Laboratory Exercise Week 6

Ryan Estes | 006, 10 pts

9/27/2017

*Directions*:

* Write your R code inside the code chunks after each question.
* Write your answer comments after the # sign.
* To generate the word document output, click the button Knit and wait for the word document to appear.
* RStudio will prompt you (only once) to install the knitr package.
* Submit your completed laboratory exercise using Blackboard's Turnitin feature. Your Turnitin upload link is found on your Blackboard Course shell under the Laboratory folder.

For this exercise, you will need to use the package mosaic to find numerical and graphical summaries.

library(mosaic) # load the package mosaic to use its functions

## Loading required package: dplyr

##   
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':  
##   
## filter, lag

## The following objects are masked from 'package:base':  
##   
## intersect, setdiff, setequal, union

## Loading required package: lattice

## Loading required package: ggformula

## Loading required package: ggplot2

##   
## New to ggformula? Try the tutorials:   
## learnr::run\_tutorial("introduction", package = "ggformula")  
## learnr::run\_tutorial("refining", package = "ggformula")

## Loading required package: mosaicData

## Loading required package: Matrix

##   
## The 'mosaic' package masks several functions from core packages in order to add   
## additional features. The original behavior of these functions should not be affected by this.  
##   
## Note: If you use the Matrix package, be sure to load it BEFORE loading mosaic.

##   
## Attaching package: 'mosaic'

## The following object is masked from 'package:Matrix':  
##   
## mean

## The following objects are masked from 'package:dplyr':  
##   
## count, do, tally

## The following objects are masked from 'package:stats':  
##   
## binom.test, cor, cor.test, cov, fivenum, IQR, median,  
## prop.test, quantile, sd, t.test, var

## The following objects are masked from 'package:base':  
##   
## max, mean, min, prod, range, sample, sum

1. A [national survey](http://www.pewinternet.org/2012/11/01/how-teens-do-research-in-the-digital-world/) in 2012 showed that 75% of middle and high school teachers believe their students use Wikipedia for typical research statements. Let X equal the number of teachers in a random sample of size n = 40 who believe their students use Wikipedia.
   1. How is X distributed?
   2. Use plotDist() to draw the probability distribution of X.
   3. Compute the probability that X is equal to 30.
   4. Compute the probability that X is at most 30. What is the difference between parts (iii) and (iv).
   5. Compute the probability that X is anywhere from 30 to 35.

### Code chunk

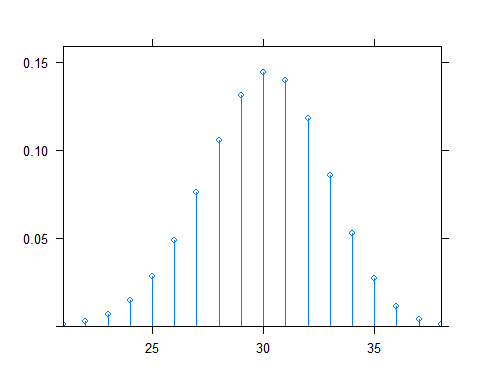
# start your code  
  
# i)  
(X <- rbinom(n = 1, size = 40, prob = 0.75))

## [1] 32

favstats(X)

## min Q1 median Q3 max mean sd n missing  
## 32 32 32 32 32 32 NA 1 0

# X is a bell curve with the top of the bell curve being around the 30 mark for a given sample.  
  
# ii)  
plotDist('binom', size = 40, prob = 0.75)



# iii)  
dbinom(x = 30, size = 40, prob = 0.75) #14.4%

## [1] 0.1443643

# iv)  
pbinom(q = 30, size = 40, prob = 0.75) #56%

## [1] 0.5604603

# v)  
pbinom(q = 35, size = 40, prob = 0.75) - pbinom(q = 29, size = 40, prob = 0.75) #56.8%

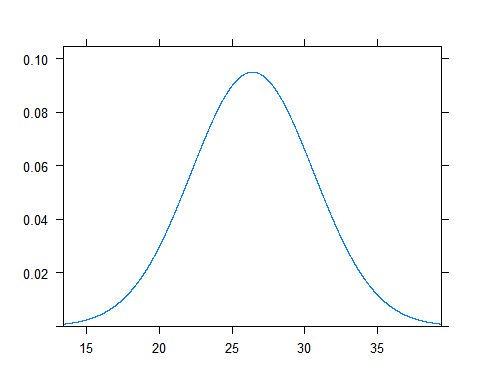
## [1] 0.5678618

# last R code line

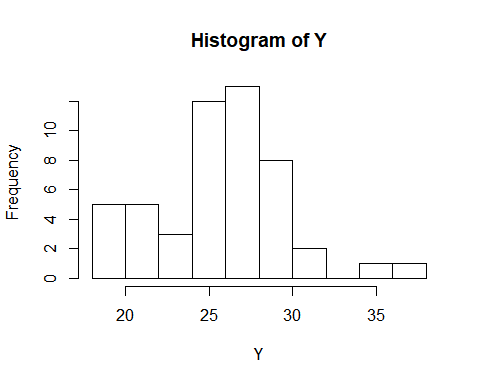
1. The size of the left upper chamber of the heart is one measure of cardiovascular health. When the upper left chamber is enlarged, the risk of heart problems is increased. The paper ("Left atrial size increases with body mass index in children")[<https://www.ncbi.nlm.nih.gov/pubmed/19147240>] (International Journal of Cardiology [2009]) described a study in which the left atrial size was measured for a large number of children age 5 to 15 years. Based on this data, the authors concluded that for healthy children, the left atrial diameter was approximately normally distributed with a mean of 26.4 mm and a standard deviation of 4.2 mm.
   1. Use plotDist to draw the probability density function of the left atrial diameter.
   2. Generate 50 sample measurements of the left atrial diameters and plot the measurements using a histogram.
   3. Approximately what proportion of healthy children have left atrial diameters less than 23.5 mm?
   4. Approximately what proportion of healthy children have left atrial diameters greater than 32.5 mm?
   5. Approximately what proportion of healthy children have left atrial diameters between 24 and 29 mm?
   6. For healthy children, what is the value for which only about 25% have larger left atrial diameter?

### Code chunk

# start your code  
# i)  
plotDist('norm', mean = 26.4, sd = 4.2)



# ii)  
Y <- rnorm(n = 50, mean = 26.4, sd = 4.2)  
hist(Y)



# iii)  
pnorm(q = 23.5, mean = 26.4, sd = 4.2) #24.5%

## [1] 0.2449474

# iv)  
1 - pnorm(q = 32.5, mean = 26.4, sd = 4.2) #0.08%

## [1] 0.07319785

# v)  
pnorm(q = 29, mean = 26.4, sd = 4.2) - pnorm(q = 24, mean = 26.4, sd = 4.2) #44.8%

## [1] 0.4482029

# vi)  
qnorm(p = 0.75, mean = 26.4, sd = 4.2) #75th percentile, 29.23%

## [1] 29.23286

# last R code line