

Work 3

Henrique Araujo Lima

02 de maio de 2018

Descrição:

Reproduza o Ex. 4.23 (Exemplo das moedas). Para a geração dos dados, gere idades aleatórias entre 0 e 42 para as 500 moedas.

Descrição detalhada:

Existem 500 moedas com idades entre 0 e 42 anos com média $\mu = 13,468$ e desvio padrão $\sigma = 11,164$. O objetivo é simular a geração das 500 moedas aleatórias. E posteriormente retirar 25000 amostras de tamanhos 5, 10 e 25. Logo em seguida deve-se calcular as médias, medianas e desvio padrão para cada tamanho de amostras e comparar os resultados.

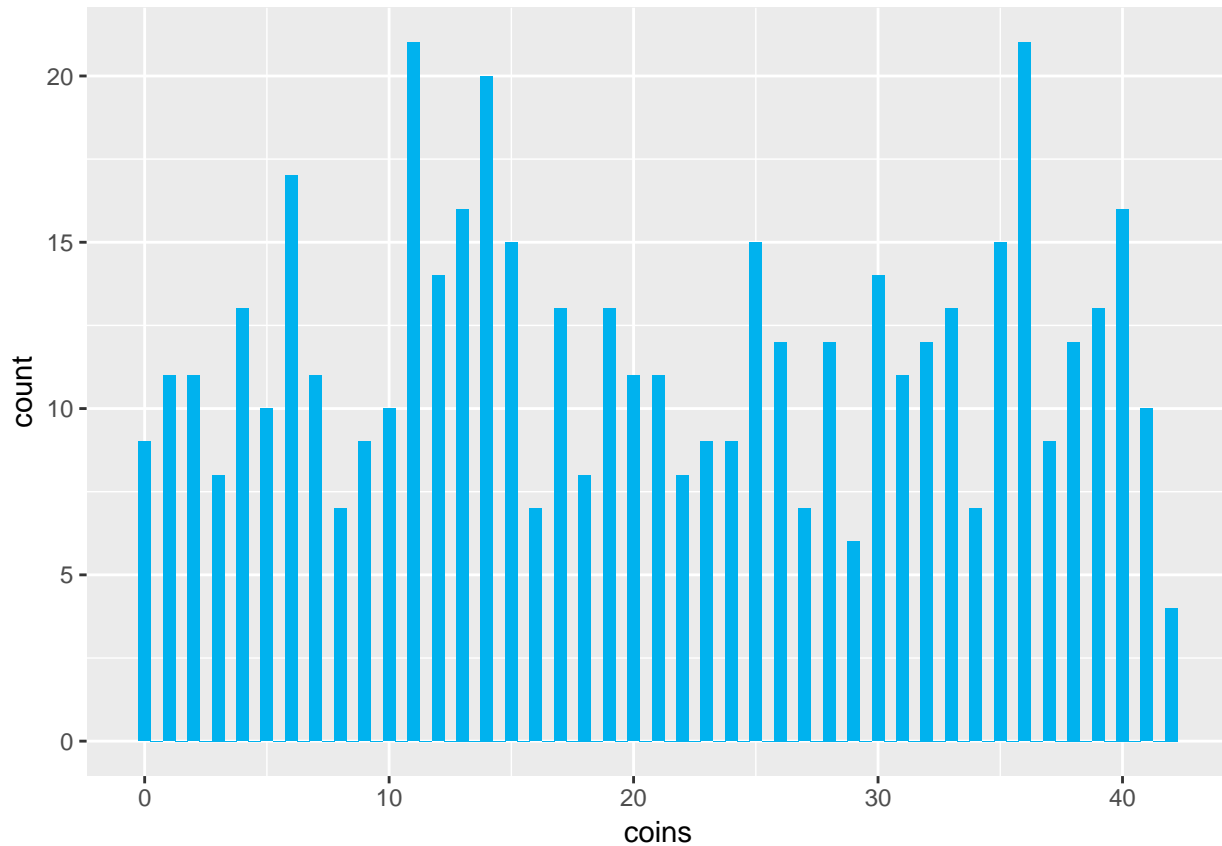
Desenvolvimento:

Geração aleatória das moedas

```
library(ggplot2)
ages = c(0:42) #cria um vetor com os valores (0,1,2,...,42)
coins = sample(ages, 500, replace=T)
coins
```

```
## [1] 36 7 40 33 38 1 12 40 14 17 23 3 35 3 13 25 35 41 21 26 38 13 2
## [24] 36 18 30 2 14 4 17 18 25 38 2 7 28 21 0 36 18 17 40 36 20 25 7
## [47] 36 38 11 8 26 33 14 15 2 23 9 30 12 15 7 30 26 13 30 1 20 28 12
## [70] 42 20 41 36 22 6 11 4 29 22 1 30 5 4 1 19 30 41 25 28 10 11 26
## [93] 10 6 42 19 6 30 38 16 15 33 30 24 7 11 17 31 39 15 1 21 33 36 25
## [116] 7 41 11 8 32 31 40 6 19 24 5 10 0 9 4 22 25 14 37 36 11 23 25
## [139] 21 4 10 37 3 15 12 38 20 34 36 40 29 17 12 17 11 13 10 2 1 28 6
## [162] 34 3 6 33 14 26 0 11 33 4 27 4 40 6 25 30 26 20 3 32 19 11 22
## [185] 11 32 9 11 30 20 12 39 21 16 6 37 14 27 19 24 34 32 14 11 37 31 11
## [208] 26 23 0 22 41 28 35 6 6 11 31 41 0 15 2 39 36 13 2 22 1 16 10
## [231] 14 18 14 37 40 17 40 21 17 35 0 8 39 35 31 42 16 36 32 9 14 5 6
## [254] 14 40 7 20 14 41 7 11 41 36 16 11 13 27 1 37 39 8 32 17 4 5 15
## [277] 38 38 11 33 5 19 12 35 39 17 12 0 36 23 35 36 29 18 27 3 3 38 5
## [300] 2 12 19 4 30 17 9 2 31 14 14 5 17 30 27 32 9 39 23 41 13 8 12
## [323] 19 13 21 33 28 36 15 18 24 24 39 15 35 1 37 14 28 1 27 15 39 15 9
## [346] 25 19 14 5 5 7 18 25 33 13 32 39 36 31 21 28 24 10 18 35 23 33 32
## [369] 20 37 13 30 26 27 5 19 23 36 21 31 32 16 11 20 39 26 4 38 12 19 9
## [392] 26 37 15 35 40 10 40 35 13 40 24 35 35 21 15 32 12 4 29 20 31 34 29
## [415] 14 23 28 13 36 4 0 2 13 17 3 21 24 33 11 6 7 28 2 11 25 33 42
## [438] 13 25 6 30 10 36 40 6 22 28 35 34 25 26 14 6 40 36 26 22 15 28 9
## [461] 38 34 15 13 4 32 19 25 35 33 16 36 0 34 6 8 12 40 6 40 8 39 20
## [484] 39 24 14 41 10 14 7 38 25 29 12 19 1 11 31 31 13
```

```
qplot(coins, geom="histogram", binwidth=0.5, bins=42, fill=I("deepskyblue2"))
```



Geração das amostras de tamanho $n = 5$, $10e20$ e calculo dos valores da média, mediana e desvio padrão.

Tamanho $n = 5$

```
r=25000
n=5
y=rep(0,n)
ybarMean5=rep(0,r)
ybarMedian5=rep(0,r)
ybarSd5=rep(0,r)
for(i in 1:r){
  y = sample(coins, n, replace=T)
  ybarMean5[i] = floor(mean(y))
  ybarMedian5[i] = floor(median(y))
  ybarSd5[i] = floor(sd(y))
}
```

Tamanho $n = 10$

```
r=25000
n=10
y=rep(0,n)
ybarMean10=rep(0,r)
ybarMedian10=rep(0,r)
ybarSd10=rep(0,r)
for(i in 1:r){
  y = sample(coins, n, replace=T)
  ybarMean10[i] = floor(mean(y))
  ybarMedian10[i] = floor(median(y))
  ybarSd10[i] = floor(sd(y))
}
```

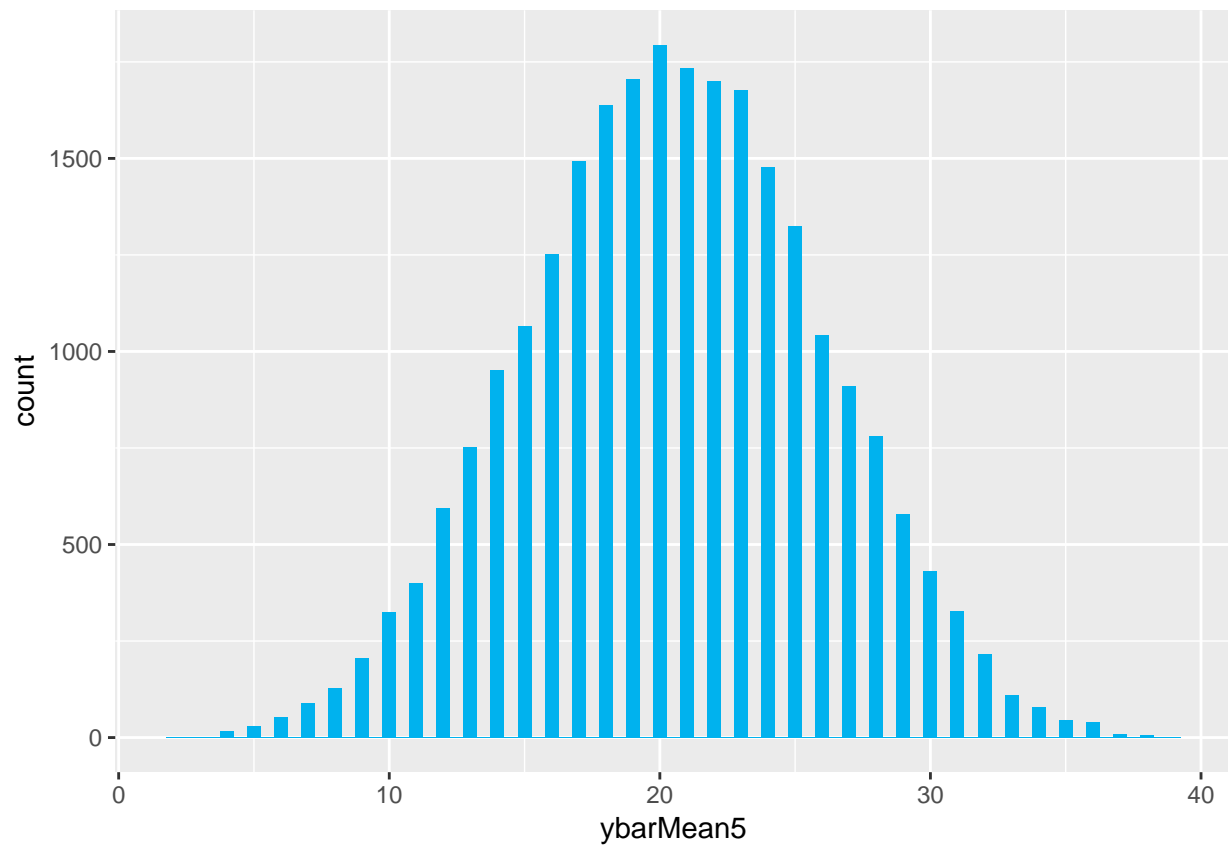
Tamanho $n = 25$

```
r=25000
n=25
y=rep(0,n)
ybarMean25=rep(0,r)
ybarMedian25=rep(0,r)
ybarSd25=rep(0,r)
for(i in 1:r){
  y = sample(coins, n, replace=T)
  ybarMean25[i] = floor(mean(y))
  ybarMedian25[i] = floor(median(y))
  ybarSd25[i] = floor(sd(y))
}
```

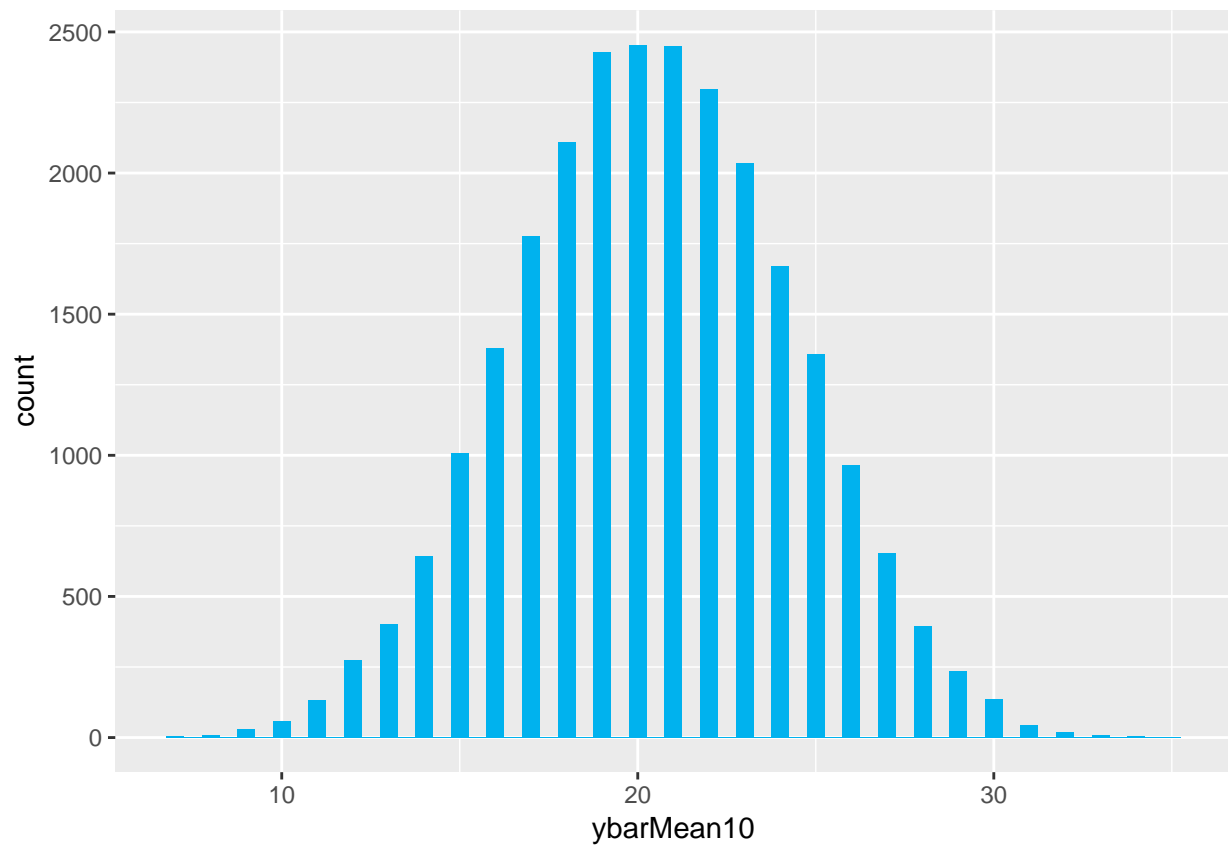
Plotagem dos gráficos

Médias

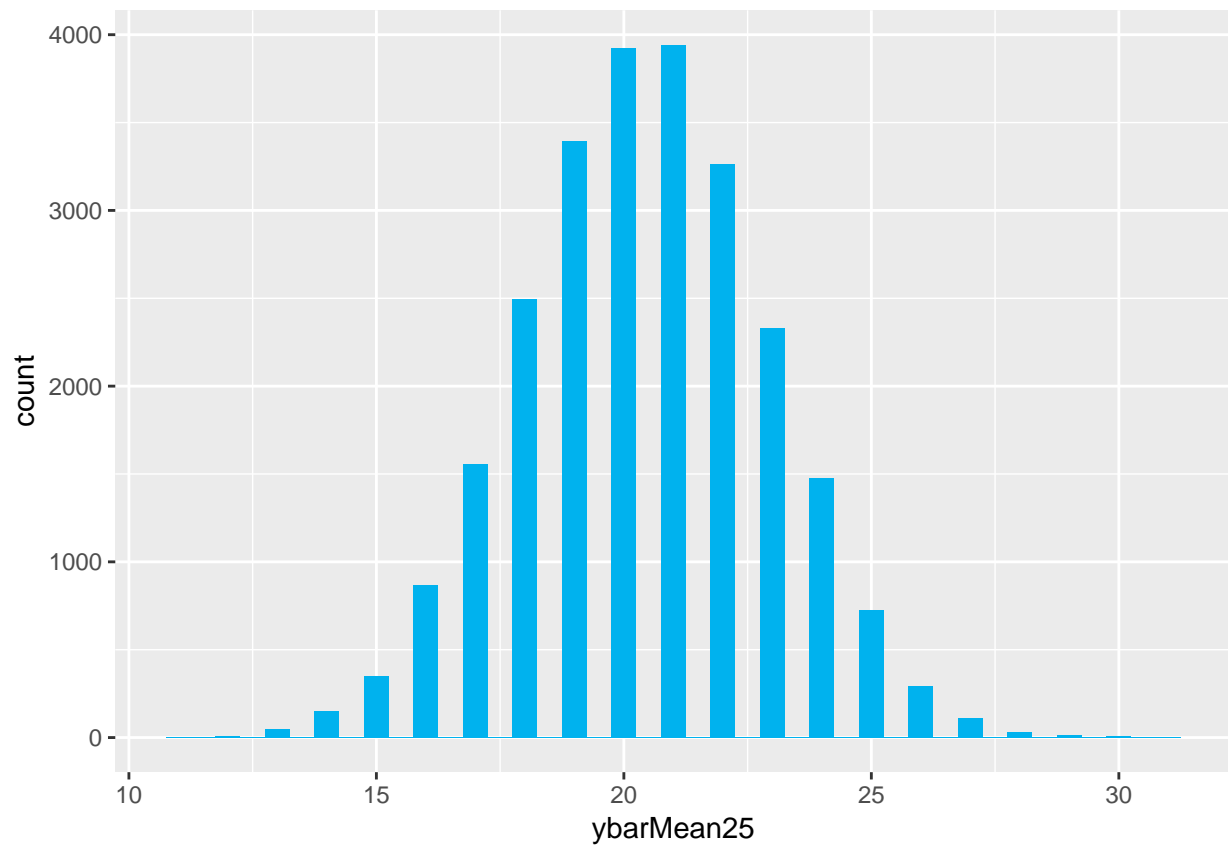
```
qplot(ybarMean5, geom="histogram", binwidth=0.5, fill=I("deepskyblue2"))
```



```
qplot(ybarMean10, geom="histogram", binwidth=0.5, fill=I("deepskyblue2"))
```

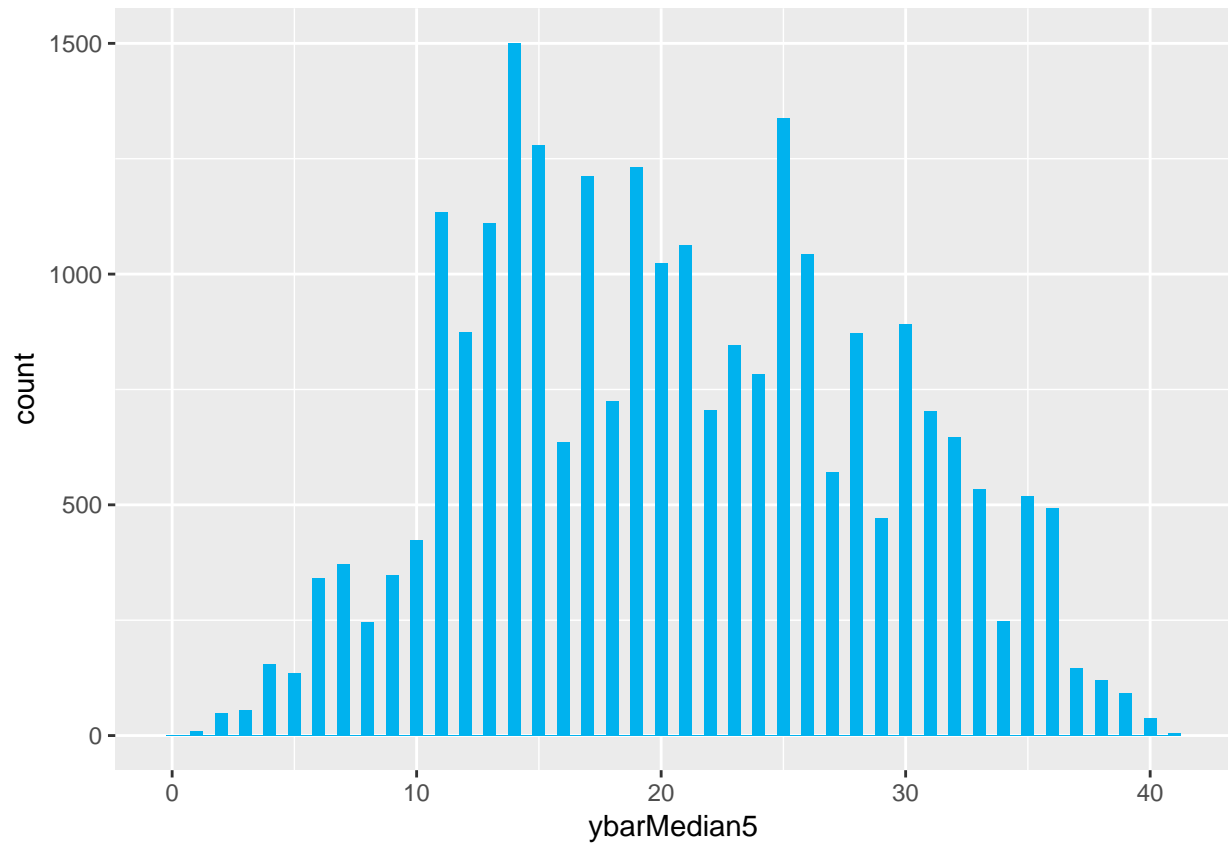


```
qplot(ybarMean25, geom="histogram", binwidth=0.5, fill=I("deepskyblue2"))
```

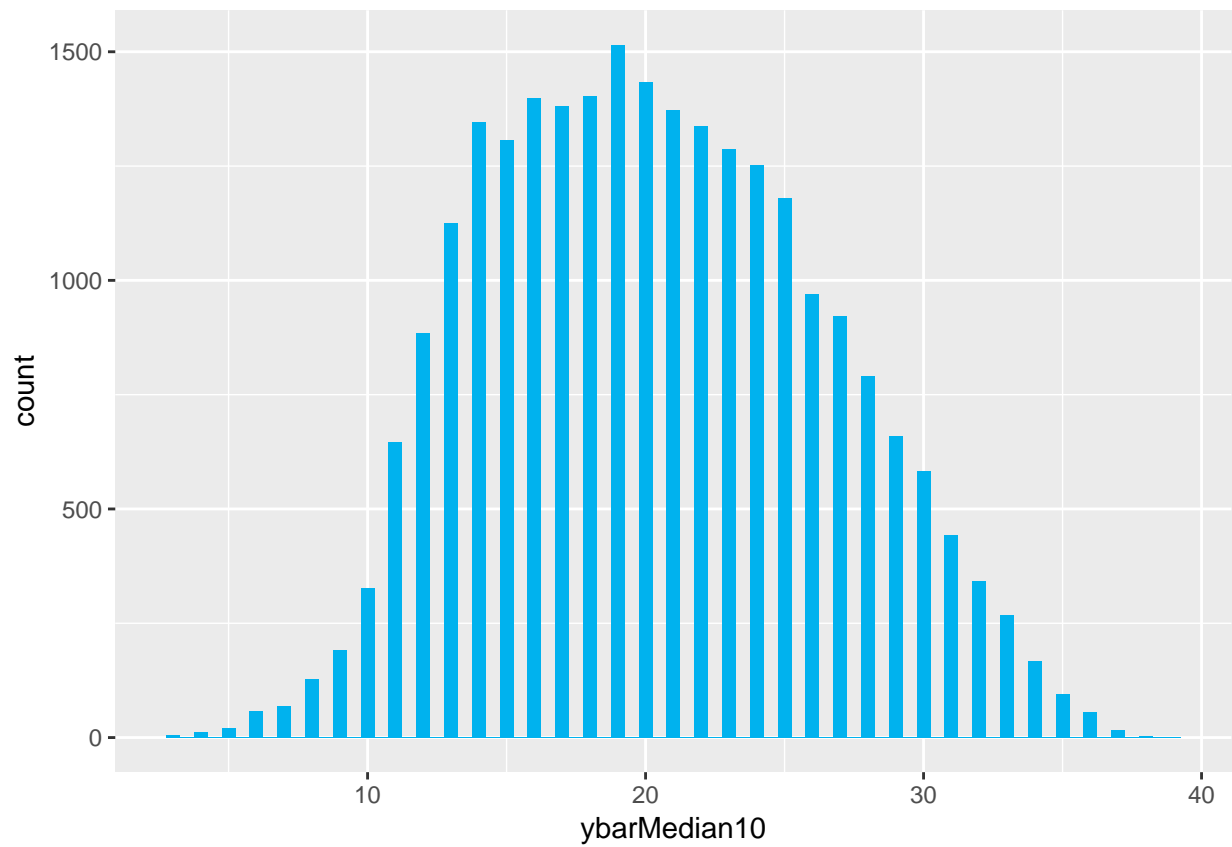


Medianas

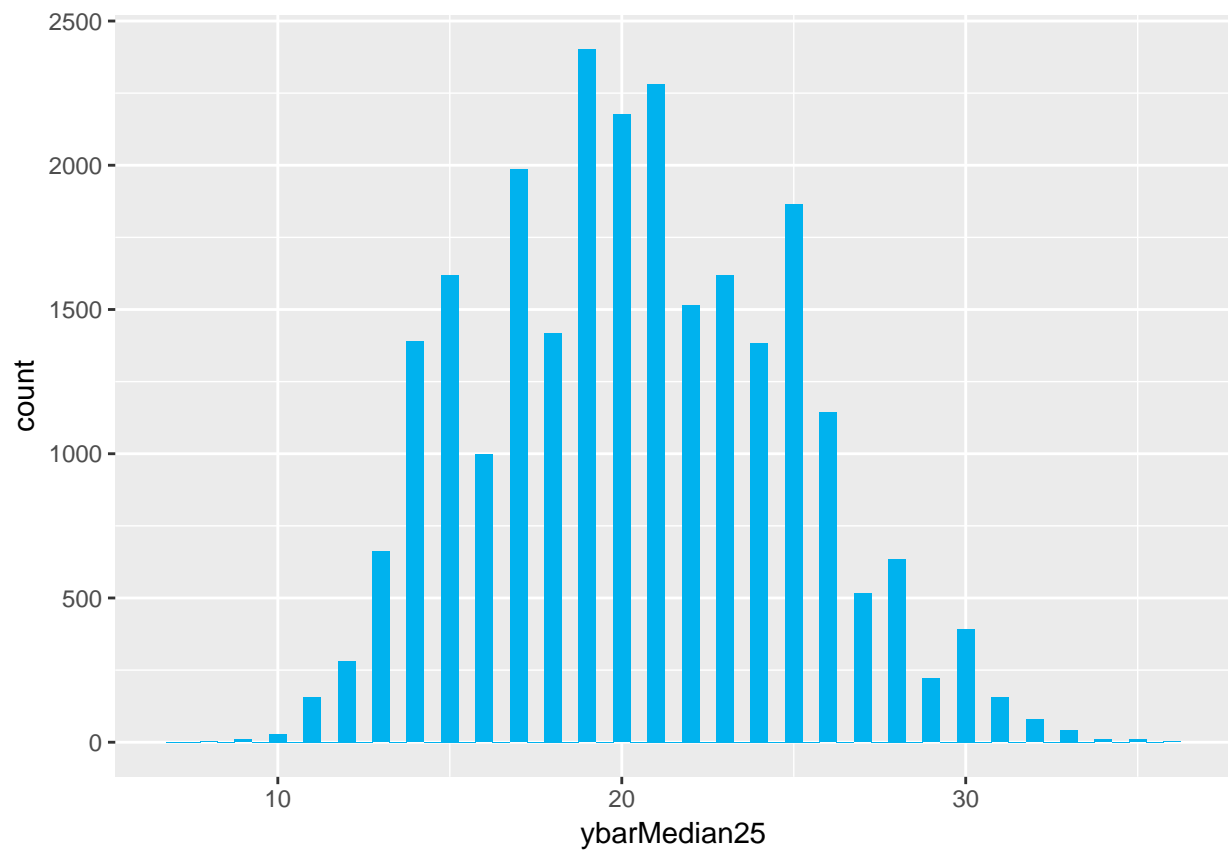
```
qplot(ybarMedian5, geom="histogram", binwidth=0.5, fill=I("deepskyblue2"))
```



```
qplot(ybarMedian10, geom="histogram", binwidth=0.5, fill=I("deepskyblue2"))
```

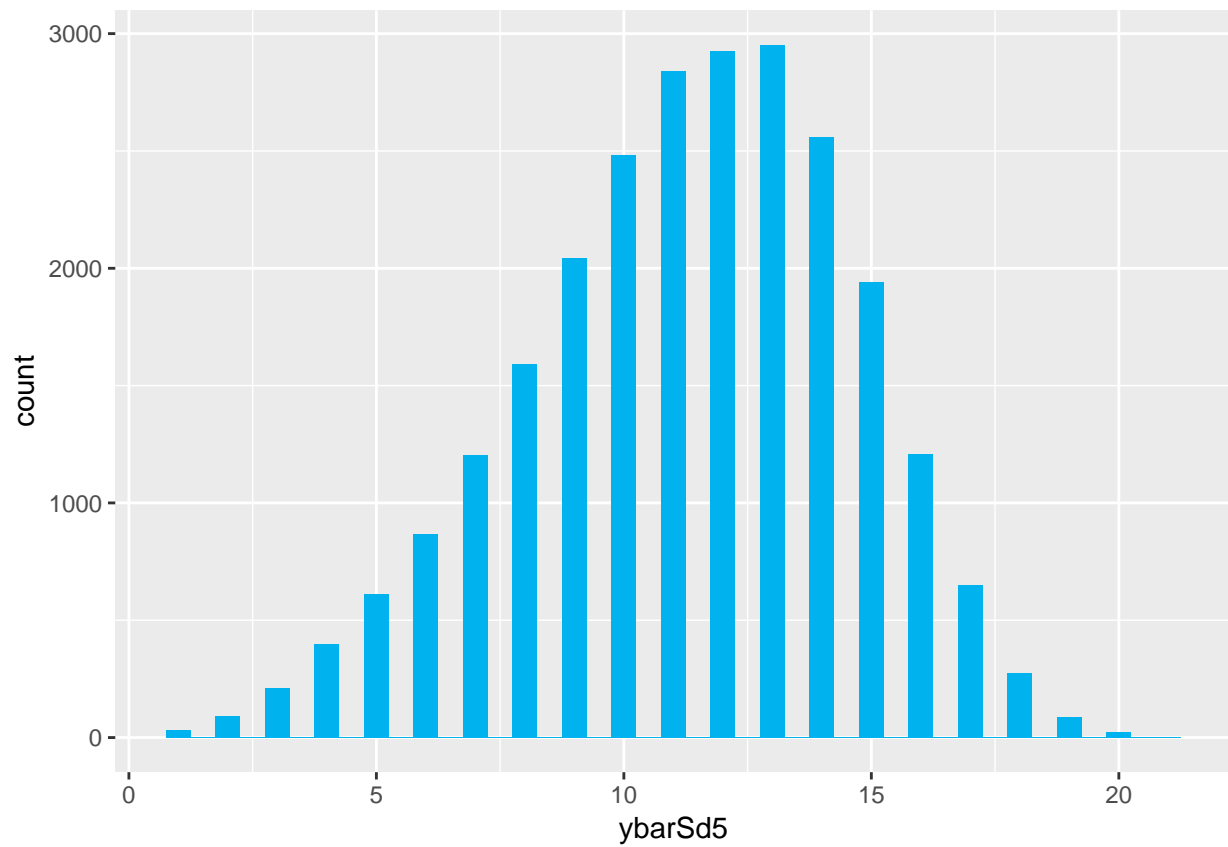


```
qplot(ybarMedian25, geom="histogram", binwidth=0.5, fill=I("deepskyblue2"))
```

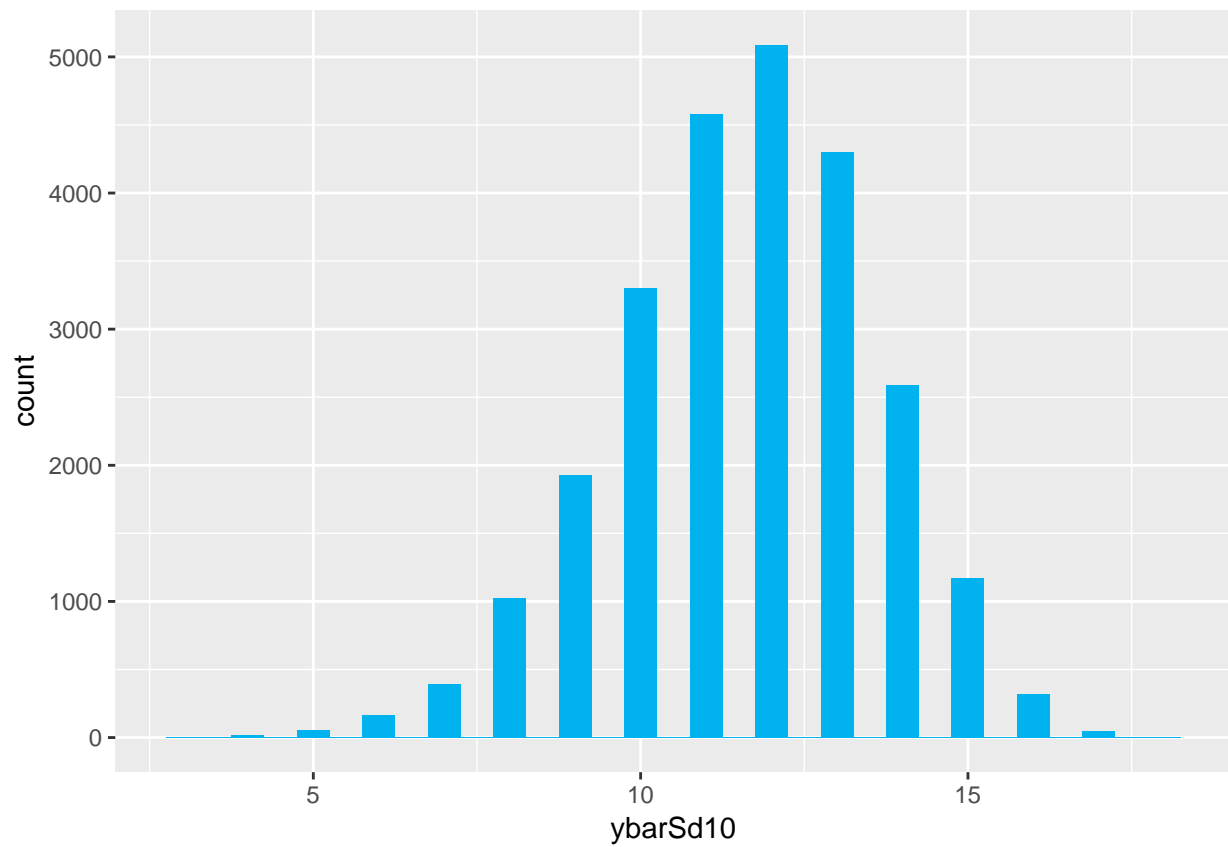



Desvio Padrão

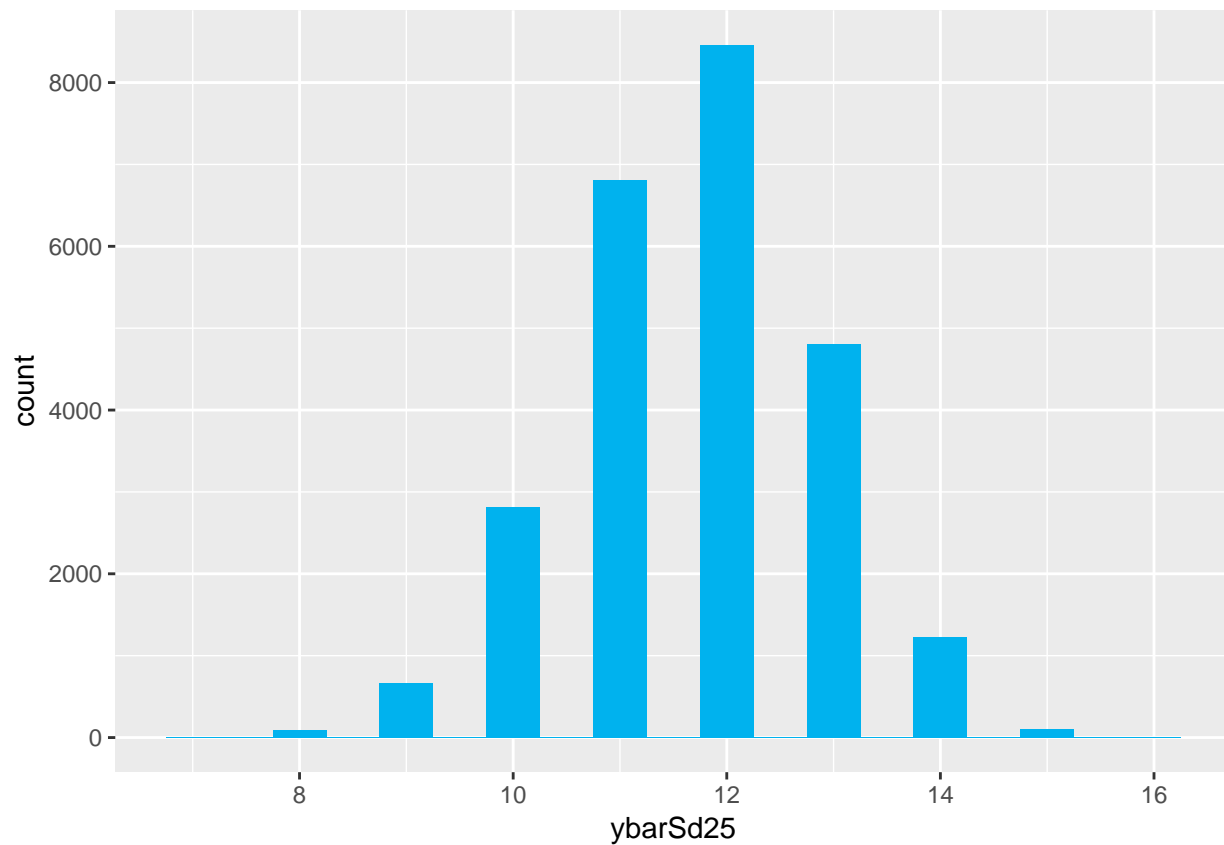
```
qplot(ybarSd5, geom="histogram", binwidth=0.5, fill=I("deepskyblue2"))
```



```
qplot(ybarSd10, geom="histogram", binwidth=0.5, fill=I("deepskyblue2"))
```



```
qplot(ybarSd25, geom="histogram", binwidth=0.5, fill=I("deepskyblue2"))
```



Sumário

```
n <- c("standard", "5", "10", "25")
```

```
medias <- c(
  mean(coins),
  mean(ybarMean5),
  mean(ybarMean10),
  mean(ybarMean25))
```

```
desvio_padrao <- c(
  sd(coins),
  sd(ybarSd5),
  sd(ybarSd10),
  sd(ybarSd25))
```

```
data.frame(n, medias, desvio_padrao)
```

```
##      n  medias desvio_padrao
## 1 standard 20.90800    12.262147
## 2      5 20.50564     3.305260
## 3     10 20.43532     1.991034
## 4     25 20.41436     1.173431
```

Conclusão

Podemos observar que quanto menor o tamanho da amostra, mais os dados (média, mediana e desvio padrão) ficam esparsos. Além disso, podemos visualizar na tabela de sumário que o fato de aumentarmos o tamanho da amostra não implica em mudanças bruscas no valor da média. No entanto, se aumentarmos o tamanho da amostra o desvio padrão diminui consideravelmente.