univariateML: An R package for maximum likelihood estimation of univariate densities

Jonas Moss

1 University of Oslo

Summary

univariateML is an R (R Core Team, 2019) package for user-friendly univariate maximum likelihood estimation (Cam, 1990). It supports more than 20 densities, the most popular generic functions such as plot, AIC, and confint, and a simple parametric bootstrap (Efron & Tibshirani, 1994) interface.

When looking at univariate data it is natural to ask if there is a known parametric density that fits the data well. The following example uses the egypt (Pearson, 1902) data set included in the package and a plot of the Weibull and Gamma densities (Johnson, Kotz, & Balakrishnan, 1995, Chapter 17 & 21).

```r
# install.packages("univariateML")
library("univariateML")
hist(egypt$age, freq = FALSE, main = "Mortality", xlab = "Mortality")
lines(mlweibull(egypt$age))  # Plots a Weibull fit.
lines(mlgamma(egypt$age), col = "red")  # Plots a Gamma fit.
```
A natural question to ask is which among several models fits the data best. This can be done using tools of model selection such as the AIC (Akaike, 1998).

\[
\text{AIC(mlweibull(egypt$age),} \\
\text{mlgamma(egypt$age))}
\]

```
## df  AIC
## mlweibull(egypt$age) 2 1230.229
## mlgamma(egypt$age) 2 1234.772
```

Problems involving estimation of univariate densities are common in statistics. Estimation of univariate densities is used in for instance exploratory data analysis, in the estimation of copulas (Ko, Hjort, & Hobæk Haff, 2019), as parametric starts in density estimation (Hjort & Glad, 1995; Moss & Tveten, 2019), and is of interest in and of itself.

Analytic formulas for the maximum likelihood estimates are used whenever they exist. Most estimators without analytic solutions have a custom made Newton-Raphson solver. This is in contrast to the \texttt{mle} function in the built-in \texttt{R} package \texttt{stats4}, which supports more general maximum likelihood estimation through numerical optimization on a supplied negative log-likelihood function.

\texttt{Rfast} (Papadakis et al., 2019) is an \texttt{R} package with fast Newton-Raphson implementations of many univariate density estimators. \texttt{univariateML} differs from \texttt{Rfast} mainly in focus: While \texttt{univariateML} is focused on user-friendly univariate density estimation, \texttt{Rfast} aims to have the fastest possible implementations of many kinds of functions.

**References**


