

<u>Course</u> > <u>Course Wrap-up an...</u> > <u>Comprehensive Ass...</u> > Brexit poll analysis ...

# Brexit poll analysis - Part 1 Directions

There are 12 multi-part problems in this comprehensive assessment that review concepts from the entire course. The problems are split over 3 pages. Make sure you read the instructions carefully and run all pre-exercise code.

For numeric entry problems, you have 10 attempts to input the correct answer. For true/false problems, you have 2 attempts.

If you have questions, visit the "Brexit poll analysis" discussion forum that follows the assessment.

IMPORTANT: Some of these exercises use **dslabs** datasets that were added in a July 2019 update. Make sure your package is up to date with the command update.packages("dslabs"). You can also update all packages on your system by running update.packages() with no arguments, and you should consider doing this routinely.

#### Overview

In June 2016, the United Kingdom (UK) held a referendum to determine whether the country would "Remain" in the European Union (EU) or "Leave" the EU. This referendum is commonly known as Brexit. Although the media and others interpreted poll results as forecasting "Remain" (p>0.5), the actual proportion that voted "Remain" was only 48.1% (p=0.481) and the UK thus voted to leave the EU. Pollsters in the UK were criticized for overestimating support for "Remain".

In this project, you will analyze real Brexit polling data to develop polling models to forecast Brexit results. You will write your own code in R and enter the answers on the edX platform.

# Important definitions

# Data Import

Import the brexit polls polling data from the **dslabs** package and set options for the analysis:

```
# suggested libraries and options
library(tidyverse)
options(digits = 3)
```

```
# load brexit_polls object
library(dslabs)
data(brexit_polls)
```

# **Final Brexit parameters**

Define p=0.481 as the actual percent voting "Remain" on the Brexit referendum and d=2p-1=-0.038 as the actual spread of the Brexit referendum with "Remain" defined as the positive outcome:

```
p <- 0.481  # official proportion voting "Remain"
d <- 2*p-1  # official spread</pre>
```

# Question 1: Expected value and standard error of a poll

6.0/6.0 points (graded)

The final proportion of voters choosing "Remain" was p=0.481. Consider a poll with a sample of N=1500 voters.

What is the expected total number of voters in the sample choosing "Remain"?

722 **Answer**: 722

## **Explanation**

You can calculate the expected number of "Remain" voters with the following code:

```
p <- 0.481
N <- 1500
N*p
```

What is the standard error of the total number of voters in the sample choosing "Remain"?

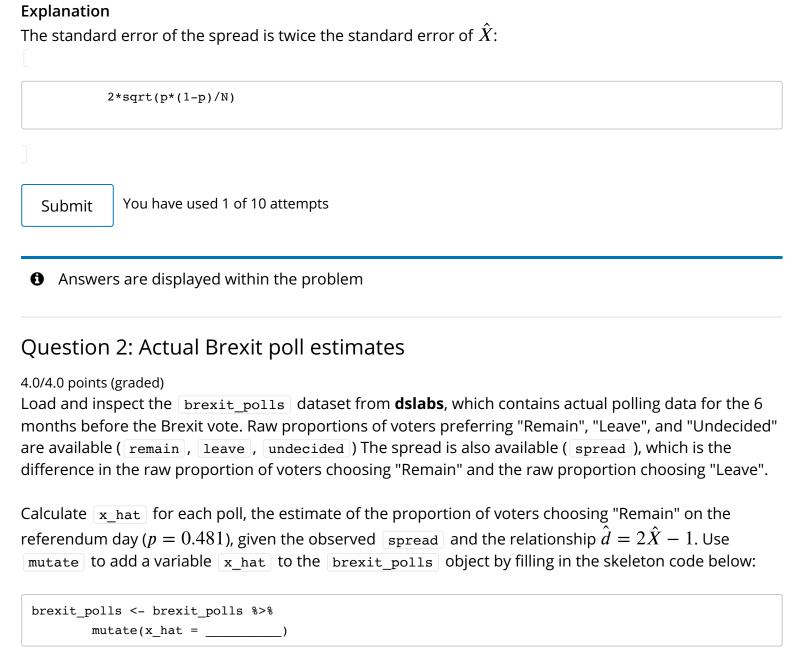
```
19.4 Answer: 19.4
```

## **Explanation**

You can calculate the standard error of the expected number of "Remain" voters with the following code:

```
sqrt(N*p*(1-p))
```

What is the expected value of $\hat{X}$ , the proportion of "Remain" voters?	
0.481 <b>✓ Answer:</b> 0.481	
0.481	
Explanation	
The expected value of $\hat{X}$ is $p = 0.481$ .	
What is the standard error of $\hat{X}$ , the proportion of "Remain" voters?	
0.0129 <b>✓ Answer:</b> 0.0129	
0.0129	
Explanation	
You can calculate the standard error $\hat{X}$ with the following code:	
cgrt (n*(1 n)/N)	
sqrt(p*(1-p)/N)	
What is the expected value of $d$ , the spread between the proportion of "Remain" voters and "Leav voters?	/e''
-0.038 <b>✓ Answer:</b> -0.038	
-0.038	
0.030	
<b>Explanation</b> Given the proportion $p$ , the expected value of the spread is $2p-1$ :	
2*p-1	
What is the standard error of $d$ , the spread between the proportion of "Remain" voters and "Leav	⁄e"
voters?	
0.0258 <b>✓ Answer:</b> 0.0258	
0.0258	



What is the average of the observed spreads ( spread )?

0.0201

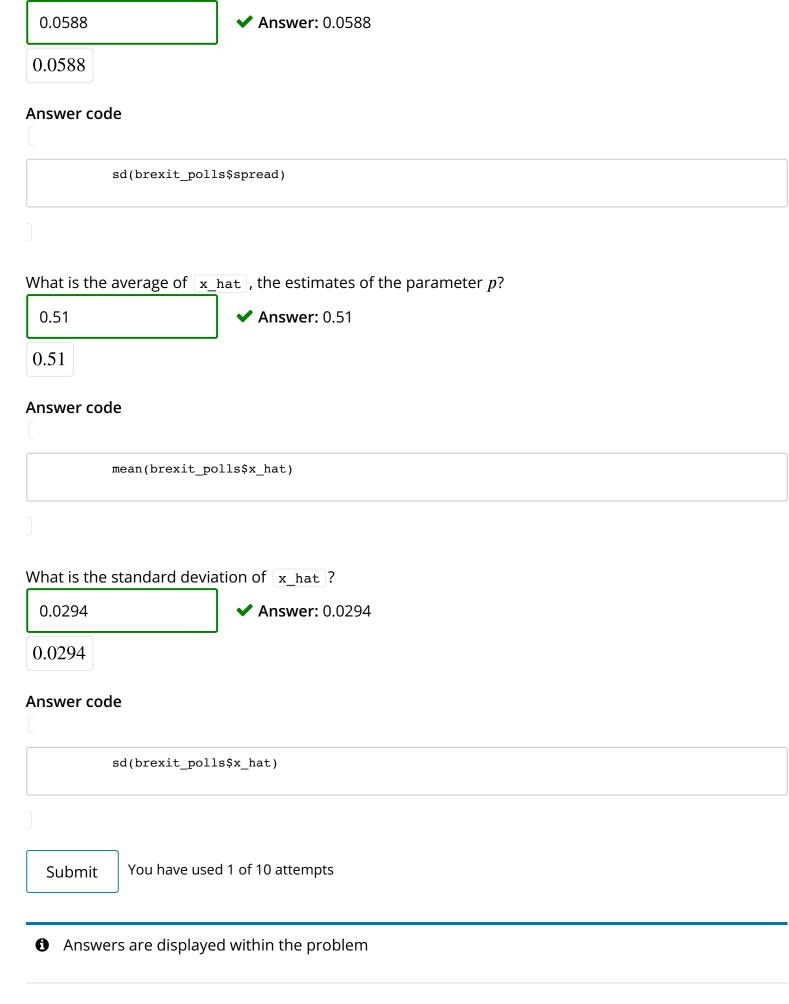
**✓ Answer:** 0.0201

0.0201

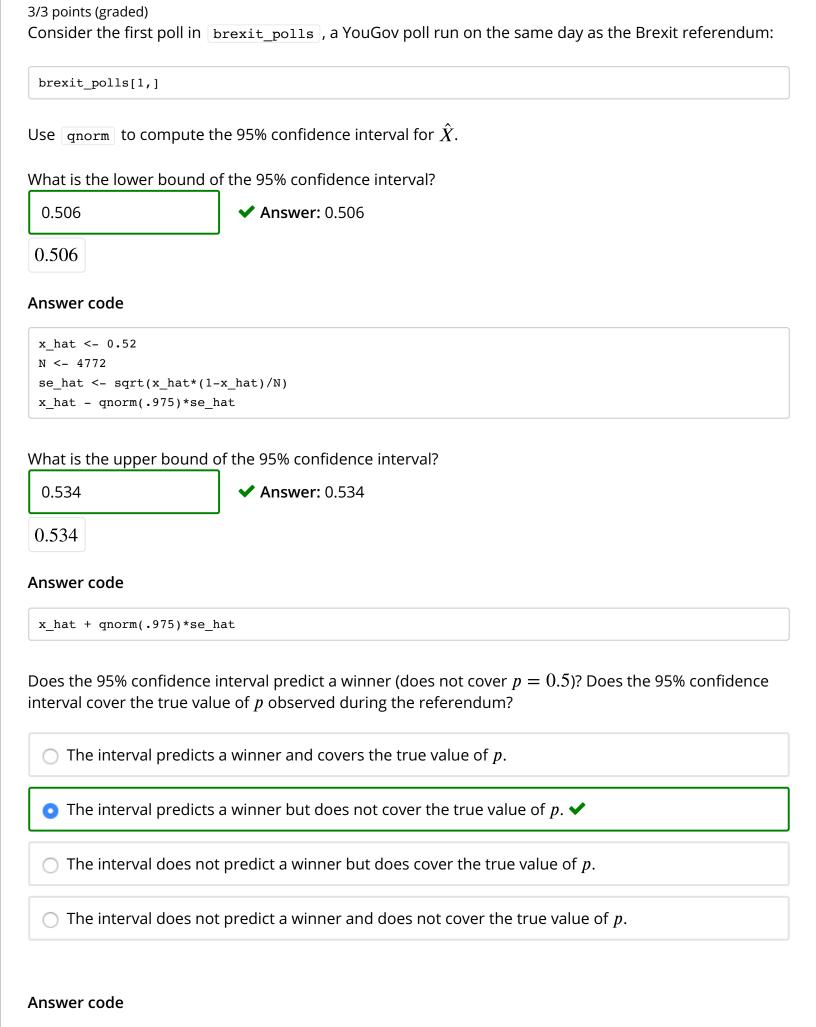
#### Answer code

```
brexit_polls <- brexit_polls %>%
mutate(x_hat = (spread + 1)/2)
mean(brexit_polls$spread)
```

What is the standard deviation of the observed spreads?



Question 3: Confidence interval of a Brexit poll



!between(0.5, x\_hat - qnorm(.975)\*se\_hat, x\_hat + qnorm(.975)\*se\_hat) # predicts winner between(0.481, x\_hat - qnorm(.975)\*se\_hat, x\_hat + qnorm(.975)\*se\_hat) # does not cover p

Submit

You have used 1 of 10 attempts

**1** Answers are displayed within the problem

Continue the comprehensive assessment on the next page.

© All Rights Reserved