Umasathvik Pantham

Efficacy of Vaccines vs. Different Variants of COVID-19

Dataset

	Effectiveness at preventing											
	Ancestral		Alpha		Beta		Gamma		Delta		Omicron	
Vaccine	Severe disease	Infection	Severe disease	Infection	Severe disease	Infection	Severe disease	Infection	Severe disease	Infection	Severe disease	Infection
AstraZeneca	94%	63%	94%	63%	94%	69%	94%	69%	94%	69%	71%	36%
CanSino	66%	62%	66%	62%	64%	61%	64%	61%	64%	61%	48%	32%
CoronaVac	50%	47%	50%	47%	49%	46%	49%	46%	49%	46%	37%	24%
Covaxin	78%	73%	78%	73%	76%	72%	76%	72%	76%	72%	57%	38%
Johnson & Johnson	86%	72%	86%	72%	76%	64%	76%	64%	76%	64%	57%	33%
Moderna	97%	92%	97%	92%	97%	91%	97%	91%	97%	91%	73%	48%
Novavax	89%	83%	89%	83%	86%	82%	86%	82%	86%	82%	65%	43%
Pfizer/BioNTech	95%	86%	95%	86%	95%	84%	95%	84%	95%	84%	72%	44%
Sinopharm	73%	68%	73%	68%	71%	67%	71%	67%	71%	67%	53%	35%
Sputnik-V	92%	86%	92%	86%	89%	85%	89%	85%	89%	85%	67%	44%
Other vaccines	75%	70%	75%	70%	73%	69%	73%	69%	73%	69%	55%	36%
Other vaccines (mRNA)	91%	86%	91%	86%	88%	85%	88%	85%	88%	85%	67%	45%

	Ancestral	Alpha	Beta	Gamma	Delta	Omicron	Average
AstraZeneca	63	63	69	69	69	36	61.5
CanSino	62	62	61	61	61	32	56.5
CoronaVax	47	47	46	46	46	24	42.67
Covaxin	73	73	72	72	72	38	66.67
J&J	72	72	64	64	64	33	61.5
Moderna	92	92	91	91	91	48	84.17
Novavax	83	83	82	82	82	43	75.83
Pfizer/BNT	86	86	84	84	84	44	78
Sinopharm	68	68	67	67	67	35	62
Sputnik-V	86	86	85	85	85	44	78.5
Average	73.2	73.2	72.1	72.1	72.1	37.7	Total Avg: 66.73

Methods

Q-Q Plot, 2-Way Anova w/o Replication, F-Test, Tukey's Procedure

Problem

We want to see if there is a major difference between vaccine efficacy against different variants of COVID 19.

Abstract

There are many types of COVID vaccines in the world, and while they all serve the same purpose, the efficacy of each vaccine varies for a number of reasons. The goal of this research paper is to analyze the degree to which each vaccine performs against multiple variants of COVID-19. This paper looks over 10 different types of vaccines and 5 different variants of the virus. To achieve this, we carried out an F test with a significance level of 0.05, with our findings shown on an ANOVA table. Tukey's method was also used to compare different variants and their effects on the vaccines. After removing the Omicron data set, our findings showed that their is no significant difference in vaccine efficacy between different variants of COVID 19. The overall importance of this study shows whether or not different COVID variants affect vaccine efficacy.

Introduction

The United States is racing to vaccinate a population of 328 million (JHU), but local health offices are progressing at vastly different rates. As more data on COVID-19 vaccines became available, statisticians began conducting an ongoing systematic review of studies related to vaccine efficacy. Vaccine efficacy is measured as (1) a vaccine's efficacy at stopping transmission of the virus from one person to another, meaning an exposed person will not contract the virus and they will also not develop symptoms or disease and (2) a vaccine's efficacy at preventing an exposed person from developing serious symptoms that often require hospitalization and lead to death. In this project, we will be analyzing the effectiveness of major vaccines by percentages against the different variants of COVID-19 and their severity in order to

determine if there is a major difference between vaccine efficacy against different variants of COVID 19.

Data Description

The dataset that we have for this project consists of various vaccines that have been scientifically tested to figure out their efficacy for each and every variant of COVID that we've had so far. The data is experimental as various producers of the vaccines conducted multiple tests and procedures to come up with the exact efficacy rates for each variant. To further analyze, efficacy is the performance of a vaccine under clinical trials under certain conditions such as different variants. To expand on this, the dataset is divided into rows and columns where each row is a different type of vaccine (ex. Moderna, CanSino, etc.) whereas each column is a different type of Covid variant (ex. Ancestral,Delta, etc.). Through these data points for each and every specific combination, we were able to figure out the averages of efficacy for each variant, vaccine, and the combined total.

Modeling

In this project, we want to test if different variants of COVID 19 affect vaccine efficacy and to what extent. We assume there is no difference between the effectiveness of the different vaccines and the different types of COVID 19. Calculations involving ANOVA and summary statistics were done in excel. In order to test for normality, we plotted all data points in a QQ plot to test whether or not the data follows a straight line. We also check whether or not the data was normally distributed by plotting a histogram of the data. If both graphs showed a normal distribution we concluded that the data fits the model assumptions of normality.

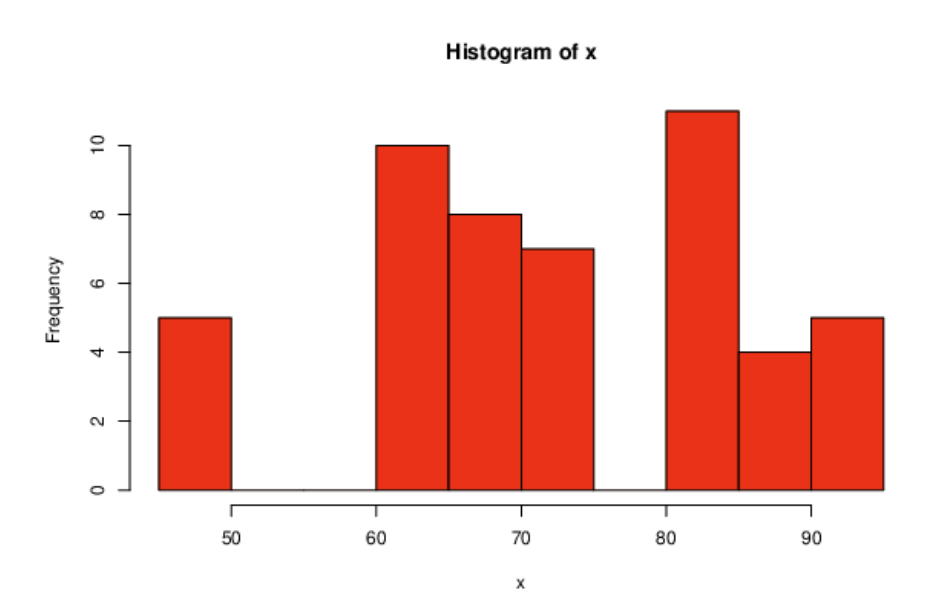
We also want to test to see which variants of COVID 19 are significantly different in changing the effectiveness of the different types of vaccines. Using a Significance Level of

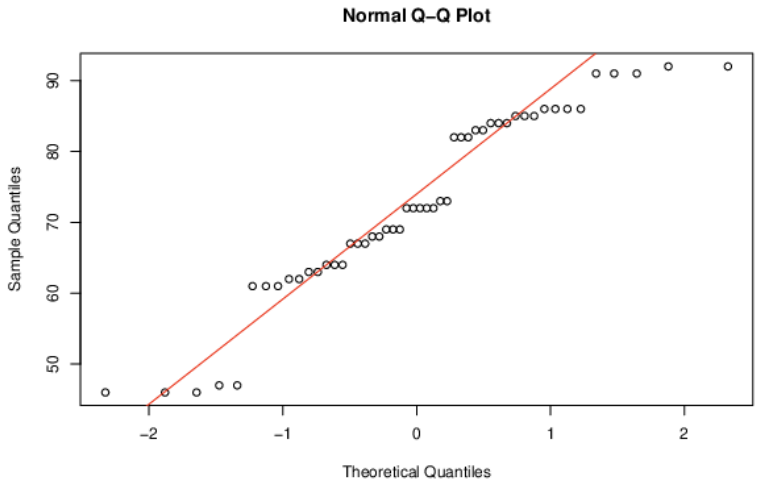
$\alpha = 0.05$ and other relevant information, we are able to see which variants of COVID 19 have significantly different effects on vaccine efficacy.

Results

Assumptions of Normality:

After cleaning out the data (removing Omicron Variant), we can see that the data set is approximately normally distributed. Therefore we can conduct a 2 way anova test without replication to test whether or not different variants of the COVID 19 vaccine affect vaccine efficacy.





We plotted the data points using a histogram and when looking at it, it can be seen that it fits the characteristics of a normal distribution graph. To further establish this point, we created a Normal Q-Q Plot where when analyzing it, we can see that it fits the assumptions of normality because the points follow a straight line throughout the graph.

Sidenote: Calculated and graphed using http://www.wessa.net/rwasp_varia1.wasp

Hypothesis:

$H_o: \beta_{ancestral} = \beta_{alpha} = \beta_{beta} = \beta_{gamma} = \beta_{delta} = 0$

$H_a: at\ least\ one\ \beta_j \neq 0$

Significance Level: $\alpha = 0.05$

ANOVA Table:

ANOVA						
Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Vaccine	8331.22	9	925.6911	280.7961	1.38E-30	2.152607
Varriant	14.52	4	3.63	1.101112	0.370961	2.633532
Error	118.68	36	3.296667			
Total	8464.42	49				

(See Excel Document for Work)

Significance Tests (Tukey's Method) w/o Omicron:

$Q_{0.05,10,36} = 4.7762$ (According to the Q-Table)

$w = Q_{0.05,10,36} \times \sqrt{\frac{MSE}{I}} = 4.7762 \times \sqrt{\frac{3.297}{5}} = 3.878$

No variants are Significantly Different.

(See Calculations Below)

Conclusion

Through the tests we conducted through ANOVA and Tukey's Method, we can confidently say that different variants of COVID do not affect vaccine efficacy rates (Fail to Reject the Null Hypothesis) as long as it doesn't include the Omicron variant. From our data and computation of our *p-value*, we can conclude that because our *p-value* is greater than our significance level, thus we should fail to reject the null hypothesis. With further testing through Tukey's method, we can also defend our result of failing to reject the null hypothesis because there are no variants that are significantly different from each other. This means that, in terms of our research, different COVID-19 variants do not affect vaccine efficacy.

References:

1) https://www.healthdata.org/covid/covid-19-vaccine-efficacy-summary
2) COVID-19 vaccine doses administered by manufacturer, United States (ourworldindata.org)
3) Understanding Vaccination Progress - Johns Hopkins Coronavirus Resource Center (jhu.edu)

Calculations:

Tukey Calculations:

$Q_{0.05,10,36} = 4.7762$ (According to the Q-Table)

$w = Q_{0.05,10,36} \times \sqrt{\frac{MSE}{I}} = 4.7762 \times \sqrt{\frac{3.297}{5}} = 3.878$

$\bar{x}_{.ancestral}$ and $\bar{x}_{.alpha}$: 73.2-73.2 = 0 < 3.878

$\bar{x}_{.ancestral}$ and $\bar{x}_{.beta}$: 73.2-72.1 = 1.1 < 3.878

$\bar{x}_{.ancestral}$ and $\bar{x}_{.gamma}$: 73.2-72.1 = 1.1 < 3.878

$\bar{x}_{.ancestral}$ and $\bar{x}_{.delta}$: 73.2-72.1 = 1.1 < 3.878

$\bar{x}_{.alpha}$ and $\bar{x}_{.beta}$: 73.2-72.1 = 1.1 < 3.878

$\bar{x}_{.alpha}$ and $\bar{x}_{.gamma}$: 73.2-72.1 = 1.1 < 3.878

$\bar{x}_{.alpha}$ and $\bar{x}_{.delta}$: 73.2-72.1 = 1.1 < 3.878

$\bar{x}_{.beta}$ and $\bar{x}_{.gamma}$: 72.1-72.1 = 0 < 3.878

$\bar{x}_{.beta}$ and $\bar{x}_{.delta}$: 72.1-72.1 = 0 < 3.878

$\bar{x}_{.gamma}$ and $\bar{x}_{.delta}$: 72.1-72.1 = 0 < 3.878