Applied Analysis 6

Rohan Hore

August 2024

1 Effect of open review policy on reviewers behavior towards a journal paper

The data set consists of data from the peer review process for papers submitted to academic journals. When a paper is submitted, the journal editor sends it to potential reviewers. Each reviewer can choose to accept or decline the invitation to review the paper. If the invitation is accepted, then they need to (1) write a review (which can be any length) and (2) choose a recommendation for what the journal should do with the paper(accept / request minor revisions / request major revisions / reject).

Typically, the reviewers are anonymous in the peer review process (i.e., the author will not see the names of the reviewers). However, the journals in this study implemented an open review policy several years ago, meaning that reviewers can choose to attach their names to the review. The goal of this paper is to examine changes in reviewer behavior that resulted from this change in policy. The data set contains data from years before and after this change was implemented. The data and analysis scripts from the paper were downloaded from nature article. The data set contains the following variables:

- `id` and `journal` are unique identifiers for the paper and for the journal it was submitted to.
- `invitation.date` and `year` indicate the date/year that the reviewer was invited to review the paper.
- `open.review` indicates whether the journal is o ering an open review option at the time of this paper
- `review.complete` indicates whether the reviewer submitted the review.
- `name.published` indicates whether the reviewer chose to publish their name.
- `recommendation` is what the reviewer recommended for the paper: Accept, Minor revisions, Major revisions, or Reject
- `accepted` indicates whether the reviewer accepted the invitation to review the paper (note: this does not mean that the reviewer recommends acceptance of the paper).
- `review.time` is the number of days between when the reviewer was invited to review, and when the review was submitted.
- `polarity` and `subjectivity` are variables computed via natural language processing. polarity takes values in $[-1, 1]$, where positive and negative values indicate positive and negative sentiments (e.g.,great or terrible). subjectivity takes values in $[0, 1]$, with larger values indicating an opinion(subjective) while smaller values indicate factual information (objective).
- `nchar` is the length of the submitted review (# of characters).
- `reviewer.status` takes values Professor, Dr, and other, recording whether the reviewer is a professor/faculty, or they have their PhD but are not a professor/faculty, or they do not have a PhD.
- `gender` is the gender of the reviewer. This information is not provided by the reviewer, but was imputed based on the name of the reviewer.

Data Reading

```r
#packages to load
library(lmerTest)
library(ordinal)
library(lubridate)
library(dplyr)
library(tidyr)

# Data upload and preparation
round1 <- read.csv("RevData.csv")
round1$id <- as.character(round1$id)
round1$journal <- factor(round1$journal, labels=c("Journal 1","Journal 2", "Journal 3",
                "Journal 4", "Journal 5"))
round1$open.review <- factor(round1$open.review, labels=c("No", "Yes"))
round1$review.complete <- factor(round1$review.complete, labels=c("No", "Yes"))
round1$name.published <- factor(round1$name.published, labels=c("No", "Yes"))
round1$recommendation <- factor(round1$recommendation, labels=c("Reject", "Major revisions",
                    "Minor revisions", "Accept"))
round1$accepted <- factor(round1$accepted, labels=c("No", "Yes"))
round1$reviewer.status <- factor(round1$reviewer.status, labels=c("Professor", "Other", "Dr."))
round1$gender <- factor(round1$gender, labels=c("Female", "Male", "Uncertain"))
```

```
id   journal invitation.date year open.review review.complete name.published
1 405 Journal 1      2010-01-01    0          No             Yes             No
2 405 Journal 1      2010-01-01    0          No             Yes             No
3 406 Journal 1      2010-01-01    0          No             Yes             No
4 406 Journal 1      2010-01-01    0          No              No             No
5 406 Journal 1      2010-01-01    0          No             Yes             No
6 407 Journal 1      2010-02-01    0          No              No             No

   recommendation accepted review.time   polarity subjectivity nchar reviewer.status
1          Reject      Yes          28 0.12838763    0.4085349  4110       Professor
2 Major revisions      Yes          16 0.08102662    0.4350710  4797           Other
3          Reject      Yes           9 0.10333333    0.4083333   687             Dr.
4            <NA>       No          NA 0.00000000    0.0000000     0             Dr.
5          Reject      Yes          39 0.13453609    0.5527891  3904             Dr.
6            <NA>       No          NA 0.00000000    0.0000000     0       Professor

     gender
1 Uncertain
2      Male
3      Male
4      Male
5      Male
6      Male
```

2 Possible Questions

Problem 1

Based on the model summary below, it was argued that

> the pure effect of the open review condition was not statistically significant. Furthermore, although several referee characteristics had an effect on the willingness of reviewing, only the interaction effect with the "other" status was significant.

How would you strengthen this statistical analysis?

Table 1 Mixed-effects logistic model on the acceptance of editors' invitation by referees

Fixed effects	Estimate	Std. error	z-value	p-value
(Intercept)	−0.193	0.214	−0.901	0.368
Open review	−0.025	0.073	−0.343	0.713
Status: Other	−0.476	0.050	−9.476	<0.001
Status: Dr	−0.135	0.030	−4.436	<0.001
Gender: Male	0.277	0.049	5.643	<0.001
Gender: Uncertain	0.338	0.055	6.164	<0.001
Year	−0.121	0.008	−14.415	<0.001
Open review × Status: Other	0.278	0.069	4.020	<0.001
Open review × Status: Dr	0.012	0.042	0.279	0.781
Open review × Gender: Male	−0.014	0.062	−0.219	0.827
Open review × Gender: Uncertain	0.005	0.070	0.074	0.941
Std. Dev. of random effects:				
Submission (intercept)	0.491			
Journal (intercept)	0.463			
No. of observations	62,790.0			
Log likelihood	−38,311.9			
AIC	76,649.8			

The reference class for the referees' status is "Professor", while for gender is "Female"

Figure 1: Model summary from paper

Problem 2

The authors fit a gaussian linear mixed model for polarity and subjectivity with open review, the recommendation by referees, the (log of) the number of characters of the report, the year, and the gender and status of the referees (along with their interactions) respectively were included as fixed effects along with the random effects corresponding to submission and journal IDs, where

- polarity denotes the tone of the report was mainly negative or positive (varying in the $[-1, 1]$, with larger numbers indicating a more positive tone).
- subjectivity denotes whether the style used in the reports was predominantly objective (takes value in $[0, 1]$, higher numbers indicating more subjective reports.

Do you have any suggestions towards improving this model?

Problem 3

Attempt to reproduce Figure 2 of the paper. Based on visual inspection alone, comment on whether the degree of smoothing provided by the authors' Loess lines appears appropriate.

Problem 4

In Table 1 of the paper, the authors used a logistic regression model with interactions to examine the effects of the open review policy on the acceptance probability of review invitations. An alternative approach is to run a logistic regression on each of the 9 subgroups separately (3 status levels $*$ 3 gender categories). For simplicity, in this question let's omit the Year variable and the random effects terms of journal and submission in both approaches.

Can we find a regression model with interactions that has the same model assumptions as a set of simple logistic regression models for each of the 9 subgroups separately? If yes, will the estimates and confidence intervals of the open review effect on each subgroup be different from the two approaches? Provide an analytical justification and also check your conclusions numerically.

Problem 5

Answer the same questions for the cumulative-logit model in Table 2 of the paper. Here, the response is recommendation which is treated as an ordered categorical variable in the paper; here we compare the full interaction cumulative-logit model with the cumulative-logit model on each subgroup separately, while we ignore the other variables.

Problem 6

As the open review policy is not randomized, the open review effect is confounded with year/time. The paper adjusts for the confounding year effect by adding a linear fixed effect term of year in their regression models. Assuming that the year effect is linear can be a strong assumption. For instance, our reproduced plot for proportion of accepted papers clearly suggests that the year effect could be non-linear.

In this question, we will use only the data on 3 journals Journals 1, 3, and 5 from years $2010 - 2014$ (before the pilot study starts for Journal 3/5). We focus on estimating the policy effect on review time(days) for Journal 1. Instead of assuming a shared linear effect of year as in Table 3, we assume that the Year effect (mean review time differences across years, after controlling for all other variables) is the same for all 3 journals. Perform an analysis to estimate the average effect (averaged across the reviewers who have accepted and completed the review) of the open review policy on the review time for Journal 1 after adjusting for Year and test whether the average effect is 0 or not.

Problem 7

In this question, we will examine how the probability that a potential reviewer accepts the review invitation varies among papers in each journal. Call

$$p_j = \text{the probability that an invited reviewer accepts to review paper } j$$

and, we assume that it is a property of the paper (and the journal it was submitted to), but not dependent on reviewer characteristics.

You may model the reviewer acceptance/non-acceptance data for each paper j using either a binomial or negative binomial model, with success probability p_j. For each journal, assess the variation in p_j across papers. For which journal(s) is there strong evidence that p_j is not constant across papers? Which journals appear to have greatest variability in p_j?

Problem 8

Using an Empirical Bayes approach, or otherwise, obtain an approximate posterior mean and 90% credible interval for each p_j. Compare the posterior mean estimates with the maximum likelihood estimates.