Math 490 HW #3

Maxwell Levin

January 26, 2018

Question 1.

Consider choosing n = 36 numbers at random from 1, 2, …, 12 with replacement, and let $\overline{Y}_{36}$ be the sample mean.

a. What are the mean and standard deviation of $\overline{Y}_{36}$?

Since the mean $\overline{Y}_{36}$ does not change with our sample size, $\overline{Y}_{36}$ is the same as $\overline{Y}_1$. In other words, we can compute $\overline{Y}_{36}$ by simply computing the expectation value of choosing a single number at random from 1, 2, …, 12:

$$\overline{Y}_{36} = \frac{1}{12}(1 + 2 + \cdots + 12) = 6.5.$$

Alternatively, we can use R to calculate our mean:

```
X = c(1:12)
mean(X)
```

```
[1] 6.5
```

Similarly, we've shown in class that variance for $\overline{Y}_{36}$ decreases proportionally to the sample size. That is, $\overline{Y}_{36}$ is $\frac{1}{36}Var(X)$, where X is a single choice from 1, 2, …, 12. We can use R to calculate $Var(X)$ and then use our result to calculate the variance of $\overline{Y}_{36}$:

```
X = c(1:12)
var(X)
```

```
[1] 13
```

$$Var(\overline{Y}_{36}) = \frac{1}{36}Var(X) \approx 0.361$$

To compute the standard deviation of $\overline{Y}_{36}$ we can take the square root of our variance:

$$SD(\overline{Y}_{36}) = \sqrt{Var(\overline{Y}_{36})} \approx 0.601$$

We can double-check this using the fact that the standard deviation of $\overline{Y}_{36}$ decreases by a factor of the square root of the sample size and some simple R code:

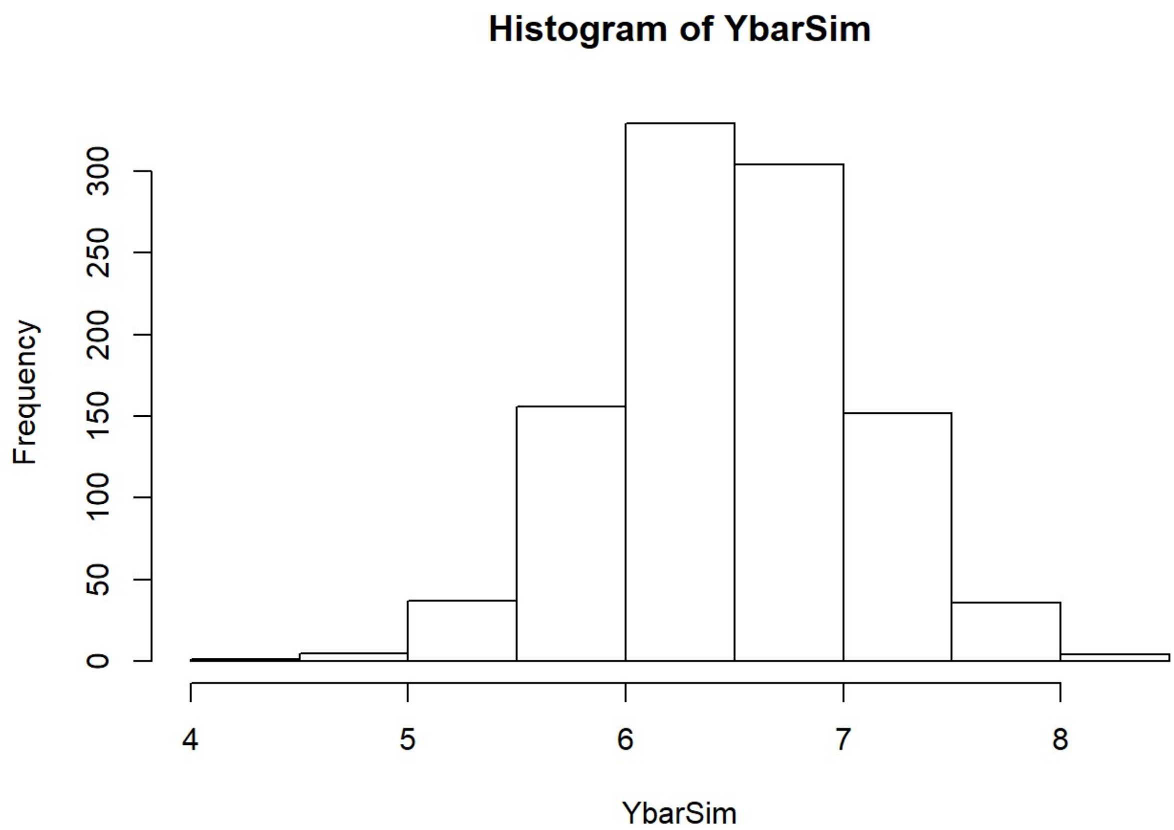
```
X = c(1:12)
SD = sd(X)
SD/6
```

```
[1] 0.6009252
```

b. Use R to generate 1024 sample means $\overline{Y}_{36}$, make a histogram, and report the mean and standard deviation of the simulated sample means.

```
func = function(sam, rep) {
  # sam is the sample size
  # rep is the number of repititions
  obs = NULL
  for (i in 1:rep) {
    ybar = mean(sample(1:12, sam, replace=T))
    obs = c(obs, ybar)
  }
  obs;
}

YbarSim = func(36, 1024)
hist(YbarSim)
```



The mean of our simulation is:

```
[1] 6.500949
```

The standard deviation of our simulation is:

```
[1] 0.5765095
```

c. For your simulation in part (b), what is the proportion of simulated sample means falling within one standard deviation from the center?

To compute this, we can run the following code in R:

```
portion = YbarSim[abs(YbarSim - mean(YbarSim)) <= sd(YbarSim)]
length(portion) / length(YbarSim)
```

```
[1] 0.6845703
```

Thus we see that the proportion of simulated sample means that fall within one standard deviation of the mean is close to 68% . If we modify this code to see the proportion of simulated sample means that fall within two standard deviations of the center, we should see that we get a number close to 95% :

```
portion = YbarSim[abs(YbarSim - mean(YbarSim)) <= 2*sd(YbarSim)]
length(portion) / length(YbarSim)
```

```
[1] 0.9550781
```

Wonderful!