we will call x_1, then $\hat{\theta} = x_1$. Therefore, the fitted cdf is:

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

As shown in Fig. 1, the largest vertical distance between the empirical cdf $F_1(x)$ and $\widehat{F}(x)$ occurs at x_1 and has the value $1 - 1/e$, regardless of the value of x_1. Thus, the distribution of D_1 is degenerate at $1 - 1/e$ with cdf:

$$F_{D_1}(d) = \begin{cases} 0 & d \leq 1 - 1/e \\ 1 & d > 1 - 1/e. \end{cases}$$

1.2. *Distribution of D_2 for Exponential Sampling*

If there are $n = 2$ sample data values, then the maximum likelihood estimate (mle) is $\hat{\theta} = (x_1 + x_2)/2$, and thus the fitted cdf is:

$$\widehat{F}(x) = 1 - e^{-x/\hat{\theta}} = 1 - e^{-2x/(x_1+x_2)} \quad x > 0.$$

A maximal scale invariant statistic (Lehmann, 1959, p. 215) is:

$$\frac{x_1}{x_1 + x_2}.$$

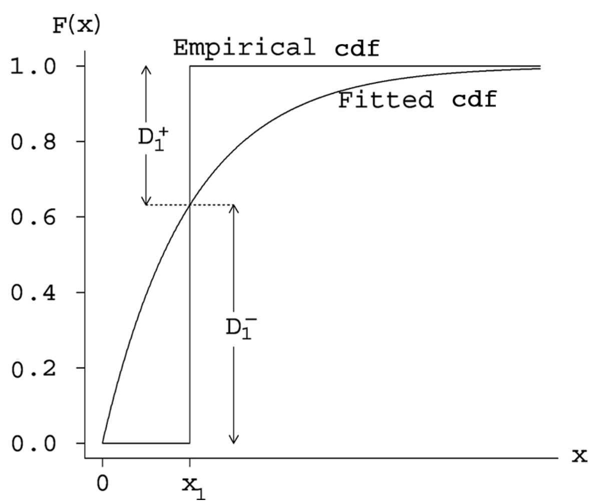

Figure 1. The empirical and fitted exponential distribution for one data value x_1, where $D_1^- = 1 - 1/e$ and $D_1^+ = 1/e$. (*Note*: The riser of the empirical cdf in this and other figures has been included to aid in comparing the lengths of D_1^- and D_1^+).

The associated statistic which is invariant to re-ordering is:

$$y = \frac{x_{(1)}}{x_{(1)} + x_{(2)}},$$

where $x_{(1)} = \min\{x_1, x_2\}$, $x_{(2)} = \max\{x_1, x_2\}$, and $0 < y \leq 1/2$ since $0 < x_{(1)} \leq x_{(2)}$. The fitted cdf $\widehat{F}(x)$ at the values $x_{(1)}$ and $x_{(2)}$ is:

$$\widehat{F}(x_{(1)}) = 1 - e^{-2x_{(1)}/(x_{(1)}+x_{(2)})} = 1 - e^{-2y}.$$

and

$$\widehat{F}(x_{(2)}) = 1 - e^{-2x_{(2)}/(x_{(1)}+x_{(2)})} = 1 - e^{-2(1-y)}.$$

It is worth noting that the fitted cdf $\widehat{F}(x)$ always intersects the second riser of the empirical cdf $F_2(x)$. This is due to the fact that $\widehat{F}(x_{(2)})$ can range from $1 - 1/e \cong 0.6321$ (when $y = 1/2$) to $1 - 1/e^2 \cong 0.8647$ (when $y = 0$), which are both included in the second riser's extension from 0.5 to 1. Conversely, the fitted cdf $\widehat{F}(x)$ *may* intersect the first riser of the empirical cdf $F_2(x)$, depending on the value of y. When $0 < y \leq \frac{\ln(2)}{2} \cong 0.3466$, the first riser is intersected by $\widehat{F}(x)$ (as displayed in Fig. 2), but when $\frac{\ln(2)}{2} < y \leq 1/2$, $\widehat{F}(x)$ lies entirely above the first riser (as subsequently displayed in Fig. 6).

Define the random lengths A, B, C, and D according to the diagram in Fig. 2. With $y = x_{(1)}/(x_{(1)} + x_{(2)})$, the lengths A, B, C, and D (as functions of y) are:

$$A = (1 - e^{-2y}) - 0 = 1 - e^{-2y} \quad 0 < y \leq 1/2,$$

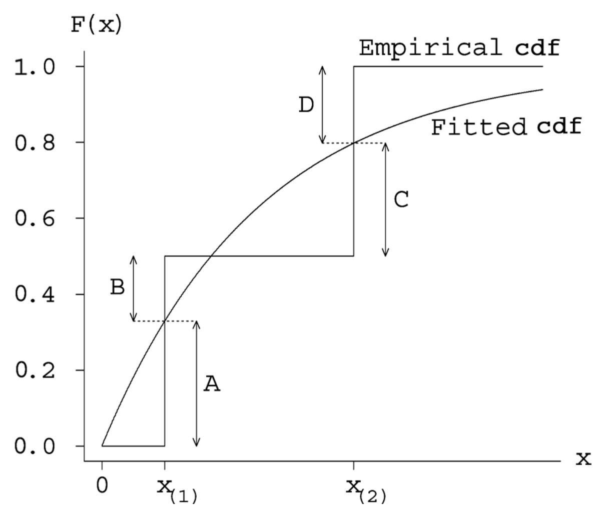

Figure 2. The empirical and fitted exponential distribution for two data values $x_{(1)}$ and $x_{(2)}$. In this particular plot, $0 < y \leq \frac{\ln(2)}{2}$, so the first riser of the empirical cdf $F_2(x)$ is intersected by the fitted cdf $\widehat{F}(x)$.

$$B = \left|\frac{1}{2} - (1 - e^{-2y})\right| = \begin{cases} e^{-2y} - \frac{1}{2} & 0 < y \le \frac{\ln(2)}{2}, \\ \frac{1}{2} - e^{-2y} & \frac{\ln(2)}{2} < y \le 1/2, \end{cases}$$

$$C = (1 - e^{-2(1-y)}) - \frac{1}{2} = \frac{1}{2} - e^{-2(1-y)} \quad 0 < y \le 1/2,$$

$$D = 1 - (1 - e^{-2(1-y)}) = e^{-2(1-y)} \quad 0 < y \le 1/2,$$

where absolute value signs are used in the definition of B to cover the case in which $\widehat{F}(x)$ does not intersect the first riser.

Figure 3 is a graph of the lengths A, B, C, and D plotted as functions of y, for $0 < y \le 1/2$. For any $y \in (0, 1/2]$, the K–S test statistic is $D_2 = \max\{A, B, C, D\}$. Since the length D is less than $\max\{A, B, C\}$ for all $y \in (0, 1/2]$, only A, B, and C are needed to define D_2.

Two particular y values of interest (indicated in Fig. 3) are y^* and y^{**} since:

1. for $0 < y < y^*$, $B = \max\{A, B, C, D\}$;
2. for $y^* < y < y^{**}$, $C = \max\{A, B, C, D\}$;
3. for $y^{**} < y \le 1/2$, $A = \max\{A, B, C, D\}$;
4. $B(y^*) = C(y^*)$ and $C(y^{**}) = A(y^{**})$.

The values of y^*, y^{**}, $C(y^*)$, and $C(y^{**})$ are given below:

- The smallest value y^* in $(0, 1/2]$ such that $B(y^*) = C(y^*)$ is:

$$y^* = 1 + \frac{1}{2}\ln\left(\frac{1}{2} - \frac{1}{2}\sqrt{1 - \frac{4}{e^2}}\right) \cong 0.0880,$$

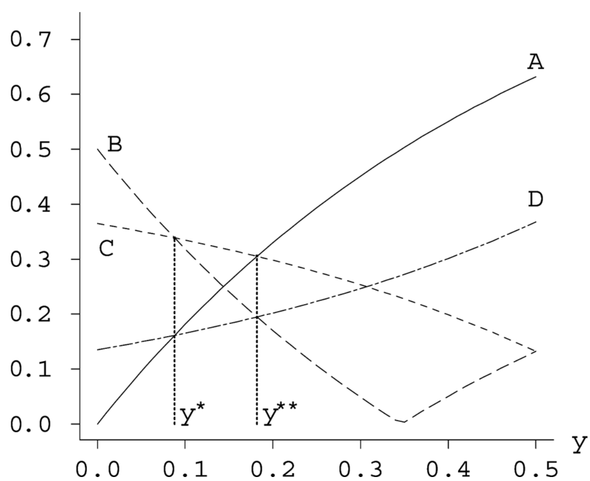

Figure 3. Lengths A, B, C, and D from Fig. 2 for $n = 2$ and $0 < y \le 1/2$.

which results in

$$C(y^*) = \frac{1}{2}\sqrt{1 - \frac{4}{e^2}} \cong 0.3386,$$

- The only value y^{**} in $(0, 1/2]$ such that $A(y^{**}) = C(y^{**})$ is:

$$y^{**} = 1 + \frac{1}{2}\ln\left(\frac{1}{4}\sqrt{1 + \frac{16}{e^2}} - \frac{1}{4}\right) \cong 0.1821,$$

which results in

$$C(y^{**}) = \frac{3}{4} - \frac{1}{4}\sqrt{1 + \frac{16}{e^2}} \cong 0.3052.$$

Thus, the largest vertical distance D_2 is computed using the length formula for $A(Y)$, $B(Y)$, or $C(Y)$ depending on the value of the random variable $Y = X_{(1)}/(X_{(1)} + X_{(2)})$, i.e.,

$$D_2 = \begin{cases} B(Y) & 0 < Y \le y^* \\ C(Y) & y^* < Y \le y^{**} \\ A(Y) & y^{**} < Y \le 1/2. \end{cases}$$

1.2.1. Determining the Distribution of $Y = X_{(1)}/(X_{(1)} + X_{(2)})$. Let X_1, X_2 be a random sample drawn from a population having pdf

$$f(x) = \frac{1}{\theta}e^{-x/\theta} \quad x > 0,$$

for $\theta > 0$. In order to determine the distribution of D_2, we must determine the distribution of $Y = X_{(1)}/(X_{(1)} + X_{(2)})$, where $X_{(1)} = \min\{X_1, X_2\}$ and $X_{(2)} = \max\{X_1, X_2\}$.

Using an order statistic result from Hogg et al. (2005, p. 193), the joint pdf of $X_{(1)}$ and $X_{(2)}$ is:

$$g(x_{(1)}, x_{(2)}) = 2! \cdot \frac{1}{\theta}e^{-x_{(1)}/\theta} \cdot \frac{1}{\theta}e^{-x_{(2)}/\theta} = \left(\frac{2}{\theta^2}\right)e^{-(x_{(1)}+x_{(2)})/\theta} \quad 0 < x_{(1)} \le x_{(2)}.$$

In order to determine the pdf of $Y = X_{(1)}/(X_{(1)} + X_{(2)})$, define the dummy transformation $Z = X_{(2)}$. The random variables Y and Z define a one-to-one transformation that maps $\mathcal{A} = \{(x_{(1)}, x_{(2)}) \mid 0 < x_{(1)} \le x_{(2)}\}$ to $\mathcal{B} = \{(y, z) \mid 0 < y \le 1/2, z > 0\}$. Since $x_{(1)} = yz/(1-y)$, $x_{(2)} = z$, and the Jacobian of the inverse transformation is $z/(1-y)^2$, the joint pdf of Y and Z is:

$$h(y, z) = \frac{2}{\theta^2}e^{-(z+yz/(1-y))/\theta} \cdot \left|\frac{z}{(1-y)^2}\right| = \frac{2z}{\theta^2(1-y)^2}e^{-z/(1-y)\theta} \quad 0 < y \le 1/2, \; z > 0.$$

Integrating the joint pdf by parts yields the marginal pdf of Y:

$$f_Y(y) = \frac{2}{\theta^2(1-y)^2}\int_0^\infty ze^{-z/(1-y)\theta}\,dz = 2 \quad 0 < y \leq 1/2,$$

i.e., $Y \sim U(0, 1/2)$.

The final step in determining the distribution of D_2 is to project $\max\{A, B, C\}$ (displayed in Fig. 4) onto the vertical axis, weighting appropriately to account for the distribution of Y. Since $\lim_{y\downarrow 0} B(y) = 1/2$, in order to determine the cdf for D_2, we must determine the functions F_α, F_β, and F_γ associated with the following intervals for the cdf of D_2:

$$F_{D_2}(d) = \begin{cases} 0 & d \leq C(y^{**}) \\ F_\alpha(d) & C(y^{**}) < d \leq C(y^*) \\ F_\beta(d) & C(y^*) < d \leq \dfrac{1}{2} \\ F_\gamma(d) & \dfrac{1}{2} < d \leq 1 - \dfrac{1}{e} \\ 1 & d > 1 - \dfrac{1}{e}. \end{cases}$$

1.2.2. Determining the Distribution of D_2. In order to determine F_α, F_β, and F_γ, it is necessary to find the point of intersection of a horizontal line of height $d \in [C(y^{**}), 1 - 1/e]$ with $A(y)$, $B(y)$, and $C(y)$, displayed in Fig. 4. These points of intersection will provide integration limits for determining the distribution of D_2.

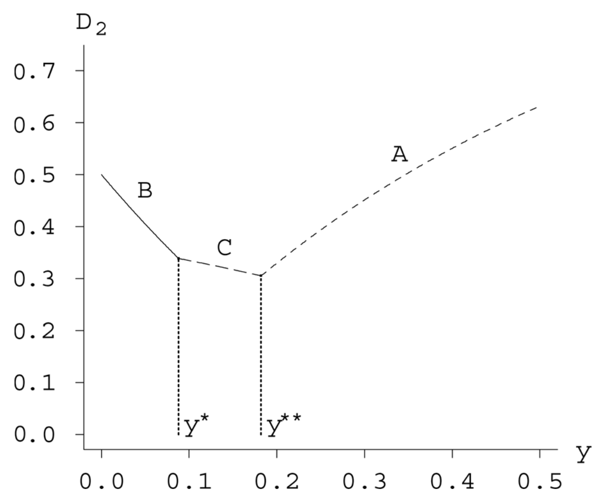

Figure 4. $D_2 = \max\{A, B, C\}$ for $n = 2$ and $0 < y \leq 1/2$.

Solving $B(y) = d$, where d satisfies $B(y^*) \le d < 1/2$ yields:

$$y = -\frac{1}{2}\ln\left(d + \frac{1}{2}\right).$$

Solving $C(y) = d$, where d satisfies $C(y^{**}) \le d \le C(y^*)$, yields:

$$y = 1 + \frac{1}{2}\ln\left(\frac{1}{2} - d\right).$$

Finally, solving $A(y) = d$ where d satisfies $A(y^{**}) \le d \le 1 - \frac{1}{e}$, yields:

$$y = -\frac{1}{2}\ln(1 - d).$$

The following three calculations yield the limits of integration associated with the functions F_α, F_β, and F_γ. For $C(y^{**}) < d \le C(y^*)$:

$$\begin{aligned} F_{D_2}(d) &= F_\alpha(d) \\ &= \Pr(D_2 \le d) \\ &= \int_{1+\frac{1}{2}\ln(\frac{1}{2}-d)}^{-\frac{1}{2}\ln(1-d)} f_Y(y)\,dy \\ &= -2 - \ln[(1/2 - d)(1 - d)]. \end{aligned}$$

For $C(y^*) < d \le 1/2$:

$$\begin{aligned} F_{D_2}(d) &= F_\beta(d) \\ &= \Pr(D_2 \le d) \\ &= \int_{-\frac{1}{2}\ln(d+\frac{1}{2})}^{-\frac{1}{2}\ln(1-d)} f_Y(y)dy \\ &= \ln\left(\frac{d + 1/2}{1 - d}\right). \end{aligned}$$

For $1/2 < d \le 1 - 1/e$:

$$\begin{aligned} F_{D_2}(d) &= F_\gamma(d) \\ &= \Pr(D_2 \le d) \\ &= \int_{0}^{-\frac{1}{2}\ln(1-d)} f_Y(y)dy \\ &= -\ln(1 - d). \end{aligned}$$

Putting the pieces together, the cdf of D_2 is:

$$F_{D_2}(d) = \begin{cases} 0 & d \leq C(y^{**}) \\ -2 - \ln(1/2 - d) - \ln(1 - d) & C(y^{**}) < d \leq C(y^{*}) \\ \ln(d + 1/2) - \ln(1 - d) & C(y^{*}) < d \leq \dfrac{1}{2} \\ -\ln(1 - d) & \dfrac{1}{2} < d \leq 1 - \dfrac{1}{e} \\ 1 & d > 1 - \dfrac{1}{e}. \end{cases}$$

Differentiating with respect to d, the pdf of D_2 is:

$$f_{D_2}(d) = \begin{cases} \dfrac{1}{1-d} + \dfrac{1}{\frac{1}{2} + d} + \dfrac{2d}{\left(\frac{1}{2} + d\right)\left(\frac{1}{2} - d\right)} & C(y^{**}) < d \leq C(y^{*}) \\ \dfrac{1}{1-d} + \dfrac{1}{\frac{1}{2} + d} & C(y^{*}) < d \leq \dfrac{1}{2} \\ \dfrac{1}{1-d} & \dfrac{1}{2} < d \leq 1 - \dfrac{1}{e}, \end{cases}$$

which is plotted in Fig. 5. The percentiles of this distribution match the tabled values from Durbin (1975).

The distribution of D_2 can also be derived using A Probability Programming Language (APPL) (Glen et al., 2001). The distribution's exact mean, variance, skewness (expected value of the standardized, centralized third moment), and kurtosis (expected value of the standardized, centralized fourth moment) can be

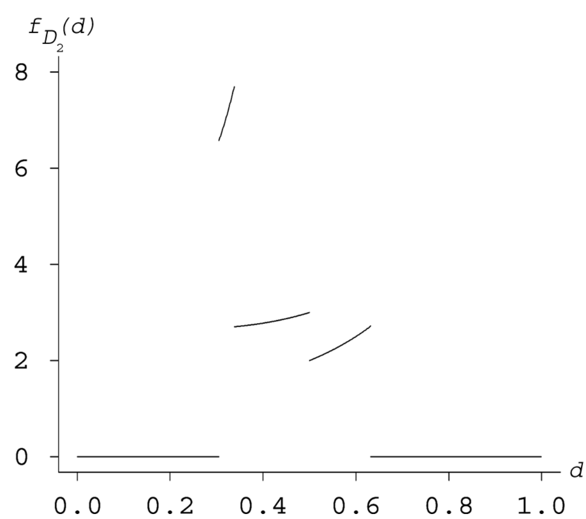

Figure 5. The pdf of D_2.

determined in Maple with the following APPL statements:

$$
\begin{aligned}
Y &:= \text{UniformRV}(0, 1/2);\\
A &:= 1 - \exp(-2 * y);\\
B &:= \exp(-2 * y) - 1/2;\\
C &:= 1/2 - \exp(-2 * (1 - y));\\
ys &:= \text{solve}(B = C, y)[1];\\
yss &:= \text{solve}(A = C, y)[1];\\
g &:= [[\text{unapply}(B, y), \text{unapply}(C, y), \text{unapply}(A, y)], [0, ys, yss, 1/2]];\\
D2 &:= \text{Transform}(Y, g);\\
&\ \text{Mean}(D2);\\
&\ \text{Variance}(D2);\\
&\ \text{Skewness}(D2);\\
&\ \text{Kurtosis}(D2);
\end{aligned}
$$

The Maple solve procedure is used to find y^* and y^{**} (the variables ys and yss) and the APPL Transform procedure transforms the random variable Y to D_2 using the piecewise segments B, C, and A. The expressions for the mean, variance, skewness, and kurtosis are given in terms of radicals, exponentials, and logarithms, e.g., the expected value of D_2 is:

$$\frac{2 - r + 6e\ln(2) - 2e + 2e\ln(e^2 + er) + e\ln(er - e^2) - e\ln(e^2 - es) - 2s + e\ln(e^2 + es)}{2e},$$

where $r = \sqrt{e^2 + 16}$ and $s = \sqrt{e^2 - 4}$. The others are too lengthy to display here, but the decimal approximations for the mean, variance, skewness, and kurtosis are, respectively, $E(D_2) \cong 0.4430$, $V(D_2) \cong 0.0100$, $\gamma_3 \cong 0.2877$, and $\gamma_4 \cong 1.7907$. These values were confirmed by Monte Carlo simulation.

Example 1.1. Suppose that two data values, $x_{(1)} = 95$ and $x_{(2)} = 100$, constitute a random sample from an unknown population. The hypothesis test

$$
\begin{aligned}
H_0 &: F(x) = F_0(x)\\
H_1 &: F(x) \neq F_0(x),
\end{aligned}
$$

where $F_0(x) = 1 - e^{-x/\theta}$, is used to test the legitimacy of modeling the data set with an exponential distribution. The mle is $\hat{\theta} = (95 + 100)/2 = 97.5$. The empirical distribution function, fitted exponential distribution, and corresponding lengths A, B, C, and D are displayed in Fig. 6.

The ratio $y = x_{(1)}/(x_{(1)} + x_{(2)}) = 95/195$ corresponds to A being the maximum of A, B, C, and D. This yields the test statistic

$$d_2 = 1 - e^{-2(95/195)} \cong 0.6226,$$

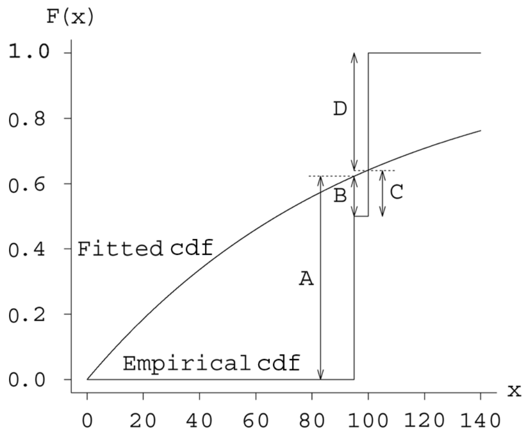

Figure 6. The empirical and fitted exponential distribution for two data values $x_{(1)} = 95$ and $x_{(2)} = 100$. In this example, $y > \frac{\ln(2)}{2}$, so the first riser of the empirical cdf is not intersected by the fitted cdf.

which falls in the right-hand tail of the distribution of D_2, as displayed in Fig 7. (The two breakpoints in the cdf are also indicated in Fig. 7.) Hence, the test statistic provides evidence to reject the null hypothesis for the goodness-of-fit test. Since large values of the test statistic lead to rejecting H_0, the p-value associated with this particular data set is:

$$p = 1 - F_{D_2}\left(1 - e^{-2(95/195)}\right) = 1 + \ln\left(1 - \left[1 - e^{-2(95/195)}\right]\right)$$
$$= 1 - \frac{190}{195} = \frac{1}{39} \cong 0.02564.$$

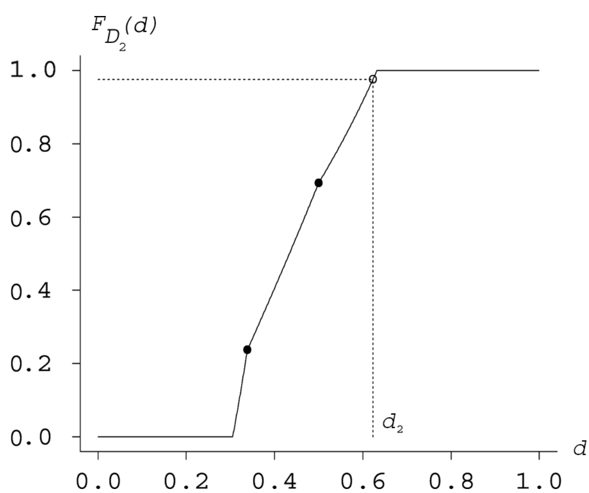

Figure 7. The cdf of D_2 and the test statistic $D_2 = 1 - e^{-2(95/195)}$.

Using the exact pdf of D_2 to determine the p-value is superior to using tables (e.g., Durbin, 1975) since the approximation associated with linear interpolation is avoided. The exact pdf is also preferred to approximations (Law and Kelton, 2000; Stephens, 1974), which often do not perform well for small values of n. Since the distribution of D_2 is not a function of θ, the power of the hypothesis test as a function of θ is constant with a value of $1-\alpha$. Since there appears to be a pattern to the functional forms associated with the three segments of the pdf of D_2, we derive the distribution of D_3 in the Appendix in an attempt to establish a pattern.

The focus of the article now shifts to investigating the distributions of other goodness-of-fit statistics.

2. Other Measures of Fit

The K–S test statistic measures the distance between $F_n(x)$ and $\widehat{F}(x)$ by using the L_∞ norm. The square of the L_2 norm gives the test statistic:

$$L_2^2 = \int_{-\infty}^{\infty} \left(F_n(x) - \widehat{F}(x)\right)^2 dx,$$

which, for exponential sampling and $n = 1$ data value, is:

$$\begin{aligned} L_2^2 &= \int_0^{x_1} \left(1 - e^{-x/x_1}\right)^2 dx + \int_{x_1}^{\infty} e^{-2x/x_1}\, dx \\ &= \left(\frac{4-e}{2e}\right) x_1. \end{aligned}$$

Since $X_1 \sim \text{exponential}(\theta)$, $L_2^2 \sim \text{exponential}\left(\frac{4-e}{2e}\theta\right)$. Unlike the K–S test statistic, the square of the L_2 norm is dependent on θ. For exponential sampling with $n = 2$, the square of the L_2 norm is also dependent on θ:

$$\begin{aligned} L_2^2 &= \int_0^{\infty} \left(F_2(x) - \widehat{F}(x)\right)^2 dx \\ &= \int_0^{x_{(1)}} \widehat{F}(x)^2 dx + \int_{x_{(1)}}^{x_{(2)}} \left(\widehat{F}(x) - \frac{1}{2}\right)^2 dx + \int_{x_{(2)}}^{\infty} \left(1 - \widehat{F}(x)\right)^2 dx \\ &= \int_0^{x_{(1)}} \left(1 - e^{-2x/(x_1+x_2)}\right)^2 dx + \int_{x_{(1)}}^{x_{(2)}} \left(\frac{1}{2} - e^{-2x/(x_1+x_2)}\right)^2 dx + \int_{x_{(2)}}^{\infty} \left(e^{-2x/(x_1+x_2)}\right)^2 dx \\ &= \int_0^{x_{(1)}} \left(1 - 2e^{-2x/(x_1+x_2)}\right) dx + \int_{x_{(1)}}^{x_{(2)}} \left(\frac{1}{4} - e^{-2x/(x_1+x_2)}\right) dx + \int_0^{\infty} e^{-4x/(x_1+x_2)}\, dx \\ &= -\frac{x_{(2)}}{2} + \frac{x_{(1)} + x_{(2)}}{2} \cdot \left(e^{-2x_{(1)}/(x_{(1)}+x_{(2)})} + e^{-2x_{(2)}/(x_{(1)}+x_{(2)})}\right), \end{aligned}$$

where $X_{(1)} \sim \text{exponential}(2\theta)$ and $X_{(2)}$ has pdf $f_{X_{(2)}}(x) = 2(1 - e^{-x/\theta})\left(\frac{1}{\theta}e^{-x/\theta}\right)$, $x > 0$.

Unlike the square of the L_2 norm, the Cramer–von Mises and Anderson–Darling test statistics (Lawless, 2003) are distribution-free. They can be defined as:

$$W_n^2 = n \int_{-\infty}^{\infty} \left(F_n(x) - \widehat{F}(x)\right)^2 d\widehat{F}(x)$$

and

$$A_n^2 = n \int_{-\infty}^{\infty} \frac{\left(F_n(x) - \widehat{F}(x)\right)^2}{\widehat{F}(x)[1 - \widehat{F}(x)]} \, d\widehat{F}(x),$$

where n is the sample size. The computational formulas for these statistics are:

$$W_n^2 = \sum_{i=1}^{n} \left(\widehat{F}(x_{(i)}) - \frac{i - 0.5}{n}\right)^2 + \frac{1}{12n}$$

and

$$A_n^2 = -\sum_{i=1}^{n} \frac{2i - 1}{n} \left(\ln(\widehat{F}(x_{(i)})) + \ln(1 - \widehat{F}(x_{(n+1-i)}))\right) - n.$$

2.1. *Distribution of W_1^2 and A_1^2 for Exponential Sampling*

When $n = 1$ and sampling is from an exponential population, the Cramer–von Mises test statistic is:

$$W_1^2 = \left(\frac{1}{2} - \frac{1}{e}\right)^2 + \frac{1}{12} = \frac{1}{3} - \frac{1}{e} + \frac{1}{e^2}.$$

Thus, the Cramer–von Mises test statistic is degenerate for $n = 1$ with cdf:

$$F_{W_1^2}(w) = \begin{cases} 0 & w \le \frac{1}{3} - \frac{1}{e} + \frac{1}{e^2} \\ 1 & w > \frac{1}{3} - \frac{1}{e} + \frac{1}{e^2}. \end{cases}$$

When $n = 1$ and sampling is from an exponential population, the Anderson–Darling test statistic is:

$$A_1^2 = -\ln(1 - e^{-1}) - \ln(e^{-1}) - 1 = 1 - \ln(e - 1).$$

It is also degenerate for $n = 1$ with cdf:

$$F_{A_1^2}(a) = \begin{cases} 0 & a \le 1 - \ln(e - 1) \\ 1 & a > 1 - \ln(e - 1). \end{cases}$$

2.2. *Distribution of W_2^2 and A_2^2 for Exponential Sampling*

When $n = 2$ and sampling is from an exponential population, the Cramer–von Mises test statistic is:

$$W_2^2 = \left(e^{-x_{(1)}/\hat{\theta}} - \frac{3}{4}\right)^2 + \left(e^{-x_{(2)}/\hat{\theta}} - \frac{1}{4}\right)^2 + \frac{1}{24},$$

where $\hat{\theta} = (x_1 + x_2)/2$. The Anderson–Darling test statistic is:

$$A_2^2 = 2 - \frac{1}{2}\ln\left(e^{x_{(1)}/\hat{\theta}} - 1\right) - \frac{3}{2}\ln\left(e^{x_{(2)}/\hat{\theta}} - 1\right).$$

If we let $y = x_{(1)}/(x_{(1)} + x_{(2)})$, as we did when working with D_2, we obtain the following formulas for W_2^2 and A_2^2 in terms of y:

$$W_2^2 = \left(e^{-2y} - \frac{3}{4}\right)^2 + \left(e^{-2(1-y)} - \frac{1}{4}\right)^2 + \frac{1}{24},$$

and

$$A_2^2 = 2 - \frac{1}{2}\ln(e^{2y} - 1) - \frac{3}{2}\ln(e^{2(1-y)} - 1),$$

for $0 < y \le 1/2$. Graphs of D_2, W_2^2, and A_2^2 are displayed in Fig. 8. Although the ranges of the three functions are quite different, they all share similar shapes.

Each of the three functions plotted in Fig. 8 achieves a minimum between $y = 0.15$ and $y = 0.2$. The Cramer–von Mises test statistic W_2^2 achieves a minimum at y^{***} that satisfies

$$4e^{-4y} - 3e^{-2y} - 4e^{-4(1-y)} + e^{-2(1-y)} = 0$$

for $0 < y \le 1/2$. This is equivalent to fourth-degree polynomial in e^{2y} that can be solved exactly using radicals. The minimum is achieved at $y^{***} \cong 0.1549$. Likewise, the Anderson–Darling test statistic A_2^2 achieves a minimum at y^{****} that satisfies

$$e^{2y} + 2e^2 - 3e^{2(1-y)} = 0$$

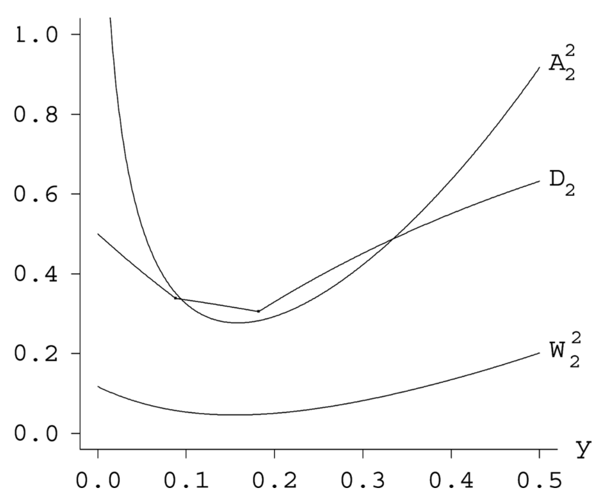

Figure 8. Graphs of D_2, W_2^2, and A_2^2 for $n = 2$ and $0 < y \le 1/2$.

for $0 < y \le 1/2$, which yields

$$y^{****} = \frac{1}{2} + \frac{1}{2}\ln\left(\sqrt{e^2+3} - e\right) \cong 0.1583.$$

These values and other pertinent values associated with D_2, W_2^2, and A_2^2 are summarized in Table 1.

Figure 8 can be helpful in determining which of the three goodness-of-fit statistics is appropriate in a particular application. Consider, for instance, a reliability engineer who is interested in detecting whether reliability growth or reliability degradation is occurring for a repairable system. One would expect a shorter failure time followed by a longer failure time if reliability growth were occurring; one would expect a longer failure time followed by a shorter failure time if reliability degradation were occurring. In either case, these correspond to a small value of y, so Fig. 8 indicates that the Anderson–Darling test is preferred due to the vertical asymptote at $y = 0$.

For notational convenience below, let $D_2(y)$, $W_2^2(y)$, and $A_2^2(y)$ denote the values of D_2, W_2^2, and A_2^2, respectively, corresponding to a specific value of y. For example, $W_2^2(1/4)$ denotes the value of W_2^2 when $y = 1/4$.

Example 2.1. Consider again the data from Example 1.1: $x_{(1)} = 95$ and $x_{(2)} = 100$. Since $y = 95/195 = 19/39$, the Cramer–von Mises test statistic is:

$$w_2^2 = \left(e^{-38/39} - 3/4\right)^2 + \left(e^{-40/39} - 1/4\right)^2 + 1/24 \cong 0.1923.$$

The p-value for this test statistic is the same as the p-value for the K–S test statistic, namely:

$$\int_{19/39}^{1/2} f_Y(y)dy = \int_{19/39}^{1/2} 2\,dy = 1/39 \cong 0.0256.$$

More generally, for each value of y such that both $D_2(y) > D_2(0)$ and $W_2^2(y) > W_2^2(0)$:

$$F_{D_2}(D_2(y)) = F_Y(y) = F_{W_2^2}\left(W_2^2(y)\right).$$

Table 1
Pertinent values associated with the test statistics D_2, W_2^2, and A_2^2

Test statistic	Value when $y = 0$	Minimized at	Global minimum on $(0, 1/2]$	Value when $y = 1/2$
D_2	$\frac{1}{2}$	$y^{**} \cong 0.1821$	$D_2(y^{**}) \cong 0.3052$	$1 - \frac{1}{e} \cong 0.6321$
W_2^2	$\frac{1}{6} + \frac{1}{e^4} - \frac{1}{2e^2}$ $\cong 0.1173$	$y^{***} \cong 0.1549$	$W_2^2(y^{***}) \cong 0.04623$	$\frac{2}{e^2} - \frac{2}{e} + \frac{2}{3} \cong 0.2016$
A_2^2	$+\infty$	$y^{****} \cong 0.1583$	$A_2^2(y^{****}) \cong 0.2769$	$2 - 2\ln(e-1)$ $\cong 0.9174$

The Anderson–Darling test statistic for $x_{(1)} = 95$ and $x_{(2)} = 100$ is:

$$a_2^2 = 2 - \frac{1}{2}\ln(e^{38/39} - 1) - \frac{3}{2}\ln(e^{40/39} - 1) \cong 0.8774.$$

Since the value of $A_2^2(y)$ exceeds the test statistic $a_2^2 \cong 0.8774$ only for $y < y' = 0.02044$ (where y' is the first intersection point of $A_2^2(y)$ and the horizontal line with height $A_2^2(19/39)$) and for $y > 19/39$, the p-value for the Anderson–Darling goodness-of-fit test is given by:

$$\begin{aligned} p &= \int_0^{y'} 2\,dy + \int_{19/39}^{1/2} 2\,dy \\ &= 2y' + (1 - 38/39) \\ &\cong 0.06654. \end{aligned}$$

2.2.1. *Determining the Distribution of W_2^2 and A_2^2.* As was the case with D_2, we can find exact expressions for the pdfs of W_2^2 and A_2^2. Consider W_2^2 first. For w values in the interval $W_2^2(y^{***}) \le w < W_2^2(0)$, the cdf of W_2^2 is:

$$\begin{aligned} F_{W_2^2}(w) &= P(W_2^2 \le w) \\ &= \int_{y_1}^{y_2} f_Y(y)dy \\ &= 2(y_2 - y_1), \end{aligned}$$

where y_1 and y_2 are the ordered solutions to $W_2^2(y) = w$. For w values in the interval $W_2^2(0) \le w < W_2^2(1/2)$, the cdf of W_2^2 is:

$$\begin{aligned} F_{W_2^2}(w) &= P(W_2^2 \le w) \\ &= \int_0^{y_1} f_Y(y)dy \\ &= 2y_1, \end{aligned}$$

where y_1 is the solution to $W_2^2(y) = w$ on $0 < y < 1/2$. The following APPL code can be used to find the pdf of W_2^2:

```
Y := UniformRV(0, 1/2);
W := (exp(-2 * y) - 3/4)^2 + (exp(-2 * (1 - y)) - 1/4)^2 + 1/24;
ysss := solve(diff(W, y) = 0, y)[1];
g := [[unapply(W, y), unapply(W, y)], [0, ysss, 1/2]];
W2 := Transform(Y, g).
```

The Transform procedure requires that the transformation g be input in piecewise monotone segments. The resulting pdf for W_2^2 is too lengthy to display here.

Now consider A_2^2. For a value in the interval $A_2^2(y^{****}) \le a < A_2^2(1/2)$, the cdf of A_2^2 is:

$$\begin{aligned} F_{A_2^2}(a) &= P(A_2^2 \le a) \\ &= \int_{y_1}^{y_2} f_Y(y)dy \\ &= 2(y_2 - y_1), \end{aligned}$$

where y_1 and y_2 are the ordered solutions to $A_2^2(y) = a$ on $0 < y \le 1/2$. For a values in the interval $A_2^2(1/2) \le a < \infty$, the cdf of A_2^2 is:

$$\begin{aligned} F_{A_2^2}(a) &= P(A_2^2 \le a) \\ &= \int_{y_1}^{1/2} f_Y(y)dy \\ &= 1 - 2y_1, \end{aligned}$$

where y_1 is the solution to $A_2^2(y) = a$ on $0 < y < 1/2$. The following APPL code can be used to find the pdf of A_2^2:

$$\begin{aligned} Y &:= \text{UniformRV}(0, 1/2); \\ A &:= 2 - \ln(\exp(2 * y) - 1)/2 - 3 * \log(\exp(2 * (1 - y)) - 1)/2; \\ yssss &:= \text{solve}(\text{diff}(A, y) = 0, y)[1]; \\ g &:= [[\text{unapply}(A, y), \text{unapply}(A, y)], [0, yssss, 1/2]]; \\ A_2 &:= \text{Transform}(Y, g). \end{aligned}$$

The resulting pdf for A_2^2 is again too lengthy to display here.

3. Applications

Although statisticians prefer large sample sizes because of the associated desirable statistical properties of estimators as the sample size n becomes large, there are examples of real-world data sets with only $n = 2$ observations where the fit to an exponential distribution is important. In this section, we focus on two applications: U.S. Space Shuttle flights and the world-wide commercial nuclear power industry. Both applications involve significant government expenditures associated with decisions that must be made based on limited data. In both cases, "events" are failures and the desire is to test whether a homogeneous Poisson process model or other (e.g., a non homogeneous Poisson process) model is appropriate, i.e., determining whether failures occur randomly over time. Deciding which of the models is appropriate is important to reliability engineers since a non homogeneous Poisson process with a decreasing intensity function may be a sign of reliability growth or improvement over time (Rigdon and Basu, 2000).

Example 3.1. NASA's Space Shuttle program has experienced $n = 2$ catastrophic failures which have implications for the way in which the United States will pursue future space exploration. On January 28, 1986, the *Challenger* exploded 72 seconds after liftoff. Failure of an O-ring was determined as the most likely cause of the accident. On February 1, 2003, Shuttle *Columbia* was lost during its return to Earth. Investigators believed that tile damage during ascent caused the accident. These two failures occurred on the 25th and 113th Shuttle flights. A goodness-of-fit test is appropriate to determine whether the failures occurred randomly, or equivalently, whether a Poisson process model is appropriate. The hope is that the data will *fail* this test due to the fact that reliability growth has occurred due to the many improvements that have been made to the Shuttle (particularly after the *Challenger* accident), and perhaps a non homogeneous Poisson process with a decreasing intensity function is a more appropriate stochastic model for failure times. Certainly, large amounts of money have been spent and some judgments about the safety and direction of the future of the Shuttle program should be made on the basis of these two data values.

The appropriate manner to model *time* in this application is nontrivial. There is almost certainly increased risk on liftoff and landing, but the time spent on the mission should also be included since an increased mission time means an increased exposure to internal and external failures while a Shuttle is in orbit. Because of this inherent difficulty in quantifying time, we do our numerical analysis on an example in an application area where time is more easily measured.

Example 3.2. The world-wide commercial nuclear power industry has experienced $n = 2$ core meltdowns in its history. The first was at the Three Mile Island nuclear facility on March 28, 1979. The second was at Chernobyl on April 26, 1986. As in the case of the Space Shuttle accidents, it is again of interest to know whether the meltdowns can be considered to be events from a Poisson process. The hypothesis test of interest here is whether the two times to meltdown are independent observations from an exponential population with a rate parameter estimated from data. Measuring time in this case is not trivial because of the commissioning and decommissioning of facilities over time. The first nuclear power plant was the Calder Hall I facility in the United Kingdom, commissioned on October 1, 1956. Figure 9 shows the evolution of the number of active commercial reactors between that date and the Chernobyl accident on April 26, 1986. The commissioning and decommissioning dates of all commercial nuclear reactors is given in Cho and Spiegelberg-Planer (2002). Downtime for maintenance has been ignored in determining the times of the two accidents.

Using the data illustrated in Fig. 9, the time of the two accidents measured in cumulative commercial nuclear reactor years is found by integrating under the curve. The calendar dates were converted to decimal values using Julian dates, adjusting for leap years. The two accidents occurred at 1548.02 and 3372.27 cumulative operating years, respectively. This means that the hypothesis test is to see whether the times between accidents, namely 1548.02 and $3372.27 - 1548.02 = 1824.25$ can be considered independent observations from an exponential population. The maximum likelihood estimator for the mean time between core meltdowns is $\hat{\theta} = 1686.14$ years. This results in a y-value of $y = 1548.02/(1548.02 +$

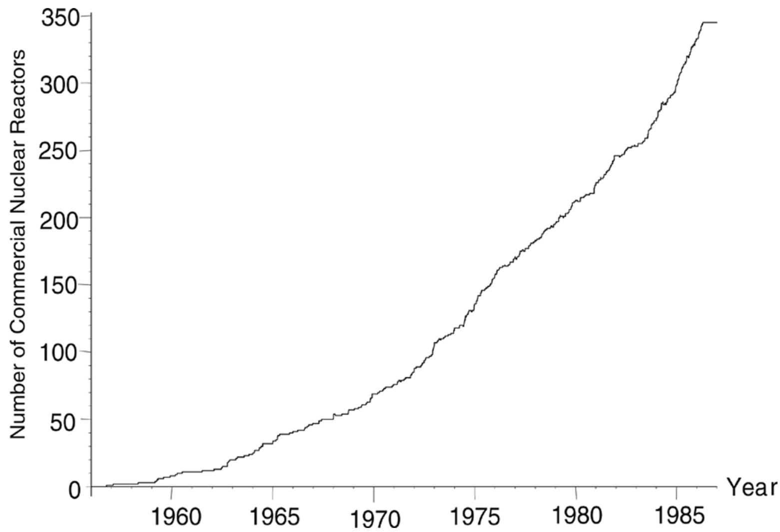

Figure 9. Number of operating commercial nuclear power plants world-wide between October 1, 1956 and April 26, 1986.

1824.25) $= 0.459$ and a K–S test statistic of $d_2 = 0.601$. This corresponds to a p-value for the associated goodness-of-fit test of $p = 1 + \ln(1 - d_2) = 0.082$. There is *not* enough statistical evidence to conclude a nonhomogeneous model is appropriate here, so it is reasonable to model nuclear power accidents as random events. Figure 10 shows a plot of the fitted cumulative distribution function and associated values of A, B, C, and D for this data set.

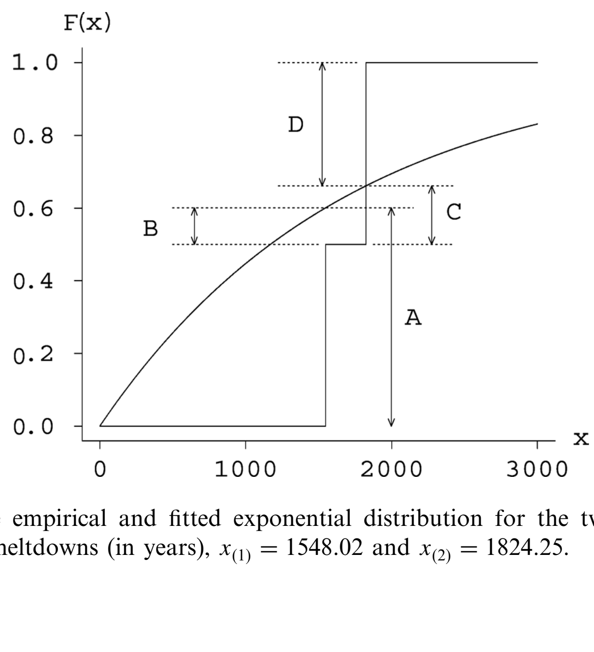

Figure 10. The empirical and fitted exponential distribution for the two times between nuclear reactor meltdowns (in years), $x_{(1)} = 1548.02$ and $x_{(2)} = 1824.25$.

Appendix: Distribution of D_3 for Exponential Sampling

This Appendix contains a derivation of the distribution of the K–S test statistic when $n = 3$ observations x_1, x_2, and x_3 are drawn from an exponential population with fixed, positive, unknown mean θ. The maximum likelihood estimator is $\hat{\theta} = (x_1 + x_2 + x_3)/3$, which results in the fitted cdf:

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

Analogous to the $n = 2$ case, define

$$y = \frac{x_{(1)}}{x_{(1)} + x_{(2)} + x_{(3)}}$$

and

$$z = \frac{x_{(2)}}{x_{(1)} + x_{(2)} + x_{(3)}}$$

so that

$$1 - y - z = \frac{x_{(3)}}{x_{(1)} + x_{(2)} + x_{(3)}}.$$

The domain of definition of y and z is:

$$\mathscr{D} = \{(y, z) \mid 0 < y < z < (1 - y)/2\}.$$

The values of the fitted cdf at the three order statistics are:

$$\widehat{F}(x_{(1)}) = 1 - e^{-x_{(1)}/\hat{\theta}} = 1 - e^{-3y},$$
$$\widehat{F}(x_{(2)}) = 1 - e^{-x_{(2)}/\hat{\theta}} = 1 - e^{-3z},$$

and

$$\widehat{F}(x_{(3)}) = 1 - e^{-x_{(3)}/\hat{\theta}} = 1 - e^{-3(1-y-z)}.$$

The vertical distances A, B, C, D, E, and F (as functions of y and z) are defined in a similar fashion to the $n = 2$ case (see Fig. 2):

$$A = 1 - e^{-3y}$$
$$B = \left|\frac{1}{3} - \left(1 - e^{-3y}\right)\right| = \left|e^{-3y} - \frac{2}{3}\right|$$
$$C = \left|\left(1 - e^{-3z}\right) - \frac{1}{3}\right| = \left|e^{-3z} - \frac{2}{3}\right|$$
$$D = \left|\frac{2}{3} - \left(1 - e^{-3z}\right)\right| = \left|e^{-3z} - \frac{1}{3}\right|$$

$$E = \left| \left(1 - e^{-3(1-y-z)}\right) - \frac{2}{3} \right| = \left| e^{-3(1-y-z)} - \frac{1}{3} \right|$$
$$F = 1 - \left(1 - e^{-3(1-y-z)}\right) = e^{-3(1-y-z)}$$

for $(y, z) \in \mathcal{D}$.

Figure 11 shows the regions associated with the maximum of A, B, C, D, E, F for $(y, z) \in \mathcal{D}$. In three dimensions, with $D_3 = \max\{A, B, C, D, E, F\}$ as the third axis, this figure appears to be a container with the region E at the bottom of the container and with each of the other four sides rising as they move away from their intersection with E. The absolute value signs that appear in the final formulas for B, C, D, and E above can be easily removed since, over the region $\mathcal{D}$ associated with D_3, the expressions within the absolute value signs are always positive for B and D, but always negative for C and E. The distance F is never the largest of the six distances for any $(y, z) \in \mathcal{D}$, so it can be excluded from consideration. Table 2 gives the functional forms of the two-way intersections between the five regions shown in Fig. 11. Note that the BC and AD curves, and the AC and BD curves, are identical.

In order to determine the breakpoints in the support for D_3, it is necessary to find the (y, z) coordinates of the three-way intersections of the five regions in Fig. 11 and the two-way intersections of the regions on the boundary of $\mathcal{D}$. Table 3 gives the values of y and z for these breakpoints on the boundary of $\mathcal{D}$, along with the value of $D_3 = \max\{A, B, C, D, E, F\}$ at these values, beginning at $(y, z) = (0, 1/2)$ and proceeding in a counterclockwise direction. One point has been excluded from

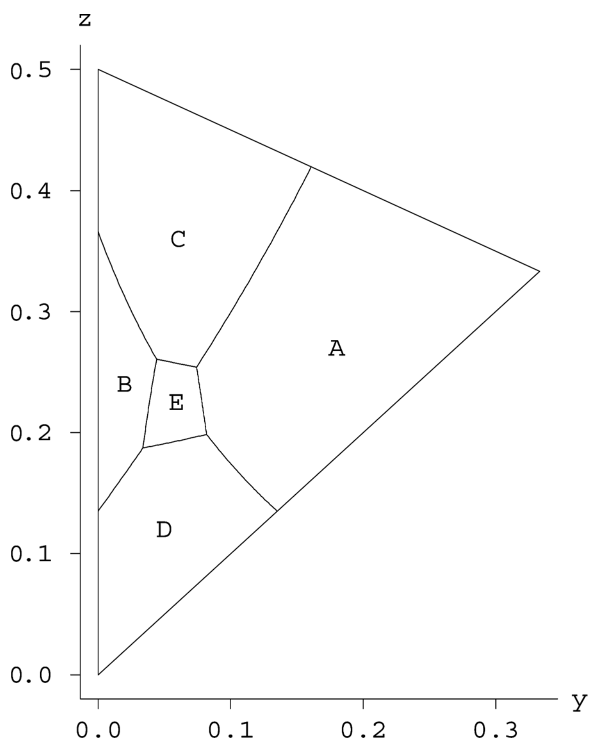

Figure 11. Regions associated with $\max\{A, B, C, D, E, F\}$ over $(y, z) \in \mathcal{D}$.

Table 2
Intersections of regions A, B, C, D, and E in $\mathcal{D}$

AD	$z = -\frac{1}{3}\ln\left(\frac{4}{3} - e^{-3y}\right)$
BD	$z = -\frac{1}{3}\ln\left(e^{-3y} - \frac{1}{3}\right)$
BC	$z = -\frac{1}{3}\ln\left(\frac{4}{3} - e^{-3y}\right)$
AC	$z = -\frac{1}{3}\ln\left(e^{-3y} - \frac{1}{3}\right)$
AE	$z = \frac{1}{3}\ln\left[e^{3(1-y)}\left(e^{-3y} - \frac{2}{3}\right)\right]$
DE	$z = \frac{1}{3}\ln\left[\frac{1}{3}e^{3(1-y)}\left(1 - \sqrt{1 - 9e^{-3(1-y)}}\right)\right]$
BE	$z = \frac{1}{3}\ln\left[e^{3(1-y)}\left(1 - e^{-3y}\right)\right]$
CE	$z = \frac{1}{3}\ln\left[\frac{1}{6}e^{3(1-y)}\left(-1 + \sqrt{1 + 36e^{-3(1-y)}}\right)\right]$

Table 3 because of the intractability of the values (y, z). The three-way intersection between regions A, C, and the line $z = (1-y)/2$ can only be expressed in terms of the solution to a cubic equation. After some algebra, the point of intersection is the decimal approximation $(y, z) \cong (0.1608, 0.4196)$ and the associated value of D_3 is $2/3$ minus the only real solution to the cubic equation

$$3d^3 + d^2 - 3e^{-3} = 0,$$

which yields

$$d_{AC} = \frac{7}{9} - \frac{1}{18}(2916e^{-3} - 8 + c)^{1/3} - \frac{2}{9}(2916e^{-3} - 8 + c)^{-1/3} \cong 0.3827,$$

where $c = 108\sqrt{729e^{-6} - 4e^{-3}}$.

The three-way intersection points in the interior of $\mathcal{D}$ are more difficult to determine than those on the boundary. The value of D_3 associated with each of these four points is the single real root of a cubic equation on the support of D_3. These equations and approximate solution values, in ascending order, are given in Table 4. For example, consider the value of the maximum at the intersection of regions A,

Table 3
Intersection points along the boundary of $\mathcal{D}$

y	z	D_3
0	$1/2$	$2/3 - e^{-3/2} \cong 0.4435$
0	$\ln(3)/3$	$1/3 \cong 0.3333$
0	$\ln(3/2)/3$	$1/3 \cong 0.3333$
0	0	$2/3 \cong 0.6667$
$\ln(3/2)/3$	$\ln(3/2)/3$	$1/3 \cong 0.3333$
$1/3$	$1/2$	$1 - 1/e \cong 0.6321$

Table 4
Three-way interior intersection points of regions A, B, C, D, and E in $\mathscr{D}$

Regions	Cubic equation	Approximate solution
ACE	$e^3(1-d)\left(\frac{2}{3}-d\right)\left(\frac{1}{3}-d\right)=1$	$d_{ACE} \cong 0.2000$
BCE	$e^3\left(\frac{1}{3}-d\right)\left(d+\frac{2}{3}\right)\left(\frac{2}{3}-d\right)=1$	$d_{BCE} \cong 0.2091$
ADE	$e^3\left(\frac{1}{3}-d\right)\left(d+\frac{1}{3}\right)(1-d)=1$	$d_{ADE} \cong 0.2178$
BDE	$e^3\left(d+\frac{2}{3}\right)\left(d+\frac{1}{3}\right)\left(\frac{1}{3}-d\right)=1$	$d_{BDE} \cong 0.2366$

C, and E in Fig. 11. The value of D_3 must satisfy the cubic equation:

$$e^3(1-d)\left(\frac{2}{3}-d\right)\left(\frac{1}{3}-d\right)=1,$$

which yields

$$\begin{aligned} d_{ACE} &= \frac{(243+c)^{2/3}12^{2/3}c - 243(243+c)^{2/3}12^{2/3} + 144e^5 - 12^{4/3}e^4(243+c)^{1/3}}{216e^5} \\ &\cong 0.19998, \end{aligned}$$

where $c = \sqrt{59049 - 12e^6}$.

The largest value of $D_3 = \max\{A, B, C, D, E\}$ on $\mathscr{D}$ occurs at the origin ($y = 0$ and $z = 0$) and has value $2/3$, which is the upper limit of the support of D_3. The smallest value of D_3 on $\mathscr{D}$ occurs at the intersection ACE and is $d_{ACE} \cong 0.19998$, which is the lower limit of the support of D_3.

Determining the Joint Distribution of Y and Z

The next step is to determine the distribution of $Y = X_{(1)}/(X_{(1)} + X_{(2)} + X_{(3)})$ and $Z = X_{(2)}/(X_{(1)} + X_{(2)} + X_{(3)})$. Using an order statistic result from Hogg et al. (2005, p. 193), the joint pdf of $X_{(1)}$, $X_{(2)}$, and $X_{(3)}$ is:

$$g(x_{(1)}, x_{(2)}, x_{(3)}) = \frac{3!}{\theta^3} e^{-(x_{(1)}+x_{(2)}+x_{(3)})/\theta} \quad 0 < x_{(1)} \le x_{(2)} \le x_{(3)}.$$

In order to determine the joint pdf of $Y = X_{(1)}/(X_{(1)} + X_{(2)} + X_{(3)})$ and $Z = X_{(2)}/(X_{(1)} + X_{(2)} + X_{(3)})$, define the dummy transformation $W = X_{(3)}$. The random variables Y, Z, and W define a one-to-one transformation from $\mathscr{A} = \{(x_{(1)}, x_{(2)}, x_{(3)}) \,|\, 0 < x_{(1)} \le x_{(2)} \le x_{(3)}\}$ to $\mathscr{B} = \{(y, z, w) \,|\, 0 < y < z < (1-y)/2, w > 0\}$. Since $x_{(1)} = yw/(1-y-z)$, $x_{(2)} = zw/(1-y-z)$, and $x_{(3)} = w$, and the Jacobian of the inverse transformation is $w^2/(1-y-z)^3$, the joint pdf of Y, Z, and W on $\mathscr{B}$ is:

$$h(y, z, w) = \frac{6}{\theta^3} e^{-\left(\frac{yw+zw}{1-y-z}+w\right)/\theta} \cdot \left| \frac{w^2}{(1-y-z)^3} \right|$$

$$= \frac{6}{\theta^3} \cdot \frac{w^2}{(1-y-z)^3} \cdot e^{-\frac{w}{(1-y-z)\theta}} \quad (y, z, w) \in \mathcal{B}.$$

Integrating by parts, the joint density of Y and Z on $\mathcal{D}$ is:

$$f_{Y,Z}(y, z) = \frac{6}{\theta^3(1-y-z)^3} \int_0^\infty w^2 e^{-\frac{w}{(1-y-z)\theta}}\, dw = 12 \quad (y, z, w) \in \mathcal{D},$$

i.e., Y and Z are uniformly distributed on $\mathcal{D}$.

Determining the Distribution of D_3

The cdf of D_3 will be defined in a piecewise manner, with breakpoints at the following ordered quantities: d_{ACE}, d_{BCE}, d_{ADE}, d_{BDE}, $1/3$, d_{AC}, $\frac{2}{3} - e^{-3/2}$, $1 - \frac{1}{e}$, and $2/3$. The cdf $F_{D_3}(d) = \Pr(D_3 \le d)$ is found by integrating the joint pdf of Y and Z over the appropriate limits, yielding:

$$F_{D_3}(d) = \begin{cases} 0 \quad d \le d_{ACE} \\ \frac{2}{3}\left[\ln\left(e^3[1-d]\left[\frac{2}{3}-d\right]\left[\frac{1}{3}-d\right]\right)\right]^2 \quad d_{ACE} < d \le d_{BCE} \\ \frac{2}{3}\ln\left[e^6(1-d)\left(\frac{2}{3}-d\right)^2\left(\frac{2}{3}+d\right)\left(\frac{1}{3}-d\right)^2\right]\ln\left(\frac{1-d}{2/3+d}\right) \\ \quad d_{BCE} < d \le d_{ADE} \\ \frac{4}{3}\ln\left(\frac{d+1/3}{2/3-d}\right)\ln\left(\frac{d+2/3}{1-d}\right) - \frac{2}{3}\left[\ln\left(e^3\left[d+\frac{2}{3}\right]\left[d+\frac{1}{3}\right]\left[\frac{1}{3}-d\right]\right)\right]^2 \\ \quad d_{ADE} < d \le d_{BDE} \\ \frac{4}{3}\ln\left(\frac{d+1/3}{2/3-d}\right)\ln\left(\frac{d+2/3}{1-d}\right) \quad d_{BDE} < d \le \frac{1}{3} \\ \frac{4}{3}\ln\left(\frac{2/3-d}{d+1/3}\right)\ln(1-d) - \frac{2}{3}\left[\ln\left(\frac{d+1/3}{1-d}\right)\right]^2 \quad \frac{1}{3} < d \le d_{AC} \\ 1 - \frac{2}{3}\left[\ln\left(d+\frac{1}{3}\right)\right]^2 - [1+\ln(1-d)]^2 - 3\left[1+\frac{2}{3}\ln\left(\frac{2}{3}-d\right)\right]^2 \\ \quad d_{AC} < d \le \frac{2}{3} - e^{-3/2} \\ 1 - \frac{2}{3}\left[\ln\left(d+\frac{1}{3}\right)\right]^2 - [1+\ln(1-d)]^2 \quad \frac{2}{3} - e^{-3/2} < d \le 1 - e^{-1} \\ 1 - \frac{2}{3}\left[\ln\left(d+\frac{1}{3}\right)\right]^2 \quad 1 - e^{-1} < d \le \frac{2}{3} \\ 1 \quad d > \frac{2}{3}, \end{cases}$$

which is plotted in Fig. 12. Dots have been plotted at the breakpoints, with each of the lower four tightly-clustered breakpoints from Table 4 corresponding to

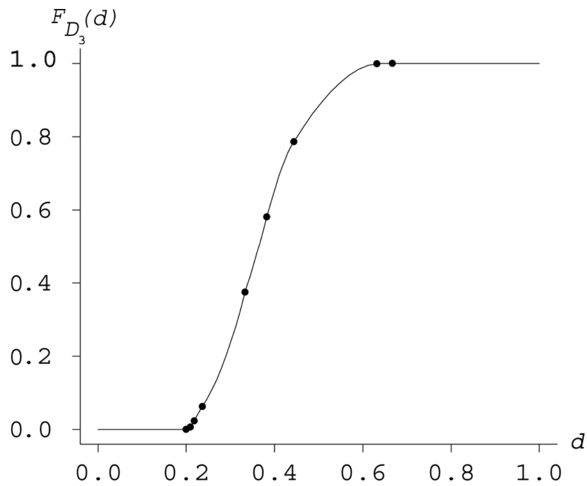

Figure 12. The cdf of D_3.

a horizontal plane intersecting one of the four corners of region E in Fig. 11. Percentiles of this distribution match the tabled values from Durbin (1975). We were not able to establish a pattern between the cdf of D_2 and the cdf of D_3 that might lead to a general expression for any n.

APPL was again used to calculate moments of D_3. The decimal approximations for the mean, variance, skewness, and kurtosis, are, respectively, $E(D_3) \cong 0.3727$, $V(D_3) \cong 0.008804$, $\gamma_3 \cong 0.4541$, and $\gamma_4 \cong 2.6538$. These values were confirmed by Monte Carlo simulation. Although the functional form of the eight-segment PDF of D_3 is too lengthy to display here, it is plotted in Fig. 13, with the only non obvious breakpoint being on the initial nearly-vertical segment at $(d_{BCE}, f_{D_3}(d_{BCE})) \cong (0.2091, 1.5624)$.

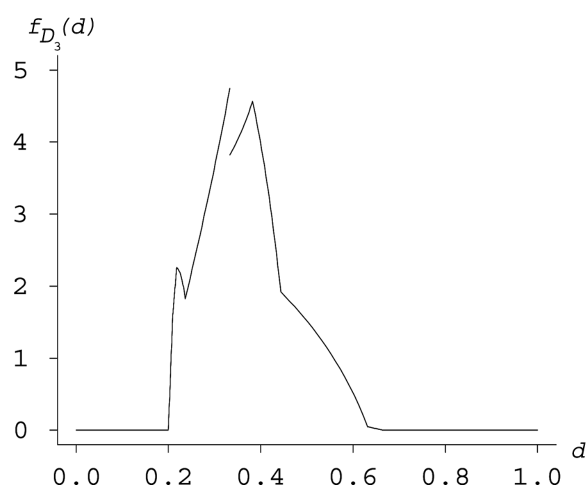

Figure 13. The pdf of D_3.

Acknowledgments

The first author acknowledges summer support from Rose–Hulman Institute of Technology. The second and third authors acknowledge FRA support from the College of William & Mary. The authors also acknowledge the assistance of Bill Griffith, Thom Huber, and David Kelton in selecting data sets for the case studies.

References

Cho, S. K., Spiegelberg-Planer, R. (2002). Country nuclear power profiles. http://www-pub.iaea.org/MTCD/publications/PDF/cnpp2003/CNPP_Webpage/PDF/2002/index.htm Accessed August 1, 2004.

D'Agostino, H., Stephens, M. (1986). *Goodness-of-Fit Techniques*. New York: Marcel Dekker.

Drew, J. H., Glen, A. G., Leemis, L. M. (2000). Computing the cumulative distribution function of the Kolmogorov–Smirnov statistic. *Computational Statistics and Data Analysis* 34:1–15.

Durbin, J. (1975). Kolmogorov–Smirnov tests when parameters are estimated with applications to tests of exponentiality and tests on spacings. *Biometrika* 62:5–22.

Glen, A. G., Evans, D. L., Leemis, L. M. (2001). APPL: a probability programming language. *The American Statistician* 55:156–166.

Hogg, R. V., McKean, J. W., Craig, A. T. (2005). *Mathematical Statistics*. 6th ed. Upper Saddle River, NJ: Prentice–Hall.

Law, A. M., Kelton, W. D. (2000). *Simulation Modeling and Analysis*. 3rd ed. New York: McGraw-Hill.

Lawless, J. F. (2003). *Statistical Models and Methods for Lifetime Data*. 2nd ed. New York: John Wiley & Sons.

Lehmann, E. L. (1959). *Testing Statistical Hypotheses*. New York: John Wiley & Sons.

Lilliefors, H. (1969). On the Kolmogorov–Smirnov test for the exponential distribution with mean unknown. *Journal of the American Statistical Association* 64:387–389.

Marsaglia, G., Tsang, W. W., Wang, J. (2003). Evaluating Kolmogorov's distribution. *Journal of Statistical Software* 8:18. URL: http://www.jstatsoft.org/v08/i18/

Rigdon, S., Basu, A. P. (2000). *Statistical Methods for the Reliability of Repairable Systems*. New York: John Wiley & Sons.

Stephens, M. A. (1974). EDF statistics for goodness of fit and some comparisons. *Journal of the American Statistical Association* 69:730–737.