On the Tukey Distribution and the Man Who Proposed it.

Here is a link for a R package that provides quintiles, probabilities and density, and random numbers for the generalized lambda distribution, also called the generalised Tukey distribution.
http://www.cran.r-project.org/src/contrib/PACKAGES.html#gld

The generalised lambda distribution is interesting because:
1. It is extremely versatile, being able to assume a wide variety of shapes. The list of distributions it can approximate includes the Gaussian, the Uniform, Beta, the Students-t, the Exponential, Weibull, Gamma, Lognormal, Pareto, F, Chi-square, Logistic, Double Exponential and Extreme Value distributions

2. The generalised Lambda is defined in terms of its quintile function or inverse distribution function, as opposed to the density function or the distribution function. So wherever we are interested in dealing with the percentile, (e.g. risk management) the generalised lambda distribution is very convenient.

The origins of this distribution lie in the one-parameter Lambda distribution proposed by John Tukey in 1960 which was generalized to a four parameter distribution by other researchers.

Here is a website with java applets that shows the shapes that this distribution can assume: http://www.ens.gu.edu.au/robertk/gld/index.html

One good book on the subject is “Fitting Statistical Distributions: The Generalized Lambda Distribution and Generalized Bootstrap Methods” by Zaven A. Karian, Edward J. Dudewicz.

When of the most versatile statisticians of the twentieth century was John Tukey. He made seminal contributions to exploratory data analysis, spatio-temporal clustering, and occurrences of extreme values and examination of patterns within data. Here is a link on the Bell Labs website that describes the life and work of this most versatile of modern statisticians. http://stat.bell-labs.com/who/tukey/

I will quote from a book that contains an essay on his life.

Excerpts from the essay on the life and work of John Tukey

From chapter 22 titled "The Picasso of Statistics" of the book The Lady Tasting Tea by David Salsburg

Pages 231 to 232

"Pablo Picasso astonished the world of art in the early years of the twentieth century with his protean output. For a while, he played with monochromatic paintings, then he invented cubism, then he examined classicism, then he went on to ceramics. Each of these excursions resulted in a revolutionary change in art, which others continued to exploit after Picasso went on to other things. So it has been with John Tukey. In the 1950's he became involved with Andrei Kolmogorov's ideas on stochastic processes and invented a technique for analysing long strings of correlated results, known as the Fast Fourier Transforms.

Like Picasso and Cubism, Tukey could have accomplished nothing more and his influence on science would have been immense.... the FFT is a smoothing technique, that to use Tukey's expression, borrows strength from neighbouring frequencies so that vast amounts of data are not needed to get good estimates...."

From page 233

"In the 1960s and the early 1970s engineers and statisticians at Bell telephone Laboratories pioneered the analysis of huge amounts of data. Monitoring telephone lines for random errors and problems led to millions of data items in a single computer file. How can you look at so much data? How can you structure it so that it can be examined... Can we find methods for examining these vast collections of numbers and learn something from them without imposing assumptions about their distributions?

Although a great deal of scientific research involves collecting data and fitting those data to some pre-conceived mould of some specific distribution, it is often useful and important just to collect data and examine them for unexpected events.

As Eric Temple Bell, the American mathematician and historian of science pointed out, "Numbers do not lie, but they have the propensity to tell the truth with the he intent to deceive." The human being is prone to seeing patterns and will often see patterns where only random noise exists...[in epidemiology] An examination of data often finds a "cluster" of diseases in a certain place or time. Suppose we find an unusually high number of children with leukaemia in a small town in Massachusetts. Does this mean there is some cause of cancer operating in this town? Or is this just a random cluster, which happened to appear here and could as easily have appeared elsewhere... In a more general sense, to what extent can we examine data with an eye towards patterns and expect to find anything more than a random will-o'- the-wisp? In the 1960's John Tukey gave serious considerations to these problems."

From page 235:
"...It was Tukey who coined the words bit (for binary digit) and software. ...Nothing was too mundane for Tukey to attack with original insight and nothing is too sacrosanct for him to question. Take the simple process of Tallying. Most readers have been exposed to the use of the tall figure when counting something. The usual one, presented to us by generations of teachers, is to make four vertical marks and a fifth one crossing out the

four.... It is a foolish way totally, said John Tukey. Consider how easy it is to make a mistake. You might put the cross over three instead of four lines, or you might put down five lines and then a cross. The incorrect tally is hard to spot. Tukey proposed a ten-mark tally. You mark four dots to make the corners of a box. You then connect the four dots with four lines completing the box, and finally you make two diagonals marking a cross within the box."