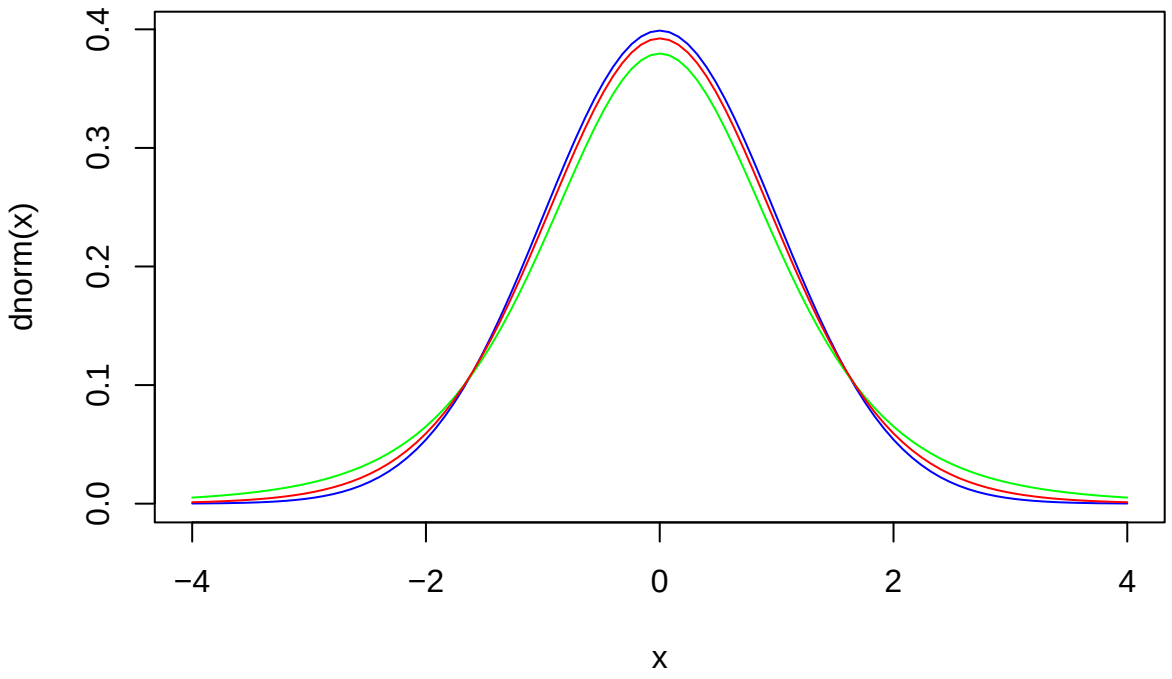
Week 9: Confidence Intervals

Example 1

William S. Goset of the Guinness brewery in Dublin developed a more accurate way to estimate the population; it involves the Student's t-distribution. Below we compare the normal t-distributions of degrees of freedom of 5 and 15.

```
curve(dnorm, -4, 4, col = "blue")
curve(dt(x, 5), -4, 4, add = TRUE, col = "green")
curve(dt(x, 15), -4, 4, add = TRUE, col = "red")
```



Consider the example on page 457 of the OpenStax textbook: a researcher measured sensory rates for 15 subjects. We manually enter the data below, and find the sample mean and standard deviation.

```
x <- c(8.6, 9.4, 7.9, 6.8, 8.3, 7.3, 9.2, 9.6, 8.7, 11.4, 10.3, 5.4, 8.1, 5.5, 6.9)
mean(x)
```

```
## [1] 8.226667
```

```
sd(x)
```

```
## [1] 1.672238
```

To find the 95% confidence interval, we need to find $t_{\alpha/2}$. The degrees of freedom is 14.

```
qt(0.05/2, 14)
```

```
## [1] -2.144787
```

Note that qt give the area to the left of the t-score, but in many introductory statistics textbook $t_{\alpha/2}$ represents the cut on the right tail, and in such a case $t_{\alpha/2} = 2.145$. A better notation for the right cut is $t_{1-\alpha/2}$.

```
qt(1-0.05/2, 14)
```

```
## [1] 2.144787
```

The lower and upper bounds of the confidence interval are the following.

```
8.2267 - 2.145*1.6722/sqrt(15)
```

```
## [1] 7.300574
```

```
8.2267 + 2.145*1.6722/sqrt(15)
```

```
## [1] 9.152826
```

We estimate with 95% confidence that the population mean sensory rate is between 7.30 and 9.15.

It's possible to find the confidence interval directly by using a pre-defined function.

```
t.test(x)$conf.int
```

```
## [1] 7.300612 9.152721
## attr(,"conf.level")
## [1] 0.95
```

Example 2

On page 459 of the OpenStax textbook, the Human Toxome Project tested 20 newborn infants. The number of targeted chemicals found in each infant's cord blood is shown below.

```
x <- c(79, 145, 147, 160, 116, 100, 159, 151, 156, 126,
       137, 83, 156, 94, 121, 144, 123, 114, 139, 99)
mean(x)
```

```
## [1] 127.45
```

```
sd(x)
```

```
## [1] 25.9645
```

The critical value for a 90% confidence interval is found in the following. The degrees of freedom is 19.

```
qt(0.1/2, 19)
```

```
## [1] -1.729133
```

The lower and upper bounds of the confidence interval are below.

```
127.45 - 1.729*25.9645/sqrt(20)
```

```
## [1] 117.4117
```

```
127.45 + 1.729*25.9645/sqrt(20)
```

```
## [1] 137.4883
```

We estimate with 90% confidence that the mean number of all targeted industrial chemicals found in cord blood is between 117.4 and 137.5.

Again we can use a pre-defined function to obtain the same result.

```
t.test(x, conf.level = 0.9)$conf.int
```

```
## [1] 117.4109 137.4891
## attr(,"conf.level")
## [1] 0.9
```

Follow Up

On page 488 Homework problem 113, six brands of chocolate chip cookies were randomly selected at the supermarket.
The grams of fat per serving are: 8, 8, 10, 7, 9, 9. Construct a 90% confidence interval for the population mean grams of fat per serving of chocolate chip cookies sold in supermarkets.