Example 2.3 Find the method of moments estimate of the unknown parameter θ for a random sample $X_1, X_2, \ldots, X_n$ drawn from a population with probability density function

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

where θ is a positive unknown parameter.

The first step is to recognize that the distribution being fitted is an exponential distribution with population mean θ. We again assume that the preliminary work of collecting the data and drawing a histogram has already been completed. The conclusion is that the histogram approximates the shape of the probability density function of an exponential random variable to a reasonable degree, so we can proceed with parameter estimation by the method of moments. There is only a single ($k = 1$) unknown parameter θ to estimate in this case. The population mean of the exponential random variable is $\mu = \theta$. So, equating the population mean to the sample mean results in

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

This equation is already solved for θ, so the final remaining step is to simply place a hat on θ:

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

For this population probability distribution, the population mean is estimated by the sample mean, which is intuitively appealing.

Now that the method of moments estimator has been determined, consider fitting an actual data set to the exponential distribution. To illustrate the point estimation process, fit the $n = 23$ ball bearing failure times (measured in 10^6 revolutions)

17.88	28.92	33.00	41.52	42.12	45.60	48.48	51.84
51.96	54.12	55.56	67.80	68.64	68.64	68.88	84.12
93.12	98.64	105.12	105.84	127.92	128.04	173.40	

to the exponential distribution using the method of moments technique. Assume that the analyst is lazy and does not first plot a histogram of the data to see if the data is likely to be drawn from an exponential population. Rather, the fitting of the exponential distribution is performed before the histogram is drawn. The method of moments estimate for θ is

$$\hat{\theta} = \frac{1}{n} \sum_{i=1}^{n} x_i = \frac{1}{23} \cdot 1661.16 \cong 72.22$$

million revolutions. One way to visually assess how well the fitted distribution approximates the data is to plot the empirical cumulative distribution function along with the fitted cumulative distribution function on its support:

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

These functions are plotted in Figure 2.1. In this case the fitted exponential distribution does *not* do an adequate job of approximating the data. The cumulative distribution

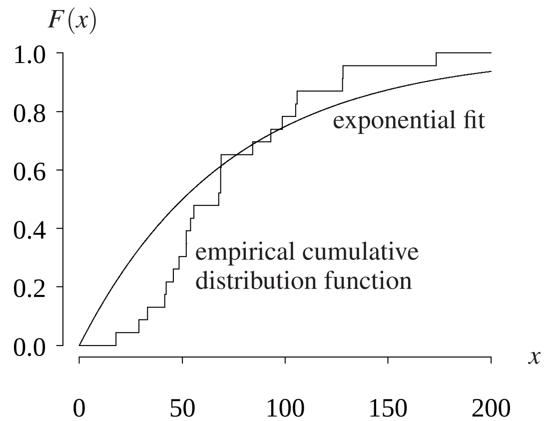

Figure 2.1: Empirical and fitted cumulative distribution function for the ball bearing data.

function for the exponential distribution is simply not capable of bending in a fashion to approximate the empirical cumulative distribution function. The fact that the exponential distribution does not make a good model for the ball bearing lifetimes could also have been recognized by plotting a histogram of the data, as shown in Figure 2.2. The histogram shape is clearly not consistent with an exponential population, as evidenced by the fitted probability density function. Further evidence that an exponential distribution will not be a good model for ball bearing failure times can be made using the memoryless property. Recall that the exponential distribution is the only continuous distribution having the memoryless property. If ball bearing failure times have the memoryless property, then used ball bearings are as good as new ball bearings. But this should not logically be the case, because the ball bearings are wearing out. The lesson to be learned from this analysis is that the exponential distribution can be easily fit to any data set of positive observations, so care must be taken to ensure that it is being applied appropriately. The next example illustrates that a two-parameter distribution,

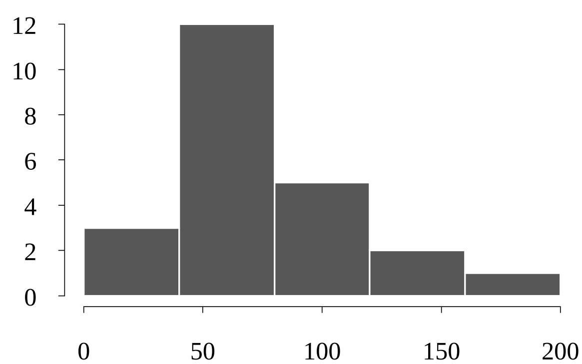

Figure 2.2: Histogram of the ball bearing failure times.

in particular the gamma distribution, is capable of more adequately modeling the ball bearing failure times.

The previous two examples involved fitting probability distributions with just a single ($k = 1$) parameter. The next example involves a parametric distribution with $k = 2$ parameters.

Example 2.4 Find the method of moments estimates of the unknown positive parameters λ and κ for a random sample $X_1, X_2, \ldots, X_n$ drawn from a gamma(λ, κ) population. Fit the gamma distribution to the ball bearing failure times.

The histogram in Figure 2.2 provides evidence that the gamma distribution, among others, is a potential model for the ball bearing failure times. There are $k = 2$ unknown parameters to be estimated, namely λ and κ, so a 2×2 set of equations must be solved. The generic equations for the method of moments procedure are

$$E[X] = \frac{1}{n}\sum_{i=1}^{n} x_i$$

$$E\left[X^2\right] = \frac{1}{n}\sum_{i=1}^{n} x_i^2.$$

For a gamma(λ, κ) population, these equations become

$$\frac{\kappa}{\lambda} = m_1 = \frac{1}{n}\sum_{i=1}^{n} x_i$$

$$\frac{\kappa(\kappa+1)}{\lambda^2} = m_2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2.$$

The value of $E\left[X^2\right]$ for the gamma distribution is found by the shortcut formula for the population variance: $V[X] = E\left[X^2\right] - E[X]^2$. Fortunately, this 2×2 set of equations has a closed-form solution that is easily calculated for a data set. Squaring both sides of the first equation and then dividing the second equation by the first equation gives

$$\frac{\kappa}{\kappa+1} = \frac{m_1^2}{m_2}$$

or

$$m_2\kappa = m_1^2\kappa + m_1^2.$$

Solving for κ in this equation, then plugging this estimator into the first original equation, the method of moments estimators are

$$\hat{\lambda} = \frac{m_1}{m_2 - m_1^2} \qquad \text{and} \qquad \hat{\kappa} = \frac{m_1^2}{m_2 - m_1^2}.$$

These method of moments estimators are random variables in the sense that they will differ from one data set to the next.

Returning to the ball bearing failure times for which the exponential distribution provided a dismal fit in the previous example, the following R statements can be used to calculate the method of moments estimates of λ and κ.

```
x  = c(17.88, 28.92, 33.00, 41.52, 42.12, 45.60, 48.48, 51.84, 51.96,
       54.12, 55.56, 67.80, 68.64, 68.64, 68.88, 84.12, 93.12, 98.64,
       105.12, 105.84, 127.92, 128.04, 173.40)
n  = length(x)
m1 = mean(x)
m2 = sum(x ^ 2) / n
lambda.hat = m1 / (m2 - m1 ^ 2)
kappa.hat  = m1 ^ 2 / (m2 - m1 ^ 2)
```

The method of moments estimators for this particular data set are

$$\hat{\lambda} = 0.05373 \qquad \text{and} \qquad \hat{\kappa} = 3.8804.$$

The empirical cumulative distribution function and the fitted cumulative distribution function are plotted in Figure 2.3. These reveal that the gamma distribution fit is far superior to the exponential distribution fit to the ball bearing failure times. Even though the fit to the ball bearing failure times is vastly improved by the gamma distribution over the exponential distribution, there may be even better parametric models. The right-hand tails of all gamma random variables are asymptotically exponentially distributed, which may not be consistent with the physics of failure associated with ball bearings. The ball bearings are likely to continue to wear away as they age. A continuous distribution with a lighter right-hand tail, such as the Weibull distribution with a shape parameter greater than one, might be a more appropriate model.

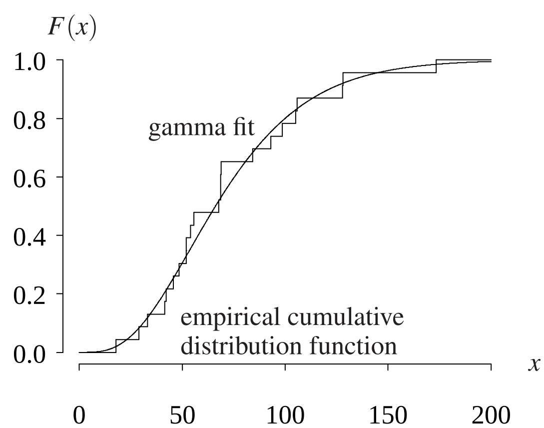

Figure 2.3: Empirical and fitted cumulative distribution function for the ball bearing data.

The next example of the method of moments estimation technique for parameter estimation fits a normal distribution to a data set.

Example 2.5 Find the method of moments estimates of the unknown parameters μ and σ^2 for a random sample $X_1, X_2, \ldots, X_n$ drawn from a $N\left(\mu, \sigma^2\right)$ population. Fit the normal distribution to Michelson's speed of light data from Example 1.7.