Testing Hypotheses About Proportions (Chapter 17)

Patrick Frenett and Nicholas Horton (nhorton@amherst.edu)

June 15, 2016

Introduction and background

This document is intended to help describe how to undertake analyses introduced as examples in the Fourth Edition of *Intro Stats* (2013) 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://www.amherst.edu/~nhorton/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).

Note that some of the figures in this document may differ slightly from those in the IS4 book due to small differences in datasets. However in all cases the analysis and techniques in R are accurate.

Chapter 17: Testing Hypotheses About Proportions

Section 17.1: Hypotheses

We can reproduce the calculation in Figure 17.1 (page 451).

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

```
## [1] 0.02
```
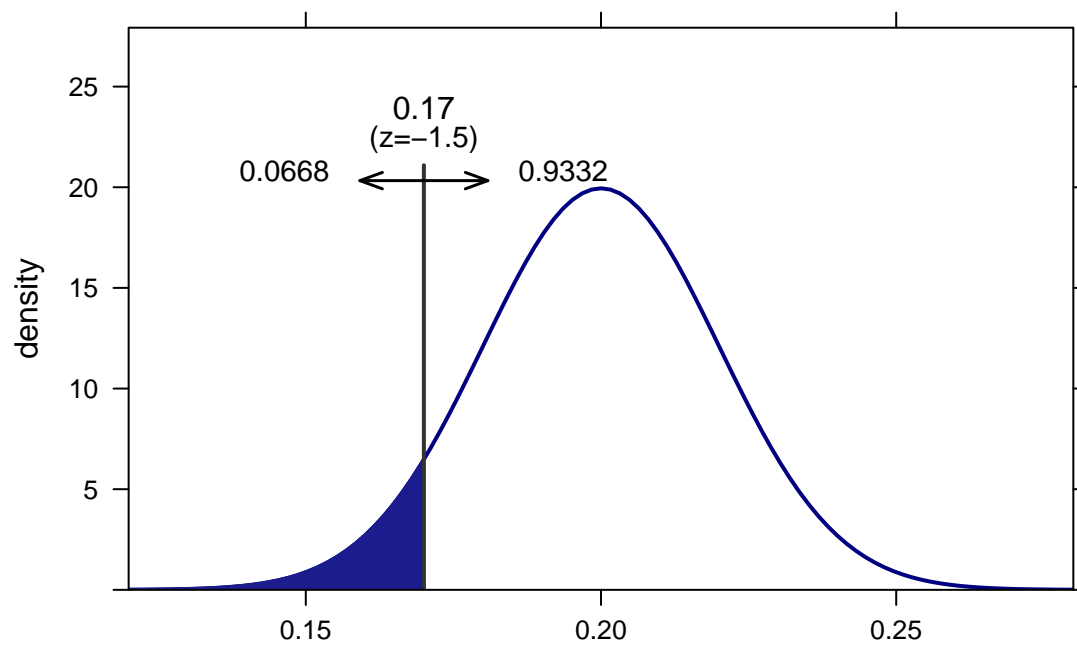
```
xpnorm(0.17, mean=0.20, sd=sdp)
```

```
##
## If X ~ N(0.2, 0.02), then
##
##   P(X <= 0.17) = P(Z <= -1.5) = 0.0668072
##   P(X >  0.17) = P(Z >  -1.5) = 0.9331928
```

```
## [1] 0.0668072
```

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

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

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

```
## [1] 0.0668072
```

Section 17.3: Reasoning of hypothesis testing

The "For Example (page 455)" lays out how to find a p-value for the one proportion z-test.
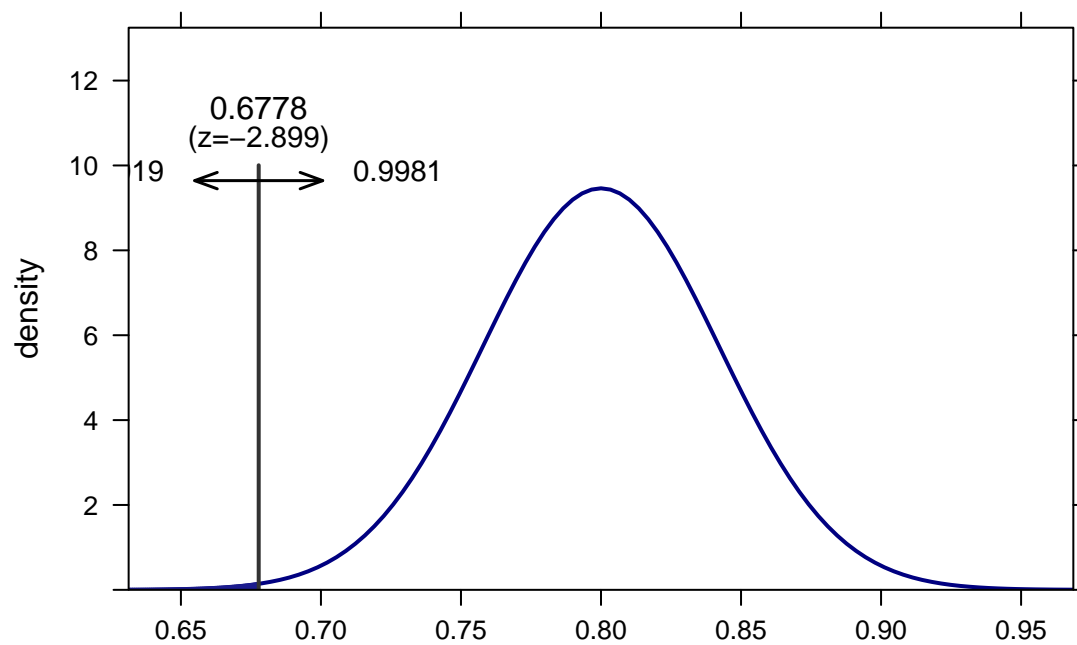
```
y <- 61; n <- 90; phat <- y/n; phat
```

```
## [1] 0.6777778
```

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

```
## [1] 0.0421637
```

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

```
##
## If X ~ N(0.8, 0.0421637), then
##
##   P(X <= 0.6777778) = P(Z <= -2.898755) = 0.00187324
##   P(X >  0.6777778) = P(Z >  -2.898755) = 0.9981268
```

```
## [1] 0.00187324
```

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

```
## [1] 0.003746481
```

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 = 61, number of trials = 90, p-value =
## 0.005511
## alternative hypothesis: true probability of success is not equal to 0.8
## 95 percent confidence interval:
##  0.5709996 0.7724797
## sample estimates:
## probability of success
##              0.6777778
```