Homework02

Problem 1: i)

```r
pbinom(8.25, 20, 0.4)
```

```
## [1] 0.5955987
```

```r
pbinom(8.25, 30, 0.4)
```

```
## [1] 0.09401122
```

```r
pbinom(8.25, 50, 0.4)
```

```
## [1] 0.0002305229
```

```r
pbinom(8.25, 75, 0.4)
```

```
## [1] 1.826106e-08
```

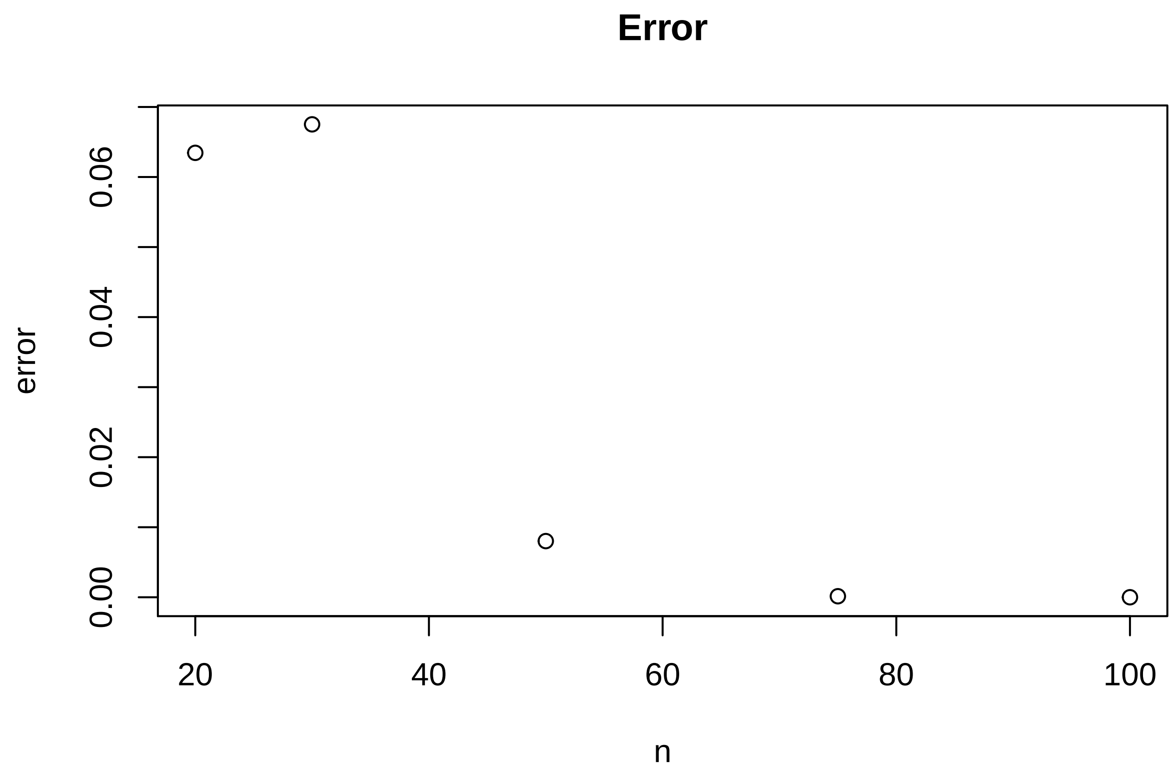
```r
pbinom(8.25, 100, 0.4)
```

```
## [1] 5.431127e-13
```

ii) p = .4; x = 8.25 n = 20; phi(.11411) = 0.53215 n = 30; phi(-1.3975) = 0.16153 n = 50; phi(-3.3919) = 0.00823 n = 75; phi(-5.1265) = 0.0001445 n = 100; phi(-6.48094) = 0.000002295

iii) n = 20; absolute(0.5955987 - 0.53215) = 0.0634487 n = 30; absolute(0.09401122 - 0.16153) = 0.067519 n = 50; absolute(0.0002305229 - 0.00823) = 0.0080195 n = 75; absolute(1.826106e-08 - 0.0001445) = 0.0001444817 n = 100; absolute(5.431127e-13 - 0.000002295) = 0.0000022949995

```r
n = c(20, 30, 50, 75, 100)
error = c(0.0634487, 0.067519, 0.0080195, 0.0001444817, 0.0000022949995)
plot(n, error, main = "Error")
```



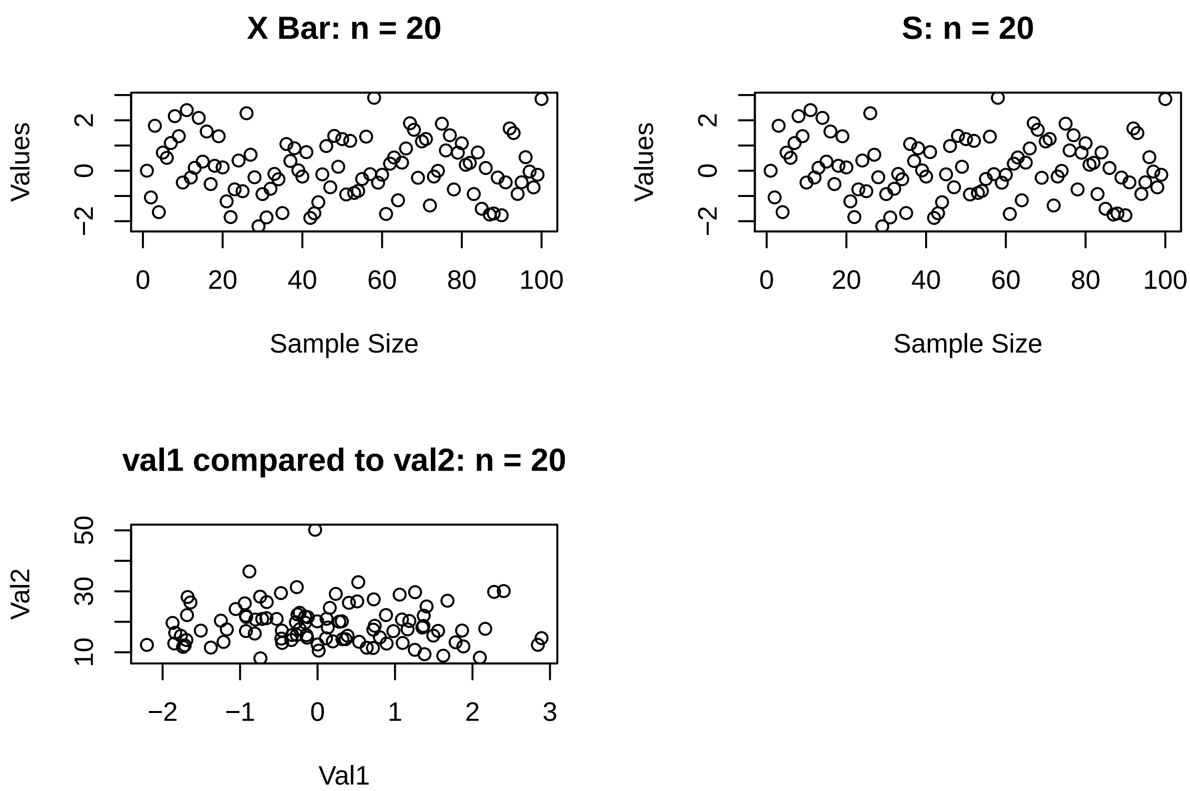
iv) As my n increases, the error approaches the true binomial distribution probability. At n = infinity, the estimation will be the true value

Problem 2:

i)

```
val1 <- 1:100
val2 <- 1:100
for(i in 1:100){
  srs = c(rnorm(20, 2, 3))
  mean = mean(srs)
  variance = var(srs)
  val1[i] = (mean - 2) / (sqrt(9 / 20))
  val2[i] = ((20 - 1)*(variance)) / 9
}
par(mfrow=c(2,2))
plot(val1, xlab = "Sample Size", ylab = "Values", main = "X Bar: n = 20")
plot(val1, xlab = "Sample Size", ylab = "Values", main = "S: n = 20")
plot(val1, val2, xlab = "Val1", ylab = "Val2", main = "val1 compared to val2: n = 20")
cor(val1, val2)
```

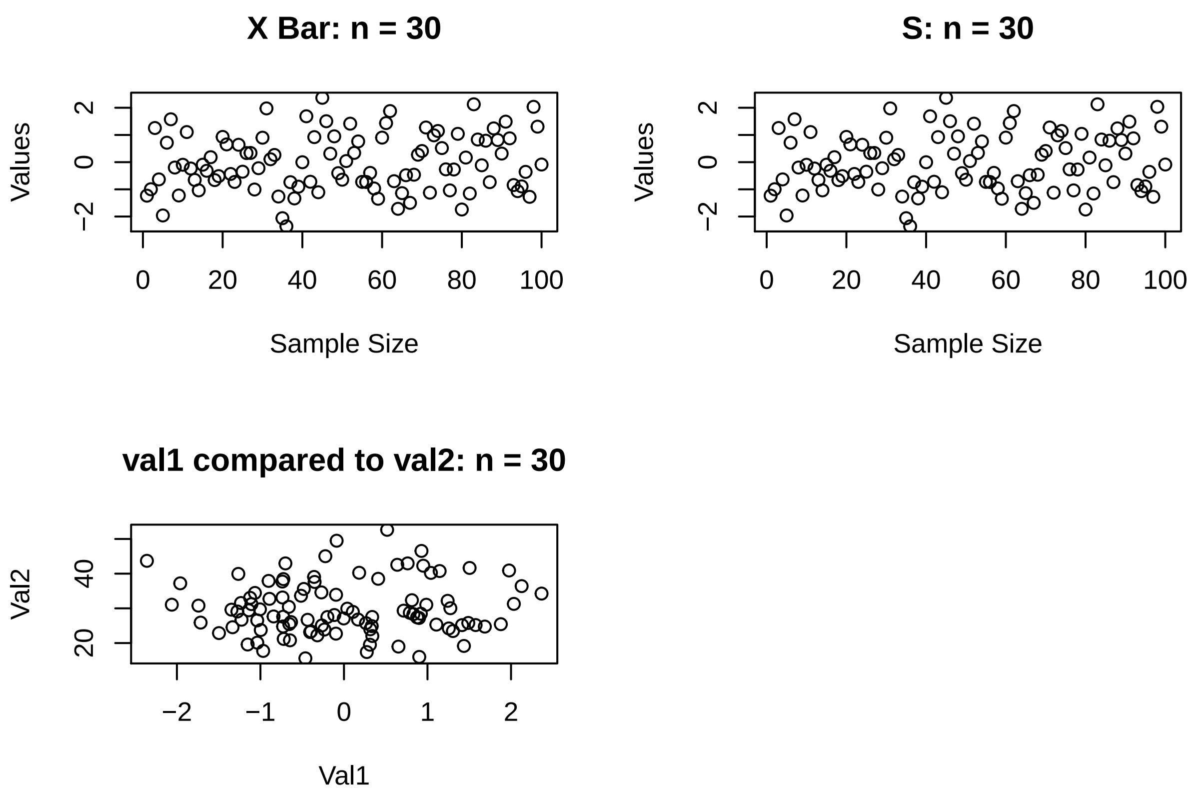
```
## [1] -0.03373181
```



ii)

```
val1 <- 1:100
val2 <- 1:100
for(i in 1:100){
  srs = c(rnorm(30, 2, 3))
  mean = mean(srs)
  variance = var(srs)
  val1[i] = (mean - 2) / (sqrt(9 / 30))
  val2[i] = ((30 - 1)*(variance)) / 9
}
par(mfrow=c(2,2))
plot(val1, xlab = "Sample Size", ylab = "Values", main = "X Bar: n = 30")
plot(val1, xlab = "Sample Size", ylab = "Values", main = "S: n = 30")
plot(val1, val2, xlab = "Val1", ylab = "Val2", main = "val1 compared to val2: n = 30")
cor(val1, val2)
```

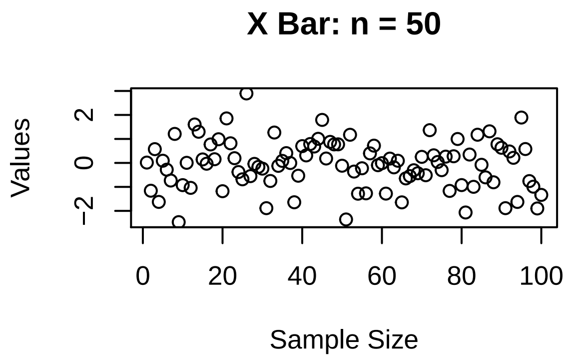
```
## [1] 0.04209735
```

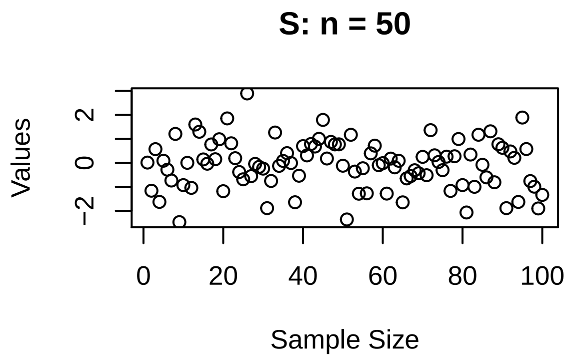


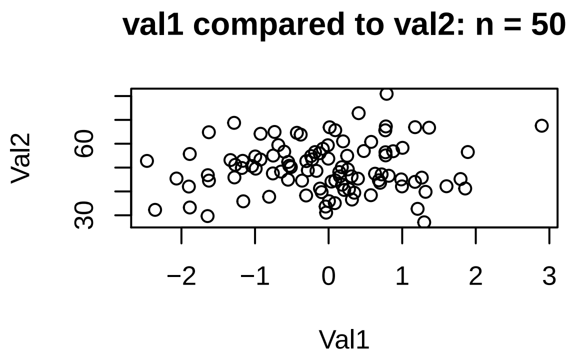
iii)

```
val1 <- 1:100
val2 <- 1:100
for(i in 1:100){
  srs = c(rnorm(50, 2, 3))
  mean = mean(srs)
  variance = var(srs)
  val1[i] = (mean - 2) / (sqrt(9 / 50))
  val2[i] = ((50 - 1)*(variance)) / 9
}
par(mfrow=c(2,2))
plot(val1, xlab = "Sample Size", ylab = "Values", main = "X Bar: n = 50")
plot(val1, xlab = "Sample Size", ylab = "Values", main = "S: n = 50")
plot(val1, val2, xlab = "Val1", ylab = "Val2", main = "val1 compared to val2: n = 50")
cor(val1, val2)
```

```
## [1] 0.08332031
```



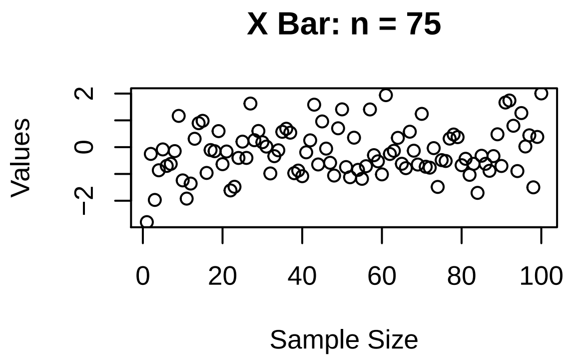


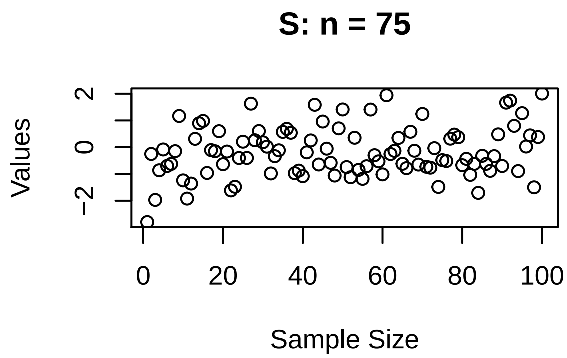


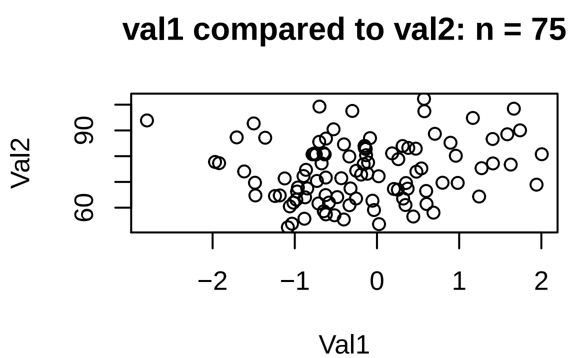
iv)

```
val1 <- 1:100
val2 <- 1:100
for(i in 1:100){
  srs = c(rnorm(75, 2, 3))
  mean = mean(srs)
  variance = var(srs)
  val1[i] = (mean - 2) / (sqrt(9 / 75))
  val2[i] = ((75 - 1)*(variance)) / 9
}
par(mfrow=c(2,2))
plot(val1, xlab = "Sample Size", ylab = "Values", main = "X Bar: n = 75")
plot(val1, xlab = "Sample Size", ylab = "Values", main = "S: n = 75")
plot(val1, val2, xlab = "Val1", ylab = "Val2", main = "val1 compared to val2: n = 75")
cor(val1, val2)
```

```
## [1] 0.1626165
```







v) There appears to be an average value of 0, and data points ranging from approximately -2.5 to 2.5. The mean is between 500 and 600 for ((n-1)(sˆ2))/9. As the sample size increases, the distribution appears to be more normal.

vi) The correlation coefficients are 0.002634042 (n = 20), 0.04597207 (n = 30), -0.04646729 (n = 50), 0.1393645 (n = 75). So basically independant.