Distribution of Random Variables Homework3

Sharon Morris

September 16, 2016

3.2 Area under the curve, Part II

What percent of a standard normal distribution $N(\mu = 0, \sigma = 1)$ is found in each region? Be sure to draw a graph.

a. Z > -1.13

Use pnorm function to determine the area under the left tail of the normal distribution by subtracting 1 from pnorm to determine the area under the right tail

```
library(IS606)
```

```
##
## Welcome to CUNY IS606 Statistics and Probability for Data Analytics
## This package is designed to support this course. The text book used
## is OpenIntro Statistics, 3rd Edition. You can read this by typing
## vignette('os3') or visit www.OpenIntro.org.
##
## The getLabs() function will return a list of the labs available.
##
## The demo(package='IS606') will list the demos that are available.
```

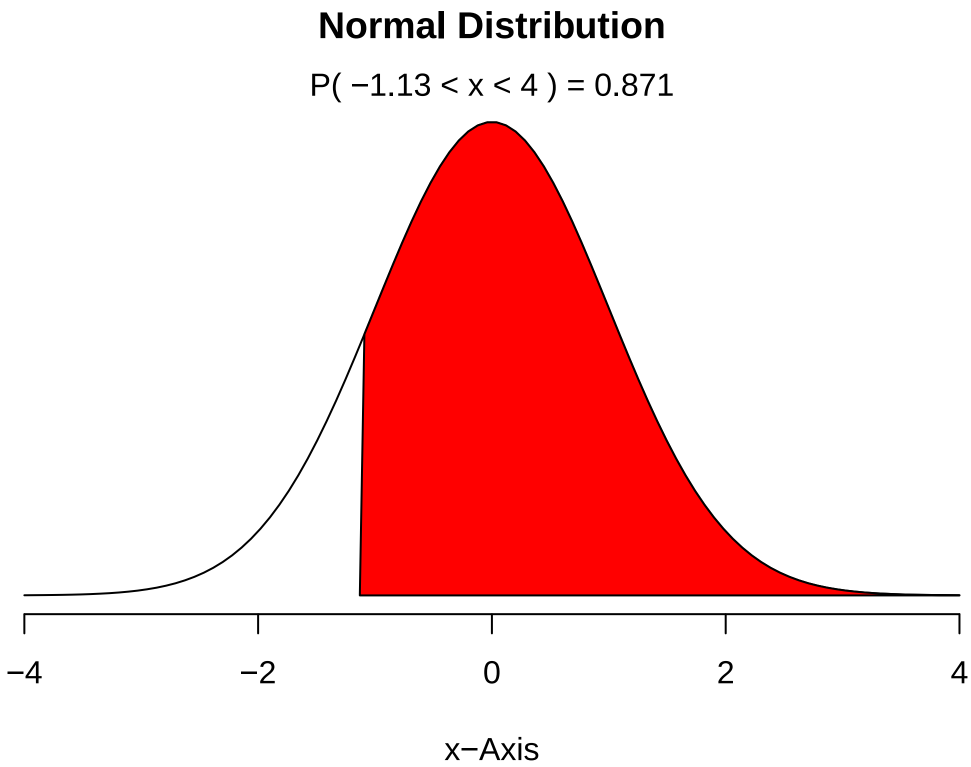
```
##
## Attaching package: 'IS606'
```

```
## The following object is masked from 'package:utils':
##
##     demo
```

```
p <- 1 - pnorm(-1.13)
p
```

```
## [1] 0.8707619
```

```
normalPlot(bounds = c(-1.13,4))
```

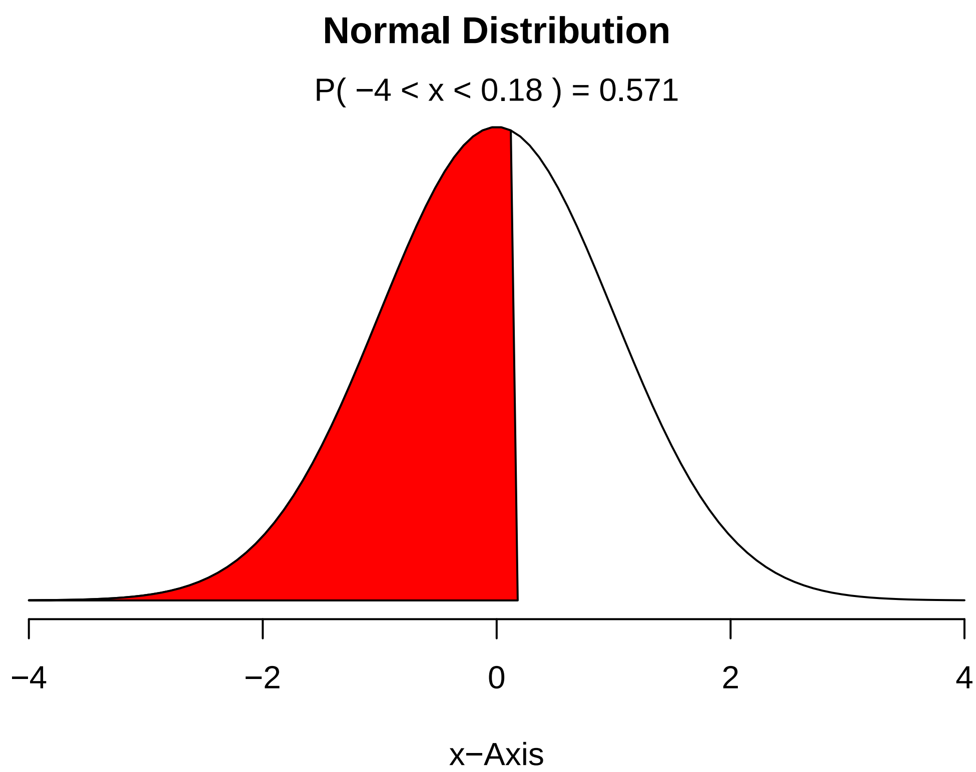


b. $Z < 0.18$

```
library(IS606)
p <- pnorm(0.18)
p
```

```
## [1] 0.5714237
```

```
normalPlot(bounds = c(-4,0.18))
```

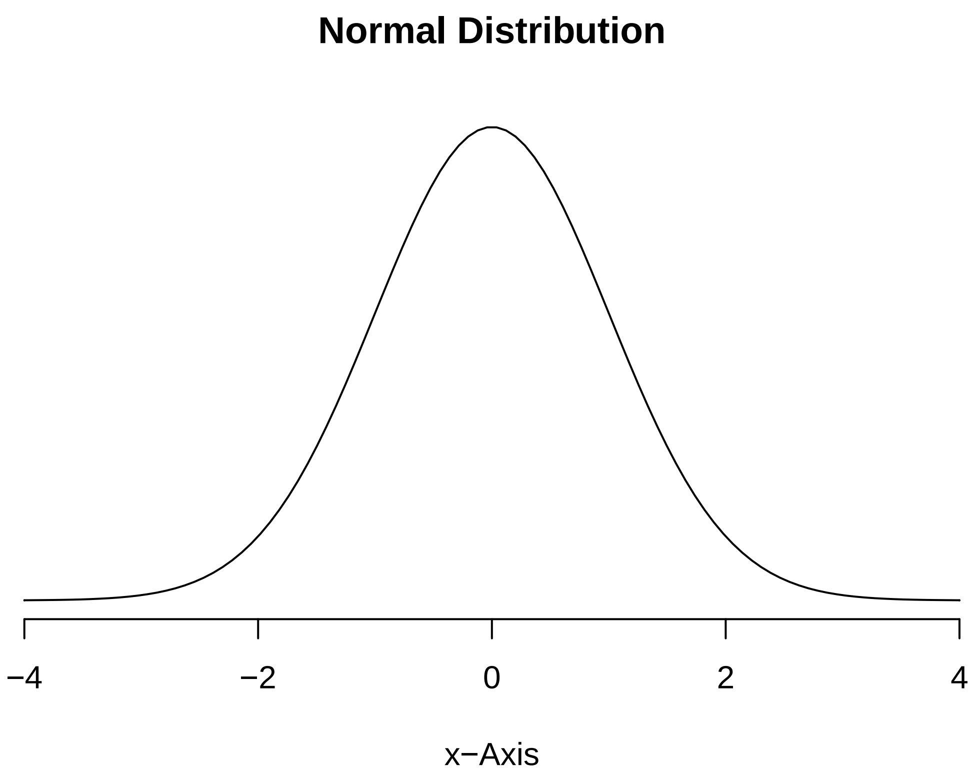


c. $Z > 8$

```
library(IS606)
p <- 1 - pnorm(8)
p
```

```
## [1] 6.661338e-16
```

```
normalPlot(bounds = c(-4,4), tails=TRUE)
```

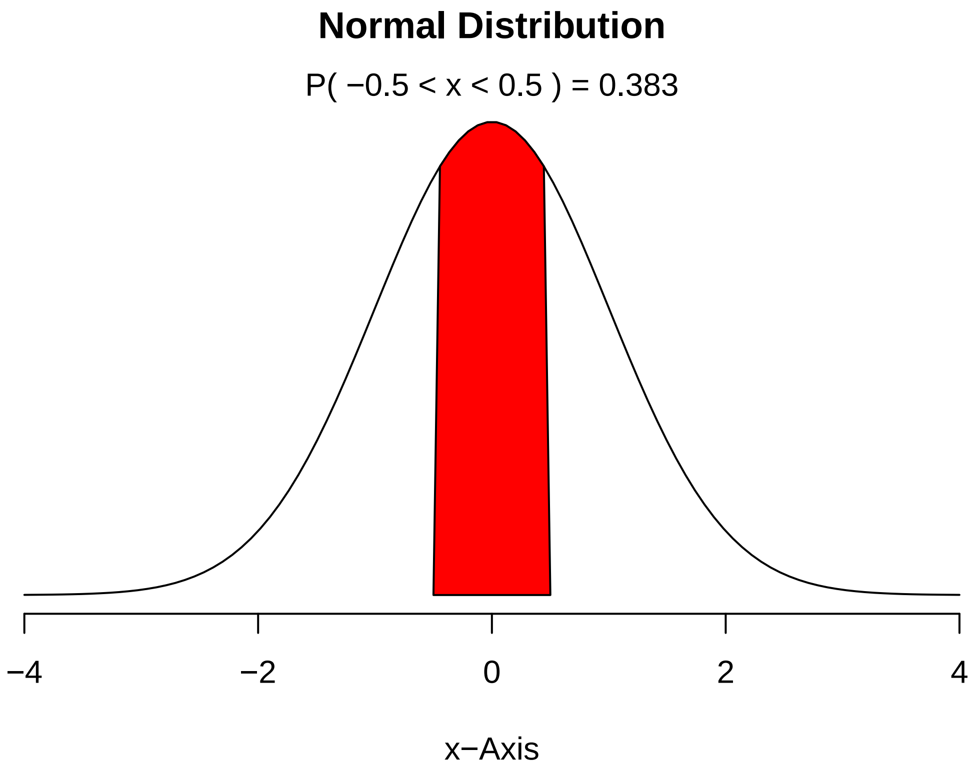


d. $|Z| < .05$

```r
pnorm(0.5) - pnorm(-0.5)
```

```
## [1] 0.3829249
```

```r
normalPlot(bounds=c(-0.5, 0.5))
```



3.4 Triathlon times, Part I.

a. Write down the short-hand for these two normal distributions. *Men, Ages 30 - 34* $N(\mu = 4313, \sigma = 583)$ *Women, Ages 25 - 29* $N(\mu = 5261, \sigma = 807)$

b. What are the Z-scores for Leo's and Mary's finishing times? What do these Z-scores tell you?

The Z-scores below indicate, Leo's finishing time is 1.09 standard diviations above the mean and Mary's finishing times is 0.31 standard diviations above the mean. Thus Mary did better in her group than Leo.

Leo's Z-score

```
zs_leo <- (4948-4313)/583
zs_leo
```

```
## [1] 1.089194
```

Mary's Z-score

```
zs_mary <- (5513-5261)/807
zs_mary
```

```
## [1] 0.3122677
```

c. Did Leo or Mary rank better in thier respective groups? Explain your reasoning.

Mary did better than Leo. Since Mary's mean score is closer to the mean than Leo.

d. What percent of the triathletes did Leo finish faster than in his group?

```
pnorm(1.09,lower.tail=FALSE)
```

```
## [1] 0.1378566
```

e. What percent of the triathletes did Mary finish faster than in her group?

```r
pnorm(0.31,lower.tail=FALSE)
```

```
## [1] 0.3782805
```

If the distributions of finising times are not nearly normal, would your answers to parts (b) - (e) change?d Explain your reasoning.

No, my answers would not change. My answers would not change because Z-scores are used and Z-scores are relative to the other performers. Thus, the percentile where each performer falls under a normal distribution does not change if the distribution is not nearly normal.

3.18 Heights of female college students

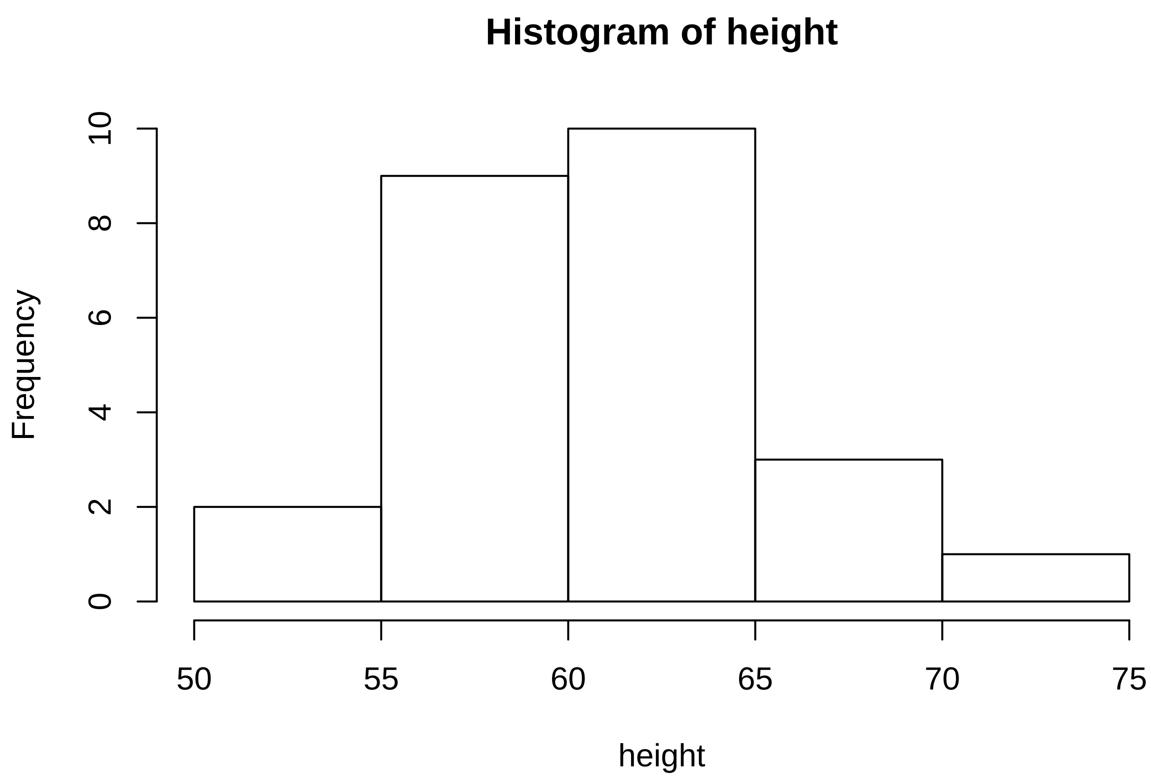
a. The mean height is 61.52 inches with a standard deviation of 4.58 inches. Use this information to determine if the heights approximately follow the 68-95-99% Rule.

The information below implies the heights do not approximatley follow the 68-95-99% rule. The heights appear to be skewed to the left.

```r
height <- c(54, 55, 56, 56, 57, 58, 58, 59, 60, 60, 60, 61, 61, 62, 62, 63, 63, 63, 64, 65, 65, 67, 67,
summary(height)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   54.00   58.00   61.00   61.52   64.00   73.00
```

```r
hist(height)
```




```r
pnorm(61.52+4.58,mean=61.52,sd=4.58)
```

```
## [1] 0.8413447
```

```r
pnorm(61.52+2*4.58,mean=61.52,sd=4.58)
```

```
## [1] 0.9772499
```

b. Do these data appear to follow a normal distriution? Explain your reasoning using the graphs provided below.

No the data does not appear to follow a normal distribution. Becuase the data are skewed to the left it does not follow the 68-95-99.7% rule.

3.22 Defective rate A machine that produces a special type of transistor (a component of computers) has a 2% defective rate. The production is considered a random process where each transistor is independed of the others.

a. What is the probability that the 10th transistor produced is the first with a defect?

```
pgeom(10-1,0.02)
```

```
## [1] 0.1829272
```

b.What is the probability that the machine produces no defective transistors in a batch of 100?

```
pgeom(100,0.02)
```

```
## [1] 0.8700328
```

c. On average, how many transistors workd you expect to be produced before the first with a defect? What is the standard deviation?

```
# Mean
1/0.02
```

```
## [1] 50
```

```
# Standard deviation
sqrt((1-0.02)/0.02^2)
```

```
## [1] 49.49747
```

d. Another machine that also produces transistors has a 5% defective rate where each transistor is produced independent of the others. On average how many transistors would you expect to be produced with this machine before the first with a defect? What is the standard deviation?

```
#Mean
1/0.05
```

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

```
# Standard deviation
sqrt((1-0.05)/0.05^2)
```

```
## [1] 19.49359
```

e. Based on your answers to parts (c) and (d), how does increasing the probability of an event affect the man and standard deviation of the wait time until success?

Increasing the probability decreases the wait time for success.

3.38 Male children.

While it is often assumed that the probabilities of having a boy or a girl are the same, the actual probability of having a boy is slightly higher at 0.51. Suppose a couple plans to have 3 kids.

a. Use the binomial model to calculate the probability that two of them will be boys.

```
dbinom(2,3,0.51)
```

```
## [1] 0.382347
```

b. Write out all possible orderings of 3 children, 2 of whom are boys. Use these scenarios to calculate the same probability from part (a) but using the addition rule for disjoint outcomes. Confirm that your answers from parts (a) and (b) match.

```
# BBG GBB BGB

(0.51^2)*0.49*3
```

```
## [1] 0.382347
```

c. If we wanted to calculate the probability that a couple who plans to have 8 kids will have 3 boys, briefly describe why the approach from part (b) would be more tedious than the approach from part (a).

The approach for part (b) is more tedious than the approach for part (a) because the combinatiions in part (b) have to be done manually. If there were more combinations this would take a great deal of time. In part (a), using the function made it much simpilier.

3.42 Serving in volleyball A not-so-skilled volleyball player has a 15% chance of making the serve, which involves hitting the ball so it passes over the net on a trajectory such that it will land in the opposing team's court. Suppose that her serves are independent of each other.

a. What is the probability that on the 10th try she will make her 3rd successful serve?

```
choose(9,2)*0.15^3*0.85^7
```

```
## [1] 0.03895012
```

b. Suppose she has made two successful serves in nine attempts. What is the probability that her 10th serve will be successful?

All probability trials are independed. Thus, the probablity that her 10th serve will be successful is the same as her 1st to 9th serve – 15%

c. Even though parts (a) and (b) discuss the same scenario, the probabilities you calculated should be different. Can you explain the reason for the discrepancy?

I don't beleive there should be a difference since, they both refer to independent outcomes.