

# Single-Sample Statistical Tests with a Binary Dependent Variable

Riley Smith

06 Nov 2016

## Contents

z-Score Test 1

References 3

See [Newsom 2016-CDA Handout-3](#)

## z-Score Test

Here we are taking one of the favorability proportions (i.e. “success proportions”) and comparing it to the *Null Hypothesis* ( $H_0$ ) represented by  $\pi_0$  (i.e., 0.50).<sup>1</sup>

<sup>1</sup> Agresti, *Categorical Data Analysis*; Agresti and Coull, “Approximate Is Better Than ‘Exact’ for Interval Estimation of Binomial Proportions.”  
Data: [2016 Polling data](#)

```
n <- 1231

zp <- .55 ## 55% Clinton favorability ("p") ##
zpi <- .50 ## Null Hypothesis difference between samples in the population "\pi_{0}" ##

z1 <- zp - zpi ## numerator: p-\pi_{0} ##

z2a <- 1 - zpi ## "1-\pi_{0}" ##
z2b <- zpi*z2a ## "\pi(1-\pi_{0})" ##
z2c <- z2b/n ## "\pi(1-\pi_{0}) / n" ##
z2 <- sqrt(z2c) ## denominator: "\sqrt{\pi(1-\pi_{0}) / n}" ##

zscore <- z1/z2 ##
##      p-\pi_{0}      ##
## -----      ##
## (\sqrt{\pi(1-\pi_{0}) / n})      ##

zscore

[1] 3.508561
```

## z-Score Test: Lower and Upper Confidence Limits

Confidence limits are calculated by the favorability proportion ( $p$ )  $\pm$  the  $z_{critical}$  value multiplied by the standard error of the estimate ( $SE_{\pi}$ ).

$$CI = p \pm (z_{critical})(SE_{\pi})$$

```
zcr <- 1.96
SEa <- 1 - zp
SEb <- zp*SEa
SEc <- SEb/n
SE <- sqrt(SEc)
```

```
LCL <- .55 - (zcr)*(SE)
UCL <- .55 + (zcr)*(SE)
```

*Goodness-of-Fit Tests ( $\chi^2$ )*

Evaluate the observed  $\chi^2$  value to the  $\chi^2$  distribution ( $f_k(x)$ ).

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$$

```
Oc <- 677
Ot <- 554
```

The  $\chi^2$  test's flexibility allows for additional comparison analyses.  
The Likelihood Ratio  $\chi^2$  is similar to the Pearson  $\chi^2$ .

References<sup>2</sup>

Agresti, Alan. *Catagorical Data Analysis*. Third Edition. New York, NY: Wiley & Sons, 2013.

Agresti, Alan, and Brent A Coull. "Approximate Is Better Than 'Exact' for Interval Estimation of Binomial Proportions." *The American Statistician* 52, no. 2 (1998): 119–26.

Allaire, JJ, Joe Cheng, Yihui Xie, Jonathan McPherson, Winston Chang, Jeff Allen, Hadley Wickham, Aron Atkins, and Rob Hyndman. *rmarkdown: Dynamic Documents for R*, 2016. <https://CRAN.R-project.org/package=rmarkdown>.

Arnold, Jeffrey B. *Ggthemes: Extra Themes, Scales and Geoms for Ggplot2*, 2016. <https://CRAN.R-project.org/package=ggthemes>.

Aust, Frederik, and Marius Barth. *Papaja: Create APA Manuscripts with RMarkdown*, 2015. <https://github.com/crsh/papaja>.

Boettiger, Carl. *knitcitations: Citations for Knitr Markdown Files*, 2015. <https://CRAN.R-project.org/package=knitcitations>.

Chang, Winston. *Extrafont: Tools for Using Fonts*, 2014. <https://CRAN.R-project.org/package=extrafont>.

Daroczi, Gergely, and Roman Tsegelskyi. *Pander: An R Pandoc Writer*, 2015. <https://CRAN.R-project.org/package=pander>.

Francois, Romain. *Bibtex: Bibtex Parser*, 2014. <https://CRAN.R-project.org/package=bibtex>.

———. *Highlight: Syntax Highlighter*, 2015. <https://CRAN.R-project.org/package=highlight>.

Qiu, Yixuan. *Showtext: Using Fonts More Easily in RGraphs*, 2015. <https://CRAN.R-project.org/package=showtext>.

Qiu, Yixuan, and others. *Sysfonts: Loading System Fonts into R*, 2015. <https://CRAN.R-project.org/package=sysfonts>.

R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing, 2016. <https://www.R-project.org/>.

Wickham, Hadley. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York, 2009. <http://ggplot2.org>.

———. *Gtable: Arrange Grobs in Tables*, 2016. <https://CRAN.R-project.org/package=gtable>.

———. *Scales: Scale Functions for Visualization*, 2016. <https://CRAN.R-project.org/package=scales>.

———. "The Split-Apply-Combine Strategy for Data Analysis." *Journal of Statistical Software* 40, no. 1 (2011): 1–29. <http://www.jstatsoft.org/v40/i01/>.

———. *Tidyr: Easily Tidy Data with Spread() and Gather() Functions*, 2016. <https://CRAN.R-project.org/package=tidyr>.

Wickham, Hadley, and Winston Chang. *Devtools: Tools to Make*

<sup>2</sup> **Note:** This document was created using R-v3.3.2 R Core Team, R, and the following R-packages: *base-v3.3*. R Core Team, R, *bibtex-v0.4*. Francois, *Bibtex*, *dplyr-v0.5*. Wickham and Francois, *Dplyr*, *DT-v0.2*. Xie, *DT*, *extrafont-v0.17*. Chang, *Extrafont*, *ggplot2-v2.1*. Wickham, *Ggplot2*, *knitcitations-v1.0*. Boettiger, *knitcitations*, *knitr-v1.14*. Xie, *Dynamic Documents with R and Knitr*, *pander-v0.6*. Daroczi and Tsegelskyi, *Pander*, *papaja-v0.1*. Aust and Barth, *Papaja*, *plyr-v1.8*. Wickham, "The Split-Apply-Combine Strategy for Data Analysis.", *rmarkdown-v1.1*. Allaire et al., *rmarkdown*, *scales-v0.4*. Wickham, *Scales*, *tidyr-v0.6*. Wickham, *Tidyr*, *ggthemes-v3.2*. Arnold, *Ggthemes*, *gtable-v0.2*. Wickham, *Gtable*, *kableExtra-v0.0*. Zhu, *KableExtra*, *tuftes-v0.2*. Xie and Allaire, *Tuftes*, *devtools-v1.12*. Wickham and Chang, *Devtools*, *highlight-v0.4*. Francois, *Highlight*, *sysfonts-v0.5*. Qiu and others, *Sysfonts*, and *showtext-v0.4*. Qiu, *Showtext*

*Developing R Packages Easier*, 2016. <https://CRAN.R-project.org/package=devtools>.

Wickham, Hadley, and Romain Francois. *Dplyr: A Grammar of Data Manipulation*, 2015. <https://CRAN.R-project.org/package=dplyr>.

Xie, Yihui. *DT: A Wrapper of the Javascript Library Datatables*, 2015. <https://CRAN.R-project.org/package=DT>.

———. *Dynamic Documents with R and Knitr*. 2nd ed. Boca Raton, Florida: Chapman; Hall/CRC, 2015. <http://yihui.name/knitr/>.

Xie, Yihui, and JJ Allaire. *Tufte: Tufte's Styles for Rmarkdown Documents*, 2016. <https://CRAN.R-project.org/package=tufte>.

Zhu, Hao. *KableExtra: Decorate Kable Output Using Pipe Syntax*, 2016. <https://github.com/haozhu233/kableExtra>.