

Random Samples

David Gerard

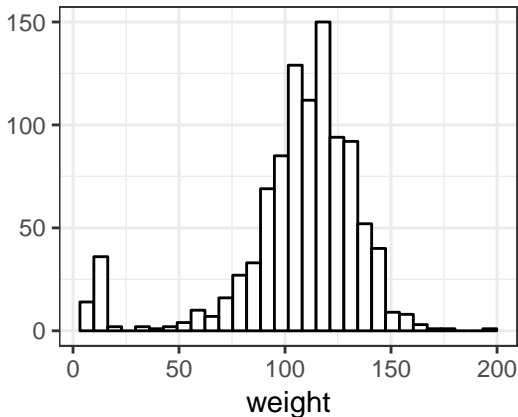
2018-12-07

Learning Objectives

- Gain intuitive understanding of CLT

Birthweight Data

```
library(ggplot2)
weight <- read.csv("../data/BIRTHWEIGHT.csv")$weight
qplot(x = weight, geom = "histogram", fill = I("white"),
      color = I("black"), bins = 30)
```



Get Sampling Distributions of Size 5

```
xbarvec5 <- replicate(n = 10000,  
                      expr = {  
                        mean(sample(weight, size = 5,  
                                   replace = TRUE))  
                      })
```

Get Sampling Distributions of Size 10

```
xbarvec10 <- replicate(n = 10000,  
  expr = {  
    mean(sample(weight, size = 10,  
      replace = TRUE))  
  })
```

Get Sampling Distributions of Size 50

```
xbarvec50 <- replicate(n = 10000,  
  expr = {  
    mean(sample(weight, size = 20,  
      replace = TRUE))  
  })
```

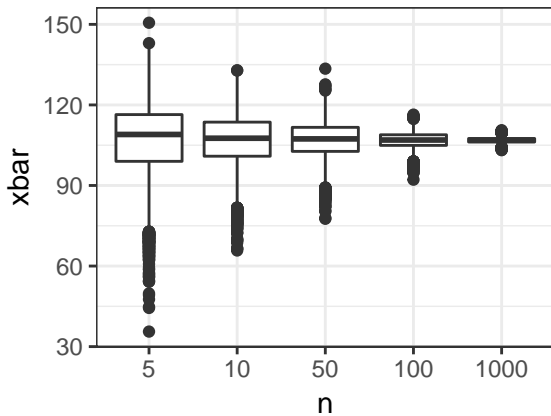
Get Sampling Distributions of Size 100

```
xbarvec100 <- replicate(n = 10000,  
                        expr = {  
                          mean(sample(weight, size = 100,  
                                       replace = TRUE))  
                        })
```

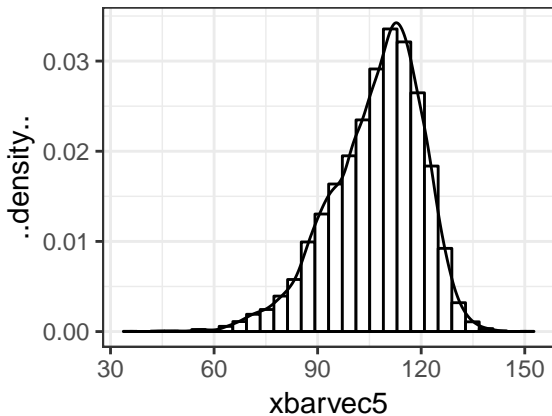
Get Sampling Distributions of Size 1000

```
xbarvec1000 <- replicate(n = 10000,  
  expr = {  
    mean(sample(weight, size = 1000,  
      replace = TRUE))  
  })
```

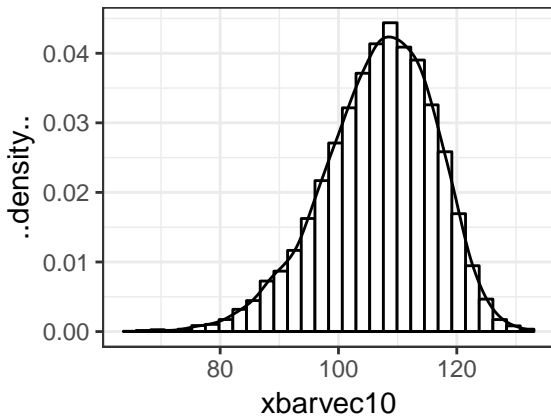

Variances Decrease over Time



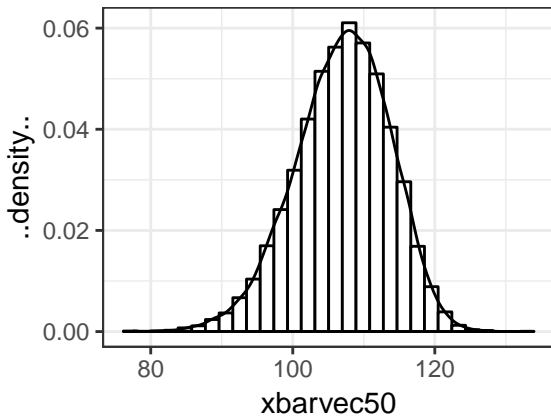
Not Very Normal ($n = 5$)



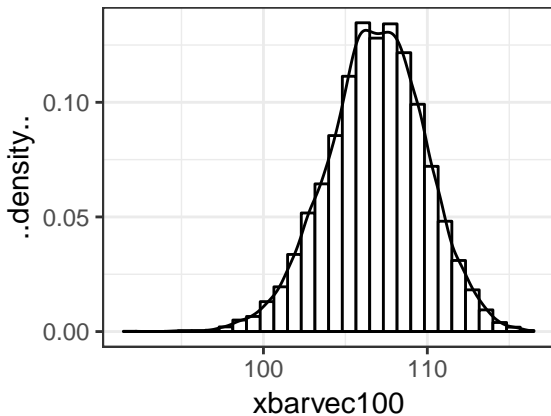
More Normal ($n = 10$)



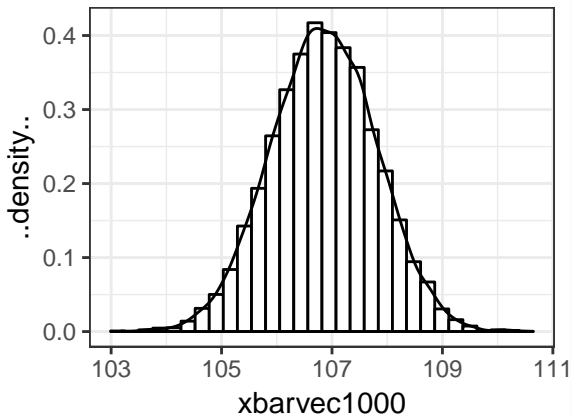
More Normal ($n = 50$)



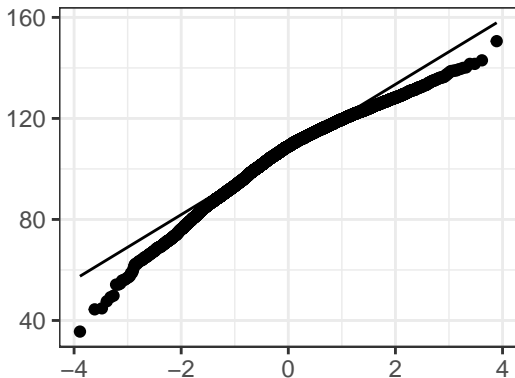
Very Normal (n = 100)



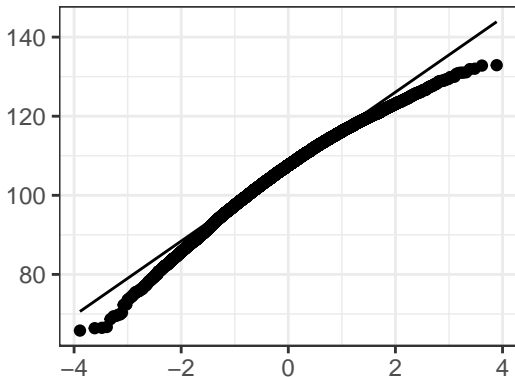
Oh-my-god-can-it-be-any-more Normal (n = 1000)



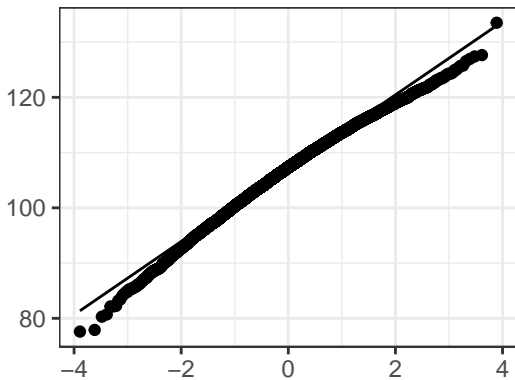
Not Very Normal ($n = 5$)



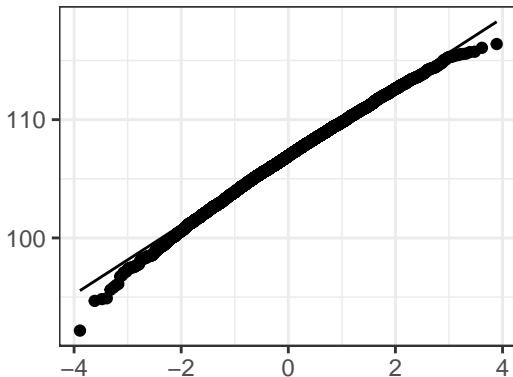
More Normal ($n = 10$)



More Normal ($n = 50$)



More Normal ($n = 100$)



Oh-my-god-can-it-be-any-more Normal (n = 1000)

