

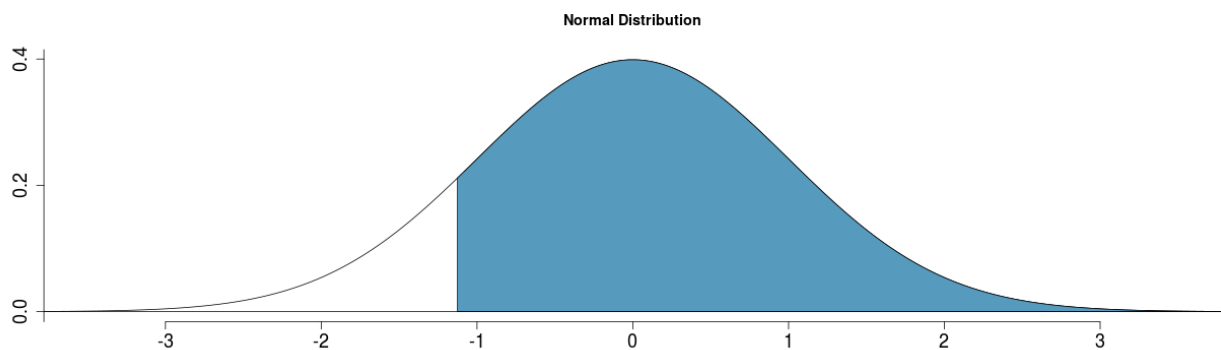
Homework Week 4

Steven Simonsen

2024-02-05

4.2 Area under the curve, Part II. What percent of a standard normal distribution $N(\text{mean} = 0; \text{sd} = 1)$ is found in each region? Be sure to draw a graph.

(a) $Z > -1.13$



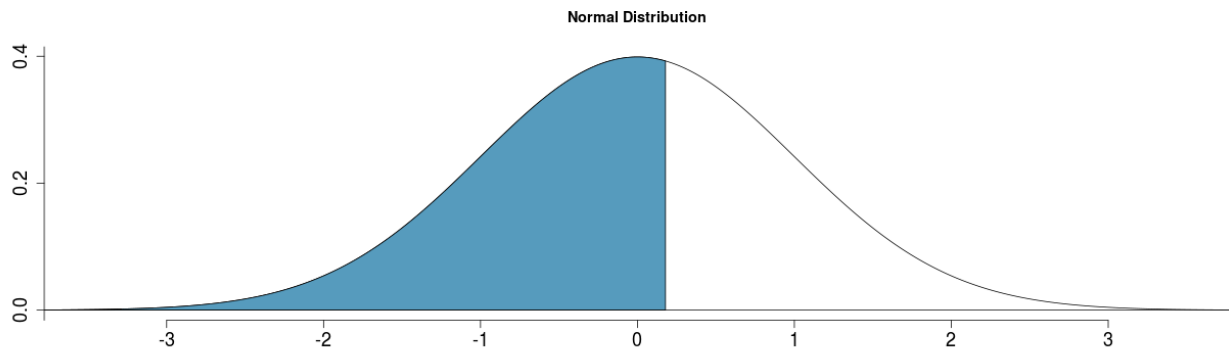
$$P(X > -1.13) = 0.871$$

```
{#id .class width=50% height=50%}
```

```
library(scales)
a <- 1 - pnorm(-1.13,0,1)
percent(a, accuracy=0.01) #Answer
```

```
## [1] "87.08%"
```

(b) $Z < 0.18$



$$P(X < 0.18) = 0.571$$

```
{#id .class width=50% height=50%}
```

```
b <- pnorm(0.18,0,1)
percent(b,accuracy=0.01) #Answer
```

```
## [1] "57.14%"
```

(c) $Z > 8$

Normal Distribution
 $X \sim N(\mu, \sigma)$

$\mu =$

$\sigma =$

$x =$

$P(X > x) =$

A plot of a Normal Distribution with mean 0 and standard deviation 1. The x-axis ranges from -4 to 4, and the y-axis ranges from 0.0 to 0.4. The area under the curve to the right of 8 is shaded in red. The title of the plot is "Normal Distribution".

$\mu = E(X) = 0 \quad \sigma = SD(X) = 1 \quad \sigma^2 = Var(X) = 1$

Help

©2021 Matt Bognar
 Department of Statistics and Actuarial Science
 University of Iowa

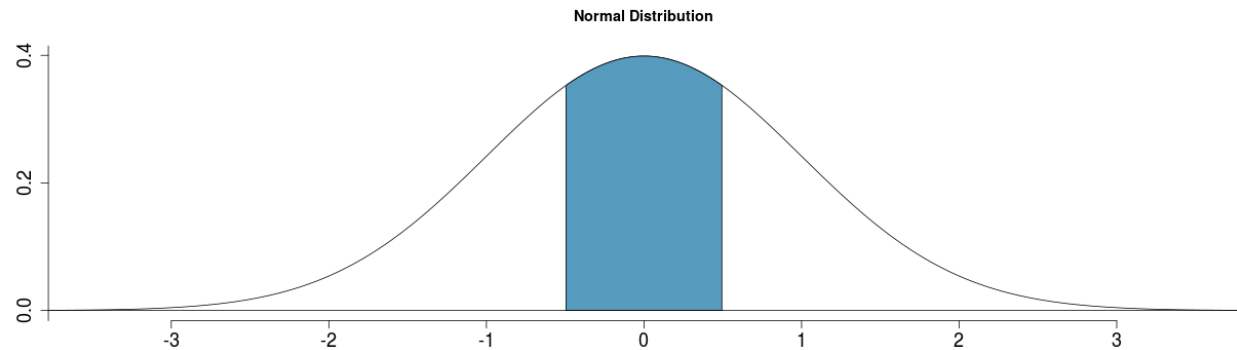
```
{#id .class width=50% height=50%}
```

```
c <- 1 - pnorm(8,0,1)
percent(c, accuracy=0.01) #Answer
```

```
## [1] "0.00%"
```

I used a different website as the one covered in class only went up to $Z=4$. <https://homepage.divms.uiowa.edu/~mbognar/applets/normal.html>

(d) $\text{absolute}(Z) < 0.5$



$$P(-0.5 < X < 0.5) = 0.383$$

```
{#id .class width=50% height=50%}
```

```
x <- pnorm(0.5,0,1) #percentile for z<0.5
y <- 1-x
z <- pnorm(-0.5,0,1) # percentile for abs
q <- 1-y-z #1 minus each of the tail values gets the value in the middle, or
#-0.5<Z<0.5
percent(q, accuracy=0.01) #Answer
```

```
## [1] "38.29%"
```

4.4 Triathlon times, Part I: In triathlons, it is common for racers to be placed into age and gender groups. Friends Leo and Mary both completed the Hermosa Beach Triathlon, where Leo competed in the Men, Ages 30 - 34 group while Mary competed in the Women, Ages 25 - 29 group. Leo completed the race in 1:22:28 (4948 seconds), while Mary completed the race in 1:31:53 (5513 seconds). Obviously Leo finished faster, but they are curious about how they did within their respective groups. Can you help them? Here is some information on the performance of their groups:

- The finishing times of the Men, Ages 30 - 34 group has a mean of 4313 seconds with a standard deviation of 583 seconds.
- The finishing times of the Women, Ages 25 - 29 group has a mean of 5261 seconds with a standard deviation of 807 seconds.
- The distributions of finishing times for both groups are approximately Normal.

(a) Write down the short-hand for these two normal distributions.'

Men: $N(\mu=4313s, sd=583s)$ Women: $N(\mu=5261s, sd=807s)$

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

```
(4948-4313)/583 #Leo Z-score
```

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

```
(5513-5261)/807 #Mary Z-score
```

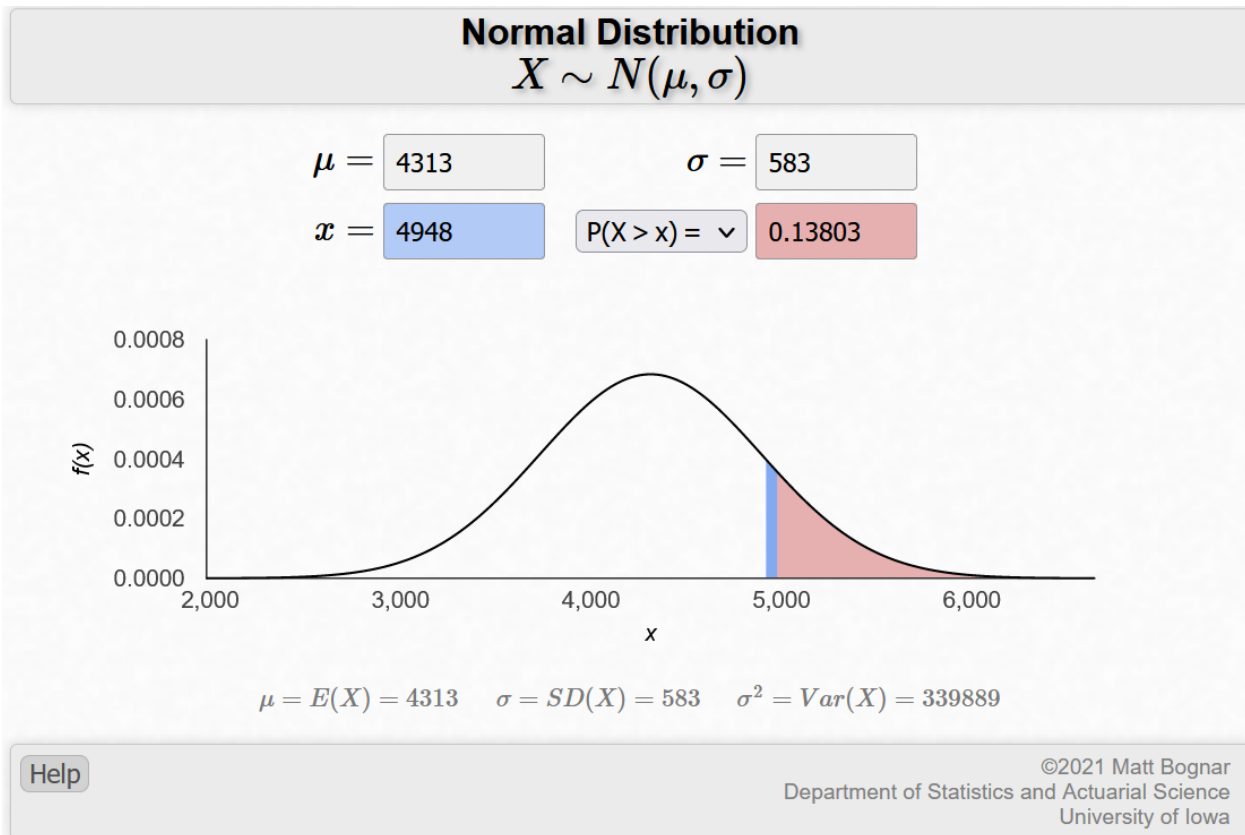
```
## [1] 0.3122677
```

```
##These scores tell me that Leo finished about 1.08 standard deviations above  
#the mean, and Mary finished about 0.31 standard deviations above the mean.
```

(c) Did Leo or Mary rank better in their respective groups? Explain your reasoning.

Mary finished better in her respective group. As scores increase in standard deviation above the mean, this means the runner finished slower than the average runner in the respective group. Because Mary had a lower Z-Score, she finished closer to the mean than Leo did, and therefore had a better rank respective to her group.

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

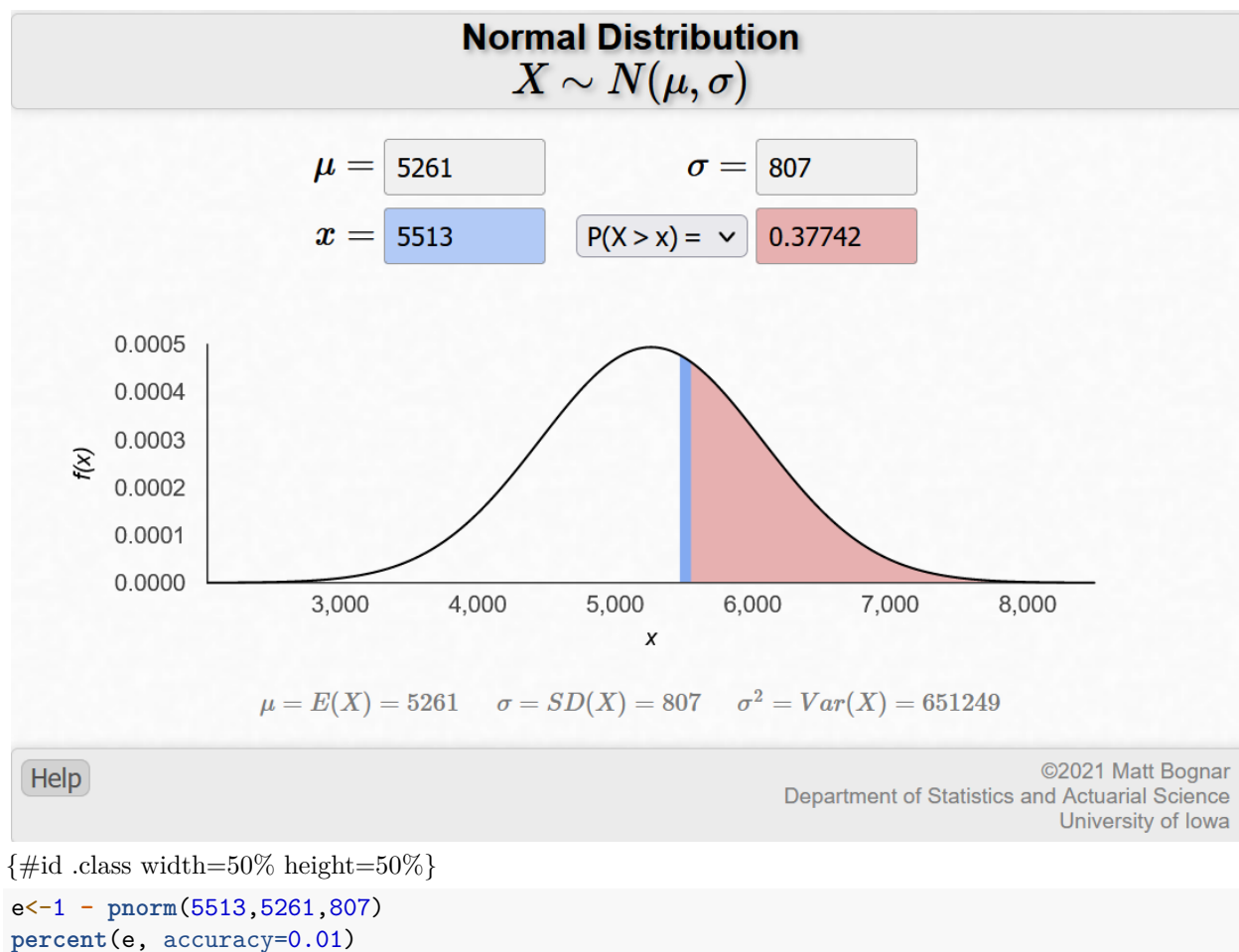


```
{#id .class width=50% height=50%}
```

```
d<-1-pnorm(4948,4313,583)
percent(d, accuracy=0.01)
```

```
## [1] "13.80%"
```

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

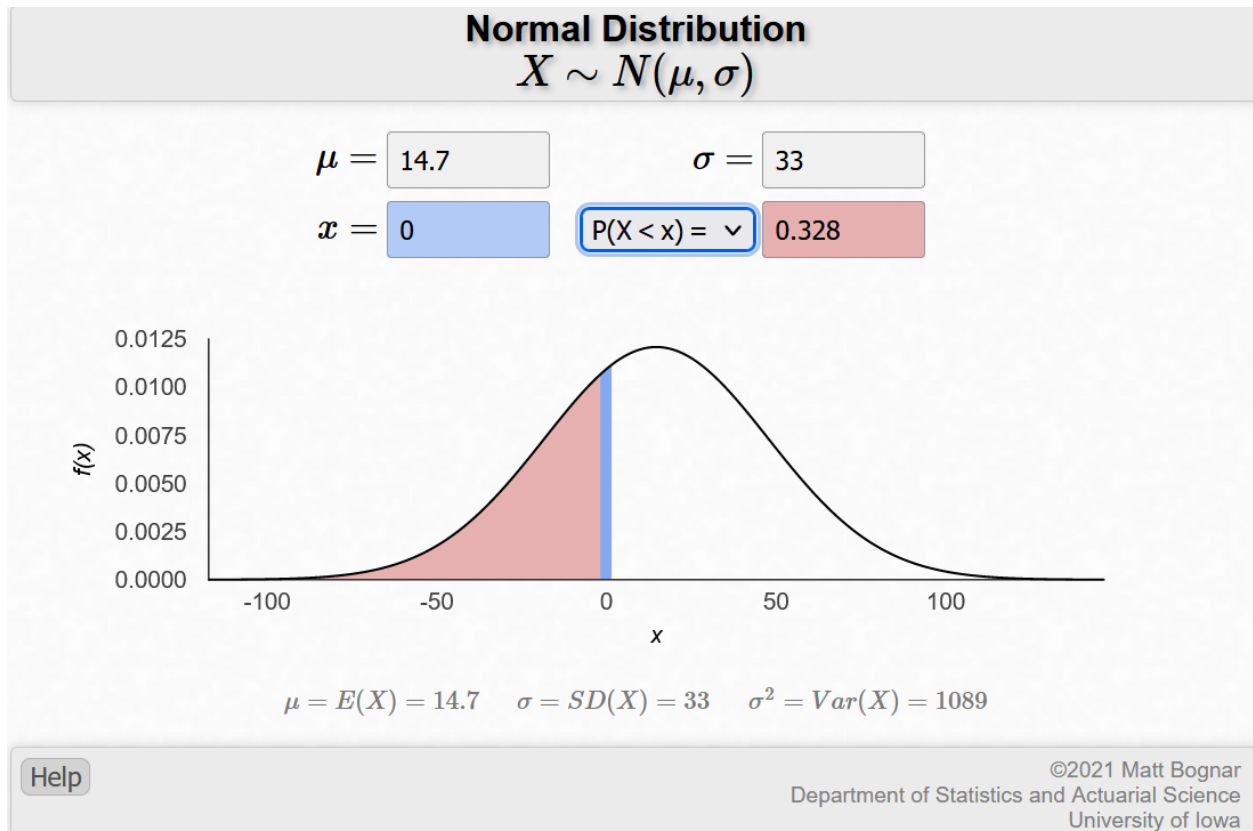


(f) If the distributions of finishing times are not nearly normal, would your answers to parts (b) - (e) change? Explain your reasoning.

Yes, if the distributions of finishing times are not nearly normal, then the mean and standard deviation for the groups would have been impacted positively or negatively. Because of this, Mary and Leo may have finished closer to, or further away from the mean. In other words, there Z-scores may have indicated more or less standard deviations away from the mean. If this was the case, either of them could have finished better or worse relative to the percent of other triathletes, and therefore would have also ranked better or worse in their respective groups.

4.8 CAPM: The Capital Asset Pricing Model (CAPM) is a financial model that assumes returns on a portfolio are normally distributed. Suppose a portfolio has an average annual return of 14.7% (i.e. an average gain of 14.7%) with a standard deviation of 33%. A return of 0% means the value of the portfolio doesn't change, a negative return means that the portfolio loses money, and a positive return means that the portfolio gains money.

(a) What percent of years does this portfolio lose money, i.e. have a return less than 0%?



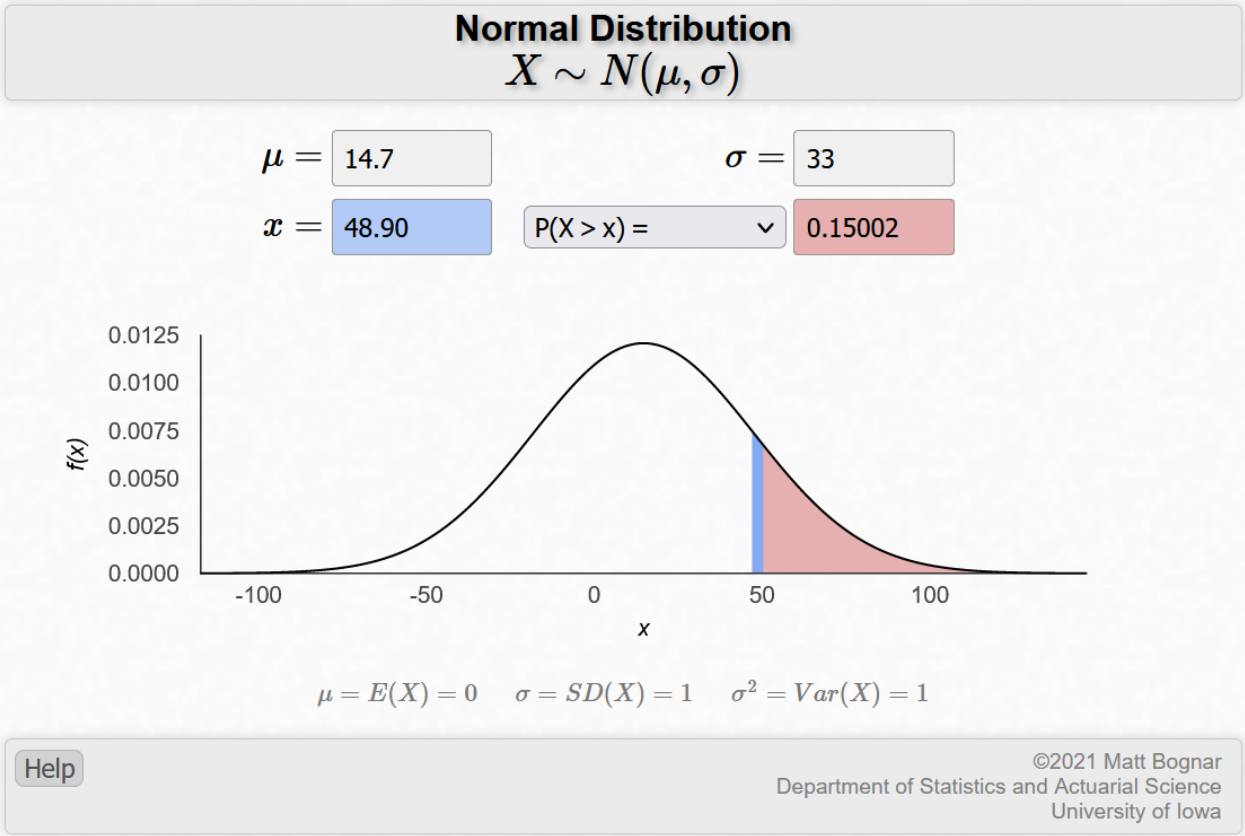
```
{#id .class width=50% height=50%}
percent(pnorm(0,14.7,33),accuracy=0.01)

## [1] "32.80%"
```

(b) What is the cutoff for the highest 15% of annual returns with this portfolio?

```
f<-qnorm(.85)
(f*33)+14.7

## [1] 48.9023
##The cutoff is 48.90 percent growth
```



{#id .class width=50% height=50%}