Three-parameter Gamma MLE

The three-parameter Gamma distribution, also referred to as a Pearson Type III distribution, is often used in hydrology to model the log of flow data for river systems. The likelihood function for this distribution can be written as:

\[ L(\alpha, \beta, \gamma | y) = \sum_{i=1}^{n} \left[ (\alpha - 1) \ln(y_i - \gamma) - \frac{y_i - \gamma}{\beta} - \alpha \ln \beta - \ln \Gamma(\alpha) \right], \]

where \( \alpha \) is the shape parameter, \( \beta \) is the scale parameter, and \( \gamma \) is the shift or location parameter. The gradient for this log likelihood function has components:

\[ \frac{\partial L}{\partial \alpha} = \sum_{i=1}^{n} \ln(y_i - \gamma) - n \ln(\beta) - n \psi(\alpha) \]
\[ \frac{\partial L}{\partial \beta} = \sum_{i=1}^{n} \frac{y_i - \gamma}{\beta^2} - \frac{n \alpha}{\beta} \]
\[ \frac{\partial L}{\partial \gamma} = - \sum_{i=1}^{n} \frac{\alpha - 1}{y_i - \gamma} + \frac{n}{\beta}, \]

where \( \psi(\cdot) \) is the “psi” function, i.e., the derivative of \( \ln(\Gamma(\alpha)) \) with respect to \( \alpha \).

If we set the gradient to 0, the resulting likelihood equations can be intractable. Note that for \( \hat{\alpha} \) less than 1, some of the \( y_i \) must be less than \( \hat{\gamma} \), a logical inconsistency. In this case, the MLE does not exist. Many alternate approaches have been suggested to resolve this problem with the MLE, but we will proceed as though a solution exists.

From the gradient, we can compute the Hessian:

\[ H(\alpha, \beta, \gamma) = \begin{bmatrix}
- n \psi'(\alpha) & - \sum_{i=1}^{n} \frac{y_i - \gamma}{\beta^2} & - \sum_{i=1}^{n} \frac{1}{y_i - \gamma} \\
- \sum_{i=1}^{n} \frac{1}{y_i - \gamma} & \sum_{i=1}^{n} \frac{y_i - \gamma}{\beta^3} + \frac{n \alpha}{\beta^2} & - \sum_{i=1}^{n} \frac{n}{\beta^2} \left( \frac{\alpha - 1}{(y_i - \gamma)^2} \right) \\
- \sum_{i=1}^{n} \frac{1}{y_i - \gamma} & - \sum_{i=1}^{n} \frac{n}{\beta^2} \left( \frac{\alpha - 1}{(y_i - \gamma)^2} \right) & \sum_{i=1}^{n} \frac{1}{\beta^2} \left( \frac{1}{(y_i - \gamma)^2} \right)
\end{bmatrix} \]

If you work through the math, the information matrix actually has the rather straightforward form:

\[ I(\alpha, \beta, \gamma) = n \begin{bmatrix}
\psi'(\alpha) & \frac{1}{\beta} & \frac{1}{(\alpha - 1) \beta^2} \\
\frac{1}{\beta^2} & \frac{1}{\beta^2} & \frac{1}{(\alpha - 1) \beta^2} \\
\frac{1}{(\alpha - 1) \beta} & \frac{1}{(\alpha - 1) \beta} & \frac{1}{(\alpha - 1) \beta^2}
\end{bmatrix} \]

This information could then be supplied to \textbf{R} to find MLE’s and their standard errors.