SDM4 in R: Testing Hypotheses about Proportions (Chapter 19)

Nicholas Horton (nhorton@amherst.edu)

June 4, 2018

Introduction and background

This document is intended to help describe how to undertake analyses introduced as examples in the Fourth Edition of *Stats: Data and Models* (2014) by De Veaux, Velleman, and Bock. More information about the book can be found at http://wps.aw.com/aw_deveaux_stats_series. This file as well as the associated R Markdown reproducible analysis source file used to create it can be found at http://nhorton.people.amherst.edu/sdm4.

This work leverages initiatives undertaken by Project MOSAIC (http://www.mosaic-web.org), an NSF-funded effort to improve the teaching of statistics, calculus, science and computing in the undergraduate curriculum. In particular, we utilize the `mosaic` package, which was written to simplify the use of R for introductory statistics courses. A short summary of the R needed to teach introductory statistics can be found in the mosaic package vignettes (http://cran.r-project.org/web/packages/mosaic). A paper describing the mosaic approach was published in the *R Journal*: https://journal.r-project.org/archive/2017/RJ-2017-024.

Chapter 19: Testing hypotheses for proportions

Section 19.1: Hypotheses

We can reproduce the calculation in Figure 19.1 (page 495).

```
sdp <- sqrt(.2*.8/400)
sdp
```

```
## [1] 0.02
```

```
xpnorm(0.17, mean = 0.20, sd = sdp)
```

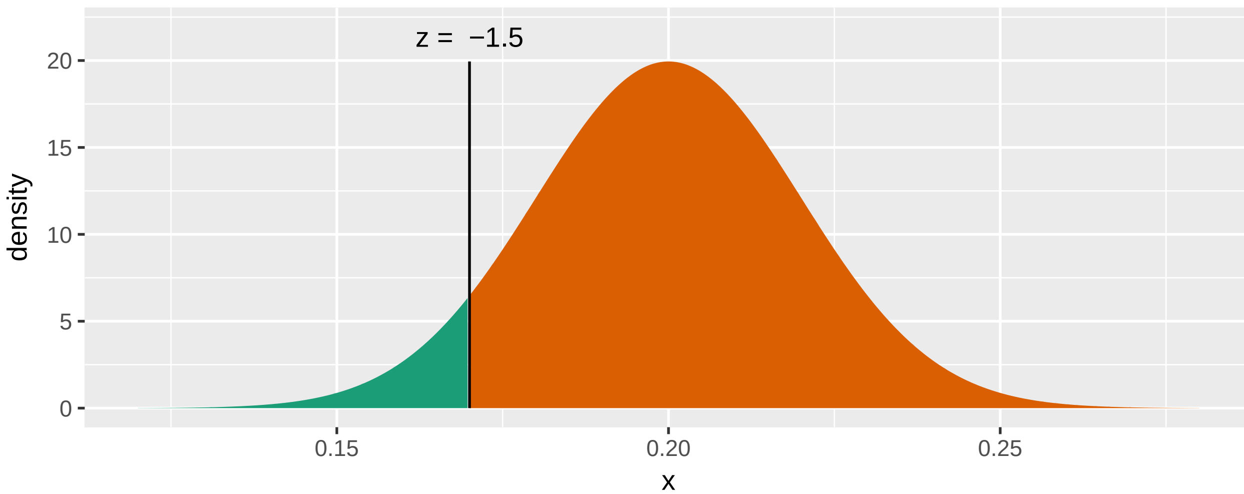
```
##

## If X ~ N(0.2, 0.02), then

##  P(X <= 0.17) = P(Z <= -1.5) = 0.06681

##  P(X >  0.17) = P(Z >  -1.5) = 0.9332

##
```




```
## [1] 0.0668
```

```
zval <- (0.17 - 0.20)/sdp
zval
```

```
## [1] -1.5
```

```
pnorm(zval, mean = 0, sd = 1)
```

```
## [1] 0.0668
```

Section 19.3: Reasoning of hypothesis testing

The "For Example (page 499)" lays out how to find a p-value for the one proportion z-test.

```
y <- 61
n <- 90
phat <- y/n
phat
```

```
## [1] 0.678
```

```
nullp <- 0.8
sdp <- sqrt(nullp*(1-nullp)/n)
sdp
```

```
## [1] 0.0422
```

```
onesidep <- xpnorm(phat, mean = nullp, sd = sdp)
```

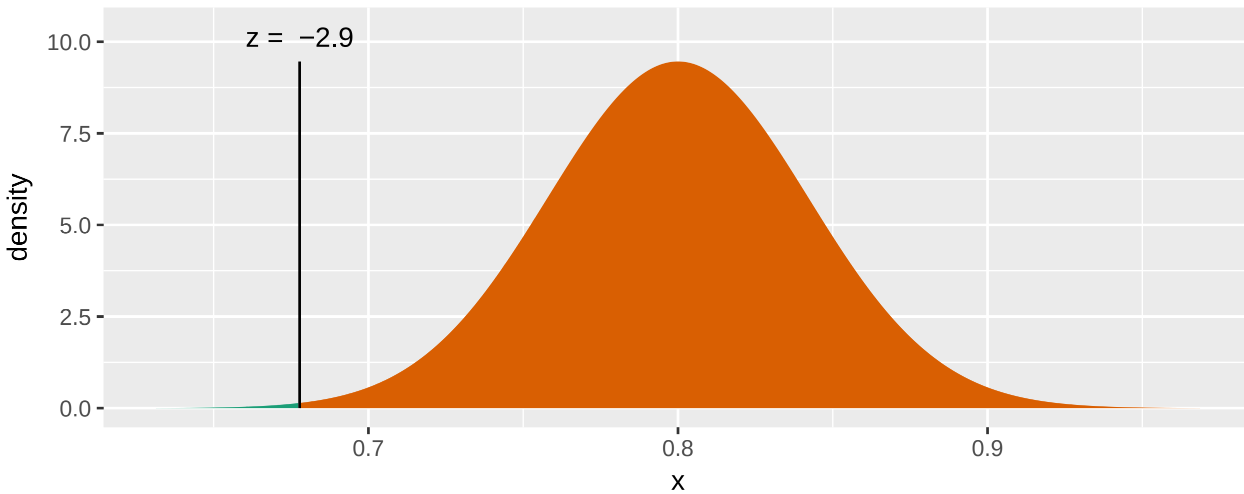
```
##

## If X ~ N(0.8, 0.04216), then
```

```
##  P(X <= 0.6778) = P(Z <= -2.899) = 0.001873

##  P(X >  0.6778) = P(Z >  -2.899) = 0.9981

##
```




```
onesidep
```

```
## [1] 0.00187
```

```
twosidep <- 2*onesidep
twosidep
```

```
## [1] 0.00375
```

or we can carry out the exact test (not described by the book):

```
binom.test(y, n, p = nullp)
```

```
##
##
##
## data:  y out of n
## number of successes = 60, number of trials = 90, p-value = 0.006
## alternative hypothesis: true probability of success is not equal to 0.8
## 95 percent confidence interval:
##  0.571 0.772
## sample estimates:
## probability of success
##                  0.678
```