Lynx & Binomial RV Simulation

Oliver Hannaoui

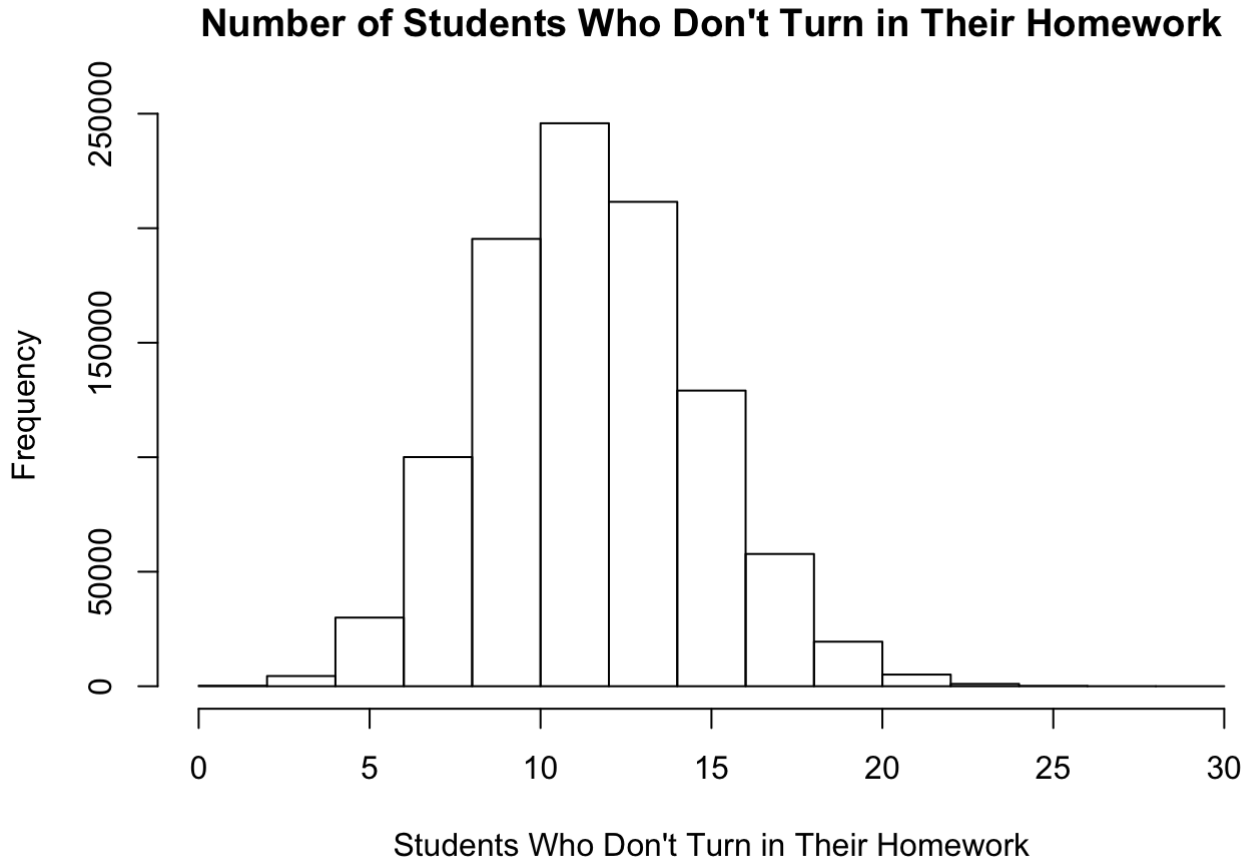
5/27/2019

Simulating a binomial random variable. Considering a class of 80 students and that the probability that a student does not turn in a homework is 0.15. Assuming independence and no change in probability as well.

a. Using the sample function, the number of students who don't turn in their homework is 5.

```
## [1] 5
```

b. Due to the fact that n>= 30 the distribution of the histogram is symmetric and has a normal shape.



c. The average and standard number of successes in 80 trials:

```
## [1] 11.99925
```

```
## [1] 3.195229
```

d. The probability that all students turn in their homework based on our simulation:

```
## [1] 4e-06
```

e. The probability that at least four students did not turn in their homework:

```
## [1] 0.998804
```

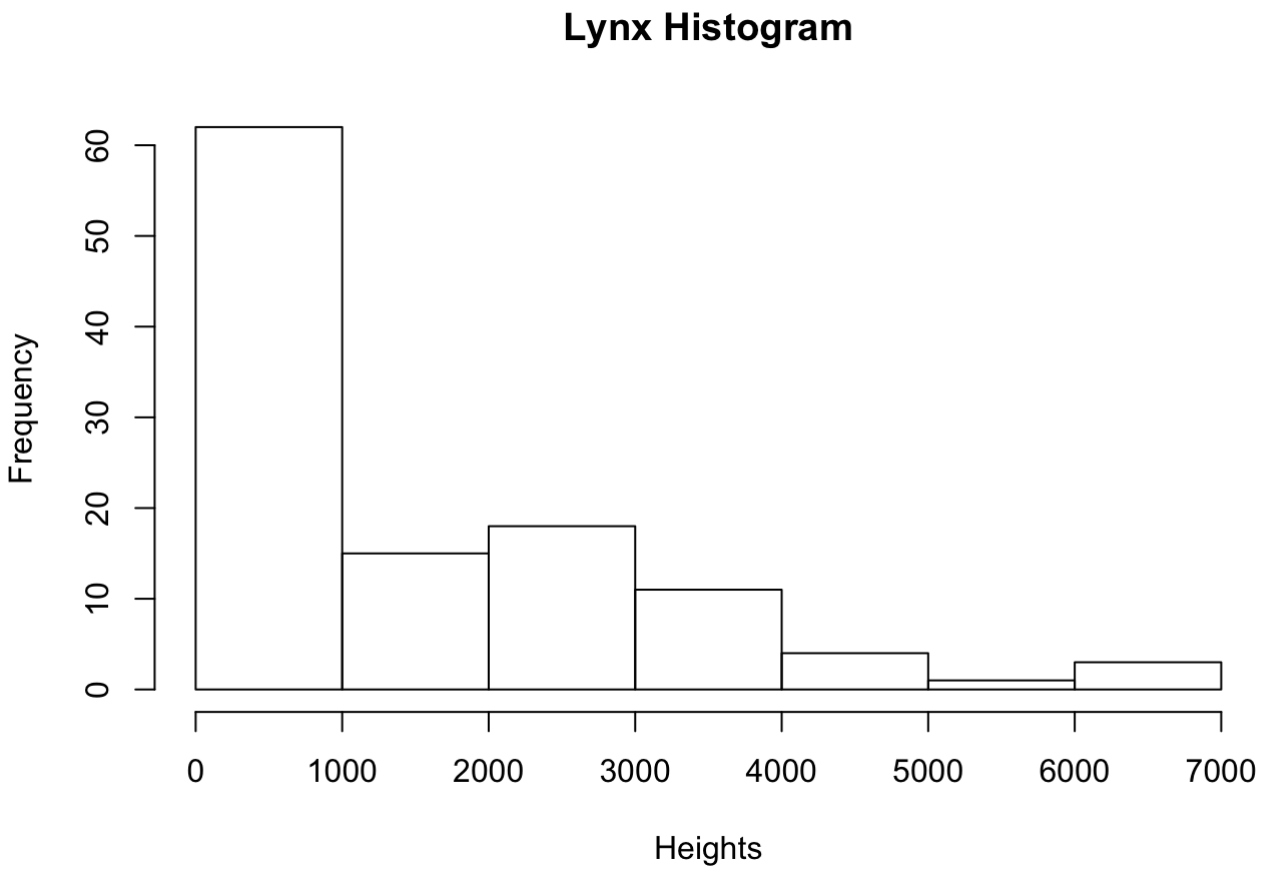
f. The median number of students who will forget their homework:

```
## [1] 12
```

II.

a. Data is used from built in “lynx” dataset.

The distribution of the histogram is skewed right.



Mean and Variance

```
## [1] 1538.018
```

```
## [1] 2514901
```

b. Mean of the random sample of 10.

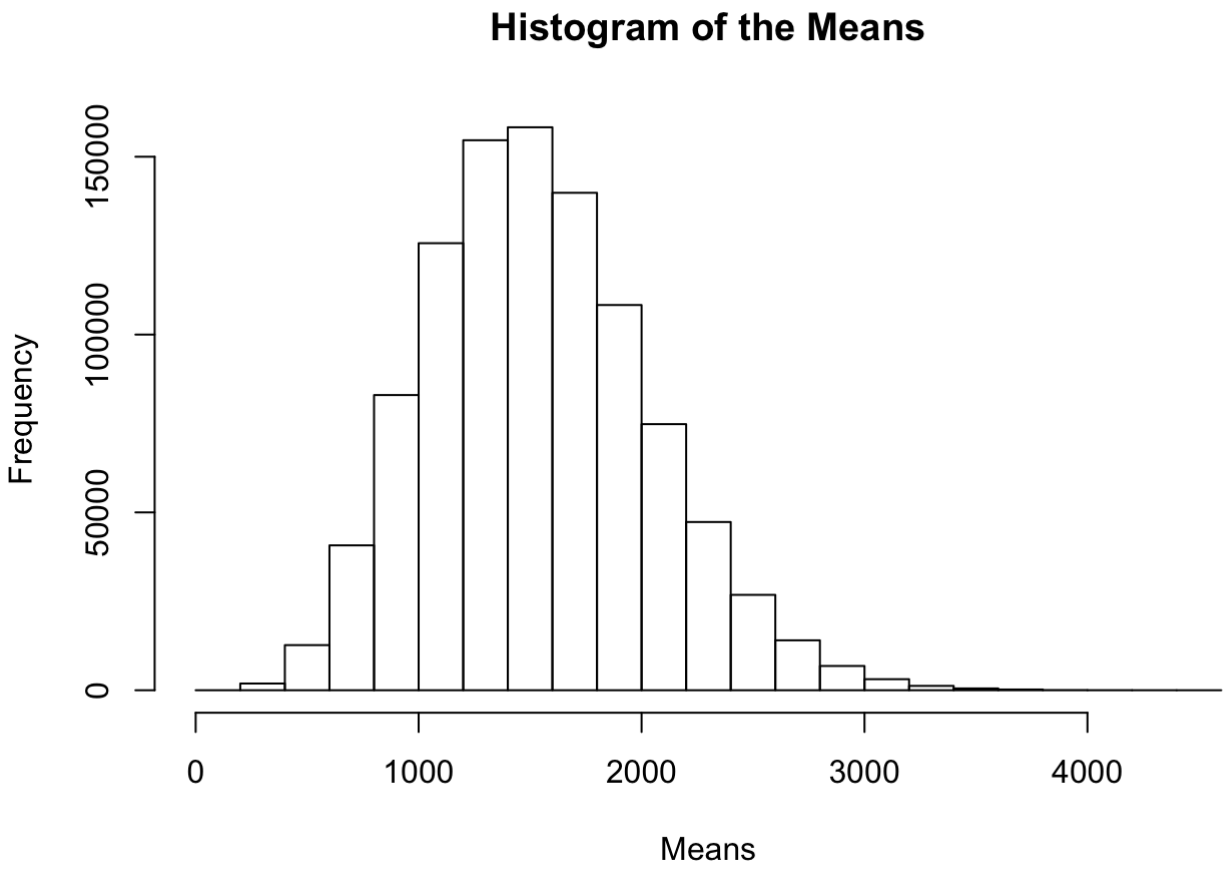
```
## [1] 1349.3
```

c. Mean and standard deviation of the 1000000 vector means.

```
## [1] 1537.531
```

```
## [1] 498.9672
```

The distribution is slightly skewed right and the means are spread out in terms of frequency.



d. Mean of the sample size of 50:

```
## [1] 1535.886
```

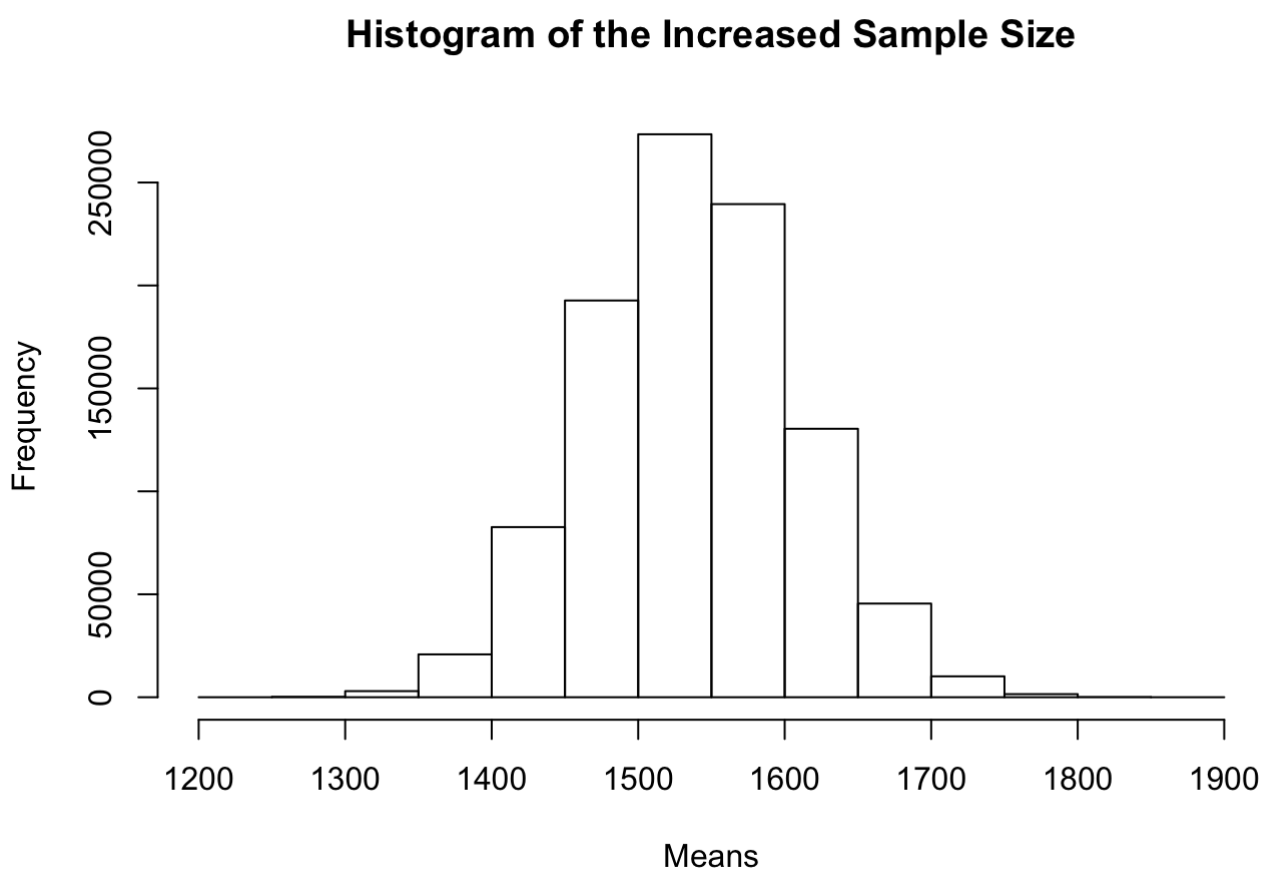
e. Mean stays approximately the same as before, however our standard deviation decreases substantially.

```
## [1] 1537.611
```

```
## [1] 70.60939
```

f. The distribution is approximately normally distributed and the means are a lot less spread out than our previous histogram, as the sample size increased our histogram become more normally distributed and our

standard deviation decreased.



Appendix:

```
knitr::opts_chunk$set(echo = TRUE)
set.seed(10)
#1(a)
experiment = sample(c("X","F"),80,prob=c(0.15,0.85),replace=TRUE)
sum(experiment=="X")
many.binom = sapply(1:1000000, function(i) {
  binom = sample(c("X","F"), 80,
                 prob = c(0.15, 0.85), replace = TRUE)
  num.of.success = sum(binom == "X")
  return(num.of.success)
  })

hist1 = hist(many.binom, main = "Number of Students Who Don't Turn in Their Homework", x
lab = "Students Who Don't Turn in Their Homework", ylab = "Frequency")
#1(c)
mean(many.binom)
sd(many.binom)
#1(d)
prob1 = sum(many.binom == 0)
prob1/length(many.binom)
#1(e)
prob3 = sum(many.binom >= 4)
prob3/length(many.binom)
#1(f)
median(many.binom)
#2(a)
hist(lynx,main="Lynx Histogram",xlab="Heights")

mean(lynx)
var(lynx)
#2b
experiment = sample(lynx, 10, replace=TRUE)
mean(experiment)
#2c
 many.experiments = sapply(1:1000000, function(i){
  experiment = sample(lynx,10, replace = TRUE)
  sample.mean= mean(experiment)
  return(sample.mean)
})
mean(many.experiments)
sd(many.experiments)

hist(many.experiments, xlab='Means', main= 'Histogram of the Means')
#2d
experiment = sample(many.experiments, 50, replace=TRUE)
mean(experiment)
#2e
 more.experiments = sapply(1:1000000, function(i){
  experiment = sample(many.experiments,50, replace = TRUE)
  sample.mean= mean(experiment)
  return(sample.mean)
})
mean(more.experiments)
```

```
sd(more.experiments)
#2f
hist(more.experiments, xlab='Means', main= 'Histogram of the Increased Sample Size')
```