Renjun Zhu
Professor: Aldous
May 6th, 2016

Stat 157 Course Project:
Thoughts About the Prediction of the Best Picture Oscar

1. Introduction:

Each year, in February, the Academy Awards of Merit, nicknamed “Oscar,” is bestowed on people who participated in certain important roles which contributed to the production and release of a film during the previous year. The awards have been a symbol of excellence in the field of filmmaking since 1928. The awards are given in a variety of categories such as best picture, director, actor/actress, cinematography and 16 others. In order to be eligible for the awards, films produced from the previous year have to go through the process of being placed on the nomination list first, and only those films are voted on by the members of the Academy of Motion Picture Arts and Sciences. The members of the Academy must have worked in the film industry and must be invited by leaders of specialized areas of filmmaking.

The award season is not just the dream for filmmakers, movie fans and businesses commercializing the industry also hold stakes in these awards. For example, culturally, many people have followed their favorite celebrities who are involved in the film industry. The sponsors and movie studios also highly prize the Oscar awards to boost the box office revenue and make new records for their investments. Many consumers rely on online reviews, ratings and awards information to decide which movie is worth their time and money. Winning an Oscar is not just an honor, but also has a cascading effect through the consumer industry.

Whether it is before the award ceremony or after, like betting a soccer game or a horse race, there are many voices who vehemently forecast their predictions. “Oscar Buzz”, is a social phenomenon carried on all the media websites. There is even a well-established blog that covers only topics related to the Oscars from the past to the present, and many people rely on this information to predict future winners. Besides social network speculation, experts from different fields apply their methods to make speculations concerning the winners. The debate whether Oscar awards are predictable or not as well as if they are worth predicting is controversial and stimulating. Some might argue that if the Oscar award could be predicted, the value of the awards would be lost. Some might advocate different math formulae to forecast the awards, and still others might use the available data to argue on the sides of social justice issues. This course project is purely made from my own interests. It is a study of other scholars' insights on the same topic that brings me joy.

2. Literature Review:

There are many ways to examine the Oscar awards. In the series of updated papers *Applying Discrete Choice Models to Predict Academy Award Winners* by Iain Pardoe and Dean Simonton (2005, 2007, 2008), the authors examine Oscar winners from 1938 to 2006 by applying multinomial logit modelling based on Nerlove and Press (1973):

$$\Pr(Y=j|\, X_i) = \frac{e^{(\beta^T X_{ij})}}{\Sigma_{\, h \in C_i} e^{(\beta^T X_{ih})}}$$

$$log[\frac{Pr(Y=a|Xi)}{Pr(Y=b|Xi)}] = \beta^T(X_{ia} - X_{ib})$$

Where *i* is the experiment (category), *j* is the choice (movie), and *β* is the parameter. Conditional on the choice of each nominee (a or b), repeatedly test each factor and see if such a factor has an effect on each of the nominees. Through this logit model, they found Oscar awards results are associated with the results of other major awards: Golden Globe for Best Picture for Drama and Musical, Screen Actors Guild (SAG), Directors Guild of America (DGA) & Producers Guild of America (PGA). Additionally, they discovered a high correlation between nominees for Best Picture and Best Director, as well as a correlation between winning and the total number of nominations in other categories. Also, this model provides a prediction accuracy of 70% for the Best Picture award.

Unlike Iain Pardoe and Dean Simonton, the article *Predicting Movie Success and Academy Awards through Sentiment and Social Network Analysis* by Peter Gloor (2009) gives another insight by analyzing the discussion of content from the online community to demonstrate the tendency for a movie to succeed. Gloor's model,

Oscar Model = a*δ+b*γ+c*λ | a+b+c=1

has three parameters: δ= intensity, γ= positivity, and λ= time noise, which represent the frequency of a movie having been mentioned, the tone of the reviewer, and the movie release time respectively. He found the best fit for this model with a=0.5, b=0.3, c=0.2; and there is a strong correlation (r = 0.75) when applying this linear model.

Brendan Bettinger (2011) in *Cinemath: What Makes a Best Picture, A Look at Rating, Runtime, and Genre Over 80 Years of Oscars* reveals several factors that were previously thought to be irrelevant, yet might be considered as a part of the prediction of Oscar winners such as running time and genre. This basic research leaves many unanswered questions and yet provides a strong foothold to investigate potentially uncovered factors that may aid in future predictions.

3. Data and Method:

After reading several articles, it appeared prudent to investigate several factors that could be considered relevant to the prediction of the oscars. To begin, data pertaining to the Academy Awards was collected from IMDb and Wikipedia. From that data, 10 charts were created with information dating from 2006 to 2016 and containing Oscar nominations and awards for all categories. Additionally, potentially relevant data was added, such as ratings from the website IMDb and rough categorizations of the films into genres: drama, documentary, and animation. During the data collection process, it was noted that most of the people appeared repeatedly on the list, especially the producers in the Best Director category. It may be natural to assume that these people may serve as a guarantee that the film productions in which they are involved will maintain a high level of quality, so in turn they produce films with high standards and position their films with a competitive edge over their potential competition in the film industry. This observation matches much of the literature reviewed, so this will be evaluated by generating a graph to visualize the intersection of winners in the Best Director and Best Picture categories. Similar observations also apply to total number of prior nominations for each nominee in other categories and the winning Best Picture. The analysis will cover IMDb rating scores and other awards, not just the four major awards included in Pardoe and Simonton's paper. To add other potentially impacting factors, Best Actor and Actress, Best Supporting Actor and Actress, and ratings by the Motion Picture Association of America (R, PG, PG-13) were also taken consideration. Additionally, budgets for films nominated for Best Picture were captured since there were a couple of years (2006 and 2013) which stood out and called for further testing. Finally, gross earnings within the first month of release in the United States was captured to see if that had any major effect on winning.

Data is a way to present ideas, however prediction does not always rely only on data. Since prediction is not a certainty, it can be swayed by incorrect assumptions and factors. It seems that the lay definition of prediction is to make a knowledgeable guess, not by intuition and observation, but using the scientific method, that hopefully produces a reasonable and meaningful conclusion. In this analysis, knowledge from the Concepts of Probability course and from the Seminar on Topics in Probability and Statistics at UC Berkeley was used to produce the linear model to predict winners from nominees in the Best Picture category. Generally, there are 5 to 12 nominees that appear in this category, and each nominee is independent from the others since they are filmed by different directors and production teams. Of course, the chance of each nominee winning cannot be calculated as $1/n$, because they are not uniformly distributed and affected by several factors. Therefore the primary goal was to find factors (indicators), assign them reasonable weights, and apply a linear model to make a prediction for a future winner.

To examine the hypotheses from literature and observation, a larger set of data was collected from the IMDb, beginning the year Oscar was born in 1928 to 2015, a total of 83 years (with 4 years missing). First, an examination of each factor was completed to see which ones could be indicators for my prediction. Those factors are total number of nominations (N), nominees who were also nominated in best director category (D), nominees who also had nominations in at least one of the four acting categories (A), public rating in IMDb (R), MPAA rating for a particular genre (r), other awards (O), budget for making the film (B), and the gross from the first month after a film was released (G). For a better visualization, several R packages were used to highlight the variation between nominees and winners. Secondly, after picking suitable indicators, each indicator needed to be rescaled into a comparable type. By taking the highest value in each group as a measurement, each nominee was given a corresponding value divided by the highest value. Notice that some of the indicators will have zero value, such as indicator D and A. This does not mean the probability is zero after rescaling. In this case, using complement probability, zero values will treated by subtracting from the probability of winning in that category from 1. For example, if P (D) = 0.9, then P (not D) = 1- 0.9=0.1. To get a value of prediction, apply the formula:

$$P\text{ (Winner)} = \sum \text{coefficient} * \text{indicator}$$

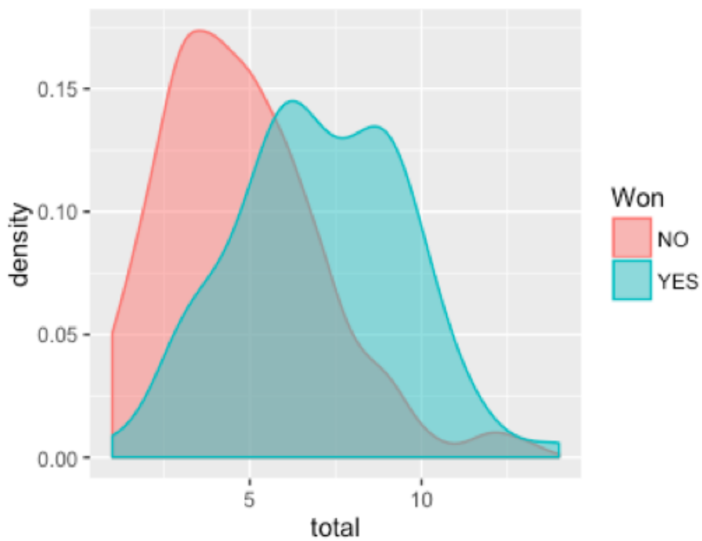
From the formula, the next goal is to find the coefficient for each indicator, and see how much each indicator should weigh for this prediction. It is also important that the sum of the coefficients must equal 1. To find each coefficient, data was taken from a ten year set (2006-2016), and winners were taken from a random sampling from this 10 year set. Using the concepts of Linear Algebra, if each indicator is assumed to be independent from each other, and they are all the indicators of concern, this means no indicator can affect the others, and therefore these indicators can be a basis for generating a winner from nominees. Comparing the values of their probability after such calculation between each nominee, the one with highest value would be the potential winner.

Assign the rescaled value from the selected indicators to an n-by-n matrix A, n depends on the number of indicators are being selected, and then take n winners to form the coefficients. Let v be a one by n vector, which has n components, which represent the coefficients of n indicators respectively. Winners determined from past ceremonies have 100 percent certainty because the event has already occurred, so a value of 1 is assigned, then the goal is to calculate $v = \langle \alpha, \beta, \gamma, \eta, \omega, \delta .... \rangle^T$ through $A*v = \langle 1,1,1,1,1,1 \rangle^T$. This method is based on the past event, so there will be error involved when predicting the future winner, because future events have not happened yet, so it should not have 1 as probability after calculating the result from coefficients. Since there will be many possible combinations ($10Cn$, where $n \leq 10$) to form such a linear system of equations even though this is using only 10 years of data. In order to improve the accuracy of the coefficients, this process was applied many times since it is a random sample, then a mean value from all the sets of processes was formed.

4. Result:

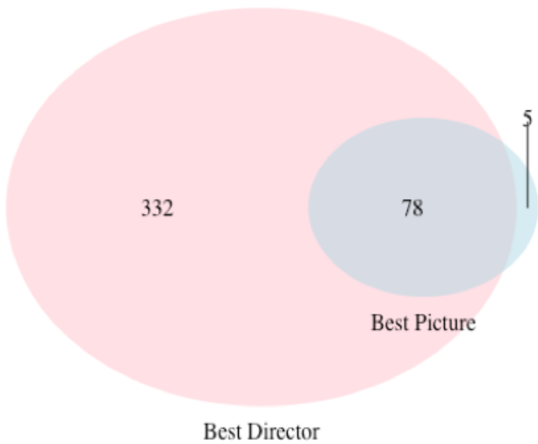
4.1 Indicator: N

From the data for all 83 years of the Academy Awards, a density graph was produced below which highlights the difference in densities of total nominations between films which won the Best Picture award and those that were nominated but did not win. From the graph, a difference is discernible, corroborated by the data where median values for number of nominations (N) for winners, 7, is much higher than the median number of N for nominees, 3. The huge deviation suggests that N should be kept as an indicator.



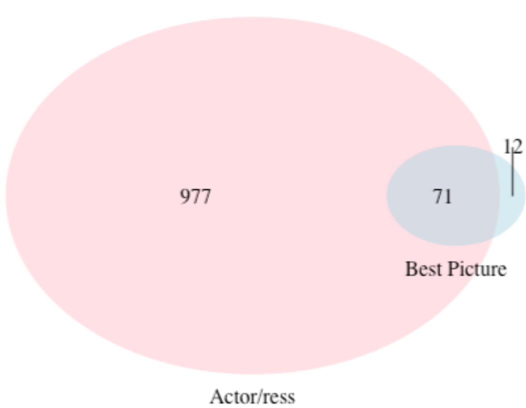
4.2 Indicator: D

Films are made by people, so it is reasonable to look at the statistics related to their directors when making predictions for the best picture category. The Venn diagram below helps draw a certain conclusion what type of consideration we should make for the intersection of those two categories and the power that one may have over the other. Throughout the Oscar's history, only 5 among 83 winning films did not also have their director nominated in the Best Director category. Applying simple calculation for the impact of being nominated for Best Director on the probability of winning Best Picture we get $P(D) = (83-5)/83 = 0.9$, rounding to one significant digit. This also indicates that the chance for the films which are not also nominated in Best Director category, but still win Best Picture is $1 - 0.9 = 0.1$. This large discrepancy supports using nominations for Best Director as an indicator.



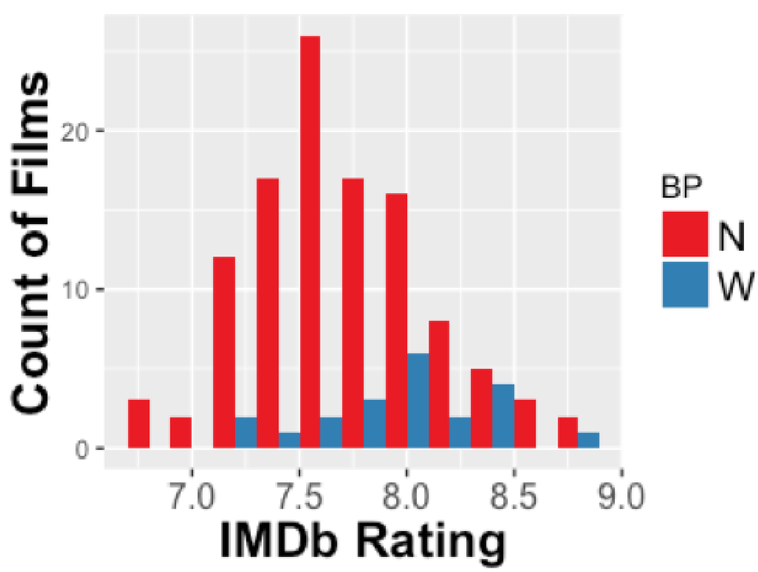
4.3 Indicator: A

Besides connecting films to their director, another group of people with power over the success of a film are the actors and actresses. Since there are 4 major categories of acting: Best Actor, Best Actress, Best Supporting Actor, and Best Supporting Actress, only 12 winners did not also receive one of these four acting nomination. Applying the same calculation as applied to indicator D, P(A) = (71-12)/83 = 0.8, once again rounding to one significant digit. This implies that indicator A should also be drawn into consideration when predicting the Best Picture winner.



4.4 Indicator: R

Bettinger concludes that “The Best picture was one of the top 2 highest-rated nominee about 60% of the time” (web, 2011), which encouraged taking a closer look at this set of data. Due the the limited relevancy of the data from IMDb -- it was born in 1990 and only very widely used after 1996 -- to examine this indicator, the data set of years in consideration had to be shrunk the range from 1996 to present. The histogram below compared the normal shape of nominees and winners in the 21 year time frame. Although the deviation of the rating scales is subtle, Bettinger’s suggestion to consider it as an indicator can be taken for the time being. Later, its impact can be examined by looking at the value of its weight (coefficient) to see how much it affects the predictability of a winner.



4.5 Indicator: r

Using the Motion Picture Association of America’s (MPAA) ratings and the same set of data as above, the table below was generated. It is more likely for a film to win Best Picture with an R rating, which is restricted for youth under 17 years old without a parent due violence or language or sexual content, than for a film with a PG-13 rating, restriction for youth under 13 years old without a parent because of sexual or violent material involved. The data also shows, among nominees in the most recent 20 years R took more than half of the Oscars for Best

Picture. Bayes' Rule is applied to calculate the conditional probability for R films P(W|R)=15/(15+59)=0.2 and for PG-13 films, P (W| PG-13)=6/(6+41)=0.13. Relatively, on a scale of 1, the chance of winning with an R-rating compared to the chance of winning with a PG-13-rating is 0.61:0.39. This does not mean that other categories would have zero chance, even though in the span of 20 years of data, there is no winner from the G and PG categories. By rounding down the value of P(R), and P(PG-13), the probability of winning Best Picture where a film is in one of the other ratings is: P(others) = 1- 0.6-0.3=0.1.

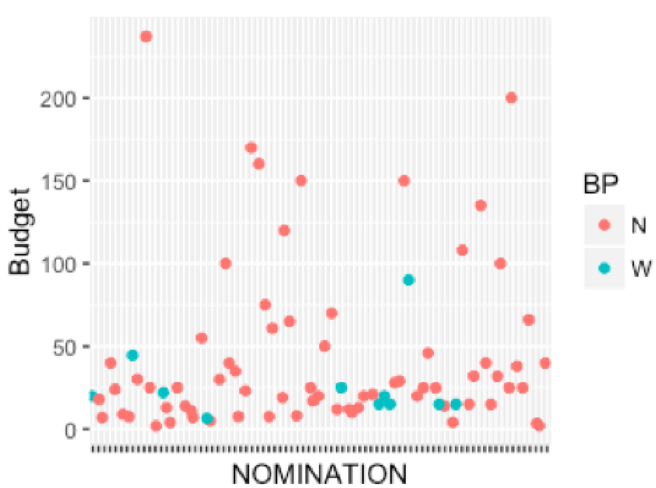
	BP	RATED	total
1	N	G	2
2	N	PG	9
3	N	PG-13	41
4	N	R	59
5	W	PG-13	6
6	W	R	15

4.6 Indicator: O

Inspired by Pardoe and Simonton's analysis using other awards, it was promising to investigate this area in more detail. Instead of considering only four major awards (Golden Globe, SAG, DGA, and PGA), all other awards prior to the Academy Awards in the areas related to Best Picture and Director were collected for each nominee. The boxplot below draws a nice comparison between nominees and winners. On the left, the nominees present with several outliers, and have a mean value of 4. However, on the right the winners, without outliers, have a mean value of 10. Even though the notch on the plot to the right shows the confidence interval being a little bit weak-- the number of winners is always less than the number of nominees-- and the relationship between having more other awards to winning is very obvious. Therefore this indicator should also be kept.

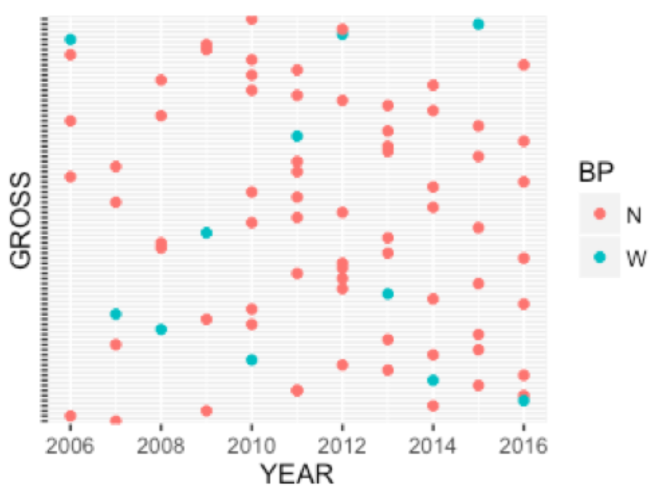
4.7 Indicator: B

The budget seems has no effect on winning as the scatter plot below did not show such any direct correlation. This makes sense because films are made with a wide range of budgets depending on the type of film. For example, documentary types of films do not need as much of a budget as some of the animation or science fiction films, which use lots of advanced technology. Some notable films like Frozen, Avatar, the Harry Potter series, and the Transformers series have had enormous budgets. One conclusion that can be made from this 10-year set is that the budget of most of the films are below 50 million dollars, and that is especially true for the winners. In accordance with this summary, indicator B can not serve as a good indicator.



4.8 Indicator G:

In the same vane as budget, whether a film succeeded in attracting large audiences and earning large box-office revenue warranted further investigation. To gather data that could be used across all films, required some cleaning of the data. Due to the fact that some of the films are released only a few months before the Oscar ceremony, around the end of November or December of previous year, it would be impossible to compare them with films that had been out almost a full year. Therefore, it appeared more fair to compare one month of box office revenue, with the expectation that films with higher grosses will have a higher chance of winning. However, the plot below shows a very scattered state each year. No tendency between winners and gross could be established, therefore it was concluded that this factor should not contribute as an indicator in the prediction.



Taking each selected indicator as described above and placing them into the formula previously established:

$$P\ (\text{Winner}) = \sum \text{coefficient} * \text{indicator}$$

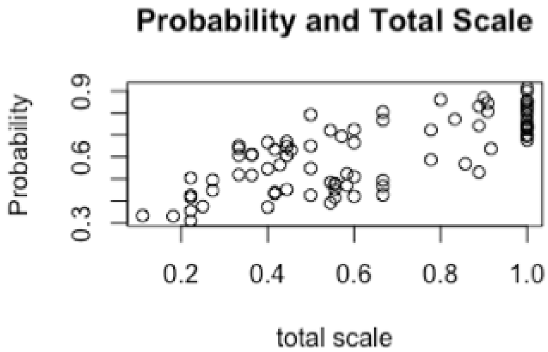
It can be reduced to $P\ (\text{Winner}) = \alpha*N + \beta*D + \gamma*A + \eta*R + \omega*O + \delta*r + \varepsilon$. Where α, β, γ, η, ω and δ are the coefficients for each Indicator, which together sum to 1, and ε is the error, which should be very small. Here, indicator B and G are excluded due to the lack of evidence to support them as a good candidates for the prediction. Apply this in the form of a linear equation with $A*v = <1,1,1,1,1,1>^T$, and $v=<\alpha, \beta, \gamma, \eta, \omega, \delta>^T$. Note that Σcoefficient = 1, where each coefficient is greater than zero, because a coefficient of zero means that indicator has no effect. This would be the case for indicators B and G, which forces them to vanish; Ideally, each coefficient should not be negative, because here each coefficient is a weight for each indicator, so it should be in favor of the prediction. A negative coefficient would be read as devaluing the corresponding indicator, which will have the effect of forcing the overall probability in another direction, diminishing the chance to win. Repeating the process will arrive at a more accurate result; using WolframAlpha.com to do the calculation 10 times and subsequently the vector *v* a couple of times and take the mean value to find the best fit. Using the period from 2005 to 2015, determine the weighted coefficient ranking for each corresponding indicator: $\alpha>\beta>\eta>\delta>\gamma>\omega>0$. After rounding, $\alpha=0.237$, $\beta=0.224$, $\eta=0.208$, $\delta=0.14$, $\gamma=0.111$, $\omega=0.08$. Using this result, the prediction was first tested on the set of years used to generate it. Subsequently it was tested on another set of films from the 10 year period from 1996 to 2005.

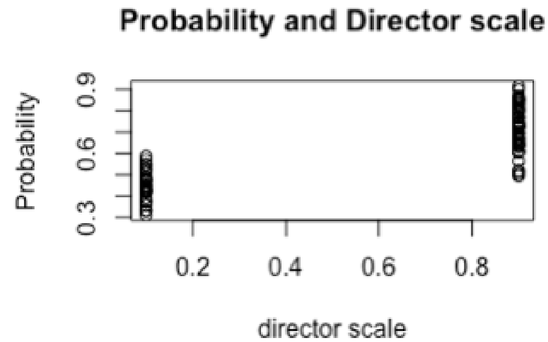
For the set of films from 2006 to 2015, 8 out of 10 times, films that received the highest value corresponded to the winner in that year. The maximum probabilities of winning in this decade is in the range 0.76 to 0.89, with a mean value of 0.82 and standard deviation of 0.038. For the two cases that the winners are not corresponding to the highest P(Win), they still maintained a high P(Win) value, coming in second highest in the year's list. In 2006, Crash is the winner with 0.78 calculated value compared to the expected winner, Brokeback Mountain, with 0.89. Unfortunately, when testing this year’s recipient, it also fails. Once again, though, the winner, Spotlight, ranked second in the list with probability 0.778.

	YEAR	NOMINATION	Predict
1	2016	The Revenant	0.8804534
2	2015	Birdman	0.8422706
3	2014	12 Years a Slave	0.8542801
4	2013	Silver Linings Playbook	0.7606039
5	2012	The Artist	0.8332866
6	2011	The King's Speech	0.8404909
7	2010	The Hurt Locker	0.8222245
8	2009	Slumdog Millionaire	0.8112545
9	2008	No Country for Old Men	0.8476000
10	2007	The Departd	0.8440000
11	2006	Brokeback Mountain	0.8941342
12	2005	The Aviator	0.7647618
13	2004	The Lord of the Ring: The Return of the King	0.7908000
14	2003	Chicago	0.7815379
15	2002	The Lord of the Ring :The Fellowship of the Ring	0.8056000
16	2001	Gladiator	0.8523412
17	2000	American Beauty	0.8969529
18	1999	Saving Private Ryan	0.8265474
19	1998	Titanic	0.8275639
20	1997	The English Patient	0.8298765
21	1996	Braveheart	0.7728000

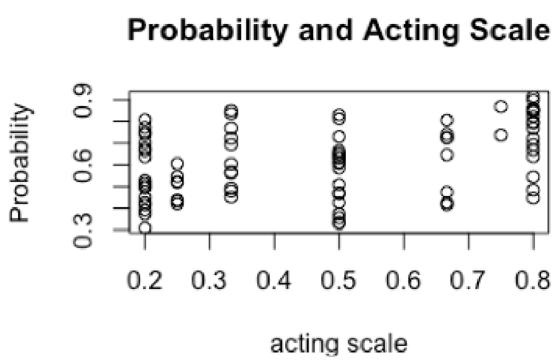
Using the same coefficients to test the second decade set, there are also 8 out of 10 predictions which match the awarded film. For the two mispredictions, the values were very close to the values of the one which was expected to be the winner. For example, in 1999, the actual winner, Shakespeare in Love, had a probability of 0.798, compared to the expected winner, Saving Private Ryan, with a probability of 0.826. The 0.028 gap is still within one standard deviation(0.038), so such a misprediction can almost be ignored.

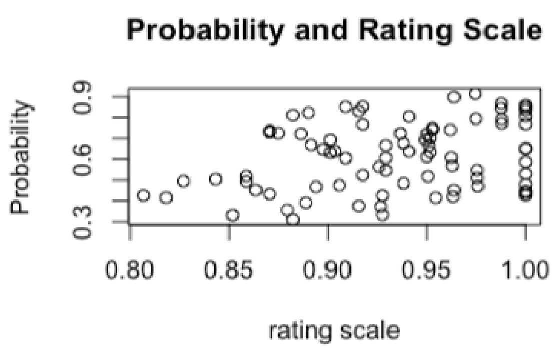
In terms of the correlation between winning probability and each indicator, it is not surprising that some of the indicators serve as good indicators by themselves, considering their the correlation values were close to 1. It turns out that there is a strong correlation(corr = 0.7421049) between winning Best Picture and total number of nominations. Even stronger is the correlation (corr = 0.8136651) with Best Director nominees.

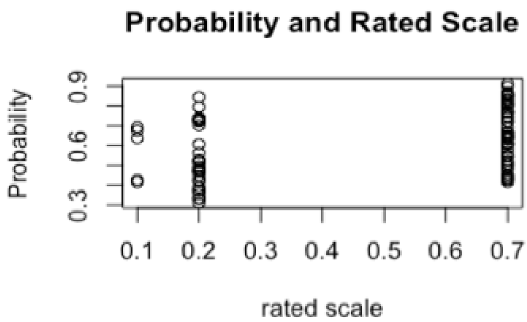


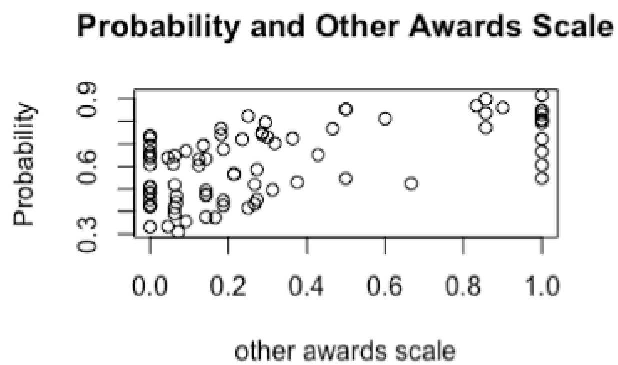


There are also weak correlations between some indicators and winning Best Picture, such as acting (corr = 0.3950825), IMDb rating(0.3133149), and rating by MPAA(corr = 0.395463). While it appeared that the quality of the acting would have some relationship to how the film was received, it also appeared that all the other nominees also had actors or actresses being nominated as well. Comparing this indicator with the indicators for number of nominations(N) and Best Director nominees(D), it seems inessential. The same argument can be made about the indicator: rating by MPAA(r). Additionally, the earlier histogram addressed that the deviation of ratings between nominees and winners was very subtle. Therefore, this indicator does not appear to have a strong relationship.









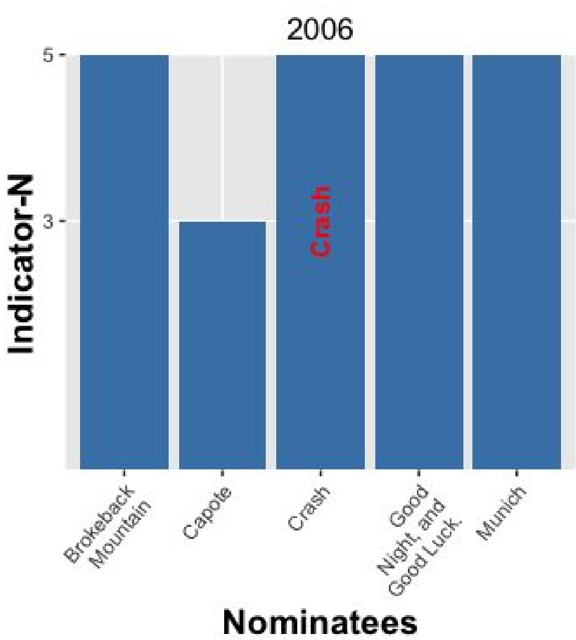
Besides the strong and weak correlations for some indicators, there is also a moderately correlative indicators, such as having other awards (O) (corr = 0.6048236) before the Oscar ceremony. Since this indicator was including many more thant the major four awards that Pardoe and Simonton mentioned, it may have reduced the correlation. This effect may be due to enlarging the set and creating a wider range of variation in number of awards between films.

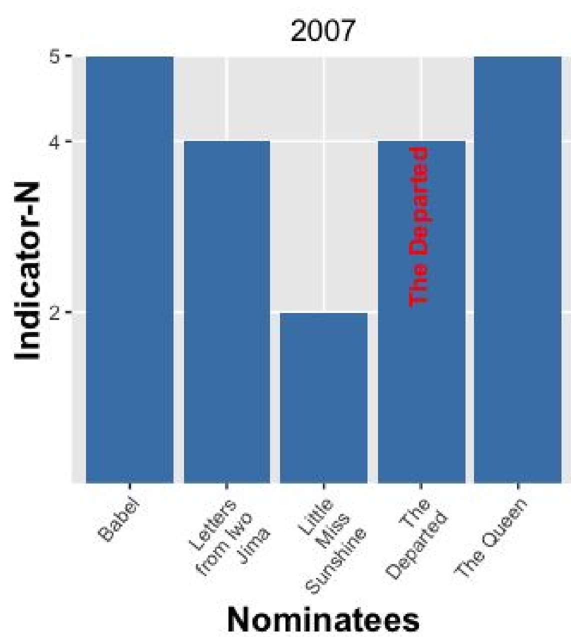
To compare the two different types of models, we begin with the logit model presented in Iain Pardoe and Dean Simonton's article. Their model gives a broad picture in multiple categories, not just Best Picture. The unique coefficients they develop for their variables (indicators) are tailored to the films, and can depend completely on the other factors used in the predictions for that film. In addition their logit model allows for the comparison of two films using the same variable. In contrast, the linear model developed here is adopted from the ideas in Peter Gloor's paper. While it does not benefit from the tailored coefficients, it provides a moderately strong prediction, e.g. 8 out of 10 for the two sets of decades, for the Best Picture category.

5. Restriction:

No prediction is perfect. Mathematical models always rely on some quantifiable circumstances. Sometimes assumptions only behave according to an ideal world, and reality can not always accommodate all the assumptions that have been based on those ideals. The linear model developed here, also has many flaws. Even though the indicator N has a strong correlation with winning, the principle of “correlation does not imply causation” tells us that we should not regard such indicators with strong correlation to regard it as a cause for winning Best Picture.

For instance, in 2006, there were a small number of nominees, four in total; and three of them had the same value for total number of nominations, five for each. In this case, indicator N can not differentiate much between the films. The same issue arises for the year 2007, within the small group in competition for the Best Picture award. The winner, The Departed, can be seen as receiving the second highest N among the group; meanwhile, since there were only five nominees, and two of them had the same N as others, it can also be viewed as The Departed receiving the second lowest N among the group. Therefore, having a high number of N does not guarantee the winning.





The assumption that each indicator is independent from each other might be unrealistic. For example, IMDb rating score(R) might be affected by other awards(O) because these processes can simultaneously build on each other. The total number of nominations(N) might be affected by other awards(O) as well. There is a 0.39 correlation between indicator O and indicator N. However, apply a t-test and the result has a 0.0798 p-value, which is greater than the 0.05 significance value. This is insufficient to say that indicator O and N has a direct dependency.

Nominees for Best Picture that are also nominated in Best Director category (D) is a “to be or not to be” case. That is, it might set up a case which is too extreme, and as a result create an artificially large gap between each nominee. In this model, Argo won the year 2012, and this is one of the five movies among 88 winners in the history that won the Best Picture without also being nominated in the Best Director category. When the P(Argo)=0.52, is compared against the highest P(Silver Linings Playbook)=0.76, this gap is not invisible. It might treated as an outlier at the moment, since as the time shifts, outliers are accumulated, and one day they will become another interesting group to analyze.

The law of large numbers from probability theory preserves the accuracy of the coefficient for each indicator by doing repeated experiments. With improved programming skills, it may have been possible to get a better prediction with automated repetitions of the process. For example, to check accuracy, the coefficient test was applied in the R software environment, and the resulting coefficient is presented below. This is not an exact match to the coefficient from the above calculations that involve the matrices. There is a 0.05207 error which appears. This number is small, but not too small. Finding a more accurate coefficient to diminish this value involves repeating the process.

Coefficients:

(Intercept)	S_r	S_R	S_N
0.05207	0.09449	0.14177	0.25946
	S_O	S_A	D
	0.09574	0.15325	0.19369

In addition to a repeating process, the confines of a data set is also important. Instead of drawing from a fixed, small set of data, if it could be replaced by a random drawing from a much larger set data, such as a span of 50 years, it would provide additional accuracy.

As a fact that there are only 83 winners in the data set, and 88 in the history, so the data can not be large enough to speak about the future winners confidently. Within these 83 winners, the resources are also limited. For example, for some Indicator, such as Rating(R), there was no relevant data for years prior to IMDb widely being used. Additionally, besides those Indicators that can convert to numbers, there are other indicators that are difficult to measure; for instance, some popular topics (war, race, gender, politics, etc.) might hold particular power over the success of those films in a fixed time frame. This type of power would surely reject the assumption that the indicators chosen for this model are the only ones which can generate a prediction for the Best Picture Oscar.

6. Conclusion

In conclusion, it is joy to take this subject as a course project, since the subject of films has always interested me. Additionally, this was the first time that I have had a chance to look at the entire film industry in depth. Applying the knowledge I learned from probability, linear algebra, and different methods of modelling from this seminar class has expanded my understanding of probability in the real life. For further study, in addition to linear models, there are other models such as Markov chain, Massey's method, Colley's, that could be evaluated for their power in making predictions in this area. Beyond the scope of modeling, opinions from other fields that do not involve any math or programming are still valid. However, I think here we are bridging the gap between the world of the arts and the world of science.

Works Cited

Bettinger, Brendan. "*Cinemath: What Makes a Best Picture, A Look at Rating, Runtime, and Genre Over 80 Years of Oscars"* (2011): Web.

Gloor, Peter. "*Predicting Movie Success and Academy Awards through Sentiment and Social Network Analysis"* (2009): Journal of *16th European Conference on Information Systems.* ECIS2008-SOC-REG-014.R1. Print.

Pardoe, Iain. "Just How Predictable Are the Oscars?" *CHANCE* 18.4 (2005): 32–39. *Taylor and Francis+NEJM*. Web.

Pardoe, Iain. "Predicting Oscar Winners." *Significance* 4.4 (2007): 168–173. *EBSCOhost*. Web.

Pardoe, Iain, and Dean K. Simonton. "Applying Discrete Choice Models to Predict Academy Award Winners." *Journal of the Royal Statistical Society. Series A (Statistics in Society)* 171.2 (2008): 375–394. Print.