Genre-Specific Analysis of Ratings and Academy Award Nominations

Team 19
Sheridan B. Green,[1] Daming Li,[1] Grant L. Schumacher,[1] and Youyuan Zhang[2]
[1]Department of Physics, Yale University, [2]Department of Mathematics, Courant Institute, New York University

Highlights

- **Genre-by-genre descriptive analysis of film ratings**
- **Predictive modeling of Oscar nominations using rating statistics and genre information**
- **Causal study of the effect of Oscar nominations on movie ratings**

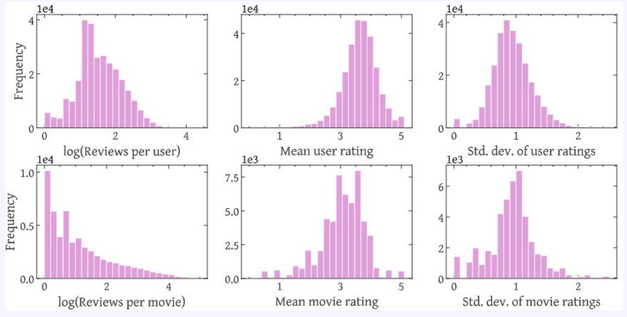
Questions

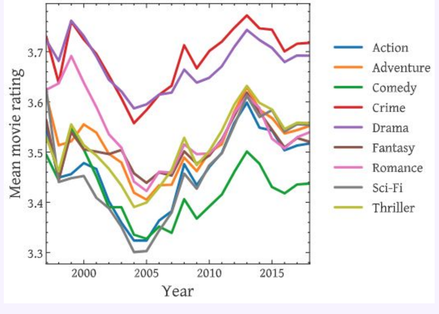
- **Are there any genre-dependent temporal evolution patterns and statistical features present in consumer movie ratings?**
- **How predictive are user rating statistics to Academy Award nominations, and, conversely, how do such nominations end up reshaping future user rating statistics?**

Data

Our analysis is based on the *MovieLens* database (user reviews and film metadata) and Academy Award Nominations.

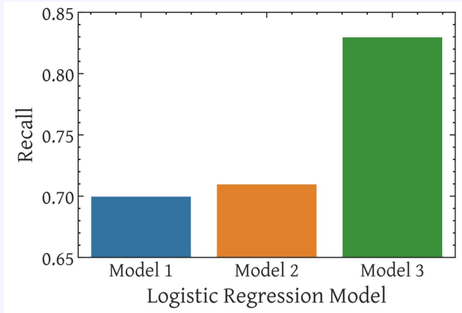
Descriptive visualizations of the distributions of ratings by user and by movie, as well as time evolution of rating by genre is provided below.





Can we predict Oscar Nominations?

Adding the standard deviation of ratings and genre information significantly improves the prediction accuracy of nominations



Logistic regression model used for nomination predictions.

Model 1: number of reviews + mean ratings

Model 2: Model 1 + standard deviation of ratings

Model 3: Model 2 + genre information

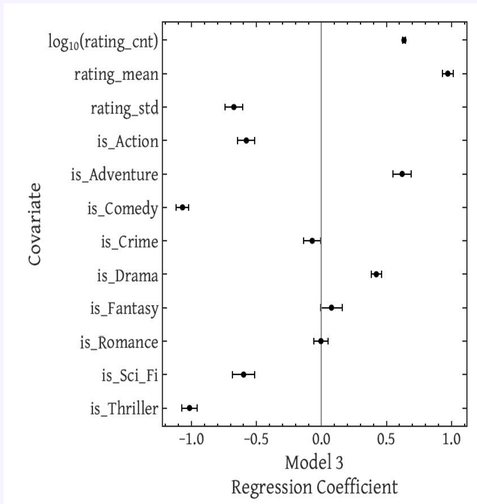
Recall = fraction of correctly predicted true Oscar nominations in the test set

Examination of the regression coefficients

Probability of nominations is hugely genre-dependent.

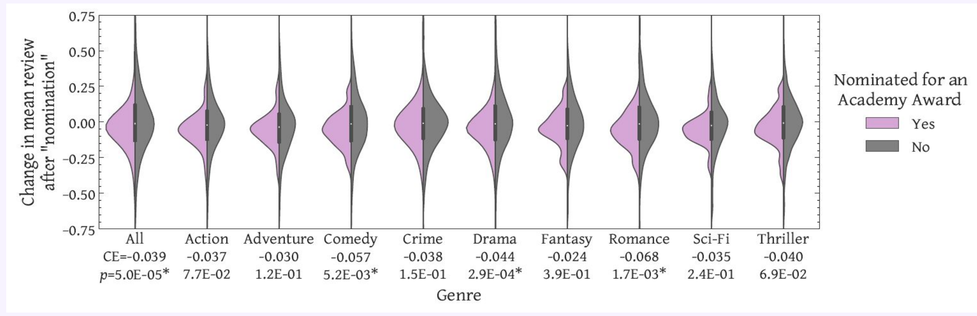
The standard deviation of ratings across users for a given movie is *negatively* associated with Oscar nominations chance.

More "controversial" films (those with a wider range of reviews) are less likely to be nominated for Oscars.



Is an Oscar nomination worth it?

By analyzing ratings before and after an Oscar nomination, we see that a film's ratings might slightly suffer as a result.



Causal effect = Change in mean rating relative to movies not nominated

The drop in average rating right after nominations is weak but significant

Drop is present across all genres with small differences

To win an Oscar, focus on producing dramas and adventure films and target a wide audience, rather than producing comedies or thrillers and targeting a specific group of viewers.

Future directions

By combining the user ratings and movie tag genome scores from *MovieLens*, one could potentially identify non-trivial user groups using clustering techniques.

This would allow engineers to build a better recommendation system that, we suspect, will become crucial as the distribution model for films and TV moves further toward streaming, subscription-based services.