Beta Distributions

INTRODUCTION

14.1 Definition

A continuous random variable X has a *Beta distribution* if its pdf has the form

$$f(x; \lambda_1, \lambda_2) = \frac{\Gamma(\lambda_1 + \lambda_2)}{\Gamma(\lambda_1)\Gamma(\lambda_2)} x^{\lambda_1 - 1}(1 - x)^{\lambda_2 - 1}; \quad 0 < x < 1; \quad \lambda_1, \lambda_2 > 0; \tag{14.1}$$

where Γ represents the *gamma function* (see Section 13.1). This model does not have location-scale structure (see Section 9.2). Hence, both λ_1 and λ_2 are *shape* parameters, which are symmetrically related by:

$$f(x; \lambda_1, \lambda_2) = f(1 - x; \lambda_2, \lambda_1). \tag{14.2}$$

This distribution arose as the theoretical model of various statistics and statistical functions. It is now an important statistical model of random variables whose values are restricted to the unit range.

PROPERTIES: TWO-PARAMETER MODEL

14.2 Beta Variable

The cdf of a Beta variable, as defined by (14.1), cannot be expressed in closed form:

$$F(x; \lambda_1, \lambda_2) = \frac{\Gamma(\lambda_1 + \lambda_2)}{\Gamma(\lambda_1)\Gamma(\lambda_2)} \int_0^x z^{\lambda_1 - 1}(1 - z)^{\lambda_2 - 1}\, dz. \tag{14.3}$$

The inverse of the ratio of gamma functions in the above expression is called the *beta function*:

$$B(\lambda_1, \lambda_2) = \frac{\Gamma(\lambda_1)\Gamma(\lambda_2)}{\Gamma(\lambda_1 + \lambda_2)}. \tag{14.4}$$

The integral in (14.3) is termed the *incomplete beta function*:

$$B_x(\lambda_1, \lambda_2) = \int_0^x z^{\lambda_1 - 1}(1 - z)^{\lambda_2 - 1}\, dz. \tag{14.5}$$

Note that $B_x \to B$ as $x \to 1$. The cdf can thus be expressed as the *incomplete beta function ratio*

$$F(x; \lambda_1, \lambda_2) = \frac{B_x(\lambda_1, \lambda_2)}{B(\lambda_1, \lambda_2)}. \tag{14.6}$$

From (14.2) it follows that

$$F(x; \lambda_1, \lambda_2) = 1 - F(1 - x; \lambda_2, \lambda_1).$$

Tables[1] are available for beta functions. However, these functions are readily computed numerically. Note that Mathcad 6+ features a built-in function that returns the Beta cdf for positive shape parameters. When λ_1 and λ_2 are both integer valued, the Beta cdf can be evaluated as a Binomial sum by using the identity

$$F(x; \lambda_1, \lambda_2) = 1 - \sum_{i=0}^{\lambda_1 - 1} \binom{\lambda_1 + \lambda_2 - 1}{i} x^i (1 - x)^{\lambda_1 + \lambda_2 - 1 - i}. \tag{14.7}$$

14.3 Properties

The rth moment of X about the origin is given by

$$\mu'_r(X) = \frac{\Gamma(\lambda_1 + \lambda_2)\Gamma(\lambda_1 + r)}{\Gamma(\lambda_1)\Gamma(\lambda_1 + \lambda_2 + r)}. \tag{14.8}$$

Thus, the expected value of X is

$$\mu'_1(X) = \frac{\lambda_1}{\lambda_1 + \lambda_2}. \tag{14.9}$$

The variance of X is

$$\mu_2(X) = \frac{\lambda_1 \lambda_2}{(\lambda_1 + \lambda_2)^2(\lambda_1 + \lambda_2 + 1)}. \tag{14.10}$$

The coefficient of variation of X is thus

$$cv(X) = \sqrt{\frac{\lambda_2}{\lambda_1(\lambda_1 + \lambda_2 + 1)}}. \tag{14.11}$$

For $\lambda_1 + \lambda_2 > 2$ and $\lambda_1 \geq 1$, the mode value is

$$x_m = \frac{\lambda_1 - 1}{\lambda_1 + \lambda_2 - 2}. \tag{14.12}$$

[1] Pearson, E. S., Johnson, N. L., *Tables of the Incomplete Beta Function*, Cambridge University Press, Cambridge, UK, 1968.

The quantile of order q is defined by

$$F(x_q; \lambda_1, \lambda_2) = q \tag{14.13}$$

and is easily computed with an equation solver. Note that Mathcad 6+ has a built-in function that gives the Beta quantile for positive shape parameters.

The first shape factor (see Section 1.10) is

$$\gamma_1 = \frac{2(\lambda_2 - \lambda_1)}{\lambda_1 + \lambda_2 + 2}\sqrt{\frac{\lambda_1 + \lambda_2 + 1}{\lambda_1\lambda_2}}. \tag{14.14}$$

If $\lambda_1 = \lambda_2$, then $\gamma_1 = 0$, and the pdf is symmetrical. If $\lambda_2 > \lambda_1$ then $\gamma_1 > 0$, and the pdf is skewed to the right. Similarly, $\lambda_2 < \lambda_1$ gives $\gamma_1 < 0$ for left skew. The second shape factor is

$$\gamma_2 = \frac{3(\lambda_1 + \lambda_2 + 1)[2(\lambda_1 + \lambda_2)^2 + \lambda_1\lambda_2(\lambda_1 + \lambda_2 - 6)]}{\lambda_1\lambda_2(\lambda_1 + \lambda_2 + 2)(\lambda_1 + \lambda_2 + 3)}. \tag{14.15}$$

Because both shape factors are symmetrical functions of λ_1 and λ_2, interchanging the parameters in a pdf yields its mirror image. For the symmetrical case ($\lambda_1 = \lambda_2$), $\gamma_2 \to 3$ as λ becomes large, and the Beta pdf approaches the Normal model.

Since the Beta distribution features two shape parameters, there is high shape flexibility. Figures 14.1 to 14.5 indicate this flexibility for various combinations of parameter values. The resulting shapes include pdfs that are symmetrical: $\lambda_1 = \lambda_2$; skewed: $\lambda_1 \neq \lambda_2$; U-shaped: $\lambda_1, \lambda_2 < 1$; and J-shaped: $(\lambda_1 - 1) \cdot (\lambda_2 - 1) < 0$.

The matrix of minimum-variance-bounds (see Section 3.2) for estimators of λ_1 and λ_2 is

$$\begin{bmatrix} V_{\lambda_1\lambda_1} & V_{\lambda_1\lambda_2} \\ V_{\lambda_1\lambda_2} & V_{\lambda_2\lambda_2} \end{bmatrix} = \frac{1}{n\{ab - c(a+b)\}}\begin{bmatrix} b - c & c \\ c & a - c \end{bmatrix}, \tag{14.16}$$

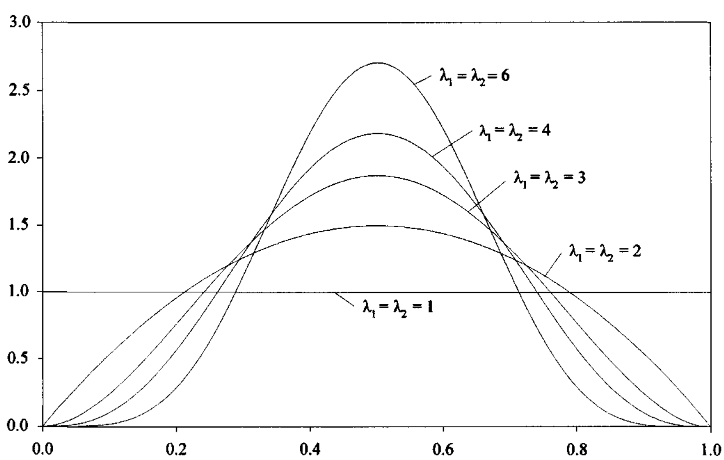


Figure 14.1. Unimodal, symmetric Beta pdfs.

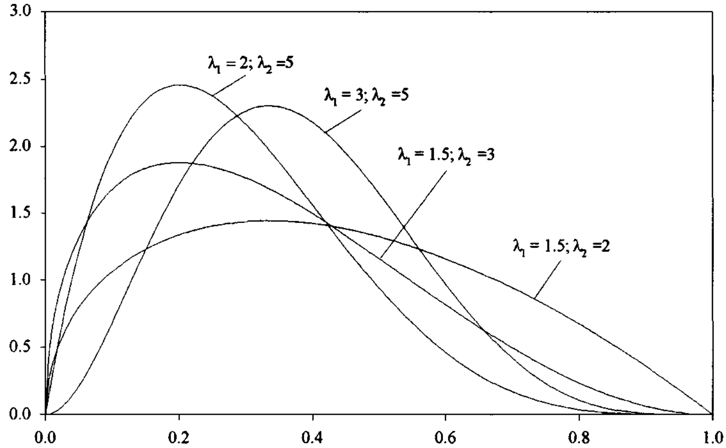


Figure 14.2. Unimodal, skewed Beta pdfs.

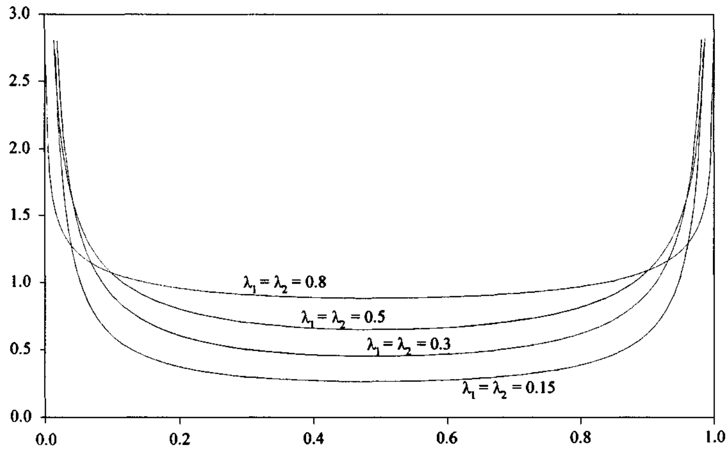


Figure 14.3. *U*-shaped, symmetrical Beta pdfs.

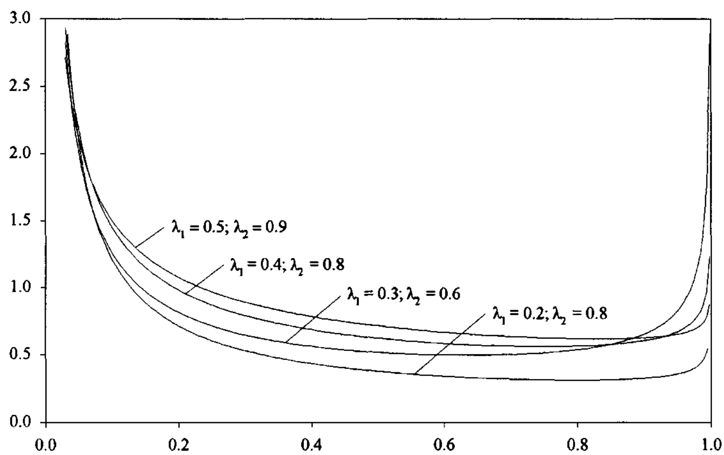


Figure 14.4. *U*-shaped, skewed Beta pdfs.

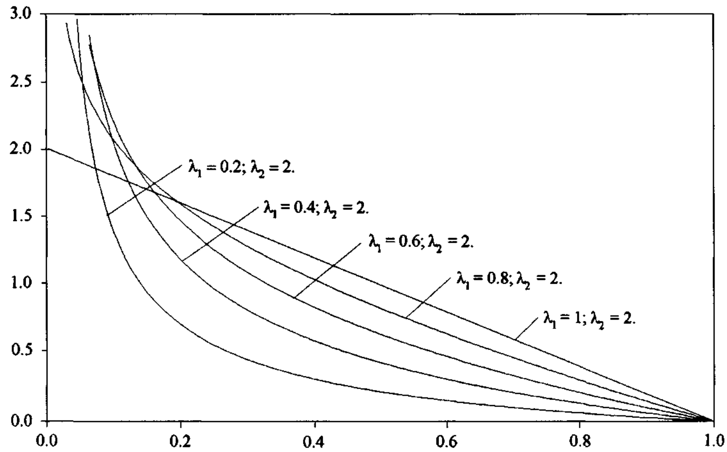


Figure 14.5. J-shaped Beta pdfs.

where $a = \psi'(\lambda_1), b = \psi'(\lambda_2)$, and $c = \psi'(\lambda_1 + \lambda_2)$. See Section 13.2 for the definition of the *trigamma* function $\psi'(\lambda)$.

14.4 Simulation

Random observations x_i from a Beta process with known parameters λ_1, λ_2 can be simulated by a simple *rejection technique*[2] for unimodal pdfs with $\lambda_1, \lambda_2 > 1$:

1. Calculate the mode x_m from (14.12) and its density value $M = f(x_m; \lambda_1, \lambda_2)$.
2. Generate two random numbers u_1, u_2 on the interval (0, 1).
3. If $u_2 \leq f(u_1; \lambda_1, \lambda_2)/M$, then accept u_1 as a random observation from f; otherwise reject u_1 and repeat the process.

This scheme is based on the recognition that the probability of u_2 being less than or equal to $f(x)/M$ is equal to $f(x)/M$. Hence, the pdf of accepted observations x will be $f(x)$.

For $\lambda_1 \leq 1$ and $\lambda_2 \leq 1$, the following scheme can be used.[3] Generate two Uniform random numbers u and v on the interval $(0, 1)$ until the condition

$$u^{1/\lambda_1} + v^{1/\lambda_2} \leq 1$$

[2] More efficient but more complicated rejection schemes are given in "Beta Variate Generation via Exponential Majorizing Functions," Schmeiser, B. and Babu, A. J. G., *Operations Research*, Vol. 28, pp. 917–926, 1980.

[3] Jöhnk, M. D., "Erzeugung von Betaverteilten und Gammaverteilten Zufallszahlen." *Metrika*, Vol. 8, pp. 5–15, 1964.

is satisfied. Then

$$x = \frac{u^{1/\lambda_1} + v^{1/\lambda_2}}{u^{1/\lambda_1}}$$

is a random observation from the Beta distribution $F(x; \lambda_1, \lambda_2)$. This scheme can also be used when $\lambda_1 > 1$ and/or $\lambda_2 > 1$, but it is inefficient for $(\lambda_1 + \lambda_2)$ large.

For small simulated samples it is perhaps more convenient to invert the given cdf directly: A simulated observation x_i from the Beta cdf (14.3) is simply the u_i-quantile defined by $F(x_i; \lambda_1, \lambda_2) = u_i$. In any case, it is advisable to check the adequacy of a simulated sample by comparing at least its first two moments with those of the given model. See Example 14.1 for an illustration. Note that Mathcad 6+ has a built-in function that generates random Beta observations for all values $\lambda_1 > 0$ and $\lambda_2 > 0$.

PROPERTIES: FOUR-PARAMETER MODEL

14.5 Definition and Properties

The two-parameter beta pdf (14.1) can model engineering variables on the unit range (0, 1); a *proportion* is an important example. More generally, the finite sample space is arbitrary.

> For example, the cost of engineering projects of a given type, scope, and complexity is a random variable that clearly exhibits a minimum bound different from zero. As well, an upper bound necessarily exists, usually dictated by the available budget, and is different from 1.

The Beta pdf (14.1) is generalized to accommodate different finite sample spaces by introducing two location parameters μ_1 and μ_2, with $\mu_1 < \mu_2$, to give

$$f(x; \mu_1, \mu_2, \lambda_1, \lambda_2) \\ = \frac{\Gamma(\lambda_1 + \lambda_2)}{\Gamma(\lambda_1)\Gamma(\lambda_2)} \left(\frac{x - \mu_1}{\mu_2 - \mu_1}\right)^{\lambda_1 - 1} \left(1 - \frac{x - \mu_1}{\mu_2 - \mu_1}\right)^{\lambda_2 - 1} \frac{1}{\mu_2 - \mu_1}, \qquad (14.17)$$

for $\mu_1 \leq x \leq \mu_2$. We see that $\sigma = \mu_2 - \mu_1$ is a scale parameter. Thus, two-parameter location measures are relocated by μ_1 and rescaled by σ. The expected value is

$$\mu_1'(X) = \mu_1 + (\mu_2 - \mu_1)\frac{\lambda_1}{\lambda_1 + \lambda_2}, \qquad (14.18)$$

and for $\lambda_1 + \lambda_2 > 2, \lambda_1 \geq 1$ the mode is

$$x_{\mathrm{m}} = \mu_1 + (\mu_2 - \mu_1)\frac{\lambda_1 - 1}{\lambda_1 + \lambda_2 - 2}. \qquad (14.19)$$

Two-parameter dispersion measures are rescaled by σ, so that the variance is

$$\mu_2(X) = \frac{(\mu_2 - \mu_1)^2 \lambda_1 \lambda_2}{(\lambda_1 + \lambda_2)^2 (\lambda_1 + \lambda_2 + 1)} \tag{14.20}$$

and the coefficient of variation becomes

$$cv(X) = \frac{(\mu_2 - \mu_1)\sqrt{\lambda_1 \lambda_2}}{\sqrt{\lambda_1 + \lambda_2 + 1}(\mu_1 \lambda_2 + \mu_2 \lambda_1)}. \tag{14.21}$$

Provided that $\lambda_1, \lambda_2 > 2$, the *expected* information matrix (see Section 3.1) is

$$\begin{bmatrix} I_{\mu_1\mu_1} = \frac{n\lambda_2(\lambda_1+\lambda_2-1)}{(\lambda_1-2)\sigma^2} & I_{\mu_1\mu_2} = \frac{n(\lambda_1+\lambda_2-1)}{\sigma^2} & I_{\mu_1\lambda_1} = \frac{n\lambda_2}{(\lambda_1-1)\sigma} & I_{\mu_1\lambda_2} = \frac{-n}{\sigma} \\ I_{\mu_1\mu_2} & I_{\mu_2\mu_2} = \frac{n\lambda_1(\lambda_1+\lambda_2-1)}{(\lambda_2-2)\sigma^2} & I_{\mu_2\lambda_1} = \frac{n}{\sigma} & I_{\mu_2\lambda_2} = \frac{-n\lambda_1}{(\lambda_2-1)\sigma} \\ I_{\mu_1\lambda_1} & I_{\mu_2\lambda_1} & I_{\lambda_1\lambda_1} = n[\psi'(\lambda_1) - \psi'(\lambda_1+\lambda_2)] & I_{\lambda_1\lambda_2} = -n\psi'(\lambda_1+\lambda_2) \\ I_{\mu_1\lambda_2} & I_{\mu_2\lambda_2} & I_{\lambda_1\lambda_2} & I_{\lambda_2\lambda_2} = n[\psi'(\lambda_2) - \psi'(\lambda_1+\lambda_2)] \end{bmatrix}, \tag{14.22}$$

where $\psi'(\lambda)$ is the *trigamma function* (see Section 13.2) and $\sigma = \mu_2 - \mu_1$. The minimum-variance-bounds are obtained from the inverse of (14.22). Alternatively, the *local* information matrix can be used (see Section 3.7), provided $\lambda_1, \lambda_2 > 1$. This matrix is identical to (14.22), except for two diagonal elements:

$$\begin{aligned} I_{\mu_1\mu_1} &= n(\lambda_1 - 1)s_1 - \frac{n(\lambda_1 + \lambda_2 - 1)}{\sigma^2} \quad \text{and} \\ I_{\mu_2\mu_2} &= n(\lambda_2 - 1)s_2 - \frac{n(\lambda_1 + \lambda_2 - 1)}{\sigma^2}, \end{aligned} \tag{14.23}$$

where

$$s_1 = \frac{1}{n}\sum_i (x_i - \mu_1)^{-2} \quad \text{and} \quad s_2 = \frac{1}{n}\sum_i (\mu_2 - x_i)^{-2}.$$

SPECIAL CASE: UNIFORM DISTRIBUTIONS

14.6 Definition and Properties

When the shape parameters of the Beta pdf (14.1) take the value $\lambda_1 = \lambda_2 = 1$, the *Uniform* or *Rectangular* distribution on (0, 1) results, meaning that all possible values x are equally likely:

$$f(x) = 1; \quad 0 \le x \le 1, \tag{14.24}$$

with cdf

$$F(x) = x. \tag{14.25}$$

The expected value of X is

$$\mu_1'(X) = \frac{1}{2}, \tag{14.26}$$

and the variance is

$$\mu_2(X) = \frac{1}{12}. \tag{14.27}$$

The quantile of order q is

$$x_q = q. \tag{14.28}$$

This is the distribution of "random numbers." The above properties can be used as a first check on the adequacy of a simulated random number set or of a random number generator itself.

The pdf (14.24) generalizes to cover an arbitrary interval (μ_1, μ_2):

$$f(x; \mu_1, \mu_2) = \frac{1}{\mu_2 - \mu_1}; \quad \mu_1 \le x \le \mu_2, \quad 0 \le \mu_1 < \mu_2. \tag{14.29}$$

The cdf is

$$F(x; \mu_1, \mu_2) = \frac{x - \mu_1}{\mu_2 - \mu_1}. \tag{14.30}$$

The expected value of X is

$$\mu_1'(X) = \frac{\mu_1 + \mu_2}{2}, \tag{14.31}$$

and the variance is

$$\mu_2(X) = \frac{(\mu_2 - \mu_1)^2}{12}. \tag{14.32}$$

The quantile of order q is

$$x_q = \mu_1 + (\mu_2 - \mu_1)q. \tag{14.33}$$

Neither distribution is useful as a *measurement* model of engineering random variables, since such variables practically never exhibit equiprobable values over their sample spaces. The importance of the Uniform distribution rests on the fact that *any* cdf, considered as a random function, is itself distributed according to (14.24) (see Section 2.8). This pdf is therefore at the core of all *Monte Carlo* simulation work.

PROBABILITY PLOT

14.7 Plotting Procedure

A probability plot of ordered data (see Section 3.13) should always be used to check the distributional assumption of the Beta model. Since the Beta cdf (14.3) cannot be expressed in closed form, it is not possible to linearize it algebraically. However, once the model has been estimated, it can be linearized numerically over the sample space (0, 1), and the data plotting positions can be adjusted by the same amounts as the estimated model. The result is a linearized graphical comparison of data and estimated model that provides visual information on how well the model fits the data.

Linearizing the *estimated* Beta cdf over (0, 1) gives the straight-line model ordinate as $x_{(i)}$ from expression (13.27) for the general linearization scheme, since for $a = 0$ and $b = 1$ we have $F(a) = 0$ and $F(b) = 1$. The adjustment from the estimated cdf to its linearized version is therefore

$$\Delta_i = F\left(x_{(i)}\right) - x_{(i)}, \tag{14.34}$$

so that the median data plotting position is adjusted to

$$p_i = \frac{i - 0.3}{n + 0.4} - \Delta_i. \tag{14.35}$$

The plot of p_i versus $x_{(i)}$ will roughly follow the linear model plot, if $X_{(i)}$ came from the estimated two-parameter Beta process (see Section 3.14). See Example 14.2 for an illustration (following model estimation).

For the four-parameter model, the observations x_i are reduced to the two-parameter case by the estimated location parameters $\widehat{\mu}_1$ and $\widehat{\mu}_2$:

$$z_i = \frac{x_i - \widehat{\mu}_1}{\widehat{\mu}_2 - \widehat{\mu}_1}, \tag{14.36}$$

and z takes the place of x in the above plotting procedure.

POINT ESTIMATES: TWO-PARAMETER MODEL

14.8 Maximum Likelihood Estimates

The likelihood function of a sample of n independent observations on a Beta variable X is

$$L(\lambda_1, \lambda_2) = B^{-n}(\lambda_1, \lambda_2) \prod_{i=1}^{n} x_i^{\lambda_1 - 1} \prod_{i=1}^{n} (1 - x_i)^{\lambda_2 - 1}, \tag{14.37}$$

where $B(\lambda_1, \lambda_2)$ is the beta function (14.4). The maximum likelihood equations (see Section 3.6) are

$$\psi(\widehat{\lambda}_1) - \psi(\widehat{\lambda}_1 + \widehat{\lambda}_2) = s_3 \tag{14.38}$$

and

$$\psi(\widehat{\lambda}_2) - \psi(\widehat{\lambda}_1 + \widehat{\lambda}_2) = s_4, \tag{14.39}$$

where $s_3 = \frac{1}{n}\sum_i \ln(x_i)$, $s_4 = \frac{1}{n}\sum_i \ln(1 - x_i)$, and $\psi(\lambda)$ is the *digamma function*, defined in Section 13.2.

The solution is easily obtained with an equation solver. To locate starting values for the solution process, one can use *moment* estimates (next section) as ballpark values. A display of the log-likelihood function (or its contour plot) in the neighborhood of these values may then show the location of the maximum, which serves as the starting point for the numerical solution. The desired plotting function is

$$\begin{aligned} LLF(\lambda_1, \lambda_2) = {} & n\ln[\Gamma(\lambda_1 + \lambda_2)] - n\ln[\Gamma(\lambda_1)] - n\ln[\Gamma(\lambda_2)] \\ & + n(\lambda_1 - 1)s_3 + n(\lambda_2 - 1)s_4 + C, \end{aligned} \tag{14.40}$$

where C is an arbitrary constant, chosen to give small positive values of the function near the maximum. See Example 14.2 for an illustration of the computations.

Approximate standard errors of these estimates are obtained from the matrix of minimum-variance-bounds (14.16). These error values are used as well to determine approximate standard errors of parameter functions $g(\lambda_1, \lambda_2)$ using the error propagation formula (see Section 3.3):

$$\mathrm{Var}(g) \doteq \left(\frac{\partial g}{\partial \lambda_1}\right)^2 V_{\lambda_1\lambda_1} + \left(\frac{\partial g}{\partial \lambda_2}\right)^2 V_{\lambda_2\lambda_2} + 2\left(\frac{\partial g}{\partial \lambda_1}\right)\left(\frac{\partial g}{\partial \lambda_2}\right) V_{\lambda_1\lambda_2}. \tag{14.41}$$

Statistical tests-of-fit are not available for Beta models. A linearized model/data plot (see Section 14.7) is recommended as a reliable graphical check on the Beta postulate.

14.9 Moment Estimates

In order to provide starting values for the ML solution process (preceding section), moment estimates may be obtained by equating the first two distribution moments, from (14.8), to corresponding data moments about the origin, and solving for the parameters

$$\widetilde{\lambda}_1 = \frac{m_1^2 - m_1 m_2}{m_2 - m_1^2} \tag{14.42}$$

and

$$\widetilde{\lambda}_2 = \frac{m_1 - m_2}{m_2 - m_1^2} - \widetilde{\lambda}_1, \tag{14.43}$$

where $m_r = \frac{1}{n}\sum_i x_i^r$. See Example 14.2 for an illustration.

INTERVAL ESTIMATES: TWO-PARAMETER MODEL

14.10 Normal Approximation

For large samples, approximate confidence intervals on the parameters λ_1 and λ_2 can be constructed from the asymptotic sampling distributions of the ML estimates (see Section 3.7). That is, the sampling pdf of $\widehat{\theta}$ is asymptotically Normal with mean θ and variance MVB_θ:

$$f_N(\widehat{\theta}; \theta, \sqrt{MVB_\theta}), \tag{14.44}$$

where θ stands for λ_1 or λ_2. Thus, the $(1-\alpha)$-level confidence interval on θ is obtained as

$$(l_1, l_2) = \theta \pm z_{\frac{\alpha}{2}} \sqrt{MVB_\theta}, \tag{14.45}$$

such that

$$Pr(l_1 \leq \theta \leq l_2) = 1 - \alpha.$$

See Example 14.3 for an illustration.

Approximate confidence intervals for parameter functions $g(\lambda_1, \lambda_2)$ are similarly obtained, with the error propagation formula (14.41) providing the function's variance estimate. Again see Example 14.3.

14.11 Likelihood Ratio Approximation

For small to moderate-sized samples, likelihood ratio methods tend to give more accurate results than the above Normal method. Recall from Section 3.9 that the statistic

$$LR(\theta) = 2\ln[L(\widehat{\theta})] - 2\ln[L(\theta)] \tag{14.46}$$

is approximately Chi-squared distributed with $\nu = 1$ degree of freedom. Thus, a $(1-\alpha)$-level confidence interval on θ comprises those values θ for which $LR(\theta) \leq \chi^2_{1,1-\alpha}$. For a two-parameter model, one parameter in (14.46) must be expressed in the other parameter by its ML equation. Thus, a confidence interval on λ_1 is obtained from

$$LR(\lambda_1) = 2\ln[L(\widehat{\lambda}_1, \widehat{\lambda}_2)] - 2\ln[L(\lambda_1, \lambda_2\{\lambda_1\})], \tag{14.47}$$

where $\lambda_2\{\lambda_1\}$ is defined by ML equation (14.39). Similarly, a confidence interval on λ_2 is obtained from

$$LR(\lambda_2) = 2\ln[L(\widehat{\lambda}_1, \widehat{\lambda}_2)] - 2\ln[L(\lambda_1\{\lambda_2\}, \lambda_2)], \tag{14.48}$$

where $\lambda_1\{\lambda_2\}$ is defined by the ML equation (14.38). See Example 14.3 for an illustration.

POINT ESTIMATES: FOUR-PARAMETER MODEL

14.12 Maximum Likelihood Estimates

The likelihood function of a sample of n independent observations from a four-parameter Beta process is

$$L(\mu_1, \mu_2, \lambda_1, \lambda_2) \\ = B^{-n}(\lambda_1, \lambda_2) \prod_{i=1}^{n} (x_i - \mu_1)^{\lambda_1 - 1} \prod_{i=1}^{n} (\mu_2 - x_i)^{\lambda_2 - 1} (\mu_2 - \mu_1)^{n(1-\lambda_1-\lambda_2)}, \tag{14.49}$$

where $B(\lambda_1, \lambda_2)$ is the beta function (14.4). The ML equations (see Section 3.6), are

$$\psi(\widehat{\lambda}_1) - \psi(\widehat{\lambda}_1 + \widehat{\lambda}_2) = \frac{1}{n} \sum_{i=1}^{n} \ln \left(\frac{x_i - \widehat{\mu}_1}{\widehat{\mu}_2 - \widehat{\mu}_1} \right), \tag{14.50}$$

$$\psi(\widehat{\lambda}_2) - \psi(\widehat{\lambda}_1 + \widehat{\lambda}_2) = \frac{1}{n} \sum_{i=1}^{n} \ln \left(\frac{\widehat{\mu}_2 - x_i}{\widehat{\mu}_2 - \widehat{\mu}_1} \right), \tag{14.51}$$

$$\frac{1 - \widehat{\lambda}_1 - \widehat{\lambda}_2}{\widehat{\mu}_2 - \widehat{\mu}_1} + \frac{\widehat{\lambda}_1 - 1}{n} \sum_{i=1}^{n} (x_i - \widehat{\mu}_1)^{-1} = 0, \tag{14.52}$$

$$\frac{1 - \widehat{\lambda}_1 - \widehat{\lambda}_2}{\widehat{\mu}_2 - \widehat{\mu}_1} + \frac{\widehat{\lambda}_2 - 1}{n} \sum_{i=1}^{n} (\widehat{\mu}_2 - x_i)^{-1} = 0, \tag{14.53}$$

where $\psi(\lambda)$ is the *digamma function* (see Section 13.2). Solving (14.52) and (14.53) for $\widehat{\lambda}_1$ and $\widehat{\lambda}_2$ gives

$$\widehat{\lambda}_1 = \frac{s_5(\sigma s_6 - 1)}{s_6(\sigma s_5 - 1) - s_5} \tag{14.54}$$

and

$$\widehat{\lambda}_2 = \frac{s_6(\sigma s_5 - 1)}{s_6(\sigma s_5 - 1) - s_5}, \tag{14.55}$$

where $s_5 = \frac{1}{n} \sum_i (x_i - \widehat{\mu}_1)^{-1}$, $s_6 = \frac{1}{n} \sum_i (\widehat{\mu}_2 - x_i)^{-1}$, and $\sigma = \widehat{\mu}_2 - \widehat{\mu}_1$. Substituting (14.54) and (14.55) into (14.50) and (14.51) gives two expressions in $\widehat{\mu}_1$ and $\widehat{\mu}_2$ only:

$$\psi\left(\frac{s_5(\sigma s_6 - 1)}{s_6(\sigma s_5 - 1) - s_5} \right) - \psi\left(1 + \frac{\sigma s_5 s_6}{s_6(\sigma s_5 - 1) - s_5} \right) = s_7 - \ln(\sigma) \tag{14.56}$$

and

$$\psi\left(\frac{s_6(\sigma s_5 - 1)}{s_6(\sigma s_5 - 1) - s_5} \right) - \psi\left(1 + \frac{\sigma s_5 s_6}{s_6(\sigma s_5 - 1) - s_5} \right) = s_8 - \ln(\sigma), \tag{14.57}$$

where $s_7 = \frac{1}{n} \sum_i \ln(x_i - \widehat{\mu}_1)$ and $s_8 = \frac{1}{n} \sum_i \ln(\widehat{\mu}_2 - x_i)$.

In principle, the solution $(\widehat{\mu}_1, \widehat{\mu}_2)$ is obtained with an equation solver, followed by $(\widehat{\lambda}_1, \widehat{\lambda}_2)$ from (14.54) and (14.55). However, a solution may not exist, as typically happens when the sample size is small. Even when a solution does exist, it may not be simple to locate it: Good starting values for μ_1 and μ_2 are essential. Moment estimates (next section) may provide ballpark values, although they are often some distance from the ML solution.

It is convenient to display the log-likelihood function, expressed in μ_1 and μ_2, so that one can visually search for the maximum and obtain close starting values for the ML solution process:

$$\begin{aligned} LLF(\mu_1, \mu_2) = {} & -n \ln[B(\lambda_1, \lambda_2)] + n(1 - \lambda_1 - \lambda_2)\ln(\sigma) \\ & + n(\lambda_1 - 1)s_7 + n(\lambda_2 - 1)s_8 + C. \end{aligned} \tag{14.58}$$

The terms λ_1 and λ_2 are given by (14.54) and (14.55), $\sigma = \mu_2 - \mu_1$, and C is an arbitrary constant, chosen to produce small positive values of the function near its maximum. The difficulty with finding a solution is that the likelihood function does not appear to hold much information on the location parameters. This is indicated by the typically shallow surface of that function.

Approximate standard errors of the estimates can be obtained from the inverse of the *expected* information matrix (14.22), provided $\lambda_1 > 2$ and $\lambda_2 > 2$, or the *local* version (14.23), provided $\lambda_1 > 1$ and $\lambda_2 > 1$. Although these formulas look intimidating, they are readily evaluated numerically at the ML estimates, using a modern computational tool (e.g., Mathcad). See Example 14.4. for an illustration of these computations. Approximate standard errors of parameter functions $g(\mu_1, \mu_2, \lambda_1, \lambda_2)$ are obtained from the error propagation formula (see Section 3.3).

Approximate confidence intervals on parameters can be constructed from the asymptotic formula (14.45) to give a rough idea of the uncertainties associated with estimated quantities. Small-sample methods for confidence intervals are not available.

Statistical tests-of-fit are not available for the four-parameter Beta model. A linearized model/data plot is recommended as a reliable check on the distributional postulate and the model fit; see Example 14.4.

14.13 Moment Estimates

When ML estimation fails, moment estimates may be used. The resulting model fit is usually acceptable, although standard errors are difficult to obtain for these estimates.

It is convenient to use data moments M_r about the mean, which are defined in terms of moments m_r about the origin (see Section 1.8) as

$$\begin{aligned} M_2 &= m_2 - m_1^2, \\ M_3 &= m_3 - 3m_1m_2 + 2m_1^3, \\ M_4 &= m_4 - 4m_1m_3 + 6m_1^2m_2 - 3m_1^4, \end{aligned} \tag{14.59}$$

where $m_r = \frac{1}{n}\sum_i x_i^r$. To estimate the shape parameters the shape factors (14.14) and (14.15) provide two equations in $\tilde{\lambda}_1$ and $\tilde{\lambda}_2$:

$$\frac{M_3}{M_2^{3/2}} = \frac{2(\tilde{\lambda}_2 - \tilde{\lambda}_1)}{(\tilde{\lambda}_1 + \tilde{\lambda}_2 + 2)}\sqrt{\frac{\tilde{\lambda}_1 + \tilde{\lambda}_2 + 1}{\tilde{\lambda}_1\tilde{\lambda}_2}} \tag{14.60}$$

and

$$\frac{M_4}{M_2^2} = \frac{3(\tilde{\lambda}_1 + \tilde{\lambda}_2 + 1)[2(\tilde{\lambda}_1 + \tilde{\lambda}_2)^2 + \tilde{\lambda}_1\tilde{\lambda}_2(\tilde{\lambda}_1 + \tilde{\lambda}_2 - 6)]}{\tilde{\lambda}_1\tilde{\lambda}_2(\tilde{\lambda}_1 + \tilde{\lambda}_2 + 2)(\tilde{\lambda}_1 + \tilde{\lambda}_2 + 3)}. \tag{14.61}$$

The estimates $\tilde{\lambda}_1$ and $\tilde{\lambda}_2$ are readily computed with an equation solver. Location-parameter estimates are then obtained from the mean value (14.18), $m_1 = \mu_1'(X)$, and the variance (14.20), $M_2 = \mu_2(X)$, as

$$\tilde{\mu}_1 = m_1 - \sqrt{M_2}\sqrt{\frac{\tilde{\lambda}_1(\tilde{\lambda}_1 + \tilde{\lambda}_2 + 1)}{\tilde{\lambda}_2}} \tag{14.62}$$

and

$$\tilde{\mu}_2 = m_1 + \sqrt{M_2}\sqrt{\frac{\tilde{\lambda}_2(\tilde{\lambda}_1 + \tilde{\lambda}_2 + 1)}{\tilde{\lambda}_1}}. \tag{14.63}$$

See Example 14.5 where a small sample did not lead to a ML solution, but the moment estimates produced a reasonable model fit.

14.14 Conditional Inferences

Sometimes the values of the location parameters are known: μ_1^* and μ_2^*. It is then simple to transform the data x_i to their reduced equivalents by

$$z_i = \frac{x_i - \mu_1^*}{\mu_2^* - \mu_1^*} \tag{14.64}$$

and to use two-parameter methods on the reduced data z_i. See Example 14.6.

APPLICATIONS

14.15 Engineering

Because of its limited sample space, $\mu_1 \le x \le \mu_2$, the general Beta distribution (14.17) serves as a useful measurement model for engineering variables for which the assumptions of an unlimited upper tail and the lower tail terminating at the origin are inappropriate. Applications include cost variables, task completion times, and load variables subject to inherent or imposed (e.g., legal) limits. The two-parameter distribution (14.1), however, is a natural candidate for modeling engineering *ratios*, for example efficiency measures, which vary over the unit range. Furthermore, the exceptional shape flexibility of the Beta distribution makes it attractive as a general measurement model and thus it is increasingly used in engineering work for variables subject to range limitations.

An interesting application of the Beta distribution occurs in the coordination of complex engineering projects that involve tasks of uncertain durations X_i. Such projects are often controlled by PERT (Project Evaluation and Control Technique). PERT considers task durations X_i as Beta random variables and requires input estimates of the expected value and variance for each project task i. Practically, these estimates are difficult to obtain directly. Hence, intuitively more accessible estimates are obtained instead, namely a *most likely* task duration m_i, an *optimistic* time a_i, and a *pessimistic* time b_i. These quantities are equivalent to the mode value x_m and to the location parameters μ_1 and μ_2, respectively.

The expected value and variance of task duration are then calculated using the following assumptions:

$$\mathrm{E}\{X\} = \frac{1}{3}\left[2m + \frac{1}{2}(a+b)\right] \tag{14.65}$$

and

$$\mathrm{Var}(X) = \frac{(b-a)^2}{36}. \tag{14.66}$$

By expressing the expected value (14.18) in terms of the mode value (14.19) and comparing with (14.65), we see that the shape parameters are constrained by $\lambda_1 + \lambda_2 = 6$. Similarly, approximating the variance (14.20) by (14.66) implies that $\lambda_1 \lambda_2 = 7$. Hence, PERT implicitly specifies the Beta shape parameters as (1.59, 4.41) for positive skew or (4.41, 1.59) for negative skew. Figure 14.6 shows these two Beta pdfs, reduced to the unit range. Despite this restriction, practice with

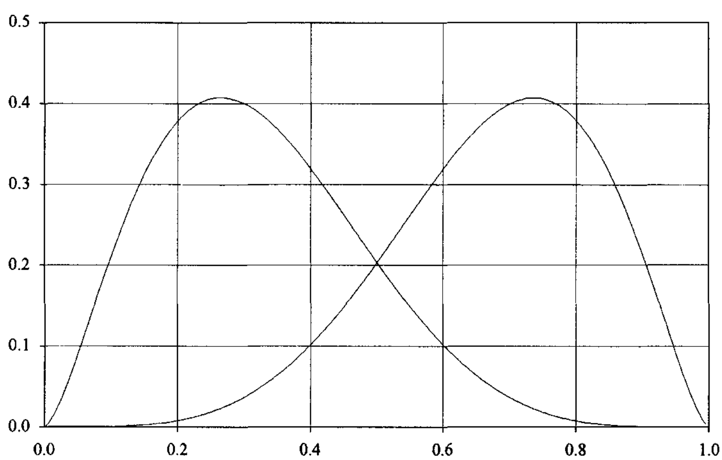


Figure 14.6. The two Beta pdfs implied by PERT.

PERT has shown that (14.65) and (14.66) give reasonable results, particularly in view of the inaccuracies that are inherent in the elicited estimates *m*, *a*, and *b*.

The expected value (14.65) and variance (14.66) of *task* durations are then used to construct a Normal probability distribution of *project* duration. This is done by invoking a central limit theorem (see Section 2.4) on the string of critical tasks that control the project duration. For large projects there are many such critical tasks, and thus the Normal distribution proves to be a valid assumption.

EXAMPLE 14.1

Generate 20 observations from a Beta distribution with $\lambda_1 = 3$ and $\lambda_2 = 5$.

$\lambda1 := 3 \qquad \lambda2 := 5 \qquad$ Beta function: $B := \dfrac{\Gamma(\lambda1) \cdot \Gamma(\lambda2)}{\Gamma(\lambda1 + \lambda2)}$

$n := 20 \qquad i := 1 \,..\, n \qquad u_i := rnd(1)$

$$F(x) := \frac{1}{B} \cdot \int_0^x z^{\lambda1-1} \cdot (1-z)^{\lambda2-1}\, dz \qquad x := 0.5 \qquad x_i := \mathrm{root}(F(x) - u_i, x)$$

Sample check:

mean value: $xb := \dfrac{1}{n} \cdot \sum_i x_i \qquad xb = 0.401$

This should be close to

$$\frac{\lambda1}{\lambda1 + \lambda2} = 0.375$$

variance: $S2 := \dfrac{1}{n} \cdot \sum_i (x_i - xb)^2 \qquad S2 = 0.021$

This should be close to

$$\frac{\lambda1 \cdot \lambda2}{(\lambda1 + \lambda2)^2 \cdot (\lambda1 + \lambda2 + 1)} = 0.026$$

x_i
0.461
0.432
0.237
0.113
0.526
0.278
0.275
0.309
0.67
0.428
0.556
0.402
0.472
0.226
0.632
0.533
0.309
0.417
0.495
0.241

EXAMPLE 14.2

For the simulated data of Example 14.1, estimate the Beta parameters and their standard errors. Also, estimate the 5-percentile and its standard error.

From Example 14.1:

$n := 20 \qquad i := 1 \,..\, n$

$x_i :=$ $\quad y := \mathrm{sort}(x)$

x_i
.461
.432
.237
.113
.526
.278
.275
.309
.67
.428
.556
.402
.472
.226
.632
.533
.309
.417
.495
.241

1. Moment estimates

$$m1 := \frac{1}{n} \cdot \sum_i x_i \qquad m2 := \frac{1}{n} \cdot \sum_i (x_i)^2$$

$$L1 := \frac{m1^2 - m1 \cdot m2}{m2 - m1^2} \qquad L1 = 4.222$$

$$L2 := \frac{m1 - m2}{m2 - m1^2} - L1 \qquad L2 = 6.317$$

2. Display the log-likelihood function

$$s3 := \frac{1}{n} \cdot \sum_i \ln(x_i) \qquad s4 := \frac{1}{n} \cdot \sum_i \ln(1 - x_i)$$

$$k := 1 \,..\, 10 \qquad p := 1 \,..\, 10$$

$$r_k := 3.6 + 0.1 \cdot k \qquad l_p := 6 + 0.1 \cdot p$$

$$B_{k,p} := \frac{\Gamma(r_k) \cdot \Gamma(l_p)}{\Gamma(r_k + l_p)} \qquad s3 = -0.9935$$

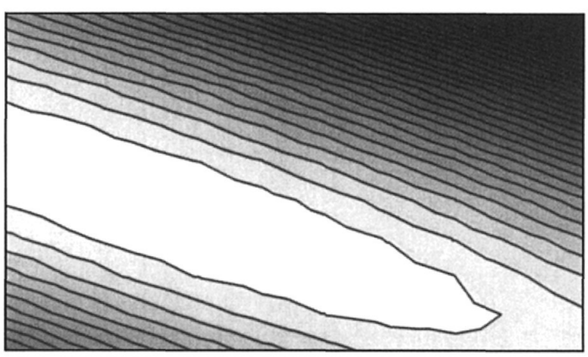
$$s4 = -0.5428$$

$$L_{k,p} := -n \cdot \ln(B_{k,p}) + n \cdot (r_k - 1) \cdot s3 + n \cdot (l_p - 1) \cdot s4 - 10$$

$$l^T = (6.100 \quad 6.200 \quad 6.300 \quad 6.400 \quad 6.500 \quad 6.600 \quad 6.700 \quad 6.800 \quad 6.900 \quad 7.000)$$

$$L = \begin{bmatrix} 0.32 & 0.24 & 0.146 & 0.04 & -0.08 & -0.213 & -0.357 & -0.514 & -0.683 & -0.863 \\ 0.439 & 0.381 & 0.308 & 0.222 & 0.123 & 0.011 & -0.113 & -0.25 & -0.399 & -0.559 \\ 0.52 & 0.483 & 0.431 & 0.366 & 0.287 & 0.195 & 0.09 & -0.027 & -0.156 & -0.297 \\ 0.563 & 0.547 & 0.516 & 0.471 & 0.412 & 0.34 & 0.255 & 0.158 & 0.048 & -0.073 \\ 0.571 & 0.575 & 0.564 & 0.539 & 0.501 & 0.449 & 0.383 & 0.305 & 0.215 & 0.112 \\ 0.543 & 0.568 & 0.577 & 0.573 & 0.554 & 0.521 & 0.475 & 0.417 & 0.345 & 0.261 \\ 0.483 & 0.528 & 0.557 & 0.572 & 0.573 & 0.56 & 0.533 & 0.493 & 0.441 & 0.376 \\ 0.39 & 0.455 & 0.504 & 0.539 & 0.559 & 0.565 & 0.557 & 0.536 & 0.503 & 0.456 \\ 0.266 & 0.351 & 0.42 & 0.474 & 0.513 & 0.538 & 0.549 & 0.547 & 0.532 & 0.504 \\ 0.113 & 0.217 & 0.305 & 0.378 & 0.437 & 0.481 & 0.511 & 0.527 & 0.53 & 0.52 \end{bmatrix} \qquad r = \begin{bmatrix} 3.7 \\ 3.8 \\ 3.9 \\ 4 \\ 4.1 \\ 4.2 \\ 4.3 \\ 4.4 \\ 4.5 \\ 4.6 \end{bmatrix}$$

$$M_{p,11-k} := L_{k,p}$$



M

starting values: $\quad L1 := 4.2 \qquad L2 := 6.3$

3. Maximum likelihood estimates

digamma function: $\psi(L) := \frac{d}{dL}\ln(\Gamma(L))$

GIVEN $\psi(L1) - \psi(L1 + L2) = s3$

$\psi(L2) - \psi(L1 + L2) = s4$ $\quad \begin{pmatrix}\lambda 1\\ \lambda 2\end{pmatrix} := \text{FIND}\,(L1, L2)$

ML parameter estimates $\lambda 1 = 4.192$

$\lambda 2 = 6.305$

4. Standard errors

trigamma function: $\psi'(L) := \frac{d}{dL}\psi(L)$

$a := \psi'(\lambda 1)$ $\quad b := \psi'(\lambda 2)$ $\quad c := \psi'(\lambda 1 + \lambda 2)$ $\quad d := n \cdot (a \cdot b - c \cdot (a + b))$

$SE\lambda 1 := \sqrt{\frac{b-c}{d}}$ $\quad SE\lambda 2 := \sqrt{\frac{a-c}{d}}$ $\quad SE\lambda 12 := \sqrt{\frac{c}{d}}$ $\quad SE\lambda 1 = 1.283$

$SE\lambda 2 = 1.969$

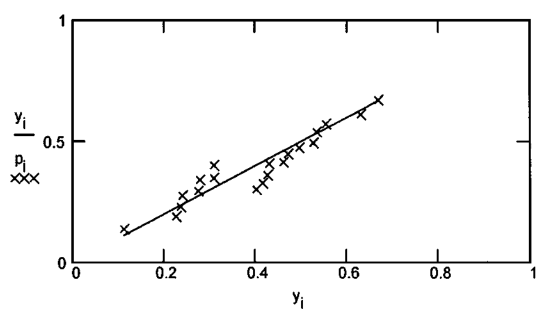
$SE\lambda 12 = 1.513$

5. Linearized model/data plot

$$F(x, \lambda 1, \lambda 2) := \frac{\Gamma(\lambda 1 + \lambda 2)}{\Gamma(\lambda 1) \cdot \Gamma(\lambda 2)} \cdot \int_0^x z^{\lambda 1 - 1} \cdot (1 - z)^{\lambda 2 - 1}\, dz$$

model ordinates: y_i

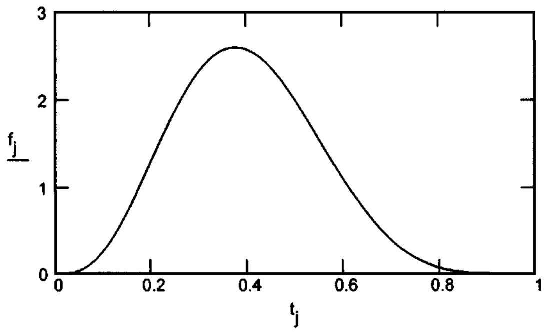
data ordinates: $p_i := \frac{i - 0.3}{n + 0.4} - F(y_i, \lambda 1, \lambda 2) + y_i$



The estimated model fits the data reasonably well.

6. Density plot

$j := 1\,..\,199$ $\quad t_j := \frac{j}{200}$ $\quad f_j := \frac{\Gamma(\lambda 1 + \lambda 2)}{\Gamma(\lambda 1) \cdot \Gamma(\lambda 2)} \cdot (t_j)^{\lambda 1 - 1} \cdot (1 - t_j)^{\lambda 2 - 1}$



7. Percentile

$x := 0.2 \qquad x05 := \mathrm{root}(F(x, \lambda 1, \lambda 2) - 0.05, x) \qquad\qquad x05 = 0.173$

standard error: $d1 := \dfrac{d}{d\lambda 1} F(x05, \lambda 1, \lambda 2) \qquad d2 := \dfrac{d}{d\lambda 2} F(x05, \lambda 1, \lambda 2)$

$\mathrm{SE}x := \sqrt{(d1 \cdot \mathrm{SE}\lambda 1)^2 + (d2 \cdot \mathrm{SE}\lambda 2)^2 + 2 \cdot d1 \cdot d2 \cdot \mathrm{SE}\lambda 12^2} \qquad\qquad \mathrm{SE}x = 0.036$

EXAMPLE 14.3

For the data of Example 14.1, calculate 90% confidence limits on the shape parameters and on the 5-percentile.

From Example 14.2:

$$\begin{array}{lllll} & \lambda 1 := 4.192 & \lambda 2 := 6.305 & & s3 := -0.9935 \\ n := 20 & \mathrm{SE}\lambda 1 := 1.283 & \mathrm{SE}\lambda 2 := 1.969 & \mathrm{SE}\lambda 12 := 1.513 & s4 := -0.5428 \end{array}$$

1. Parameters

i) Normal approximation: Standard Normal 95-percentile is $z95 := 1.645$

$$\begin{array}{ll} L\lambda 1 := \lambda 1 - z95 \cdot \mathrm{SE}\lambda 1 & L\lambda 1 = 2.08 \\ U\lambda 1 := \lambda 1 + z95 \cdot \mathrm{SE}\lambda 1 & U\lambda 1 = 6.30 \\ L\lambda 2 := \lambda 2 - z95 \cdot \mathrm{SE}\lambda 2 & L\lambda 2 = 3.07 \\ U\lambda 2 := \lambda 2 + z95 \cdot \mathrm{SE}\lambda 2 & U\lambda 2 = 9.54 \end{array}$$

ii) Likelihood ratio method: Chi-squared 90-percentile at $\nu = 1$ is $\mathrm{K} := 2.71$

$$B(a, b) := \frac{\Gamma(a) \cdot \Gamma(b)}{\Gamma(a + b)} \qquad \psi(a) := \frac{d}{da} \ln(\Gamma(a))$$

log of likelihood function:

$$LL(a, b) := -n \cdot \ln(B(a, b)) + n \cdot (a - 1) \cdot s3 + n \cdot (b - 1) \cdot s4$$

shape parameter λ_1: $ML(a, b) := \psi(b) - \psi(a+b) - s4$

$b := \lambda 2 \qquad I(a) := \mathrm{root}(ML(a, b), b)$

$\qquad LR(a) := 2 \cdot LL(\lambda 1, \lambda 2) - 2 \cdot LL(a, I(a))$

$a := 2 \qquad \lambda 1L := \mathrm{root}(LR(a) - K, a) \qquad \lambda 1L = 2.43$

$a := 6 \qquad \lambda 1U := \mathrm{root}(LR(a) - K, a) \qquad \lambda 1U = 6.68$

shape parameter λ_2: $ML(a, b) := \psi(a) - \psi(a+b) - s3$

$a := \lambda 1 \qquad I(b) := \mathrm{root}(ML(a, b), a)$

$\qquad LR(b) := 2 \cdot LL(\lambda 1, \lambda 2) - 2 \cdot LL(I(b), b)$

$b := 3 \qquad \lambda 2L := \mathrm{root}(LR(b) - K, b) \qquad \lambda 2L = 3.60$

$b := 10 \qquad \lambda 2U := \mathrm{root}(LR(b) - K, b) \qquad \lambda 2U = 10.13$

2. Percentile

estimate:

$$q := 0.05 \qquad F(x, \lambda 1, \lambda 2) := \frac{1}{B(\lambda 1, \lambda 2)} \cdot \int_0^x t^{\lambda 1 - 1} \cdot (1-t)^{\lambda 2 - 1} dt$$

$x := 0.1 \qquad x05(\lambda 1, \lambda 2) := \mathrm{root}(F(x, \lambda 1, \lambda 2) - q, x) \qquad x05(\lambda 1, \lambda 2) = 0.17$

standard error:

$$d1 := \frac{d}{d\lambda 1} x05(\lambda 1, \lambda 2) \qquad d2 := \frac{d}{d\lambda 2} x05(\lambda 1, \lambda 2)$$

$$\mathrm{var}x := (d1 \cdot \mathrm{SE}\lambda 1)^2 + (d2 \cdot \mathrm{SE}\lambda 2)^2 + 2 \cdot d1 \cdot d2 \cdot \mathrm{SE}\lambda 12^2 \qquad \mathrm{SE}x := \sqrt{\mathrm{var}x}$$

$\mathrm{SE}x = 0.04$

Normal interval approximation:

$Lx := x05(\lambda 1, \lambda 2) - z95 \cdot \mathrm{SE}x \qquad Lx = 0.11$

$Ux := x05(\lambda 1, \lambda 2) + z95 \cdot \mathrm{SE}x \qquad Ux = 0.23$

EXAMPLE 14.4

A sample of 182 observations is available from a test program on the compressive strength (kg/cm^2) of a concrete. The sample is too large to present in print; it is read from a computer file. Assuming a four-parameter Beta distribution for strength, estimate the parameters and their standard errors.

$n := 182 \qquad i := 1..n$

Read the data file called "*df*": $y_i := \mathrm{READ}(df)$

$x := \mathrm{sort}(y) \qquad x_1 = 209.000 \qquad mx := 209 \qquad x_n = 335.000 \qquad Mx := 335$

1. Display of log-likelihood function

$j := 1..10 \qquad a_j := 206.6 + 0.2 \cdot j \qquad k := 1..10 \qquad b_k := 338 + k$

$$s5_j := \frac{1}{n} \cdot \sum_i (x_i - a_j)^{-1} \qquad s6_k := \frac{1}{n} \cdot \sum_i (b_k - x_i)^{-1}$$

$$s7_j := \frac{1}{n} \cdot \sum_i \ln(x_i - a_j) \qquad s8_k := \frac{1}{n} \cdot \sum_i \ln(b_k - x_i)$$

$$(14.54): \quad L1_{j,k} := \frac{s5_j \cdot [(b_k - a_j) \cdot s6_k - 1]}{s6_k \cdot [(b_k - a_j) \cdot s5_j - 1] - s5_j}$$

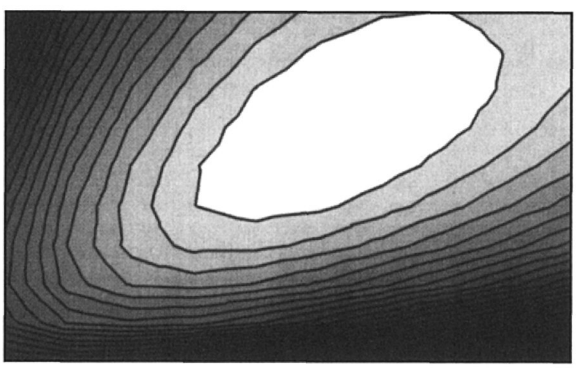
$$(14.55): \quad L2_{j,k} := \frac{s6_k \cdot [(b_k - a_j) \cdot s5_j - 1]}{s6_k \cdot [(b_k - a_j) \cdot s5_j - 1] - s5_j} \qquad B_{j,k} := \frac{\Gamma(L1_{j,k}) \cdot \Gamma(L2_{j,k})}{\Gamma(L1_{j,k} + L2_{j,k})}$$

$$LLF_{j,k} := -n \cdot \ln(B_{j,k}) + n \cdot (1 - L1_{j,k} - L2_{j,k}) \cdot \ln(b_k - a_j) + n \cdot (L1_{j,k} - 1) \cdot s7_j + n \cdot (L2_{j,k} - 1) \cdot s8_k + 866$$

$b^T = (339.0 \quad 340.0 \quad 341.0 \quad 342.0 \quad 343.0 \quad 344.0 \quad 345.0 \quad 346.0 \quad 347.0 \quad 348.0)$

$$LLF = \begin{bmatrix} -0.044 & 0.255 & 0.445 & 0.568 & 0.645 & 0.690 & 0.711 & 0.715 & 0.704 & 0.683 \\ 0.043 & 0.327 & 0.505 & 0.617 & 0.684 & 0.720 & 0.734 & 0.730 & 0.712 & 0.684 \\ 0.127 & 0.394 & 0.559 & 0.659 & 0.716 & 0.742 & 0.747 & 0.734 & 0.710 & 0.675 \\ 0.206 & 0.455 & 0.604 & 0.691 & 0.737 & 0.753 & 0.748 & 0.727 & 0.693 & 0.651 \\ 0.277 & 0.507 & 0.639 & 0.711 & 0.744 & 0.748 & 0.732 & 0.701 & 0.659 & 0.607 \\ 0.338 & 0.544 & 0.657 & 0.713 & 0.731 & 0.722 & 0.694 & 0.652 & 0.599 & 0.537 \\ 0.380 & 0.560 & 0.651 & 0.687 & 0.688 & 0.663 & 0.621 & 0.565 & 0.500 & 0.427 \\ 0.392 & 0.539 & 0.602 & 0.615 & 0.594 & 0.551 & 0.491 & 0.419 & 0.339 & 0.252 \\ 0.342 & 0.446 & 0.473 & 0.455 & 0.407 & 0.338 & 0.256 & 0.163 & 0.063 & -0.043 \\ 0.140 & 0.181 & 0.155 & 0.091 & 0.002 & -0.103 & -0.219 & -0.343 & -0.472 & -0.605 \end{bmatrix} \qquad a = \begin{bmatrix} 206.8 \\ 207.0 \\ 207.2 \\ 207.4 \\ 207.6 \\ 207.8 \\ 208.0 \\ 208.2 \\ 208.4 \\ 208.6 \end{bmatrix}$$

$$M_{j,11-k} := LLF_{k,j}$$



M

starting values: $a := 207.4 \qquad b := 344$

2. Maximum likelihood estimates

digamma function: $\psi(t) := \frac{d}{dt}\ln(\Gamma(t))$

$$s5(a) := \frac{1}{n}\cdot\sum_i (x_i - a)^{-1} \qquad s6(b) := \frac{1}{n}\cdot\sum_i (b - x_i)^{-1}$$

$$s7(a) := \frac{1}{n}\cdot\sum_i \ln(x_i - a) \qquad s8(b) := \frac{1}{n}\cdot\sum_i \ln(b - x_i)$$

GIVEN

$$\psi\left[\frac{s5(a)\cdot((b-a)\cdot s6(b)-1)}{s6(b)\cdot((b-a)\cdot s5(a)-1)-s5(a)}\right] - \psi\left[1+\frac{(b-a)\cdot s5(a)\cdot s6(b)}{s6(b)\cdot((b-a)\cdot s5(a)-1)-s5(a)}\right] = s7(a) - \ln(b-a)$$

$$\psi\left[\frac{s6(b)\cdot((b-a)\cdot s5(a)-1)}{s6(b)\cdot((b-a)\cdot s5(a)-1)-s5(a)}\right] - \psi\left[1+\frac{(b-a)\cdot s5(a)\cdot s6(b)}{s6(b)\cdot((b-a)\cdot s5(a)-1)-s5(a)}\right] = s8(b) - \ln(b-a)$$

constraints on the location parameters: $a < mx \qquad b > Mx$

$$\begin{pmatrix}\mu 1\\ \mu 2\end{pmatrix} := \mathrm{FIND}(a, b)$$

location parameters: $\mu 1 = 207.4$

$\mu 2 = 344.1$

$$\lambda 1 := \frac{s5(\mu 1)\cdot((\mu 2-\mu 1)\cdot s6(\mu 2)-1)}{s6(\mu 2)\cdot((\mu 2-\mu 1)\cdot s5(\mu 1)-1)-s5(\mu 1)}$$

shape parameters: $\lambda 1 = 1.625$

$$\lambda 2 := \frac{s6(\mu 2)\cdot((\mu 2-\mu 1)\cdot s5(\mu 1)-1)}{s6(\mu 2)\cdot((\mu 2-\mu 1)\cdot s5(\mu 1)-1)-s5(\mu 1)}$$

$\lambda 2 = 2.400$

3. Standard errors

$$s1 := \frac{1}{n}\cdot\sum_i (x_i - \mu 1)^{-2} \qquad s2 := \frac{1}{n}\cdot\sum_i (\mu 2 - x_i)^{-2}$$

trigamma function: $\psi'(t) := \frac{d}{dt}\psi(t) \qquad \sigma := \mu 2 - \mu 1$

local information matrix:

$$I := n\cdot\begin{bmatrix} (\lambda 1-1)\cdot s1 - \frac{\lambda 1+\lambda 2-1}{\sigma^2} & \frac{\lambda 1+\lambda 2-1}{\sigma^2} & \frac{\lambda 2}{(\lambda 1-1)\cdot\sigma} & -\frac{1}{\sigma} \\ \frac{\lambda 1+\lambda 2-1}{\sigma^2} & (\lambda 2-1)\cdot s2 - \frac{\lambda 1+\lambda 2-1}{\sigma^2} & \frac{1}{\sigma} & -\frac{\lambda 1}{(\lambda 2-1)\cdot\sigma} \\ \frac{\lambda 2}{(\lambda 1-1)\cdot\sigma} & \frac{1}{\sigma} & \psi'(\lambda 1)-\psi'(\lambda 1+\lambda 2) & -\psi'(\lambda 1+\lambda 2) \\ -\frac{1}{\sigma} & -\frac{\lambda 1}{(\lambda 2-1)\cdot\sigma} & -\psi'(\lambda 1+\lambda 2) & \psi'(\lambda 2)-\psi'(\lambda 1+\lambda 2) \end{bmatrix}$$

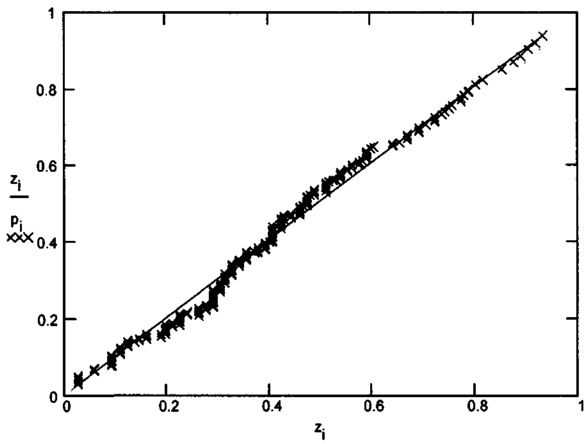
$$V := 1^{-1} \qquad V = \begin{pmatrix} 5.612 & -10.849 & -0.639 & -0.991 \\ -10.849 & 89.772 & 2.160 & 5.536 \\ -0.639 & 2.160 & 0.110 & 0.192 \\ -0.991 & 5.536 & 0.192 & 0.426 \end{pmatrix}$$

$$SE\mu 1 := \sqrt{V_{1,1}} \qquad SE\mu 1 = 2.369$$
$$SE\mu 2 := \sqrt{V_{2,2}} \qquad SE\mu 2 = 9.475$$
$$SE\lambda 1 := \sqrt{V_{3,3}} \qquad SE\lambda 1 = 0.332$$
$$SE\lambda 2 := \sqrt{V_{4,4}} \qquad SE\lambda 2 = 0.652$$

4. Linearized model/data plot

$$B := \frac{\Gamma(\lambda 1)\cdot\Gamma(\lambda 2)}{\Gamma(\lambda 1+\lambda 2)} \qquad \text{model ordinates: } z_i := \frac{x_i-\mu 1}{\mu 2-\mu 1}$$

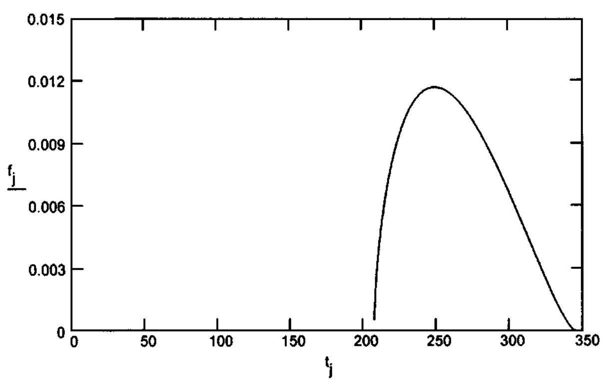
$$F(t) := \frac{1}{B}\cdot\int_0^t z^{\lambda 1-1}\cdot(1-z)^{\lambda 2-1}dz \qquad \text{data ordinates: } p_i := \frac{i-0.3}{n+0.4} - F(z_i) + z_i$$



The estimated model fits the data quite well.

5. pdf plot

$$j := 1..999 \qquad t_j := \mu 1 + \frac{\sigma\cdot j}{1000} \qquad f_j := \frac{1}{B\cdot\sigma}\cdot\left(\frac{j}{1000}\right)^{\lambda 1-1}\cdot\left(1-\frac{j}{1000}\right)^{\lambda 2-1}$$



EXAMPLE 14.5

The following data are available on the modulus of elasticity (1,000,000 psi) of a certain size, grade, and species of lumber:

1.73, 1.50, 1.56, 1.89, 1.54, 1.68, 1.39, 1.64, 1.49, 1.43, 1.68, 1.61, 1.62.

Assuming a four-parameter Beta distribution for this quantity, estimate the parameters.

1. Maximum likelihood approach

$n := 13 \qquad i := 1..n$

$y_i \qquad x := \mathrm{sort}(y) \qquad x_1 = 1.390 \qquad x_n = 1.890$

y_i
1.73
1.5
1.56
1.89
1.54
1.68
1.39
1.64
1.49
1.43
1.68
1.61
1.62

Display of likelihood function:

$j := 1..10 \qquad a_j := 1.369 + 0.002 \cdot j$

$k := 1..10 \qquad b_k := 1.88 + 0.012 \cdot k$

$$s5_j := \frac{1}{n} \cdot \sum_i (x_i - a_j)^{-1} \qquad s6_k := \frac{1}{n} \cdot \sum_i (b_k - x_i)^{-1}$$

$$s7_j := \sum_i \ln(x_i - a_j) \qquad s8_k := \sum_i \ln(b_k - x_i)$$

$$(14.54): \quad L1_{j,k} := \frac{s5_j \cdot [(b_k - a_j) \cdot s6_k - 1]}{s6_k \cdot [(b_k - a_j) \cdot s5_j - 1] - s5_j}$$

$$(14.55): \quad L2_{j,k} := \frac{s6_k \cdot [(b_k - a_j) \cdot s5_j - 1]}{s6_k \cdot [(b_k - a_j) \cdot s5_j - 1] - s5_j}$$

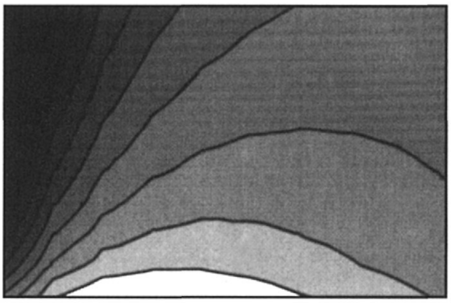
$$B_{j,k} := \frac{\Gamma(L1_{j,k}) \cdot \Gamma(L2_{j,k})}{\Gamma(L1_{j,k} + L2_{j,k})}$$

$$LLF_{j,k} := -n \cdot \ln(B_{j,k}) + n \cdot (1 - L1_{j,k} - L2_{j,k}) \cdot \ln(b_k - a_j) + [(L1_{j,k} - 1) \cdot s7_j + (L2_{j,k} - 1) \cdot s8_k] - 9.1$$

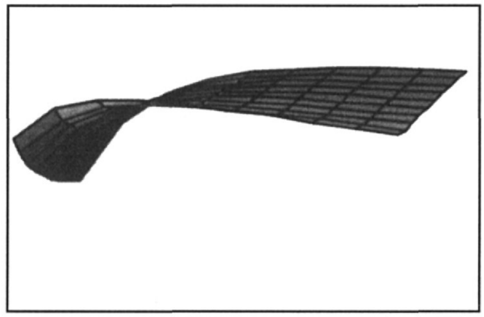
$b^T = 1.892 \quad 1.904 \quad 1.916 \quad 1.928 \quad 1.940 \quad 1.952 \quad 1.964 \quad 1.976 \quad 1.988 \quad 2.000)$

$$LLF = \begin{bmatrix} -1.193 & -0.735 & -0.499 & -0.377 & -0.309 & -0.271 & -0.250 & -0.240 & -0.235 & -0.236 \\ -1.128 & -0.684 & -0.458 & -0.342 & -0.280 & -0.247 & -0.229 & -0.221 & -0.220 & -0.222 \\ -1.058 & -0.629 & -0.414 & -0.306 & -0.250 & -0.221 & -0.207 & -0.203 & -0.204 & -0.209 \\ -0.981 & -0.569 & -0.367 & -0.267 & -0.218 & -0.194 & -0.184 & -0.184 & -0.188 & -0.196 \\ -0.897 & -0.505 & -0.315 & -0.226 & -0.184 & -0.166 & -0.161 & -0.164 & -0.172 & -0.183 \\ -0.803 & -0.433 & -0.259 & -0.181 & -0.147 & -0.136 & -0.137 & -0.144 & -0.156 & -0.171 \\ -0.697 & -0.353 & -0.197 & -0.132 & -0.107 & -0.104 & -0.111 & -0.124 & -0.140 & -0.159 \\ -0.573 & -0.261 & -0.126 & -0.075 & -0.063 & -0.068 & -0.082 & -0.101 & -0.123 & -0.146 \\ -0.425 & -0.150 & -0.040 & -0.008 & -0.008 & -0.024 & -0.047 & -0.073 & -0.101 & -0.130 \\ -0.232 & -0.001 & -0.079 & -0.090 & -0.074 & -0.046 & -0.013 & -0.022 & -0.057 & -0.092 \end{bmatrix} \qquad a = \begin{bmatrix} 1.371 \\ 1.373 \\ 1.375 \\ 1.377 \\ 1.379 \\ 1.381 \\ 1.383 \\ 1.385 \\ 1.387 \\ 1.389 \end{bmatrix}$$

$M_{j,11-k} := LLF_{k,j}$ Conclusion: It appears that the ML estimate of μ_1 converges to $x_1 = 1.39$.



M



LLF

2. Moment estimates

$r := 1..4 \qquad m_r := \frac{1}{n} \cdot \sum_i (x_i)^r \qquad M2 := m_2 - (m_1)^2$

$M3 := m_3 - 3 \cdot m_1 \cdot m_2 + 2 \cdot (m_1)^3$

$M4 := m_4 - 4 \cdot m_1 \cdot m_3 + 6 \cdot (m_1)^2 \cdot m_2 - 3 \cdot (m_1)^4$

Shape parameters:

$a := 1 \quad b := 1 \quad \text{GIVEN} \quad \frac{M3}{M2^{1.5}} = 2 \cdot \frac{b-a}{a+b+2} \cdot \sqrt{\frac{a+b+1}{a \cdot b}}$

$$\frac{M4}{M2^2} = \frac{3 \cdot (a+b+1) \cdot [2 \cdot (a+b)^2 + a \cdot b \cdot (a+b-6)]}{a \cdot b \cdot (a+b+2) \cdot (a+b+3)}$$

$\begin{pmatrix} \lambda 1 \\ \lambda 2 \end{pmatrix} := \text{FIND}(a, b) \qquad \lambda 1 = 4.088$

$\lambda 2 = 10.417$

Location parameters:

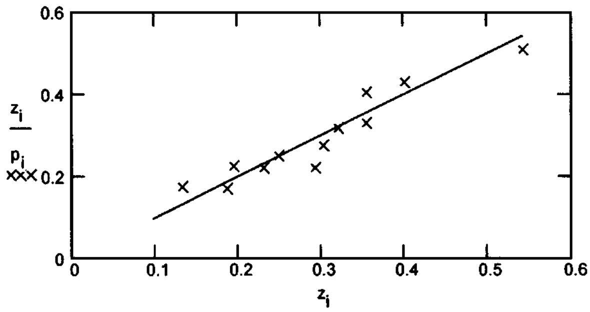
$\mu 1 := m_1 - \sqrt{M2} \cdot \sqrt{\frac{\lambda 1 \cdot (\lambda 1 + \lambda 2 + 1)}{\lambda 2}} \qquad \mu 1 = 1.279$

$\mu 2 := m_1 + \sqrt{M2} \cdot \sqrt{\frac{\lambda 2 \cdot (\lambda 1 + \lambda 2 + 1)}{\lambda 1}} \qquad \mu 2 = 2.407$

3. Linearized model/data plot

$B := \frac{\Gamma(\lambda 1) \cdot \Gamma(\lambda 2)}{\Gamma(\lambda 1 + \lambda 2)}$ model ordinates: $z_i := \frac{x_i - \mu 1}{\mu 2 - \mu 1}$

$F(t) := \frac{1}{B} \cdot \int_0^t z^{\lambda 1 - 1} \cdot (1-z)^{\lambda 2 - 1} dz$ data ordinates: $p_i := \frac{i - 0.3}{n + 0.4} - F(z_i) + z_i$

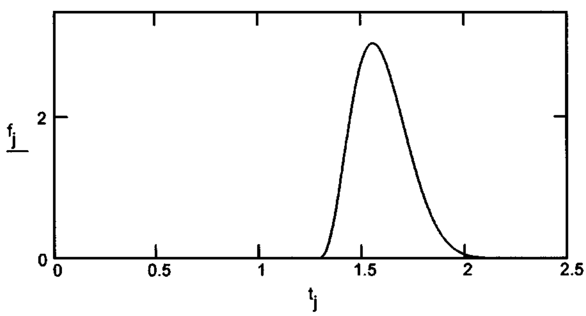


The estimated model fits the data reasonably well.

4. pdf plot

$$j := 1..199 \qquad t_j := \mu1 + \frac{(\mu2 - \mu1) \cdot j}{200}$$

$$f_j := \frac{1}{B \cdot (\mu2 - \mu1)} \cdot \left(\frac{j}{200}\right)^{\lambda1-1} \cdot \left(1 - \frac{j}{200}\right)^{\lambda2-1}$$



EXAMPLE 14.6

A sample of 16 observations was obtained on the time it takes to assemble a certain product:

27.0, 28.7, 29.2, 28.6, 30.8, 27.5, 30.1, 31.2,
29.8, 28.3, 27.3, 29.1, 27.9, 26.5, 30.0, 31.4.

Assembly time is thought to be limited to the range (25,32). Estimate the Beta parameters and their standard errors. Estimate the coefficient of variation and its standard error.

$n := 16 \qquad i := 1..n$

$y_i :=$ $\quad x := \text{sort}(y) \quad z_i := \dfrac{x_i - 25}{32 - 25}$

27.0
28.7
29.2
28.6
30.8
27.5
30.1
31.2
29.8
28.3
27.3
29.1
27.9
26.5
30.0
31.4

1. Moment estimates

$$m1 := \frac{1}{n} \cdot \sum_i z_i \qquad m2 := \frac{1}{n} \cdot \sum_i (z_i)^2$$

$$L1 := \frac{m1^2 - m1 \cdot m2}{m2 - m1^2} \qquad L1 = 2.635$$

$$L2 := \frac{m1 - m2}{m2 - m1^2} - L1 \qquad L2 = 2.020$$

2. Display the log-likelihood function

$$s3 := \sum_i \ln(z_i) \qquad s4 := \sum_i \ln(1 - z_i)$$

$$k := 1..10 \qquad p := 1..10$$

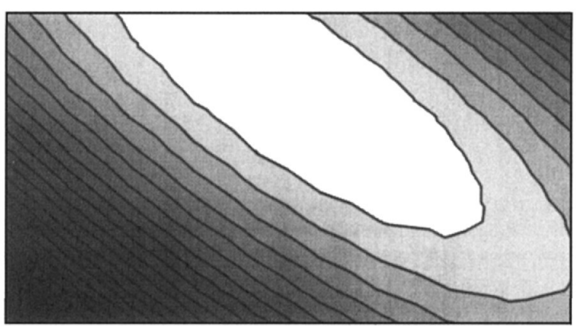
$$r_k := 2.6 + 0.05 \cdot k \qquad l_p := 1.8 + 0.05 \cdot p \qquad B_{k,p} := \frac{\Gamma(r_k) \cdot \Gamma(l_p)}{\Gamma(r_k + l_p)}$$

$$L_{k,p} := -n \cdot \ln(B_{k,p}) + (r_k - 1) \cdot s3 + (l_p - 1) \cdot s4 - 3$$

$$l^T = (1.850 \quad 1.900 \quad 1.950 \quad 2.000 \quad 2.050 \quad 2.100 \quad 2.150 \quad 2.200 \quad 2.250 \quad 2.300)$$

$$L = \begin{bmatrix} 0.183 & 0.230 & 0.259 & 0.271 & 0.267 & 0.247 & 0.213 & 0.164 & 0.101 & 0.025 \\ 0.153 & 0.210 & 0.249 & 0.270 & 0.276 & 0.265 & 0.240 & 0.200 & 0.147 & 0.080 \\ 0.115 & 0.182 & 0.230 & 0.261 & 0.276 & 0.275 & 0.259 & 0.228 & 0.184 & 0.126 \\ 0.070 & 0.146 & 0.204 & 0.244 & 0.268 & 0.276 & 0.269 & 0.248 & 0.212 & 0.162 \\ 0.017 & 0.102 & 0.169 & 0.219 & 0.252 & 0.270 & 0.271 & 0.259 & 0.232 & 0.191 \\ -0.044 & 0.051 & 0.128 & 0.187 & 0.229 & 0.255 & 0.266 & 0.262 & 0.243 & 0.211 \\ -0.111 & -0.007 & 0.079 & 0.147 & 0.198 & 0.233 & 0.252 & 0.257 & 0.247 & 0.223 \\ -0.185 & -0.072 & 0.022 & 0.099 & 0.159 & 0.203 & 0.231 & 0.244 & 0.243 & 0.228 \\ -0.266 & -0.144 & -0.041 & 0.045 & 0.114 & 0.166 & 0.203 & 0.225 & 0.232 & 0.225 \\ -0.354 & -0.223 & -0.110 & -0.016 & 0.062 & 0.123 & 0.168 & 0.198 & 0.213 & 0.215 \end{bmatrix} \qquad r = \begin{bmatrix} 2.650 \\ 2.700 \\ 2.750 \\ 2.800 \\ 2.850 \\ 2.900 \\ 2.950 \\ 3.000 \\ 3.050 \\ 3.100 \end{bmatrix}$$

$$M_{p,11-k} := L_{k,p}$$



M

starting values: $\quad r_3 = 2.750 \qquad l_5 = 2.050$

3. Maximum likelihood estimates

$L1 := 2.75 \qquad L2 := 2.05$

digamma function: $\psi(L) := \dfrac{d}{dL}\ln(\Gamma(L))$

GIVEN

$\psi(L1) - \psi(L1 + L2) = \dfrac{s3}{n}$

$\psi(L2) - \psi(L1 + L2) = \dfrac{s4}{n}$

$\begin{pmatrix} \lambda 1 \\ \lambda 2 \end{pmatrix} := \text{FIND}(L1, L2)$ ML parameter estimates: $\lambda 1 = 2.754$

$\lambda 2 = 2.074$

4. Standard errors

trigamma function: $\psi'(L) := \dfrac{d}{dL}\psi(L)$

$a := \psi'(\lambda 1) \quad b := \psi'(\lambda 2) \quad c := \psi'(\lambda 1 + \lambda 2) \quad d := n \cdot (a \cdot b - c \cdot (a + b))$

$\text{SE}\lambda 1 := \sqrt{\dfrac{b - c}{d}} \quad \text{SE}\lambda 2 := \sqrt{\dfrac{a - c}{d}} \quad \text{SE}\lambda 12 := \sqrt{\dfrac{c}{d}}$

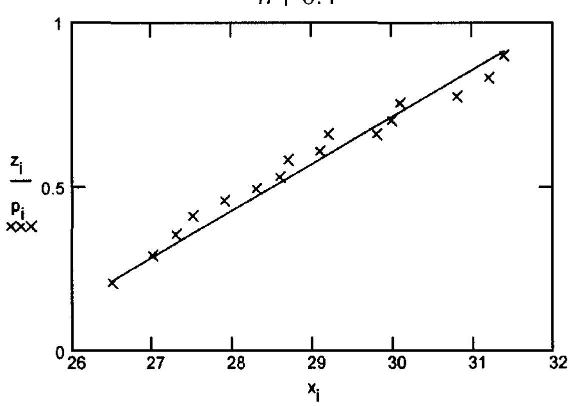
$\text{SE}\lambda 1 = 0.946$
$\text{SE}\lambda 2 = 0.692$
$\text{SE}\lambda 12 = 0.730$

5. Linearized model/data plot

$F(x, \lambda 1, \lambda 2) := \dfrac{\Gamma(\lambda 1 + \lambda 2)}{\Gamma(\lambda 1) \cdot \Gamma(\lambda 2)} \cdot \int_0^x z^{\lambda 1 - 1} \cdot (1 - z)^{\lambda 2 - 1} dz$

model ordinates: z_i

data ordinates: $p_i := \dfrac{i - 0.3}{n + 0.4} - F(z_i, \lambda 1, \lambda 2) + z_i$

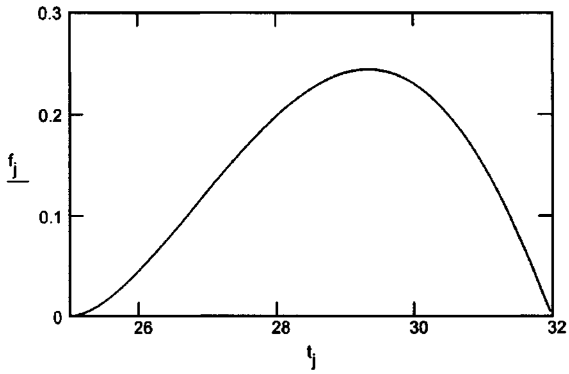


The estimated model fits the data quite well.

6. Density plot

$$j := 1..199 \quad t_j := 25 + \frac{7 \cdot j}{200}$$

$$f_j := \frac{\Gamma(\lambda 1 + \lambda 2)}{\Gamma(\lambda 1) \cdot \Gamma(\lambda 2) \cdot 7} \cdot \left(\frac{j}{200}\right)^{\lambda 1 - 1} \cdot \left(1 - \frac{j}{200}\right)^{\lambda 2 - 1}$$



7. Coefficient of variation

$$(14.21): \quad \mathrm{cv}(\lambda 1, \lambda 2) := \frac{7 \cdot \sqrt{\lambda 1 \cdot \lambda 2}}{\sqrt{\lambda 1 + \lambda 2 + 1} \cdot (25 \cdot \lambda 2 + 32 \cdot \lambda 1)} \qquad \mathrm{cv}(\lambda 1, \lambda 2) = 0.050$$

$$\text{standard error:} \quad d1 := \frac{d}{d\lambda 1}\mathrm{cv}(\lambda 1, \lambda 2) \qquad d2 := \frac{d}{d\lambda 2}\mathrm{cv}(\lambda 1, \lambda 2)$$

$$\mathrm{SE}x := \sqrt{(d1 \cdot \mathrm{SE}\lambda 1)^2 + (d2 \cdot \mathrm{SE}\lambda 2)^2 + 2 \cdot d1 \cdot d2 \cdot \mathrm{SE}\lambda 12^2} \qquad \mathrm{SE}x = 0.007$$