Midterm

Shabir Gazzaly

2024-10-29

R Markdown

Question 1:

Null hypothesis- Infection rate with Omicron variant will be the same in Pfizer vaccinated and non vaccinated participants.(Fourth dose of Pfizer vaccine doesn't protect against infection with the omicron variant)

Alternative hypothesis- Infection rate with Omicron variant will be lower in Pfizer vaccinated participants than non vaccinated participants.(Fourth dose of Pfizer vaccine protects against infection with the omicron variant)

Question 2: Statistic will be the difference between estimated number of infected participants in the vaccinated group under null hypothesis and the observed number of infected participants. Therefore, if the infection rate is lower than estimation(under null hypothesis), statistic will be positive.

Estimated infection rate under null hypothesis:

```r
80/ 308
```

```
## [1] 0.2597403
```

Can be rounded off to 0.26. Estimated number of infected participants in the vaccinated group under null hypothesis:

```r
0.26*154
```

```
## [1] 40.04
```

```r
statistic<-function(x){40.04-x}
```

Question 3: Generating a null distribution:

```r
nulld<-replicate(100000,statistic(rbinom(1,154,0.26)))
```
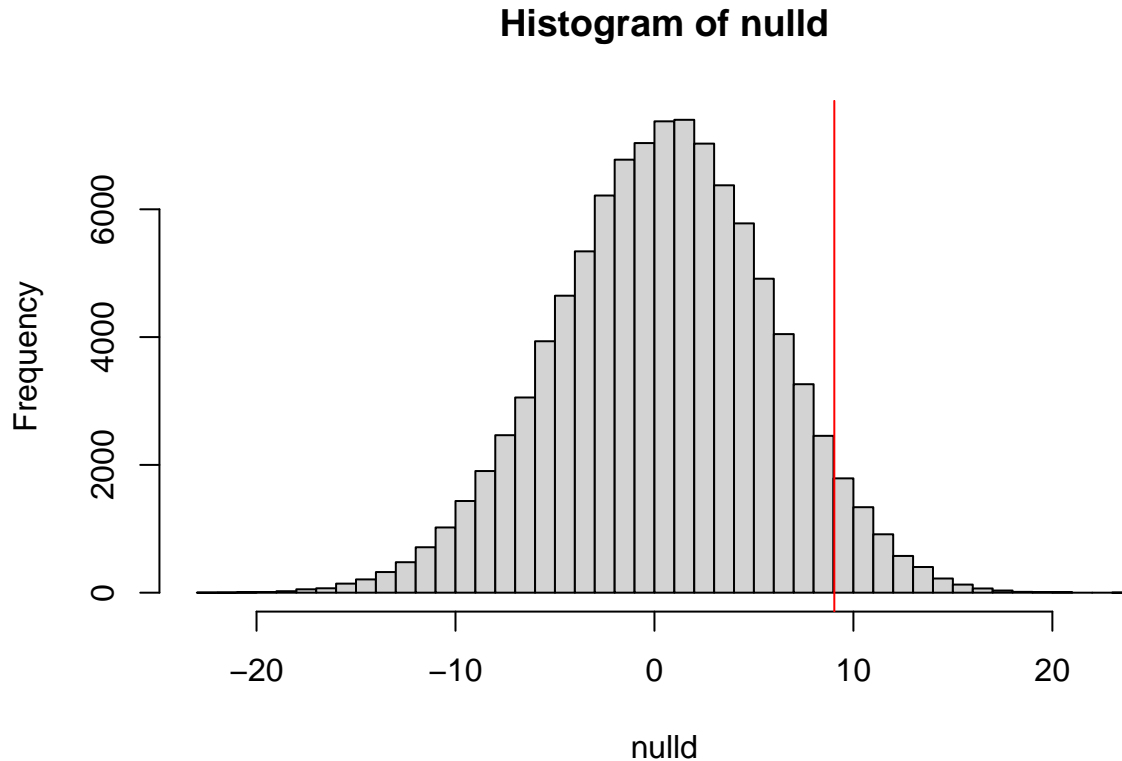
We will use a level of 5%, because the sample size is small. Finding threshold:

```r
quantile(nulld,0.95)
```

```
##  95%
## 9.04
```

```
hist(nulld,breaks=50)
abline(v=9.04,col="red")
```

Histogram of nulld

Question 4: Statistic of observed value:

```
statistic(28)
```

```
## [1] 12.04
```

Observed statistic is greater than threshold (12.04>9.04), therefore, we can reject the null hypothesis and conclude that the infection rate is lower in Pfizer vaccinated participants compared to non vaccinated participants, and the fourth dose of Pfizer protects against the omicron variant.

Question 5:

p-value of the test:

```
mean(nulld>=12.04)
```

```
## [1] 0.01457
```

Question 6: VE is the vaccination efficacy (in %), where "v" is the infection rate in vaccinated participants and "uv" is the infection rate in non vaccinated participants.

```r
VE<-function(v,uv){(uv-v)*100/uv}

v<-28/154
uv<-80/308

VE(v,uv)
```

```
## [1] 30
```

efficacy is 30%.

at 50% efficacy, infection rate in vaccinated participants would be $0.26/2 = 0.13$

Null hypothesis: Vaccination efficacy of Pfizer would be 50% (or less) Alternative hypothesis: Vaccination efficacy of Pfizer would be greater than 50%
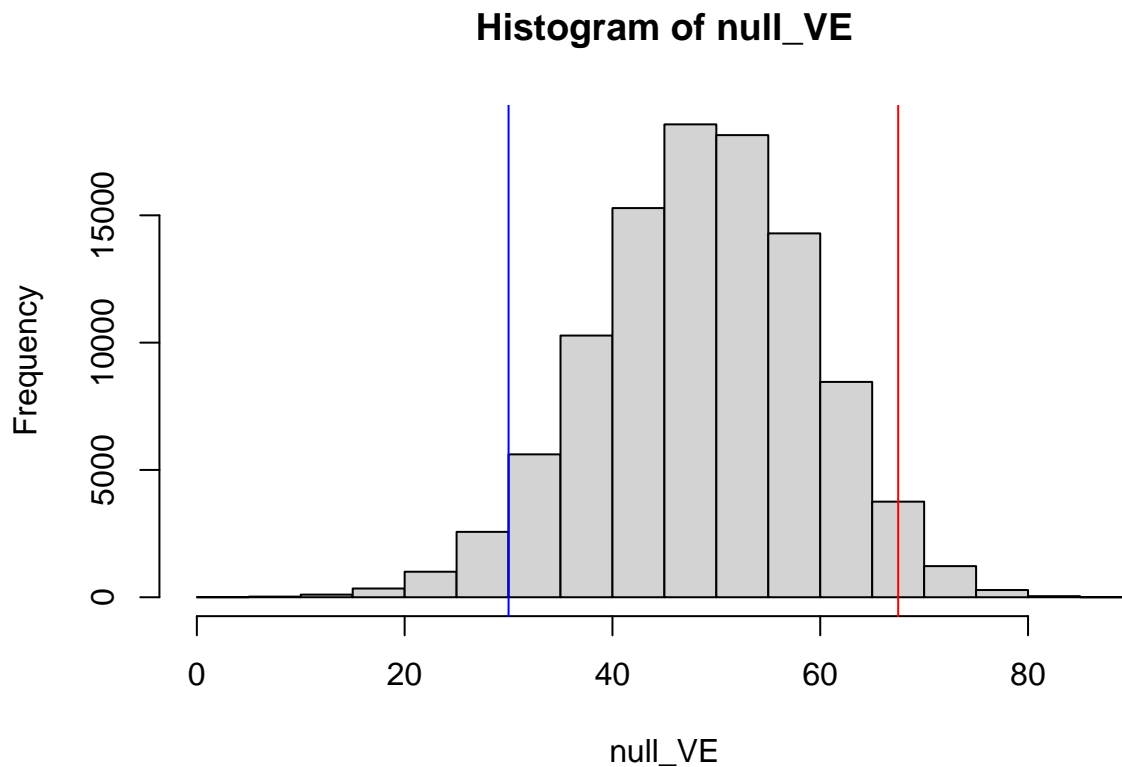
Null distribution:

```r
null_VE=c()
nullv=rbinom(100000,154,0.13)/154
for (i in nullv){null_VE=c(null_VE,VE(i,uv))}
```

threshold to reject null hypothesis:

```r
quantile(null_VE,0.95)
```

```
##  95%
## 67.5
```

```r
hist(null_VE,breaks=30)
abline(v=quantile(null_VE,0.95),col="red")
abline(v=30,col="blue")
```

Histogram of null_VE

Red line shows the threshold and blue line shows the observed efficacy.

The observed vaccination efficacy is less than threshold (30<67.5). p-value:

```r
mean(null_VE>=30)
```

```
## [1] 0.95954
```

Based on the analysis, we can conclude that the Pfizer vaccination efficacy is less than 50%. Although the fourth dose of Pfizer protects against the infection with omicron variant, it does not significantly reduce the risk of getting infected (Efficacy needs to be greater than 50%).