Smart Tools for Academic Submission Decisions: Waiting Times Modeling

Strumenti "Smart" per sottoporre i manoscritti accademici: modelli per i tempi di attesa

Francesca De Battisti - Giancarlo Manzi

Abstract This paper illustrates the results of a multilevel analysis aiming at the determinants of peer-review waiting times until acceptance across some of the top statistics & probability journals. Classical multilevel tools are used for analyzing waiting times until acceptance for around 3,500 articles. Results reveal the importance of the number of authors, their academic level and nationality, together with the quality level of the journal as the most important factors affecting waiting times.
Abstract *Questo articolo presenta i risultati di un'analisi multilivello sui fattori che determinano il tempo di attesa per il processo di peer review in alcune delle principali riviste accademiche di statistica e probabilità. Strumenti classici della metodologia multilivello vengono utilizzati per analizzare i tempi di attesa fino all'accettazione di circa 3.500 articoli. I risultati rivelano che il numero degli autori, il loro livello accademico, la loro nazionalità e il livello della rivista sono i fattori più importanti che influenzano i tempi di attesa.*

Key words: Peer review, Statistical journals, Waiting times, Hierarchical models

1 Introduction

In recent years academic peer-review waiting times are constantly increasing. Figure 1 shows the yearly average number of days between submission and acceptance of published articles in the Advances in Data Analysis and Classification (ADAC)

Francesca De Battisti
Department of Economics, Management and Quantitative Methods, Università degli Studi di Milano, Via Conservatorio 7, 20122 Milan, Italy. e-mail: francesca.debattisti@unimi.it

Giancarlo Manzi
Department of Economics, Management and Quantitative Methods, Università degli Studi di Milano, Via Conservatorio 7, 20122 Milan, Italy. e-mail: giancarlo.manzi@unimi.it

journal[1] between 2011 and 2016, and depicts a scenario where it has more than doubled from 224 days in 2013 to 474 days in 2016. This is a situation that not only ADAC but also other journals are experiencing, due to lack of peer reviewers, increasing number of submissions, among other reasons.

The scientific peer review process has always been labour intensive, costly, and often slow, with resulting delays in publication [9]. Nowadays we are witnessing the progressive introduction of stricter rules and guidelines by editors and publishers with the latter increasingly favoring pay-to-publish process, with the side effect of a general value downsizing of academic research [12].

As for the impact of academic research, still the impact factor (IF) of academic journals is considered one of the cornerstones for the evaluation of journals [1, 11], whereas for the personal reputation of authors, citation-based indexes like the Hirsh index *does seem to be able to identify good scientists* [2, 8]. However, they both have been severely criticized and are also regarded as not being able to reflect a broader impact of research [3, 7, 10]. Notwithstanding this drawbacks, they reflect the value of papers and can be referred as important measures to evaluate their quality.

The structure *publisher*→*journal*→*article* might in truth be viewed as a multilevel structure and therefore exploited to reveal hidden characteristics of the scientific research production. Therefore, this paper illustrates the results from multiple multilevel models focusing on finding determinants affecting peer-review waiting times until acceptance across around 3,500 articles published between 2011 and 2016 in 8 statistics & probability journals (chosen among the top 38 journals in

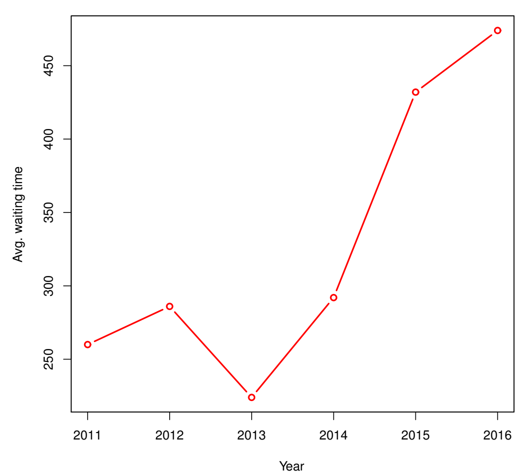

Fig. 1 Average waiting times for articles published in ADAC - Years 2011-2016 (*Source*: Author's calculations based on Springer ADAC website)

[1] ADAC was ranked 31st out of 123 journals in the Web of Science Journal Citation Report Impact Factor 2017 ranking in the Statistics & Probability category.

the 2017 Thomson Web of Science Statistics & Probability ranking) from different publishers. The 2-level structure *journal*$\rightarrow$*article* is chosen for this analysis.

2 Journal, variable and article selection

In constructing our data set, we aimed at top statistical journals. We chose 8 journals among the top 40 statistical journals ranked in the Statistics & Probability category of the 2017 Thomson Journal Citation Report published by multiple publishers. The chosen journals were the *Journal of Statistical Software* (J Stat Softw - ranked 1st), *Fuzzy Sets and Systems* (Fussy Set Syst - ranked 8th), the *Journal of the Royal Statistical Society - Series A - Statistics in Society* (JRSSA - ranked 11th), the *Journal of the American Statistical Association* (JASA - ranked 16th), the *Annals of Probability* (Ann Prob - ranked 19th), the *Journal of Business Economic Statistics* (JBES - ranked 20th), *Advanced in Data Analysis and Classification* (ADAC - ranked 31st), and *Biostatistics* (ranked 38th). This selection was done in order to cover multiple fields of statistics and probability, and allow for multiple regional covering, considering both data analysis journals and theoretical journals.

In selecting potential predictors and response variables for our models, we used both manual and automatic web scraping and data parsing on the journals' websites. For each article (uniquely identified by the Document Object Identifiyer - DOI) we extracted the date of submission (always available) as the date of the "birth" of an article in the peer-reviewing process, and, as the date of the "death" of an article, the date of acceptance when this was available, otherwise the date of online publication or the date of final revision. The difference between the date of acceptance and the dates of online publication or the final revision might be considered negligible with respect to the average waiting times for acceptance. It ranged between a few days and one or two months.

We also searched for the keyword "Bayes" in the articles and classified them as "Bayesian articles" (value equal to 1) if it was clear from the occurrences that the Bayesian method was used extensively in the article (variable *BAYES*). We operationalized a variable indicating whether the article was openly accessible or not (*OPEN*). From the Scopus database (www.scopus.com) we extracted the number of citations and the h index for each author. Computed variables entering the models were the length in days of the waiting times between submission and acceptance (*AGE*), the average h index among authors (*AVG_HI*), the standard deviation of the h index of the authors (*SD_HI*), the average monthly number of Scopus citations per article (*MONTH_SCOP_CIT*), the number of authors of the articles (*NUMBER_AUTHORS*), two dichotomous variables representing respectively the presence of young or non-expert researchers among authors (expressed by a Scopus h index lower than 5 - variable *JUNIOR_LESS5*), and the presence of expert researchers (expressed by a Scopus h index greater than 20 - variable *SENIOR_MORE20*), a dichotomous variable equal to 1 if all the authors were affiliated to an US institute and 0 otherwise (*USA_ALL*), a dichotomous variable equal to 1 if at least one of the authors was affiliated to an US institute and 0 otherwise (*USA*), a dichotomous variable equal to 1 if all the authors were affiliated to an institute of the same country (*SAME_COUNTRY*), a dichotomous variable equal to 1 if the main author (detected by the highest *h* index among the authors) was affiliated to an institute of the same country of the institute to which the editor(s) in chief of the journal was affiliated (*SAME_NATIONALITY_EDS*), the country of the institute to which each author was affiliated, a categorical variable representing the continent of the institute to which the most expert author belongs. Journal-level variables were the 2017 Thomson Reuters impact factor (*IF*), the journal frequency in days (*PERIODICITY*), the age of the journal since its foundation (*AGE_JOURNAL*), the AI index (i.e. the article influence score measuring the average influence of a journal article over the first five years after publication - variable *AI*).

Articles considered in the analysis were only non-solicited research articles not included in special issues. We excluded presidential addresses, editorials, book reviews, conference proceedings, code snippets, comments, rejoinders and corrections. After this inclusion procedure, a total number of around 3,500 articles formed our database subdivided across journals as reported in Table 1.

Table 1 Distribution of articles across journals

Journal	*Publisher*	*Country*	*Eds' country*	*No. of articles in the database*	*%*
JRSSA	Royal Stat.Soc./Wiley	UK	UK	246	7.03
ADAC	Springer	GER	ITA, GER, JAP	127	3.63
JBES	Am.Stat.Ass./ Taylor&Francis	USA	USA	252	7.20
Ann Prob	Inst.Math.Stats./ Bernoulli Society	USA	USA	485	13.85
Biostatistics	Oxford Uni. Press	UK	NED, USA	334	9.54
Fuzzy Set Syst	Elsevier	NED	BEL, FRA, GER, SPA	982	28.05
JASA	Am.Stat.Ass./ Taylor&Francis	USA	USA	734	20.97
J Stat Softw	UCLA Dept.Stats	USA	AUT, SWI, GER	341	9.74

3 Model used and results

A standard random intercept-random slope multilevel model with two levels (article=level 1; journal=level 2) without interactions between first and second level predictors was used to model the responses "days from submission to acceptance":

$$Y_{ij} = \gamma_{00} + \gamma_{10}\mathbf{X}_{ij} + \gamma_{01}\mathbf{Z}_j + u_{0j} + u_{1j}\mathbf{X}_{ij} + \varepsilon_{ij}, \qquad (1)$$

where i is the index for article, j is the index for journal, $\mathbf{X}$ is a vector of level 1 predictors, and $\mathbf{Z}$ is a vector of 2-level predictors.

Table 2 shows the results of model (1) applied to our data set. The dependent variable "time between submission and acceptance" is regressed towards various independent variables. The null model (column I), the model with only 1-level predictors (column II), the full model (column III) and the best model when only *AI* is retained among 2nd-level predictors (column IV) are shown.

Table 2 Multilevel model (1) applied to the article data set. Dependent variable: time between submission and acceptance. SE in parentheses. Significance: * 0.10; ** 0.05; *** 0.01

Variable	I (null)	II (1-level only)	III (full model)	IV (best model)
Intercept	460.21(57.89)***	452.43(59.58)***	270.07(85.37)***	302.20(25.63)***
First level predictors:				
BAYES		5.26(11.78)	5.27(11.79)	4.79(11.67)
OPEN		1.21(17.39)	-3.03(17.60)	
MONTH_SCOPUS_CIT		-1.46(1.54)	-1.48(1.54)	
AVG_HI		-0.22(0.73)	-0.22(0.73)	-0.93(0.46)**
SD_HI		-0.58(0.92)	-0.59(0.92)	
JUNIOR_LESS5		9.53(9.74)	9.58(9.74)	5.71(8.80)
SENIOR_MORE20		-11.92(11.89)	-12.05(11.89)	
NUMBER_AUTHORS		8.05(3.88)**	8.11(3.89)**	6.00(3.46)*
USA_ALL		0.60(17.16)	-0.05(17.16)	
USA		2.99(14.11)	2.34(14.11)	
SAME_COUNTRY		0.53(11.38)	0.97(11.40)	
SAME_NATIONALITY_EDS		-17.56(10.19)*	-16.92(10.20)*	-18.53(9.32)*
Second level predictors:				
PERIODICITY			-0.28(0.66)	
AGE_JOURNAL			-0.43(0.77)	
IF			-9.33(7.24)	
AI			93.64(19.68)***	66.88(8.59)***
AIC:	48066.8	47275. 8	47280.8	48021.9

4 Discussion on results and future work

The contribution of this paper is twofold. Firstly, to our knowledge this is the first time a database focused on the time interval between submission and editorial decision and comprehensive of many features of articles contained in more than one journal is built. The work by Bornmann and Daniel [4, 5] was focused on the time interval between submission and editorial decision of 1899 communication manuscripts, but only considering one journal, the *Angewandte Chemie International Edition*. Secondly, it represents the basis for further and more appealing analysis to be implemented with more sophisticated statistical tools like the multilevel excess hazard survival model proposed by Charvat et al. [6], which will be one important task to be carried out in the future. Results of our analysis show that to some extent the number of authors forming the research team seems to increase the waiting time for acceptance. This could be interpreted as a wasting time from the side of researchers due to more time needed for meetings, taking decisions etc. If the most important researcher among the authors is of the same nationality of one of the editors in chief the waiting time for acceptance decreases. This could be imputed to the attitude of editors being facilitated in communicating with researchers. Among second level predictor, *AI* seems to be the best: the higher the quality level of the journal the longer the waiting time. There is a lot of future work to do. Among many things, further levels comprising publishers and authors must be considered, the use of a multilevel survival analysis to be compared with the simple multilevel analysis performed here and a research impact analysis based on the Scopus citations per article.

References

[1] Archambault, E., Larivire, V. (2009) History of the journal impact factor: Contingencies and consequences. *Scientometrics*, 79(3), 635-649.
[2] Ball, P. (2007) Achievement index climbs the ranks. *Nature*, 448, 737.
[3] Bornmann, L., Daniel, H.D. (2008) What do citation counts measure? A review of studies on citing behaviour. *J Doc*, 64, 45-80.
[4] Bornmann, L., Daniel, H.-D. (2010a) How Long is the Peer Review Process for Journal Manuscripts? A Case Study on Angewandte Chemie International Edition. *CHIMIA*, 64(1-2), 72-77.
[5] Bornmann, L., Daniel, H.-D. (2010b) The manuscript reviewing process: Empirical research on review requests, review sequences, and decision rules in peer review. *Library & Information Science Research*, 32: 5-12.
[6] Charvat, H., Remontet, L., Bossard, N., Roche, L., Dejardin, O., Rachet, B., Launoy, G., Belot, A. (2016) A multilevel excess hazard model to estimate net survival on hierarchical data allowing for non-linear and non-proportional effects of covariates. *Stats in Med*, 35, 3066-3084.
[7] Didegah, F., Bowman, T.D., Holmberg, T. (2018) On the Differences Between Citations and Altmetrics: An Investigation of Factors Driving Altmetrics Versus Citations for Finnish Articles. *J Assoc Inf Sci Tech*, 69(6), 832-843.
[8] Egghe, L. (2010) The Hirsch index and related impact measures. *Information Science and Technology*, 44(1), 65-114.
[9] Hames, I. (2007) *Peer Review and Manuscript Management in Scientific Journals. Guidelines for Good Practice*. Malden, MA: Blackwell Publishing.
[10] Holmberg, K., Didegah, F., Bowman, T.D. (2015) The different meanings and levels of impact of altmetrics. Proceedings of the 11th International Conference on Webometrics, Informetrics and Scientometrics & 16th COLLNET Meeting, 26?28 November, New Delhi, India.
[11] Seiler, C., Wohlrabe, K. (2014) How robust are journal rankings based on the impact factor? Evidence from the economic sciences *J Informetr*, 8, 904-911.
[12] Teixeira da Silva, J. A. (2017) Fake peer reviews, fake identities, fake accounts, fake data: beware! *AME Med J*, 2(28), 1-3.