Re: Goodness of fit measures for a distribution

Source: http://sci.tech−archive.net/Archive/sci.stat.math/2005−04/msg00220.html

- *From*: Aleks Jakulin <a_jakulin..@..hotmail..com>
- *Date*: Tue, 12 Apr 2005 23:49:07 +0200

Reef Fish wrote:

> Chi−square is based on histograms -- it's worthless.
>
> Kolmogorov uses on ONE POINT in the difference between the empirical and theoretical cdfs, the point of maximum departure.
>
> Your EYEBALLS can do an infinitely better job than that, looking at the plot of the entire cdfs.

In principle yes. But what do you do when you have a few dozen or hundred variables in a complex data mining task? Then you do want a one−number summary: you have no time to manually inspect a few hundred thousand QQ plots.

You can use the p−value or some VaR quantification as a summary of the tail, use it order the fits, and let the eyes inspect the deviant ones first.

--
mag. Aleks Jakulin
http://kt.ijs.si/aleks/
Department of Knowledge Technologies,
Jozef Stefan Institute, Ljubljana, Slovenia.
.