SDM4 in R: Randomness and Probability (Chapter 15)

Nicholas Horton (nhorton@amherst.edu) and Sarah McDonald

June 5, 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 15: Randomness and Probability

Section 15.1: Center (the Expected Value)

We can replicate the calculation on page 390:

```
library(mosaic)
library(readr)
options(digits = 3)
x <- c(10000, 5000, 0)
prob <- c(1/1000, 2/1000, 997/1000)
sum(prob)   # sums to 1
```

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

```
expect <- sum(x*prob)
expect  # expected value
```

```
## [1] 20
```

Section 15.2: Spread (The Standard Deviation)

We can continue with the example from page 392:

```
xminmu <- x - expect
xminmu
```

```
## [1] 9980 4980  -20
```

```
myvar <- sum(xminmu^2*prob)
myvar
```

```
## [1] 149600
```

```
sd <- sqrt(myvar)
sd
```

```
## [1] 387
```

Section 15.3: Shifting and Combining Random Variables

Let's replicate the values from the example on page 394:

```
ex <- 5.83
varx <- 8.62^2
ed <- ex+5
ed
```

```
## [1] 10.8
```

```
vard <- varx
vard
```

```
## [1] 74.3
```

```
sqrt(vard)
```

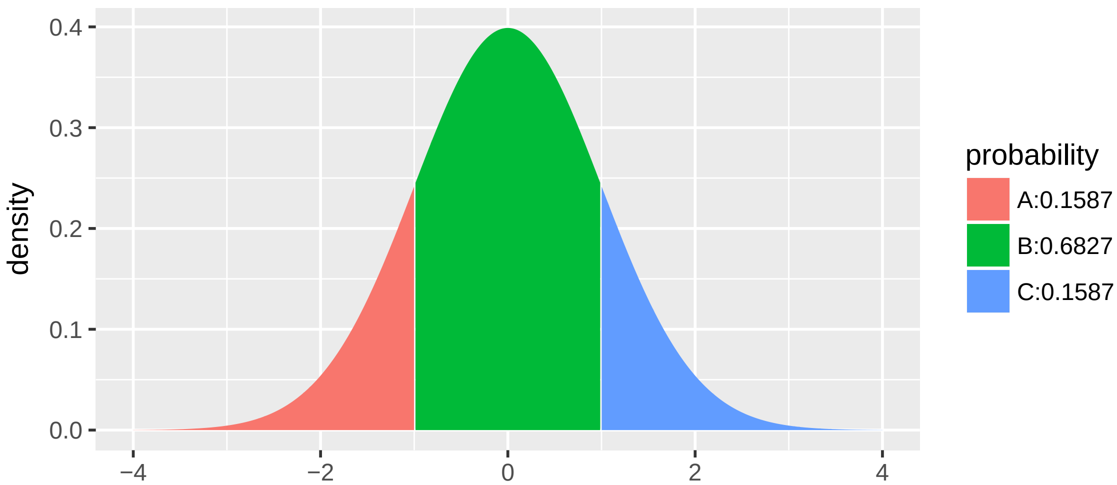
```
## [1] 8.62
```

Section 15.5: Continuous random variables

Let's replicate Figure 15.1 (page 400):

```
xpnorm(c(-1, 1), mean=0, sd = 1)
```

```
##
## If X ~ N(0, 1), then
##  P(X <= -1) = P(Z <= -1) = 0.1587    P(X <=  1) = P(Z <=  1) = 0.8413
##  P(X >  -1) = P(Z >  -1) = 0.8413    P(X >   1) = P(Z >   1) = 0.1587
##
```



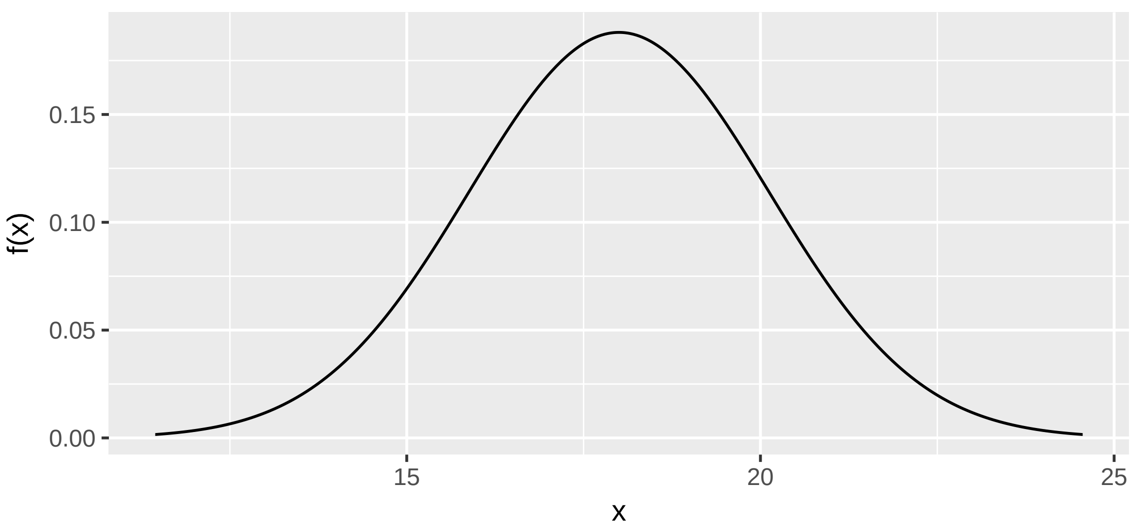

```
## [1] 0.159 0.841
```

and the Think/Show/Tell/Think on pages 402 and 403:

```
sdval <- sqrt(4.50)
sdval
```

```
## [1] 2.12
```

```
gf_dist("norm", params = list(18, sdval), xlab = "x", ylab = "f(x)")
```




```
xpnorm(20, mean = 18, sd = sdval)  # note how exact value is different from the table!
```

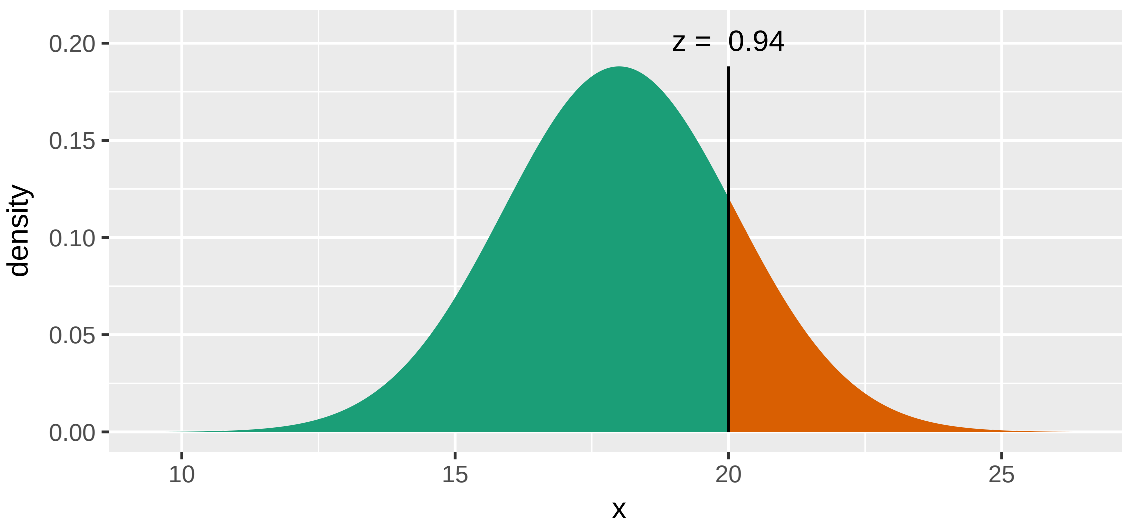
```
##

## If X ~ N(18, 2.121), then

##  P(X <= 20) = P(Z <= 0.9428) = 0.8271
```

```
##  P(X >  20) = P(Z >  0.9428) = 0.1729

##
```




```
## [1] 0.827
```

```
zval <- (20-18)/sdval
zval
```

```
## [1] 0.943
```

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

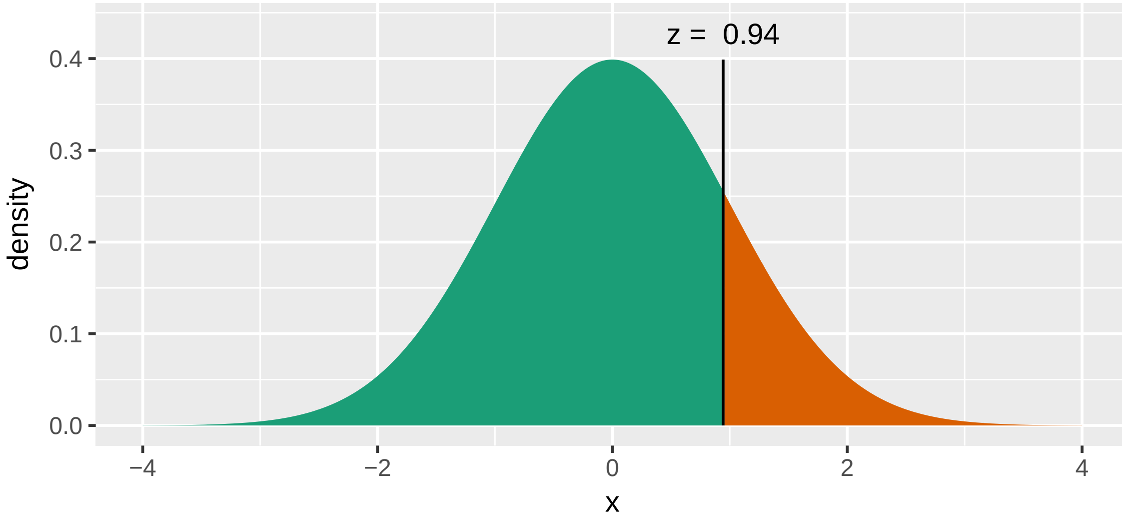
```
##

## If X ~ N(0, 1), then

##  P(X <= 0.9428) = P(Z <= 0.9428) = 0.8271

##  P(X >  0.9428) = P(Z >  0.9428) = 0.1729

##
```




```
## [1] 0.827
```