Worksheet 4

**Solution 1.a)**

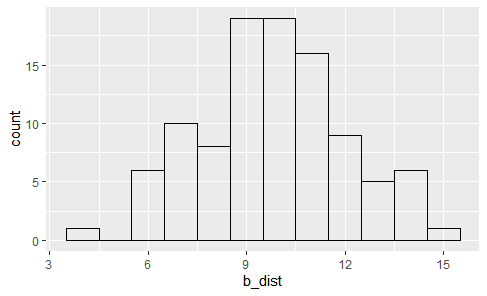
> library(ggplot2)

> b\_dist = rbinom(100,20,0.5)

> df\_b = data.frame(b\_dist)

> p\_b= ggplot(df\_b) + geom\_histogram(mapping=aes(x=b\_dist), color="black", fill="NA", binwidth=1)

> p\_b



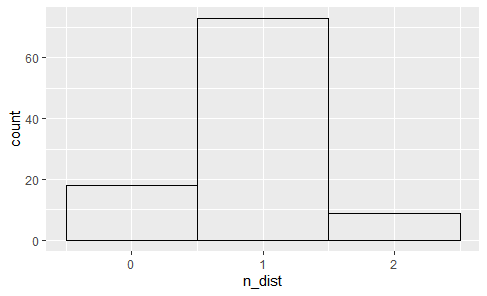
**Solution 1.b)**

> n\_dist = rnorm(100,1,0.5)

> df\_n = data.frame(n\_dist)

> p\_n= ggplot(df\_n) + geom\_histogram(mapping=aes(x=n\_dist), color="black", fill="NA", binwidth=1)

> p\_n

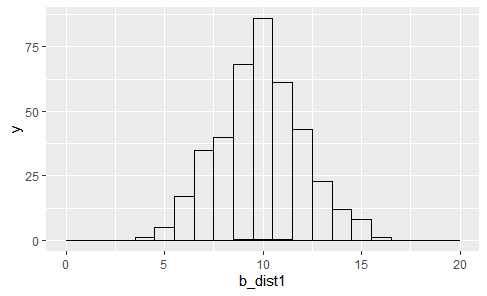


> b\_dist1 = rbinom(400,20,0.5)

> df\_b1 = data.frame(b\_dist1)

> p\_b1 = ggplot(df\_b1) + geom\_histogram(mapping=aes(x=b\_dist1), color="black", fill="NA", binwidth=1) + xlim(0,20) + geom\_function(fun=density, args=list(a=10,s=(5)^(0.5)))

> p\_b1

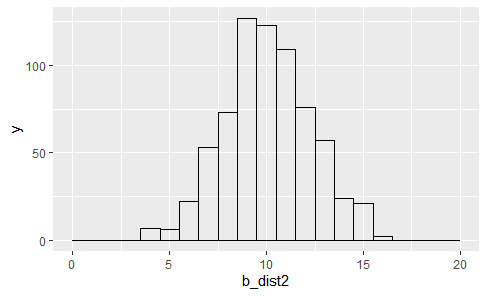


> b\_dist2 = rbinom(700,20,0.5)

> df\_b2 = data.frame(b\_dist2)

> p\_b2 = ggplot(df\_b2) + geom\_histogram(mapping=aes(x=b\_dist2), color="black", fill="NA", binwidth=1) + xlim(0,20) + geom\_function(fun=density, args=list(a=10,s=(5)^(0.5)))

> p\_b2



> b\_dist3 = rbinom(1000,20,0.5)

> df\_b3 = data.frame(b\_dist3)

> p\_b3 = ggplot(df\_b3) + geom\_histogram(mapping=aes(x=b\_dist3), color="black", fill="NA", binwidth=1) + xlim(0,20) + geom\_function(fun=density, args=list(a=10,s=(5)^(0.5)))

> p\_b3



> b\_dist4 = rbinom(5000,20,0.5)

> df\_b4 = data.frame(b\_dist4)

> p\_b4 = ggplot(df\_b4) + geom\_histogram(mapping=aes(x=b\_dist4), color="black", fill="NA", binwidth=1) + xlim(0,20) + geom\_function(fun=density, args=list(a=10,s=(5)^(0.5)))

> p\_b4



> #As we increase the number of samples, the plot for the Binomial Distribution after scaling about the mean starts to look similar to the plot for the Normal Distribution.

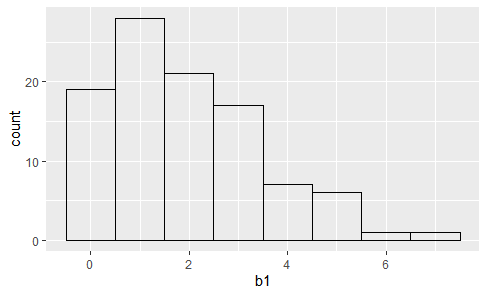
**Solution 2.a)**

> b1 = rbinom(100,2000,0.001)

> dfbinomial = data.frame(b1)

> p1 = ggplot(dfbinomial) + geom\_histogram(mapping=aes(x=b1), color="black", fill="NA", binwidth=1)

> p1



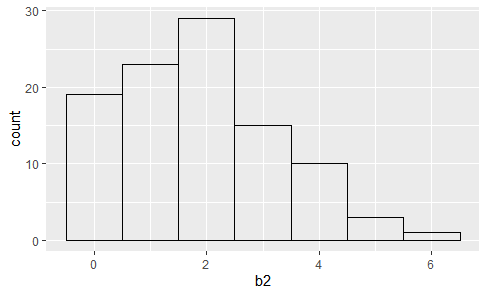
**Solution 2.b)**

> b2 = rpois(100,2)

> dfnormal = data.frame(b2)

> p2 = ggplot(dfbinomial) + geom\_histogram(mapping=aes(x=b2), color="black", fill="NA", binwidth=1)

> p2

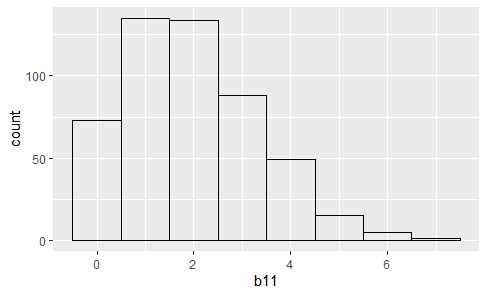


> b11 = rbinom(500,2000,0.001)

> dfbinomial1 = data.frame(b11)

> p11 = ggplot(dfbinomial1) + geom\_histogram(mapping=aes(x=b11), color="black", fill="NA", binwidth=1)

> p11

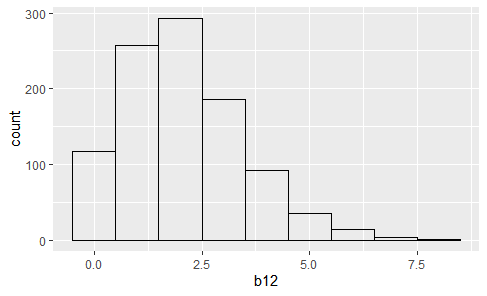


> b12 = rbinom(1000,2000,0.001)

> dfbinomial2 = data.frame(b12)

> p12 = ggplot(dfbinomial2) + geom\_histogram(mapping=aes(x=b12), color="black", fill="NA", binwidth=1)

> p12

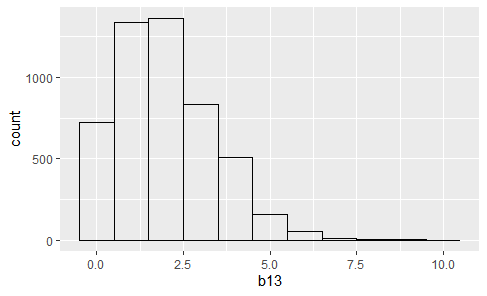


> b13 = rbinom(5000,2000,0.001)

> dfbinomial3 = data.frame(b13)

> p13 = ggplot(dfbinomial3) + geom\_histogram(mapping=aes(x=b13), color="black", fill="NA", binwidth=1)

> p13



> #As we increase the number of samples, the plot for the Binomial Distribution starts to look similar to the plot for the Poisson Distribution since 'p' is very small. It has an increase initially and then a slow decrease.