Hellenic Complex Systems Laboratory

Inferences about the Difference between Two Proportions

Technical Report XVIII

Theodora Chatzimichail and Aristides T. Hatjimihail
2018

Inferences about the Difference between Two Proportions

Theodora Chatzimichail [a] and Aristides T. Hatjimihail (Aristeidis T. Chatzimichail) [a]

[a] Hellenic Complex Systems Laboratory

Short Description of the Demonstration

This Demonstration explores statistical inferences regarding the difference between two proportions of populations obeying a condition (or trait). It calculates the statistical significance and the confidence intervals for the difference and plots the confidence intervals versus *p*-value. This is done for differing populations obeying a condition or without the condition and for differing *p*-values for the estimation of the confidence intervals.

first proportion	0.900
second proportion	0.850
difference (point estimation)	−0.050
statistical significance of the difference	0.886

99. percent confidence interval for the difference	
lower bound	−0.127
upper bound	0.027

first population

population obeying the condition
⊞ 90

population without the condition
⊞ 10

second population

population obeying the condition
⊞ 850

population without the condition
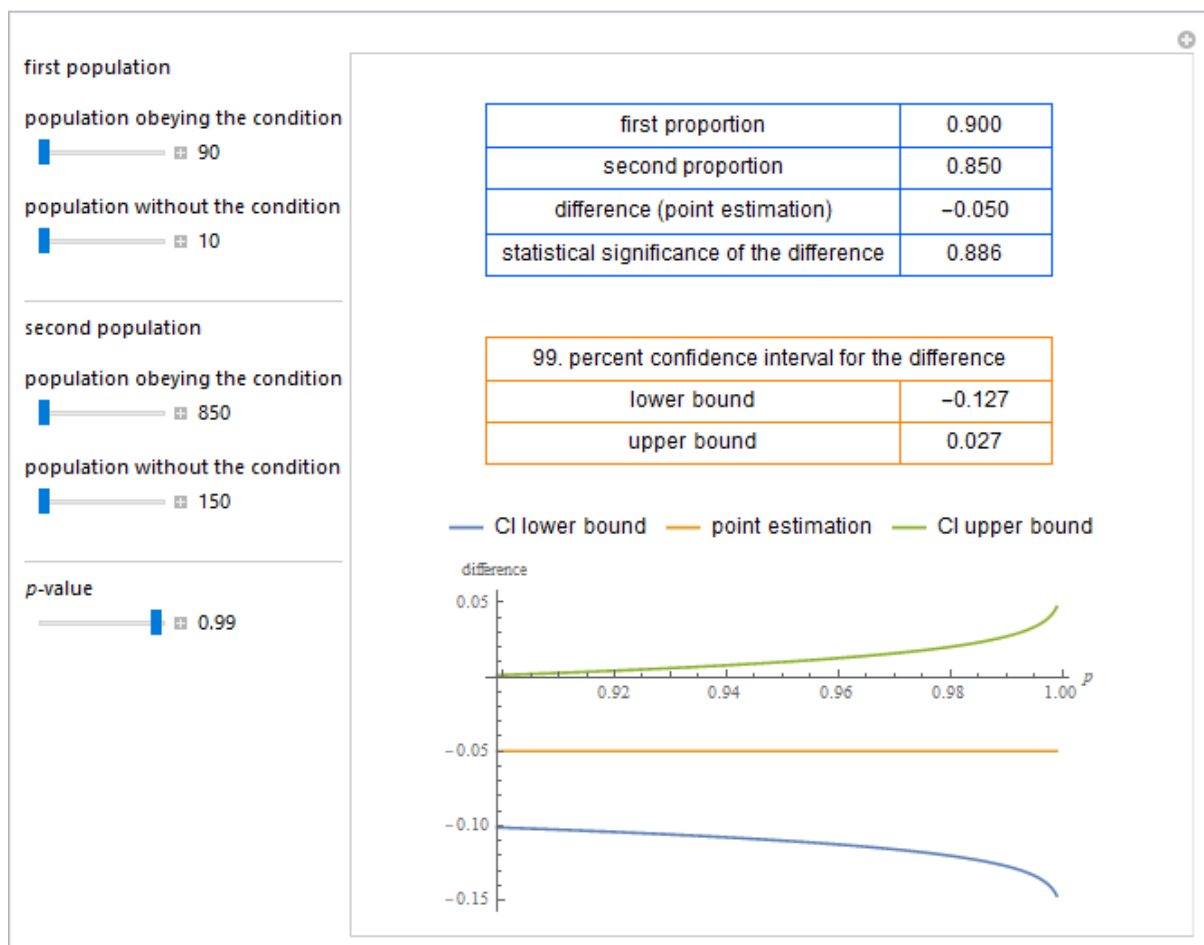⊞ 150

p-value
⊞ 0.99

Figure 1: *Point estimation and 99% confidence interval for the difference of two proportions of populations obeying a condition, as well as their plots versus p-value, with the settings shown at the left.*
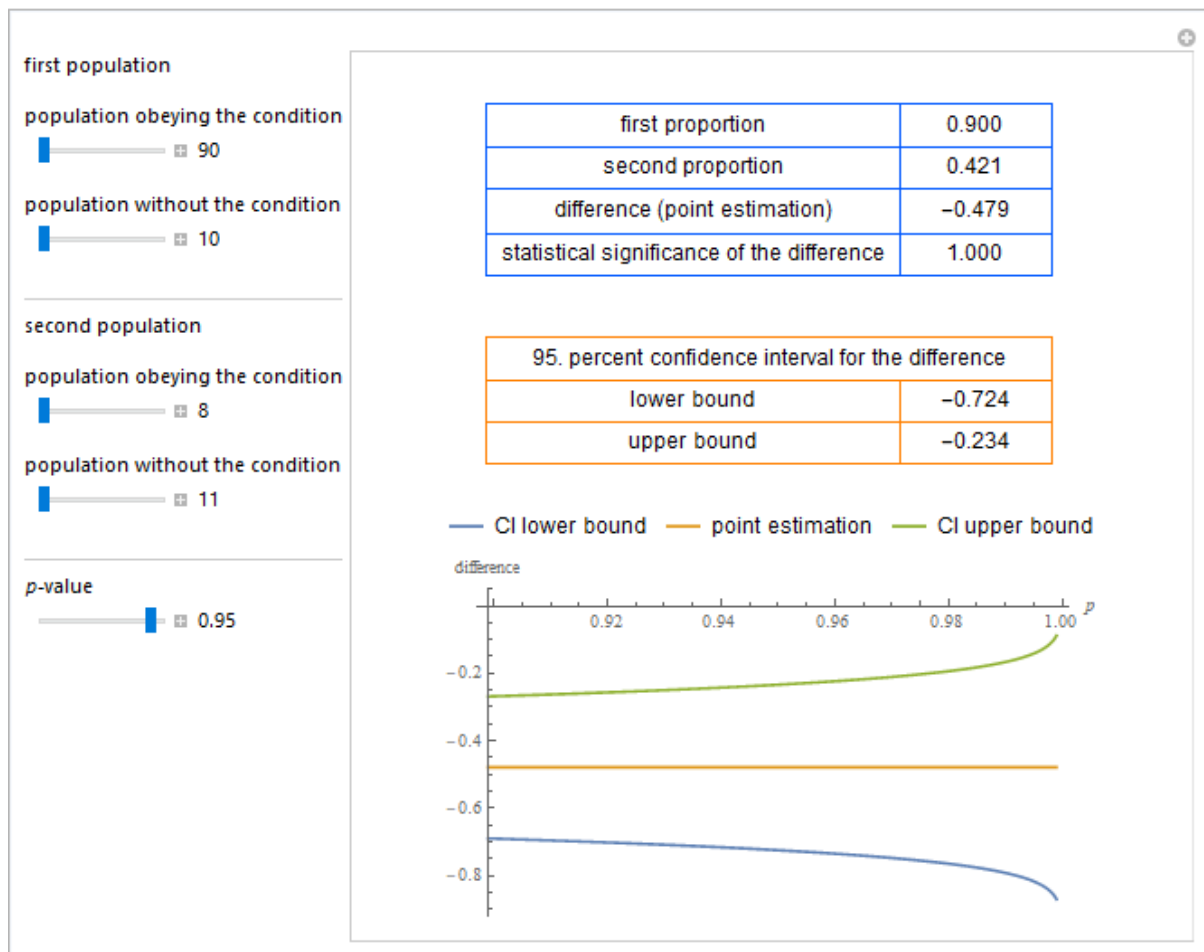
Figure 2: Point estimation and 95% confidence interval for the difference of two proportions of populations obeying a condition, as well as their plots versus p-value, with the settings shown at the left.

Figure 3: *Point estimation and 95% confidence interval for the difference of two proportions of populations obeying a condition, as well as their plots versus p-value, with the settings shown at the left.*
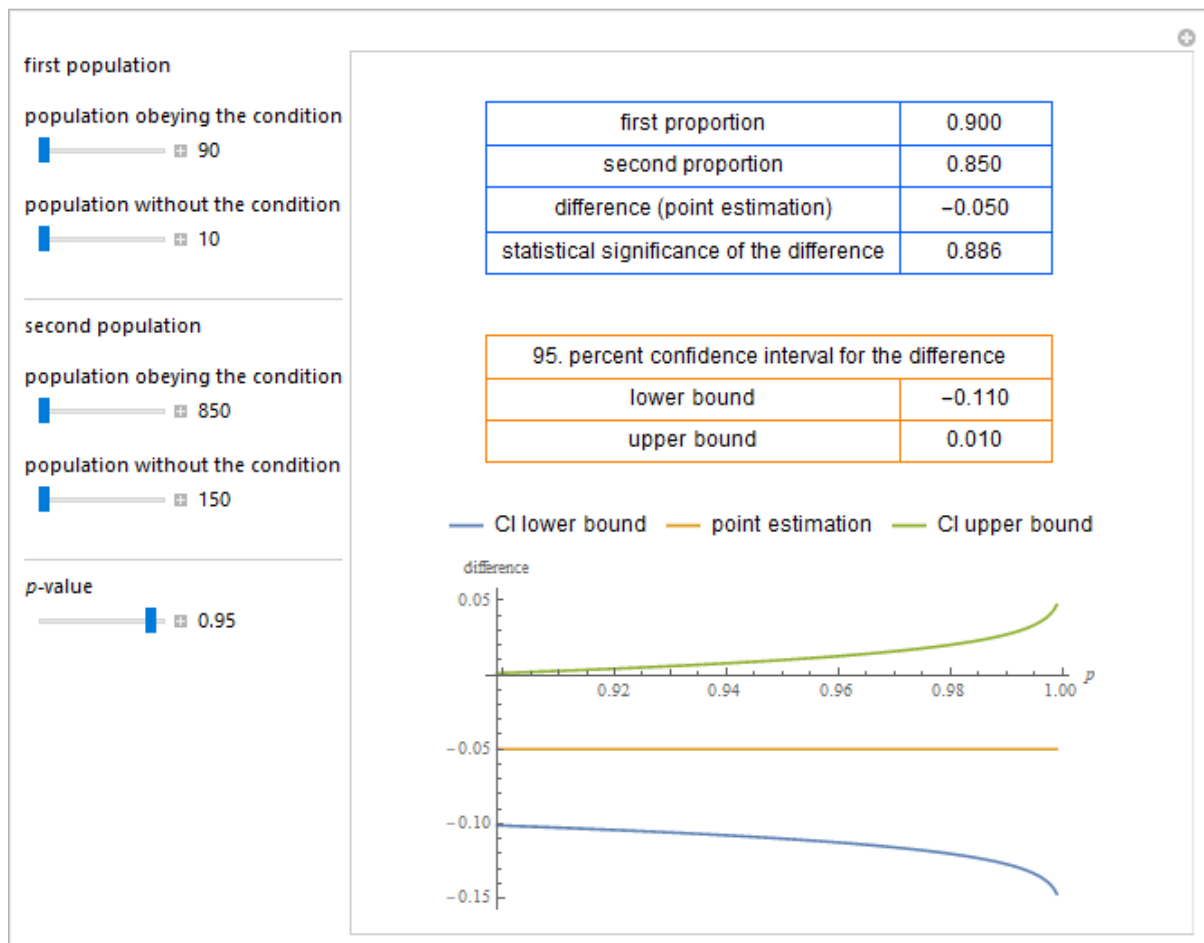
first population		
population obeying the condition		95
population without the condition		15

second population

population obeying the condition		88
population without the condition		27

p-value   0.99

first proportion	0.864
second proportion	0.765
difference (point estimation)	−0.098
statistical significance of the difference	0.958

99. percent confidence interval for the difference	
lower bound	−0.164
upper bound	−0.032

— CI lower bound — point estimation — CI upper bound

Figure 4: *Point estimation and 99% confidence interval for the difference of two proportions of populations obeying a condition, as well as their plots versus p-value, with the settings shown at the left.*
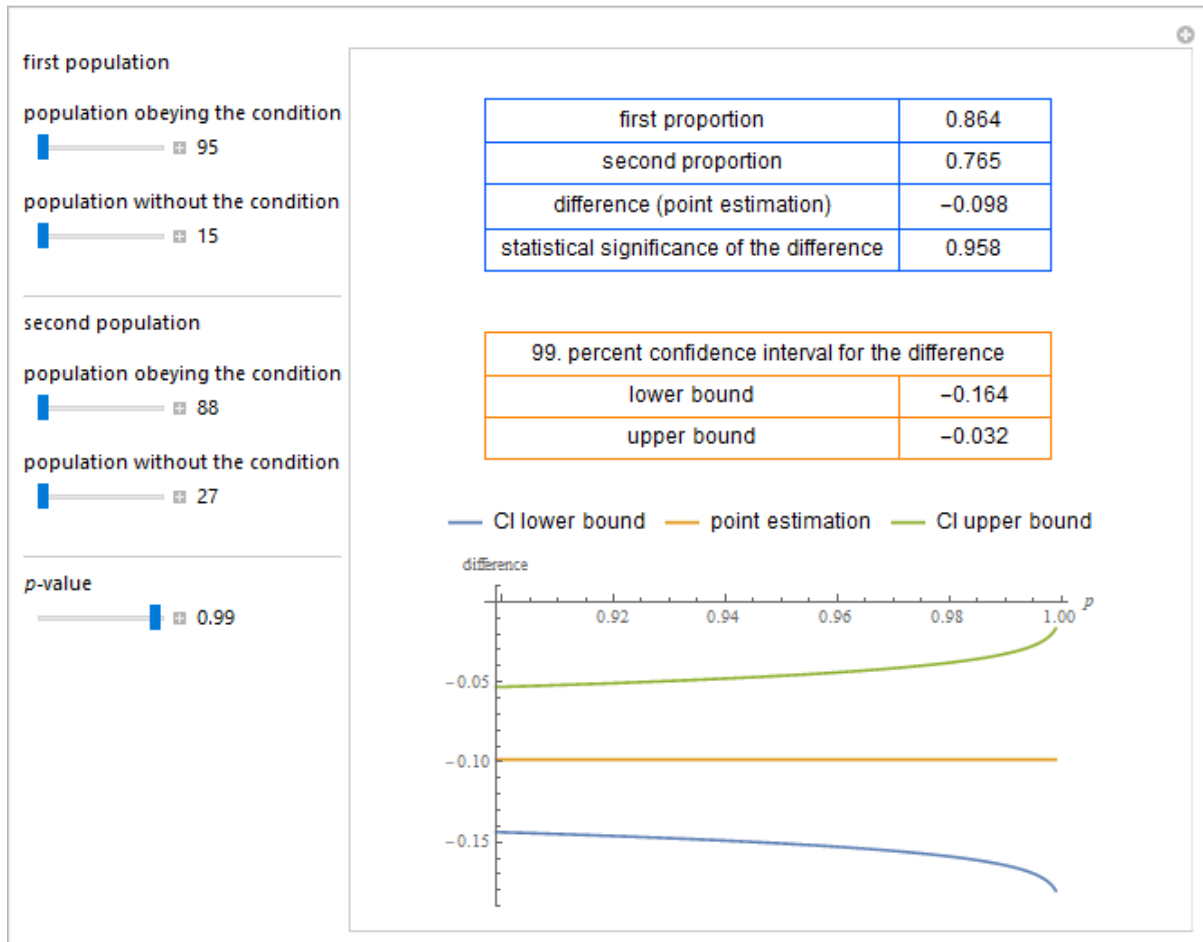
5

Details

The statistical significance of the difference between the two proportions is assessed by means of the z-score (or critical ratio statistic) [1]. The confidence interval for the difference between the proportions is calculated as in [2].

References

[1] J. L. Fleiss, B. Levin and M. C. Paik. Methods for Generating a Fourfold Table. *Statistical Methods for Rates and Proportions*, 3rd ed., Hoboken, NJ: J. Wiley, 2003 pp. 51–52.

[2] J. L. Fleiss, B. Levin and M. C. Paik. A Simple Confidence Interval for the Difference between Two Independent Proportions. *Statistical Methods for Rates and Proportions*, 3rd ed., Hoboken, NJ: J. Wiley, 2003 pp. 60–61.

Source Code

The updated Wolfram Mathematica© source code is available at:

https://www.hcsl.com/Tools/InferencesAboutTheDifferenceBetweenTwoProportions-author.nb

Permanent Citation:

Chatzimichail T, Hatjimihail AT. Inferences about the Difference between Two Proportions. Wolfram Demonstrations Project, Champaign: Wolfram Research, Inc., 2018. Available at:

http://demonstrations.wolfram.com/InferencesAboutTheDifferenceBetweenTwoProportions/

First Published: August 2 2018