Real World Mini-Project 2: Modeling Assumptions

The tools of this class are useful to computer scientists, but many of them are useful beyond just "classic" computer science. In order to use the powerful tools of probability, we need to make assumptions to let our mathematical tools model the real world. Things like "this coin is perfectly fair" or "the coin flips are all independent." These are usually not perfectly true[1]. Indeed, occasionally these assumptions are ways that people "lie with statistics" or provide evidence for claims that aren't actually true.

In this project you will critique the modelling assumptions made in an analysis and see if other modeling assumptions would lead to a different result.

THIS PROJECT IS DELIBERATELY OPEN-ENDED. PLEASE VISIT THE STAFF AT OFFICE HOURS EARLY FOR BRAINSTORMING IDEAS.

Submission: You must upload a **pdf** of your written solutions to Gradescope under "Real World Mini-Project 2". The use of LaTeX is *highly recommended*. (Note that if you want to hand-write your solutions, you'll need to scan them. We will take off points for hand-written solutions that are difficult to read due to poor handwriting and neatness.)

Please put each numbered problem on its own page of the pdf (this will make selecting pages easier when you submit), and ensure that your pdfs are oriented correctly (e.g. not upside-down or sideways).

Remember that you must tag your written problems on Gradescope.

Due Date: This assignment is due on Wednesday, August 11 at 11:59 PM (Seattle time, i.e. GMT-7).

Collaboration Policy: You are to conduct your own search and analysis for this assignment. While you may get feedback from other students on your writing, you cannot just use the results of another student's search.

1. Find a Model

Find an analysis that uses probability and statistics tools you're familiar with from this course. By "analysis," we mean any estimate of a "real-world" probability, along with the assumptions that lead to that number. You might want to look at the examples in the final section for what we mean.

We expect most of the answers to this section will be short (2-3 sentences), but you are free to write more if your resource is more complicated.

(a) Provide a link to (or somehow let us access) the analysis you're critiquing.

(b) What is the fundamental claim of the analysis? I.e., what conclusion do they draw at the end of their analysis?

(c) What modelling assumptions do they use? (For example, do they assume some occurrences are independent? Do they assume a set of events all have equal probability? Do they assume they know the probability? Do they use a variable from the zoo?)

2. Improve the model

Now, see if their modelling assumptions are reasonable or if other ones would lead to a different conclusion. We expect parts a, b, d will be a few sentences each (though you can write more if you have more to say).

(a) Identify at least one weakness of the modelling assumptions they have made (e.g. a potential dependence on events that are supposed to be independent).

[1] e.g., if you flip a coin repeatedly, the result of the last flip is probably how the coin will appear on your hand before you flip it, which will make the results not quite independent.

(b) Now create your own model for the same problem. You might do this by coming up with different probability estimates for the events, or by using a different random variable (e.g. a binomial distribution instead of a Poisson), or by incorporating some outside knowledge about the problem that you think sheds more light. Briefly describe what your model will be, and how it differs from the previous one.

(c) Under your new model, calculate the probability of the event your source calculated.

(d) Does the calculation change significantly? If it does, does the conclusion of the analysis change?

3. Sample Analysis

(a) (i) The dataset we use is about the upcoming Chess championship: Carlsen vs. Nepomniachtchi: What Do The Numbers Say?

(ii) The analysis is about the upcoming 2021 world chess championship tournament in which current title-holder Magnus Carlsen will play challenger Ian Nepomniachtchi. The tournament works as follows: they play 14 games, and the winner of each game wins a point, whereas a tie is worth 0.5 points. A player needs 7.5 points to win the series. The article concludes that, based on elo-ratings, Magnus has a 72% chance of winning (this assumes they do not tie in the 14-game series).

(iii) The article's final analysis assumes that the FIDE rating, which aggregates scores between a player's entire match career, is a good predictor of future performance against a specific opponent. They assume that each future match is a direct consequence of current FIDE rating - e.g. that each previous match result is independent of future matches given the FIDE rating of both players.

(b) (i) One thing the analysis does not take into account is past player to player match history, which is not necessarily captured by overall FIDE rating, which summarizes a player's total past performance against all opponents. That is, a future match between Nepo and Carlsen may not be independent of past matches between those two players given the current FIDE rating of both players, because personal play style factors into match results.

(ii) In previous head-to-head matches, Nepo has a 4-1-6 match history against Magnus in the classical setting, meaning that he has won against Magnus 4 times, lost once, and drawn 6 times in classical games. If we take into account only personal history, Nepo has a $\frac{4}{11} \approx 0.36$ chance to win, $\frac{1}{11} \approx 0.1$ chance to lose and $\frac{6}{11} \approx 0.54$ chance to tie.

(iii) We will assume that each game has an independent outcome of 0.54 chance to tie, 0.1 chance for a win by Magnus, and a 0.36 chance to win by Nepo.

(iv) To compute the probability that Magnus wins the overall 14-game series, we need to break it down into all the possible ways he can win. If he needs 7.5 points that means: he can get at least 8 wins, or he can get 7 wins and at least 1 tie, or 6 wins and at least 3 ties ... and so on. In order to write this out, we will use $B(n, p, x)$ to denote the PDF of a binomial random variable with parameters n, p is equal to x and $B_{\geq}(n, p, x)$ to denote the probability that a binomial random variable with parameters n, p is greater than or equal to x. We will further let p_w, p_t, p_l denote the probabilities that Magnus wins, ties and loses a single game respectively. Using this notation, the probability that Magnus wins is:

$$B_{\geq}(14, p_w, 8) + \Sigma_{i=1}^{7} B(14, p_w, i) \cdot B_{\geq}(14, p_t, 15 - 2i)$$

That is, the probability that he either wins 8 or more games outright, or the probability that he wins exactly some number of games between 1 and 7 and ties at least enough games to total 7.5 point (remember a tie is worth 0.5 points and a win is worth 1 point). Using our computed values of $p_w = 0.1$ and

$p_t = 0.54$ from the previous part, we can use a computer to calculate the above equation to get approximately 0.08, while the probability of Nepo winning the championship would be 0.77. If we condition on the fact that they don't tie, the probability that Magnus wins is $\frac{0.08}{0.08 + 0.77} \approx 0.09$

The program used to calculate the above probabilities:

```
from scipy.stats import binom

def B(n, p, x):
    return binom.pmf(x, n, p)

def B_g(n, p, x):
    return 1.0 - binom.cdf(x - 1, n, p)

def compute_win_prob(num_games=14, p_w=0.1, p_t=0.54):
    score = B_g(num_games, p_w, num_games // 2 + 1)
    for i in range(1, num_games // 2 + 1):
        score += B(num_games, p_w, i) * \
                 B_g(num_games, p_t, num_games + 1 - 2*i)
    return score

magnus_win_prob = compute_win_prob()
nepo_win_prob   = compute_win_prob(p_w=0.36)
tie_prob        = 1.0 - magnus_win_prob - nepo_win_prob
print(magnus_win_prob, nepo_win_prob, tie_prob)
```

The calculation definitely makes the likelihood of a win by Nepo much more likely — essentially it says that Nepo will crush Magnus. In reality, this is unlikely as we are basing our model on **very** little data (11 games is a very small sample size). For example, our model does not take into account the FIDE rating which is the overall stats for a chess player against all opponents. A more balanced and possibly accurate model would take into account both FIDE rating and past history.

4. Some Ideas

You are free (and encouraged!) to find your own examples outside this list if you have a topic you are passionate about, but if you can't think of anything, you may use any of these as starting points. In many cases, there are already critiques of poor statistical/probability analyses online — it's ok to look at these critiques, as long as you tell us if you're using any and still do the new probability calculation independently and put everything in your own words.

- Every year millions of people predict the outcomes of the NCAA men's basketball tournament. It is commonly said that the probability of a perfect bracket is $\frac{1}{2^{63}}$, (since there are 2^{63} ways the 63 games could play out) and therefore no one will ever predict a perfect bracket. Here is a source using that number
- Shortly after the 2020 presidential election, there were many assertions (including by people with PhD's...) that the probability of the election night shift was so low as to be impossible. Page 20 here is one example.
- A video-game streamer named "Dream" did a speedrun of Minecraft where they had incredible luck in a few parts of the game. So lucky, that a speedrunning organization declared that Dream had to be using a modified version of the game, and that the run was therefore invalid. The analysis that lead to the rejection of the run.
- For all MCU fans, there is an analysis of deaths and revival of Avengers. As an Avenger, Dying Is Part of The Job
- One might be wondering how dangerous lightning really is and if you can die. How Dangerous is Lightning?
- Navigating through the board game of life to make a Monopoly stake and become a real estate tycoon may lead to a visit to prison occasionally. Probability of Going to Jail in Monopoly