Findings and Curiosities from Published Movie Data (Wallisch & Whritner, 2017)

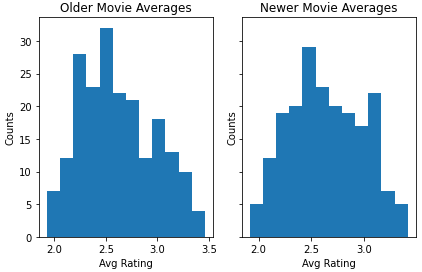
as reported by: Guilherme Albertini

Summary

Several hypothesis tests were run to address concerns posed by Management using a dataset that queried 1097 participants on 400 films using a Likert scale. The individual responses for each film rating spanned an ordinal measure of 0.0 for finding a film wholly unsatisfactory to 5.0 for personal favorites. Half-point increments were allowed. Additionally, several questions were posed to participants and will be discussed below. To cut down on false positives, the per-test significance level 𝛼 is set to 0.005 (as per Benjamin et al., 2018).

Findings

*Are movies that are more popular (operationalized as having more ratings) rated higher than movies that are less popular?* To answer this question, a median-split of movie popularity produced two datasets of ratio measure that were assumed to be independent for a Welch’s t-test, as variances were not equal for a two-sample t-test. As the p-value was near zero at the level of significance, I found that there is *insufficient evidence to suggest that that movies that are more popular have ratings that are higher than movies that are less popular.*

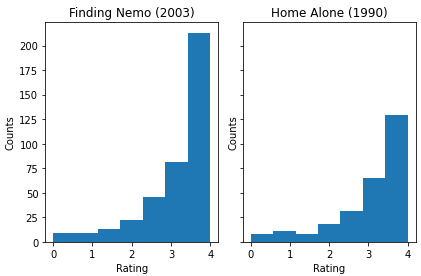
*Are movies that are newer rated differently than movies that are older?* A median-split of movie release years generated two datasets of ratio measure: one for newer films and one for older films. Using the Welch’s t-test (variances not assumed equal) the p-value of 0.552 was greater than the critical value; thus, I found that t*here is insufficient evidence to suggest that that newer movies are rated differently than older films.*

*Is enjoyment of Shrek (2001) gendered, i.e. do female and male viewers rate it differently? What proportion of films show a gendered effect?* Here, a non parametric test (Mann-Whitney U, or Wilcoxon rank-sum test) had movie ratings split between males and females for one of our best-performing films. As the p-value of 0.044 was greater than the critical value, there was i*nsufficient evidence to claim that the enjoyment of Shrek (2001) is gendered. In fact, using the same methodology, only 13.25% (53 films) show gendered preferences.*

*Do people who are only children enjoy ‘The Lion King (1994)’ more than people with siblings? What proportion of films show this possible single-child effect?* To test the hypotheses, the U-test was once again used for ordinal data: one set for the ratings from only children and one for ratings of those with siblings . For the Lion King (1994), I found that the p-value of 0.978 was greater than the critical value, so there is *insufficient evidence to suggest that only children enjoy the film more than those with siblings. In fact, only 1.75% (7 films) show a single-child effect.*

*Do people who like to watch movies socially enjoy ‘The Wolf of Wall Street (2013)’ more than those who prefer to watch them alone? What proportion of these films show this possible social watching effect?* Another U-test performed on the datasets (one for those who rated the film and preferred watching it alone, and the other for those preferring to view with others) showed that the p-value of 0.944 was greater than the critical value and thus there is insufficient evidence to suggest that people enjoy *The Wolf of Wall Street (2013) more than those who prefer to watch them alone. At large, only 2.5% (10 films) showed a social watching effect.*

*Is the ratings distribution of ‘Home Alone (1990)’ different than that of ‘Finding Nemo (2003)’?* To test whether the claim that the two distributions are different, a Kolmogorov–Smirnov hypothesis test was run for the two distributions of ratings data. *As the p-value is near 0, there is ample evidence to suggest that the distributions are indeed different.*

 *There are ratings on movies from several franchises ([‘Star Wars’, ‘Harry Potter’, ‘The Matrix’, ‘Indiana Jones’, ‘Jurassic Park’, ‘Pirates of the Caribbean’, ‘Toy Story’, ‘Batman’]) in this dataset. How many of these are of inconsistent quality, as experienced by viewers?* In answering this question, the Kolmogorov-Smirnov test was run through all the franchise films, analyzing whether the distributions of ratings (a measure of quality) were indeed from the same distribution. The test uncovered that *only the Harry Potter franchise showed consistent quality ratings across its films.*

*As an additional question posed by Management for the upcoming season: Do men enjoy horror thrillers more than women?* To conduct such a test, I opt for the U-test once more for the horror thriller films included below. After performing this test, I find that o*nly 3 of these horror thrillers (about 9.1%) were enjoyed more by males than females; even if we double the per-test level of significance, only 4 of these horror thrillers (about 12.1%) seem to be enjoyed more by males than females.*

*Films considered:*

*'Alien (1979)','Alien (1979)','Black Swan (2010)','A Nightmare on Elm Street (1984)','Psycho (1960)','Night of the Living Dead (1968)', 'The Silence of the Lambs (1991)', 'The Others (2001)', 'The Texas Chainsaw Massacre (1974)', 'The Omen (1976)', 'Poltergeist (1982)', 'Poltergeist (1982)', 'The Conjuring (2013)', 'The Thing (1982)', 'The Shining (1980)', 'Shutter Island (2010)', 'Carrie (1976)', "Rosemary's Baby (1968)", 'The Evil Dead (1981)',The Blair Witch Project (1999)', 'Cloverfield (2008)', "Bram Stoker's Dracula (1992)", 'What Lies Beneath (2000)', 'Friday the 13th Part III (1982)', 'Hannibal (2001)', 'American Psycho (2000)', 'Predator (1987)', 'Aliens (1986)', "You're Next (2011)", 'Saw (2004)', 'Halloween (1978)', 'The Mist (2007)', 'The Descent (2005)'*