Inference on a single proportion using normal distributions

Overview

Questions about anything?

Hypothesis tests for a single proportion using normal distributions

Confidence for a single proportion using normal distributions

Confidence for a single proportion using the bootstrap

Hypothesis tests for a single proportion using normal distributions

Does the AstraZeneca vaccine cause blood clots?

A BBC news article from April 8th discussed the AstraZeneca vaccine.

The article stated…

The Medicines and Health Regulatory Agency (MHRA) looked into UK cases of rare blood clots in people who had recently received the Oxford-AstraZeneca vaccine…

Covid: How does the Oxford-AstraZeneca vaccine work?

4 days ago

Coronavirus pandemic

People under the age of 30 are to be offered an alternative to the Oxford-AstraZeneca vaccine, after a review into a possible link with extremely rare blood clots in adults.

Does the AstraZeneca vaccine cause blood clots?

Let's we run a hypothesis test to assess if people who received the vaccine had a higher rate of blood clots than expected

It found that 79 people experienced clots after receiving a first vaccine dose.

More than 20 million AstraZeneca vaccines doses had been administered across the UK by the end of March.

- The first doses were administered in December.

About four people in a million would normally be expected to develop this particular kind of blood clot - though the fact they are so rare makes the usual rate hard to estimate.

- Note: this 4 in a million is the rate that occurs over a year.

Does the AstraZeneca vaccine cause blood clots?

Step 1: State the null and alternative hypotheses

The MHRA found that 79 people experienced clots after receiving a first vaccine dose.

More than 20 million AstraZeneca vaccines doses had been administered across the UK by the end of March.

- The first doses were administered in December.

About four people in a million would normally be expected to develop this particular kind of blood clot - though the fact they are so rare makes the usual rate hard to estimate.

- Note: this 4 in a million is the rate that occurs over a year.

Does the AstraZeneca vaccine cause blood clots?

Step 1: State the null and alternative hypotheses

About four people in a million would normally be expected to develop this particular kind of blood clot - though the fact they are so rare makes the usual rate hard to estimate.

- Note: This is 4 in a million over the course of a year
- The first doses were administered in December (4 months ago)

H_0: $\pi = 1/1{,}000{,}000$

H_A: $\pi > 1/1{,}000{,}000$

Does the AstraZeneca vaccine cause blood clots?

Step 2: Calculate the observed statistic

The MHRA found that 79 people experienced clots after receiving a first vaccine dose.

More than 20 million AstraZeneca vaccines doses had been administered across the UK by the end of March.

- The first doses were administered in December.

About four people in a million would normally be expected to develop this particular kind of blood clot - though the fact they are so rare makes the usual rate hard to estimate.

- Note: this 4 in a million is the rate that occurs over a year.

Does the AstraZeneca vaccine cause blood clots?

Step 2: Calculate the observed statistic

$\hat{p}$ = 79/20,000,000 = 3.96×10^{-6}

Can you calculate the corresponding z-statistic?

$$z = \frac{\hat{p} - \pi_0}{\sqrt{\frac{\pi_0(1-\pi_0)}{n}}}$$

$n = 20 \times 10^{6}$

$\pi_0 = 10^{-6}$

Does the AstraZeneca vaccine cause blood clots?

```
n <- 20 * 10^6

pi_0 <- 1/10^6

p_hat <- 79/(20 * 10^6)

SE <- sqrt(   (pi_0  * (1 - pi_0))/n    )

z_stat <- (p_hat - pi_0)/SE
```

$$z = \frac{\hat{p} - \pi_0}{\sqrt{\frac{\pi_0(1-\pi_0)}{n}}}$$

13.19

Does the AstraZeneca vaccine cause blood clots?

Is a normal distribution valid to use as a null distribution here?

The normal approximation is reasonable good when we see 10 “positive” outcomes and 10 “negative” outcomes

$n\pi \geq 10$ and $n(1 - \pi) \geq 10$

Are these conditions met?

n * pi_0 n * (1 - pi_0)

Does the AstraZeneca vaccine cause blood clots?

Step 3: Create the null distribution

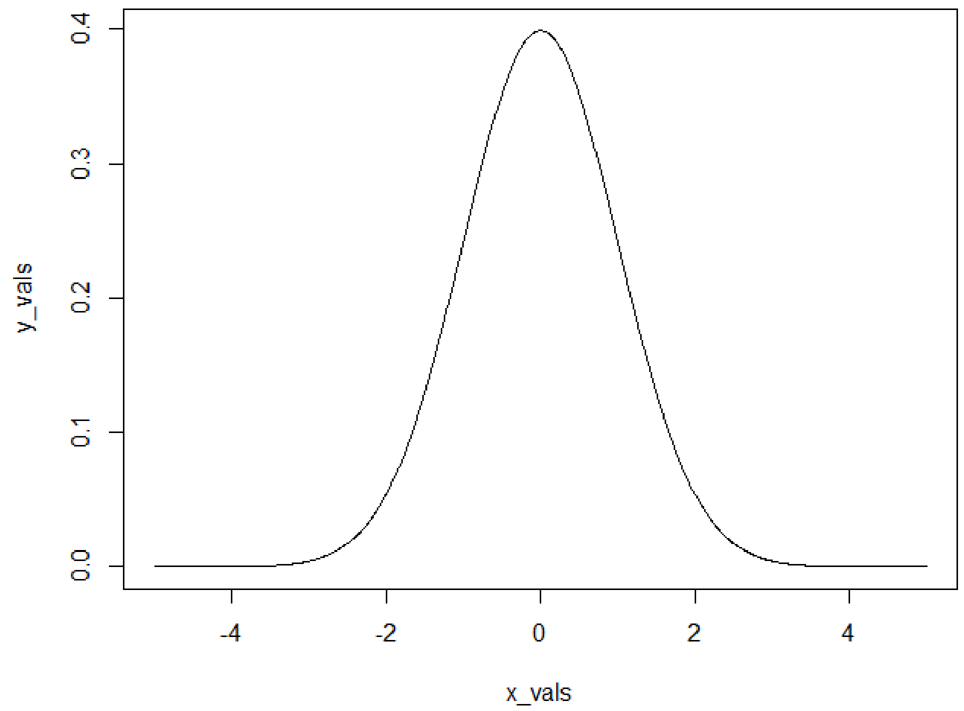
What type of distribution is the appropriate distribution?

Can you plot this distribution in R?

Does the AstraZeneca vaccine cause blood clots?

Step 3: Create the null distribution

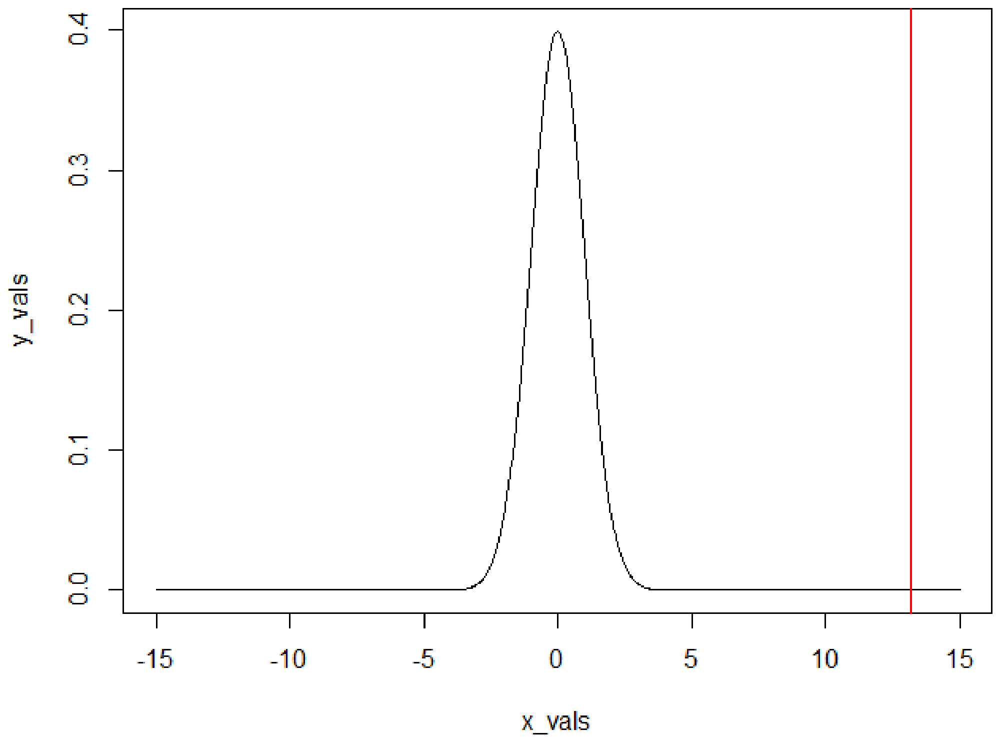
```
x_vals <- seq(-5, 5, length.out = 1000)

y_vals <- dnorm(x_vals, 0, 1)

plot(x_vals, y_vals, type = "l")
```



Does the AstraZeneca vaccine cause blood clots?

Step 4: Calculate the p-value

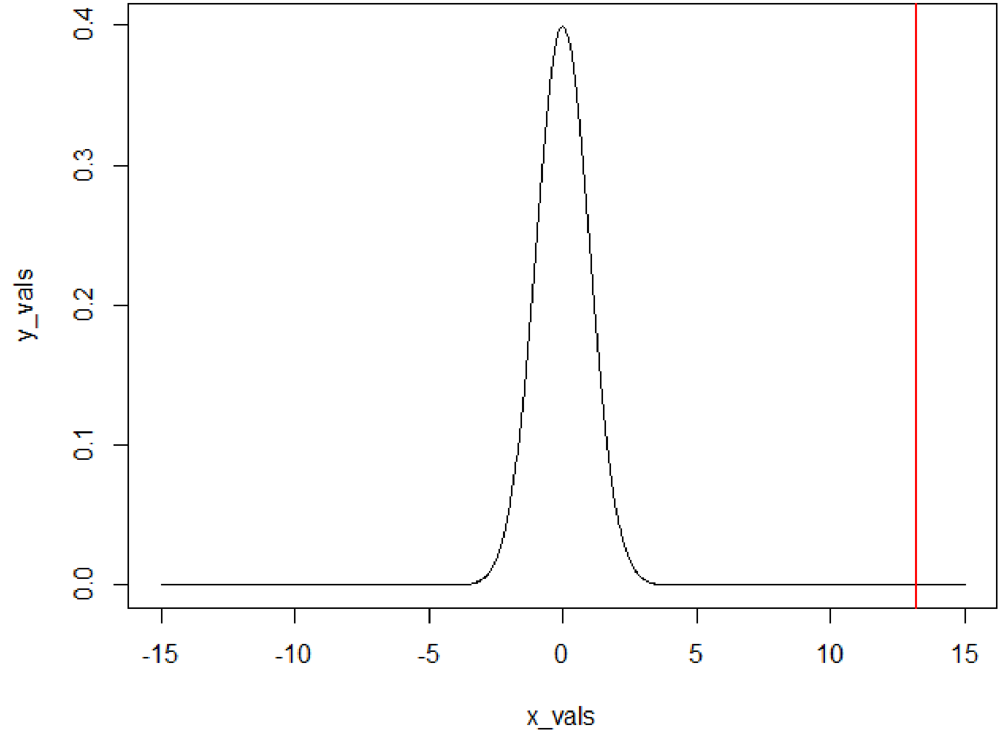
```
x_vals <- seq(-15, 15, length.out = 1000)

y_vals <- dnorm(x_vals, 0, 1)

plot(x_vals, y_vals, type = "l")

abline(v = z_stat, col = "red")
```



Does the AstraZeneca vaccine cause blood clots?

Step 4: Calculate the p-value

```
pnorm(z_stat, 0, 1, lower.tail = FALSE)
```



Does the AstraZeneca vaccine cause blood clots?

Step 5: Make a decision

Does the AstroZeneca vaccine cause blood clots?

Should you take the AstroZeneca vaccine?

Confidence intervals for a single proportion using normal distributions

What is the probability of having a blood clot if you take the AstraZeneca vaccine?

Let's create a 95% confidence interval for the likelihood of having a blood clot if you take the AstraZeneca vaccine

The MHRA found that 79 people experienced clots after receiving a first vaccine dose.

More than 20 million AstraZeneca vaccines doses had been administered across the UK by the end of March.

- The first doses were administered in December.

About four people in a million would normally be expected to develop this particular kind of blood clot - though the fact they are so rare makes the usual rate hard to estimate.

- Note: this 4 in a million is the rate that occurs over a year.

What is the probability of having a blood clot if you take the AstraZeneca vaccine?

$$\hat{p} \pm z^* \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$\hat{p} = 79/20{,}000{,}000 = 3.96 \times 10^{-6}$

$n = 20 \times 10^{6}$

What is the probability of having a blood clot if you take the AstraZeneca vaccine?

$$\hat{p} \pm z^* \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

```
z_star <- qnorm(.975)

p_hat - z_star * SE
p_hat + z_star * SE
```

[3.51×10^{-6} 4.39×10^{-6}] in 4 months

[1.05×10^{-5} 1.31×10^{-5}] in a year

Does the AstraZeneca vaccine cause blood clots?

Make a decision!

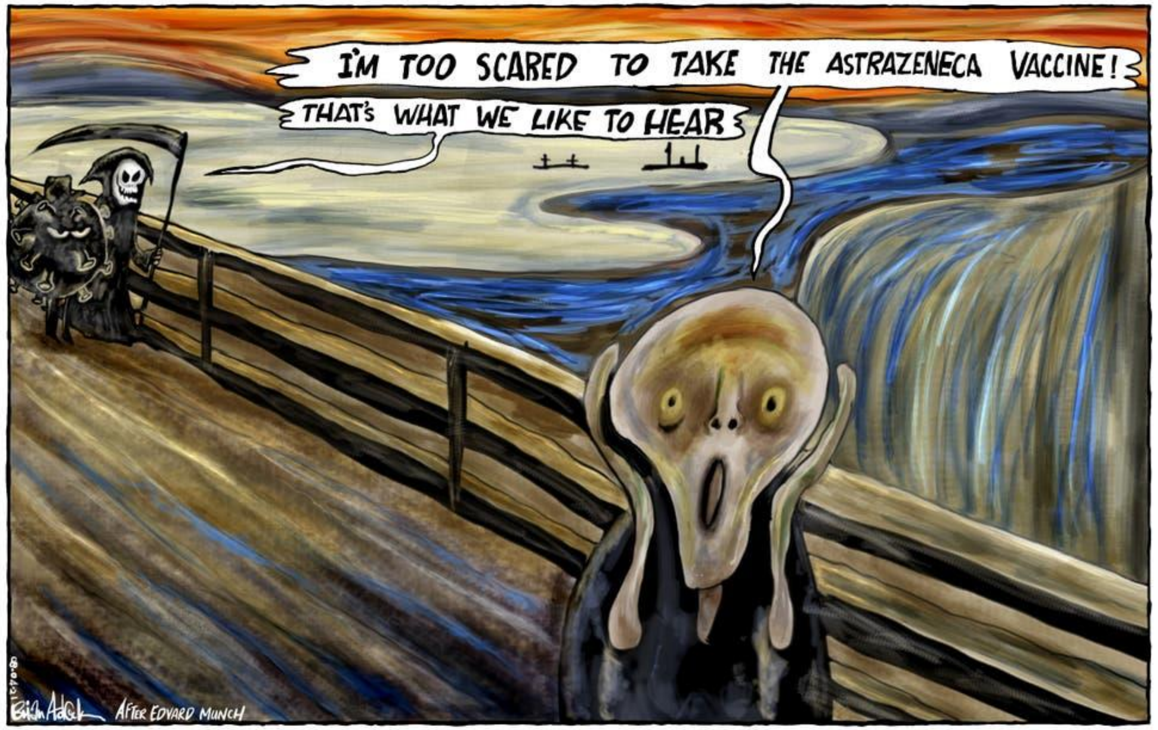
Should you take the AstroZeneca vaccine?

INDEPENDENT PREMIUM

Fear of blood clots must not undermine the case for vaccinations

Editorial: The possibility of getting fatal blood clots from Covid are vastly higher than from the Oxford-AstraZeneca jab. We must avoid a catastrophic third wave, economic recession, and help build herd immunity

4 days ago | 43 comments



Confidence intervals for a single proportion using the bootstrap

What is the probability of having a blood clot if you take the AstraZeneca vaccine?

Let's create a 95% confidence interval for the probability of having a blood clot if you take the AstraZeneca vaccine using the bootstrap

The MHRA found that 79 people experienced clots after receiving a first vaccine dose.

More than 20 million AstraZeneca vaccines doses had been administered across the UK by the end of March.

- The first doses were administered in December.

About four people in a million would normally be expected to develop this particular kind of blood clot - though the fact they are so rare makes the usual rate hard to estimate.

- Note: this 4 in a million is the rate that occurs over a year.

What is the probability of having a blood clot if you take the AstraZeneca vaccine?

How should we start this analysis?

```
library(SDS100)

has_clot <- rep(TRUE, 79)
no_clot <- rep(FALSE, (20 * 10^6) - 79)

data_vec <- c(has_clot, no_clot)
```

What is the probability of having a blood clot if you take the AstraZeneca vaccine?

Let's create **one** bootstrap statistic:

```
boot_data <- sample(data_vec, replace = TRUE)

boot_proportion <- mean(boot_data)
```

What is the probability of having a blood clot if you take the AstraZeneca vaccine?

Let's do it 10 times!

```
boot_dist <- do_it(10) * {

  boot_data <- sample(data_vec, replace = TRUE)
  boot_proportion <- mean(boot_data)

}
```

Would doing this 10,000 times be reasonable?

What is the probability of having a blood clot if you take the AstraZeneca vaccine?

```
install.packages(tictoc)
library(tictoc)

tic()

boot_dist <- do_it(10) * {

  boot_data <- sample(data_vec, replace = TRUE)
  boot_proportion <- mean(boot_data)

}

toc()
```

What is the probability of having a blood clot if you take the AstraZeneca vaccine?

It took me around 19 seconds for 10 runs

So it would take over 5 hours for the typical 10,000 runs we use

Any other questions about anything?