Revista Colombiana de Estadística
ISSN: 0120-1751
revcoles_fcbog@unal.edu.co
Universidad Nacional de Colombia
Colombia

Marmolejo-Ramos, Fernando
Current Topics in Statistical Graphics
Revista Colombiana de Estadística, vol. 37, núm. 2, diciembre, 2014
Universidad Nacional de Colombia
Bogotá, Colombia

Revista Colombiana de Estadística
Current Topics in Statistical Graphics
Diciembre 2014, volumen 37, no. 2
DOI: http://dx.doi.org/10.15446/rce.v37n2spe.48058

Editorial

Current Topics in Statistical Graphics

Fernando Marmolejo-Ramos[a]

Department of Psychology, Stockholm University, Sweden

Traditionally, statistical graphics have been associated with the representation and exploration of data. For example, in meteorology, wind boxplots (also known as wind roses) are used to depict distributions of wind speed and directional data (e.g. see Figure 2 in VanDerWal, Murphy, Kutt, Perkins, Bateman, Perry & Reside, 2013), and, in computational biology, dendrograms are used to display arrangements of clusters found by hierarchical clustering methods (e.g. see Figure 2 in Garcia, Pinho, Rodrigues, Bastos & Ferreira, 2011). As with statistical tests, statistical graphics are ubiquitous across many scientific fields. For example, just as ANOVA is used in different research areas, boxplots are also found in different research areas (e.g. boxplots used in neuroscience; Ghosh, Kakunoori, Augustinack, Nieto-Castanon, Kovelman, Gaab, Christodoulou, Triantafyllou, Gabrieli & Fischl, 2010, Figure 4; and boxplots used in biomedical engineering; Kreja, Liedert, Schlenker, Brenner, Fiedler, Friemert & Dürselen, 2012, Figure 2). However, statistical graphics go beyond the selection and representation of data, they are also used to represent statistical concepts and models (indeed, the colours used in visual displays do have an influence on the interpretation of statistical models; e.g. Zeileis, Hornik & Murrell, 2009). Statistical graphics can also be used to test assumptions, select estimators, and detect outliers.

Thus, statistical graphics are tools capable of exploring the content of data, finding structure in data, checking assumptions in statistical models, representing statistical concepts, and communicating the results of analyses. Figure 1 illustrates the capabilities of statistical graphics. Figure 1A shows how statistical graphics assist in exploring the content and structure of data: ECDFs representing three data sets of equal sample size but with discrepant distributions (source: The author). Figure 1B exemplifies the role of statistical graphics in checking statistical models: A circular boxplot representing circular data (L_F = lower fence, Q_1 = first quantile, ϕ = median direction, Q_3 = third quantile, and U_F = upper fence)

[a]Invited Editor.
E-mail: fernando.marmolejo.ramos@psychology.su.se

(source: Figure 1 in Abuzaid, Mohamed & Hussin, 2012). Figure 1C illustrates how statistical graphics aid in representing statistical concepts: the effects of the replacement of missing values (N) on the mean (squares) and dispersion (solid circles) of conditions (C) around the grand mean (crosses) (source: Figure 4 in Lachaud & Renaud, 2011. It is worth mentioning that these researchers borrowed visual concepts from physics, specifically centre of gravity and area, to represent the effects of replacing missing values on experimental conditions means and variances). Finally, Figure 1D exemplifies how statistical graphics are valuable in communicating the results of analyses: Comparison of the performance of two robust estimators of central tendency (γ) on a Gamma distribution (with combinations of different values in the parameters α and β) of sample size five (source: Figure 2 in Vélez & Correa, 2014).

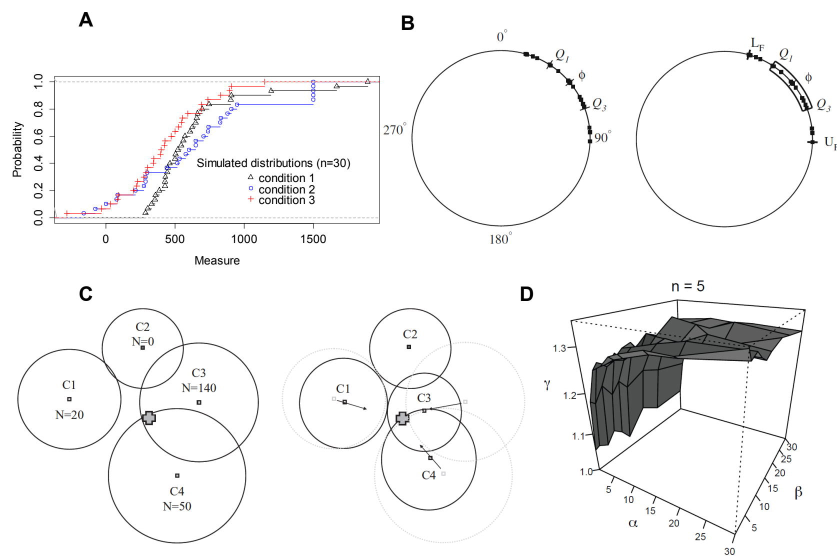

Figure 1: Some examples of the capabilities of statistical graphics.

Given new developments in methods for the visualisation of quantitative and categorical data, it is important to become aware of the latest advances in order to probe potential cross-field implementations or further the capabilities of the graphical methods. This special issue has the purpose of bringing together current advances and uses of well-known and novel statistical graphic methods from different research areas to enable the reader to find potential applications to his/her own research field.

An overview of the history of dynamic-interactive statistical graphics is presented by *Valero-Mora and Ledesma*. *Friendly and Sigal* present a review of the methods used in the visualisation of multivariate linear models and *Makela, Si and Gelman* report methods for representing sample survey data. Two contributions relate to the case of psychological data. *Filiz, Trumpower and Vanapalli* present a method to visualise concept maps and that is dependent on a specific type of algorithm and *Campitelli and Macbeth* propose ideas that help graphing Bayesian models. *Castro-Kuriss, Leiva and Athayde*; *Nieto, Galindo, Leiva and Vicente-Galindo*; and *Ospina, Larangeiras and Frery*, respectively, propose new graphical tests based on Goodness-of-Fit techniques, inferential biplots based on bootstrap confidence intervals, and a combination of boxplots that enhances the visualisation of skewed data. Two contributions focus on data emerging in socio-economic and one on geographic research. *Arcagni and Porro* study the graphical representation of inequality and *Fattore, Arcagni and Barberis* study the visualisation of partial order sets (or posets). *Symanzik, Dai, Weber, Payton and McManus* study graphical methods for displaying statistical summaries associated with regional spatial units (these graphical methods are known as Linked Micromaps). These authors use South American geographic data to illustrate these methods. In the last contribution reporting a study that relies on accelerometer data, *Teknomo and Estuar* demonstrate how graphical methods assist in the analysis of gait patterns.

Since statistical graphics have a decisive role in making the most of data (see Marmolejo-Ramos & Tian, 2010; Marmolejo-Ramos & Matsunaga, 2009), they often constitute a necessary tool in the proper conduct of research (Friendly, 2008; Wainer & Velleman, 2001). It is expected the topics considered in this special issue will serve to motivate statisticians to develop new advances in statistical graphics and help researchers to access the latest developments in this fascinating arena.

Acknowledgments

The author thanks Professor John Dunn for proofreading this manuscript.

References

Abuzaid, A. H., Mohamed, I. B. & Hussin, A. G. (2012), 'Boxplot for circular variables', *Computational Statistics* **27**, 381–392.

Friendly, M. (2008), 'The golden age of statistical graphics', *Statistical Science* **23**(4), 502–535.

Garcia, S. P., Pinho, A. J., Rodrigues, J. M. O. S., Bastos, C. A. C. & Ferreira, P. J. S. G. (2011), 'Minimal absent words in prokaryotic and eukaryotic genomes', *PLoS ONE* **6**(1). e16065. doi:10.1371/journal.pone.0016065.

Ghosh, S. S., Kakunoori, S., Augustinack, J., Nieto-Castanon, A., Kovelman, I., Gaab, N., Christodoulou, J. A., Triantafyllou, C., Gabrieli, J. D. E. & Fischl, B. (2010), 'Evaluating the validity of volume-based and surface-based brain image registration for developmental cognitive neuroscience studies in children 4 to 11 years of age', *NeuroImage* **6**(1), 85–93.

Kreja, L., Liedert, A., Schlenker, H., Brenner, R. E., Fiedler, J., Friemert, B. & Dürselen, L. (2012), 'Anita ignatius effects of mechanical strain on human mesenchymal stem cells and ligament fibroblasts in a textured poly(l-lactide) scaffold for ligament tissue engineering', *Journal of Materials Science: Materials in Medicine* **23**(10), 2575–2582.

Lachaud, C. M. & Renaud, O. (2011), 'A tutorial for analyzing human reaction times: How to filter data, manage missing values, and choose a statistical model', *Applied Psycholinguistics* **32**(2), 389–416.

Marmolejo-Ramos, F. & Matsunaga, M. (2009), 'Getting the most from your curves: Exploring and reporting data using informative graphical techniques', *Tutorials in Quantitative Methods for Psychology* **5**(2), 40–50.

Marmolejo-Ramos, F. & Tian, S. (2010), 'The shifting boxplot. A boxplot based on essential summary statistics around the mean', *International Journal of Psychological Research* **3**(1), 37–46.

VanDerWal, J., Murphy, H. T., Kutt, A. S., Perkins, G. C., Bateman, B. L., Perry, J. J. & Reside, A. E. (2013), 'Focus on poleward shifts in species distribution underestimates the fingerprint of climate change', *Nature Climate Change* **3**, 239–243.

Vélez, J. I. & Correa, J. C. (2014), 'Should we think of a different median estimator?', *Comunicaciones en Estadística* **7**(1), 11–17.

Wainer, H. & Velleman, P. F. (2001), 'Statistical graphics: Mapping the pathways of science', *Annual Review of Psychology* **52**, 305–335.

Zeileis, A., Hornik, K. & Murrell, P. (2009), 'Escaping rgbland: Selecting colors for statistical graphics', *Computational Statistics and Data Analysis* **53**(9), 3259–3270.