Chapter 3

Fitting Distributions and Data with the GLD via the Method of Moments

In most practical applications, when constructing a statistical model we do not know the appropriate probability distribution (or do not know it fully). If the appropriate probability distribution is fully known (e.g., if it is known that $X$ follows the normal distribution with mean 5 feet and standard deviation 6 inches), then this distribution should be used in the model. However, if a variable such as the height of females in a certain population is stated in the literature to be normal in distribution with population mean 5 feet and standard deviation 6 inches, very often these are only estimates obtained from a sample. Note that the normal distribution for height would be inappropriate formally since it gives $P(X < 0) > 0$, when we know that $X$ cannot be negative. This need not be a reason to reject the normal model since for this model

$$P(X < 0) = \Phi(-10) = 0.7619855 \times 10^{-23}$$

(National Bureau of Standards (1953)), and a model that comes close to the true value of zero for the probability is acceptable as long as it is used intelligently by realizing that we have a close approximation to the true model, and do not make claims such as “persons of negative height will occur once each $1/(0.7619855 \times 10^{-23}) = 1.3 \times 10^{23}$ people.” While such a claim is absurd, and no reasonable person would think a model bad because an unreasonable person could use it to make such a statement, the adage “What is so uncommon as common sense?” is one we have found to have much truth to it.

Even if we know, by theoretical derivation from reasonable assumptions, for example, that $X$ is normal in distribution, we will often not know its parameters. These will need to be estimated from whatever data is available. Suppose that we have a set of data $X_1, X_2, \ldots, X_n$ that are independent and identically distributed
random variables. If the data is normally distributed, i.e., if $X_i \sim N(\mu, \sigma^2)$ for $i = 1, 2, \ldots, n$, then the mean $\mu$ and the variance $\sigma^2$ are usually estimated, respectively, by

$$\hat{\mu} \equiv \bar{X} = \frac{X_1 + \ldots + X_n}{n}$$

and

$$\hat{\sigma}^2 \equiv s^2 = \frac{\sum_{i=1}^{n} (X_i - \bar{X})^2}{n - 1}.$$ 

Note that a random variable $X$ that is $N(\mu, \sigma^2)$ has

\begin{align*}
\text{(measure of center)} &= E(X) = \mu, \\
\text{(measure of variability)} &= E((X - \mu)^2) = \sigma^2, \\
\text{(measure of skewness)} &= E((X - \mu)^3)/\sigma^3 = 0, \\
\text{(measure of kurtosis)} &= E((X - \mu)^4)/\sigma^4 = 3.
\end{align*}

If a random variable $Y$ has a distribution other than the normal, we might attempt to approximate it by a random variable $X$ that is $N(\mu, \sigma^2)$ for some $\mu$ and $\sigma^2$. We can do this successfully for center $E(Y)$ and variability $\text{Var}(Y) = E((Y - E(Y))^2)$ by choosing $\mu$ and $\sigma^2$ in such a way as to match the center and variability of $Y$ with the same center and variability for $X$. However, after that is done there are no free parameters in the distribution of $X$, and (unless the skewness and kurtosis of $Y$ are, respectively, 0 and 3) we will not be able to match them in $X$. Hence, the normal family of distributions cannot be used to match data successfully unless the data is symmetric (so its skewness is 0) and has tail weight similar to that of the normal (so that its kurtosis is near 3). For this reason, families of distributions with additional parameters are often used, allowing us to match more than the center and the variability of $Y$.

In order to find a moment-based GLD fit to a given dataset $X_1, X_2, \ldots, X_n$, we determine the first four moments ($\bar{X}$, and the second, third, and fourth central moments) of $X_1, X_2, \ldots, X_n$, set these equal to their $\text{GLD}(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ counterparts, and solve the resulting equations for $\lambda_1, \lambda_2, \lambda_3, \lambda_4$.

In Section 3.1 we consider the first four moments of $\text{GLD}(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ distributions and in Section 3.2 we determine the possible values that these moments can attain. Fitting a $\text{GLD}(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ through the method of moments is developed in Section 3.3. Applications of these results for approximating some well-known distributions, and for fitting a $\text{GLD}(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ to a dataset, are developed in Sections 3.4 and 3.5, respectively.

### 3.1 The Moments of the GLD Distribution

In this section we develop expressions for the moments of $\text{GLD}(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ random variables. We start by setting $\lambda_1 = 0$ to simplify this task; next, we
3.1 The Moments of the GLD Distribution

obtain the non-central moments of the GLD($\lambda_1, \lambda_2, \lambda_3, \lambda_4$); and finally, we derive the central GLD($\lambda_1, \lambda_2, \lambda_3, \lambda_4$) moments.

**Theorem 3.1.1.** If $X$ is a GLD($\lambda_1, \lambda_2, \lambda_3, \lambda_4$) random variable, then $Z = X - \lambda_1$ is GLD(0, $\lambda_2, \lambda_3, \lambda_4$).

**Proof.** Since $X$ is GLD($\lambda_1, \lambda_2, \lambda_3, \lambda_4$),

$$Q_X(y) = \lambda_1 + \frac{y^{\lambda_3} - (1 - y)^{\lambda_4}}{\lambda_2},$$

and

$$F_{X - \lambda_1}(x) = P[X - \lambda_1 \leq x] = P[X \leq x + \lambda_1] = F_X(x + \lambda_1). \quad (3.1.2)$$

If we set $F_X(x + \lambda_1) = y$, we obtain

$$x + \lambda_1 = Q_X(y) = \lambda_1 + \frac{y^{\lambda_3} - (1 - y)^{\lambda_4}}{\lambda_2}, \quad x = Q_{X - \lambda_1}(y). \quad (3.1.3)$$

From (3.1.2) we also have $F_{X - \lambda_1}(x) = y$ which with (3.1.3) yields

$$Q_{X - \lambda_1}(y) = x = \frac{y^{\lambda_3} - (1 - y)^{\lambda_4}}{\lambda_2},$$

proving that $X - \lambda_1$ is GLD(0, $\lambda_2, \lambda_3, \lambda_4$).

Having established $\lambda_1$ as a location parameter, we now determine the non-central moments (when they exist) of the GLD($\lambda_1, \lambda_2, \lambda_3, \lambda_4$).

**Theorem 3.1.4.** If $Z$ is GLD(0, $\lambda_2, \lambda_3, \lambda_4$), then $E(Z^k)$, the expected value of $Z^k$, is given by

$$E(Z^k) = \frac{1}{\lambda_2^k} \sum_{i=0}^{k} \left[ \binom{k}{i} (-1)^i \beta(\lambda_3(k - i) + 1, \lambda_4i + 1) \right] \quad (3.1.5)$$

where $\beta(a, b)$ is the beta function defined by

$$\beta(a, b) = \int_0^1 x^{a-1}(1 - x)^{b-1} \, dx. \quad (3.1.6)$$

**Proof.**

$$E(Z^k) = \int_{-\infty}^{\infty} z^k f(z) \, dz = \int_0^1 (Q(y))^k \, dy \quad (3.1.7)$$

$$= \int_0^1 \left( \frac{y^{\lambda_3} - (1 - y)^{\lambda_4}}{\lambda_2} \right)^k \, dy = \frac{1}{\lambda_2^k} \int_0^1 \left( y^{\lambda_3} - (1 - y)^{\lambda_4} \right)^k \, dy.$$
Chapter 3: Fitting Distributions and Data with the GLD via Moments

By the binomial theorem,
\[
(y^\lambda_3 - (1 - y)^\lambda_4)^k = \sum_{i=0}^{k} \left[ \binom{k}{i} (y^\lambda_3)^{k-i} (-1)^{i} (1 - y)^{\lambda_4} \right]^i.
\] (3.1.8)

Using (3.1.8) in the last expression of (3.1.7), we get
\[
E(Z^k) = \frac{1}{\lambda_2^k} \sum_{i=0}^{k} \left[ \binom{k}{i} (-1)^i \int_0^1 y^{\lambda_3(k-i)} (1 - y)^{\lambda_4i} dy \right]
\]
\[
= \frac{1}{\lambda_2^k} \sum_{i=0}^{k} \left[ \binom{k}{i} (-1)^i \beta(\lambda_3(k-i) + 1, \lambda_4i + 1) \right],
\]
completing the proof of the theorem.

Before continuing with our investigation of the GLD(\lambda_1, \lambda_2, \lambda_3, \lambda_4) moments, we note three properties of the beta function that will be useful in our subsequent work.

Properties of the beta function

1. The integral in (3.1.6) that defines the beta function will converge if and only if \(a\) and \(b\) are positive (this can be verified by choosing \(c\) from the \((0, 1)\) interval and considering the integral over the subintervals \((0, c)\) and \((c, 1))\).
2. When \(a\) and \(b\) are positive, \(\beta(a, b) = \beta(b, a)\). Using the substitution \(y = 1 - x\) in the integral for \(\beta(a, b)\) will transform it to the integral for \(\beta(b, a)\).
3. By direct evaluation of the integral in (3.1.6), it can be determined that for \(u > -1\),
\[
\beta(u + 1, 1) = \beta(1, u + 1) = \frac{1}{u + 1}.
\] (3.1.9)

The first of these observations, along with (3.1.5) of Theorem 3.1.4, helps us determine the conditions under which the GLD(\lambda_1, \lambda_2, \lambda_3, \lambda_4) moments exist.

Corollary 3.1.10. The \(k\)-th GLD(\lambda_1, \lambda_2, \lambda_3, \lambda_4) moment exists if and only if \(\lambda_3 > -1/k\) and \(\lambda_4 > -1/k\).

Proof. From Theorem 3.1.1, \(E(X^k)\) will exist if and only if \(E(Z^k) = E((X - \lambda_1)^k)\) exists, which, by Theorem 3.1.4, will exist if and only if
\[
\lambda_3(k-i) + 1 > 0 \text{ and } \lambda_4i + 1 > 0, \quad \text{for } i = 0, 1, \ldots, k.
\]
This condition will prevail if and only if \(\lambda_3 > -1/k\) and \(\lambda_4 > -1/k\).

Since, ultimately, we are going to be interested in the first four moments of the GLD(\lambda_1, \lambda_2, \lambda_3, \lambda_4), we will need to impose the condition \(\lambda_3 > -1/4\).
and \( \lambda_4 > -1/4 \) throughout the remainder of this chapter. The next theorem gives an explicit formulation of the first four centralized GLD\((\lambda_1, \lambda_2, \lambda_3, \lambda_4)\) moments.

**Theorem 3.1.11.** If \( X \) is GLD\((\lambda_1, \lambda_2, \lambda_3, \lambda_4)\) with \( \lambda_3 > -1/4 \) and \( \lambda_4 > -1/4 \), then its first four moments, \( \alpha_1, \alpha_2, \alpha_3, \alpha_4 \) (mean, variance, skewness, and kurtosis, respectively), are given by

\[
\alpha_1 = \mu = E(X) = \lambda_1 + \frac{A}{\lambda_2}, \tag{3.1.12}
\]

\[
\alpha_2 = \sigma^2 = E[(X - \mu)^2] = \frac{B - A^2}{\lambda_2^2}, \tag{3.1.13}
\]

\[
\alpha_3 = E(X - E(X))^3/\sigma^3 = \frac{C - 3AB + 2A^3}{\lambda_2^3\sigma^3}, \tag{3.1.14}
\]

\[
\alpha_4 = E(X - E(X))^4/\sigma^4 = \frac{D - 4AC + 6A^2B - 3A^4}{\lambda_2^4\sigma^4}, \tag{3.1.15}
\]

where

\[
A = \frac{1}{1 + \lambda_3} - \frac{1}{1 + \lambda_4}, \tag{3.1.16}
\]

\[
B = \frac{1}{1 + 2\lambda_3} + \frac{1}{1 + 2\lambda_4} - 2\beta(1 + \lambda_3, 1 + \lambda_4), \tag{3.1.17}
\]

\[
C = \frac{1}{1 + 3\lambda_3} - \frac{1}{1 + 3\lambda_4} - 3\beta(1 + 2\lambda_3, 1 + \lambda_4)
+ 3\beta(1 + \lambda_3, 1 + 2\lambda_4), \tag{3.1.18}
\]

\[
D = \frac{1}{1 + 4\lambda_3} + \frac{1}{1 + 4\lambda_4} - 4\beta(1 + 3\lambda_3, 1 + \lambda_4)
+ 6\beta(1 + 2\lambda_3, 1 + 2\lambda_4) - 4\beta(1 + \lambda_3, 1 + 3\lambda_4). \tag{3.1.19}
\]

**Proof.** Let \( Z \) be a GLD\((0, \lambda_2, \lambda_3, \lambda_4)\) random variable. By Theorem 3.1.1,

\[E(X^k) = E((Z + \lambda_1)^k).\]

We first express \( E(Z^i) \), for \( i = 1, 2, 3, \) and \( 4 \), in terms of \( A, B, C, \) and \( D \). To do this for \( E(Z) \), we use Theorem 3.1.4 to obtain

\[E(Z) = \frac{1}{\lambda_2} (\beta(\lambda_3 + 1, 1) - \beta(1, \lambda_4 + 1)),\]

and from (3.1.9) we get

\[E(Z) = \frac{1}{\lambda_2} \left( \frac{1}{\lambda_3 + 1} - \frac{1}{\lambda_4 + 1} \right) = \frac{A}{\lambda_2}. \tag{3.1.20}\]
Chapter 3: Fitting Distributions and Data with the GLD via Moments

For \( E(Z^2) \) we again use Theorem 3.1.4 and the simplification allowed by (3.1.9) to get

\[
E(Z^2) = \frac{1}{\lambda_2^2} (\beta(2\lambda_3 + 1, 1) - \beta(\lambda_3 + 1, \lambda_4 + 1) + \beta(1, 2\lambda_4 + 1))
\]

\[
= \frac{1}{\lambda_2^2} \left( \frac{1}{2\lambda_3 + 1} - \frac{1}{2\lambda_4 + 1} - 2\beta(\lambda_3 + 1, \lambda_4 + 1) \right) = \frac{B}{\lambda_2^2}. \quad (3.1.21)
\]

Similar arguments, with somewhat more complicated algebraic manipulations, for \( E(Z^3) \) and \( E(Z^4) \) produce

\[
E(Z^3) = \frac{C}{\lambda_3^2} \quad (3.1.22)
\]

\[
E(Z^4) = \frac{D}{\lambda_4^2} \quad (3.1.23)
\]

We now use (3.1.20) to derive (3.1.12):

\[
\alpha_1 = E(X) = E(Z + \lambda_1) = \lambda_1 + E(Z) = \lambda_1 + \frac{A}{\lambda_2}.
\]

Next, we consider (3.1.13):

\[
\alpha_2 = E(X^2) - \alpha_1^2 = E((Z + \lambda_1)^2) - \alpha_1^2
\]

\[
= E(Z^2) + 2\lambda_1 E(Z) + \lambda_1^2 - \alpha_1^2. \quad (3.1.24)
\]

Substituting \( A/\lambda_2 \) for \( E(Z) \) and \( \lambda_1 + A/\lambda_2 \) for \( \alpha_1 \) in (3.1.24) and using (3.1.21), we get

\[
\alpha_2 = E(Z^2) - \frac{A^2}{\lambda_2^2} = \frac{B - A^2}{\lambda_2^2}.
\]

The derivations of (3.1.14) and (3.1.15) are similar but algebraically more involved. ■

Corollary 3.1.25. If \( \alpha_1, \alpha_2, \alpha_3, \alpha_4 \) are the first four moments of \( \text{GLD}(\lambda_1, \lambda_2, \lambda_3, \lambda_4) \), then the first four moments of \( \text{GLD}(\lambda_1, \lambda_2, \lambda_4, \lambda_3) \) will be

\[
\alpha_1 - \frac{2A}{\lambda_2^2}, \quad \alpha_2, \quad -\alpha_3, \quad \alpha_4. \quad (3.1.26)
\]

Proof. The exchange of \( \lambda_3 \) and \( \lambda_4 \) in the expressions for \( A, B, C, \) and \( D \) in (3.1.16) through (3.1.20) changes the signs of \( A \) and \( C \) and leaves \( B \) and \( D \) intact.
Thus, from (3.1.12), the first moment of $\text{GLD}(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ will be $\lambda_1 - A/\lambda_2 = \alpha_1 - 2A/\lambda_2$.

Since $B$ and $A^2$ are not affected by the exchange of $\lambda_3$ and $\lambda_4$, from (3.1.13), $\alpha_2$ will be the second moment of $\text{GLD}(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$. $C$, $A$, and $A^3$ of (3.1.14) all change signs when $\lambda_3$ and $\lambda_4$ are switched, making $-\alpha_3$ the third moment of $\text{GLD}(\lambda_1, \lambda_2, \lambda_4, \lambda_3)$. Since $D$, $AC$, $A^2B$, and $A^4$ of (3.1.15) are not affected by the exchange of $\lambda_3$ and $\lambda_4$, $\text{GLD}(\lambda_1, \lambda_2, \lambda_4, \lambda_3)$ will have $\alpha_4$ for its fourth moment.

3.2 The $(\alpha_3^2, \alpha_4)$-Space Covered by the GLD Family

If a random variable $Y$ has a distribution other than the GLD, we might try to approximate it by a random variable $X$ that is $\text{GLD}(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ for some $\lambda_1, \lambda_2, \lambda_3, \lambda_4$. Suppose that the first four moments of $Y$ are $\alpha_1 = \mu$, $\alpha_2 = \sigma^2$, $\alpha_3$, and $\alpha_4$. If we can choose $\lambda_3$, $\lambda_4$ so that a $\text{GLD}(0, 1, \lambda_3, \lambda_4)$ has third and fourth moments $\alpha_3$ and $\alpha_4$, then we can let $\lambda_1$ and $\lambda_2$ be solutions of the equations

$$\mu = \lambda_1 + \frac{A}{\lambda_2}, \quad \sigma^2 = \frac{B - A^2}{\lambda_2^2}.$$  \hspace{1cm} (3.2.1)

It follows from Theorem 3.1.11 that the resulting $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ specify a GLD with the desired first four moments. Here, we note that $A, B, C, D$ are functions only of $\lambda_3, \lambda_4$, and that (3.2.1) can be solved for any $\mu$ and any $\sigma^2 > 0$. We have, therefore, established the following consequence of Theorem 3.1.11.

**Corollary 3.2.2.** The $\text{GLD}(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ can match any first two moments $\mu$ and $\sigma^2$, and some third and fourth moments $\alpha_3$ and $\alpha_4$.

The larger the set of $(\alpha_3, \alpha_4)$ that the $\text{GLD}(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ can generate, the more useful the $\text{GLD}(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ family will be in fitting a broad range of datasets and approximating a variety of other random variables. So we next consider the spectrum of values that $\alpha_3$ and $\alpha_4$ can attain. From Corollary 3.1.25, we know that if $(\alpha_3, \alpha_4)$ can be attained, then so can $(\alpha_3, \alpha_4)$ (by switching $\lambda_3$ and $\lambda_4$), allowing us to consider the $(\alpha_3^2, \alpha_4)$-space associated with the $\text{GLD}(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$.

Figures 3.2–1, 3.2–2, and 3.2–3 show the $(\alpha_3^2, \alpha_4)$ contour plots for $(\lambda_3, \lambda_4)$ from Regions 3, 4, and 5 and 6, respectively (recall that these regions were defined in Section 2.2 and illustrated in Figure 2.2–1). The curves in Figure 3.2–1 are associated with a sequence of values of $\lambda_4$, with $\lambda_3$ ranging on the interval $(0, 15)$ for each of these choices of $\lambda_4$. For example, the curve labeled “0.02” is obtained by plotting the $(\alpha_3^2, \alpha_4)$ pairs when $\lambda_4$ is set to 0.02 and $\lambda_3$ is taken from the interval $(0, 15)$. All the curves of Figure 3.2–1 are obtained in a similar manner.
with $\lambda_1 = 0.02, 0.07, 0.12, \ldots, 0.52, 0.6, 0.7, \ldots, 1.0$ and $\lambda_3$ from the interval $(0, 15)$.

The construction of Figure 3.2–2 is similar to that of Figure 3.2–1, with $(\lambda_3, \lambda_4)$ taken from Region 4. Note that in this case we must have

$$-1/4 < \lambda_3 < 0 \quad \text{and} \quad -1/4 < \lambda_4 < 0$$

since otherwise (see Corollary 3.1.10) $\alpha_3$ or $\alpha_4$ or both may not exist. Some of the $\lambda_4$ values associated with these curves are given in Figure 3.2–2; the other values of $\lambda_4$ are $-0.0125, -0.025, \ldots, -0.075$. In Figure 3.2–3 the roles of $\lambda_3$ and $\lambda_4$ are reversed in the sense that $\lambda_3$ is fixed (to the values shown in the figure) and $\lambda_4$ is allowed to range upward (in the case of Region 5) from the boundary that defines the region (see Theorem 2.2.33).

Figure 3.2–4 gives a comprehensive view of the connection between the GLD $(\alpha_2^2, \alpha_4)$-space and the regions of $(\lambda_3, \lambda_4)$. The area marked “Impossible Region” is where $\alpha_4 \leq 1 + \alpha_3^2$, an impossibility since the inequality

$$\alpha_4 > 1 + \alpha_3^2 \quad (3.2.3)$$

holds for all distributions (it is less well-known than the inequality $E(X^2) \geq \mu^2$, which follows from $\text{Var}(X) = E(X^2) - \mu^2 \geq 0$). Moreover, since $\alpha_4 > 1 + \alpha_3^2$ always holds, $\alpha_4$ necessarily exceeds 1. This result, given in some classical books, is well-known in the field of distribution fitting but is not covered in most texts on probability and statistics. Moreover, we are not aware of any texts that give a simple proof. For a brief indication of how an advanced proof may proceed, see Kendall and Stuart (1969), p. 92, Exercise 3.19. We now state and prove this result.

**Theorem 3.2.4.** For any r.v. $X$ for which these moments exist, $\alpha_4 > 1 + \alpha_3^2$. (Note that equality is not possible for any continuous distribution.)

**Proof.** Since

$$\alpha_i = \frac{E(X - E(X))^i}{\sigma_X^i} = E \left( \frac{X - E(X)}{\sigma_X} \right)^i,$$

we can assume without loss of generality that $E(X) = 0$ and $\sigma_X = 1$, so that $E(X^2) = 1$ (because $\alpha_i$ involves only $X^* = (X - E(X))/\sigma_X$, which has $E(X^*) = 0$ and $E(X^*^2) = 1$). Now the Schwarz inequality (see, for example, Dudewicz and Mishra (1988), p. 240, Theorem 5.3.23) says that for any r.v.s, $X$ and $Y$, for which the expectation exists, $(E(XY))^2 \leq E(X^2)E(Y^2)$. If we take the two r.v.s in the Schwarz inequality to be $X$ and $X^2 - 1$, then

$$\left( E(X(X^2 - 1)) \right)^2 \leq E(X^2)E((X^2 - 1)^2)$$

$$\left( E(X^3 - X) \right)^2 \leq 1 \cdot E(X^4 - 2X^2 + 1)$$

$$\left( E(X^3) - 0 \right)^2 \leq E(X^4) - 2 + 1$$

$$\alpha_3^2 \leq \alpha_4 - 1.$$
3.2 The \((\alpha_3^2, \alpha_4)\)-Space Covered by the GLD Family

Figure 3.2–1. \((\alpha_3^2, \alpha_4)\)-space generated by \((\lambda_3, \lambda_4)\) from Region 3 (see (2.2.13)).

Figure 3.2–2. \((\alpha_3^2, \alpha_4)\)-space generated by \((\lambda_3, \lambda_4)\) from Region 4 (see (2.2.14)).

Figure 3.2–3. \((\alpha_3^2, \alpha_4)\)-space generated by \((\lambda_3, \lambda_4)\) from Regions 5 and 6 (defined in Theorem 2.2.33).
The proof will be complete when we show that equality is not possible for any continuous distribution.

The Schwarz inequality also asserts that equality occurs if and only if for some constant, \( a \), we have \( Y = aX \). In our case, this implies that equality holds if and only if \( X^2 - 1 = aX \) or

\[
X = \frac{a \pm \sqrt{a^2 + 4}}{2}.
\]

Thus, we can have equality if and only if \( X \) is a r.v. that takes on only the two values

\[
\frac{a - \sqrt{a^2 + 4}}{2} \quad \text{and} \quad \frac{a + \sqrt{a^2 + 4}}{2}.
\]

Suppose these values are attained with probabilities \( p \) and \( 1 - p \), respectively \( (0 \leq p \leq 1) \). Then \( E(X) = 0 \) implies

\[
p = \frac{1}{2} \left( 1 + \frac{a}{\sqrt{a^2 + 4}} \right).
\]  
(3.2.5)

Since \( 0 \leq p \leq 1 \), \( -1 \leq a/\sqrt{a^2 + 4} \leq 1 \). And \( E(X^2) = 1 \) also implies (3.2.5). So

\[
\alpha_3^2 = \alpha_4 - 1 \quad \text{if and only if for some } a
\]

\[
P \left( X = \frac{a - \sqrt{a^2 + 4}}{2} \right) = \frac{1}{2} \left( 1 + \frac{a}{\sqrt{a^2 + 4}} \right)
\]
3.2 The \((\alpha^2_3, \alpha_4)\)-Space Covered by the GLD Family

and

\[
P \left( X = \frac{a + \sqrt{a^2 + 4}}{2} \right) = \frac{1}{2} \left( 1 - \frac{a}{\sqrt{a^2 + 4}} \right).
\]

For general \(a\), \((\alpha_3, \alpha_4) = (a, a^2 + 1)\). When \(a = 0\), \(P(X = -1) = 0.5 = P(X = +1)\) and \((\alpha_3, \alpha_4) = (0, 1)\).

Immediately above the Impossible Region in Figure 3.2–4 there is a narrow “sliver” marked by “X.” The GLD\((\lambda_1, \lambda_2, \lambda_3, \lambda_4)\) does not produce \((\alpha^2_3, \alpha_4)\) in this region. We saw in Figures 3.2–1, 3.2–2, and 3.2–3 that Regions 3, 4, 5, and 6 cannot yield points in area X. That Regions 1 and 2 also cannot follows from the fact that \(\lambda_3 \leq -1\) (Region 1) and \(\lambda_4 \leq -1\) (Region 2) violate the conditions \(\lambda_3 > -0.25\) and \(\lambda_4 > -0.25\) needed for the third and fourth moments to exist (see Corollary 3.1.10). Other distributions (e.g., the beta distribution) do have their \((\alpha^2_3, \alpha_4)\) in this area. In Chapter 4 we give an extension of the GLD that covers this portion of \((\alpha^2_3, \alpha_4)\)-space. (While area X of Figure 3.2–4 may look “small,” it is important in a variety of applications.)

The remaining portions of Figure 3.2–4 are marked with “R3,” designating that \((\lambda_3, \lambda_4)\) has to be chosen from Region 3 for this portion of \((\alpha^2_3, \alpha_4)\)-space, and with “R4, R5, R6,” designating that \((\lambda_3, \lambda_4)\) is to be chosen from one of Regions 4, 5, or 6 to generate \((\alpha^2_3, \alpha_4)\) in this area. The boundaries between the various portions of Figure 3.2–4 are drawn reasonably accurately, except for the boundaries that enclose the “R4, R5, R6” area; these are rough approximations obtained through numeric computations from the curves in Figure 3.2–3.

The \((\alpha^2_3, \alpha_4)\)-space covered by the GLD already includes the moment combinations of such distributions as the uniform, Student’s \(t\), normal, Weibull, gamma, lognormal, exponential, and some beta distributions, among others. Thus, it is a rich class in terms of moment coverage. To put this in context, we show in Figure 3.2–5 the \((\alpha^2_3, \alpha_4)\) pairs associated with a number of well-known distributions. The shaded region is the region covered by the \((\alpha^2_3, \alpha_4)\) pairs of the GLD

In Figure 3.2–5, the lines that are designated by “W,” “L-N,” “G,” and “S” (the latter refers to the line defined by \(\alpha^2_3 = 0\)) show the \((\alpha^2_3, \alpha_4)\) pairs for the Weibull, lognormal, gamma, and Student’s \(t\) distributions, respectively. The area designated by “B E T A  R E G I O N” shows the \((\alpha^2_3, \alpha_4)\) points that can be produced by the beta distribution. This region extends from the Impossible Region to slightly beyond the line marked for the lognormal distribution. The point designated by a small square with label “u” and located at \((0, 1.8)\) represents the uniform distribution; the point at \((0, 3)\) labeled with “n” represents the normal distribution, \(N(\mu, \sigma^2)\); and the point located at \((4, 9)\) and labeled with “e” gives the \((\alpha^2_3, \alpha_4)\) point associated with the exponential distribution.
3.3 Fitting the GLD through the Method of Moments

As stated at the beginning of the chapter, our intention is to fit a GLD to a dataset by equating \( \alpha_1, \alpha_2, \alpha_3, \alpha_4 \) to \( \hat{\alpha}_1, \hat{\alpha}_2, \hat{\alpha}_3, \hat{\alpha}_4 \), the sample statistics corresponding to \( \alpha_1, \alpha_2, \alpha_3, \alpha_4 \), and solving the equations for \( \lambda_1, \lambda_2, \lambda_3, \lambda_4 \). For a dataset \( X_1, X_2, \ldots, X_n \), the sample moments corresponding to \( \alpha_1, \alpha_2, \alpha_3, \alpha_4 \) are denoted \( \hat{\alpha}_1, \hat{\alpha}_2, \hat{\alpha}_3, \hat{\alpha}_4 \) and are defined by

\[
\hat{\alpha}_1 = \bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i / n, \tag{3.3.1}
\]

\[
\hat{\alpha}_2 = \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X})^2 / n, \tag{3.3.2}
\]

\[
\hat{\alpha}_3 = \frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X})^3 / (n \hat{\sigma}^3), \tag{3.3.3}
\]

\[
\hat{\alpha}_4 = \frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X})^4 / (n \hat{\sigma}^4). \tag{3.3.4}
\]

These are not the maximum likelihood estimators (those would have some \( n \)s replaced by \( n - 1 \)), but correspond to method-of-moments estimators.
3.3 Fitting the GLD through the Method of Moments

Solving the system of equations

\[ \alpha_i = \hat{\alpha}_i \quad \text{for } i = 1, 2, 3, 4 \]  (3.3.5)

for \( \lambda_1, \lambda_2, \lambda_3, \lambda_4 \) is simplified somewhat by observing that \( A, B, C, D \) of (3.1.16) through (3.1.19) are free of \( \lambda_1 \) and \( \lambda_2 \), and \( \lambda_2 \) drops out of (3.1.14) and (3.1.15) because (see (3.1.13)) \( \lambda_2 \sigma^i = (B - A^2)^{i/2} \) for \( i = 3 \) and \( 4 \). Thus, \( \alpha_3 \) and \( \alpha_4 \) depend only on \( \lambda_3 \) and \( \lambda_4 \). Hence, if \( \lambda_3 \) and \( \lambda_4 \) can be obtained by solving the subsystem

\[ \alpha_3 = \hat{\alpha}_3 \quad \text{and} \quad \alpha_4 = \hat{\alpha}_4 \]  (3.3.6)

of two equations in the two variables \( \lambda_3 \) and \( \lambda_4 \), then using (3.1.13) and (3.1.12) successively will yield \( \lambda_2 \) and \( \lambda_1 \).

Unfortunately, (3.3.6) is complex enough to prevent exact solutions, forcing us to appeal to numerical methods to obtain approximate solutions. Algorithms for finding numerical solutions to systems of equations such as (3.3.6) are generally designed to “search” for a solution by checking if an initial set of values \( (\lambda_3 = \lambda_3^*, \lambda_4 = \lambda_4^*) \) in the case of (3.3.6)) can be considered an approximate solution. This determination is made by checking if

\[ \max(|\alpha_3 - \hat{\alpha}_3|, |\alpha_4 - \hat{\alpha}_4|) < \epsilon, \]  (3.3.7)

when \( \lambda_3 = \lambda_3^* \) and \( \lambda_4 = \lambda_4^* \). The positive number \( \epsilon \) represents the accuracy associated with the approximation; if it is determined that the initial set of values \( \lambda_3 = \lambda_3^*, \lambda_4 = \lambda_4^* \) does not provide a sufficiently accurate solution, the algorithm searches for a better choice of \( \lambda_3 \) and \( \lambda_4 \) and iterates this process until a suitable solution is discovered (i.e., one that satisfies (3.3.7)). In algorithms of this type there is no assurance that the algorithm will terminate successfully nor that greater accuracy will be attained in successive iterations. Therefore, such searching algorithms are designed to terminate (unsuccessfully) if (3.3.7) is not satisfied after a fixed number of iterations.

If table values of approximate solutions to (3.3.11) are readily available, an alternate grid-based algorithm can be used. In this case, the algorithm first conducts a “table lookup” to determine values of \( \lambda_{3a}, \lambda_{3b}, \lambda_{4a}, \lambda_{4b} \), so that the desired solution (or at least some solution) satisfies \( \lambda_{3a} \leq \lambda_3 \leq \lambda_{3b} \) and \( \lambda_{4a} \leq \lambda_4 \leq \lambda_{4b} \). Next, an \( n \times n \) grid is constructed on the rectangle \([\lambda_{3a}, \lambda_{3b}] \times [\lambda_{4a}, \lambda_{4b}]\) and for every point of that grid, \( \max(|\hat{\alpha}_3 - \alpha_3|, |\hat{\alpha}_4 - \alpha_4|) \) is evaluated. The algorithm returns the \((\lambda_3, \lambda_4)\) point on the grid that produces the smallest \( \max(|\hat{\alpha}_3 - \alpha_3|, |\hat{\alpha}_4 - \alpha_4|) \). Karian and Dudewicz (2007) describe how this seemingly inefficient algorithm can be improved by repeatedly “zooming in” and reinitializing the grid.

3.3.1 Fitting through Direct Computation

The outcome of searching algorithms (success or failure, and in the former case a particular solution) depends on the equations themselves, the \( \epsilon \) of (3.3.7), the
maximum number of iterations allowed, and the initial starting point for the search. Such algorithms usually have built-in specifications for \( \epsilon \) and the maximal number of iterations, leaving the choice of the starting point as the only real option for the user. To get some insight into where to look for solutions (i.e., how to choose a starting point), consider a specific case where

\[ \hat{\alpha}_1 = 0, \quad \hat{\alpha}_2 = 1, \quad \hat{\alpha}_3 = 0.025, \quad \hat{\alpha}_4 = 2. \]  

(3.3.8)

It is clear from Figures 3.2–1, 3.2–2, and 3.2–3 that the only region of \((\lambda_3, \lambda_4)\)-space that can produce a solution is Region 3 (this can also be observed from Figure 3.2–4). The equation \( \alpha_4 = 2 \), represented by a curve in \((\lambda_3, \lambda_4)\)-space, is shown in Figure 3.3–1. The \((\lambda_3, \lambda_4)\) points on this curve satisfy the second equation; hence, they represent potential solutions to the fourth equation in (3.3.8). The actual solutions will also have to be on the curve specified by \( \alpha_3^2 = 0.025 \). We see from the intersection of the curves of Figure 3.3–1 with the contour curves of \( \alpha_3^2 = 0.025 \) (shown in Figure 3.3–2), that there seem to be four solutions with \((\lambda_3, \lambda_4)\) roughly

\[ (0.03, 0.75), \ (0.8, 0.5), \ (0.95, 0.1), \ (3.25, 2.25). \]  

(3.3.9)

There are also four additional solutions,

\[ (0.75, 0.03), \ (0.5, 0.8), \ (0.1, 0.95), \ (2.25, 3.25), \]

when \( \lambda_3 \) and \( \lambda_4 \) are exchanged. The symmetry of solutions about the \( \lambda_3 = \lambda_4 \) line is due to the presence of two values of \( \hat{\alpha}_3 = \pm \sqrt{0.025} \) (the rough solutions of (3.3.9) are associated with \( \hat{\alpha}_3 = +\sqrt{0.025} \) and the other four with \( \hat{\alpha}_3 = -\sqrt{0.025} \)).
3.3 Fitting the GLD through the Method of Moments

In practice, of course, we have either \( \hat{\alpha}_3 = +\sqrt{0.025} \) or \( \hat{\alpha}_3 = -\sqrt{0.025} \) and then find four solutions for the appropriate case.

The R procedure \texttt{FindLambdasM}, that is described in Appendix A, was devised specifically to produce solutions to (3.3.5). This program uses the table lookup and grid-based search described in the previous section. The only argument of \texttt{FindLambdasM} is the vector \((\hat{\alpha}_1, \hat{\alpha}_2, \hat{\alpha}_3, \hat{\alpha}_4)\) and it returns the \(\lambda_1, \lambda_2, \lambda_3, \lambda_4\) of the fitted distribution. If such a solution cannot be found with \(\max(|\hat{\alpha}_3 - \alpha_3|, |\hat{\alpha}_4 - \alpha_4|) < 10^{-5}\), then \texttt{FindLambdasM} returns 0 0 0 0. More detailed information about \texttt{FindLambdasM}, as well as the other programs included with this book is given in Appendix A.

For the illustration at hand, the use of

\[
> A \leftarrow c(0, 1, \sqrt{0.025}, 2)
> \text{FindLambdasM}(A)
\]

produces

\[
[1] 2.562450e-01 4.992390e-01 8.014867e-01 4.640761e-01 2.478080e-08
\]

which corresponds to the second \((\lambda_3, \lambda_4) = (0.8, 0.5)\) point of (3.3.9). We designate this fit by

\[
\text{GLD}_2(0.2562, 0.4992, 0.8015, 0.4641)
\]

and note that \(\max(|\hat{\alpha}_3 - \alpha_3|, |\hat{\alpha}_4 - \alpha_4|) = 2.5 \times 10^{-8}\).

To search for the other solutions we use the R program \texttt{RefineSearchGLDM}. This program needs 5 arguments: the vector \((\hat{\alpha}_1, \hat{\alpha}_2, \hat{\alpha}_3, \hat{\alpha}_4)\), the minimum and maximum values, represented as a vector, of \(\lambda_3\) to be used in the search for a solution, a similar vector to be used for \(\lambda_4\), the number of grid partitions to used during the search, and the number of iterations where these partitions are to
be applied. *RefineSearchGLDM* returns the the $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ of the fit with the smallest value of $\max(|\hat{\alpha}_3 - \alpha_3|, |\hat{\alpha}_4 - \alpha_4|)$ and the value of $\max(|\hat{\alpha}_3 - \alpha_3|, |\hat{\alpha}_4 - \alpha_4|)$.

To find a solution that is close to the first $(\lambda_3, \lambda_4) = (0.03, 0.75)$ point given in (3.3.9), we use

> RefineSearchGLDM(A,c(0,0.1),c(0.5,1),50,4)

and get

[1] -1.332727e+00 2.919003e-01 2.981683e-02 7.181446e-01 2.548631e-07

indicating the fit

\[
\text{GLD}_1(-1.3327, 0.2919, 0.02982, 0.7181)
\]

and $\max(|\hat{\alpha}_3 - \alpha_3|, |\hat{\alpha}_4 - \alpha_4|) = 2.5 \times 10^{-7}$. Similar searches for the remaining, third and fourth, $(\lambda_3, \lambda_4)$ points of (3.3.9) give the fits

\[
\text{GLD}_3(1.1896, 0.34639, 0.9384, 0.07762)
\]

and

\[
\text{GLD}_4(0.1267, 0.5357, 3.2257, 2.2835).
\]

The p.d.f.s of these GLD fits are shown in Figure 3.3–3 where the GLD$_i$ p.d.f. is labeled with $(i)$. In GLD$_2$, GLD$_3$ and GLD$_4$, $\lambda_3 > \lambda_4$ but in GLD$_1$, $\lambda_3 < \lambda_4$.

Note that had the original $\hat{\alpha}_3^2$ and $\hat{\alpha}_4$ been 0.5 and 2, respectively, we would easily determine from Figures 3.2–1, 3.2–2, and 3.2–3 (or, with some difficulty from Figure 3.2–4) that (3.3.6) would not have any solutions. This can be seen even more convincingly in Figure 3.3–4 which shows that the contour curves of the two equations do not intersect when $(\lambda_3, \lambda_4)$ is in Region 3.

Figure 3.3–5 shows a family of contour curves for Region 3 for $\alpha_4$ with values

1.825, 1.85, 1.9, 2, 2.1, 2.25, 2.5, 2.75, 3, 3.5, 4.

The curve associated with $\alpha_4 = 1.825$ consists of the innermost oval and the lowest branches along the $\lambda_3$ and $\lambda_4$ axes, the curve for $\alpha_4 = 1.85$ consists of the next larger oval and the next higher branches along the axes, and so on. Figure 3.3–6 gives contour curves for Region 3 for $\alpha_3^2$ with values

0.005, 0.01, 0.015, 0.025, 0.05, 0.1, 0.2, 0.3, 0.5, 0.75, 1, 1.25, 1.5, 2.

The curve closest to the line $\lambda_3 = \lambda_4$ (not shown) and on either side of this line is associated with $\alpha_3^2 = 0.005$ and subsequent curves moving away from $\lambda_3 = \lambda_4$ represent increasing values of $\alpha_3^2$. The “dense” set of curves in the lower left corner are branches of the curves in the larger portion of Figure 3.3–6. The one farthest from the origin is associated with $\alpha_3^2 = 0.005$ and is actually connected with the rest of the curve for $\alpha_3^2 = 0.005$. As $\alpha_3^2$ increases the curves become
3.3 Fitting the GLD through the Method of Moments

Figure 3.3–3. Four GLD fits for the \( \hat{\alpha}_1, \hat{\alpha}_2, \hat{\alpha}_3, \hat{\alpha}_4 \) specified by (3.3.8).

disconnected and move closer to the origin. These two families of curves (the \( \alpha_4 \) curves of Figure 3.3–5 and the \( \alpha_3^2 \) curves of Figure 3.3–6) are shown in Figure 3.3–7 and provide a rough guide for determining initial searching points when Region 3 solutions are sought.

The contour plots for Region 4 with \( \alpha_4 \) taking values

\[ 6, 6.5, 7, 8, 9, 10, 12, 14, 17, 20, 25, 30, 45, 65 \]

and \( \alpha_3^2 \) taking values

\[ 0.2, 0.4, 0.7, 1, 1.5, 2, 3, 5, 7, 10, 13, 17 \]

are given in Figure 3.3–8. The curves that are open near the origin are associated with \( \alpha_4 \) and those that are open away from the origin are the curves for \( \alpha_3^2 \). In both cases, the smallest values (of \( \alpha_4 \) or \( \alpha_3^2 \)) produce curves that are closest to \( \lambda_3 = \lambda_4 \) with the curves moving away from this line with increasing values. Figure 3.3–8 not only provides starting points for the search of solutions to (3.3.6), but it also indicates that, if a solution exists, it will be unique, except for the interchange of \( \lambda_3 \) and \( \lambda_4 \).

It seems from Figure 3.3–8 that there should be a unique (up to symmetry) GLD fit with \( (\lambda_3, \lambda_4) \) from Region 4 for the \( \hat{\alpha}_1, \hat{\alpha}_2, \hat{\alpha}_3, \hat{\alpha}_4 \) specified in (3.3.8). Moreover, the \( (\lambda_3, \lambda_4) \) associated with this fit should be close to the origin. Several attempts made through \texttt{RefineSearchGLDM} to find a solution in Region 4 failed (i.e., gave large values of \( \max(|\hat{\alpha}_3 - \alpha_3|, |\hat{\alpha}_3 - \alpha_3|) \)), but gave an indication that if there were to be a solution, it would have to be close to the origin. Eventually,

\[ \texttt{RefineSearchGLDM}(A, c(-0.0002, 0), c(-0.0002, 0), 100, 4) \]
Chapter 3: Fitting Distributions and Data with the GLD via Moments

Figure 3.3–4. The contour curves for and $\alpha_3^2 = 0.5$ and $\alpha_4 = 2$.

Figure 3.3–5. Contour curves of $\alpha_4$ with $(\lambda_3, \lambda_4)$ from Region 3.
3.3 Fitting the GLD through the Method of Moments

Figure 3.3–6. Contour curves of $\alpha_3^2$ with $(\lambda_3, \lambda_4)$ from Region 3.

Figure 3.3–7. Contour curves of $\alpha_3^2$ and $\alpha_4$ with $(\lambda_3, \lambda_4)$ from Region 3.
Chapter 3: Fitting Distributions and Data with the GLD via Moments

Figure 3.3–8. Contour curves of $\alpha_3^2$ and $\alpha_4$ with $(\lambda_3, \lambda_4)$ from Region 4.

yielded a fifth fit,

$$\text{GLD}_5(0.07325, -0.0001710, -0.0001005, -0.00008796)$$

with $\max(|\hat{\alpha}_3 - \alpha_3|, |\hat{\alpha}_3 - \alpha_3|) = 2.1 \times 10^{-3}$.

When search algorithms such as the one implemented in FindLambdasM fail, they fail because the surfaces associated with the equations have sharp corners or points where differentiability fails. This is not the case here. FindLambdasM and RefineSearchGLDM have difficulty finding a solution in Region 4 because of the proximity of the solution to the origin. When $\alpha_3$ and $\alpha_4$, particularly $\alpha_4$, are calculated with $(\lambda_3, \lambda_4)$ near the origin, unless very high levels of computational precision are used, the computational errors at intermediate levels could get magnified throughout the search path of the algorithm. This phenomenon is illustrated in Figures 3.3–9 and 3.3–10. In Figure 3.3–9 the surface $\alpha_4$ is plotted using 25 digits of precision for all computations (this is well beyond the precision allowed by most hardware-based floating point operations). We can see that the surface is smooth and well-behaved. In Figure 3.3–10 the same surface is plotted with only 10 digits of precision, a rather common level of precision in most computing environments. It is clear that substantial errors are produced when $(\lambda_3, \lambda_4)$ is near the origin. If FindLambdasM is unable to obtain a solution within the specified number of iterations and error tolerance, $\epsilon$, it gives the approximate
3.3 Fitting the GLD through the Method of Moments

Figure 3.3–9. The surface $\alpha_4$, plotted with high precision computation, near the origin in Region 4.

Figure 3.3–10. The surface $\alpha_4$, plotted with ordinary precision computation, near the origin in Region 4.
values of \( \lambda_1, \lambda_2, \lambda_3, \lambda_4 \) that it has computed, along with an appropriate warning.

A final word of caution: when FindLambdasM returns \( \lambda_1, \lambda_2, \lambda_3, \lambda_4 \), there is no assurance that the GLD associated with \( \lambda_1, \lambda_2, \lambda_3, \lambda_4 \) is valid. Thus, one needs to check that \( (\lambda_3, \lambda_4) \) is in one of the valid regions of Figure 2.2–1.

### 3.3.2 Fitting by the Use of Tables

Some readers may not have sufficient expertise in programming or adequate programming support to use the type of analysis that was illustrated in Section 3.3.1. For this reason, a number of investigators have provided tables for the estimation of \( \lambda_1, \lambda_2, \lambda_3, \lambda_4 \). The first of these was given by Ramberg and Schmeiser (1974); more comprehensive tables (in the sense of coverage of \((\alpha_2^2, \alpha_4^-)\)-space) were provided subsequently by Ramberg, Tadikamalla, Dudewicz, and Mykytka (1979); Cooley (1991) used greater computational precision to improve previous tables; and Dudewicz and Karian (1996) provide the most accurate and comprehensive tables to date. The latter are given in Appendix B. To capture as much precision as possible within the table of Appendix B, the notation \( a^b \) is used for the entries of the table to mean \( a \times 10^{-b} \). For example, an entry of 0.14172 represents 0.001417.

Unless some simplifications are used, tabulated results for determining \( \lambda_1, \lambda_2, \lambda_3, \lambda_4 \) from \( \hat{\lambda}_1, \hat{\lambda}_2, \hat{\lambda}_3, \hat{\lambda}_4 \) would require a “four-dimensional” display, a decidedly impractical undertaking. To make the tabulation manageable, we first use not \((\hat{\lambda}_1, \hat{\lambda}_2, \hat{\lambda}_3, \hat{\lambda}_4)\) but \((0, 1, |\hat{\lambda}_3|, \hat{\lambda}_4)\) and obtain a solution \((\lambda_1(0, 1), \lambda_2(0, 1), \lambda_3, \lambda_4)\) to (3.3.5). Note that interchanging \( \lambda_3 \) and \( \lambda_4 \) would change the signs of \( A \) and \( C \) in (3.1.12) through (3.1.15), changing the sign of \( \alpha_3 \) and necessitating a sign change for \( \lambda_1(0, 1) \). Therefore, when \( \hat{\alpha}_3 < 0 \), we interchange \( \lambda_3 \) and \( \lambda_4 \) and change the sign of \( \lambda_1(0, 1) \). Next, we obtain the solution to (3.3.5) associated with \( \hat{\alpha}_1, \hat{\alpha}_2, \hat{\alpha}_3, \hat{\alpha}_4 \) by setting

\[
\lambda_1 = \lambda_1(0, 1)\sqrt{\hat{\alpha}_2} + \hat{\alpha}_1 \quad \text{and} \quad \lambda_2 = \lambda_2(0, 1)/\sqrt{\hat{\alpha}_2}.
\]

We summarize this process in the GLD–M algorithm below.

**Algorithm GLD–M: Fitting a GLD distribution to data by the method of moments.**

GLD–M–1. Use (3.3.1) through (3.3.4) to compute \( \hat{\alpha}_1, \hat{\alpha}_2, \hat{\alpha}_3, \hat{\alpha}_4 \);

GLD–M–2. Find the entry point in a table of Appendix B closest to \((|\hat{\alpha}_3|, \hat{\alpha}_4)\);

GLD–M–3. Using \((|\hat{\alpha}_3|, \hat{\alpha}_4)\) from Step GLD–M–2 extract \( \lambda_1(0, 1), \lambda_2(0, 1), \lambda_3, \lambda_4 \) from the table;

GLD–M–4. If \( \hat{\alpha}_3 < 0 \), interchange \( \lambda_3 \) and \( \lambda_4 \) and change the sign of \( \lambda_1(0, 1) \);

GLD–M–5. Compute \( \lambda_1 = \lambda_1(0, 1)\sqrt{\hat{\alpha}_2} + \hat{\alpha}_1 \quad \text{and} \quad \lambda_2 = \lambda_2(0, 1)/\sqrt{\hat{\alpha}_2} \).
To illustrate the use of Algorithm GLD–M and the table of Appendix B suppose that \( \hat{\alpha}_1, \hat{\alpha}_2, \hat{\alpha}_3, \hat{\alpha}_4 \) have been computed to have values

\[
\hat{\alpha}_1 = 2, \quad \hat{\alpha}_2 = 3, \quad \hat{\alpha}_3 = -\sqrt{0.025}, \quad \hat{\alpha}_4 = 2. \tag{3.3.10}
\]

Note that \( \hat{\alpha}_4 \) has been taken to be the same as, and \( \hat{\alpha}_3 \) has been taken to be the negative of, the previous values from (3.3.8) used in the \texttt{FindLambdasM} procedure.

Step GLD–M–1 is taken care of since \( \hat{\alpha}_1, \hat{\alpha}_2, \hat{\alpha}_3, \hat{\alpha}_4 \) is given. For Step GLD–M–2, we observe that \( \hat{\alpha}_3 = -0.15811 \); hence, the closest point to \((|\hat{\alpha}_3|, \hat{\alpha}_4)\) in the Table of Appendix B is \((0.15, 2.0)\), giving us

\[
\lambda_1(0, 1) = -1.3231, \quad \lambda_2(0, 1) = 0.2934, \quad \lambda_3 = 0.03145, \quad \lambda_4 = 0.7203.
\]

The instructions on the use of the table in Appendix B indicate that a superscript of \( b \) in a table entry designates a factor of \( 10^{-b} \). In this case, an entry of 0.3145\(^1\) for \( \lambda_3 \) indicates a value of \( 0.3145 \times 10^{-1} = 0.03145 \). Since \( \alpha_3 < 0 \), Step GLD–M–4 readjusts these to

\[
\lambda_1(0, 1) = 1.3231, \quad \lambda_2(0, 1) = 0.2934, \quad \lambda_3 = 0.7203, \quad \lambda_4 = 0.03145.
\]

With the computations in Step GLD–M–5 we get

\[
\lambda_1 = 4.2917, \quad \lambda_2 = 0.1694, \quad \lambda_3 = 0.7203, \quad \lambda_4 = 0.03145.
\]

### 3.3.3 Limitations of the Method of Moments

The wide applicability of the methods developed in Sections 3.3.1 and 3.3.2 will be apparent when we use GLD distributions to approximate a number of commonly encountered distributions (Section 3.4) and when we fit GLD distributions to several datasets (Section 3.5). However, it is worth keeping in mind that most methods have limitations and we discuss the limitations associated with fitting a GLD through the method of moments here.

Algorithm GLD–M, through the table of Appendix B, enables us to fit a GLD when \((\alpha_3^2, \alpha_4)\) is confined by

\[
1.8(\hat{\alpha}_3^2 + 1) \leq \hat{\alpha}_4 \leq 1.8\hat{\alpha}_3^2 + 15. \tag{3.3.11}
\]

The upper restriction \( \hat{\alpha}_4 \leq 1.8\hat{\alpha}_3^2 + 15 \) is forced on us by limitations of table space and difficulties associated with computations when this restriction is removed. We can see from Figures 3.2–1 through 3.2–3 that the GLD is capable of generating distributions with \((\alpha_3^2, \alpha_4)\) beyond this constraint (see also the shaded region of possible \((\alpha_3^2, \alpha_4)\) pairs in Figure 3.2–5). If needed, it is quite likely that, perhaps with some difficulty, we would be able to fit a suitable GLD in this region.

The lower restriction of (3.3.11), \( 1.8(\hat{\alpha}_3^2 + 1) \leq \hat{\alpha}_4 \), is based on computational results that are depicted in Figures 3.2–1 through 3.2–3 and thus represent an
approximation of the true boundary. Recall (Theorem 3.2.4) that for all distributions we must have $\alpha_4 > 1 + \alpha_2^2$. Thus, while the upper restriction of (3.3.11) may be overcome through greater computational effort, the lower restriction eliminates the possibility of fitting a GLD($\lambda_1, \lambda_2, \lambda_3, \lambda_4$) when

$$1 + \hat{\alpha}_3^2 < \hat{\alpha}_4 < 1.8(\hat{\alpha}_3^2 + 1).$$

While analyses of actual data by Wilcox (1990), Pearson and Please (1975), and Micceri (1989), indicate that values of $|\hat{\alpha}_3|$ up to 4 and values of $\hat{\alpha}_4$ up to 50 are realistic, it is most common for data to produce ($\alpha_2, \alpha_4$) with

$$1 + \hat{\alpha}_3^2 < \hat{\alpha}_4 \leq 1.8\hat{\alpha}_3^2 + 15,$$

making the lower constraint of (3.3.11) a more serious limitation. In Chapter 4 we develop the EGLD system, the Extended GLD, to address this problem.

A different problem arises when we try to find a GLD($\lambda_1, \lambda_2, \lambda_3, \lambda_4$) approximation to a distribution when (some of) the first four moments of the distribution do not exist. This type of fitting problem will also arise in a less obvious form if we encounter data that is a random sample from such a distribution. We address this difficulty in Chapter 5 by devising a GLD($\lambda_1, \lambda_2, \lambda_3, \lambda_4$) fitting method that depends on percentiles rather than moments.

In terms of a preference between the two approaches discussed in Sections 3.3.1 (direct computation) and 3.3.2 (use of tables), we note that the unavailability of the proper computing environment and, to a lesser extent, simplicity are the principle advantages of using tables. There are, however, two disadvantages:

1. Because of length limitations, existing tables provide at most one solution even when multiple solutions may exist.

2. Results obtained through Algorithm GLD–M, because of their dependence on tables, are less accurate than estimations of $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ by direct computation. We know, for example, from the GLD$_1$ fit of Section 3.3.1 that 0.7181 and 0.02982 would be more precise values for $\lambda_3$ and $\lambda_4$, respectively.

If one has access to a computational system that can provide solutions to equations like (3.3.5), it is possible to use the tables of Appendix B to determine a good starting point for the search so that an accurate solution may be obtained with considerable efficiency.

Of course, the ultimate criterion is goodness of fit of the fitted distribution to the true (unknown) distribution. Methods of assessing this are discussed in Section 3.5.1.

### 3.4 GLD Approximations of Some Well-Known Distributions

In Section 2.3 we saw the large variety of shapes that the GLD($\lambda_1, \lambda_2, \lambda_3, \lambda_4$) p.d.f. can attain. For the GLD($\lambda_1, \lambda_2, \lambda_3, \lambda_4$) to be useful for fitting distributions to
data, it should be able to provide good fits to many of the distributions the data may come from. In this section we see that the GLD($\lambda_1, \lambda_2, \lambda_3, \lambda_4$) fits well many of the most important distributions.

We apply five checks on each occasion where we fit a GLD to a distribution. The first check considers the closeness of $\hat{f}(x)$, the approximating GLD p.d.f., to $f(x)$, the p.d.f. of the distribution being approximated. The proximity of $\hat{f}(x)$ to $f(x)$ is determined by approximating $\sup |\hat{f}(x) - f(x)|$. The p.d.f. of the distribution we will be approximating, $f(x)$, is available to us; therefore, there is no difficulty with computing $f(x)$. To compute $\hat{f}(x)$,

- we take 2249 equispaced points $y_i = i/2250$, for $i = 1, 2, \ldots, 2249$ from the interval $(0, 1)$;
- using (2.1.1) we compute $x_i = Q(y_i)$ for $i = 1, 2, \ldots, 2249$;
- using (2.1.3) we compute $\hat{f}(x_i)$ for $i = 1, 2, \ldots, 2249$.

We now use the approximation

$$\sup |\hat{f}(x) - f(x)| \approx \max_{1 \leq i \leq 2249} |\hat{f}(x_i) - f(x_i)|.$$ 

In actual practice we found that using 2249 points does not limit the accuracy of the $\sup |\hat{f}(x) - f(x)|$ that we compute. We have obtained essentially the same (i.e., within $10^{-4}$) values for $\sup |\hat{f}(x_i) - f(x_i)|$ when the 2249 points have been increased to 4999 points.

For a second check, we look at the proximity of the approximating and approximated d.f.s, $\hat{F}(x)$ and $F(x)$, respectively. While p.d.f. differences are less easy to interpret, differences in d.f.s have an immediate meaning in the probability assigned to easy-to-interpret events. To check the closeness of $\hat{F}(x)$ to $F(x)$, we follow the same idea (and use the same points $x_i$) as in the computation of $\hat{f}(x)$ and use the approximation

$$\sup |\hat{F}(x) - F(x)| \approx \max_{1 \leq i \leq 2249} |\hat{F}(x_i) - F(x_i)|.$$ 

Again, in practice, $\sup |\hat{F}(x) - F(x)|$ does not change much (less than $10^{-5}$) when a much larger number of points is used.

To get a quantitative measure of the quality of approximations we also consider the “distances” between $g(x)$, the p.d.f. of the distribution being approximated, and $f(x)$, the p.d.f. of the approximating GLD($\lambda_1, \lambda_2, \lambda_3, \lambda_4$). There is a large body of literature involving evaluation of the estimate $f(x)$ of a p.d.f. $g(x)$, through nonnegative divergence or pseudodistance measures designated by $D(f,g)$, which rely on the $L_p$-norm, which we now define.

**Definition 3.4.1.** The $L_p$ norm of an integrable function $f$ over $R$ is

$$\|f\|_p = \left( \int_R |f(x)|^p dx \right)^{1/p}.$$
and the $L_p$ distance between to integrable functions $f$ and $g$ over $R$ is $\|f - g\|_p$.

The $L_p$ distance is a commonly used form of $D(f, g)$ (see Györfi, Liese, Vajda, and van der Meulen (1998) for details). We will concentrate on the cases $p = 1$ and $p = 2$.

The case $p = 1$ has a natural interpretation in terms of probability. The overlapping coefficient, $\Delta(f, g)$, of any two p.d.f.s, $f(x)$ and $g(x)$, is defined as the area that is under both p.d.f.s and above the horizontal axis; equivalently, $\Delta(f, g)$ is the area above the horizontal axis and below the function $\min(f(x), g(x))$. (For an introductory discussion of the history and literature of this subject, see Mishra, Shah, and Lefante (1986) and for examples, see Dudewicz and Mishra (1988).)

The relationship of the $L_1$ distance, $\|f - g\|_1$, to $\Delta(f, g)$ is shown in Figure 3.4–1 where $\|f - g\|_1$ is the area between the two illustrated p.d.f.s and $\Delta(f, g)$ is the area under $\min(f(x), g(x))$ which is shown in heavy print. It can be seen from Figure 3.4–1 that

$$\Delta(f, g) + \|f - g\|_1 = \int_{-\infty}^{\infty} \max(f(x), g(x)) \, dx. \tag{3.4.2}$$

Hence, the $\|f - g\|_1$ and $\Delta(f, g)$ are two sides of the same coin: $\|f - g\|_1$ measures the difference, while $\Delta(f, g)$ measures the commonality, of two p.d.f.s

In Point 6 of Section 3.5.1 we will discuss some of the many statistical tests of the hypothesis that given data comes from a specified d.f., $G(x)$. Many of these tests rely on the sample (empirical) d.f. of the data. However, a number of the tests have an interpretation based on the closeness of an $f(x)$, the estimating p.d.f., to $g(x)$, the p.d.f. being estimated (see Dudewicz and van der Meulen (1981) for such an interpretation with the entropy test). Since it is of interest
3.4 GLD Approximations of Some Well-Known Distributions

to consider distance measures between p.d.f.s \(f(x)\) and \(g(x)\), based on \(f(x)\) and \(g(x)\) themselves (and not their d.f.s), we begin with the definition of \(L_1\) and \(L_2\) norms.

When there is complete agreement between the functions \(f\) and \(g\),
\[
||f - g||_1 = ||f - g||_2 = 0.
\]

Moreover, because they are p.d.f.s, the integrals of \(f\) and \(g\) are equal to 1 and
\[
||f - g||_1 \leq 2.
\]

The integrations that lead to \(||f - g||_1\) and \(||f - g||_2\) must be done numerically because we do not have a closed-form expression for \(f(x)\), the p.d.f. of the approximating GLD(\(\lambda_1, \lambda_2, \lambda_3, \lambda_4\)). The results, therefore, will be numerical approximations. The following algorithm will produce approximate \(L_1\) and \(L_2\) distances.

**Algorithm \(L_1L_2\): Approximations to \(L_1\) and \(L_2\) Distances.**

\(L_1L_2\)-1. Input
- \(n\), a positive integer \(\geq 3\).
- \(\lambda_1, \lambda_2, \lambda_3, \lambda_4\), the parameters of the GLD(\(\lambda_1, \lambda_2, \lambda_3, \lambda_4\)) fit with p.d.f. \(f(x)\).
- \(g(x)\), the p.d.f. of the distribution being fitted.

\(L_1L_2\)-2. Compute values \(p_i = i/n\) for \(i = 1, 2, \ldots, n - 1\).

\(L_1L_2\)-3. Compute the \(n - 1\) percentile points \(\pi_i = Q(p_i)\) for \(i = 1, 2, \ldots, n - 1\) (see (2.1.1)).

\(L_1L_2\)-4. Compute the \(n - 1\) \(y\)-coordinates of the GLD(\(\lambda_1, \lambda_2, \lambda_3, \lambda_4\)) p.d.f. using
\[
y_i = \frac{\lambda_2}{\lambda_3 p_i^{\lambda_3 - 1} + \lambda_4 (1 - p_i)^{\lambda_4 - 1}},
\]
for \(i = 1, 2, \ldots, n - 1\) (see (2.1.3)). The points \((\pi_i, y_i)\) are on the graph of \(f\).

\(L_1L_2\)-5. Compute the \(n - 1\) values on \(g(x)\), the function being fitted, by \(Y_i = g(\pi_i)\) for \(i = 1, 2, \ldots, n - 1\), making sure that when \(p_i\) is outside of the support of \(g(x)\), \(Y_i\) is assigned a value of zero. The points \((\pi_i, Y_i)\) are on the graph of \(g\).

\(L_1L_2\)-6. Let \(\Delta_i = \pi_{i+1} - \pi_i\) for \(i = 1, 2, \ldots, n - 2\).

\(L_1L_2\)-7. Compute the sums
\[
S_1 = \sum_{i=1}^{n-2} \Delta_i |y_i - Y_i| \quad \text{and} \quad S_2 = \sum_{i=1}^{n-2} \Delta_i (y_i - Y_i)^2.
\]
Chapter 3: Fitting Distributions and Data with the GLD via Moments

$L_1 L_2 - 8$. Compute

\[ T_1 = \int_{-\infty}^{\pi_1} g(x) \, dx + \int_{\pi_{n-1}}^{\infty} g(x) \, dx \]

and

\[ T_2 = \int_{-\infty}^{\pi_1} g^2(x) \, dx + \int_{\pi_{n-1}}^{\infty} g^2(x) \, dx. \]

$L_1 L_2 - 9$. $S_1 + T_1$ approximates the $L_1$ distance between $f$ and $g$ and $\sqrt{S_2 + T_2}$ approximates the $L_2$ distance between $f$ and $g$.

In Algorithm $L_1 L_2$, $S_1$ and $S_2$ are Riemann sums for the integrals

\[ \int_{\pi_1}^{\pi_{n-1}} |f(x) - g(x)| \, dx \quad \text{and} \quad \int_{\pi_1}^{\pi_{n-1}} (f(x) - g(x))^2 \, dx, \tag{3.4.3} \]

respectively. Therefore, $S_1$ and $S_2$ converge to the values of these definite integrals as $n \to \infty$. Since the accumulated probability of $f(x)$ on the intervals $(-\infty, \pi_1)$ and $(\pi_{n-1}, \infty)$ is $2/n$ and $f(x)$ and $g(x)$ are non-negative,

\[
\int_{-\infty}^{\pi_1} |f(x) - g(x)| \, dx + \int_{\pi_{n-1}}^{\infty} |f(x) - g(x)| \, dx \\
\leq \int_{-\infty}^{\pi_1} (f(x) + g(x)) \, dx + \int_{\pi_{n-1}}^{\infty} (f(x) + g(x)) \, dx \\
\leq \frac{2}{n} + T_1. \tag{3.4.4}
\]

Therefore,

\[ \left\| \hat{f} - g \right\|_1 - (S_1 + T_1) \leq \frac{2}{n} \tag{3.4.5} \]

and

\[ \left\| \hat{f} - g \right\|_1 = \lim_{n \to \infty} (S_1 + T_1). \tag{3.4.6} \]

It can be similarly established that

\[ \left\| \hat{f} - g \right\|_2 = \lim_{n \to \infty} \sqrt{S_2 + T_2}, \tag{3.4.7} \]

justifying the conclusions in Step $L_1 L_2 - 9$.

We use the $L_1$ and $L_2$ distance measures, $\left\| \hat{f} - f \right\|_1$ and $\left\| \hat{f} - f \right\|_2$, as our third and fourth checks, respectively. For a fifth check we consider the closeness of the $\hat{\alpha}_1, \hat{\alpha}_2, \hat{\alpha}_3, \hat{\alpha}_4$ of the fitted GLD to the $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ of the distribution that is approximated. Since in all cases $\text{FindLambdasM}$ will be used to obtain the approximating GLD, we are assured, subject to a warning message from $\text{FindLambdasM}$, that both $|\hat{\alpha}_3 - \alpha_3|$ and $|\hat{\alpha}_4 - \alpha_4|$ are less than $10^{-5}$. From
(3.1.12) and (3.1.13) we know that there are no difficulties associated with the computations of \( \hat{\alpha}_1 \) and \( \hat{\alpha}_2 \); therefore, we expect \(|\hat{\alpha}_1 - \alpha_1|\) and \(|\hat{\alpha}_2 - \alpha_2|\) to be no larger than \(10^{-4}\). Since the use of \texttt{FindLambdasM} takes care of this check, we do not explicitly mention it as we look for fits to distributions in the following sections.

### 3.4.1 The Normal Distribution

The normal distribution, with mean \(\mu\) and variance \(\sigma^2 \ (\sigma > 0)\), \(N(\mu, \sigma^2)\), has p.d.f.

\[
f(x) = \frac{1}{\sigma \sqrt{2\pi}} \exp \left[ -\frac{(x - \mu)^2}{2\sigma^2} \right], \quad -\infty < x < \infty.
\]

Since all normal distributions can be obtained by a location and scale adjustment to \(N(0, 1)\), we consider a GLD\((\lambda_1, \lambda_2, \lambda_3, \lambda_4)\) fit to \(N(0, 1)\) for which \(\alpha_1 = 0, \ \alpha_2 = 1, \ \alpha_3 = 0, \ \alpha_4 = 3\).

Appendix B suggests the existence of a solution with \((\lambda_3, \lambda_4)\) near \((0.13, 0.13)\). Using \texttt{FindLambdasM} we obtain the approximation

\[
\text{GLD}( -2.2 \times 10^{-9}, 0.1975, 0.1349, 0.1349).
\]

For our first check of the fit, we observe that the graphs of the \(N(0, 1)\) and the fitted p.d.f.s, given in Figure 3.4–2 (a), show the two p.d.f.s to be “nearly identical” (the \(N(0, 1)\) p.d.f. is slightly higher at the center). Specifically, we compute

\[
\sup |\hat{f}(x) - f(x)| = 0.002812,
\]

where \(f(x)\) and \(\hat{f}(x)\) are the p.d.f.s of the \(N(0, 1)\) and the fitted distributions, respectively. As a second check of the fit, we observe that the graphs of the \(N(0, 1)\) and the fitted d.f.s, given in Figure 3.4–2 (b), cannot be distinguished. Specifically,

\[
\sup |\hat{F}(x) - F(x)| = 0.001087,
\]

where \(F(x)\) and \(\hat{F}(x)\) are the d.f.s of the \(N(0, 1)\) and the fitted distributions, respectively. This means that the probability that \(X\) is at most \(x\) differs from its approximation by no more than 0.001087 at any \(x\).

For our third and fourth checks we note that the \(L_1\) and \(L_2\) distances for this approximation are

\[
||\hat{f} - f||_1 = 0.006650 \quad \text{and} \quad ||\hat{f} - f||_2 = 0.003231.
\]

Since \((\lambda_3, \lambda_4)\) is from Region 3, GLD\(( -2.2 \times 10^{-9}, 0.1975, 0.1349, 0.1349)\) has finite support (see Theorem 2.3.23) in the form of the interval \([-5.06, 5.06]\). This may or may not be desirable — see the discussion at the beginning of this chapter.
3.4.2 The Uniform Distribution

The continuous uniform distribution on the interval \((a, b)\) with \(a < b\) has p.d.f.

\[
f(x) = \begin{cases} 
\frac{1}{b - a}, & \text{if } a < x < b \\
0, & \text{otherwise.}
\end{cases}
\]

For simplicity we consider the uniform distribution on the interval \((0, 1)\) for which

\[
\alpha_1 = \frac{1}{2}, \quad \alpha_2 = \frac{1}{12}, \quad \alpha_3 = 0, \quad \alpha_4 = \frac{9}{5}.
\]

For this distribution, \texttt{FindLambdasM} yields

\[
\text{GLD}(0.5, 2.0000, 1.0000, 1.0000).
\]

In this case, the fit is perfect because using these values of \(\lambda_1, \lambda_2, \lambda_3, \lambda_4\), in (2.1.1) gives \(Q(y) = y\) where \(Q(y)\) is the inverse distribution function of the fitted GLD. Therefore, we must also have \(F(y) = y\), matching the distribution function of the uniform distribution on the \((0, 1)\) interval. Hence, the results of our checks are:

\[
\sup |\hat{f}(x) - f(x)| = 0, \quad \sup |\hat{F}(x) - F(x)| = 0, \quad ||\hat{f} - f||_1 = 0, \quad ||\hat{f} - f||_2 = 0.
\]

For the general uniform distribution on \((a, b)\) we have

\[
\alpha_1 = \frac{a + b}{2}, \quad \alpha_2 = \frac{(b - a)^2}{12}, \quad \alpha_3 = 0, \quad \alpha_4 = \frac{9}{5}.
\]

Since \(\alpha_3\) and \(\alpha_4\) do not change, to fit the uniform distribution on \((a, b)\) we need only adjust \(\lambda_1\) and \(\lambda_2\).
3.4 GLD Approximations of Some Well-Known Distributions

If \( X \) is a uniform random variable on \([a, b]\), then its distribution function is
\[
F(x) = \frac{x - a}{b - a} \quad (\text{with } a \leq x \leq b)
\]
and its quantile function is
\[
Q_X(y) = a + (b - a)y \quad (\text{with } 0 \leq y \leq 1).
\]
Setting \( Q_X(y) \) equal to the GLD quantile function gives
\[
\lambda_1 + \left[ y^{\lambda_3} - (1 - y)^{\lambda_4} \right] / \lambda_2 = a + (b - a)y.
\]
If we let \( \lambda_3 = \lambda_4 = 1 \), this reduces to
\[
(\lambda_1 - 1/\lambda_2) + (2/\lambda_2)y = a + (b - a)y,
\]
from which we get \( \lambda_1 = a + 1/\lambda_2 \) and \( \lambda_2 = 2/(b - a) \), equivalently, \( \lambda_1 = (a + b)/2 \) and \( \lambda_2 = 2/(b - a) \). We thus have the exact GLD fit GLD((a+b)/2, 2/(b-a), 1, 1).

Setting \( \lambda_3 = \lambda_4 = 1 \) is not our only option. As another possibility, we can take \( \lambda_3 = 1 \) and \( \lambda_4 = 0 \) to get, by equating quantile functions,
\[
\lambda_1 + (y - 1)/\lambda_2 = a + (b - a)y,
\]
and eventually, \( \lambda_1 = b \) and \( \lambda_2 = 1/(b - a) \). This leads to a second exact fit GLD(b, 1/(b-a), 1, 0) (or, GLD(1, 1, 1, 0) in the case of the uniform distribution on \([0, 1]\)). Setting \( \lambda_3 = 0 \) and \( \lambda_4 = 1 \) produces yet a third GLD representation of the uniform distribution as GLD(a, 1/(b-a), 0, 1) (or, GLD(0, 1, 0, 1) in the case of the uniform distribution on \([0, 1]\)).

3.4.3 The Student’s \( t \) Distribution

The Student’s \( t \) distribution with \( \nu \) degrees of freedom, \( t(\nu) \), has p.d.f.
\[
f(x) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi\nu} \Gamma\left(\frac{\nu}{2}\right) \left(1 + \frac{x^2}{\nu}\right)^{(\nu+1)/2}}, \quad -\infty < x < \infty.
\]
The specification of the \( t(\nu) \) p.d.f. uses the gamma function, \( \Gamma(t) \), which for \( t > 0 \) is defined by
\[
\Gamma(t) = \int_0^{\infty} y^{t-1} e^{-y} \, dy.
\]
Some properties of \( \Gamma(t) \) will be developed in Section 3.1 (for a more detailed discussion see Artin (1964)).

The existence of the \( i \)-th moment of \( t(\nu) \) depends on the relative sizes of \( i \) and \( \nu \). For the \( i \)-th moment to exist, the integral
\[
\int_{-\infty}^{\infty} \frac{x^i \Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi\nu} \Gamma\left(\frac{\nu}{2}\right) \left(1 + \frac{x^2}{\nu}\right)^{(\nu+1)/2}} \, dx
\]
must converge. The power of \(x\) in the integrand is \(i - \nu - 1\). Therefore, the integral will converge if and only if \(i - \nu - 1 < -1\) or \(i < \nu\).

Since we need the first four moments of a distribution to apply the method of moments, we can only consider \(t\) distributions with \(\nu \geq 5\). (In Chapter 5 we develop other methods that can be used for approximating \(t\) distributions with small \(\nu\).) We expect that the limiting case \(\nu = 5\) may be the one where a GLD(\(\lambda_1, \lambda_2, \lambda_3, \lambda_4\)) fit will be most difficult.

The first four moments of \(t(\nu)\) are

\[
\begin{align*}
\alpha_1 &= 0, & \alpha_2 &= \frac{\nu}{\nu - 2}, & \alpha_3 &= 0, & \alpha_4 &= 3 \frac{\nu - 2}{\nu - 4} \\
\end{align*}
\]

which, when \(\nu = 5\), become

\[
\begin{align*}
\alpha_1 &= 0, & \alpha_2 &= \frac{5}{3}, & \alpha_3 &= 0, & \alpha_4 &= 9.
\end{align*}
\]

To fit \(t(5)\), we appeal to FindLambdasM and obtain

\[
\text{GLD}(0.1641 \times 10^{-9}, -0.2481, -0.1359, -0.1359).
\]

For our first check we observe the p.d.f.s of this GLD and \(t(5)\), shown in Figure 3.4–3 (a) (the one that rises higher at the center is the GLD p.d.f.), and compute

\[
\sup |\hat{f}(x) - f(x)| = 0.03581.
\]

Our second check leads us to consider the graphs of the d.f.s of \(t(5)\) and its fitted distribution. These are shown in Figure 3.4–3 (b) (the \(t(5)\) d.f. rises slightly higher on the left and is slightly lower on the right) and yield

\[
\sup |\hat{F}(x) - F(x)| = 0.01488.
\]

For our third and fourth checks we obtain

\[
||\hat{f} - f||_1 = 0.006650 \quad \text{and} \quad ||\hat{f} - f||_2 = 0.003231.
\]

While this seems to be a reasonably good fit, it does not look as good as the \(N(0, 1)\) fit. Also, unlike the \(N(0, 1)\) fit, the support of this GLD fit is \((-\infty, \infty)\) — as it is for \(t(5)\).

As \(\nu\) gets large, the GLD(\(\lambda_1, \lambda_2, \lambda_3, \lambda_4\)) fits for \(t(\nu)\) get better. For \(\nu = 6, 10,\) and 30 we get, respectively, the following fits.

\[
(\lambda_1, \lambda_2, \lambda_3, \lambda_4) = (2.4 \times 10^{-10}, -0.1376, -0.08020, -0.08020)\ 
\]

\[
\sup |\hat{f}(x) - f(x)| = 0.02311, \text{ and } \sup |\hat{F}(x) - F(x)| = 0.009513 \\
||\hat{f} - f||_1 = 0.04362 \text{ and } ||\hat{f} - f||_2 = 0.02251;
\]

\[
(\lambda_1, \lambda_2, \lambda_3, \lambda_4) = (0.0002229, 0.02340, 0.01479, 0.01479), \text{ with}
\]

\[
(\lambda_1, \lambda_2, \lambda_3, \lambda_4) = (0.0002229, 0.02340, 0.01479, 0.01479), \text{ with}
\]

\[
||\hat{f} - f||_1 = 0.04362 \text{ and } ||\hat{f} - f||_2 = 0.02251;
\]
3.4 GLD Approximations of Some Well-Known Distributions

Figure 3.4–3. The p.d.f. and d.f. of t(5) with the fitted GLD. The GLD p.d.f. rises higher at the center (a). The d.f. rises higher on the left and is lower on the right (b).}

\[
sup |\hat{f}(x) - f(x)| = 0.01039 \quad \text{and} \quad sup |\hat{F}(x) - F(x)| = 0.004165
\]

\[
||\hat{f} - f||_1 = 0.02104 \quad \text{and} \quad ||\hat{f} - f||_2 = 0.01065;
\]

\[
(\lambda_1, \lambda_2, \lambda_3, \lambda_4) = (3.4 \times 10^{-9}, 0.1452, 0.09701, 0.09701), \quad \text{with}
\]

\[
sup |\hat{f}(x) - f(x)| = 0.004544 \quad \text{and} \quad sup |\hat{F}(x) - F(x)| = 0.001766
\]

\[
||\hat{f} - f||_1 = 0.01009 \quad \text{and} \quad ||\hat{f} - f||_2 = 0.005089.
\]

3.4.4 The Exponential Distribution

The exponential distribution with parameter \( \theta > 0 \) has p.d.f.

\[
f(x) = \begin{cases} 
\frac{1}{\theta}e^{-\frac{x}{\theta}}, & \text{if } x > 0 \\
0, & \text{otherwise.}
\end{cases}
\]

For this distribution

\[
\alpha_1 = \theta, \quad \alpha_2 = \theta^2, \quad \alpha_3 = 2, \quad \alpha_4 = 9.
\]

We can see that \( \alpha_3 \) and \( \alpha_4 \) do not change because \( \theta \) is a scale parameter. Therefore, if a specific exponential distribution, say with \( \theta = 1 \), can be fitted, then other exponential distributions can be fitted using the \( \lambda_3 \) and \( \lambda_4 \) from the fit obtained for \( \theta = 1 \).

For \( \theta = 1 \),

\[
\alpha_1 = 1, \quad \alpha_2 = 1, \quad \alpha_3 = 2, \quad \alpha_4 = 9.
\]
Since \((\alpha_3, \alpha_4) = (2, 9)\) is an entry point in the table of Appendix B, we can use the table values without the concern about interpolation errors. This produces, after the adjustments mandated in Step GLD–M–5 of Algorithm GLD–M,

\[
\text{GLD}(1.7240 \times 10^{-4}, 9.2756 \times 10^{-5}, 1.1840 \times 10^{-8}, 9.2766 \times 10^{-5}).
\]

We consider the exponential p.d.f., with \(\theta = 1\), and its fitted GLD p.d.f., plotted together in Figure 3.4–4(a) and note that the two p.d.f.s seem identical. The explanation for the surprising result

\[
\sup |\hat{f}(x) - f(x)| = 0.8608
\]
is that, although not evident from Figure 3.4–4 (a), the GLD p.d.f. has support \([-10781, 10781]\) and for negative values close to 0, it assumes values near 0.8608 where the exponential p.d.f. is zero. This yields a large difference over a small range for the p.d.f.s, as shown in Figure 3.4–4 (b) (though the d.f.s are very close).

When \(\theta = 3\), the fitted p.d.f. has

\[
(\lambda_1, \lambda_2, \lambda_3, \lambda_4) = (3.1572 \times 10^{-4}, 3.0919 \times 10^{-5}, 1.1840 \times 10^{-8}, 9.2766 \times 10^{-5}),
\]

which leads to the observation that \(\lambda_3\) approaches 0 as do \(\lambda_2\) and \(\lambda_4\) but with the condition \(\lambda_4 = \theta \lambda_2\). If we set \(\lambda_3 = 0\), \(\lambda_4 = \theta \lambda_2\) and then take the limit of the GLD quantile function \(Q(y)\) as \(\lambda_2 \to 0\), we get

\[
\lim_{\lambda_2 \to 0} Q(y) = \lim_{\lambda_2 \to 0} \left( \lambda_1 + \frac{y^0 - (1 - y)^{\theta \lambda_4}}{\lambda_2} \right) = \lambda_1 - \theta \ln(1 - y),
\]

which (when \(\lambda_1 = 0\)) becomes the quantile function of the exponential distribution with parameter \(\theta\). Therefore, the exponential distribution can be realized as a limiting case of \(\text{GLD}(0, \lambda_2, 0, \theta \lambda_2)\), as \(\lambda_2 \to 0\). This allows us, at least theoretically, to make

\[
\sup |\hat{f}(x) - f(x)|, \quad \sup |\hat{F}(x) - F(x)|, \quad ||\hat{f} - f||_1, \quad ||\hat{f} - f||_2,
\]

all arbitrarily close to zero.
3.4.5 The Chi-Square Distribution

The p.d.f. of the $\chi^2(\nu)$ distribution with $\nu$ degrees of freedom is given by

$$f(x) = \begin{cases} \frac{x^{\nu/2-1}e^{-x/2}}{\Gamma(\nu/2)2^{\nu/2}}, & \text{if } x \geq 0 \\ 0, & \text{otherwise}. \end{cases}$$

The first four moments of $\chi^2(\nu)$ are

$$\alpha_1 = \nu, \quad \alpha_2 = 2\nu, \quad \alpha_3 = \frac{2\sqrt{2}}{\sqrt{\nu}}, \quad \alpha_4 = 3 + \frac{12}{\nu}.$$ 

We first illustrate the GLD($\lambda_1, \lambda_2, \lambda_3, \lambda_4$) fit for $\nu = 5$ where FindLambdasM is used to obtain

$$\text{GLD}(2.6050, 0.01756, 0.009469, 0.05422)$$

with support $[\lambda_1 - 1/\lambda_2, \lambda_1 + 1/\lambda_2] = [-54.35, 59.56]$ (Theorem 2.3.23).

The first check regarding the closeness of the $\chi^2(5)$ and the fitted p.d.f.s is shown in Figure 3.4–5 (the GLD p.d.f. rises higher at the center). We can see from Figure 3.4–5 that this GLD fits $\chi^2(5)$ reasonably well, and has long but finite left and right tails. To complete this check, we determine

$$\sup |\hat{f}(x) - f(x)| = 0.02115.$$ 

For our second check, we obtain

$$\sup |\hat{F}(x) - F(x)| = 0.01357$$

but we do not illustrate the $\chi^2(5)$ and the fitted d.f.s because the two graphs cannot be visually distinguished.

The computations associated with the third and fourth checks yield

$$||\hat{f} - f||_1 = 0.05879 \quad \text{and} \quad ||\hat{f} - f||_2 = 0.02444.$$ 

When we try to fit $\chi^2(1)$ by the methods of this chapter, we run into difficulties. For $\nu = 1$,

$$\alpha_1 = 1, \quad \alpha_2 = 2, \quad \alpha_3 = 2.8284, \quad \alpha_4 = 15,$$

placing $(\alpha_3^2, \alpha_4)$ well outside the range of the table of Appendix B and our computational capability. When $\nu = 2$, $\chi^2(2)$ is the same as the exponential distribution with $\theta = 2$ and fitting this distribution was covered in Section 3.4.4. There are no difficulties associated with fitting $\chi^2(\nu)$ distributions with $\nu \geq 3$. For $\nu = 3$, 

...
Figure 3.4–5. The $\chi^2$ p.d.f. with its fitted GLD (the GLD p.d.f. rises higher at the center).

$\nu = 10$, and $\nu = 30$, we find, respectively,

$$(\lambda_1, \lambda_2, \lambda_3, \lambda_4) = (0.8596, 0.009543, 0.002058, 0.02300), \text{ with }$$

$$\sup |\hat{f}(x) - f(x)| = 0.05833 \text{ and } \sup |\hat{F}(x) - F(x)| = 0.01589$$

$$||\hat{f} - f||_1 = 0.05934 \text{ and } ||\hat{f} - f||_2 = 0.04086;$$

$$(\lambda_1, \lambda_2, \lambda_3, \lambda_4) = (7.1747, 0.02168, 0.02520, 0.09388), \text{ with }$$

$$\sup |\hat{f}(x) - f(x)| = 0.007793 \text{ and } \sup |\hat{F}(x) - F(x)| = 0.01182$$

$$||\hat{f} - f||_1 = 0.05188 \text{ and } ||\hat{f} - f||_2 = 0.01478;$$

$$(\lambda_1, \lambda_2, \lambda_3, \lambda_4) = (26.4479, 0.01896, 0.05578, 0.1366), \text{ with }$$

$$\sup |\hat{f}(x) - f(x)| = 0.002693 \text{ and } \sup |\hat{F}(x) - F(x)| = 0.007649$$

$$||\hat{f} - f||_1 = 0.03959 \text{ and } ||\hat{f} - f||_2 = 0.007589.$$

We know that for large $\nu$, the closeness of $\chi^2(\nu)$ to $N(\nu, 2\nu)$ would produce a reasonably good fit.

### 3.4.6 The Gamma Distribution

The p.d.f. of the Gamma distribution, $\Gamma(\alpha, \theta)$, with parameters $\alpha > 0$ and $\theta > 0$ is given by

$$f(x) = \begin{cases} 
  x^{\alpha-1}e^{-x/\theta} & \text{if } x \geq 0 \\
  0 & \text{otherwise}.
\end{cases}$$
3.4 GLD Approximations of Some Well-Known Distributions

If $\alpha = 1$, this is the exponential distribution; if $\alpha = \nu/2$ and $\theta = 2$, it is the $\chi^2(\nu)$. The $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ for $\Gamma(\alpha, \theta)$ are given by

$$\alpha_1 = \alpha \theta, \quad \alpha_2 = \alpha \theta^2, \quad \alpha_3 = \frac{2}{\sqrt{\alpha}}, \quad \alpha_4 = 3 + \frac{6}{\alpha}.$$ 

For the purpose of this illustration we take $\alpha = 5$ and $\theta = 3$ for which

$$\alpha_1 = 15, \quad \alpha_2 = 45, \quad \alpha_3 = 2\sqrt{5}, \quad \alpha_4 = \frac{21}{5}.$$ 

The GLD fit that results from $\text{FindLambdasM}$ is

$$\text{GLD}(10.7621, 0.01445, 0.02520, 0.09388).$$

For the first check, we consider the two p.d.f.s shown in Figure 3.4–6 (the curve that rises higher in the middle is the GLD p.d.f.). This seems to be a reasonable fit for which

$$\sup |\hat{f}(x) - f(x)| = 0.005195.$$ 

The fit is somewhat deceptive in that its support is, by Theorem 2.3.23, the interval $[-58.43, 79.96]$. For the second check, we cannot distinguish the graphs of the d.f.s of $\Gamma(5, 3)$ and its fitted GLD and

$$\sup |\hat{F}(x) - F(x)| = 0.01182.$$ 

For our third and fourth checks we note that

$$||\hat{f} - f||_1 = 0.05188 \quad \text{and} \quad ||\hat{f} - f||_2 = 0.01207.$$ 

Figure 3.4–6. The gamma p.d.f. with $\alpha = 5$ and $\theta = 3$ and its fitted GLD (the GLD p.d.f. rises higher at the center).
3.4.7 The Weibull Distribution

A Weibull random variable with parameters $\alpha > 0$ and $\beta > 0$ has p.d.f.

$$f(x) = \begin{cases} \alpha \beta x^{\beta - 1} e^{-\alpha x^\beta}, & \text{if } x \geq 0 \\ 0, & \text{otherwise.} \end{cases}$$

An excellent reference on the Weibull distribution, its multivariate generalizations, as well as reliability-related distributions is Harter (1993), which contains several hundred references on the Weibull distribution alone.

For a Weibull r.v., $X$, it can be easily established that

$$E(X^i) = \alpha^{-i/\beta} \Gamma \left( \frac{i + \beta}{\beta} \right),$$

leading us to

$$\begin{align*}
\alpha_1 &= \alpha^{-\beta-1} \Gamma \left( \frac{\beta + 1}{\beta} \right) \\
\alpha_2 &= -\alpha^{-2\beta-1} \left( -\Gamma \left( \frac{\beta + 2}{\beta} \right) + \left( \Gamma \left( \frac{\beta + 1}{\beta} \right) \right)^2 \right) \\
\alpha_3^{3/2} \alpha_3 &= -\alpha^{-3\beta-1} \left( -\Gamma \left( \frac{3 + \beta}{\beta} \right) + 3 \Gamma \left( \frac{\beta + 2}{\beta} \right) \Gamma \left( \frac{\beta + 1}{\beta} \right) - 2 \left( \Gamma \left( \frac{\beta + 1}{\beta} \right) \right)^3 \right) \\
\alpha_2^2 \alpha_4 &= \alpha^{-4\beta-1} \left( \Gamma \left( \frac{4 + \beta}{\beta} \right) - 4 \Gamma \left( \frac{3 + \beta}{\beta} \right) \Gamma \left( \frac{\beta + 1}{\beta} \right) \right) \\
& \quad + 6 \Gamma \left( \frac{\beta + 2}{\beta} \right) \left( \Gamma \left( \frac{\beta + 1}{\beta} \right) \right)^2 - 6 \left( \Gamma \left( \frac{\beta + 1}{\beta} \right) \right)^4. 
\end{align*}$$

Note that $\alpha_3$ and $\alpha_4$ do not depend on the parameter $\alpha$. If we take $\alpha = 1$ and $\beta = 5$, the moments of the Weibull distribution will be

$$\alpha_1 = 0.91816, \quad \alpha_2 = 0.04423, \quad \alpha_3 = -0.2541, \quad \alpha_4 = 2.8802.$$

For this distribution, FindLambdasM provides the GLD approximation

$$\text{GLD}(0.9935, 1.0491, 0.2121, 0.1061).$$

The support of the fitted GLD is the interval $[0.0404, 1.947]$. For the first check we see in Figure 3.4–7 that the Weibull and the fitted p.d.f.s are “nearly identical” and calculate

$$\sup |\hat{f}(x) - f(x)| = 0.03537.$$
3.4 GLD Approximations of Some Well-Known Distributions

As our second check we note that the graphs of the Weibull and the fitted distributions appear to be identical and

$$\sup |\hat{F}(x) - F(x)| = 0.002765.$$  

For our third and fourth checks we obtain the following $L_1$ and $L_2$ measures.

$$||\hat{f} - f||_1 = 0.01477 \quad \text{and} \quad ||\hat{f} - f||_2 = 0.01678.$$  

3.4.8 The Lognormal Distribution

The p.d.f. of the lognormal distribution with parameters $\mu$ and $\sigma > 0$ is

$$f(x) = \begin{cases} \frac{1}{x\sigma\sqrt{2\pi}} \exp \left[ -\frac{(\ln(x) - \mu)^2}{2\sigma^2} \right], & \text{if } x > 0, \\ 0, & \text{otherwise.} \end{cases}$$

The moments of the lognormal distributions are

$$\alpha_1 = e^{\mu + \frac{\sigma^2}{2}},$$

$$\alpha_2 = (e^{\sigma^2} - 1)e^{2\mu + \sigma^2},$$

$$\alpha_3 = \sqrt{e^{\sigma^2} - 1} \left( e^{\sigma^2} + 2 \right),$$

$$\alpha_4 = e^{4\sigma^2} + 2e^{3\sigma^2} + 3e^{2\sigma^2} - 3.$$  

In the special case when $\mu = 0$ and $\sigma = 1/3$, the moments, obtained by direct computation, are

$$\alpha_1 = 1.0571, \quad \alpha_2 = 0.1313, \quad \alpha_3 = 1.0687, \quad \alpha_4 = 5.0974.$$
Figure 3.4–8. The lognormal p.d.f. with $\mu = 0$ and $\sigma = 1/3$ and its fitted GLD (the GLD p.d.f. rises higher at the center).

With the starting point of $(0.01, 0.03)$ indicated by Appendix B, we get, using FindLambdasM, the GLD fit

$$\text{GLD}(0.8451, 0.1085, 0.01017, 0.03422).$$

The support of this fit is $[-8.37, 10.06]$. For our **first check**, we plot the p.d.f.s of the lognormal with $\mu = 0$ and $\sigma = 1/3$ and the fitted GLD. This is shown in Figure 3.4–8 (the graph that rises higher at the center is the p.d.f. of GLD$(0.8451, 0.1085, 0.01017, 0.03422)$. We also compute

$$\sup |\hat{f}(x) - f(x)| = 0.09535.$$  

For our **second check** we note that the d.f.s of the two distributions are virtually identical and

$$\sup |\hat{F}(x) - F(x)| = 0.01235.$$  

For our **third and fourth checks** we note that

$$||\hat{f} - f||_1 = 0.05118 \quad \text{and} \quad ||\hat{f} - f||_2 = 0.05217.$$  

The choice of parameters in this, as well as previous distributions, has been quite arbitrary. Generally, GLD fits can be obtained with most, but not all, choices of parameters. In the case of the lognormal distribution, had we chosen $\mu = 0$ and $\sigma = 1$, the resulting moments would have been

$$\alpha_1 = 1.6487, \quad \alpha_2 = 4.6708, \quad \alpha_3 = 6.1849, \quad \alpha_4 = 113.9364,$$

making the search for a solution, if indeed one exists, quite difficult. In general, for $(\alpha_3^2, \alpha_4)$ to be within the range of computation, we must have $0 < \sigma \leq 0.55$. When $\sigma$ is small (e.g., $\sigma = 0.1$) the hazard rate of the lognormal is increasing;
when $\sigma$ is moderate (e.g., $\sigma = 0.5$) it increases and then slowly decreases; when $\sigma$ is large (e.g., $\sigma = 1.0$) it is decreasing. The latter case arises when $X = \ln(Y)$ and $Y$ is $N(0, 1)$, making $X$ lognormal with $\mu = 0$ and $\sigma = 1$. Thus, it is possible to fit the GLD in the ranges of most use in reliability applications (see, for example, Nelson (1982), p. 35).

### 3.4.9 The Beta Distribution

A random variable has the beta distribution if for parameters $\beta_3, \beta_4 > -1$, it has p.d.f.

$$f(x) = \begin{cases} \frac{x^{\beta_3}(1-x)^{\beta_4}}{\beta(\beta_3 + 1, \beta_4 + 1)}, & \text{if } 0 \leq x \leq 1, \\ 0, & \text{otherwise.} \end{cases}$$

The notation of $\beta_3$ and $\beta_4$ for the parameters of the beta distribution is used for reasons that will become clear in Chapter 4, when we consider a generalization of this distribution. If $\beta_3 = \beta_4 = 0$, this is the uniform distribution on $(0, 1)$. The moments of the beta distribution (these will be derived in Section 3.1) are

$$\alpha_1 = \frac{\beta_3 + 1}{\beta_3 + \beta_4 + 2},$$
$$\alpha_2 = \frac{(\beta_3 + 1)(\beta_4 + 1)}{(\beta_3 + \beta_4 + 2)^2(\beta_3 + \beta_4 + 3)},$$
$$\alpha_3 = \frac{2(\beta_4 - \beta_3)\sqrt{\beta_3 + \beta_4 + 3}}{(\beta_3 + \beta_4 + 4)(\beta_3 + 1)(\beta_4 + 1)},$$
$$\alpha_4 = \frac{3(\beta_3 + \beta_4 + 3)(\beta_3\beta_4(\beta_3 + \beta_4 + 2) + 3\beta_3^2 + 5\beta_3 + 3\beta_4^2 + 5\beta_4 + 4)}{(\beta_3 + \beta_4 + 5)(\beta_3 + \beta_4 + 4)(\beta_3 + 1)(\beta_4 + 1)}.$$

If we take the specific beta distribution with $\beta_3 = \beta_4 = 1$, we get

$$\alpha_1 = \frac{1}{2}, \quad \alpha_2 = \frac{1}{20}, \quad \alpha_3 = 0, \quad \alpha_4 = \frac{15}{7}.$$  

and using FindLambdasM, we obtain the fit

GLD(0.5000, 1.9693, 0.4495, 0.4495).

Our first check indicates that this is an excellent fit (with support the interval $[-0.978, 1.078]$) to the chosen beta distribution, as can be seen from Figure 3.4–9 where the two p.d.f.s are indistinguishable. Moreover,

$$\sup |\hat{f}(x) - f(x)| = 0.04717.$$  

The quality of this fit is confirmed by our second check where the d.f.s of the beta distribution with parameters $\beta_3 = \beta_4 = 1$ and its fitted GLD are virtually identical and

$$\sup |\hat{F}(x) - F(x)| = 0.0003842.$$
Figure 3.4–9. The beta p.d.f. with $\beta_3 = \beta_4 = 1$ and its fitted GLD (the two p.d.f.s are nearly indistinguishable).

For our third and fourth checks we have

$$||\hat{f} - f||_1 = 0.004495 \quad \text{and} \quad ||\hat{f} - f||_2 = 0.01057.$$ 

In spite of this good fit, in some cases the GLD is unable to provide a good approximation to beta distributions because the $(\alpha_2, \alpha_4)$ points of these distributions are outside the range of the GLD (they lie in the region marked X in Figure 3.2–4). Consider, for example, the beta distribution with $\beta_3 = -1/2$, $\beta_4 = 1$, for which

$$\alpha_1 = \frac{1}{5}, \quad \alpha_2 = \frac{8}{175}, \quad \alpha_3 = \frac{\sqrt{14}}{3}, \quad \alpha_4 = \frac{42}{11}.$$ 

In this case $(\alpha_2, \alpha_4)$ lies below the region that the GLD moments cover (see (3.3.11) and Figure 3.2–5). Only some of the $(\alpha_2, \alpha_4)$ points in the BETA REGION of Figure 3.2–5 are in the GLD $(\alpha_2, \alpha_4)$ region that was given in (3.3.11). This motivates (in Chapter 4) extending the GLD to cover the portion of $(\alpha_2, \alpha_4)$-space below the region described in (3.3.11).

### 3.4.10 The Inverse Gaussian Distribution

The p.d.f. of the inverse Gaussian distribution with parameters $\mu > 0$ and $\lambda > 0$ is given by

$$f(x) = \begin{cases} \sqrt{\frac{\lambda}{2\pi x^3}} \exp \left[ -\frac{\lambda(x - \mu)^2}{2\mu^2 x} \right], & \text{if } x > 0, \\ 0, & \text{otherwise.} \end{cases}$$ 

This distribution has applications to problems associated with the times required to cover fixed distances in linear Brownian motion. The reader may wish to...
3.4 GLD Approximations of Some Well-Known Distributions

![Figure 3.4–10](image)

The inverse Gaussian p.d.f. ($\mu = 0.5$ and $\lambda = 6$) and its fitted GLD (the GLD p.d.f. rises higher at the center).

consult Govindarajulu (1987), p. 611. The moments of this distribution are

\[
\alpha_1 = \mu, \quad \alpha_2 = \frac{\mu^3}{\lambda}, \quad \alpha_3 = 3\sqrt{\frac{\mu}{\lambda}}, \quad \alpha_4 = 3 + \frac{15\mu}{\lambda}.
\]

We first consider the special case of $\mu = 0.5$ and $\lambda = 6$ with

\[
\alpha_1 = \frac{1}{2}, \quad \alpha_2 = \frac{1}{48}, \quad \alpha_3 = \frac{\sqrt{3}}{2}, \quad \alpha_4 = \frac{17}{4}.
\]

Use of FindLambdasM gives the fit

GLD(0.4164, 0.6002, 0.02454, 0.08009).

We note that the support of this GLD is $[-1.25, 2.08]$ and proceed to our first check by considering the graphs of the p.d.f.s of this inverse Gaussian distribution and its fitted GLD (given in Figure 3.4–10) and determine that

\[
\sup |\hat{f}(x) - f(x)| = 0.1900.
\]

For our second check, we note that it is impossible to graphically distinguish the d.f. of this inverse Gaussian distribution from that of its fitted GLD and

\[
\sup |\hat{F}(x) - F(x)| = 0.01053.
\]

The $L_1$ and $L_2$ measures for the third and fourth checks, are

\[
||\hat{f} - f||_1 = 0.04623 \quad \text{and} \quad ||\hat{f} - f||_2 = 0.07151.
\]
3.4.11 The Logistic Distribution

The p.d.f. of the logistic distribution, with parameters $\mu$ and $\sigma > 0$, is given by

$$f(x) = \frac{e^{-\frac{(x-\mu)}{\sigma}}}{\sigma (1 + e^{-\frac{(x-\mu)}{\sigma}})^2} \quad \text{for} \quad -\infty < x < \infty.$$  

For this distribution,

$$\alpha_1 = \mu, \quad \alpha_2 = \frac{\pi^2 \sigma^2}{3}, \quad \alpha_3 = 0, \quad \alpha_4 = \frac{21}{5}.$$  

The $\alpha_3$ and $\alpha_4$ do not depend on the parameters of the logistic distribution and $(\alpha_3, \alpha_4) = (0, 4.2)$ is an entry in Table B–1 of Appendix B.

In the special case of $\mu = 0$ and $\sigma = 1$, we use Table B–1 of Appendix B to obtain the fit

$$\text{GLD}( -7.1 \times 10^{-6}, -0.0003246, -0.0003244, -0.0003244)$$

and note that the support of the fitted distribution is $(-\infty, \infty)$. Next, we observe that the p.d.f.s of this logistic distribution and its fitted GLD are graphically indistinguishable, as shown in Figure 3.4–11. Moreover,

$$\sup |\hat{f}(x) - f(x)| = 0.00008419.$$  

This takes care of our first check. For our second check, we consider the d.f.s of the two distributions and observe that they look identical and

$$\sup |\hat{F}(x) - F(x)| = 0.00006202.$$  

Figure 3.4–11. The logistic p.d.f. ($\mu = 0$ and $\sigma = 1$) and its fitted GLD. The two p.d.f.s cannot be distinguished.
For our third and fourth checks we note that
\[ ||\hat{f} - f||_1 = 0.0002744 \quad \text{and} \quad ||\hat{f} - f||_2 = 0.0001077. \]

By all of our criteria, this is an unusually good approximation and when we look at the \( \lambda_1, \lambda_2, \lambda_3, \lambda_4 \) of this approximation we see that \( \lambda_1 \) is very close to zero and \( \lambda_2, \lambda_3, \lambda_4 \) are almost equal to each other. This prompts us to consider distributions of type GLD(0, \( \lambda, \lambda, \lambda \)) that have quantile function
\[ Q(y) = \frac{y^\lambda - (1 - y)^\lambda}{\lambda}. \]

It can be shown that
\[ \lim_{\lambda \to 0} Q(y) = \ln\left(\frac{y}{1-y}\right) \]
converges to the quantile function of the logistic distribution with \( \mu = 0 \) and \( \sigma = 1 \), indicating that arbitrarily good approximations to the logistic with \( \mu = 0 \) and \( \sigma = 1 \) are attainable through the GLD.

Investigating GLD approximations to logistic distributions with other values of \( \mu \) and \( \sigma \), we conjecture that GLD(\( \mu, \lambda/\sigma, \lambda, \lambda \)) should provide a good approximation to the general logistic distribution, when \( \lambda \) is small. To establish this result, we note that the quantile function of this GLD is given by
\[ Q(y) = \mu + \sigma \frac{y^\lambda - (1 - y)^\lambda}{\lambda} \]
and
\[ \lim_{\lambda \to 0} Q(y) = \mu + \sigma \ln(y) - \sigma \ln(1 - y) = \mu + \sigma \ln\left(\frac{y}{1-y}\right) \].

To obtain the d.f. of the GLD(\( \mu, \lambda/\sigma, \lambda, \lambda \)) (in the limit as \( \lambda \to 0 \)) distribution, we set the last expression in (3.4.9) to \( x \)
\[ x = \mu + \sigma \ln\left(\frac{y}{1-y}\right), \]
and rearrange to obtain
\[ \frac{x - \mu}{\sigma} = \ln\left(\frac{y}{1-y}\right), \]
or
\[ e^{(x-\mu)/\sigma} = \frac{y}{1-y}. \]

Solving this equation for \( y \) we can represent the d.f. of GLD(\( \mu, \lambda/\sigma, \lambda, \lambda \)) (in the limit as \( \lambda \to 0 \)) by
\[ F(x) = \frac{1}{1 + e^{-(x-\mu)/\sigma}}. \]
Now differentiating with respect to \( x \) gives a p.d.f that is identical to the p.d.f. of the logistic distribution with parameters \( \mu \) and \( \sigma \). We conclude that arbitrarily
good approximations to the logistic distributions can be attained by considering the GLD(\(\mu, \lambda/\sigma, \lambda, \lambda\)) when \(\lambda\) is sufficiently small.

The logistic is used as a life distribution (see Nelson (1982)), and is not a member of the Pearson family (see Freimer, Kollia, Mudholkar, and Lin (1988), p. 3560). Since the loglogistic distribution deals with \(Y = e^X\) where \(X\) is logistic, by taking logarithms, we can reduce the loglogistic to the logistic.

### 3.4.12 The Largest Extreme Value Distribution

The largest extreme value distribution, with parameters \(\mu\) and \(\sigma > 0\), has p.d.f.

\[
f(x) = \frac{1}{\sigma} e^{-(x-\mu)/\sigma} \exp \left[ -e^{-(x-\mu)/\sigma} \right], \quad \text{for} \quad -\infty < x < \infty.
\]

This distribution can be used to describe human life (see Nelson (1982), p. 40) and has moments

\[
\alpha_1 = \mu + \gamma\sigma, \quad \alpha_2 = \frac{\pi^2\sigma^2}{6}, \quad \alpha_3 = \sqrt{\frac{1.29857}{\pi^4}}, \quad \alpha_4 = 5.4,
\]

where \(\gamma \approx 0.57722\) is Euler’s constant.

Since \(\alpha_3\) and \(\alpha_4\) are independent of the parameters of the distribution, we can consider fitting a specific distribution, knowing that fits for other distributions of this family can be obtained in a similar manner. If we set \(\mu = 0\) and \(\sigma = 1\),

\[
\alpha_1 = .5772, \quad \alpha_2 = 1.6449, \quad \alpha_3 = 1.1395, \quad \alpha_4 = 5.4.
\]

Using FindLambdasM, we obtain the fit

\[
\text{GLD( } -0.1859, 0.02109, 0.006701, 0.02284)\]

and note that the support of this GLD is \([-47.647, 47.275]\). We compare the p.d.f. of this distribution and that of its fitted GLD. This comparison, illustrated in Figure 3.4–12, yields

\[
\sup |\hat{f}(x) - f(x)| = 0.02216
\]

for our first check. Next, we compare the d.f.s of the extreme value distribution with \(\mu = 0\) and \(\sigma = 1\) to observe that the d.f.s cannot be visually distinguished and

\[
\sup |\hat{F}(x) - F(x)| = 0.01004,
\]

completing our second check. The third and fourth checks indicate

\[
||\hat{f} - f||_1 = 0.04093 \quad \text{and} \quad ||\hat{f} - f||_2 = 0.02284.
\]
3.4 GLD Approximations of Some Well-Known Distributions

3.4.13 The Extreme Value Distribution

This distribution, also called the smallest extreme value distribution, is important in some applications. For example, Weibull (1951) reported that the strength of a certain material follows an extreme value distribution with \( \mu = 108\, \text{kg/cm}^2 \) and \( \sigma = 9.27\, \text{kg/cm}^2 \) (also see Nelson (1982), p. 41).

Also, if \( X \) is Weibull, then \( \ln(X) \) has an extreme value distribution, making this relationship similar to the one between the lognormal and the normal distributions. It is also known (see Kendall and Stuart (1969), pp. 85, 335, 344) that if \( X \) has the extreme value distribution then the p.d.f. of \( (X - \mu)/\sigma \) is \( f(x) = e^{-x-e^{-x}} \). The extreme value distribution is also used in life and failure data analysis, “weakest link” situations, temperature minima, rainfall in droughts, human mortality of the aged, etc., and often represents the first failure (which fails the unit).

The extreme value distribution has parameters \( \mu \) and \( \sigma > 0 \) and p.d.f.

\[
f(x) = \frac{1}{\sigma} e^{(x-\mu)/\sigma} \exp \left[ -e^{(x-\mu)/\sigma} \right], \quad \text{for} \quad -\infty < x < \infty.
\]

If \( Y \) is extreme value with parameters \( \mu_Y \) and \( \sigma_Y \), then \( X = -Y \) is largest extreme value with parameters \( \mu_X = -\mu_Y \) and \( \sigma_X = \sigma_Y \). This follows since

\[
P(Y \leq y) = 1 - e^{-e^{(y-\mu_Y)/\sigma_Y}}.
\]

Hence,

\[
P(X \leq x) = P(-Y \leq x) = P(Y \geq -x) = 1 - P(Y \leq -x) = 1 - \left( 1 - e^{-e^{(-x-\mu_Y)/\sigma_Y}} \right) = e^{-e^{(x+\mu_Y)/\sigma_Y}} = e^{-e^{-(x-\mu_X)/\sigma_X}}.
\]
Chapter 3: Fitting Distributions and Data with the GLD via Moments

The moments of the two distributions have the relation
\[ \alpha_1(X) = -\alpha_1(Y), \quad \alpha_2(X) = \alpha_2(Y), \quad \alpha_3(X) = -\alpha_3(Y), \quad \alpha_4(X) = \alpha_4(Y). \]

Thus, to fit a largest extreme value with moments \( (5, 16, 0.5, 9.2) \) is the same as fitting an extreme value with moments \( (-5, 16, -0.5, 9.2) \). As we have already considered fitting a largest extreme value distribution, Section 3.4.12 has the details for fitting an extreme value distribution.

The moments of the distribution are
\[ \alpha_1 = \mu - \gamma \sigma, \quad \alpha_2 = \frac{1}{6} \pi^2 \sigma^2, \quad \alpha_3 = -\sqrt{1.29857}, \quad \alpha_4 = 5.4, \]
where \( \gamma \approx 0.57722 \) is Euler’s constant.

### 3.4.14 The Double Exponential Distribution

The double exponential distribution with parameter \( \lambda > 0 \) has p.d.f.
\[ f(x) = \frac{e^{-|x|/\lambda}}{2\lambda} \quad \text{for} \quad -\infty < x < \infty. \]

The moments of this distribution are
\[ \alpha_1 = 0, \quad \alpha_2 = 2\lambda^2, \quad \alpha_3 = 0, \quad \alpha_4 = 6. \]

Since \( \alpha_3 \) and \( \alpha_4 \) are constants, we will be able to find GLD\( (\lambda_1, \lambda_2, \lambda_3, \lambda_4) \) fits to this distribution if we can find a fit when \( \lambda = 1 \) and the moments are
\[ \alpha_1 = 0, \quad \alpha_2 = 2, \quad \alpha_3 = 0, \quad \alpha_4 = 6. \]

In this situation, FindLambdasM provides the fit
\[ \text{GLD}(2.8 \times 10^{-10}, -0.1129, -0.08020, -0.08020). \]

Note that the support of this GLD is \( (-\infty, \infty) \). We now compare the p.d.f.s of the double exponential distribution with \( \lambda = 1 \) to that of the fitted GLD. The graphs of these p.d.f.s are given in Figure 3.4–13 (a) where the graph of the double exponential p.d.f. rises higher in the center. Our observation that
\[ \sup |\hat{f}(x) - f(x)| = 0.1485 \]
completes the first check. For the second check, we compute
\[ \sup |\hat{F}(x) - F(x)| = 0.02871 \]
and obtain a plot of the d.f.s of the two distributions, shown in Figure 3.4–12 (b). Note that in contrast to most of the fits provided so far, there is an observable difference between the two d.f.s (the double exponential d.f. is lower immediately to the left of the center and higher immediately to the right of the center).

For our third and fourth checks we note that
\[ ||\hat{f} - f||_1 = 0.1292 \quad \text{and} \quad ||\hat{f} - f||_2 = 0.081727151. \]
3.4 GLD Approximations of Some Well-Known Distributions

Figure 3.4–13. The p.d.f.s (a) and d.f.s (b) of the double exponential (λ = 1) and its fitted GLD.

3.4.15 The F-Distribution

The p.d.f. of the $F$ distribution, with $\nu_1 > 0$ and $\nu_2 > 0$ degrees of freedom, $F(\nu_1, \nu_2)$, is given by

$$f(x) = \frac{\Gamma((\nu_1 + \nu_2)/2)(\nu_1/\nu_2)^{\nu_1/2}}{\Gamma(\nu_1/2)\Gamma(\nu_2/2)} \times \frac{x^{(\nu_1-2)/2}}{(1 + x^{\nu_1/\nu_2})^{(\nu_1+\nu_2)/2}}$$

when $x > 0$ and $f(x) = 0$ when $x \leq 0$.

The power of $x$ in the p.d.f. of the $F$ distribution is $-1 - \nu_2/2$. Therefore for the $i$-th moment to exist, we must have $i - 1 - \nu_2/2 < -1$ or $\nu_2 > 2i$. This immediately restricts the moment-based GLD($\lambda_1, \lambda_2, \lambda_3, \lambda_4$) fits to those $F$ distributions with $\nu_2 > 8$.

The moments of the $F$ distribution (when $\nu_2 > 8$), are

$$\alpha_1 = \frac{\nu_2}{\nu_2 - 2},$$

$$\alpha_2 = \frac{\nu_2^2(\nu_1 + \nu_2 - 2)}{\nu_1(\nu_2 - 4)(\nu_2 - 2)^2},$$

$$\alpha_3 = \frac{2\sqrt{\nu_1}(\nu_2 - 2 + 2\nu_1)}{\nu_1(\nu_2 - 6)\sqrt{\nu_1/\nu_2 - 2}},$$

$$\alpha_4 = \frac{3(\nu_1\nu_2^2 + 4\nu_2^2 + 8\nu_1\nu_2 + \nu_1^2\nu_2 - 16\nu_2 + 10\nu_1^2 - 20\nu_1 + 16)(\nu_2 - 4)}{\nu_1(\nu_1 + \nu_2 - 2)(\nu_2 - 6)(\nu_2 - 8)}.$$

If we set $\nu_2 = 10$, $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ will exist. However, if we also choose $\nu_1 = 5$, then we have $(\alpha_3, \alpha_4) = (14.9538, 53.8615)$, taking us outside the range of the...
table in Appendix B as well as outside our range of computation. $\alpha_3^2$ and $\alpha_4$ both decrease with increasing $\nu_1$, but when $\nu_2 = 10$, this does not help since

$$\lim_{\nu_1 \to \infty} \alpha_3^2 = 12, \quad \lim_{\nu_1 \to \infty} \alpha_4 = 45,$$

and $(\alpha_3^2, \alpha_4)$ stays out of range for all $\nu_1 > 0$.

More generally, by considering $\alpha_3^2$ as a function of $\nu_1$ and $\nu_2$ and taking the derivative of $\alpha_3^2(\nu_1, \nu_2)$ with respect to $\nu_2$, we can determine that $\alpha_3^2(\nu_1, \nu_2)$ decreases with increasing $\nu_2$ when $\nu_2 > 8$. This implies that for $8 \leq \nu_2 \leq 15$,

$$\alpha_3^2(\nu_1, \nu_2) \geq \alpha_3^2(\nu_1, 15) = \frac{88(2\nu_1 + 13)^2}{81(\nu_1(\nu_1 + 13))} \approx 4.35,$$

placing $(\alpha_3^2, \alpha_4)$ outside of our computational range whenever $\nu_2 \leq 15$.

With similar analyses, we can determine that in order to have $\alpha_3^2 \leq 4$ when $\nu_2 = 16$, we must have $\nu_1 \geq 28$; when $\nu_2 = 20$, we must have $\nu_1 \geq 6.26$; and when $\nu_2 = 30$, we must have $\nu_1 \geq 3.52$.

It is, therefore, possible to obtain fits for large $\nu_2$. For example, if we let $\nu_2 = 25$ and $\nu_1 = 6$, we get

$$\alpha_1 = 1.0869, \quad \alpha_2 = 0.5439, \quad \alpha_3 = 1.8101, \quad \alpha_4 = 9.1986$$

and the GLD($\lambda_1, \lambda_2, \lambda_3, \lambda_4$) fit

$$\text{GLD}(0.6457, -0.06973, -0.01100, -0.04020)$$

by using the entries of Appendix B. We observe that the support of this fit is $(-\infty, \infty)$. For the first check, we compare the p.d.f. of $F(6, 25)$ with that of its GLD fit (see Figure 3.4–14) and compute

$$\sup |\hat{f}(x) - f(x)| = 0.1361.$$ 

For the second check, we note that d.f.s of $F(6, 25)$ and the fitted GLD cannot be visually distinguished and

$$\sup |\hat{F}(x) - F(x)| = 0.02684.$$ 

The $L_1$ and $L_2$ distances for this approximation are

$$||\hat{f} - f||_1 = 0.09612 \quad \text{and} \quad ||\hat{f} - f||_2 = 0.08323.$$
3.4 GLD Approximations of Some Well-Known Distributions

Figure 3.4–14. The p.d.f.s of $F(6,25)$ and its fitted GLD (the one that rises higher at the center is the p.d.f. of the fitted GLD).

In similar ways, we can get approximations for some other $F$-distributions. Specifically, for $(\nu_1, \nu_2) = (6,12)$ and $(6,16)$, we have, respectively,

$$(\lambda_1, \lambda_2, \lambda_3, \lambda_4) = (0.7192, -0.2705, -0.05458, -0.1581),$$

with

$$\sup |\hat{f}(x) - f(x)| = 0.21138$$

and

$$\sup |\hat{F}(x) - F(x)| = 0.04733$$

$$||\hat{f} - f||_1 = 0.2121$$

and

$$||\hat{f} - f||_2 = 0.14566;$$

$$(\lambda_1, \lambda_2, \lambda_3, \lambda_4) = (0.6193, -0.1731, -0.02495, -0.1041),$$

with

$$\sup |\hat{f}(x) - f(x)| = 0.1347$$

and

$$\sup |\hat{F}(x) - F(x)| = 0.02428$$

$$||\hat{f} - f||_1 = 0.08817$$

and

$$||\hat{f} - f||_2 = 0.07911.$$

3.4.16 The Pareto Distribution

The Pareto distribution, with parameters $\beta > 0$ and $\lambda > 0$, has p.d.f.

$$f(x) = \begin{cases} \frac{\beta \lambda^\beta}{x^{\beta+1}}, & \text{if } x > \lambda, \\ 0, & \text{otherwise}. \end{cases}$$

The moments of the Pareto distributions are

$$\alpha_1 = \frac{\beta \lambda}{\beta - 1},$$

$$\alpha_2 = \frac{\beta \lambda^2}{(\beta - 1)^2 (\beta - 2)}.$$
Figure 3.4–15. The $\alpha_3^2$ and $\alpha_4$ of the Pareto distribution as functions of $\beta$ ($\alpha_4$ is the higher curve).

\[ \alpha_3 = \frac{2(\beta + 1)}{(\beta - 3) \sqrt{2}} \]
\[ \alpha_4 = \frac{3(3 \beta^2 + \beta + 2)(\beta - 2)}{\beta(\beta - 3)(\beta - 4)} \]

We see that $\alpha_3$ and $\alpha_4$ depend only on the single parameter $\beta$ and for $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ to exist, we must have $\beta > 4$. For almost all reasonable choices of $\beta$, $\alpha_3$ and $\alpha_4$ are out of range for us to be able to compute $\lambda_1, \lambda_2, \lambda_3, \lambda_4$. For example, when $\beta = 5$, $\alpha_3 = 4.6476$ and $\alpha_4 = 73.8000$. As $\beta$ increases both $\alpha_3$ and $\alpha_4$ decrease. However,

\[ \lim_{\beta \to \infty} \alpha_3 = 2 \quad \text{and} \quad \lim_{\beta \to \infty} \alpha_4 = 9 \]

so $\beta$ has to get quite large for $\alpha_3$ and $\alpha_4$ to come near our tabulated range. This can be seen from the graphs of $\alpha_3^2$ and $\alpha_4$ given in Figure 3.4–15 where $\alpha_4$ is the higher of the two curves.

It is worthwhile to point out that despite the difficulties associated with finding fits for the Pareto distribution, this distribution can be realized as a special case of the GLD. In fact, for all $\beta > 4$, the Pareto distribution with parameters $\lambda$ and $\beta$ is identical to

\[ \text{GLD} \left( \lambda, \frac{-1}{\lambda}, 0, \frac{-1}{\beta} \right) \]

This can be seen by comparing the quantile function of the Pareto with parameters $\lambda$ and $\beta$ with that of the GLD with parameters $\lambda, -1/\lambda, 0, -1/\beta$; both produce the quantile function

\[ Q(y) = \lambda(1 - y)^{-1/\beta} \]
3.5 Examples: GLD Fits of Data, Method of Moments

We have already seen details of the shapes of GLD distributions (Section 2.4), a GLD fitted to an actual dataset (the fit to measurements of the coefficient of friction of a metal at (2.1.6)), and fits to many of the most important distributions encountered in applications in various areas.

In this section we consider fits of the GLD to actual datasets. A number of examples will be given, from a variety of fields, in order to illustrate a spectrum of nuances that arise in the fitting process. In this section, we will fit datasets for which the estimated \((\hat{\alpha}_3, \hat{\alpha}_4)\) pair (denoted by \((\hat{\alpha}_3, \hat{\alpha}_4)\)) is in the region covered by the GLD that was given in Figure 3.2–4. The reasons for this restriction are twofold: first, as we saw in Section 3.4.9, when we attempt to fit distributions that have (skewness, kurtosis) that is not in the area covered by the GLD (wide though this is) the fit will usually not be excellent; second, in Chapter 4 we will extend the GLD to an EGLD that covers the \((\alpha_3, \alpha_4)\) points not covered by the GLD so it makes sense to consider such examples in Chapter 4.

It is true that there is variability of sampling in the estimates of (skewness\(^2\), kurtosis): even if the true \((\alpha_3, \alpha_4)\) point is in the GLD region, the estimate from the data might not be or if the true point is not in the GLD region, the estimate from the data might be. Thus, one may, in applications, end up fitting by a model that cannot cover the true distribution very well. For this reason, when we take up these additional examples in Chapter 4, after we extend the GLD to the EGLD, we will attempt to fit both a GLD and an EGLD and compare the two fits.

We should also note, as we will see in detail in Chapter 4, that the method of extension of the GLD is such that there is a zone of overlap of the \((\alpha_3, \alpha_4)\) points of the two models. This means that we in fact have a zone where both model types will fit the data well, so “wrong zone” (skewness\(^2\), kurtosis) should be less of a problem in applications when \((\hat{\alpha}_3, \hat{\alpha}_4)\) is “near” the boundary.

In the examples we discuss we have both real datasets from the literature, and datasets with simulated (Monte Carlo) data. While with a real dataset we can assess how well the model fits the data, we cannot assess how well it fits the underlying true distribution (as that distribution is not known). With simulated datasets we can do both, which is why we have included them.

3.5.1 Assessment of Goodness-of-Fit

There are a number of aspects of “goodness-of-fit,” a topic on which many papers and books have been written that are relevant to the subject area of this book. We will cover them mainly in this section. The situation we have is as follows:
Chapter 3: Fitting Distributions and Data with the GLD via Moments

S–1. We are seeking to model a phenomenon $X$ of interest in some area of research.

S–2. There are certain distributions $F_1, F_2, \ldots, F_k$ that we wish to include, in some form, among those we wish to consider to describe $X$. These may be ones that it is believed $X$ truly follows (e.g., they are derived from assumptions $X$ is believed to obey), or they may simply be ones that yielded reasonable approximations in the previous studies.

S–3. We have available data on the phenomenon $X$ in some form, for example,

1. Independent observations $X_1, X_2, \ldots, X_n$ on the phenomenon.
2. A histogram of the distribution of $X$ based on data (but the data itself is not available to us).
3. The sample moments $\hat{\alpha}_1, \hat{\alpha}_2, \hat{\alpha}_3, \hat{\alpha}_4$ based on data (but we do not have the data).
4. Some other form of information based on data is available.

S–4. We are considering using a family of distributions $G(\omega)$ to fit the phenomenon’s p.d.f./d.f./p.f., where $\omega$ is the vector of parameters, which (when chosen) picks a particular member of the family $G(\omega)$. For example, if $G$ is the GLD family then $\omega = (\lambda_1, \lambda_2, \lambda_3, \lambda_4)$.

We make the following points regarding the circumstances described in S–1 through S–4.

Point 1. In most cases, the $F_i$s of S–2 will merely be approximations to a true (and unknown) distribution of $X$. If they are truly the distributions that $X$ follows, and have one or more unknown parameters, then one should act as though $X$ is in the family $F_i(\omega)$ and use techniques which optimally choose $\omega$ for that family.

Point 2. In light of Point 1 it is not necessarily required that the fitted $G$ distribution have all of the characteristics of $F_i$ of S–2. For example, in the first paragraph of this chapter we noted that often the support (range of possible values) of $F_i$ may be ones we know cannot occur; in such a setting we certainly do not want to require that $G$ have the same support as $F_i$.

Point 3. In light of S–1, we will want the family $G(\omega)$ to be one that can come reasonably close to including $F_1, F_2, \ldots, F_k$ when $\omega$ is chosen appropriately. For this reason, in Section 3.4 we described how close the GLD family can come to various widely used distributions, using a fit based on matching the first four moments of the GLD to the first four moments of the particular $F_i$. We might also fit with the EGLD of Chapter 4, or with a Method of Percentiles as in Chapter 5, especially if we did not find the fit to $F_i$ satisfactory. If $F_i$ is
fitted with $G_0$, then the question arises: Is $G_0$ a good approximation to $F_i$? Some ways to approach this question include

- Assess the adequacy and appropriateness of the support of $G_0$.
- Compare the p.d.f.s, d.f.s, and p.f.s of $G_0$ with those of $F_i$.
- Compare $G_0$ with $F_i$ on other measures that “matter” for the application under consideration; for example, moments, probabilities of key events, hazard functions, and so on.

These are not simple questions for which there is a simple formula that can provide a correct answer — they are some of the most difficult questions that arise in modeling, and usually require serious interaction between the statistician and the subject matter specialists — but it is essential that they be carefully and fully considered if we are to have confidence in the model.

**Point 4.** With data of any of the forms described in S–3, we can use the Method of Moments to fit a GLD. In case of S–3–2, the moments are approximated using the midpoint assumption; with S–3–1 and S–3–2, the Method of Percentiles (presented in Chapter 5) can be used to fit a GLD; in case of S–3–2, the percentiles needed are approximated (an analog of case S–3–3 is to have available certain sample percentiles). What can be done with situation S–3–4 depends on the specifics of the information available.

**Point 5.** After we fit a distribution $G$ to the data the question will arise: Is this a good fit to the true underlying distribution? Since the true distribution is unknown, this is a difficult question. In case S–3–3 we can reasonably ask that $G$ have values close to the specified sample moments (or, if sample percentiles were given, to the sample percentiles). In case S–3–2 we should **overplot the histogram of the data and p.d.f. $g$ of the d.f. $G$** and examine key shape elements. In case S–3–1 we can **overplot the d.f. $G$ and the empiric d.f. of the data** and examine their closeness.

**Point 6.** There are many **statistical tests of the hypothesis** where the data, $X_1, X_2, \ldots, X_n$, come from the distribution $G$. These are in addition to the **eyeball test**, which should always be used (and in which many experimenters place the most faith). Based on the theorem that follows we assert: The hypothesis that the data come from $G$ is **equivalent to the hypothesis that $G(\frac{X_1}{\text{sample size}}), G(\frac{X_2}{\text{sample size}}), \ldots, G(\frac{X_n}{\text{sample size}})$ are independent uniform r.v.s on $(0, 1)$**.

**Theorem 3.5.1.** If a r.v. $Y$ has continuous d.f. $H(y)$, then the r.v. $Z = H(Y)$ has the uniform distribution on $(0, 1)$.

There are many statistical tests for this hypothesis such as the Kolmogorov–Smirnov $D$, Cramér–von Mises $W^2$, Kuiper $V$, Watson $U^2$, Anderson–Darling $A^2$, log–statistic $Q$, $\chi^2$, entropy, as well as other tests of uniformity. Some references
are Dudewicz and van der Meulen (1981) (where the entropy test is developed, and which has references to the literature) and Shapiro (1980) (which, while oriented to testing normality, has excellent comments on goodness-of-fit testing in general).

Of the goodness-of-fit tests, the oldest is the chi-square test proposed by Pearson (1900). The idea of the test is to divide the range of the distribution into $k$ cells and compare the observed number in each cell to the number that would be expected if the assumed distribution is true. Under certain construction of the cells, the resulting test statistic has (approximately) a chi-square distribution with degrees of freedom $k - 1 - t$ where $t$ is the number of parameters estimated (e.g., $t = 4$ for the GLD). Drawbacks of the test are

- low power (ability to reject an incorrect fit) vs. other tests;
- loss of information when the data are grouped into cells;
- arbitrariness of the choice of the cells.

For reference on the chi-square test, see Moore (1977). We illustrate the chi-square test in the examples that we consider below.

We should note why some experimenters place the most faith in the eyeball test, though they may also do a formal test of the hypothesis:

- the chi-square test itself is approximate (and assumes use of the method of maximum likelihood, rather than the method of moments);
- for any test, failure to reject does not mean the hypothesis is true; it could, for example, be that the sample size is too small to detect differences that exist between the true and hypothesized models;
- for any test, rejection does not mean that the hypothesized model is inappropriate for the purpose intended.

However, the eyeball test requires experience. Therefore, we recommend starting with a hypothesis test and if we then wish to go against the conclusion of the test based on examination of the overplot of hypothesis and data, make the case for reversal. This is analogous to a lawyer making a case at an appellate level. As Gibbons (1997, p. 81) states, “the investigator hopes to be able to accept the null hypothesis, even when it appears to be only nearly true.” The eyeball test is not a license to follow a personal whim, and it is important to guard against any lack of rigor in so important a decision as choice of distributional model for the phenomenon under study. The eyeball test is sometimes called a graphical test (see Ricer (1980), p. 18). If the plot is of the e.d.f. and the fitted $G$, we might look for criss-crossing of the two rather than one staying consistently on one side of the other. We should always be aware that there is subjectivity to the analysis.
3.5 Examples: GLD Fits of Data, Method of Moments

The Kolmogorov–Smirnov or KS test is based on the largest difference (in absolute value) between the e.d.f. and its hypothesized counterpart. This can be interpreted as a quantification of the eyeball test. Only the biggest deviation is used; if one curve is uniformly on one side of the other but does not reach very far away at any one point, the test will not reject; for this reason, the test is sometimes viewed as lacking in power. In each of the examples that follow, if we are able to obtain a fit, we give the KS statistic for that fit. Since it is more informative to have the $p$-values for such tests, the $p$-values (obtained from the table in Appendix H) are also given. A good comparison of the KS and chi-square tests is given by Gibbons (1997, pp. 80–82). The use of the KS statistic to give a confidence interval for the true d.f. (see Gibbons (1997), Section 3.2) is also of some interest.

There is a large literature on quantification of eyeball tests; one article, with references, worth considering is Gan, Koehler, and Thompson (1991).

3.5.2 Example: Cadmium in Horse Kidneys

Elinder, Jönsson, Piscator and Rahnster (1981) investigated histopathological changes due to different levels of metal concentrations in horse kidneys. We list below part of their data dealing with the average cadmium concentrations in the kidney cortex of horses.

<table>
<thead>
<tr>
<th></th>
<th>11.9</th>
<th>16.7</th>
<th>23.4</th>
<th>25.8</th>
<th>25.9</th>
<th>27.5</th>
<th>28.5</th>
<th>31.1</th>
</tr>
</thead>
<tbody>
<tr>
<td>32.5</td>
<td>35.4</td>
<td>38.3</td>
<td>38.5</td>
<td>41.8</td>
<td>42.9</td>
<td>50.7</td>
<td>52.3</td>
<td></td>
</tr>
<tr>
<td>52.5</td>
<td>52.6</td>
<td>54.5</td>
<td>54.7</td>
<td>56.6</td>
<td>56.7</td>
<td>58.0</td>
<td>60.8</td>
<td></td>
</tr>
<tr>
<td>61.8</td>
<td>62.3</td>
<td>62.5</td>
<td>62.6</td>
<td>63.0</td>
<td>67.7</td>
<td>68.5</td>
<td>69.7</td>
<td></td>
</tr>
<tr>
<td>73.1</td>
<td>76.0</td>
<td>76.9</td>
<td>77.7</td>
<td>78.2</td>
<td>80.3</td>
<td>93.7</td>
<td>101.0</td>
<td></td>
</tr>
<tr>
<td>104.5</td>
<td>105.4</td>
<td>107.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

For this data, using the program FitGLDM we find

\[ \hat{\alpha}_1 = 57.2442, \quad \hat{\alpha}_2 = 576.0741, \quad \hat{\alpha}_3 = 0.2546, \quad \hat{\alpha}_4 = 2.5257 \]

and obtain the GLD fit

\[ \text{GLD}(41.7897, 0.01134, 0.09853, 0.3606) \]

whose support is \([-46.355, 129.935]\).

Figure 3.5–1 (a) shows the p.d.f. of the fitted GLD with a histogram of the data; Figure 3.5–1 (b) shows the e.d.f. of the data with the d.f. of the fitted GLD. These figures indicate a reasonably good fit which is substantiated with a chi-square test. The test uses the classes

\[ (-\infty, 30), \quad [30, 50), \quad [50, 60), \quad [60, 70), \quad [70, 85), \quad [85, \infty) \]
whose respective frequencies are

\[ 7, 7, 9, 9, 6, 5. \]

Calculation of expected frequencies yields

\[ 5.5633, 12.3032, 6.5387, 5.8075, 6.6775, 6.1097 \]

and the chi-square statistic and the \( p \)-value that result are

\[ 5.6087 \quad \text{and} \quad 0.01787, \]

respectively. The Kolmogorov-Smirnov statistic for this GLD fit is \( KS = 0.1025 \) and \( KS\sqrt{n} = 0.1025\sqrt{43} = 0.6721 \). From the table in Appendix J we see that the \( p \)-value for this test is 0.76.

### 3.5.3 Example: Brain (Left Thalamus) MRI Scan Data

Dudewicz, Levy, Lienhart, and Wehrli (1989) give data on the brain tissue MRI scan parameter, AD. It should be noted that the term “parameter” is used differently in brain scan studies — it is used to designate what we would term random variables. In the cited study the authors show that \( AD^{-2} \) has a normal distribution while \( AD \) does not, and report the following 23 observations associated with scans of the left thalamus.

\[
\begin{align*}
108.7 &\quad 107.0 &\quad 110.3 &\quad 110.0 &\quad 113.6 &\quad 99.2 &\quad 109.8 &\quad 104.5 \\
108.1 &\quad 107.2 &\quad 112.0 &\quad 115.5 &\quad 108.4 &\quad 107.4 &\quad 113.4 &\quad 101.2 \\
98.4 &\quad 100.9 &\quad 100.0 &\quad 107.1 &\quad 108.7 &\quad 102.5 &\quad 103.3 \\
\end{align*}
\]
3.5 Examples: GLD Fits of Data, Method of Moments

The computations through FitGLDM indicate

\[ \hat{\alpha}_1 = 106.8349, \quad \hat{\alpha}_2 = 22.2988, \quad \hat{\alpha}_3 = -0.1615, \quad \hat{\alpha}_4 = 2.1061 \]

and provide the fit

\[ \text{GLD}(112.1465, 0.06332, 0.6316, 0.05350) \]

with support \([96.353, 127.940]\).

Figure 3.5–2 (a) shows the p.d.f.s of this fit and a histogram of the data; Figure 3.5–1 (b) shows the e.d.f. of the data with the d.f. of the GLD fit.

With the small sample size of this example, we are not able to perform a chi-square test but we can partition the data into classes, such as

\[ (-\infty, 103), \quad [103, 108), \quad [108, 111), \quad [111, \infty), \]

whose respective frequencies are

6, 6, 7, 4,

and calculate the chi-square statistics for this to obtain 1.4715. For this fit \( KS = 0.1438, \ KS\sqrt{n} = 0.1438\sqrt{23} = 0.6896 \), which indicates (from Appendix J) a \( p \)-value of 0.73.

3.5.4 Example: Human Twin Data for Quantifying Genetic (vs. Environmental) Variance

There is variability in most human characteristics: people differ in height, weight, and other variables. One question of interest in fields such as health, clothing
manufacture, etc. (see Dudewicz, Chen, and Taneja (1989), pp. 223–228) is how much of the variability is due to genetic makeup, and how much is due to environmental influences. There is little dispute that some of the variability is from each source; there is a great controversy about how much of the variability should be attributed to each source. One approach to quantifying this variability is through so-called “twin studies,” where human twins are studied in various scenarios. In such studies, normality of the variables is often assumed when this assumption may not always be valid (see Williams and Zhou (1998)). Some references to this interesting field include Christian, Carmelli, Castelli, Fabsitz, Grim, Meaney, Norton, Reed, Williams, and Wood (1990), and Christian, Kang, and Norton (1974). Interesting datasets in this area come from the Indiana Twin Study. We focus on one dataset\(^1\) which is given in sorted form in Table 3.5–3. 

From FitGLDM, we determine the sample moments of \(X\) to be 
\[
\hat{\alpha}_1 = 5.48585, \quad \hat{\alpha}_2 = 1.3082, \quad \hat{\alpha}_3 = -0.04608, \quad \hat{\alpha}_4 = 2.7332
\]
and obtain the fit
\[
\text{GLD}(5.5872, 0.2266, 0.2089, 0.1762).
\]
To get a sense of the quality of this fit we look at the histogram of \(X\) with the superimposed graph of the fitted p.d.f. (Figure 3.5–4 (a)) as well as the e.d.f. of \(X\) with the superimposed d.f. of the GLD fit (Figure 3.5–4 (b)). Next, we apply the chi-square test with the 8 classes 
\[(-\infty, 4), \ [4, 4.5), \ [4.5, 5), \ [5, 5.5), \ [5.5, 6), \ [6, 6.5), \ [6.5, 7), \ [7, \infty)\]
for which the observed frequencies, respectively, are
\[12, \ 15, \ 12, \ 21, \ 25, \ 11, \ 18, \ 9.\]
The chi-square test yields expected frequencies of
\[12.6118, \ 12.0610, \ 16.8281, \ 20.0075, \ 20.2773, \ 17.3429, \ 12.2937, \ 11.5777\]
for these 8 classes, giving us 
\[8.8226 \quad \text{and} \quad 0.03174\]
for the chi-square statistic and \(p\)-value for this test. We also obtain \(KS = 0.03978\) and \(KS\sqrt{n} = 0.03978\sqrt{123} = 0.4412\) for a \(p\)-value of 0.99.

With a similar analysis of \(Y\), using the same classes, we find
\[
\hat{\alpha}_1 = 5.3666, \quad \hat{\alpha}_2 = 1.2033, \quad \hat{\alpha}_3 = -0.01219, \quad \hat{\alpha}_4 = 2.7665,
\]
\(^1\)This data comes from the Ph.D. thesis of Dr. Cynthia Moore, under the supervision of Dr. Joseph C. Christian, Department of Medical and Molecular Genetics, Indiana University School of Medicine, kindly shared with us by Dr. Moore. The data collection was supported by the National Institutes of Health Individual Research Fellowship Grant—“Twin Studies in Human Development.” PHS-5-F32-HD06869, 1987–1990.
### Table 3.5–3. Birth Weights of Twins.

<table>
<thead>
<tr>
<th>Twin 1</th>
<th>Twin 2</th>
<th>Twin 1</th>
<th>Twin 2</th>
<th>Twin 1</th>
<th>Twin 2</th>
<th>Twin 1</th>
<th>Twin 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.44</td>
<td>3.00</td>
<td>3.78</td>
<td>3.15</td>
<td>2.93</td>
<td>3.17</td>
<td>4.13</td>
<td></td>
</tr>
<tr>
<td>3.63</td>
<td>3.19</td>
<td>3.68</td>
<td>5.38</td>
<td>3.69</td>
<td>3.56</td>
<td>3.74</td>
<td>3.24</td>
</tr>
<tr>
<td>3.75</td>
<td>3.16</td>
<td>3.83</td>
<td>3.83</td>
<td>3.91</td>
<td>3.81</td>
<td>3.91</td>
<td>4.60</td>
</tr>
<tr>
<td>4.00</td>
<td>3.66</td>
<td>4.00</td>
<td>4.28</td>
<td>4.10</td>
<td>5.00</td>
<td>4.12</td>
<td>4.75</td>
</tr>
<tr>
<td>4.12</td>
<td>6.31</td>
<td>4.19</td>
<td>4.31</td>
<td>4.20</td>
<td>4.75</td>
<td>4.28</td>
<td>4.50</td>
</tr>
<tr>
<td>4.31</td>
<td>3.66</td>
<td>4.31</td>
<td>3.88</td>
<td>4.31</td>
<td>4.69</td>
<td>4.38</td>
<td>3.40</td>
</tr>
<tr>
<td>4.38</td>
<td>5.00</td>
<td>4.44</td>
<td>5.13</td>
<td>4.49</td>
<td>4.15</td>
<td>4.53</td>
<td>3.83</td>
</tr>
<tr>
<td>4.56</td>
<td>4.31</td>
<td>4.56</td>
<td>5.38</td>
<td>4.63</td>
<td>4.12</td>
<td>4.69</td>
<td>4.63</td>
</tr>
<tr>
<td>4.75</td>
<td>4.56</td>
<td>4.75</td>
<td>4.63</td>
<td>4.81</td>
<td>4.38</td>
<td>4.81</td>
<td>4.44</td>
</tr>
<tr>
<td>4.91</td>
<td>5.13</td>
<td>4.94</td>
<td>4.78</td>
<td>4.95</td>
<td>4.22</td>
<td>5.00</td>
<td>5.38</td>
</tr>
<tr>
<td>5.00</td>
<td>6.16</td>
<td>5.03</td>
<td>4.94</td>
<td>5.13</td>
<td>3.81</td>
<td>5.15</td>
<td>5.00</td>
</tr>
<tr>
<td>5.16</td>
<td>5.75</td>
<td>5.16</td>
<td>6.69</td>
<td>5.19</td>
<td>4.94</td>
<td>5.22</td>
<td>4.75</td>
</tr>
<tr>
<td>5.25</td>
<td>5.38</td>
<td>5.25</td>
<td>5.63</td>
<td>5.25</td>
<td>5.81</td>
<td>5.25</td>
<td>6.10</td>
</tr>
<tr>
<td>5.25</td>
<td>6.25</td>
<td>5.31</td>
<td>4.69</td>
<td>5.38</td>
<td>5.69</td>
<td>5.38</td>
<td>6.81</td>
</tr>
<tr>
<td>5.41</td>
<td>5.69</td>
<td>5.44</td>
<td>5.75</td>
<td>5.47</td>
<td>5.13</td>
<td>5.47</td>
<td>6.75</td>
</tr>
<tr>
<td>5.50</td>
<td>5.38</td>
<td>5.56</td>
<td>4.44</td>
<td>5.56</td>
<td>5.48</td>
<td>5.56</td>
<td>6.31</td>
</tr>
<tr>
<td>5.59</td>
<td>6.22</td>
<td>5.61</td>
<td>6.18</td>
<td>5.63</td>
<td>4.69</td>
<td>5.63</td>
<td>4.88</td>
</tr>
<tr>
<td>5.63</td>
<td>4.97</td>
<td>5.66</td>
<td>5.38</td>
<td>5.72</td>
<td>5.06</td>
<td>5.75</td>
<td>5.63</td>
</tr>
<tr>
<td>5.81</td>
<td>5.19</td>
<td>5.81</td>
<td>5.94</td>
<td>5.84</td>
<td>5.56</td>
<td>5.88</td>
<td>4.88</td>
</tr>
<tr>
<td>5.88</td>
<td>5.69</td>
<td>5.88</td>
<td>5.88</td>
<td>5.88</td>
<td>5.88</td>
<td>5.88</td>
<td>5.88</td>
</tr>
<tr>
<td>5.94</td>
<td>4.81</td>
<td>5.94</td>
<td>5.41</td>
<td>5.94</td>
<td>5.75</td>
<td>5.94</td>
<td>6.31</td>
</tr>
<tr>
<td>5.97</td>
<td>5.63</td>
<td>6.06</td>
<td>5.31</td>
<td>6.10</td>
<td>5.19</td>
<td>6.10</td>
<td>6.13</td>
</tr>
<tr>
<td>6.16</td>
<td>5.85</td>
<td>6.19</td>
<td>4.44</td>
<td>6.19</td>
<td>5.50</td>
<td>6.31</td>
<td>5.81</td>
</tr>
<tr>
<td>6.31</td>
<td>6.10</td>
<td>6.33</td>
<td>8.14</td>
<td>6.38</td>
<td>5.19</td>
<td>6.38</td>
<td>6.05</td>
</tr>
<tr>
<td>6.63</td>
<td>6.19</td>
<td>6.63</td>
<td>6.38</td>
<td>6.66</td>
<td>7.10</td>
<td>6.69</td>
<td>5.81</td>
</tr>
<tr>
<td>6.81</td>
<td>6.60</td>
<td>6.81</td>
<td>7.41</td>
<td>6.88</td>
<td>6.06</td>
<td>6.88</td>
<td>6.63</td>
</tr>
<tr>
<td>6.94</td>
<td>5.50</td>
<td>6.95</td>
<td>5.72</td>
<td>7.06</td>
<td>7.31</td>
<td>7.25</td>
<td>8.00</td>
</tr>
<tr>
<td>7.31</td>
<td>4.58</td>
<td>7.31</td>
<td>7.31</td>
<td>7.46</td>
<td>7.22</td>
<td>7.69</td>
<td>7.25</td>
</tr>
<tr>
<td>7.72</td>
<td>6.44</td>
<td>8.13</td>
<td>7.75</td>
<td>8.44</td>
<td>6.31</td>
<td>8.44</td>
<td>6.31</td>
</tr>
</tbody>
</table>
Chapter 3: Fitting Distributions and Data with the GLD via Moments

Figure 3.5–4. The histogram of $X$ and its fitted p.d.f. (a); the empirical d.f. of $X$ and the d.f. of its fitted GLD (b).

Figure 3.5–5. The histogram of $Y$ and its fitted p.d.f. (a); the empirical d.f. of $Y$ and the d.f. of its fitted GLD (b).

GLD(5.3904, 0.2293, 0.1884, 0.1807).

The comparative p.d.f. and d.f. plots for $Y$ are given in Figures 3.5–5 (a) and (b). When we apply the chi-square test we obtain observed frequencies of

15, 11, 19, 19, 21, 22, 7, 9,

expected frequencies of

13.5765, 13.4925, 18.6505, 21.4160, 20.4945, 16.2641, 10.5722, 8.5335,

and chi-square statistic and $p$-value of 4.1567 and 0.2450, respectively. In this case, $KS = 0.04824$ and $KS \sqrt{n} = 0.04824\sqrt{123} = 0.5350$ for a $p$-value of 0.94.
3.5.5 Example: Rainfall Distributions

Statistical modeling plays an essential role in the study of rainfall and the relationships between rainfall at multiple sites (Shimizu (1993)). Lognormal distributions have been used extensively in this work, and in univariate cases have worked well; however, with multiple sites the many rejections of lognormality (e.g., see the Rs in “Table 2. Test for lognormality” on p. 168 of Shimizu (1993)) indicate a need for more modeling flexibility. While the data for Shimizu’s studies was from sites in Japan and is no longer available, similar data for U.S. sites is readily available from the U.S. National Oceanic and Atmospheric Administration. From that data, shown in Table 3.5–6 is data from the period May 1998 to October 1988 in Rochester, New York and Syracuse, New York. This data is for the 47 days in which both cities had positive rainfall measured. (The study of Shimizu (1993) looks at all 4 combinations of “rain, no rain” possibilities; however, for this example’s purpose we concentrate on only the case of rain at both sites.)

Table 3.5–6. Rainfall (in inches) at Rochester, N.Y. ($X$) and Syracuse, N.Y., ($Y$), from May to October of 1998, on days when both sites had positive rainfall.

<table>
<thead>
<tr>
<th>$X$</th>
<th>$Y$</th>
<th>$X$</th>
<th>$Y$</th>
<th>$X$</th>
<th>$Y$</th>
</tr>
</thead>
<tbody>
<tr>
<td>.03</td>
<td>.05</td>
<td>.11</td>
<td>.07</td>
<td>.07</td>
<td>.04</td>
</tr>
<tr>
<td>.09</td>
<td>.14</td>
<td>.01</td>
<td>.14</td>
<td>.69</td>
<td>.72</td>
</tr>
<tr>
<td>.62</td>
<td>.60</td>
<td>.08</td>
<td>.21</td>
<td>.61</td>
<td>.36</td>
</tr>
<tr>
<td>.08</td>
<td>.09</td>
<td>.50</td>
<td>.54</td>
<td>.23</td>
<td>.18</td>
</tr>
<tr>
<td>1.27</td>
<td>.22</td>
<td>.03</td>
<td>1.26</td>
<td>.09</td>
<td>.10</td>
</tr>
<tr>
<td>1.20</td>
<td>.21</td>
<td>1.65</td>
<td>.94</td>
<td>1.87</td>
<td>.17</td>
</tr>
<tr>
<td>.15</td>
<td>.05</td>
<td>2.61</td>
<td>.94</td>
<td>.85</td>
<td>.37</td>
</tr>
<tr>
<td>1.35</td>
<td>.48</td>
<td>.15</td>
<td>.28</td>
<td>.31</td>
<td>.38</td>
</tr>
<tr>
<td>.16</td>
<td>.48</td>
<td>.27</td>
<td>.31</td>
<td>.06</td>
<td>.04</td>
</tr>
<tr>
<td>.23</td>
<td>1.09</td>
<td>1.38</td>
<td>.42</td>
<td>.74</td>
<td>1.79</td>
</tr>
<tr>
<td>2.51</td>
<td>1.25</td>
<td>.13</td>
<td>.11</td>
<td>.10</td>
<td>.16</td>
</tr>
<tr>
<td>.24</td>
<td>1.05</td>
<td>.76</td>
<td>.36</td>
<td>.03</td>
<td>.42</td>
</tr>
<tr>
<td>.29</td>
<td>.08</td>
<td>.02</td>
<td>.17</td>
<td>.07</td>
<td>.02</td>
</tr>
<tr>
<td>.17</td>
<td>.19</td>
<td>.47</td>
<td>.04</td>
<td>.41</td>
<td>.79</td>
</tr>
<tr>
<td>.03</td>
<td>.37</td>
<td>.15</td>
<td>.18</td>
<td>.01</td>
<td>.09</td>
</tr>
<tr>
<td>.02</td>
<td>.19</td>
<td>.19</td>
<td>.22</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The moments of $X$ are

$$\hat{\alpha}_1 = 0.4913, \quad \hat{\alpha}_2 = 0.4074, \quad \hat{\alpha}_3 = 1.8321, \quad \hat{\alpha}_4 = 5.7347,$$
and those of \( Y \) are
\[
\hat{\alpha}_1 = .3906, \quad \hat{\alpha}_2 = .1533, \quad \hat{\alpha}_3 = 1.6163, \quad \hat{\alpha}_4 = 5.2245.
\]

In both cases, \((\hat{\alpha}_3^2, \hat{\alpha}_4)\) is located outside of the region of the table of Appendix B and outside of the region for reliable computations. We will return to this example in Chapters 4 and 5 and fit \( \text{GLD}(\lambda_1, \lambda_2, \lambda_3, \lambda_4) \) distributions to both \( X \) and \( Y \) via the EGLD and through a percentile-based method.

### 3.6 Moment-Based GLD Fit to Data from a Histogram

In the examples of Section 3.5 the actual data on the phenomenon of interest, \( X_1, X_2, \ldots, X_n \), was available to us. We acted assuming that these were independent and identically distributed observations — an assumption that should be verified. In many cases the data are given in the form of a histogram. While it would in general be preferable to have the actual data (going to a histogram involves a certain loss of information, such as the distribution of the observations within the classes), it nevertheless is possible to proceed with fitting of a distribution using data in the form of a histogram.

One could seek the original data from the authors of the histogram, but this is often difficult: the author may have moved, or may not be able to release the data without a time-consuming approval process, or may not have retained the basic data. As an example of use of histograms, one may look in virtually any scientific journal or newspaper. For example, Dahl-Jensen, Mosegaard, Gundestrup, Clow, Johnsen, Hansen, and Balling (1998) in their Figure 2 give six histograms, each based on 2000 Monte Carlo observations. As they use between 5 and 15 classes in each histogram, they are easily able to meet the rules for an informative histogram: there should be no more than two classes with a frequency less than 5, and classes should be of equal width. While histograms with classes of unequal width can be used, they are often misused. In order not to be misleading, the heights of the bars must be adjusted so the area of the bar is proportional to the frequency. Thus, if class 1 has a width of 2.5 and 150 observations, and its height is 150 in the frequency histogram, and if class 2 has width 5 and also 150 observations, then its height must be 75. Below we give two examples of the use of a histogram to fit a distribution.

Sometimes the histogram data is presented to us in the form of a table such as Table 3.6–1, which comes from Ramberg, Tadikamalla, Dudewicz, and Mykytka (1979, p. 207). The classes have equal widths (assuming the lowest class is 0.010 to 0.015 and the highest class is 0.060 to 0.065), but there are more than 2 classes with fewer than 5 observations in them, so we combine the two lowest and the two highest classes. Now there are no classes with fewer than 5 observations in them. Classes with low frequencies have highly variable estimates.
3.6 Moment-Based GLD Fit to Data from a Histogram

### Table 3.6–1. Observed Coefficient of Friction Frequencies.

<table>
<thead>
<tr>
<th>Coefficient of Friction</th>
<th>Observed Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>less than 0.015</td>
<td>1</td>
</tr>
<tr>
<td>0.015 – 0.020</td>
<td>9</td>
</tr>
<tr>
<td>0.020 – 0.025</td>
<td>30</td>
</tr>
<tr>
<td>0.025 – 0.030</td>
<td>44</td>
</tr>
<tr>
<td>0.030 – 0.035</td>
<td>58</td>
</tr>
<tr>
<td>0.035 – 0.040</td>
<td>45</td>
</tr>
<tr>
<td>0.040 – 0.045</td>
<td>29</td>
</tr>
<tr>
<td>0.045 – 0.050</td>
<td>17</td>
</tr>
<tr>
<td>0.050 – 0.055</td>
<td>9</td>
</tr>
<tr>
<td>0.055 – 0.060</td>
<td>4</td>
</tr>
<tr>
<td>0.060 or more</td>
<td>4</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>250</td>
</tr>
</tbody>
</table>

of true p.d.f. heights, due to inadequate amount of observation, and it is preferable to avoid them in order to avoid a bias in the fitting.

We use the mid-point assumption that every data point is located at the center of its class interval; for the first class, we assume that its entries are located at 0.0125 and for the last class we assume that the entries have the value 0.0625. With these assumptions, we can compute approximate moments for this data and obtain (from \( \text{FitGLD}\))

\[
\hat{\alpha}_1 = 0.03448, \quad \hat{\alpha}_2 = 9.2380 \times 10^{-5}, \quad \hat{\alpha}_3 = 0.5374, \quad \hat{\alpha}_4 = 3.2130,
\]

and the GLD fit

\[
\text{GLD}_1(0.02889, 18.1935, 0.05744, 0.1850).
\]

We note that the original data are in fact given in Hahn and Shapiro (1967), and from the original data (not a histogram) Hahn and Shapiro computed the sample moments as

\[
\hat{\alpha}_1 = .0345, \quad \hat{\alpha}_2 = 0.00009604, \quad \hat{\alpha}_3 = .87, \quad \hat{\alpha}_4 = 4.92. \quad (3.6.1)
\]

We now also fit a GLD using (3.6.1) to illustrate the change that results from having the original data vs. only the histogram of the data. The GLD\(_2\) associated with the moments given in (3.6.1) is given by

\[
\text{GLD}_2(0.03031, 1.5771, 0.005174, 0.01190).
\]

The histogram of the data with the fitted GLD\(_1\) and GLD\(_2\) p.d.f.s is given in Figure 3.6–2 (a) (the GLD\(_2\) p.d.f. rises higher at the center). The e.d.f of the
data together with the two fitted d.f.s is given in Figure 3.6–2 (b) (the GLD\textsubscript{2} d.f. extends farther to the left).

We do a chi-square test using the classes
\[ (-\infty, 0.025), \quad [0.025, 0.030), \quad [0.030, 0.035), \quad [0.035, 0.040), \]
\[ [0.040, 0.045), \quad [0.045, 0.050), \quad [0.050, \infty) \]
for which the observed frequencies are, respectively,
\[ 40, \quad 44, \quad 58, \quad 45, \quad 29, \quad 17, \quad 17. \]

We get the following expected frequencies:
\[
\text{GLD}_1 : \quad 38.8184, \quad 50.0488, \quad 52.7344, \quad 42.9687, \quad 29.6631, \quad 18.2495, \quad 17.5171 \\
\text{GLD}_2 : \quad 347900, \quad 51.3916, \quad 58.3496, \quad 44.7998, \quad 27.8931, \quad 15.6860, \quad 17.0898.
\]

These give the following chi-square statistics and \( p \)-values for the two fits:
\[
\text{GLD}_1 : \quad \chi^2 \text{ statistic } = 1.5045, \quad p\text{-value} = 0.8048 \\
\text{GLD}_2 : \quad \chi^2 \text{ statistic } = 2.0008, \quad p\text{-value} = 0.3677.
\]

For the Kolmogorov-Smirnov tests for these two fits, we have
\[
KS_1 = 0.1279 : \quad KS\sqrt{n} = 0.1279\sqrt{250} = 2.0223, \quad p\text{-value} \approx 0 \\
KS_2 = 0.1288 : \quad KS\sqrt{n} = 0.1288\sqrt{250} = 2.0365, \quad p\text{-value} \approx 0.
\]

We conclude that at least in this example (with a relatively large sample size, and a well-constructed histogram), the effect of the lack of the original data has been negligible: both GLD\textsubscript{1} and GLD\textsubscript{2} models fit the data quite well (the
very poor results from the Kolmogorov-Smirnov test are a consequence of our assumption that all data points are located at the mid-points of their intervals).

As an alternative to the mid-point assumption we consider now the, equally reasonable, assumption that the data in a given interval is spread uniformly throughout that interval. This produces

\[ \hat{\alpha}_1 = 0.03448, \quad \hat{\alpha}_2 = 9.43 \times 10^{-5}, \quad \hat{\alpha}_3 = 0.5289, \quad \hat{\alpha}_4 = 3.1886, \]

and the GLD fit

\[ \text{GLD}_3(0.02895, 18.2454, 0.05964, 0.1865), \]

with support \([-0.0259, 0.0838]\).

A chi-square test, using the same intervals as before, gives expected frequencies of

\[ 39.7949, 49.3164, 52.2461, 42.6025, 29.7241, 18.4326, 17.8833, \]

and the chi-square statistic and its associated \( p \)-value

\[ 1.5154 \quad \text{and} \quad 0.4687, \]

respectively. For this fit we also have \( KS = 0.02386 \) and \( KS\sqrt{n} = 0.02386\sqrt{250} = 0.3773 \) that gives a \( p \)-value of 0.999. Figure 3.6–3 (a) shows the GLD\_3 p.d.f. with the data histogram and Figure 3.6–3 (b) gives the e.d.f. of the data with the d.f. of the GLD\_3. We can see that GLD\_3 is also a good fit, with its chi-square \( p \)-value about the same as that for GLD\_1 and a much better \( KS \) \( p \)-value.
Chapter 3: Fitting Distributions and Data with the GLD via Moments

3.7 The GLD and Design of Experiments

In many studies including, but not limited to, simulation studies and theoretical studies of sensitivity of results to distributions, there is a distribution one thinks of as the “best fit” to the true underlying distribution. For example, in the coefficient of friction data of the example in Section 3.6, the underlying “true” state of nature is estimated to have moments

\[(\alpha_1, \alpha_2, \alpha_3, \alpha_4) = (0.03448, 9.238 \times 10^{-5}, 0.5374, 3.2130)\]  
(3.7.1)

and the GLD p.f.

\[Q(y) = 0.02889 + (y^{0.05744} - (1 - y)^{0.1850})/18.1935.\]  
(3.7.2)

However, as these are only estimates from the data, very often it will be desired to re-do the analysis (e.g., re-run the simulation program, or refine the theoretical analysis) with a changed distribution (a distribution in the neighborhood of the best estimate, but varied as much as seems reasonable in the setting using the best available engineering information).

This type of analysis is relatively simple to perform using the GLD and notions of statistical design of experiments (e.g., see Section 6.4.2 of Karian and Dudewicz (1999a)). This approach is superior to the “grab bag of distributions” approach sometimes used, wherein one haphazardly adds alternative distributions to the set used. To vary the distribution from the fitted GLD in (3.7.2), we vary the parameters \((\lambda_1, \lambda_2, \lambda_3, \lambda_4)\), but these do not have intrinsic value to us (as we may not easily understand the meaning of a change of 0.5 units in \(\lambda_4\), for example). So, instead we vary the \((\alpha_1, \alpha_2, \alpha_3, \alpha_4)\), fit a new GLD for each desired set of \(\alpha\)-values, and run our study with the new fitted GLD.

In the coefficient of friction example, suppose we are comfortable with the fitted mean \((0.03448)\) and variance \((9.2380 \times 10^{-5})\) and do not believe that perturbations of reasonable size in these values will impact the variables of interest to us. However, we believe that the shape parameters (skewness \(\alpha_3\) and kurtosis \(\alpha_4\)) may have an impact when varied within reasonable bounds. Further, suppose that we have reason to believe that \(\alpha_3\) may vary from 0.5374 by up to \(\pm 0.2\), and \(\alpha_4\) may vary from 3.2130 by up to \(\pm 0.5\). That reason can either be based on previous studies in the area, or on theoretical considerations, or on statistical reasons for which we might take these bounds to be \(\pm 2\) standard deviations (estimated) of the estimates in (3.7.1). Note here that one needs to estimate (and take twice the square root of) the variances

\[\text{Var}(\hat{\alpha}_1) = \frac{\alpha_2}{n},\]  
(3.7.3)

\[\text{Var}(\hat{\alpha}_2) = \frac{(n - 1)^2}{n^3} \mu_4 - \frac{(n - 1)(n - 3)}{n^3} \alpha_2 \approx \frac{\mu_4 - \alpha_2^2}{n},\]  
(3.7.4)
3.7 The GLD and Design of Experiments

\[
\text{Var}(\hat{\alpha}_3) \approx (\alpha_2^2\mu_6 - 3\alpha_2\mu_3\mu_5 - 6\alpha_3^3\mu_4 \\
+ 2.25\alpha_3^2\mu_4 + 8.75\alpha_2^2\mu_3^2 + 9\alpha_2^5)/(n\alpha_2^5),
\]

(3.7.5)

\[
\text{Var}(\hat{\alpha}_4) \approx (\alpha_2^2\mu_8 - 4\alpha_2\mu_4\mu_6 - 8\alpha_3^2\mu_3\mu_5 + 4\mu_4^2 \\
- \alpha_2^2\mu_4^2 + 16\alpha_2\mu_3^2\mu_4 + 16\alpha_3^3\mu_3^2)/(n\alpha_2^6).
\]

(3.7.6)

In the above we already know how to estimate \( \alpha_2 \) by \( \hat{\alpha}_2 \). The new expressions \( \mu_3, \mu_4, \mu_5, \mu_6, \) and \( \mu_8 \) are related to the \( \alpha_1, \alpha_2, \alpha_3, \alpha_4 \) that we are familiar with by the equations

\[
\mu_3 = \alpha_2^1.5\alpha_3, \quad \mu_4 = \alpha_2^2\alpha_4, \quad \mu_5 = \alpha_2^2.5\alpha_5, \quad \mu_6 = \alpha_2^3\alpha_6, \quad \mu_8 = \alpha_2^4\alpha_8.
\]

(3.7.7)

Hence, they can be estimated by replacing right-hand side terms by their estimates. The approximations in (3.7.5) and (3.7.6) are up to order \( O(n^{-3/2}) \). A good reference is Cramér (1946, pp. 354, 357). In this example, we get

\[
\begin{align*}
\alpha_5 &= 4.5553, \\
\alpha_6 &= 17.0073, \\
\alpha_8 &= 114.2131, \\
\mu_3 &= 4.7716 \times 10^{-7}, \\
\mu_4 &= 2.7142 \times 10^{-8}, \\
\mu_5 &= 3.7365 \times 10^{-10}, \\
\mu_6 &= 1.3408 \times 10^{-11}, \\
\mu_8 &= 8.3182 \times 10^{-15}, \\
\text{Var}(\alpha_3) &= 0.01600, \\
\text{Var}(\alpha_4) &= 0.07150, \\
\sqrt{\text{Var}(\alpha_3)} &= 0.1265, \\
\sqrt{\text{Var}(\alpha_4)} &= 0.2674.
\end{align*}
\]

Hence, \( 2\sqrt{\text{Var}(\alpha_3)} = 0.253 \) and \( 2\sqrt{\text{Var}(\alpha_4)} = 0.536 \). A study that varies \( \alpha_3 \) by up to \( \pm 0.25 \), and \( \alpha_4 \) by up to \( \pm 0.54 \), is reasonable. We use \( \pm 0.2 \) and \( \pm 0.5 \) for simplicity below.

We now choose the \((\alpha_3, \alpha_4)\) for our experiments using what is called a **Central Composite Design in two variables** \((\alpha_3 \text{ and } \alpha_4)\), with a multiplier of 1.5. The sample point \((0.5374, 3.2130)\) is at the center of our experiments, and is called a **Center Point**. We go out from this Center Point to a distance \( \pm 1.5d_1 = \pm 0.2 \) on \( \alpha_3 \), and \( \pm 1.5d_2 = \pm 0.5 \) on \( \alpha_4 \), so we have \( d_1 = .1333, \ d_2 = .3333 \). This sets the **Star Points** as

\[
(0.5374 - 0.2, 3.2130) = (0.3374, 3.2130)
\]

© 2011 by Taylor and Francis Group, LLC
Figure 3.7–1. The design points for the sensitivity study of the coefficient of friction example.

\[
(0.5374 + 0.02, 3.2130) = (0.7374, 3.2130) \\
(0.5374, 3.2130 - 0.5) = (0.5374, 2.7130) \quad (3.7.8) \\
(0.5374, 3.2130 + 0.5) = (0.5374, 3.7130).
\]

The Factorial Points are taken as

\[
(0.5374 - d_1, 3.2130 - d_2) = (0.4041, 2.8797) \\
(0.5374 - d_1, 3.2130 + d_2) = (0.4041, 3.5463) \\
(0.5374 + d_1, 3.2130 - d_2) = (0.6707, 2.8797) \quad (3.7.9) \\
(0.5374 + d_1, 3.2130 + d_2) = (0.6707, 3.5463).
\]

This is illustrated on coordinate axes in Figure 3.7–1. Also see Figure 6.4–2 of Karian and Dudewicz (1999b) for the general case and see Dudewicz and Karian (1985, pp. 189, 196, 206, 233) for details for more than two variables; that case would arise if we varied all of \(\alpha_1, \alpha_2, \alpha_3, \alpha_4\), and there are ways to cut the number of points used from the full \(2^k + 2k + 1\) with \(k\) variables without losing much information in a precise sense. This is desirable since when \(k = 2\) we have only 9 points, but this would, if not reduced, become 25 with \(k = 4\) and grow rapidly with \(k\).
In Figure 3.7–1 note that the factorial points lie on the corners of a box (shown by “x”), and the star points (shown by “*”) lie on perpendiculars from the center point and go out a distance ±0.2 (horizontally), ±0.5 (vertically) from the center of the design area. One should also be careful that the specified variations continue to fall in the possible region of Figure 3.2–4; if they move outside the GLD area, one may need to utilize the EGLD of Chapter 4.

We now fit a GLD for each of the 9 points (we already have this for the center point, so there are 8 additional points to be fitted). This process yields the following \((\alpha_3, \alpha_4)\) and their associated GLDs (the GLDs associated with star points are designated with the superscript “*,” those associated with factorial points are similarly marked with an “x,” and the one associated with the center point is designated by “o”).

\[
\begin{align*}
(\alpha_3, \alpha_4) &= (0.5374, 3.2130), & \text{GLD}^c(0.02902, 18.1399, 0.05850, 0.1825); \\
(\alpha_3, \alpha_4) &= (0.3374, 3.2130), & \text{GLD}^*(0.03136, 17.0249, 0.07594, 0.1412); \\
(\alpha_3, \alpha_4) &= (0.7374, 3.2130), & \text{GLD}^*(0.02595, 19.0831, 0.03251, 0.2410); \\
(\alpha_3, \alpha_4) &= (0.5374, 2.7130), & \text{GLD}^*(0.02547, 23.6743, 0.04238, 0.3404); \\
(\alpha_3, \alpha_4) &= (0.5374, 3.7130), & \text{GLD}^*(0.03080, 10.8659, 0.04175, 0.08697); \\
(\alpha_3, \alpha_4) &= (0.4041, 2.8797), & \text{GLD}^x(0.02895, 22.5726, 0.07766, 0.2452); \\
(\alpha_3, \alpha_4) &= (0.4041, 3.5463), & \text{GLD}^x(0.03160, 11.9256, 0.05164, 0.09109); \\
(\alpha_3, \alpha_4) &= (0.6707, 2.8797), & \text{GLD}^x(0.02450, 21.8367, 0.02642, 0.3223); \\
(\alpha_3, \alpha_4) &= (0.6707, 3.5463), & \text{GLD}^x(0.02881, 14.7763, 0.04340, 0.1434).
\end{align*}
\]

The graphs of these p.d.f.s are shown in Figure 3.7–2, where the fit associated with the center point is shown by diamond-shaped points. From Figure 3.7–2 we can observe the spread of the distributions chosen for the sensitivity study about the distribution associated with the center point.

As a final point, note that if we are interested in an output of an experiment (simulation or other experiment), we can now fit a metamodel via regression. That will show the effect of changes in \(\alpha_3\) and \(\alpha_4\) from the center on the output, and is a useful way of summarizing the information in the experiments run at the design points. For a detailed example, including SAS code and contour graphs for interpretation, see Section 8.3 of Karian and Dudewicz (1999a).
Problems for Chapter 3

3.1. For the GLD studied in (2.1.6), find the value of \( \alpha = (\alpha_1, \alpha_2, \alpha_3, \alpha_4) \), and identify \((\alpha_3^2, \alpha_4)\) in Figure 3.2–4.

3.2. For the two GLDs of Figure 2.3–1, find the first four moments, and identify the distributions in the plot of Figure 3.2–4.

3.3. For the 9 GLDs plotted in Figure 2.3–3a, find the first four moments, and identify their points in Figure 3.2–4.

3.4. For the 9 GLDs plotted in Figure 2.3–4a, find the first four moments, and identify their points in Figure 3.2–4.

3.5. For the 7 GLDs plotted in Figure 2.3–5, find the first four moments, and identify their points in Figure 3.2–4.

3.6. For the 4 GLDs plotted in Figure 2.3–6, find the first four moments, and identify their points in Figure 3.2–4.

3.7. For the 9 GLDs plotted in Figure 2.3–7a, find the first four moments, and identify their points in Figure 3.2–4.

3.8. For the 4 GLDs plotted in Figure 2.3–8a, find the first four moments, and identify their points in Figure 3.2–4.
3.9. For the 3 GLDs plotted in Figure 2.3–8b, find the first four moments, and identify their points in Figure 3.2–4.

3.10. For the 4 GLDs plotted in Figure 2.3–9, find the first four moments, and identify their points in Figure 3.2–4.

3.11. In Chapter 2 (e.g., see Figures 2.3–6, 2.3–7a, 2.3–8a) plots were given of GLDs with parameters in Regions 1 and 5 of Figure 2.2–1, but no plots for Regions 2 and 6. Construct similar plots for Regions 2 and 6, and find their \((\alpha_3^2, \alpha_4)\) points (if they exist). (Hint: See Theorem 2.3.22.)

3.12. Ramberg (1975) developed the model

\[ Q(y) = \lambda_1 + (\lambda_4y^{\lambda_3} - (1 - y)^{\lambda_3})/\lambda_2, \]

and gave a set of tables. Fit some examples with the GLD and with the above model and compare. Which do you prefer, and why?

3.13. In Figure 3.5–4 (b) we plotted the e.d.f. and the fitted GLD d.f. on the same axes. Plot a 95% confidence band on the same axes to quantify the quality of the fit. (The confidence band uses the e.d.f. data to yield a band in which the true d.f. will, with high confidence, lie. See Gan, Koehler, and Thompson (1991).)

References for Chapter 3


Chapter 3: Fitting Distributions and Data with the GLD via Moments


References for Chapter 3


Department of Civil Engineering, The Ohio State University, Columbus, Ohio, xi+183 pp.


