Fitting a distribution by the first two moments (partial and complete)

Article *in* Computational Statistics & Data Analysis · February 1995

DOI: 10.1016/0167-9473(94)00016-C · Source: RePEc

Haim Shore
Ben-Gurion University of the Negev

Computational Statistics & Data Analysis 19 (1995) 563–577
North-Holland

Fitting a distribution by the first two moments (partial and complete)

Haim Shore
Tel-Aviv University, Tel-Aviv, Israel

Received April 1992
Revised December 1992

Abstract: Given a sample of observations from an unknown population, a common practice to derive distributional representation for the given data is to fit a four-parameter distribution via matching of the first four moments. However, third and fourth sample moments are notorious for their large standard errors, which require sample sizes that in a typical industrial setting are rarely available. In this paper we propose an alternative approach that employs only the first two moments (partial and complete) to fit a certain four-parameter distribution to the given sample data. The fitted distribution is a mixture of two components, where each is a linear transformation of a symmetrically distributed standardized variable. Separate transformations are used for each half of the distribution. Estimation of the parameters is carried out by matching of the mean, the variance, and the first and second partial moments. This fitting procedure is shown to be approximately a least squares solution, that provides good-estimates for the fractiles of the approximated distribution. Moreover, the linear transformations may provide mathematically manageable solutions to stochastic optimization problems (like inventory problems) that would otherwise require complex solution procedures. Some numerical examples and a simulation study attest to the effectiveness of the new approach when sample data are scarce.

Keywords: Approximations; Distribution fitting; Moments; Transformations

1. Introduction

Given a sample of observations from an unknown distribution, a common practice is to fit to the data, via moment matching, a member of a four-parameter family of distributions like the Pearson or the Johnson families. This approach is methodologically valid, and its usefulness is corroborated by many studies demonstrating that different families of distributions sharing the same first four moments exhibit remarkable proximity in the values of their respective

Correspondence to: Haim Shore, Department of Industrial Engineering, Faculty of Engineering, Tel-Aviv University, Tel-Aviv, Israel.

SSDI 0167-9473(94)00016-C

fractiles (refer, for example, to Pearson, Johnson and Burr, 1979). Thus, the actual choice of the distribution to be fitted may become relatively irrelevant as long as the first four moments of the unknown distribution are preserved (given the sample data). Yet sample estimates of high moments tend to have large standard errors, and the related accuracy deteriorates rapidly as we move to higher moments. In particular, the standard errors of the third and the fourth sample moments are high (refer to Stuart and Ord, 1989, p. 338), and require sample sizes that in a typical industrial setting are rarely available.

In this paper, an alternative approach is suggested, where distribution fitting is performed separately for the two halves of the (unknown) distribution. The fitted distribution is a two-component mixture, where each component is a linear transformation of a symmetrically distributed random variable. This special structure enables us to employ moments of order two at most in the fitting procedure, and thus avoid the large standard errors associated with higher moments. Furthermore, the new procedure retains desirable least squares properties.

In the following, we first introduce (in Section 2) the new two-component mixture distribution. Its first four moments are derived, and some special cases are treated. Introducing for the approximating variable the standardized logistic variate, it is shown that the range of variation of the third and the fourth cumulants of the new distribution corresponds to cumulants' values shared by the majority of commonly encountered distributions, including many members of the Pearson, the Johnson and the Burr families. Thus, the new distribution may serve to model a wide variety of distributions while retaining their first four moments.

The formulae for the two-moment fitting procedure are developed and demonstrated in Section 3. Section 4 presents sample results from a simulation study, where the upper "3σ limit" (the 0.99865 fractile) of a process distribution is estimated using a direct sample estimate, a four-parameter fitting based on the Pearson distribution and the new two-moment fitting.

Section 5 demonstrates an application of the new transformations to derive closed form optimal solutions to stochastic models, having input distributions that are only partially specified (by the first two partial and complete moments). A closed form solution for a commonly used inventory model (Hadley-Whitin continuous-review (Q,R) model) is derived and numerically compared to an alternative solution procedure recently published (Lau and Lau, 1993).

Section 6 summarizes the results and gives some conclusions.

2. A mixture distribution that preserves the first four moments

Let X be a standardized random variable (r.v) with distribution function $F(x)$, and let Z be another standardized r.v with a symmetric distribution and distribution function $G(z)$. Let the ith partial moment of Z be denoted by M_i, namely: $M_i = \int_{1/2}^{1} z^i \, \mathrm{d}G(z)$. Let l_i be the i-th cumulant of X.

Let x and z be the two P-th fractiles of the respective r.vs, namely: $F(x) = G(z) = P$.

A four parameter linear transformation of z that approximates x (denote the approximation by $\hat{x}$) is (Shore, 1986):

$$\hat{x} = \begin{cases} A_1 z + B_1, & z < 0 \\ A_2 z + B_2, & z > 0 \end{cases} \quad A_i,\ B_i \in \mathbb{R},\ i = 1, 2. \tag{1}$$

Let $\hat{X}$ be a r.v., the P-th fractile of which is $\hat{x}$. $\hat{X}$ has a mean of: $\hat{\mu} = (A_2 - A_1)M_1 + (1/2)(B_1 + B_2)$. Subtracting $\hat{\mu}$ from $\hat{X}$ to obtain zero-mean we have for the P-th fractile:

$$\hat{x} = \begin{cases} A_1 z + (1/2)B - (A_2 - A_1)M_1, & z < 0 \\ A_2 z - (1/2)B - (A_2 - A_1)M_1, & z > 0, \end{cases} \tag{2}$$

where $B = B_1 - B_2$.

$\hat{X}$ now has a zero-mean three-parameter distribution that may be used to approximate unknown standardized distributions which are partially specified by their first four moments (skewed distributions) or by their mean, variance and fourth cumulant (symmetric distributions).

Let us first assume that the distribution of $\hat{X}$ is symmetric. By observing the odd order moments of $\hat{X}$ (in particular, the third standardized cumulant; refer to Shore (1986) or to eqs. 9 in the sequel), it is easily verified that the only parameter-set for which $\hat{X}$ also has a symmetric distribution fulfills: $A_1 = A_2 = A$. Introducing this solution into (2) we obtain:

$$\hat{x} = \begin{cases} Az + (1/2)B, & z < 0 \\ Az - (1/2)B, & z > 0. \end{cases} \tag{3}$$

The variance and the fourth standardized cumulant are, respectively:

$$\hat{V} = 2A^2 M_2 - 2ABM_1 + [(1/2)B]^2, \tag{4}$$

$$\hat{l}_4 = \left\{2A^4 M_4 - 4A^3 BM_3 + 3A^2 B^2 M_2 - AB^3 M_1 + [(1/2)B]^4\right\}/(\hat{V})^2 - 3, \tag{5}$$

(for details refer to Shore, 1986).

Introducing $\hat{V} = 1$ into (4) and the given l_4 into (5), we obtain a five moment approximation for any symmetric distribution, provided a proper Z is used that yields a feasible solution for eqs. 4 and 5.

Several candidates may serve as Z. First, let Z be the standard normal variable, for which:

$$M_1 = 1/\sqrt{(2\pi)} = 0.3989,\ M_2 = 1/2,\ M_3 = \sqrt{(2/\pi)} = 0.7979,\ M_4 = 3/2.$$

Introducing into (4) and (5) we obtain:

$$\hat{V} = A^2 - 0.7978AB + (1/4)B^2 = 1.0, \tag{4a}$$

$$\hat{l}_4 = \left[3A^4 - 3.1916A^3B + (3/2)A^2B^2 - 0.3989AB^3 + (1/16)B^4\right] - 3. \tag{5a}$$

In the range: $-2.0 < l_4 < 1.7906$, a solution for (4a) and (5a) will always be found. This is rather a limited range of variation, so an alternative Z is investigated. Let Z be the standardized logistic variate:

$$z = (\sqrt{3}/\pi)\ln[P/(1-P)] = 0.5513\ln[P/(1-P)], \tag{6}$$

the partial moments of which are (Johnson and Kotz, 1970, ch. 22):

$$M_1 = 0.3821,\ M_2 = 1/2,\ M_3 = 0.90656,\ M_4 = 2.1000.$$

The relevant equations, corresponding to (4a) and (5a), are:

$$\hat{V} = A^2 - 0.7642AB + (1/4)B^2 = 1.0, \tag{4b}$$

$$\hat{l}_4 = \left[4.2A^4 - 3.6262A^3B + (3/2)A^2B^2 - 0.3821AB^3 + (1/16)B^4\right] - 3. \tag{5b}$$

In the range: $-2.0 < l_4 < 4.554$, a solution for (4b) and (5b) may always be found. This is a satisfactory range since it is able to accommodate the majority of symmetric distributions encountered in practice.

In particular, introducing in (5b): $l_4 = 0$, we obtain a very simple five moment approximation for the standard normal inverse distribution function (see also Shore, 1982):

$$z = -0.4506\ln[(1-P)/P] + 0.2252,\ P > 1/2. \tag{7}$$

This approximation has partial moments (Shore, 1986):

$$M_1 = 0.4249,\ M_2 = 1/2,\ M_3 = 0.7738,\ M_4 = 3/2.$$

Now suppose that $\hat{X}$ is skewed. This may be modeled by using as a solution for (2): $A_1 = A - C$, $A_2 = A + C$, to obtain:

$$\hat{x} = \begin{cases} (A-C)z + (1/2)B - 2CM_1, & z < 0 \\ (A+C)z - (1/2)B - 2CM_1, & z > 0. \end{cases} \tag{8}$$

The resulting variance, third and fourth standardized cumulants are:

$$\hat{V} = \{C^2(2M_2 - 4M_1^2)\} + \{2M_2A^2 - 2ABM_1 + [(1/2)B]^2\},$$

$$\hat{l}_3 = \{6C[A^2(M_3 - 2M_2M_1) - 2AB(0.5M_2 - M_1^2)] + 2C^3[M_3 - 6M_2M_1 + 8M_1^3]\}/\hat{V}^{(3/2)},$$

$$\hat{l}_4 = \left\{K_1C^4 + K_2C^2 + \left[2A^4M_4 - 4A^3BM_3 + 3A^2B^2M_2 - B^3AM_1 + (0.5B)^4\right]\right\} / \hat{V}^2 - 3, \tag{9}$$

where

$$K_1 = 2\left(M_4 - 8M_3M_1 + 24M_2M_1^2 - 24M_1{}^4\right),$$

$$K_2 = 2\left[6A^2\left(M_4 - 4M_3M_1 + 4M_2M_1^2\right) + 6AB\left(-M_3 + 4M_2M_1 - 4M_1^3\right) + (3/2)B^2\left(M_2 - 2M_1^2\right)\right].$$

Note that C is an asymmetry parameter ($C = 0$ for a symmetric $\hat{X}$). Also note that (8) has a point of discontinuity at $z = 0$, that may disrupt the monotony of the transformation, namely: $\hat{x}$ may not be a monotonously increasing function of z, near $z = 0$. To circumvent that, it is suggested that when the transformation is applied (for example, in calculating approximate fractiles), the separating point between the two parts of the transformation should be their intersection point, namely:

$$\hat{x} = \begin{cases} (A-C)z + (1/2)B - 2CM_1, & z < \{B/(2C)\} \quad \text{or:} \quad \hat{x} < \{(AB)/(2C) - (2CM_1)\} \\ (A+C)z - (1/2)B - 2CM_1, & z \geq \{B/(2C)\} \quad \text{or:} \quad \hat{x} \geq \{(AB)/(2C) - (2CM_1)\}. \end{cases} \tag{8a}$$

Introducing into (9): $\hat{V} = 1.0$ and the known l_3 and l_4, and identifying A, B and C, we obtain a simple three-parameter approximation that preserves all first four moments of the approximated variable.

Figure 1 shows a "distribution map" which describes the position of a statistical distribution in the (l_3, l_4) space (refer for example to Hahn and

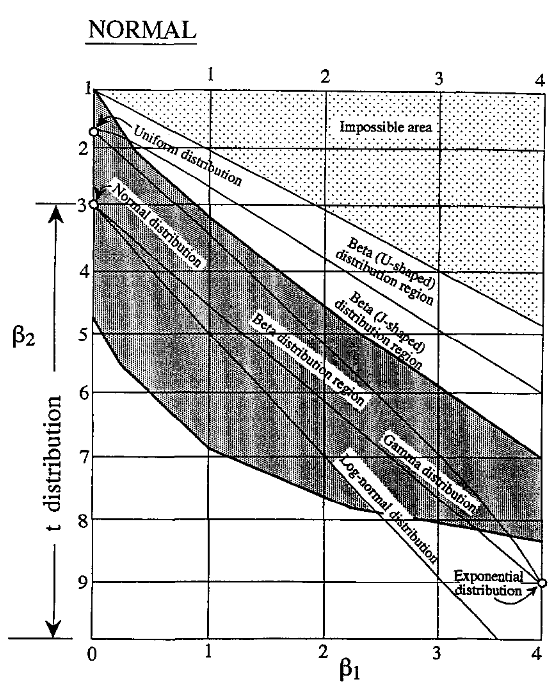

Fig. 1. Range of variation of third and fourth moments of eq. 8 in the (β_1, β_2) coordinate system ($\beta_1 = l_3^2$, $\beta_2 = l_4 + 3$; z used is the standard normal variate).

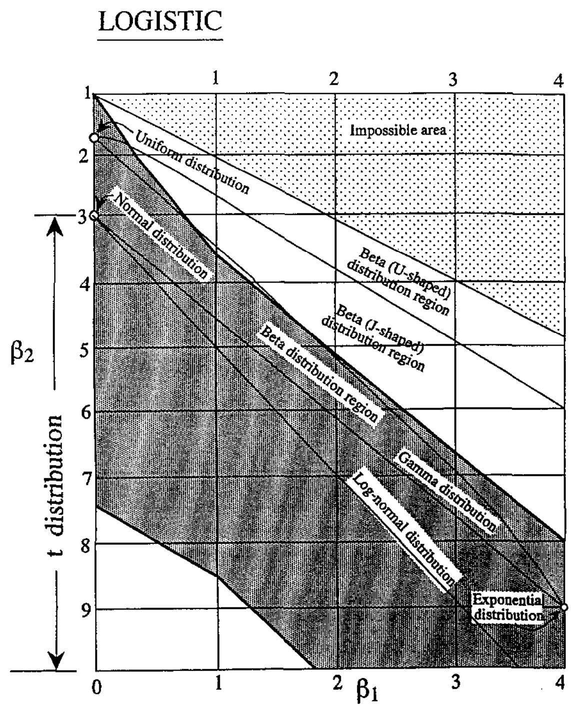

Fig. 2. Range of variation of third and fourth moments of eq. 8 in the (β_1, β_2) coordinate system ($\beta_1 = l_3^2$, $\beta_2 = l_4 + 3$; z used is the logistic variate).

Shapiro (1968, p. 197) or to Stuart and Ord (1989, p. 211 and 236). The "distribution map" describes these positions in a coordinate system, where the horizontal axis denotes: $\beta_1 = l_3^2$, the squared standardized third cumulant, and the vertical axis denotes: $\beta_2 = l_4 + 3$, which is a common measure for kurtosis.

For z given by the standard normal variate, the lined area in Figure 1 shows the ranges of variation of β_1 and β_2, where solutions for eqs. 9 may be found.

Likewise, for z given by eq. 6, Figure 2 shows the corresponding ranges of feasible solutions when β_1 and β_2 are given. In particular:

For $l_3 = 0$: $-2.00 < l_4 < 4.55$; For $l_3 = 0.5$: $-0.87 < l_4 < 4.75$;
For $l_3 = 1.0$: $0.73 < l_4 < 5.66$; For $l_3 = 1.5$: $2.69 < l_4 < 7.82$;
For $l_3 = 2.0$: $5.08 < l_4 < 9.61$.

Figure 2 clearly shows that eq. 2, with z given by (6), is flexible enough to represent a large variety of distributions, as the latter are depicted in the "distribution map".

To demonstrate the accuracy obtained, let Y (the unstandardized variable) have a gamma distribution with parameters: $\alpha = 1/2$, $r = 3/2$, namely: $E(Y) = 3$, $V(Y) = 6$, $l_3 = 1.633$, $l_4 = 4.0$. Using z given by (6) and introducing l_3 and l_4

Table 1
Comparative exact and approximate values of Gamma fractiles (Y_1 - Four moment matching: $A = 1.0233$, $B = 0.2151$, $C = 0.5131$; Y_2 - Two moment matching: $A = 0.9326$, $B = -0.0268$, $C = 0.5164$; Y_3 - The Central Limit Approximation)

P	Y Exact	Y_1	Y_2	Y_3
0.050	0.3518	0.2743	0.3456	−1.0291
0.100	0.5844	0.7891	0.7655	−0.1391
0.400	1.8691	2.0236	1.7726	2.3794
0.600	2.9461	2.6173	2.8596	3.6206
0.700	3.6649	3.5340	3.7242	4.2845
0.800	4.6415	4.6523	4.7789	5.0615
0.900	6.2526	6.3348	6.3657	6.1391
0.95	7.8148	7.8851	7.8278	7.0291
0.99	11.3362	11.3099	11.0579	8.6984
0.995	12.8247	12.7585	12.4241	9.3095
0.99865	15.5182	15.4826	15.0000	10.348

into eqs. 9, we obtain: $A = 1.023$, $B = 0.2151$, $C = 0.5131$. Table 1 presents some comparative values which may help assess the resulting accuracy.

Approximation (8) has a point of discontinuity at $z = 0$. We may wish to relinquish one degree of freedom and preserve continuity at that point. Introducing: $B = 0$, we obtain a zero-mean two-parameter distribution that may be fitted by matching of the variance and the third moment:

$$\hat{x} = \begin{cases} (A - C)z - 2CM_1, & z < 0 \\ (A + C)z - 2CM_1, & z > 0. \end{cases} \tag{10}$$

This solution has variance, third and fourth cumulants equal to, respectively:

$$\begin{aligned} &\hat{V} = \{C^2(2M_2 - 4M_1^2)\} + 2M_2A^2, \\ &\hat{l}_3 = \{6C[A^2(M_3 - 2M_2M_1)] + 2C^3[M_3 - 6M_2M_1 + 8M_1^3]\}/\hat{V}^{(3/2)}, \\ &\hat{l}_4 = \{[K_1C^4 + K_2C^2] + [2A^4M_4]\}/\hat{V}^2 - 3, \quad \text{where} \\ &K_1 = 2(M_4 - 8M_3M_1 + 24M_2M_1^2 - 24M_1^4), \\ &K_2 = 12A^2(M_4 - 4M_3M_1 + 4M_2M_1^2), \end{aligned} \tag{11}$$

(these are eqs. 9, wherein we have introduced: $B = 0$).

Introducing $\hat{V} = 1.0$, and then the resulting expression for A^2 into $\hat{l}_3$ we obtain:

$$\hat{l}_3 = 6C(M_3 - M_1) + 2C^3(12M_1^2M_3 - 4M_1^3 - 2M_3), \tag{12}$$

where we introduced $M_2 = 1/2$. From this expression C may be identified.

For the standardized logistic variate as the approximating Z, (12) yields: $\hat{l}_3 = 3.1468C - 0.8959C^3$, and the extreme values of $\hat{l}_3$ are obtained at: $C =$

± 1.0825, where $\hat{l}_3 = \pm 2.27$. The corresponding value of A will be identified from:

$$A^2 = \hat{V} - C^2(1 - 4M_1^2) = 1 - C^2(1 - 4M_1^2).$$

3. A two-moment fitting procedure

In this section, we develop a procedure to identify the parameters of the transformation by using moment-matching with moments of order two at most. Under a mild approximating assumption, this procedure may be shown to be a least squares solution. Consequently, when fractiles' values are required, the suggested procedure is expected to perform better than the regular four-moment matching, though the third and the fourth sample moments are not preserved in the transformation.

Let us define the ith partial moment of Y, the unstandardized X, at $P = 1/2$ by $M_i(Y)$, and let us rewrite (8) in terms of $\hat{y}$, the unstandardized $\hat{x}$, namely:

$$\hat{y} = \begin{cases} \sigma(A - C)z + \sigma[(1/2)B - 2CM_1] + \mu, & z < 0 \\ \sigma(A + C)z + \sigma[-(1/2)B - 2CM_1] + \mu, & z > 0 \end{cases} \tag{13}$$

or, by simpler notation

$$\hat{y} = \begin{cases} A_1 z + B_1, & z < 0 \\ A_2 z + B_2, & z > 0 \end{cases} \tag{13a}$$

To determine the parameters of (13a), let us match the first and the second partial moments (at $P = 1/2$) for the two components of (13a) with the corresponding partial moments of the approximated distribution. This will also preserve the overall mean, μ, and the standard deviation, σ, of Y.

The following set of four equations is obtained:

$$\begin{aligned} \mu - M_1(Y) &= A_1(-M_1) + (1/2)B_1, \\ M_1(Y) &= A_2 M_1 + (1/2)B_2, \\ E(Y^2) - M_2(Y) &= (\sigma^2 + \mu^2) - M_2(Y) \\ &= (1/2)A_1^2 - 2A_1B_1M_1 + (1/2)B_1^2, \\ M_2(Y) &= (1/2)A_2^2 + 2A_2B_2M_1 + (1/2)B_2^2. \end{aligned} \tag{14}$$

Solving these equations yields

$$\begin{aligned} A_1^2 &= \{(\sigma^2 + \mu^2) - M_2(Y) - 2[\mu - M_1(Y)]^2\} / [(1/2) - 2M_1^2], \\ B_1 &= 2[\mu - M_1(Y) + A_1M_1], \\ A_2^2 &= \{M_2(Y) - 2[M_1(Y)]^2\} / [(1/2) - 2M_1^2], \\ B_2 &= 2[M_1(Y) - A_2M_1]. \end{aligned} \tag{14a}$$

It may be easily verified that under the approximating assumption:

$$\int (yz)\,\mathrm{d}F = \int [y(\hat{y} - B_i)/A_i]\,\mathrm{d}F \cong \int [y(y - B_i)/A_i]\,\mathrm{d}F, \quad (i = 1,2),$$

(14a) is a least squares solution for fitting (13a). Introducing the partial moments of the normal distribution ($M_1 = 1/\sqrt{(2\pi)}$, $M_3 = \sqrt{(2/\pi)}$) into (14a) and using (6) we obtain a two-moment approximation for the standard normal variate (corresponding to the four-moment approximation given by 7):

$$z = -0.5153 \ln[(1-P)/P] + 0.08365, \quad P > 1/2 \tag{7a}$$

To demonstrate the accuracy obtained from the two-moment distribution-fitting procedure, we will employ three distributions:

(A) The exponential distribution (with parameter λ), for which closed form expressions for the first and the second partial moments (at the Pth fractile, y_P) exist. These are, respectively:

$$M_1(y_P) = [(1-P)/\lambda][1 - \ln(1-P)];$$

$$M_2(y_P) = [(1-P)/\lambda^2][2 - 2\ln(1-P) + \ln^2(1-P)].$$

Without loss of generality we will assume: $\lambda = 1$, namely: $\mu = \sigma^2 = 1$. We obtain (for $P = 1/2$, and using the previous notation for the partial moments): $M_1(Y) = 0.8467$; $M_2(Y) = 1.9334$.

Introducing into (14a) we derive for the unstandardized variable, and using z given by (6): $A_1 = 0.3066$, $B_1 = 0.5411$, $A_2 = 1.5504$, $B_2 = 0.5083$.

These values yield for the third and the fourth standardized cumulants of (13): $\hat{l}_3 = 1.763$, $\hat{l}_4 = 4.174$.

(B) The gamma distribution with parameters as aforespecified (Section 2). Partial moments were derived via simulation to yield: $M_1(Y) = 2.3893$, $M_2(Y) = 14.0379$.

The resulting parameters are (for the unstandardized variable): $A_1 = 1.0195$, $B_1 = 2.0005$, $A_2 = 3.5494$, $B_2 = 2.0662$.

These values yield third and fourth standardized cumulants equal to, respectively: $\hat{l}_3 = 1.486$, $\hat{l}_4 = 3.169$.

(C) The Weibull distribution, with parameters (α, β): (2, 10), namely: $E(Y) = 8.8623$, $V(Y) = 21.4602$, $l_3 = 0.6311$, $l_4 = 0.2451$. Partial moments were derived via simulation to yield: $M_1(Y) = 6.2812$, $M_2(Y) = 84.658$.

The resulting parameters are (for the unstandardized variable): $A_1 = 3.1147$, $B_1 = 7.5425$, $A_2 = 5.2585$, $B_2 = 8.5439$.

These values yield third and fourth standardized cumulants equal to, respectively: $\hat{l}_3 = 0.656$, $\hat{l}_4 = 1.020$.

Table 2
Comparative exact and approximate values (two-moment fitting) for the exponential and the Weibull distributions

P	Exponential		Weibull	
	Exact	Approx.	Exact	Approx.
0.05	0.0513	0.0434	2.2648	2.4865
0.10	0.1054	0.1697	3.2459	3.7696
0.20	0.2231	0.3068	4.7237	5.1620
0.40	0.5108	0.4726	7.1472	6.8463
0.60	0.9163	0.8549	9.5723	9.7193
0.80	1.6094	1.6933	12.686	12.563
0.90	2.3026	2.3864	15.174	14.914
0.95	2.9957	3.0252	17.308	17.080
0.975	3.6889	3.6398	19.206	19.165
0.995	5.2983	5.0329	23.018	23.889
0.99865	6.6077	6.1551	25.705	27.694

Tables 1 and 2 display some values obtained from (13). For comparative purposes, Table 1 presents also values obtained from the normal approximation (the parameters A, B and C all refer to the standardized variable).

The reader may compare the skewness of the resulting transformations (eq. 13) with the skewness of the approximated distributions to appreciate the fact that the parameters of (13) are derived using only first and second moments (the same moments used by the Central Limit approximation). In fact, for many distributions that we have examined, it seems that the two moment procedure yields a very close value for the skewness of the approximated distribution for small values of skewness (say, $l_3 < 1$, assuming positive skewness), while for higher values a simple linear transformation yields the correct skewness value. A demonstration to this effect for the gamma distribution is given in Table 3, where the (empirically derived) linear transformation is:

$$\tilde{l}_3 = \begin{cases} \hat{l}_3, & \hat{l}_3 < 1 \\ 1.37\hat{l}_3 - 0.38, & \hat{l}_3 > 1 \end{cases}$$

Table 3
Skewness approximation by the two-moment method (the Gamma distribution)

Gamma ($\alpha = 1/2$) $r =$	Mean $= r/\alpha$	Variance $= r/\alpha^2$	Median	$M_1(X)$	$M_2(X)$	$L_3 = 2/r^{1/2}$	$\hat{L}_3$	$\tilde{L}_3$
0.7561	1.5122	3.0244	0.9196	0.3314	0.7669	2.30	1.94	2.28
1	2	4	1.3863	0.3465	0.7386	2.00	1.752	2.02
2	4	8	3.3567	0.3718	0.6775	1.414	1.323	1.43
3	6	12	5.3481	0.3807	0.6476	1.155	1.106	1.14
4	8	16	7.3441	0.3852	0.6296	1.000	0.974	0.974
7	14	28	13.339	0.3910	0.5988	0.756	0.745	0.745
16	32	64	31.336	0.3955	0.5663	0.500	0.501	0.501
64	128	256	127.33	0.3981	0.5332	0.250	0.252	0.252

Table 4
Results of Monte-Carlo simulation (y_1 - sample fractile, y_2 - fractile estimate derived form the two-moment fitting, y_3 - fractile estimate derived from a four-moment fitting using the Pearson distribution)

Sim.	μ	σ^2	l_3	l_4	$M_1(Y)$	$M_2(Y)$	y_1	y_2	y_3
Gamma									
1	3.04	6.63	1.92	5.19	2.42	14.89	14.2	16.1 *	17.0
2	3.05	6.62	1.80	4.33	2.43	14.95	13.8	16.0 *	16.4
3	3.01	5.94	1.44	2.24	2.40	14.03	11.8	14.8 *	13.7
4	2.99	5.91	1.51	2.65	2.38	13.90	12.1	14.9 *	14.3
5	2.98	5.56	1.29	1.54	2.37	13.52	10.9	14.2 *	12.6
6	2.95	5.21	1.20	1.24	2.34	12.98	10.6	13.7 *	12.1
7	3.07	7.18	2.04	5.73	2.46	15.65	14.9 *	16.8	17.9
8	2.98	5.48	1.29	1.54	2.36	13.39	11.0	14.1 *	12.6
9	3.01	6.12	1.65	3.54	2.39	14.22	12.9	15.3 *	15.2
10	3.02	6.40	1.68	3.52	2.41	14.58	13.0	15.6 *	15.4
Weibull									
1	8.87	21.4	0.55	−0.20	6.29	84.83	20.9	27.4 *	23.5
2	8.83	20.4	0.51	−0.20	6.24	82.94	21.1	26.7 *	23.3
3	8.87	21.5	0.66	0.23	6.29	84.84	22.7	27.9	25.3 *
4	8.87	21.1	0.56	−0.22	6.29	84.51	20.6	27.3 *	23.4
5	8.86	20.6	0.51	−0.30	6.26	83.63	20.1	26.8 *	22.9
6	8.81	20.5	0.46	−0.35	6.24	82.88	19.8	26.6 *	22.8
7	8.90	23.2	0.69	0.38	6.30	85.36	23.2	28.1	26.8 *
8	8.83	20.3	0.49	−0.34	6.25	83.03	20.4	26.6 *	22.7
9	8.87	21.4	0.56	−0.16	6.29	84.67	21.1	27.5 *	23.5
10	8.91	22.3	0.69	0.36	6.32	86.22	23.3	28.5	26.5 *

The 0.99865 fractile (exact): Gamma-15.518; Weibull-25.705.
Note: The most accurate estimate is starred.

and $\tilde{l}_3$ is the approximation for skewness, based on a linear transformation of $\hat{l}_3$, the skewness obtained for (13a) from the two-moment fitting procedure.

4. A simulation study

To study the effect that the new two-moment fitting procedure has on the accuracy of sample estimates of fractiles derived thereof, a simulation study has been conducted using random numbers generated from the Gamma and the Weibull distributions (with the aforespecified parameters) *. For both distributions, samples of 50 observations each had been generated, and sample estimates for the mean, the variance, the third and fourth cumulants and the first

* Simulation was carried out on an 486-DX IBM compatible machine, using ATRISK™ (a product of Palisade Cor.).

and second partial moments were derived. The "unknown" fractile of the origin distribution was thence estimated by three methods:

(A) The sample fractile (based on appropriate interpolation) served as the required fractile estimate.
(B) The two-moment fitting procedure was employed to identify the parameters of (13), using eqs. (14a) and the sample's first two moments (partial and complete).
(C) Based on sample estimates of the first four moments and Pearson tables (adapted from the Biometrika Tables in Clements, 1989), estimates for the required fractiles were derived.

Table 4 presents some typical results obtained for the upper "three sigma limit" (the 0.99865 fractile). The latter is traditionally used in process capability analyses to determine the upper end point of the range of variation of the process distribution. Due to the relatively small sample size (50) and the far tail value of the required fractile, the latter is consistently underestimated by method A and estimated poorly. Referring to methods B and C, the accuracy of estimates derived from the two-moment fitting is in the majority of cases better than that of the four-moment procedure. This is true regardless of whether the distribution fitted by the four-moment procedure belongs to the same family of distributions that has originated the given data (as for the gamma distribution) or otherwise (the Weibull distribution).

5. An application to inventory analysis

Hadley and Whitin (1963) formulated an approximate "backorder" version of the continuous review (Q, R) model that has the periodic cost function:

$$C = KD/Q + h(Q/2 + R - \mu) + \pi DL(R)/Q, \tag{15}$$

where Q (lot size) and R (reorder point) are decision variables, K is the fixed order cost, D the average demand per period, h unit carrying cost per period, μ is the average lead-time demand and π is the unit shortage cost (independent of the shortage duration). $L(R)$ is the expected shortage per cycle (the loss function at R), defined by:

$$L(R) = \int_R^\infty (y - R) f(y)\, \mathrm{d}y = \sigma \int_u^\infty [1 - F(x)]\, \mathrm{d}x, \tag{16}$$

where $F(y)$ is the distribution function of the lead-time demand (with mean μ and standard deviation σ), and u is the standardized R, namely: $u = (R - \mu)/\sigma$. Denoting the loss function of Z at $G(z) = P$ by$L_Z(P)$, a simple approximation for $L(R)$ in terms of $L_Z(P)$ is readily derived from (1) to yield:

$$\hat{L}(R) = \begin{cases} \sigma\{A_1[L_Z(P) - L_Z(0.50)] + A_2[L_Z(0.50)]\} & \\ \quad = \sigma[A_1 L_Z(P) + (A_2 - A_1)L_Z(0.50)], & P < 1/2 \\ \sigma A_2 L_Z(P), & P > 1/2 \end{cases} \tag{17}$$

where A_1 and A_2 refer to the standardized variable (eq. 1).
Let us introduce some z that have explicit expressions for $L_Z(P)$ (refer to Shore (1982 and 1986) for details). First, for:

$$z = -0.4115\{(1-P)/P + \ln[(1-P)/P] - 1\}, \; P > 1/2, \tag{18}$$

the partial moments of which are: $M_1 = 0.4115$, $M_2 = 1/2$, $M_3 = 0.7892$, $M_4 = 1.50$, we have:

$$L_Z(P) = \begin{cases} 0.4115P/(1-P) - z, & P < 1/2 \\ 0.4115(1-P)/P, & P > 1/2, \end{cases} \tag{19}$$

or

$$L_Z(P) = \begin{cases} 0.4115\{1 - \ln[P/(1-P)]\}, & P < 1/2 \\ 0.4115(1-P)/P, & P > 1/2, \end{cases} \tag{19a}$$

where $P = G(z)$.
For z given by (6) we obtain:

$$\begin{aligned} L_Z(P) &= 0.5513\int_P^1 [1-G](\mathrm{d}z/\mathrm{d}G)\,\mathrm{d}G = 0.5513\int_P^1 (1/G)\,\mathrm{d}G \\ &= -0.5513\ln(P) \end{aligned} \tag{20}$$

Introducing into (15) in terms of P we obtain for z given by (6):

$$\begin{aligned} C &= KD/Q + h(Q/2 + R - \mu) + \pi DL(R)/Q \\ &\cong KD/Q + hQ/2 + h\sigma\{A_2 0.5513\ln[P/(1-P)] + B_2\} \\ &\quad + (\pi D/Q)\sigma A_2[-0.5513\ln(P)] \end{aligned} \tag{15a}$$

assuming the optimal P^* is larger than $1/2$.
Differentiating with respect to P and with respect to Q we obtain:

$$\begin{aligned} &\mathrm{d}C/\mathrm{d}P = 0: \quad 1 - P^* = (hQ)/(\pi D), \text{ which is the exact solution.} \\ &\mathrm{d}C/\mathrm{d}Q = 0: \quad (Q^*)^2 = 2DK/h + (2D\pi/h)\sigma A_2 \\ &\qquad\qquad\qquad \times\{-0.5513\ln[1 - (hQ^*)/(\pi D)]\}. \end{aligned} \tag{21}$$

Solving for Q^* and then for P^* the optimal reorder point may be identified by eq. (1).
For z given by (18) we obtain, in a similar manner:

$$\begin{aligned} &\mathrm{d}C/\mathrm{d}P = 0: \quad 1 - P^* = (hQ)/(\pi D), \text{ which is the exact solution.} \\ &\mathrm{d}C/\mathrm{d}Q = 0: \quad (Q^*)^2 = 2DK/h + (2D\pi/h)\sigma A_2 0.4115 \\ &\qquad\qquad\qquad \times[hQ^*/(\pi D - hQ^*)], \end{aligned} \tag{21a}$$

where the second expression may be rewritten as:

$$h(Q^*)^3 - (\pi D)(Q^*)^2 - [1 - (0.4115\pi\sigma A_2)/K](2DK)Q^* + (2D^2K\pi)/h = 0$$

Note, that A_2 refers here to the standardized variable (X).

To assess the effectiveness and accuracy of this procedure we will contrast it with a new procedure recently suggested by Lau and Lau (1993). The latter have presented the following numerical data:

$$D = 1000,\ K = 20,\ h = 1,\ \pi = 30,\ \mu = 80,\ \sigma = 8,$$

and solved it for the Normal and Weibull cases (Table 1, therein). The exact optimal solutions are:

The Normal case: $Q^* = 202.6$; $R^* = 99.76$; $P^* = 0.99325$; $C^* = 222.38$
The Weibull case: $Q^* = 201.4$; $R^* = 95.26$; $P^* = 0.99329$; $C^* = 216.62$

Note, that P^* for the Normal and the Weibull cases have been wrongly swapped in Lau & Lau (the correct values are given here).

Basing their solution procedure on the relative insensitivity of P^* and Q^* to the lead-time demand distribution, Lau and Lau obtain the following optimal solution for the normal case (and similarly we derived the optimal solution for the Weibull case, using Lau & Lau procedure):

The Normal case: $Q^* = 200.092$; $R^* = 99.80$; $P^* = 0.99333$;
$C^* = 222.40$ (deviation of 0.009% from the optimal C)
The Weibull case: $Q^* = 200.092$; $R^* = 95.27$; $P^* = 0.99333$;
$C^* = 216.63$ (deviation of 0.005%)

Alternatively, using the two parameter fitting procedure we obtain:

For the Normal case: Using the partial moments of (18), which is a five moment approximation to the standard normal variate, a two-moment approximation for the standardized normal variate yields (introduce $M_1(Y) = 0.3989$, $M_2(Y) = 1/2$, $M_1 = 0.4115$ into A_2^2 of 14a): $A_2 = 1.061$. Introducing into (21a) yields: $Q^* = 203.5$; $P^* = 0.99322$; $R^* = 99.75$, $C^* = 222.375$. R^* is obtained from a standard normal table.

For the Weibull case: Given the specified mean and standard deviation, we obtain the parameters: $\alpha = 12.153$, $\beta = 83.443$, and by Monte-Carlo simulation: $M_1(Y) = 43.12$, $M_2(Y) = 3725.4$.
Introducing into (14a) we obtain (for the standardized Y): $A_2 = 0.7896$, and from (21a): $Q^* = 202.6$; $P^* = 0.99325$; $R^* = 95.26$, $C^* = 216.62$. R^* is obtained from the inverse distribution function.

Although the two-moment procedure yields more accurate solutions then Lau & Lau's, the improvement is not meaningful (at least not for the given examples). Notwithstanding, the two-moment procedure seems to be methodologically preferable since in applications it will be based on sample estimates of the first two moments (partial and complete) of the underlying unknown distribution. Lau & Lau suggest no statistically valid alternative for the calculation of R^*, once P^* is known.

6. Concluding remarks

Fitting a distribution where partial distributional knowledge of the first two moments only is required has been shown to yield acceptable accuracy that avoids the large sampling errors associated with skewness and kurtosis estimation. Although other methods exist for fitting a distribution (like maximum likelihood methods), the simplicity of the new approach, the small standard errors of its estimates and its desirable least-squares properties seem to render it a preferred alternative in distribution fitting, in general, or in various application areas where sample data are scarce, in particular.

References

[1] Hadley, G., and Whitin, T. (1963), *Analysis of Inventory Systems*, Prentice-Hall, Englewood Cliffs, NJ.

[2] Hahn, G.J. and Shapiro, S.S. (1968), *Statistical methods in Engineering*, John Wiley, NY.

[3] Clements, A. (1989), Process capability calculations for non-normal distributions, *Quality Progress*, **22**(9), 95–100.

[4] Johnson, N.L., and Kotz, S. (1970), *Distributions in statistics*, Houghton-Mifflin, Boston.

[5] Lau, A.H., and Lau, H. (1993), A simple cost minimization procedure for the (Q, R) inventory model: Development and evaluation, *IIE Transactions*. **5**(2), 45–53.

[6] Pearson, E.S., Johnson, N.L., and Burr, I.W. (1979), Comparison of the percentage points of distributions with the same first four moments, chosen from eight different systems of frequency curves. *Communications in Statistics*. B, **9**, 81.

[7] Shore, H. (1982), Simple approximations for the inverse cumulative function, the density function and the loss integral of the normal distribution, *Applied Statistics*, **31**, 108–114.

[8] Shore, H. (1986), Simple general approximations for a random variable and its inverse distribution function based on linear transformations of a nonskewed variate, *Siam Journal on Scientific and Statistical Computing*, **7**, 1–23.

[9] Stuart, A., and Ord J.K. (1987), *Kendall's advanced theory of statistics, V.1: Distribution Theory*. Charles Griffin & Co. Ltd. London.