Lab 2

Adam Orr

September 7, 2017

Section: 91973 Friday 9am

Question 1

Script

```
nfemales <- 3
total <- 9
p <- .5

(combinations <- choose(total, nfemales))
(p_each <- p ^ (nfemales) * (1 - p) ^ (total - nfemales))
(tot_p <- combinations * p_each)
```

Output

```
## [1] 84
## [1] 0.001953125
## [1] 0.1640625
```

Answers

What are five of the possible combinations with three females and six males? Pick any five you please.

- FFFMMMMMM
- MFFFMMMMM
- MMFFFMMMM
- MMMFFFMMM
- FMFMFMMMM

Question 2

Script

```
nfemales <- seq(0,9)
total <- rep(9,10)
p <- rep(.5,10)

(combinations <- choose(total, nfemales))
(p_each <- p ^ (nfemales) * (1 - p) ^ (total - nfemales))
(tot_p <- combinations * p_each)
```

Output

```
##  [1]   1   9  36  84 126 126  84  36   9   1
##  [1] 0.001953125 0.001953125 0.001953125 0.001953125 0.001953125
##  [6] 0.001953125 0.001953125 0.001953125 0.001953125 0.001953125
##  [1] 0.001953125 0.017578125 0.070312500 0.164062500 0.246093750
##  [6] 0.246093750 0.164062500 0.070312500 0.017578125 0.001953125
```

Answers

Question 3

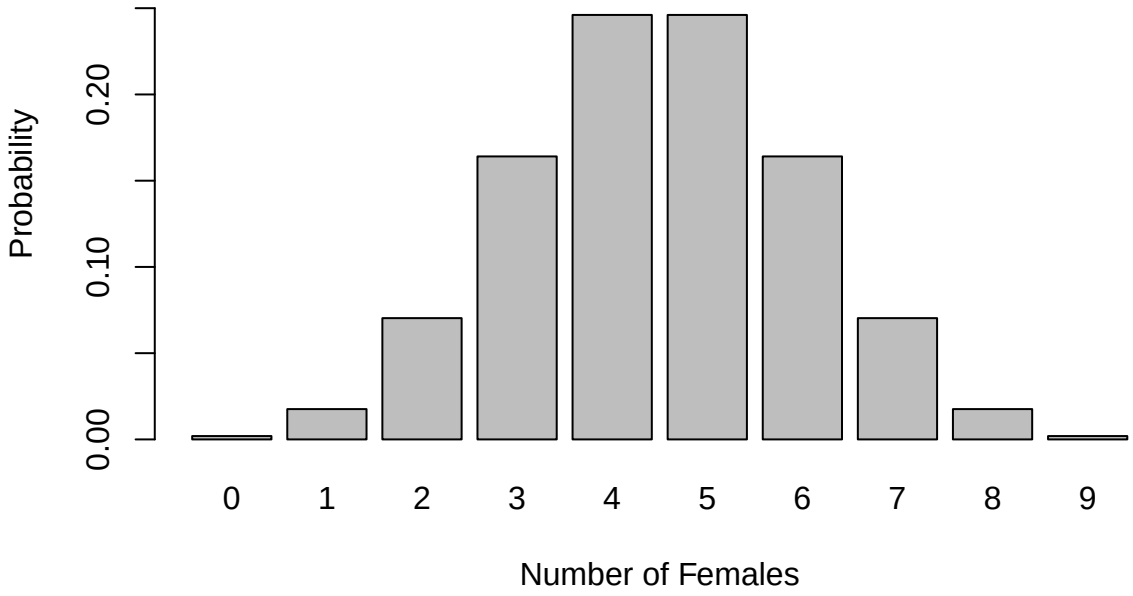
Script

```
nfemales <- seq(0,9)
total <- 9
p <- .5

(prob <- dbinom(nfemales, size = total, .5))
barplot(prob,
        names.arg = nfemales,
        xlab = "Number of Females",
        ylab = "Probability",
        ylim = c(0, max(prob + .05)))
(mu <- sum(nfemales * prob))
(mu <- total * p)
(variance <- sum((nfemales - mu)^2 * prob))
(sdev <- sqrt(variance))
(variance <- mu * (1 - p))
(sdev <- sqrt(variance))

mn <- function(total, p){ total * p }
va <- function(mn, p){ mn * (1 - p)}
ps <- c(0.1,0.3,0.5,0.7,0.9)
means <- mn(rep(total,length(ps)),ps)
variances <- va(means, ps)
data.frame(p = ps,mean = means, variance = variances)
```

Output

```
##  [1] 0.001953125 0.017578125 0.070312500 0.164062500 0.246093750
##  [6] 0.246093750 0.164062500 0.070312500 0.017578125 0.001953125
## [1] 4.5
## [1] 4.5
## [1] 2.25
## [1] 1.5
## [1] 2.25
## [1] 1.5
##     p mean variance
## 1 0.1  0.9     0.81
## 2 0.3  2.7     1.89
```

```
## 3 0.5  4.5     2.25
## 4 0.7  6.3     1.89
## 5 0.9  8.1     0.81
```



Answers

What happens to the mean and variance of the distribution as the probability changes?

As the probability increases, the mean increases as well. On the other hand, the variance until `p = .5`, then decreases again.

At what value is the variance of the distribution highest? Lowest?

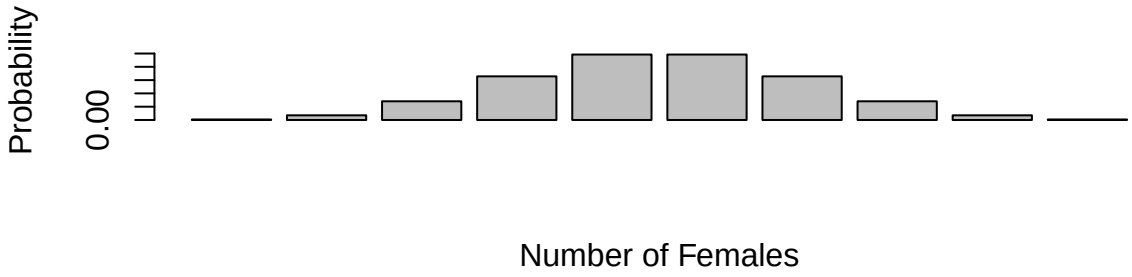
The variance of the distribution is highest when `p = .5`. The variance is lowest when `p = .1` or `p = .9`.

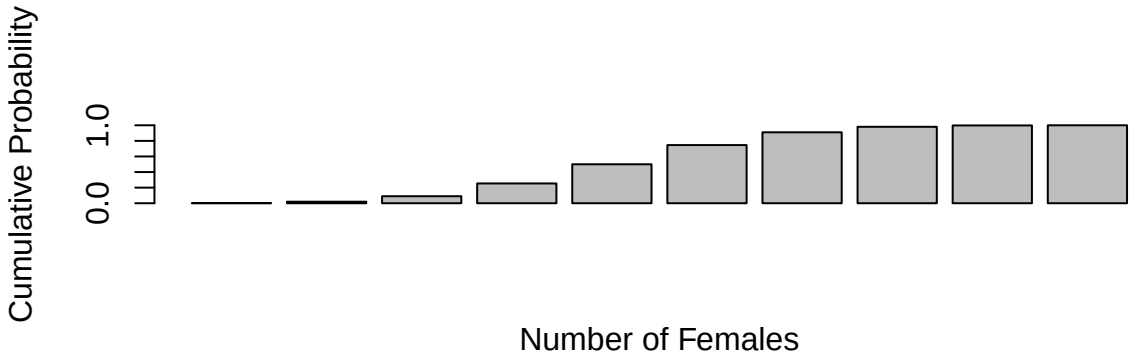
Question 4

Script

```r
par(mfrow=c(2,1))
barplot(prob, xlab = "Number of Females", ylab = "Probability", ylim = c(0, max(prob + .05)))
barplot(cumsum(prob), xlab = "Number of Females", ylab = "Cumulative Probability")
```

Output





Answers

What is the relationship between each bar of the cumulative plot and the bars of the distribution plot?

Each bar of the cumulative plot is the height of that bar in the distribution plot added to each bar before that one in the distribution plot.

Question 5

Script

```
p <- .5
dbinom(4,9,p) # exactly 4 females
dbinom(7,9,p) # exactly 7 females
pbinom(7,9,p) # 7 or fewer females
1 - pbinom(3,9,p) # 4 or more females
pbinom(5,9,p) # 4 or more males
(1 - pbinom(6,9,p)) + pbinom(2,7,p) #7 or more females + 7 or more males
```

Output

```
## [1] 0.2460938
## [1] 0.0703125
## [1] 0.9804688
## [1] 0.7460937
```

```
## [1] 0.7460937
## [1] 0.3164063
```

Answers

Question 6

Script

```
qbinom(.25,9,.5) # first quartile
qbinom(.05,9,.5) # 5th percentile
qbinom(.95,9,.5) # 95th percentile
qbinom(.5,9,.5) # median
```

Output

```
## [1] 3
## [1] 2
## [1] 7
## [1] 4
```

Answers

Question 7

Script

```
(vals <- rbinom(100,9,.5)) #100 simulated values
mean(vals) #mean of simulated values
```

Output

```
##   [1] 4 6 6 6 4 3 2 9 5 6 6 3 3 5 7 6 8 4 7 4 5 4 4 8 6 5 3 4 3 5 3 3 5 6 5
##  [36] 4 5 3 6 2 5 4 6 3 6 5 7 7 3 4 5 4 5 7 5 5 4 2 5 8 4 5 4 6 4 3 4 5 6 3
##  [71] 4 3 5 5 4 4 4 4 3 4 7 1 4 2 3 5 3 3 6 4 2 6 4 4 2 7 6 5 4 6
## [1] 4.61
```

Answers

Compare the average you calculated here with the mean you calculated in question 3. Do you expect them to be the same? Why or why not?

The average I calculated is 4.61, which is very close to the average I calculated in question 3, 4.5. I do not expect these values to be the same due to the stochastic nature of the experiments.