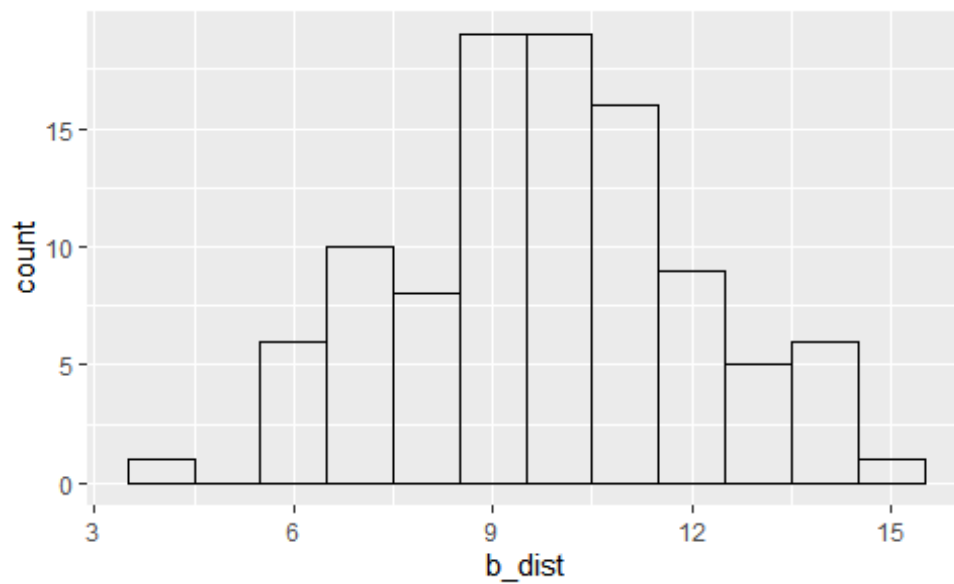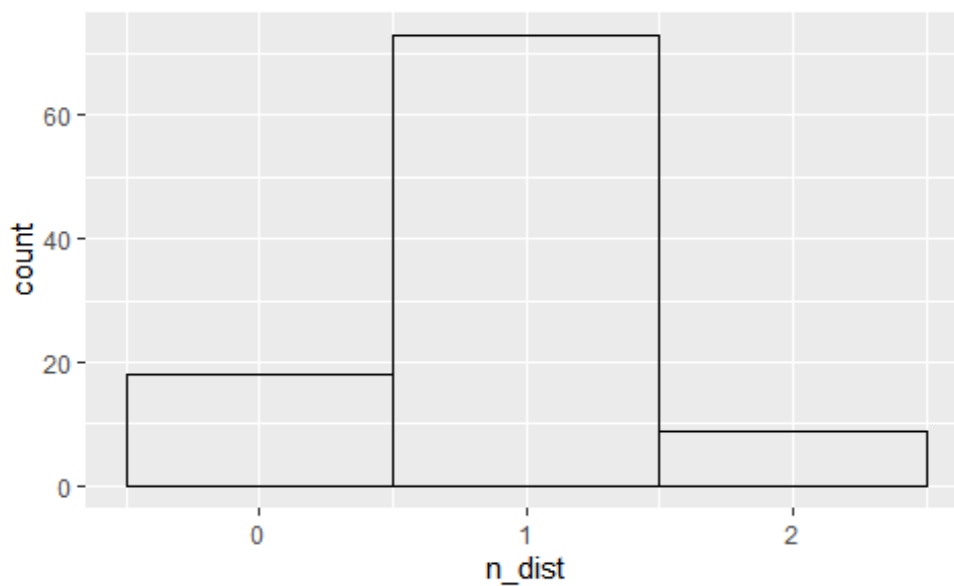# Worksheet 4

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
```
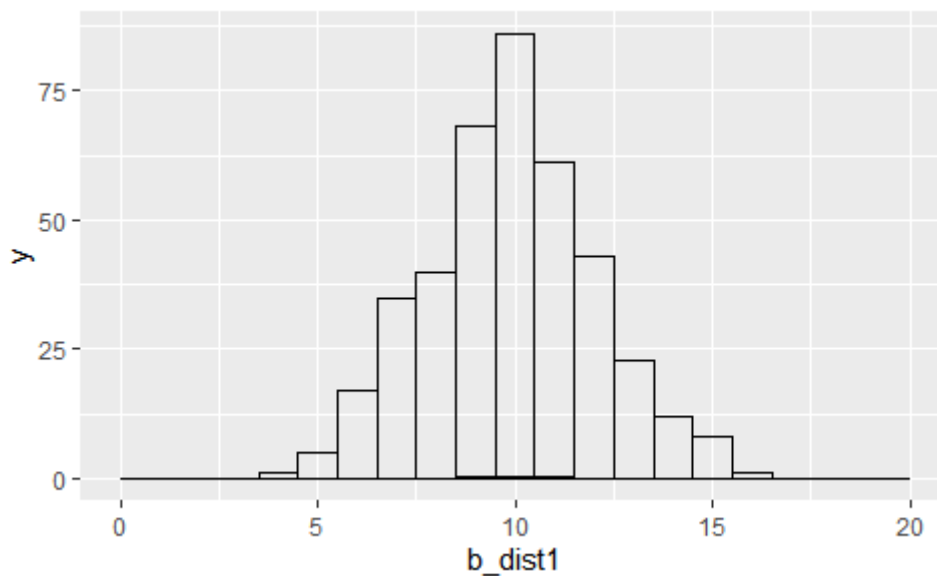

Solution 1.b)

```
> 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)))
```
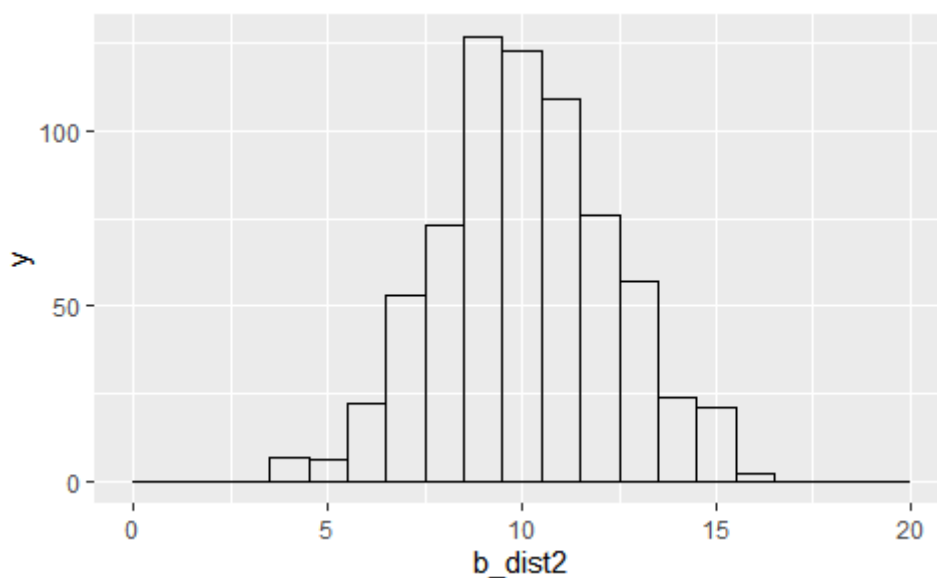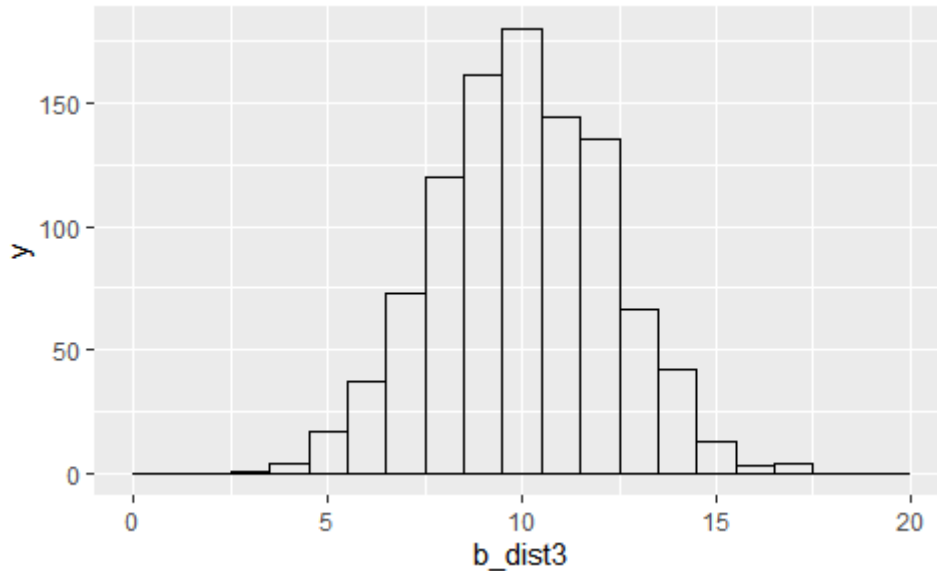
```
> p_b2
```

```
> 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
```
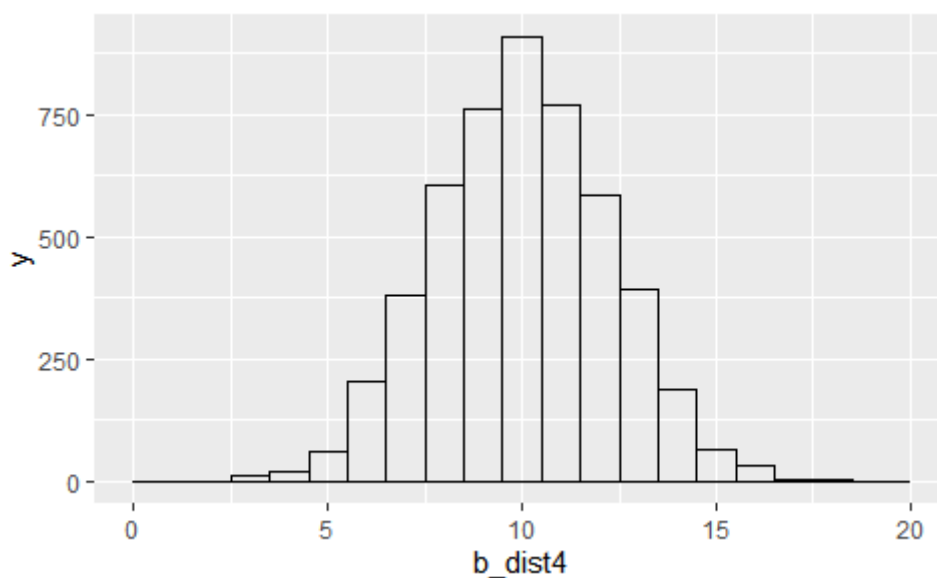


```
> b_dist4 = rbinom(5000,20,0.5)

> 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
```
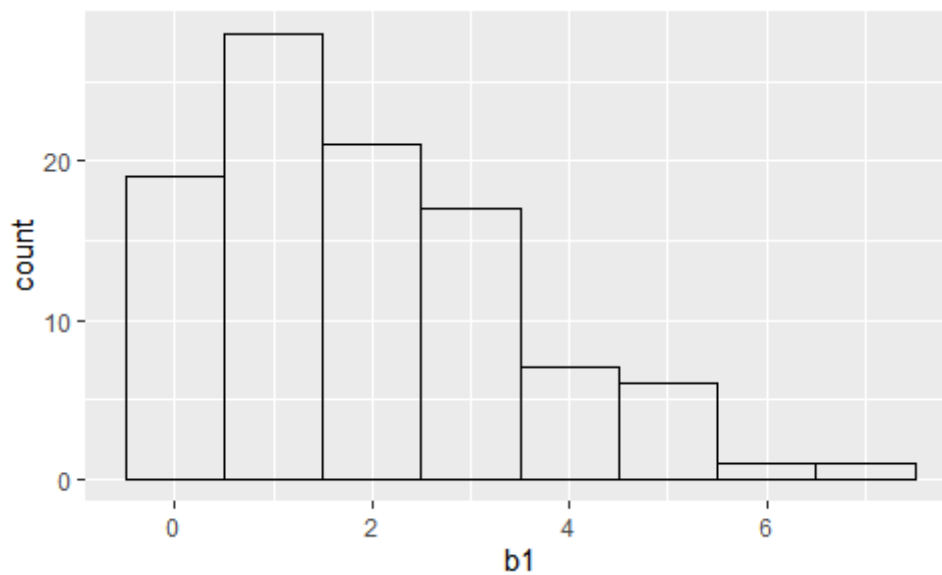


```
> #As we increase the number of samples, the plot for the Binomial Distribution after scaling about the mean starts
to look similar to the plot for the Normal Distribution.
```
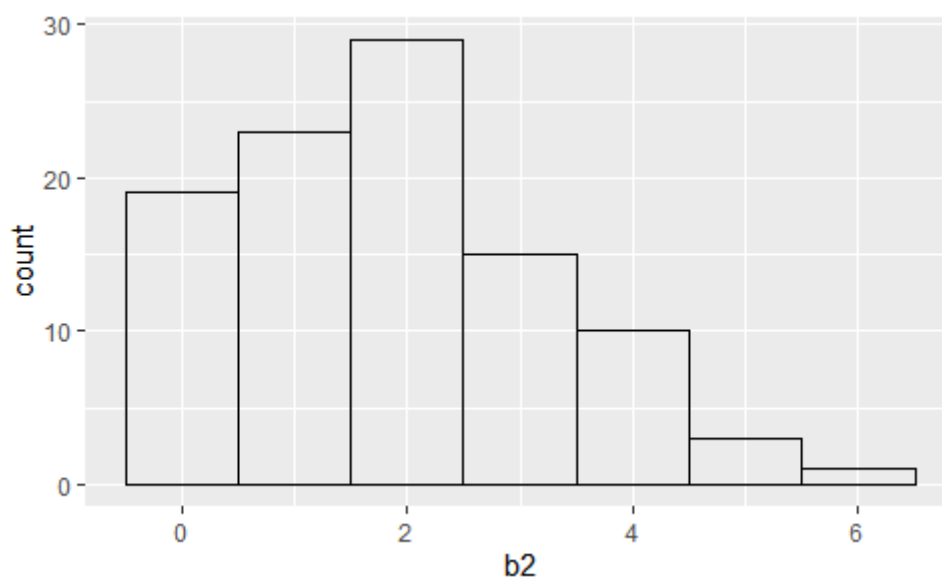
Solution 2.a)

```
> 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
```
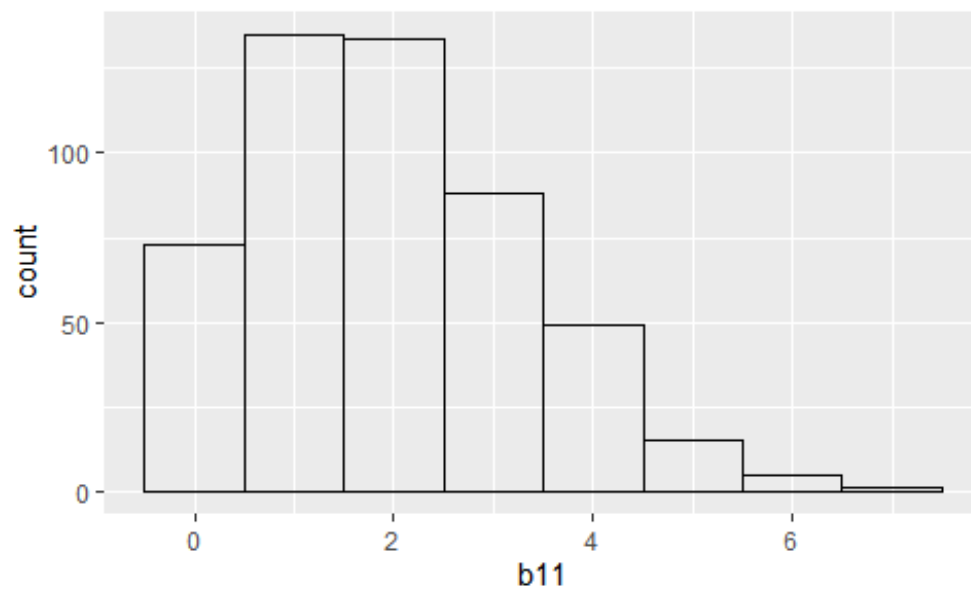

Solution 2.b)

```
> 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
```
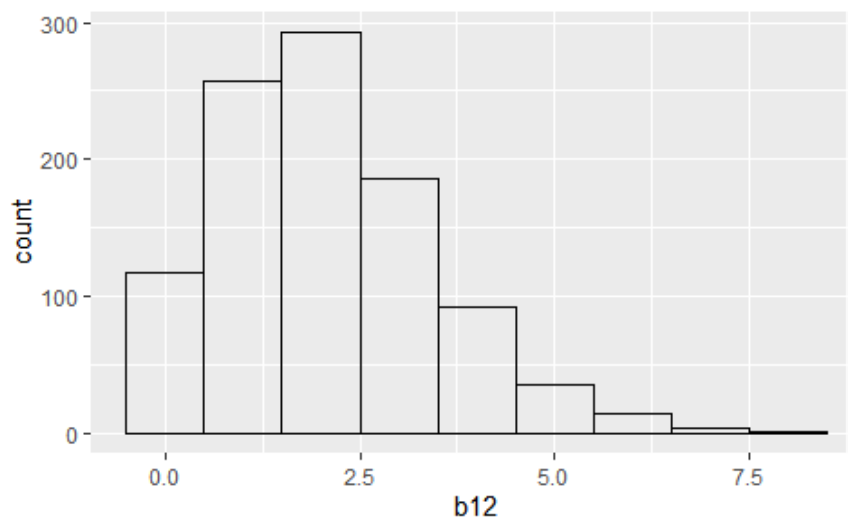
```
> b13 = rbinom(5000,2000,0.001)

> dfbinomial3 = data.frame(b13)

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

> p13
```
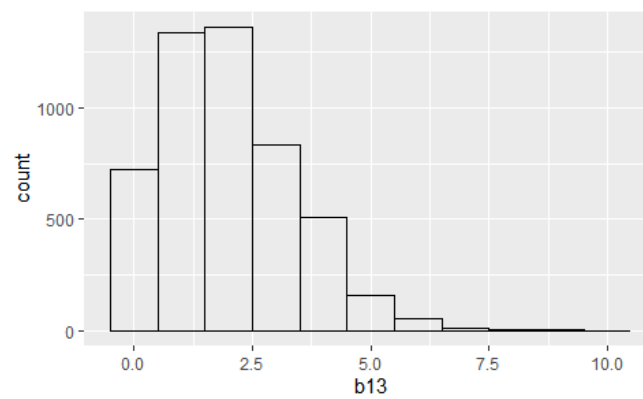


```
> #As we increase the number of samples, the plot for the Binomial Distribution starts to look similar to the plot for
the Poisson Distribution since 'p' is very small. It has an increase initially  and then a slow decrease.
```