Abstract

Using publicly available data sources, this study demonstrates that a test positive COVID-19 individual has a 66.43% chance of being from a low-income county of Washington state.

Income and Race Impact on COVID-19 Positive Cases using Conditional Probability

Claire Gong

March 2021

Contents

1 Introduction

COVID-19 is a contagious disease that can be caught by multiple factors. In this study, I want to focus on two factors: income and race.

If the patient is positive, what is the probability that the patient comes from a state with lower income?

If the patient is positive, what is the probability that the patient comes from a state with a certain race population?

When deciding the scope of this study, I have chosen to limit it to a state level instead of a country level. The data drawn and the comparisons made for this study is exclusive to Washington state and its counties in order to limit the amount of external factors that would influence people's chances of catching COVID-19. Such factors include state protocols such as mask mandates or business opening capacity restrictions.

2 Formulation

In the formulation, let H = high income and C = a positive COVID-19 case. The joint probability of H and C is equal to the conditional probability of H

given C multiplied by the probability of C.

$$P(C, H) = P(C|H)P(H) = P(H, C) = P(H|C)P(C) \tag{1}$$

Here, is the joint probability of C and H:

$$P(C, H) \tag{2}$$

The probability of C given H:

$$P(C|H) \tag{3}$$

The probability of H:

$$P(H) \tag{4}$$

The probability of H given C:

$$P(H|C) \tag{5}$$

The probability of C:

$$P(C) = P(C|H)P(H) + P(C|\neg H)P(\neg H) \tag{6}$$

where $\neg$ is the logical negation.

This study requires us to find the probability of H given C:

$$P(H|C) = \frac{P(C|H)P(H)}{P(C|H)P(H) + P(C|\neg H)P(\neg H)} \tag{7}$$

3 Data

The positive COVID-19 cases by county data is reported by the New York Times as of March 23, 2021. [1] To clarify, we consider low-income to be under the US median household income, in which according to the US Census Bureau, is $68,703 in 2019. Household earnings over that number will be considered as high-income. [2] To determine the probability of high income households, I used the population (footnote ...) and the median household income of each Washington county. [3]

4 Computation

The probability of getting COVID-19 in Washington state, EQ-6, was calculated by dividing the total number of Washington COVID-19 cases (358,762) as of March 23, 2021 with the Washington state population (7,404,107). $P(C) = 0.0485$

[1] https://www.nytimes.com/interactive/2020/us/washington-coronavirus-cases.html

[2] https://www.census.gov/library/publications/2020/demo/p60-270.html

[3] https://ofm.wa.gov/washington-data-research/economy-and-labor-force/median-household-income-estimates

The probability of a high-income household, EQ-4, was calculated by adding the population's of the county's that had a high-income (4,985,670) and dividng that by the total Washington population (7,404,107) which resulted in $P(H) = 0.6734$.

The probability of a testing positive for COVID-19 given the individual is from a high-income household, EQ-3, was computed by adding COVID-19 cases of counties that were considered to have an average high-income household (204,062) and dividing that with the total amount of COVID-19 cases in Washington state (358,762). $P(C|H) = 0.5688$

Therefore, by using EQ-7:

$$P(H|C) = \frac{0.5688 \times 0.6734}{0.5688 \times 0.6734 + 0.4312 \times 0.3266} = 0.3357 \tag{8}$$

5 Conclusion

The probability of being a high-income household county given a positive COVID-19 case is 0.3357. In other words, there is a 33.57% chance a test positive COVID-19 individual is from a high-income county. This means that there is a 66.43% chance of a test positive COVID-19 individual that is from a low-income county.