

### **Abstract**

There has been recent media attention on the phenomenon of the low-income being linked to higher COVID-19 test positive rates. I wanted to see if that was the case in Washington state. Using publicly available data sources, this study demonstrates that a test positive COVID-19 individual has a 26.88% chance of being from a low-income county of Washington state. Please note that the low-income represents a smaller population in Washington state (32.66%).

# Income Impact on COVID-19 Positive Cases using Conditional Probability

Claire Gong

March 2021

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Formulation</b>	<b>2</b>
<b>3</b>	<b>Data</b>	<b>3</b>
<b>4</b>	<b>Computation</b>	<b>3</b>
4.1	Fixed Income Threshold . . . . .	3
4.2	Variable Income Threshold . . . . .	4
<b>5</b>	<b>Conclusion</b>	<b>4</b>
<b>6</b>	<b>Appendix</b>	<b>4</b>

## 1 Introduction

COVID-19 is a contagious disease that can be caused by multiple factors. In this study, I want to focus on the following factor: income disparities.

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

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.

$$P(C, H) = P(C|H)P(H) = P(H, C) = P(H|C)P(C) \quad (1)$$

Here, is the joint probability of C and H:

$$P(C, H) \quad (2)$$

The probability of C given H:

$$P(C|H) \quad (3)$$

The probability of H:

$$P(H) \quad (4)$$

The probability of H given C:

$$P(H|C) \quad (5)$$

The probability of C:

$$P(C) = P(C|H)P(H) + P(C|\neg H)P(\neg H) \quad (6)$$

where  $\neg$  is the logical negation.

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

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

### 3 Data

The positive COVID-19 cases by county data is reported by the New York Times as of March 23, 2021. <sup>1</sup> 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. <sup>2</sup> To determine the probability of high income households, I used the population (footnote ...) and the median household income of each Washington county. <sup>3</sup>

## 4 Computation

### 4.1 Fixed Income Threshold

The fixed income threshold selection process is described in the previous section. 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

---

<sup>1</sup><https://www.nytimes.com/interactive/2020/us/washington-coronavirus-cases.html>

<sup>2</sup><https://www.census.gov/library/publications/2020/demo/p60-270.html>

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

March 23, 2021 with the Washington state population (7,404,107).  $P(C) = 0.0485$

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 dividing 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:

$$P(H|C) = \frac{0.5688 \times 0.6734}{0.5688 \times 0.6734 + 0.4312 \times 0.3266} = 0.7312 \quad (8)$$

## 4.2 Variable Income Threshold

Given that the income threshold is fixed, I wanted to study the  $P(\neg H|C)$  given the variable threshold. Specifically in 5% increments from 0 to 100 percent of the maximum median household income county in Washington state (\$ 89,881). See the appendix for a table (Figure-1) and graph (Figure-2) that elaborate on the relationship of the probability of coming from a low-income county given a positive COVID-19 case ( $P(\neg H|C)$ ).

## 5 Conclusion

The probability of being a high-income household county given a positive COVID-19 case is 0.7312. In other words, there is a 73.12% chance a test positive COVID-19 individual is from a high-income county. This means that there is a 26.88% chance of a test positive COVID-19 individual that is from a low-income county.

## 6 Appendix

median household income threshold	P(not H C)
4494	0.0000
8988	0.0000
13482	0.0000
17976	0.0000
22470	0.0000
26964	0.0000
31458	0.0000
35952	0.0000
40446	0.0000
44941	0.0001
49435	0.0140
53929	0.0626
58423	0.0892
62917	0.1161
67411	0.2334
71905	0.5299
76399	0.6397
80893	0.7452
85387	0.7452
89881	0.8795

Figure 1: Table of probability of coming from a low-income county given a positive COVID-19 case ( $P(\neg H|C)$ ) as function of income threshold

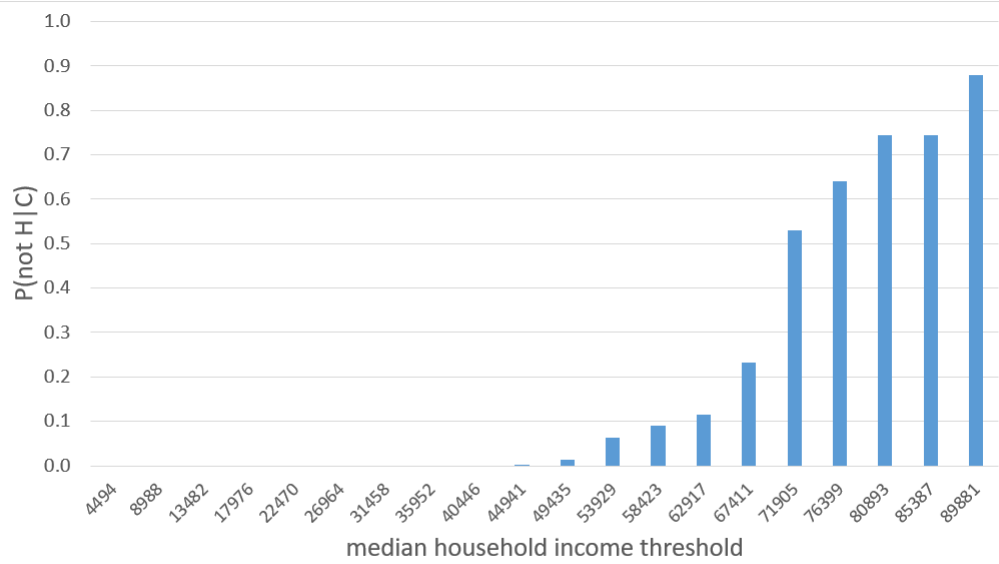


Figure 2: Chart of probability of coming from a low-income county given a positive COVID-19 case ( $P(\neg H|C)$ ) as function of income threshold