Neuroscience and Biobehavioral Reviews 68 (2016) 1004–1005

Contents lists available at ScienceDirect

Neuroscience and Biobehavioral Reviews

journal homepage: www.elsevier.com/locate/neubiorev

Discussion

The anova to mixed model transition

Matthieu P. Boisgontier[a,*], Boris Cheval[b]

[a] Movement Control and Neuroplasticity Research Group, Department of Kinesiology, Biomedical Sciences Group, KU Leuven, Leuven, Belgium
[b] Methodology and Data Analysis Research Group, Faculty of Psychology and Educational Sciences (FAPSE), Geneva University, Geneva, Switzerland

ARTICLE INFO

Article history:
Received 20 March 2016
Received in revised form 13 May 2016
Accepted 27 May 2016
Available online 27 May 2016

Keywords:
Analysis of variance
Linear mixed-effects models
Neuroscience
Statistics

ABSTRACT

A transition towards mixed models is underway in science. This transition started up because the requirements for using analyses of variances are often not met and mixed models clearly provide a better framework. Neuroscientists have been slower than others in changing their statistical habits and are now urged to act.

Recently, the reliability of neuroscience research has been seriously questioned (Nieuwenhuis et al., 2011; Button et al., 2013). To limit the accumulation of such warnings, improving statistical methods and changing habits is urgently needed. One of the most widespread and deeply entrenched statistical habits neuroscientists have is the use of the analysis of variance (ANOVA). In Google Scholar, the association of the keywords "neuroscience" and "ANOVA" reaches 28,600 hits in 2015; representing 36% of the "neuroscience" hits.

ANOVA is a useful statistical model simultaneously testing between-mean differences in more than two conditions. The conditions define the different modalities of a given factor or explanatory variable (e.g., factor = age; conditions = young adults, older adults). One of the assumptions in ANOVA is the independence of observations (Aarts et al., 2014). However, in neuroscience, multiple observations are often collected from single subjects/animals. Such observations cannot be considered independent and should therefore be analyzed using models controlling for variability within units of observations. Ignoring this nested structure of the data yields incorrect p-values, usually associated with an increase of false positives (type I error).

Repeated measures ANOVA is a statistical model that allows for both between-subject factors (when subjects are measured only in one of the conditions of the factor) and within-subject factors (when subjects are measured in each condition of the factor). This model can correctly account for non-independence in observations within subjects/animals, but only if each observation is performed in a different condition from an exhaustive set of predefined (i.e., fixed) conditions, not at several occasions for the same condition.

Moreover, in neuroscience, subjects/animals are often observed in a limited number of conditions that only represent a random sample from the whole population of possible conditions, just as subjects/animals only represent a random sample from the whole population of possible subjects/animals. Ignoring this randomness of conditions, i.e., treating both subjects/animals and conditions as fixed effects, but only treating subjects/animals as a random effect, is also likely to result in an increase of type I errors. Both ANOVA and repeated measures ANOVA disregard the sampling variability of conditions.

Linear mixed models (LMM) have been developed to take into account both the nested (multiple observations within a single subject/animal in a given condition) and crossed (subjects/animals observed in multiple conditions) structure of the data, thereby providing results with acceptable type I error rates, i.e., acceptable reliability (Baayen et al., 2008). Moreover, treating both subjects/animals and conditions as random effects allows generalizing the results to the population of subjects/animals, but also to the population of conditions (Barr et al., 2013). Finally, LMM allow incomplete and unbalanced data to be used, continuous and categorical predictors to be combined, and information loss due to averaging over observations or participants to be avoided (Judd et al., 2012). Interestingly, Generalized LMM, an extension of LMM, allow not only continuous, but also binary and count-dependent variables to be analyzed (McCulloch et al., 2008).

The aforementioned advantages of LMM over ANOVA, their easy availability in the principal statistical software (e.g., R, SAS, SPSS,

* Corresponding author.
E-mail address: matthieu.boisgontier@faber.kuleuven.be (M.P. Boisgontier).

http://dx.doi.org/10.1016/j.neubiorev.2016.05.034

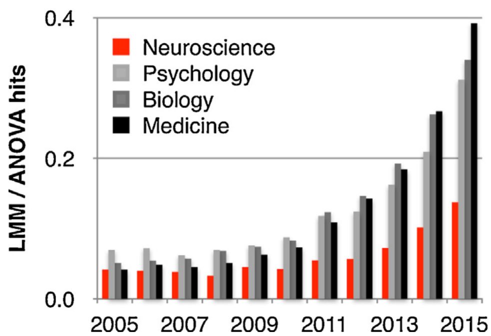

Fig. 1. Linear mixed models over ANOVA ratio. Ratio of "mixed effect" or "mixed model" (LMM) over "ANOVA" hits when associated with "neuroscience", "psychology", "biology", or "medicine" in a Google Scholar search disregarding patents and citations on March 6th, 2016.

Stata), and the fact that sticking to ANOVA may result in spurious results (Jaeger, 2008) should have resulted in a preference for LMM. This is clearly not the case so far (Fig. 1). In 2015, the ratio of "mixed effect" or "mixed model" over "ANOVA" hits was equal to 0.14 when associated with "neuroscience". In other fields such as biology, psychology and medicine, the relative use of LMM was higher, with a maximum ratio of 0.39 in medicine. These ratios are low, but have recently started to increase dramatically. The total increase over the last 5 years ranges from 150% in neuroscience to 260% in medicine. This effort should be continued towards a more consistent use of LMM in science, and especially in neuroscience. Indeed, assuming a quadratic trend, neuroscience would only reach the current ratio of psychology, biology and medicine in 2020–2021.

In sum, a transition towards LMM is underway and neuroscientists are lagging behind compared to other scientists. In order to catch up, a heightened awareness of how far neuroscience has fallen behind is urgent. If we want to minimize the risk of having the reliability of our field questioned, as was recently and justifiably done (Nieuwenhuis et al., 2011; Button et al., 2013), changing our statistical habits towards higher reliance on LMM appears mandatory.

Competing financial interests

The authors declare no competing financial interests.

Acknowledgments

Matthieu P. Boisgontier is supported by a research grant (1504015N) and a post-doctoral fellowship of the Research Foundation – Flanders (FWO). The authors are thankful to Olivier Renaud and Delphine Courvoisier for providing useful comments on the manuscript.

References

Aarts, E., Verhage, M., Veenvliet, J.V., Dolan, C.V., van der Sluis, S., 2014. A solution to dependency: using multilevel analysis to accommodate nested data. Nat. Neurosci. 17, 491–496.

Baayen, R.H., Davidson, D.J., Bates, D.M., 2008. Mixed-effects modeling with crossed random effects for subjects and items. J. Mem. Lang. 59, 390–412.

Barr, D.J., Levy, R., Scheepers, C., Tily, H.J., 2013. Random effects structure for confirmatory hypothesis testing: keep it maximal. J. Mem. Lang. 68, 255–278.

Button, K.S., Ioannidis, J.P., Mokrysz, C., Nosek, B.A., Flint, J., Robinson, E.S., Munafò, M.R., 2013. Power failure: why small sample size undermines the reliability of neuroscience. Nat. Rev. Neurosci. 14, 365–376.

Jaeger, T.F., 2008. Categorical data analysis: away from ANOVAs (transformation or not) and towards logit mixed models. J. Mem. Lang. 59, 434–446.

Judd, C.M., Westfall, J., Kenny, D.A., 2012. Treating stimuli as a random factor in social psychology: a new and comprehensive solution to a pervasive but largely ignored problem. J. Pers. Soc. Psychol. 103, 54–69.

McCulloch, P., Searle, S.R., Neuhaus, J.M., 2008. Generalized, Linear, and Mixed Models, 2nd ed. Wiley.

Nieuwenhuis, S., Forstmann, B.U., Wagenmakers, E.J., 2011. Erroneous analyses of interactions in neuroscience: a problem of significance. Nat. Neurosci. 14, 1105–1107.