Lab Assignment: Chapter 17 - Probablity Distributions

Michael Beebe

library(ggplot2)  
library(reshape2)  
library(stringr)

## 1. Probability Distribution

Using rnorm store 10 random numbers in a variable.

rand10 <- rnorm(10)  
rand10

## [1] 1.9425592 0.5021975 0.9368119 0.1746158 -0.3382840 0.4700923  
## [7] -0.8803537 1.3851012 -0.1226219 0.1701195

Find the probabilities of your numbers from above.

rand10Probs <- dnorm(rand10)  
rand10Probs

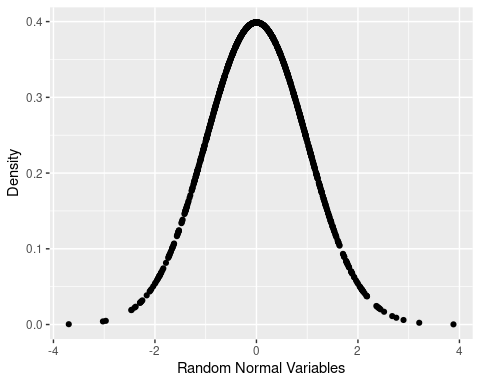
## [1] 0.06046403 0.35167787 0.25723974 0.39290639 0.37675636 0.35720983  
## [7] 0.27077965 0.15286635 0.39595425 0.39321102

Create 2 new variables - one that will store a large number of random variables, and the other that stores their density (probabilities)

rand1000 <- rnorm(1000)  
rand1000Probs <- dnorm(rand1000)

Plot the distribution using ggplot.

ggplot(  
 data.frame(x=rand1000, y=rand1000Probs)) +   
 aes(x=x, y=y) +geom\_point() +   
 labs(x="Random Normal Variables", y="Density"  
)



## 2. Binomial Distribution

Calculate the probability and cumulative probability using the Binomial Distribution using the following scenarios:

Probability of tossing a die 10 times and getting a “3” at least 2 times.

* Probability of a success is 1/6 = 0.17

dbinom(x = 2, size = 10, prob = .17)

## [1] 0.2929106

Probability and getting at least a “B” (80%) on a quiz that has 10 true/false questions.

* Probability of getting a question right is 1/2 = 0.5.
* Use dbinom function to determine probability of getting a B by randomly guessing

dbinom(x = 8, size = 10, prob = .5)

## [1] 0.04394531

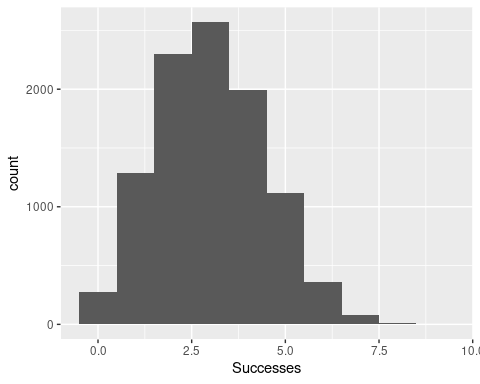
* Use pbinom function to determine probability of answering 8 or less questions correctly by random guess.

pbinom(q = 8, size = 10, prob = .5)

## [1] 0.9892578

Generate a graph based on a 10,000 random experiments, 10 trials, and 0.3 probability of success.

binomData <- data.frame (Successes= rbinom (n=10000, size=10, prob=.3))  
ggplot (data = binomData, aes (x=Successes)) + geom\_histogram (binwidth=1)



## 3. Poisson Distribution

We are interested in calculating the probability that today at least 5 callers will receive a busy signal if they call Lakeland’s Help Desk.

* The average probability of a busy signal is 0.5%
* There are 1,200 calls to the Lakeland Help Desk
* Use the ppois function with lower=FALSE

ppois(q = 1200, lambda = 0.005, lower.tail = FALSE)

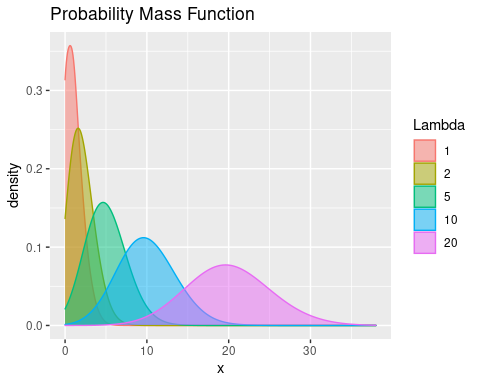
## [1] 0

Generate the plot on page 238 and explain what it shows.

pois1 <- rpois(n = 10000, lambda = 1)  
pois2 <- rpois(n = 10000, lambda = 2)  
pois5 <- rpois(n = 10000, lambda = 5)  
pois10 <- rpois(n = 10000, lambda = 10)  
pois20 <- rpois(n = 10000, lambda = 20)  
pois <- data.frame(  
 Lambda.1 = pois1,  
 Lambda.2 = pois2,  
 Lambda.5 = pois5,  
 Lambda.10 = pois10,  
 Lambda.20 = pois20  
)  
  
pois <- melt(data=pois, variable.name="Lambda", value.name="x")

## No id variables; using all as measure variables

pois$Lambda <- as.factor(  
 as.numeric(  
 str\_extract(string=pois$Lambda, pattern="\\d+")  
 )  
)  
  
  
ggplot (pois, aes (x=x)) +  
geom\_density ( aes (group=Lambda, color=Lambda, fill=Lambda), adjust=4, alpha=1/2) +  
scale\_color\_discrete () + scale\_fill\_discrete () +  
ggtitle ("Probability Mass Function")



This graph shows density plots for 10,000 draws from the Poisson distribution at varying levels of λ.