

Vaxelence: Statistical Analysis of COVID-19 Vaccine Efficacy

Frequentist, Bayesian, and Bootstrap Estimation

Jaiden Atterbury, Tanner Huck, Andy Wen

1 Abstract

In this paper, we investigated the efficacy of the Pfizer COVID-19 vaccine using data collected from the Pfizer vaccine trial from 2020. In particular, we used three avenues of inference to validate/analyze the Pfizer data, these three methods being: Frequentist inference, Bayesian inference, and bootstrapping the sampling distribution of vaccine efficacy. In order to see if the Pfizer vaccine met the FDA standards of at least 30 efficacy, we found point estimates, confidence intervals, and ran hypothesis tests. Through a skeptical lens, we used an α level of 0.01 and found that the results were robust throughout all methods. In particular, the estimated vaccine efficacy was in the lower to mid 0.9 range for all three methods and the results were all extremely significant at the given alpha level. Lastly, from our results we conclude that the experimenters at Pfizer used an uninformative prior in their analysis.

2 Keywords & Introduction

Efficacy, COVID-19 Vaccine, Frequentist, Bayesian, Bootstrap

Coronavirus disease 2019 (COVID-19), an infectious disease caused by SARS-CoV-2 virus, emerged in 2019, leading to a global pandemic ⁴. The disease's symptoms range from mild to severe, necessitating medical attention in serious cases. The World Health Organization emphasizes that anyone can succumb to severe symptoms or death, irrespective of age ⁴. Amid this crisis, the development of COVID-19 vaccines has played an integral role in mitigating the disease's spread

and impact. Efficacy, indicating a drug’s capacity to produce an effect, is pivotal in medical trials². In the context of vaccines, it represents the ability to induce desired immunological responses to prevent the disease. A significant study titled “Safety and Efficacy of the BNT162b2 mRNA COVID-19 Vaccine” was conducted by the ¹. This research involved a large participant pool and aimed to assess the BioNTech and Pfizer-developed COVID-19 vaccine’s effectiveness. The study’s outcomes underscored the vaccine’s high efficacy and consistent performance across varied demographic groups¹. Our research intends to reanalyze this data set with a primary focus on estimating the Pfizer vaccine’s efficacy. We aim to back these estimates with confidence intervals/credible intervals and hypothesis tests through three methods: Frequentist inference, Bayesian inference, and simulation. A specific area of interest is to determine whether the vaccine’s efficacy surpasses the FDA’s minimum required threshold of 30% for approval.

3 Statistical Methods

Statistical Analysis Setup:

The main goal of this paper is to analyze the data from a clinical trial run by Pfizer in the midst of the COVID-19 pandemic when trying to get their vaccine approved by the FDA, the data is shown below.

Table 1: Vaccine Efficacy against Covid 19 at least 7 days after second dose in patients without evidence of infection

Group	Cases	No. of subjects
BNT162b2	8	17,411
Placebo	162	17,511
Total	170	34,922

With this data in mind, we will use three methods to determine if Pfizer and the FDA were right in allowing this vaccine to be given to the public, as well as give out thought on if we felt their methods were optimistic or if they were skeptical with their testing. In particular we will analyze these data through three different lenses: Frequentist, Bayesian, and through the use of simulation. Below we will set up the statistical machinery needed to approach this problem and furthermore, delve into the statistical methods used in each specific method.

First off, let X denote the number of COVID-19 infections among the 17,411 patients randomly assigned to the vaccine group. $X \sim \text{Binom}(17411, \pi_v)$ where π_v is the probability of a COVID-19 infection for an individual in the vaccine group. Next, let Y denote the number of COVID-19 infections among the 17,511 patients randomly assigned to the placebo group. $Y \sim \text{Binom}(17511, \pi_p)$. Where π_p is the probability of a COVID-19 infection for an individual in the placebo group.

In this setting, we assume that X is drawn independently of Y . The parameter of interest is the vaccine efficacy ψ : $\psi = 1 - \frac{\pi_v}{\pi_p} = \frac{\pi_p - \pi_v}{\pi_p}$.

The main substance of this trial is that the FDA requires at least 30% efficacy for a new therapy to be approved.

We know that for a Binomial random variable, the optimal estimator of the success probability π is the mean of the sample. Thus, $\hat{\psi} = \frac{\hat{\pi}_p - \hat{\pi}_v}{\hat{\pi}_p} = \frac{\bar{Y} - \bar{X}}{\bar{Y}}$. However, a problem arises when we realize that the sampling distribution of $\hat{\psi}$ is not a well known distribution. We will come back to this idea in the simulation section, but for now we will instead look at this data from a different perspective; through the use conditional probability/inference.

In the Pfizer trial, they ran until they obtained $X + Y = 170$ infections, by doing this we effectively turn the problem from one of two sample, to one of only one sample. If we let T denote the number of infected patients from the $s = 170$ patients who got COVID-19 infected patients who were in the vaccine group, then because of the fact that the sample sizes $n_1 = 17,411$ and $n_2 = 17,511$ are large, and the rates π_v and π_p are small enough to keep $n_1\pi_v$ and $n_2\pi_p$ relatively constant, then we can say that $X \approx \text{Pois}(17,411)$ and $Y \approx \text{Pois}(17,411)$, after a bit of algebraic manipulations that are outside of the scope of this paper, we can see that: $T \sim \text{Binom}(170, \pi)$. Where $\pi = P(\text{vaccine} | \text{infection})$. To see what this value of π should be, we will use Bayes' rule. $\pi = P(\text{vaccine} | \text{infection}) = \frac{P(\text{infection} | \text{vaccine})}{P(\text{infection} | \text{vaccine}) + P(\text{infection} | \text{placebo}) \cdot P(\text{placebo})} = \frac{\pi_v \cdot \frac{n_1}{n_1 + n_2}}{\pi_v \cdot \frac{n_1}{n_1 + n_2} + \pi_p \cdot \frac{n_2}{n_1 + n_2}} = \frac{n_1 \pi_v}{n_1 \pi_v + n_2 \pi_p}$ where n_1 is the number of individuals in the vaccine group, and n_2 is the number in the placebo group. With that said since in our case $n_1 = 17,411$ and $n_2 = 17,511$, this implies that $n_1 \approx n_2$ and hence we can rewrite π as; $\pi = \frac{\pi_v}{\pi_v + \pi_p}$.

From there since we know the parameter of interest is the vaccine efficacy ψ , we need to find rewrite ψ in terms of this newly found value of π , with a little bit of algebra we can see that

$\pi = \frac{\pi_v}{\pi_v + \pi_p}$. Then multiplying by $(\pi_v + \pi_p)$ on both sides we get $(\pi_v + \pi_p)\pi = \pi_v$. Further simplifying we get $\pi\pi_v + \pi\pi_p = \pi_v \rightarrow (1 - \pi)\pi_v = \pi\pi_p \rightarrow \pi_v = \frac{\pi\pi_p}{(1-\pi)}$. Hence, by plugging this into the expression for ψ which is $\psi = 1 - \frac{\pi_v}{\pi_p} = \frac{\pi_p - \pi_v}{\pi_p}$, we can get this new expression for ψ . $\psi = 1 - \frac{\pi_v}{\pi_p} = 1 - \frac{\pi\pi_p}{\pi_p(1-\pi)} = 1 - \frac{\pi}{1-\pi} = \frac{1-2\pi}{1-\pi}$.

Part 1: Frequentist:

Now that we have set up the the statistical machinery needed to analyze the data, namely that $T \sim \text{Binom}(170, \pi_0)$, we can now do inference using the Frequentist approach. In particular we will do a likelihood analysis by finding the standard error and confidence interval for the maximum likelihood estimator of ψ . Furthermore, we will run a likelihood ratio test to test the following competing hypotheses: $H_0: \psi_0 = 0.3$ versus $H_1: \psi_0 \neq 0.3$ in order to check if the vaccine efficacy is up to FDA standards.

First off, in order to do any inference on the parameter ψ , we must reparametrize the PMF of T in terms of ψ . Thus, in this setting, the PMF of T is: $f(t) = \binom{s}{t} \left(\frac{1-\psi_0}{2-\psi_0}\right)^t \left(\frac{1}{2-\psi_0}\right)^{s-t}$, $t = 0, 1, \dots, s$ (Appendix 1). Hence we can write the likelihood function of ψ as: $L(\psi) = \binom{s}{t} \left(\frac{1-\psi}{2-\psi}\right)^t \left(\frac{1}{2-\psi}\right)^{s-t}$, $-\infty < \psi \leq 1$. After this, in order to make the derivative calculations easier, we can find $\ell(\psi)$ to be $\ell(\psi) = \ln\left(\binom{s}{t}\right) + t \ln(1 - \psi) - s \ln(2 - \psi)$ (Appendix 2).

Now in order to find the MLE of ψ we must, find the first derivative of $\ell(\psi)$ and set it equal to zero. We can do this because the natural logarithm is a monotonic increasing function and thus preserves order, which means the log-likelihood function has the same maximum as the likelihood function. Hence the same maximizing techniques can be used on the log-likelihood function. $\frac{d}{d\psi} \ell(\psi) = \frac{-t}{1-\psi} + \frac{s}{2-\psi}$ (Appendix 3). We can find the MLE candidate by setting the first derivative equal to 0 to get: $\hat{\psi}_0^{mle} = \frac{2t-s}{t-s}$ (Appendix 4).

Now that we have the candidate for the MLE of ψ , we need to find the second derivative of $\ell(\psi)$ in order to prove the MLE is in fact a maximum. This process is shown as: $\frac{d^2}{d\psi^2} \ell(\psi) = \frac{d}{d\psi} \left(\frac{-t}{1-\psi} + \frac{s}{2-\psi} \right) = \frac{-t}{(1-\psi)^2} + \frac{s}{(2-\psi)^2}$. By plugging in $\hat{\psi}_0^{mle} = \frac{2t-s}{t-s}$ into the second derivative we calculated above we find that: $\frac{d^2}{d\psi^2} \ell(\hat{\psi}_0^{mle}) = \left(\frac{1}{s} - \frac{1}{t}\right)(t-s)^2$ (Appendix 5).

Since $t \leq s$, we have $\frac{1}{s} - \frac{1}{t} \leq 0$. Thus, $\ell''\left(\hat{\psi}_0^{mle}\right) > 0$ and our MLE candidate is indeed a maximum. Note that this result hold as long as $s > t$.

Now that we have a point estimate for ψ , namely the maximum likelihood estimate of ψ , we now need to add some confidence around this point estimate. To do this we will need set up some more statistical machinery. In particular, if we suppose $T_1, T_2, \dots, T_{170} \sim \text{Bernoulli}(\pi)$ represents the 170 infected individuals across both groups, then since the number of patients is large, under certain regulatory conditions $\hat{\psi}_0^{mle} \sim \text{Norm}\left(\psi_0, \frac{1}{\sqrt{-\ell'(\hat{\psi}_0^{mle})}}\right)$. These regulatory conditions are that the bounds of the data do not depend on the ψ and that the true value of ψ does not fall on a boundary. The first of these is passed and the second of these will require the use of analyzing the second order Taylor approximation of the log-likelihood function of ψ . Furthermore, if the fit of this approximation is good we have assurance that the approximate normal distribution is valid.

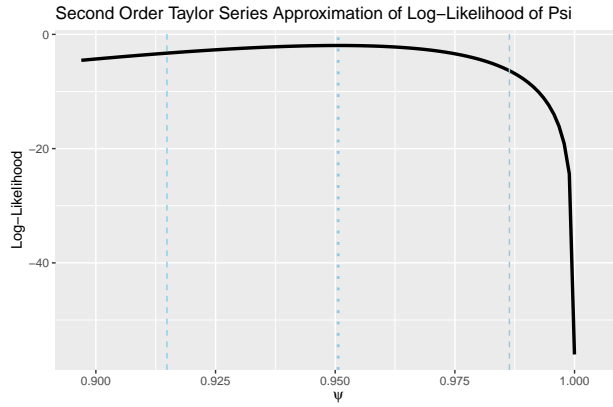


Figure 1.

As can be seen the estimated value of ψ is not on or too close to a boundary and furthermore the fit of the second order approximation and the log-likelihood function itself are nearly indistinguishable.

Thus we have assurance that the normality of the MLE of ψ holds and we can now find the estimated

$$\text{standard error of the above MLE for } \psi: \hat{SE}[\hat{\psi}_0^{mle}] = \frac{1}{\sqrt{-\ell''(\hat{\psi}_0^{mle})}} = \frac{1}{\sqrt{(\frac{1}{t} - \frac{1}{s})(t-s)^2}}$$

Now that we have an expression for the standard error of $\hat{\psi}_0^{mle}$, we can create a 99% Wald confidence interval for $\hat{\psi}_0^{mle}$ through the use of the asymptotic normality of $\hat{\psi}_0^{mle}$ itself.

Considering the large population and the importance of reducing type I errors, we've adopted a more stringent significance level of 0.01 as suggested by Lin ³. The confidence interval is then defined as: $\hat{\psi}_0^{mle} \pm z_{\alpha/2} \sqrt{\frac{1}{-\ell'(\hat{\psi}_0^{mle})}}$ where $z_{\alpha/2}$ represents the $1 - \frac{\alpha}{2}$ th quantile of the standard normal distribution.

Lastly, we will run a likelihood ratio test to test the following competing hypotheses: $H_0: \psi_0 = 0.3$

versus H_1 : $\psi_0 \neq 0.3$ in order to check if the vaccine efficacy is up to FDA standards. The likelihood ratio test statistic (LRT) is defined as $W = 2 \ln \left[\frac{L(\hat{\psi}_0^{mle})}{L(\hat{\psi}_0^{null})} \right]$. Under the null hypothesis $W \approx \chi_1^2$.

Since large values of W provide evidence in the direction of the null hypothesis, the P-value is the right tail probability under the χ_1^2 distribution. The LRT statistic takes the form of: $W =$

$$2 \ln \left[\frac{\binom{s}{t} \left(\frac{1-\hat{\psi}_0^{mle}}{2-\hat{\psi}_0^{mle}} \right)^t \left(\frac{1}{2-\hat{\psi}_0^{mle}} \right)^{s-t}}{\binom{s}{t} \left(\frac{1-\hat{\psi}_0^{null}}{2-\hat{\psi}_0^{null}} \right)^t \left(\frac{1}{2-\hat{\psi}_0^{null}} \right)^{s-t}} \right] \text{ however since we observed the values } s = 170 \text{ and } t = 8 \text{ we can}$$

$$\text{simplify the expression and rewrite this statistic in the form: } W = 2 \ln \left[\frac{\left(\frac{1-\hat{\psi}_0^{mle}}{2-\hat{\psi}_0^{mle}} \right)^8 \left(\frac{1}{2-\hat{\psi}_0^{mle}} \right)^{162}}{\left(\frac{1-\hat{\psi}_0^{null}}{2-\hat{\psi}_0^{null}} \right)^8 \left(\frac{1}{2-\hat{\psi}_0^{null}} \right)^{162}} \right]$$

we will then use this test statistic and the approximate chi-squared distribution to obtain a p-value to rigorously test our competing hypotheses.

Part 2: Bayesian:

For our second approach to inference, we decided to use Bayesian methods. In particular we will use the beta-binomial model with four different priors. With that said, we will construct 99% credible intervals and calculate the p-value to make test the hypotheses $H_0 : \psi_0 = 0.3$ vs $H_1 : \psi_0 > 0.3$ where ψ_0 is the efficacy of the vaccine.

To help with the calculation, we need to make use of the Bayes' Theorem: $h(\theta_0|x) = \frac{f(x|\theta_0)g(\theta_0)}{\int_{\theta_0} f(x|\theta_0)g(\theta_0)d\theta_0}$ where f is the likelihood function, g is the prior distribution, and h is the posterior distribution.

Since we are using beta-binomial model, we will find the posterior distribution in the general case so that we use it with our different prior distributions. In our case, $X \sim Binom(s, \pi)$ and $g(\pi_0) = Beta(\alpha_0, \beta_0)$, then the posterior distribution is $Beta(\alpha_0 + t, \beta_0 + s - t)$ (Appendix 7).

We will be covering four different prior distributions, detailed bellow.

Prior 1, Pfizer Prior The first prior we chose is $g(\pi_0) = Beta(0.700102, 1)$, the prior used in the Pfizer efficacy trials, then using the formula we have derived and our given data, the posterior distribution $g(\pi_0|s) = Beta(8.700102, 163)$.

Prior 2, Elicited Prior For our second prior, we will derive it with skeptical prior beliefs. In particular, the median of the vaccine efficacy ψ_0 (apriori) is 0 and the 99th percentile of vaccine efficacy ψ_0 (apriori) is 0.3.

This prior indicates that we think that half the time the vaccine efficacy will be 0. Meaning that 50% of the time the vaccine will either have no effect or make people worse off. We also stated that 99% of the time, the vaccine efficacy will be less than 0.3, meaning that only 1% of the time the vaccine will be approved by FDA. Using this prior information, we can find the parameters for the prior distribution of π_0 . We can do this by manipulating the probability statements as shown below. First, $P(\psi_0 < 0) = 0.5 \implies P(\frac{1-2\pi}{1-\pi} < 0) = 0.5 \implies P(1 - 2\pi < 0) = 0.5 \implies P(\pi > \frac{1}{2}) = 0.5$. This tells us that the median prior distribution of π_0 is 0.5. Second, $P(\psi_0 < 0.3) = 0.99 \implies P(\frac{1-2\pi}{1-\pi} < 0.3) = 0.99 \implies P(-1.7\pi < -0.7) = 0.99 \implies P(\pi > \frac{0.7}{1.7}) = 0.99$, which tells us that the 99th percentile of the prior distribution of π_0 is $\frac{0.7}{1.7}$. Lastly, we can use the `beta.select` function in R to use the above quantiles to find the parameters of the prior distribution of π_0 . Using the formula we derived earlier, the posterior $g(\pi_0|s) = \text{Beta}(94.04, 249.04)$.

Prior 3, Jeffrey's prior The third Prior we chose is called [Jeffrey's prior](#)⁵, which is a non-informative (objective) prior that is derived from the likelihood function alone, independent of the specific form of the parameter of interest. It is derived by taking the square root of the determinant of the Fisher information matrix, which measures the amount of information that an observable random variable carries about an unknown parameter.

Our choice to use Jeffreys' prior as our third prior is driven by our desire to balance the potential pitfalls of subjective elicitation. It's not uncommon for subjective elicitation to lead to less than ideal prior distributions due to biases that may arise during the elicitation process. Often, elicitation only uncovers a limited set of characteristics about the prior, and the rest - such as its functional form - is selected for convenience, which may not necessarily be the most appropriate choice.

Hence, it's considered good practice to match results from a subjective analysis with those generated from a non informative prior analysis, such as one involving Jeffreys' prior.

As stated in the Appendix 8, Jeffreys' prior takes in the form $\text{Beta}(\frac{1}{2}, \frac{1}{2})$. Using the formula we derived earlier, the posterior is $g(\pi_0|s) = \text{Beta}(8.5, 162.5)$.

Prior 4, Uniform Prior The fourth prior we chose is $\text{Unif}(0, 1)$ which is a special case of Beta distribution with parameter $\text{Beta}(1, 1)$. We decided to use this as an additional non-informative prior. Using the formula we derived earlier, the posterior $g(\pi_0|s) = \text{Beta}(9, 163)$.

With all the information we have, we can now generalize the process of calculating the medians for ψ_0 , 99% credible intervals for ψ_0 , and the p-value for the hypothesis testing against $H_0 : \psi_0 = 0.3, H_1 : \psi_0 > 0.3$ for all four priors.

To transform from the 99% credible interval for π_0 to ψ_0 , suppose the 99% credible interval for π_0 is $[a, b]$, we can generalize the process below.

For $P(\pi_0 < k|t) = p \implies P\left(\psi_0 < \frac{1-2k}{1-k}|t\right) = 1 - p$ (Appendix 9).

From 99% credible interval $[a, b]$ for π_0 with a median c , we can get $P(\pi_0 < a|t) = 0.005, P(\pi_0 < b|t) = 1 - P(\pi_0 > b|t) = 0.995$, then we can calculate the 99% credible interval for ψ_0 to be $\left[\frac{1-2b}{1-b}, \frac{1-2a}{1-a}\right]$, and the median of ψ_0 will be $\frac{1-2c}{1-c}$.

To Test $H_0 : \psi_0 = 0.3, H_1 : \psi_0 \neq 0.3$, we can first replace ψ_0 with π_0 and then use the posterior distribution of π_0 corresponding to the prior that we are investigating to find the p-value: $P\left(\pi_0 \geq \frac{0.7}{1.7}|t\right)$ (Appendix 10).

This is the same across all four priors since we are testing against the same null hypothesis and the only thing changing is the posterior distributions updated from different prior distributions.

Thus, for any posterior distribution of π_0 , P-value = $P(\psi_0 \leq 0.3|t) = P\left(\pi_0 \geq \frac{0.7}{1.7}|t\right)$.

Part 3: Simulation:

Next, we'll apply bootstrapping to simulate the sampling distribution of ψ_0 . Bootstrapping is robust and versatile for constructing approximate sampling distributions, thereby enabling us to estimate standard errors and construct confidence intervals, particularly when the underlying distribution is unknown. We'll aim to estimate the 99% confidence interval of ψ_0 , representing Pfizer vaccine's efficacy.

In our bootstrap approach, we view the given data as the population and take re-samples with replacement, not assuming a specific distribution. This nonparametric and distribution-free approach gives us flexibility and applicability even with little information about the distribution of ψ_0 .

Two sample bootstrap:

Firstly, we'll conduct a two-sample bootstrap to create the sampling distribution of ψ_0 concerning individuals testing positive or not for COVID-19. We'll generate 10,000 re-samples, compute ψ_0

for each and save these in a data frame. If the distribution appears skewed or symmetric, we'll use percentile or standard bootstrap confidence interval methods respectively. The resulting simulation code can be found in Appendix.

One sample bootstrap:

Next, we'll conduct a one-sample bootstrap focusing on ψ_0 for the population that tested positive for COVID-19. The procedure is similar to the two-sample bootstrap, but with one combined parameter π instead of two. The resulting simulation code can be found in Appendix.

Bootstrapped sampling distribution:

Finally, we'll create histograms of the bootstrapped sampling distributions of ψ_0 for both methods, helping us understand the variability, uncertainty, and differences based on the method used. This will guide our choice of confidence interval method.

4 Results

Recall that our goal is to analyze a clinical trial ran by Pfizer. Using the data from Table 1, we now want to follow the procedures detailed in the methods section to obtain point estimates and confidence intervals for ψ_0 as well as test the competing hypotheses.

Part 1: Frequentist:

As detailed in the methods section, we found that $\hat{\psi}_0^{mle} = \frac{2t-s}{t-s}$. Thus with our data, we calculated $\hat{\psi}_0^{mle} = \frac{2t-s}{t-s} = \frac{16-170}{8-170} = \frac{77}{81}$. Additionally, we found $\hat{SE}[\hat{\psi}_0^{mle}] = \frac{1}{\sqrt{(\frac{1}{t}-\frac{1}{s})(t-s)^2}}$, thus for our data, $\hat{SE}(\hat{\psi}_0^{mle}) = \frac{1}{\sqrt{(\frac{1}{t}-\frac{1}{s})(t-s)^2}} = \frac{1}{\sqrt{(\frac{1}{8}-\frac{1}{170})(8-170)^2}} = \sqrt{\frac{170}{531441}} \approx 0.017885$. We then constructed a 99% confidence interval for $\hat{\psi}_0^{mle}$ using these values with an α level of 0.01. The confidence interval is given by $\hat{\psi}_0^{mle} \pm Z_{\alpha/2}\hat{SE}$ thus for our data, $\frac{77}{81} \pm 2.576\sqrt{\frac{170}{531441}} \approx [0.905, 0.997]$. Thus over hypothetical replications, we found that 99% of the time, ψ_0 will be in the range $[0.905, 0.997]$. In our context, this tells us that over hypothetical replications, the efficacy of the vaccine will be in the range $[0.905, 0.997]$ 99% of the time. Since the lower bound of this interval is significantly greater than 30%, we concluded that the efficacy of the vaccine to be greater than 30%. Lastly, we ran a likelihood ratio test of $H_0 : \psi_0 = 0.3$ vs $H_1 : \psi_0 > 0.3$. We found the W test statistic to be about 121.601. With this test statistic we found the upper tail probability in a χ_1^2 distribution and

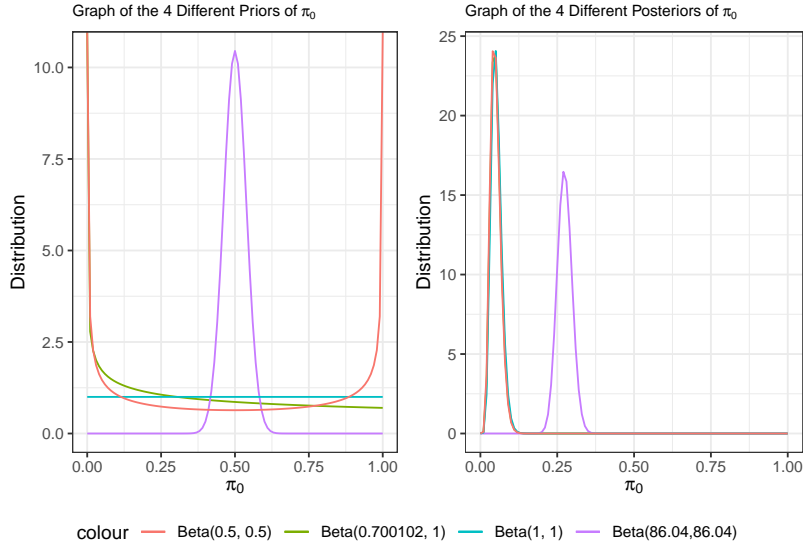
obtained the p-value $2.8222944 \times 10^{-28}$. Thus we have significant evidence at the 1% α level that $\psi_0 \neq 0.3$.

Part 2: Bayesian:

As detailed in the methods section, we have calculated the median and 99% credible intervals of ψ_0 for all four priors. Furthermore, tested the competing null hypothesis $H_0 : \psi_0 = 0.3$ against $H_1 : \psi_0 > 0.3$ and calculated the corresponding p-values:

Table 2: Bayesian Inference			
Prior	Median for ψ_0	99% Credible Interval for ψ_0	P-value
Beta(0.700102,1)	0.948551	[0.8850584, 0.9820899]	$1.9600144 \times 10^{-28}$
Beta(86.04,86.04)	0.621703	[0.4864698, 0.7253856]	6.3061231×10^{-8}
Beta(0.5,0.5)	0.949625	[0.8865568, 0.9827153]	$1.6184301 \times 10^{-28}$
Beta(1,1)	0.946707	[0.8822869, 0.9810631]	$3.7269324 \times 10^{-28}$

We then created a graph showing the four different prior and posterior distribution of π_0 to further investigate the different priors (Figure 2).



As we can see from the graph, the posterior distribution has skewed to the left in compared with the prior. This means that our data is saying that the proportion of people who got the vaccine given that they infected COVID-19 is smaller than our prior beliefs.

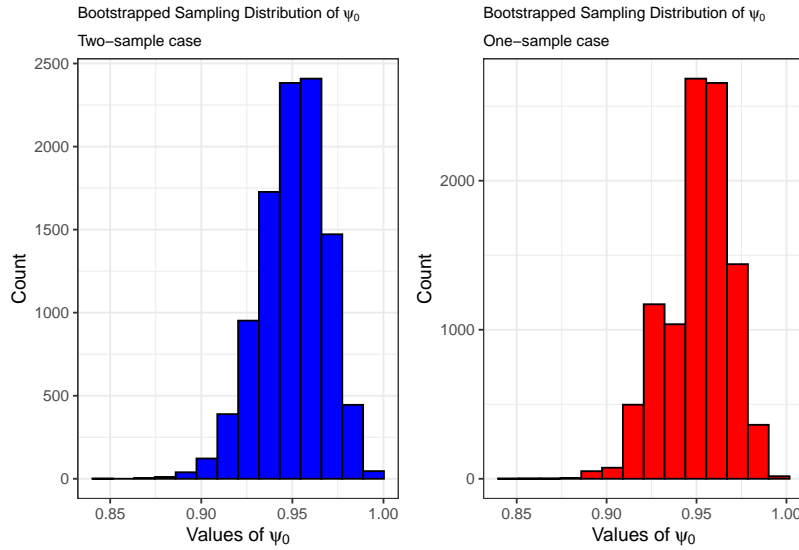
Given that the vaccine efficacy (ψ) is defined as $1 - \frac{\pi_v}{\pi_p}$, where π_v is the probability of a COVID-19 infection for an individual in the vaccine group and π_p is the probability of a COVID-19 infection

for an individual in the placebo group.

Moreover, all of the 99% credible intervals are well above 0.30 (or 30% efficacy), which is the FDA's minimum requirement for a new therapy to be approved. The p-values in all four cases are extremely small, which gives us strong evidence against the null hypothesis that the vaccine has no effect or that the vaccine efficacy ψ is zero. Given these results, we can conclude with high confidence that the Pfizer vaccine is effective against COVID-19, with an efficacy rate significantly above the FDA's minimum threshold of 30%. This affirms the decision by Pfizer and the FDA to approve this vaccine for public use.

Part 3: Simulation:

Here are the bootstrapped sampling distributions of ψ_0 , (Figure 3).



From examining the histograms visualizing the ψ distributions from the bootstrap re-samples, we can make similar conclusions. From both the one sample and two sample cases, we can see distributions representative of our confidence intervals. Majority of the ψ values lie somewhere in the range of 90% to 99%. There are slight differences in distribution based on whether we performed a one sample or two sample bootstrap, but the overall results are the same. We can see that even the left tail of the distributions are much higher than 30%. These histograms provide further evidence that our calculated confidence intervals are correct and that we can reject our null hypothesis. We can be more confident that the efficacy of the vaccine is greater than 30%. As

mentioned in the method section, because these histograms appear to be approximately normal, we must use the standard bootstrap confidence interval method to construct our confidence intervals. From the two sample bootstrapping procedure, we obtained a 99% confidence interval for the efficacy of the vaccine ψ_0 . This interval was about $[0.9039, 0.9967]$. This means that we are 99% confident that the true value of ψ_0 falls somewhere in this interval.

Our null hypothesis was that the $\psi_0 = 0.3$, we can clearly see that our confidence interval for ψ is well above 30%. In fact, we are 99% confident that ψ_0 is at least about 90%, which is much larger than 30%. In context, this tells us that we can be confident in saying that the vaccine has an efficacy in the range $[0.9039, 0.9967]$ which is well above the FDA required minimum efficacy of 30%.

From the one sample bootstrapping procedure, we found that 99% confidence interval for ψ_0 was approximately $[0.9039, 0.9968]$. This confidence interval provides further evidence supporting the efficacy of the vaccine is above 30%. From this confidence interval we can say that we are 99% confident that the true efficacy of the vaccine is in the range $[0.9039, 0.9968]$. Once again, this is much higher than 30%, which provides further evidence to again reject the null hypothesis in favor of saying that the vaccine's efficacy is greater than 30%.

5 Discussion/Conclusion

As described in the result section, for all our approaches, we got confidence and credible intervals that are far away from the FDA required 30% cutoff for the vaccine efficacy. The p-values were significantly small at an α level of 0.1. Thus we have concluded that the efficacy of the Pfizer vaccine is greater than 30% and should be distributed to the public.

In our Bayesian analysis, we employed four different priors: $Beta(0.700102, 1)$, $Beta(86.04, 86.04)$, Jeffreys' prior $Beta(0.5, 0.5)$ and uniform prior which is the same as a $Beta(1, 1)$. For $Beta(0.700102, 1)$, Jeffreys' prior $Beta(0.5, 0.5)$ and uniform prior $Beta(1, 1)$, we got similar credible interval and p-values. The Bayesian analysis with $Beta(86.04, 86.04)$, As seen in the result section, provided a somewhat lower estimate of efficacy compared to the other three priors and the other two approaches. This is due to the skeptical assumptions we made when defining the

prior. However as stated above, all estimates clearly exceeded the FDA’s required 30% efficacy for vaccine approval. The Pfizer’s prior, on the other hand, gives very similar results (p-values and credible intervals) to the Jeffreys’ prior, which is a non-informative prior we derived. If Pfizer’s prior was overly optimistic, we would expect it to significantly skew the results towards higher vaccine efficacy. However, the analysis does not reflect this as it yields similar statistics as $Beta(0.5, 0.5)$. Based on the analysis results and the comparisons made, we can see that Pfizer’s prior was not overly optimistic, and it behaved similarly to a non-informative prior.

In our study, we employ three different statistical methodologies – Frequentist, Bayesian, and Bootstrap – to provide a comprehensive analysis of the Pfizer vaccine’s efficacy. These three approaches offer complementary strengths, allowing us to interpret the data from multiple perspectives and balance potential bias or limitations of each individual method. Furthermore, because we obtained similar results, this offered validity in the results themselves and showed that for this particular vaccine results were robust regardless of the chosen method.

However, there were several challenges and limitations. One such limitation is the difficulty in directly comparing confidence intervals derived from the Frequentist approach and bootstrap approach with credible intervals from the Bayesian approach. While they may appear similar in nature, they are fundamentally different in concept and interpretation, which makes the comparison challenging. Despite this, they both lead us to the same conclusion about the vaccine efficacy. Another potential limitation is that we performed a two-sided hypothesis test (not equals) instead of a one-sided test (greater than), in the likelihood ratio test. By following this procedure, we when rejected the null hypothesis we had to assume that the vaccine efficacy was larger. This may impact our observed p-values. Finally, due to constraints in the length of the paper, we have only explored beta-binomial priors, and we have not had chance to explore other options like Poisson distributions or other possible priors (Gamma prior, poisson likelihood), which might have provided additional insights. These are potential areas for future research to enhance the robustness of our findings and conclusions. While our analysis is robust and comprehensive given the constraints, the potential for further investigation and refinement remains.

6 References

- ¹Clinical Trial Group; (2020, December). Safety and efficacy of the BNT162B2 mrna COVID-19 vaccine | Nejm. The New England Journal of Medicine
- ²Gartlehner G;Hansen RA;Nissman D;Lohr KN;Carey TS; (2006, April). Criteria for distinguishing effectiveness from efficacy trials in systematic reviews. National Center for Biotechnology Information
- ³Lin, D.-Y., Zeng, D., Mehrotra, D. V., Corey, L., & Gilbert, P. B. (2021, October 20). Evaluating the efficacy of Coronavirus Disease 2019 vaccines. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America
- ⁴World Health Organization. Coronavirus. World Health Organization
- ⁵YANG, Ruoyong and James O. BERGER, A Catalog of Noninformative Priors

7 Appendix

7.1 Methods Part 1. Frequentist:

¹ PMF of t : $f(t) = \binom{s}{t} \left(\frac{1-\psi_0}{2-\psi_0}\right)^t \left(1 - \frac{1-\psi_0}{2-\psi_0}\right)^{s-t} = \binom{s}{t} \left(\frac{1-\psi_0}{2-\psi_0}\right)^t \left(\frac{1}{2-\psi_0}\right)^{s-t}, t = 0, 1, \dots, s$

² Log Likelihood of ψ : $\ell(\psi) = \ln(L(\psi)) = \ln\left(\binom{s}{t} \left(\frac{1-\psi}{2-\psi}\right)^t \left(\frac{1}{2-\psi}\right)^{s-t}\right) = \ln\left(\binom{s}{t}\right) + \ln\left(\left(\frac{1-\psi}{2-\psi}\right)^t\right) + \ln\left(\left(\frac{1}{2-\psi}\right)^{s-t}\right) = \ln\left(\binom{s}{t}\right) + t \ln(1-\psi) - t \ln(2-\psi) + (s-t) \ln(1) - (s-t) \ln(2-\psi) = \ln\left(\binom{s}{t}\right) + t \ln(1-\psi) - s \ln(2-\psi)$

³ First derivative of Log Likelihood of ψ : $\frac{d}{d\psi} \ell(\psi) = \frac{d}{d\psi} (\ln\left(\binom{s}{t}\right) + t \ln(1-\psi) - s \ln(2-\psi)) = 0 + \frac{-t}{1-\psi} + \frac{s}{2-\psi} = \frac{-t}{1-\psi} + \frac{s}{2-\psi}$

⁴ MLE candidate: $\frac{d}{d\psi} \ell(\psi) = 0 \implies \frac{-t}{1-\psi} + \frac{s}{2-\psi} = 0 \implies \frac{s}{2-\psi} = \frac{t}{1-\psi} \implies s(1-\psi) = t(2-\psi) \implies s - s\psi = 2t - t\psi \implies (t-s)\psi = 2t - s \implies \hat{\psi}_0^{mle} = \frac{2t-s}{t-s}$

⁵ Second derivative of Log Likelihood of ψ : $\frac{d^2}{d\psi^2} \ell(\hat{\psi}_0^{mle}) = \frac{-t}{(1-\frac{2t-s}{t-s})^2} + \frac{s}{(2-\frac{2t-s}{t-s})^2} = \frac{-t}{(\frac{-t}{t-s})^2} + \frac{s}{(\frac{-s}{t-s})^2} = -\frac{1}{t}(t-s)^2 + \frac{1}{s}(t-s)^2 = \left(\frac{1}{s} - \frac{1}{t}\right)(t-s)^2$

⁶ Figure 1. Second Order Taylor Series Approximation:

```

# Create the log-likelihood function of psi:
loglik_psi <- function(psi) {
  ifelse(psi > 1, NA,
    log(choose(170, 8)) + 8*log(1-psi) - 170*log(2-psi)
  )}

# Get the second order Taylor approximation of the log-likelihood function:
psi_second <- maxLik2(loglik = loglik_psi, start = psi_mle, method = "NR")

# Make the plot of the two above functions:
plot <- plot(psi_second) %>%

  gf_labs(title = "Second Order Taylor Series Approximation of Log-Likelihood of
    Psi",

    x = expression(psi), y = "Log-Likelihood", caption = "Figure 1.")

# Plot the data:
plot

```

7.2 Methods Part 2. Bayesian:

⁷ Proof of beta-binomial posterior $h(\pi_0|x) = \frac{\binom{s}{t} \pi_0^t (1-\pi_0)^{s-t} \frac{\Gamma(\alpha_0+\beta_0)}{\Gamma(\alpha_0)\Gamma(\beta_0)} \pi_0^{\alpha_0-1} (1-\pi_0)^{\beta_0-1}}{\int_0^1 \binom{s}{t} \pi_0^t (1-\pi_0)^{s-t} \frac{\Gamma(\alpha_0+\beta_0)}{\Gamma(\alpha_0)\Gamma(\beta_0)} \pi_0^{\alpha_0-1} (1-\pi_0)^{\beta_0-1} d\pi_0} =$

$$\frac{\binom{s}{t} \frac{\Gamma(\alpha_0+\beta_0)}{\Gamma(\alpha_0)\Gamma(\beta_0)} \pi_0^{t+\alpha_0-1} (1-\pi_0)^{s-t+\beta_0-1}}{\binom{s}{t} \frac{\Gamma(\alpha_0+\beta_0)}{\Gamma(\alpha_0)\Gamma(\beta_0)} \int_0^1 \pi_0^{t+\alpha_0-1} (1-\pi_0)^{s-t+\beta_0-1} d\pi_0}.$$

Note that Beta function: $\int_0^1 u^{r-1} (1-u)^{s-1} du = \frac{\Gamma(r)\Gamma(s)}{\Gamma(r+s)}.$

$$\frac{\pi_0^{t+\alpha_0-1} (1-\pi_0)^{s-t+\beta_0-1}}{\frac{\Gamma(\alpha_0+t)\Gamma(\beta_0+s-t)}{\Gamma(\alpha_0+\beta_0+s)}} = \frac{\Gamma(\alpha_0+\beta_0+s)}{\Gamma(\alpha_0+t)\Gamma(\beta_0+s-t)} \pi_0^{t+\alpha_0-1} (1-\pi_0)^{s-t+\beta_0-1}, 0 \leq \pi_0 \leq 1 = \text{Beta}(\alpha_0+t, \beta_0+s-t)$$

⁸ Jefferey's Prior $g(\theta) \propto \sqrt{I(\theta)}$ where $I(\theta)$ is Fishers Information matrix. $g(\theta) \propto \sqrt{-E \left[\frac{d^2}{d\theta^2} \ln(f(y|\theta)) \right]}$

where $f(y|\theta)$ is the likelihood function Since in our case, $f(t|\pi)$ is $T \sim \text{Binomial}(s, \pi)$,

$E[T] = s\pi$ and we can derive the Jeffrey's prior: likelihood: $f(t|\pi) = \binom{s}{t} \pi^t (1-\pi)^{s-t}$

log-likelihood: $\ln(f(t|\pi)) = \ln \left(\binom{s}{t} \pi^t (1-\pi)^{s-t} \right) = \ln \left(\binom{s}{t} \right) + t \ln(\pi) + (s-t) \ln(1-\pi)$

first derivative: $\frac{d}{d\pi} (\ln(f(t|\pi))) = \frac{d}{d\pi} (\ln \left(\binom{s}{t} \right) + t \ln(\pi) + (s-t) \ln(1-\pi)) = \frac{t}{\pi} - \frac{s-t}{1-\pi}$

second derivative: $\frac{d^2}{d\pi^2} (\ln(f(t|\pi))) = \frac{d}{d\pi} \left(\frac{t}{\pi} - \frac{s-t}{1-\pi} \right) = \frac{-t}{\pi^2} - \frac{s-t}{(1-\pi)^2} E \left[\frac{d^2}{d\pi^2} (\ln(f(t|\pi))) \right] =$

$E \left[\frac{-t}{\pi^2} - \frac{s-t}{(1-\pi)^2} \right] = -\frac{1}{\pi^2} E[t] - \frac{s-E[t]}{(1-\pi)^2} = \frac{-s\pi}{\pi^2} - \frac{s-s\pi}{(1-\pi)^2} = -\frac{s}{\pi} - \frac{s}{1-\pi}$ Jeffrey's prior: $g(\pi) \propto \sqrt{-\left(-\frac{s}{\pi} - \frac{s}{1-\pi}\right)} \propto \sqrt{s \cdot \left(\frac{1-\pi+\pi}{\pi(1-\pi)}\right)} \propto \sqrt{s\pi}^{-\frac{1}{2}} (1-\pi)^{-\frac{1}{2}}$

⁹ General - transformation $P(\pi_0 < k|t) = P\left(\frac{1-\psi}{2-\psi} < k|t\right) = P(1 - \psi < 2k - k\psi|t) = P\left(\psi > \frac{1-2k}{1-k}|t\right) = p \implies P\left(\psi < \frac{1-2k}{1-k}|t\right) = 1 - p$

¹⁰ General - P-value $P(\psi_0 \leq 0.3|t) \implies P\left(\frac{1-2\pi_0}{1-\pi_0} \leq 0.3|t\right) \implies P(-1.7\pi_0 \leq -0.7|t) \implies P\left(\pi_0 \geq \frac{0.7}{1.7}|t\right)$

7.3 Results Part 1. Frequentist:

¹¹ Likelihood Statistics:

```
# Set up observed binomial values:
t <- 8
s <- 170

# Maximum Likelihood Estimate of Psi:
psi_mle <- (2*t - s) / (t - s)

# Estimated standard error of Psi:
psi_est_se <- 1 / sqrt((1/t - 1/s) * (t - s)^2)

# 99% confidence interval for Psi:
lower <- psi_mle - qnorm(0.995) * psi_est_se
upper <- psi_mle + qnorm(0.995) * psi_est_se

# W statistic for Likelihood Ratio Test of Psi:
w_obs <- 2*(8*log((1-psi_mle)/(2-psi_mle))+162*log((1)/(2-psi_mle))-
          8*log((1-0.3)/(2-0.3))-162*log((1)/(2-0.3)))

# P-value for the Likelihood Ratio Test of Psi:
p_val_lrt <- pchisq(q=w_obs, df=1, lower.tail=F)
```

¹² Regulatory conditions:

```
# Create the log-likelihood function of psi:
loglik_psi <- function(psi) {
  ifelse(psi > 1, NA,
         log(choose(170, 8)) + 8*log(1-psi) - 170*log(2-psi))}

```



```

# Get the second order Taylor approximation of the log-likelihood function:
psi_second <- maxLik2(loglik = loglik_psi, start = psi_mle, method = "NR")
# Make the plot of the two above functions:
plot <- plot(psi_second) %>%
  gf_labs(title = "Second Order Taylor Series Approximation of Log-Likelihood of
              Psi",
          x = expression(psi), y = "Log-Likelihood")
plot

```

7.4 Results Part 2 Bayesian:

¹³ Prior & Posterior Distribution Information

```

by1Med <- round(qbeta(0.5,8.700102,163),6)
by1CI <- c(round(qbeta(0.005,8.700102,163),6),
           round(qbeta(0.995,8.700102,163),6))
by1pval <- pbeta(0.7/1.7, 8.700102,163,lower.tail = F)
CI99 <- function(pi) {
  return((1-2*pi)/(1-pi))
}
by1CIs <- sort(CI99(by1CI))
quantile1 = list(p = 0.5, x = 0.50)
quantile2 = list(p = 0.01, x = 0.7/1.7)
by2Parm <- beta.select(quantile1,quantile2)
by2Med <- round(qbeta(0.5,86.04+8,86.04+162),6)
by2CI <- c(round(qbeta(0.005,86.04+8,86.04+162),6),
           round(qbeta(0.995,86.04+8,86.04+162),6))
by2pval <- pbeta(0.7/1.7, 86.04+8, 86.04+162,lower.tail = F)
by2CIs <- sort(CI99(by2CI))
by3Med <- round(qbeta(0.5,0.5+8,0.5+162),6)

```

```

by3CI <- c(round(qbeta(0.005,0.5+8,0.5+162),6),
           round(qbeta(0.995,0.5+8,0.5+162),6))
by3pval <- pbeta(0.7/1.7, 0.5+8,0.5+162,lower.tail = F)
by3CIs <- sort(CI99(by3CI))
by4Med <- round(qbeta(0.5,1+8,1+162),6)
by4CI <- c(round(qbeta(0.005,9,163),6), round(qbeta(0.995,9,163),6))
by4pval <- pbeta(0.7/1.7,9,163,lower.tail = F)
by4CIs <- sort(CI99(by4CI))

```

¹⁴ Figure 2. Prior & Posterior Graphs

```

g1 <- ggplot() +
  # Prior 1: Pfizer prior:
  stat_function(fun = dbeta,
               args = list(shape1 = 0.700102, shape2 = 1),
               mapping = aes(color = "Beta(0.700102, 1)")) +
  # Prior 2: Elicited prior:
  stat_function(fun = dbeta,
               args = list(shape1 = 86.04, shape2 = 86.04),
               mapping = aes(color = "Beta(86.04,86.04)")) +
  # Prior 3: Unif(0, 1) prior:
  stat_function(fun = dbeta,
               args = list(shape1 = 1, shape2 = 1),
               mapping = aes(color = "Beta(1, 1)")) +
  # Prior 4: Jeffrey's prior:
  stat_function(fun = dbeta,
               args = list(shape1 = 0.5, shape2 = 0.5),
               mapping = aes(color = "Beta(0.5, 0.5)")) +
  labs(title = expression("Graph of the 4 Different Priors of" ~ pi[0]),
       x = expression(pi[0]),

```

```

    y = "Distribution") +
  theme_bw() +
  theme(plot.title = element_text(size = 9)) +
  theme(legend.position="none")
g2 <- ggplot() +
  # Posterior 1: Pfizer posterior:
  stat_function(fun = dbeta,
               args = list(shape1 = 8.700102, shape2 = 163),
               mapping = aes(color = "Beta(8.700102, 163)")) +
  # Posterior 2: Elicited posterior:
  stat_function(fun = dbeta,
               args = list(shape1 = 94.04, shape2 = 249.04),
               mapping = aes(color = "Beta(94.04, 249.04)")) +
  # Posterior3: Unif(0, 1) posterior:
  stat_function(fun = dbeta,
               args = list(shape1 = 9, shape2 = 163),
               mapping = aes(color = "Beta(9, 163)")) +
  # Posterior 4: Jeffrey's posterior:
  stat_function(fun = dbeta,
               args = list(shape1 = 8.5, shape2 = 162.5),
               mapping = aes(color = "Beta(8.5, 162.5)")) +
  labs(title = expression("Graph of the 4 Different Posteriors of" ~ pi[0]),
       x = expression(pi[0]),
       y = "Distribution",
       caption = "Figure 2.") +
  theme_bw() +
  theme(plot.title = element_text(size = 9)) +
  theme(legend.position="none")
ggarrange(g1,g2,ncol = 2,common.legend = TRUE, legend="bottom")

```

7.5 Results Part 3 Simulation:

¹⁵ Figure 3. Sampling Distributions Histogram

```
# Create the bootstrapped sampling distribution of psi:
# Two Sample
plot1 <- ggplot(data = boot_sim_psi) +
  geom_histogram(mapping = aes(x = psi), bins = 14, color = "black",
    fill = "blue") +
  labs(title = expression("Bootstrapped Sampling Distribution of" ~ psi[0]),
    subtitle = "Two-sample case", x = expression("Values of" ~ psi[0]),
    y = "Count") +
  theme_bw()
# One Sample
plot2 <- ggplot(data = boot_sim_psi_2) +
  geom_histogram(mapping = aes(x = psi), bins = 14, color = "black",
    fill = "red") +
  labs(title = expression("Bootstrapped Sampling Distribution of" ~ psi[0]),
    subtitle = "One-sample case", x = expression("Values of" ~ psi[0]),
    y = "Count", caption = "Figure 2.") +
  theme_bw()
# Combine the plots side by side
combined_plots <- grid.arrange(plot1, plot2, ncol = 2)
```

¹⁶ Two Sample Bootstrap:

```
# Set the seed for reproducibility:
set.seed(123)
# Number of re-samples:
B = 10000
# Set up the data from the experiment:
```

```

x <- rep(c(1, 0), times=c(8, 17411-8))
y <- rep(c(1, 0), times=c(162, 17511-162))

# Compute the observed psi value:
x_bar <- mean(x)
y_bar <- mean(y)
psi_obs <- (y_bar - x_bar) / y_bar

# Begin the simulation:
boot_sim <- lapply(1:B, FUN = function(i) {
  # Generate a re-sample from x and y with replacement:
  x_sample = sample(x, size = 17411, replace = TRUE)
  y_sample = sample(y, size = 17511, replace = TRUE)

  # Find the psi from the re-sample:
  x_bar_sim = mean(x_sample)
  y_bar_sim = mean(y_sample)
  psi_sim = (y_bar_sim - x_bar_sim) / y_bar_sim

  # Add this value to a data frame:
  data.frame(psi = psi_sim)})

# Turn the simulation into a data frame:
boot_sim_psi <- do.call(rbind, boot_sim)

# Create the standard bootstrap confidence interval:
lower_2 <- psi_obs - qnorm(0.995) * sd(boot_sim_psi$psi)
upper_2 <- psi_obs + qnorm(0.995) * sd(boot_sim_psi$psi)

```

¹⁷ One Sample Bootstrap:

```

# Set the seed for reproducibility:
set.seed(63)

# Number of re-samples:
B <- 10000

# Set up the data from the experiment:

```

```

t <- rep(c(1, 0), times=c(8, 162))

# Compute the observed psi value:
pi_obs <- x_bar / (x_bar + y_bar)
psi_obs_2 <- (1 - 2*pi_obs) / (1 - pi_obs)

# Begin the simulation:
boot_sim_2 <- lapply(1:B, FUN = function(i) {
  # Re-sample from the t vector with replacement:
  t_sample = sample(t, size=170, replace=TRUE)
  # Find the pi value from the simulation:
  pi_sim = mean(t_sample)
  # Find the psi value from the simulation:
  psi_sim = (1-2*pi_sim) / (1-pi_sim)
  # Store the values of the simulation in a data frame:
  data.frame(pi=pi_sim, psi=psi_sim)})

# Turn the simulation into a data frame:
boot_sim_psi_2 <- do.call(rbind, boot_sim_2)

# Create the standard bootstrap confidence interval:
lower_3 <- psi_obs_2 - qnorm(0.995) * sd(boot_sim_psi_2$psi)
upper_3 <- psi_obs_2 + qnorm(0.995) * sd(boot_sim_psi_2$psi)

```