

Biometrics Section

[Post New Message](#)

BIOMETRICS SECTION NEWSLETTER, October 2017

[Reply to Group](#)

[Reply to Sender
via Email](#)



Oct 3, 2017 11:28 AM

[Zheyu Wang](#)

BIOMETRICS SECTION NEWSLETTER, October 2017

Edited by Zheyu Wang, Biometrics Section Publications Officer

Graph of the Month

This month's graph is contributed by section members David Shera and Christopher Tong: They introduce readers to a modern implementation of Parzen's quantile plot, also known as a "boxdot plot". When displaying individual data symbols overlaid on a boxplot, it is common to add horizontal random "jittering" to the data symbols to separate them. Jittering has some disadvantages, and there has been interest in more systematic ways of displaying the data symbols, such as Aron Eklund's beeswarm R package. Another alternative is to display the data symbols sorted horizontally, in the form of a sideways empirical distribution function (EDF), or an inverse CDF. Such a display provides a stronger guarantee that data symbols will be separated, and their pattern can be interpreted in the same way that an EDF can. An example comparing jittering with a boxdot plot is shown below.

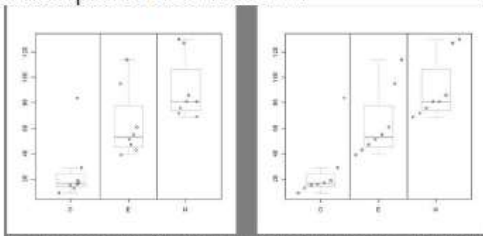


Figure: Comparison of subset of OrchardSprays data using boxplots with jittered data symbols (left) and quantile boxplots (right).

A simple R function for drawing boxdot plots may be found on GitHub:
github.com/hydrodynamicstability/boxdot.plot

References:

1. Parzen, *Journal of the American Statistical Association*, 74: 105-121 (1978).
2. M. Shera, *Computing Science and Statistics*, 23: 50-53 (1991).

If you have created or encountered an interesting data visualization you would like to share with the Biometrics section, e-mail Zheyu Wang (wangzy@jhu.edu) or Rebecca Hubbard (rhubb@upenn.edu).