

# Untitled

Getong Zhong

2023-01-30

1

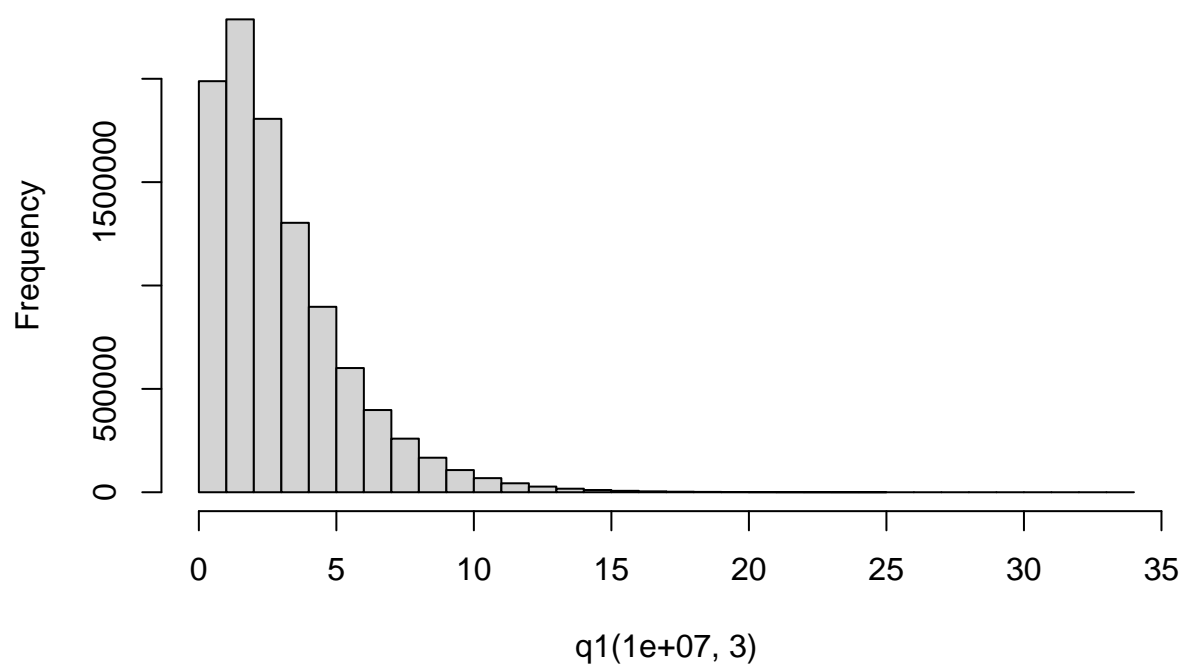
```
df <- c(1)
q1<-
  function (n,k){
    df <- c(1)
    for (b in seq(n)){
      a <- 0
      for( x in seq(k)){
        a = a + (rnorm(1))^2
      }
      df[b] = a
    }
    return(df)
  }
```

2

By function in #1

```
set.seed(1001)
hist(q1(1000000,3), main="Histogram of function from #1")
```

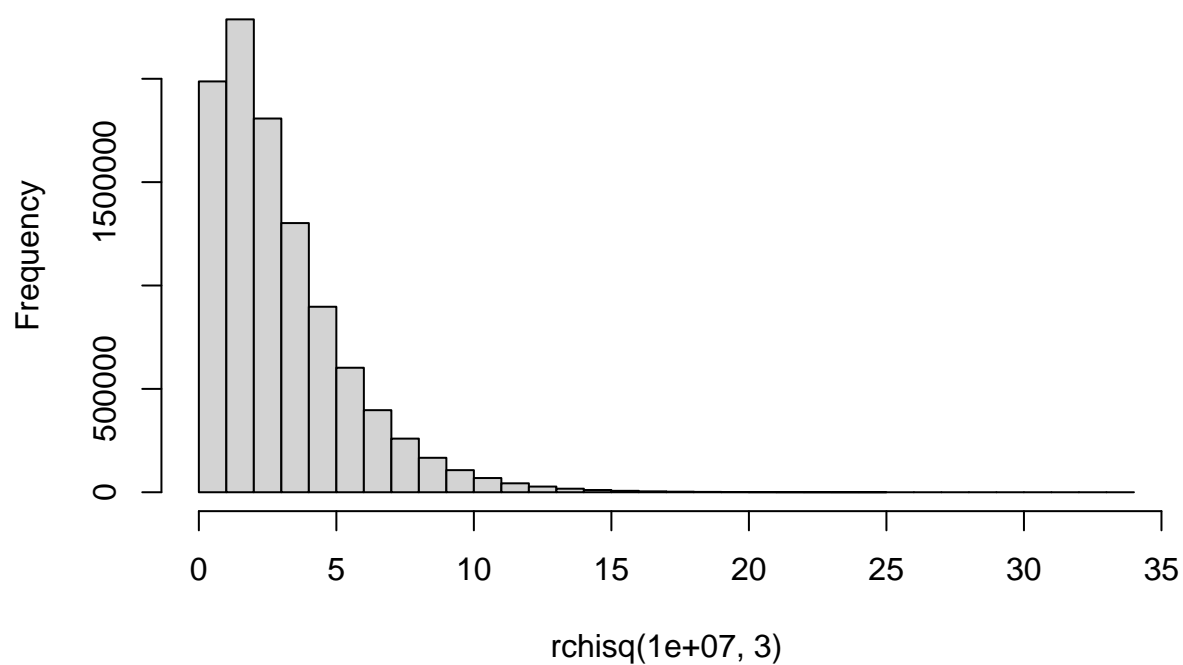
## Histogram of function from #1



By function `rchisq()`

```
set.seed(8289)
hist(rchisq(10000000,3), main="Histogram of rchisq()")
```

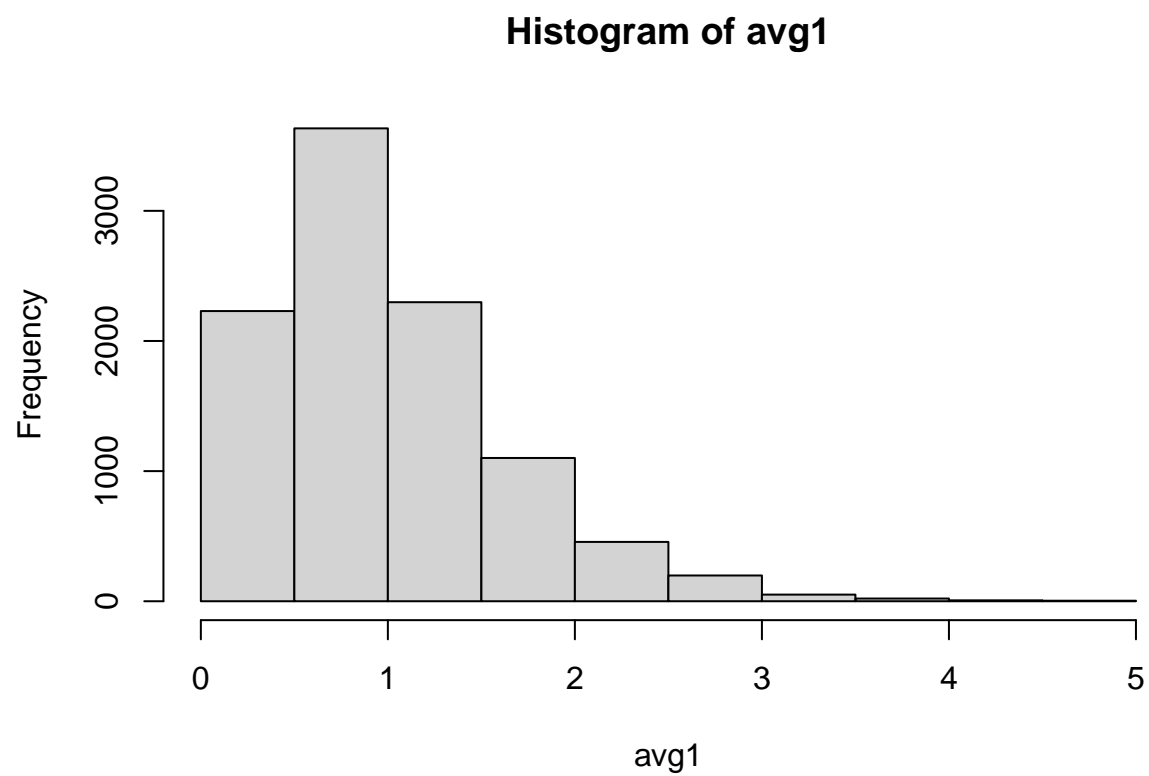
## Histogram of rchisq()



3

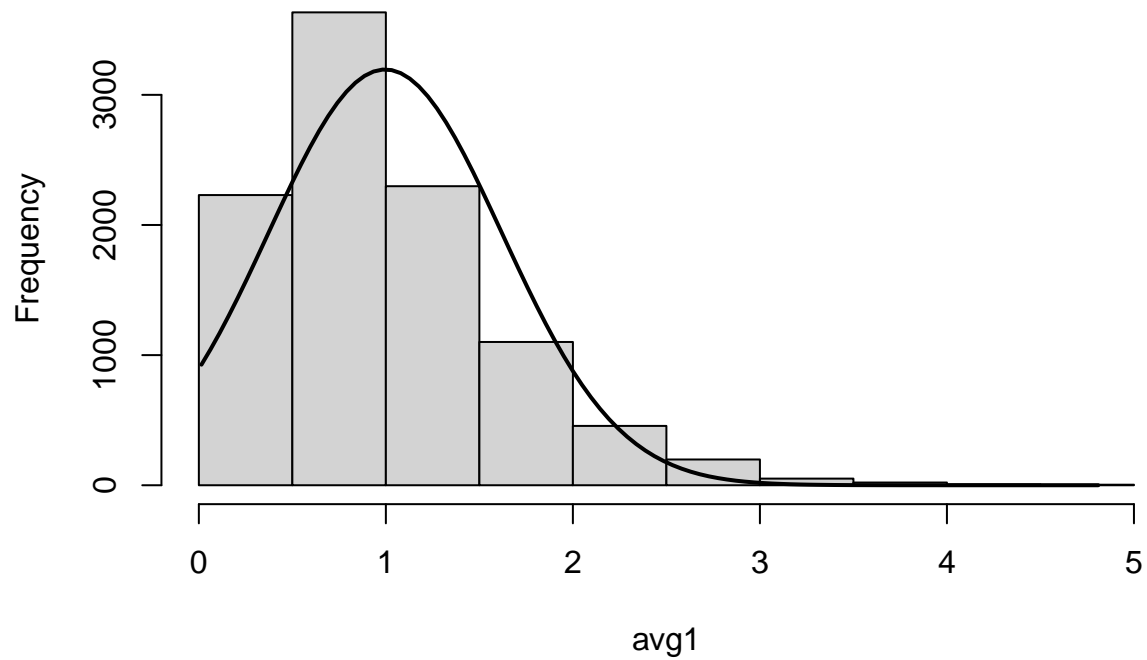
n = 5

```
set.seed(8289)
avg1=c(0)
for(x in seq(10000)){
  avg1[x] = mean(rchisq(5,1))
}
x2 <- seq(min(avg1), max(avg1), length = 100)
fun <- dnorm(x2, mean = mean(avg1), sd = sd(avg1))
y2<-fun* diff(hist(avg1)$mids[1:2])*length(avg1)
```



```
hist(avg1, main="Mean of Chi-sqr with n = 20")  
lines(x2, y2, lwd = 2)
```

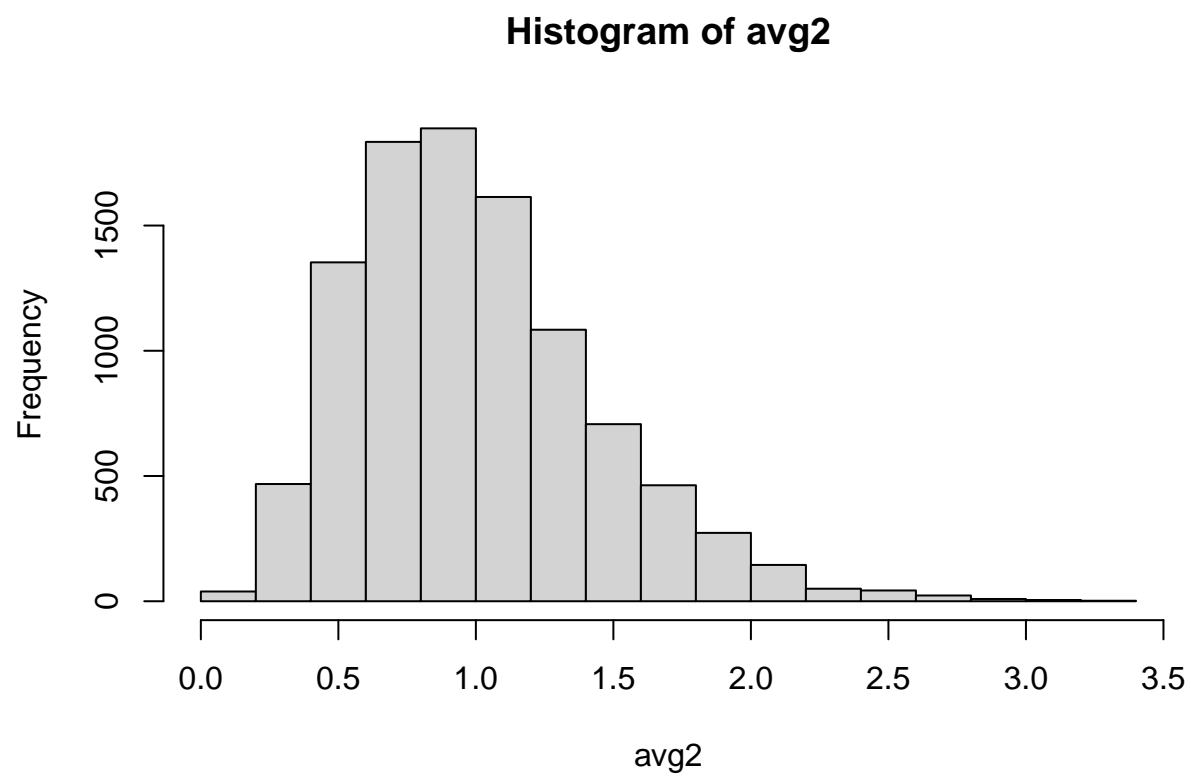
## Mean of Chi-sqr with n = 20



n = 10

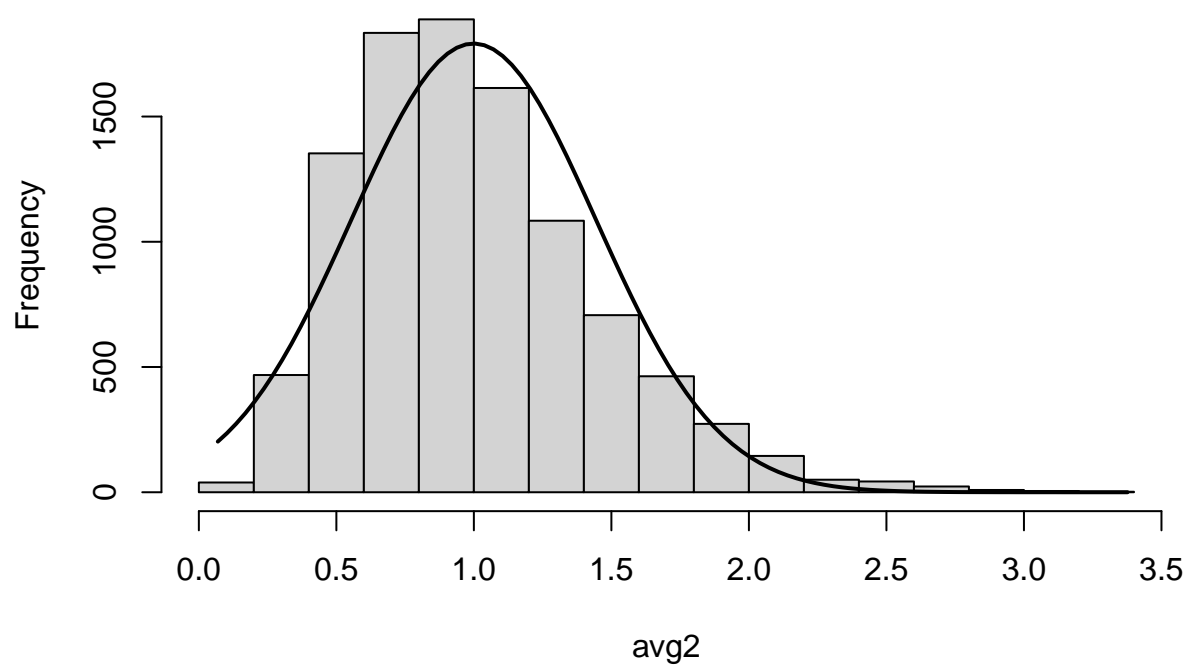
```
set.seed(8289)
avg2=c(0)
for(x in seq(10000)){
  avg2[x] = mean(rchisq(10,1))
}

x2 <- seq(min(avg2), max(avg2), length = 100)
fun <- dnorm(x2, mean = mean(avg2), sd = sd(avg2))
y2<-fun* diff(hist(avg2)$mids[1:2])*length(avg2)
```



```
hist(avg2, main="Mean of Chi-sqr with n = 20")  
lines(x2, y2, lwd = 2)
```

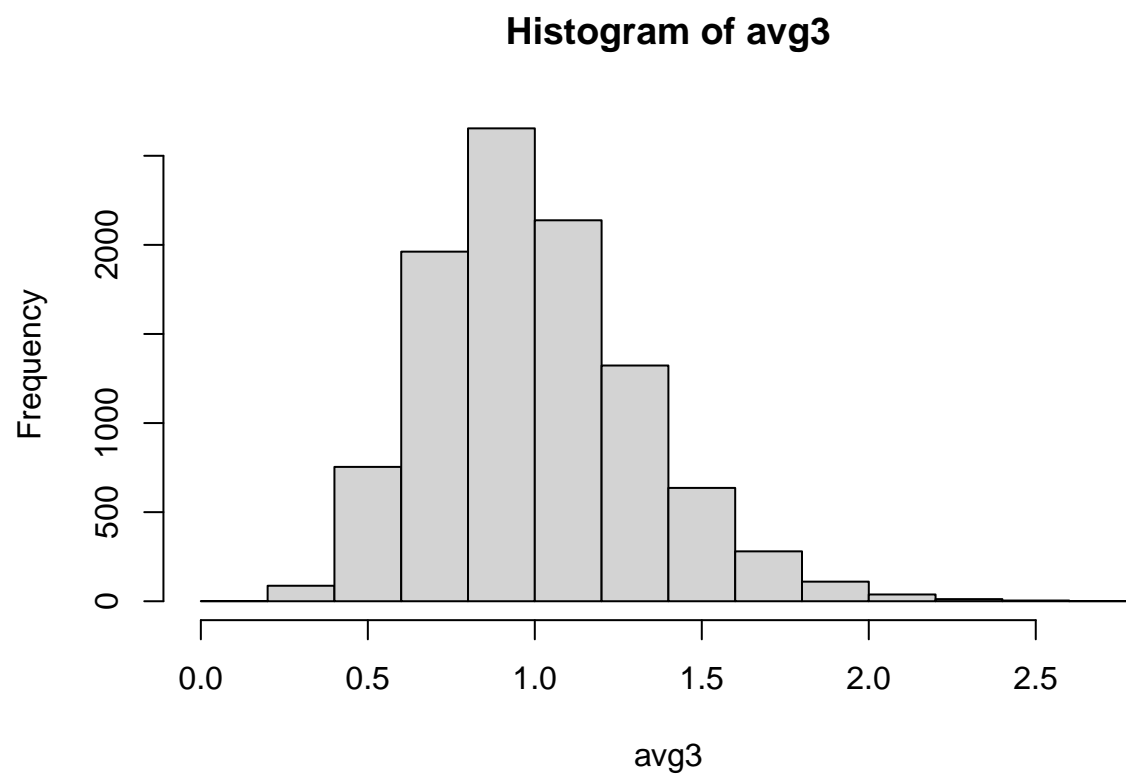
### Mean of Chi-sqr with n = 20



n = 20

```
set.seed(8289)
avg3=c(0)
for(x in seq(10000)){
  avg3[x] = mean(rchisq(20,1))
}

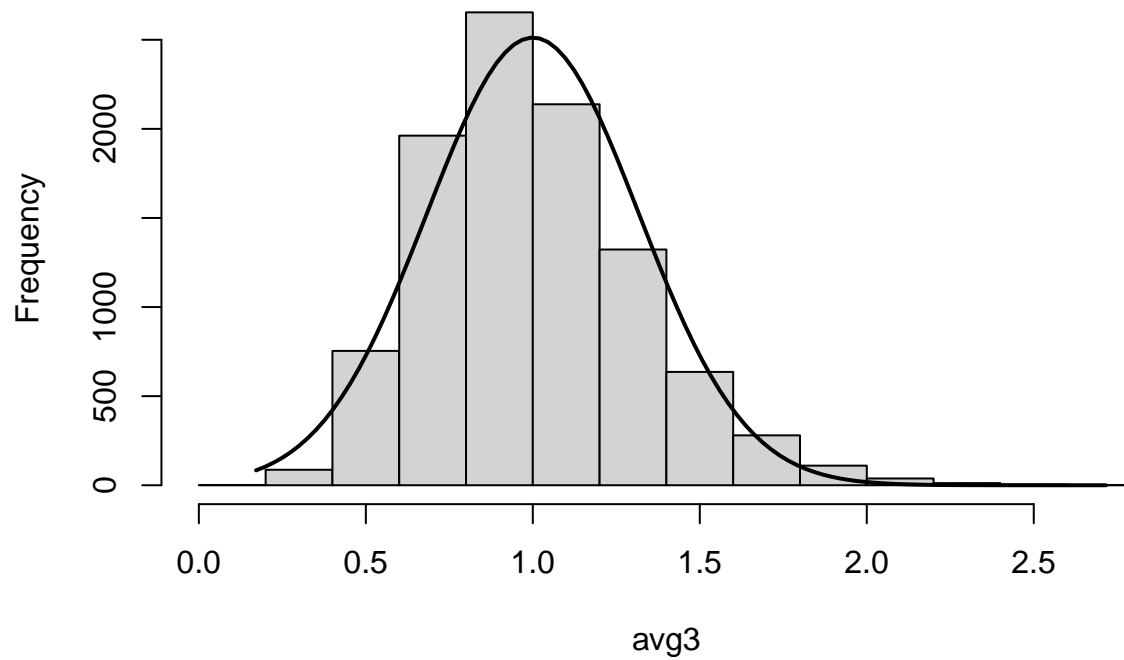
x2 <- seq(min(avg3), max(avg3), length = 100)
fun <- dnorm(x2, mean = mean(avg3), sd = sd(avg3))
y2<-fun* diff(hist(avg3)$mids[1:2])*length(avg3)
```



```
hist(avg3, main="Mean of Chi-sqr with n = 20")  
lines(x2, y2, lwd = 2)
```



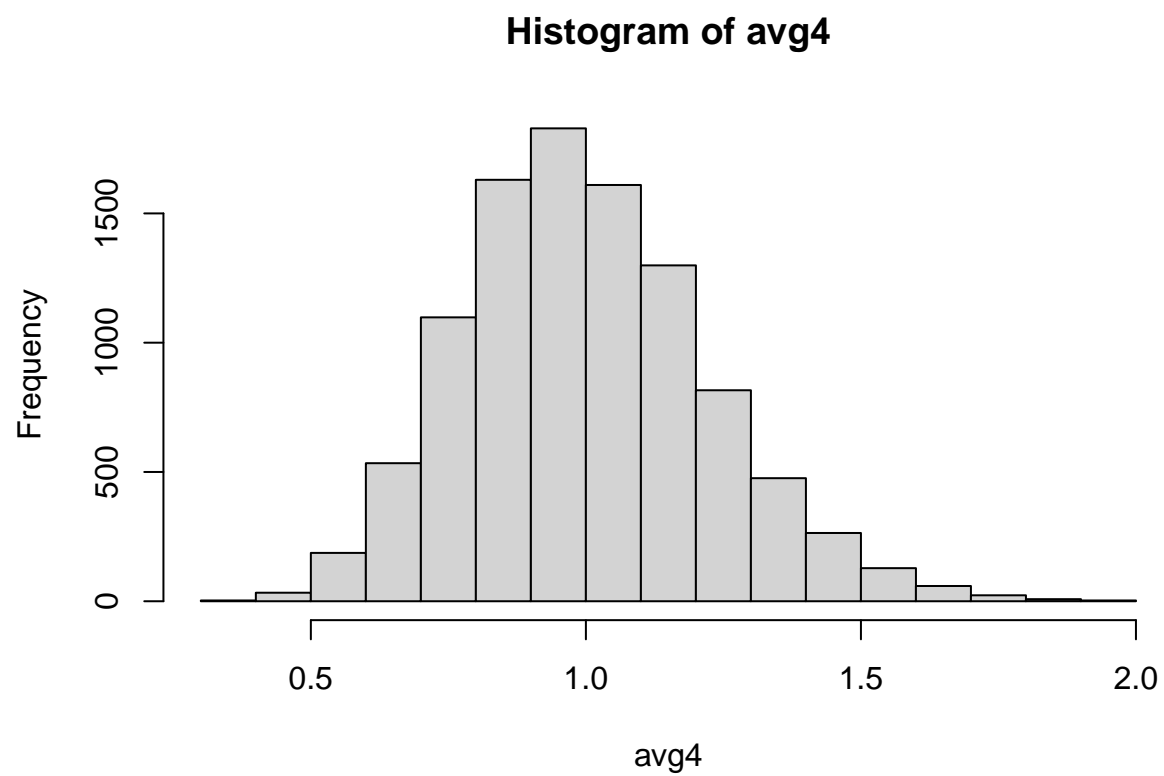
### Mean of Chi-sqr with n = 20



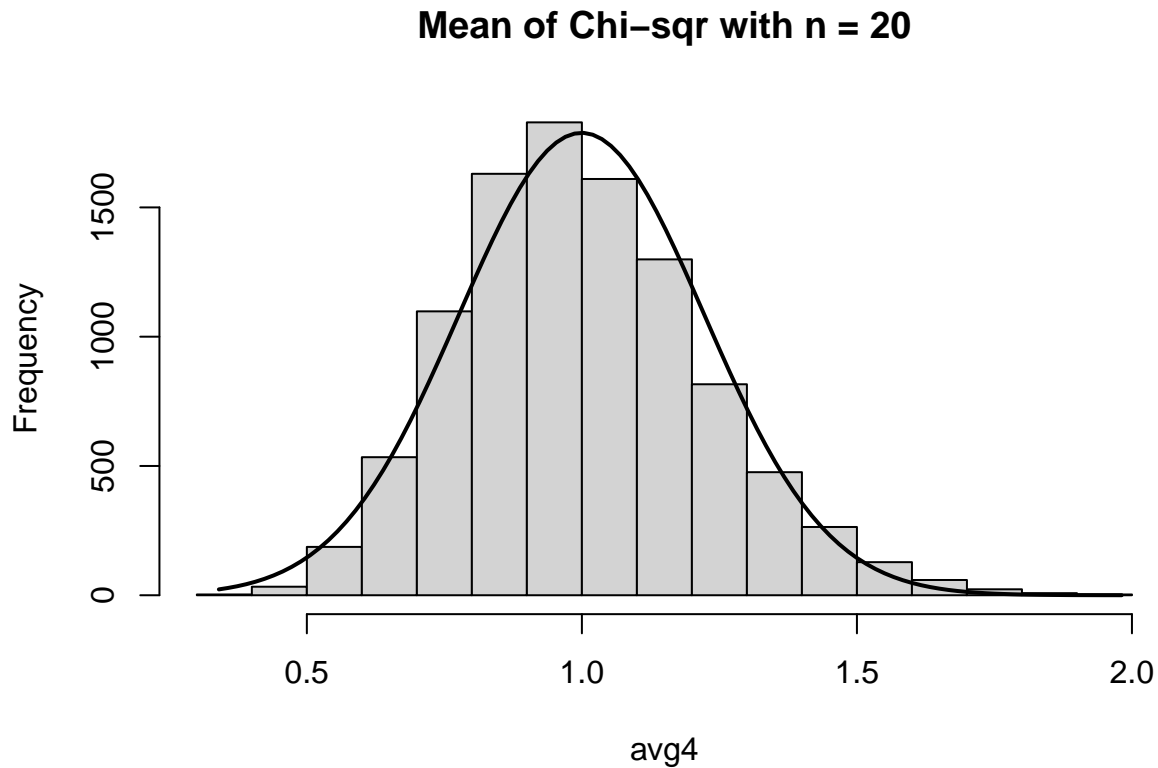
n = 40

```
set.seed(8289)
avg4=c(0)
for(x in seq(10000)){
  avg4[x] = mean(rchisq(40,1))
}

x2 <- seq(min(avg4), max(avg4), length = 100)
fun <- dnorm(x2, mean = mean(avg4), sd = sd(avg4))
y2<-fun* diff(hist(avg4)$mids[1:2])*length(avg4)
```



```
hist(avg4, main="Mean of Chi-sqr with n = 20")  
lines(x2, y2, lwd = 2)
```



## Discussion

According to the LLNs, the sample average converges as the sample size  $n$  approaches to infinity (the sample average almost converges surely). From the above progression we can tell that the histogram of sample mean for Chi-square distribution changed from very left-skewed to more and more central as sample size ( $n$ ) increased from 5 to 40.

## 4

```
sig <- matrix(c(13.6, -14.5, -103.4, -14.5, 65.2, 154, -103.4, 154, 2702), 3, 3, byrow = TRUE)
error <- rnorm(1000, 0, 0.75^2)
mvn_gen <- function(n, mu, sigma, factorization = "Cholesky") {
  d <- length(mu)
  Z <- matrix(rnorm(n*d), nrow = n, ncol = d)
  if(factorization == "Cholesky") { Q <- chol(sigma) } else {
    if(factorization == "Spectral") {
      ev <- eigen(sigma)
      lambda <- ev$values
      P <- ev$vectors
      Q <- P %*% diag(sqrt(lambda)) %*% t(P) } else {
      stop("Arg factorization must be 'Cholesky' or 'Spectral'.")
    }
  }
  mu <- matrix(mu, nrow = d, ncol = 1)
```

```

  J = matrix(1, nrow = n, ncol = 1)
  X <- Z %*% Q + J %*% t(mu)
  return(data.frame(X))
}
set.seed(1211)

mv <- mvn_gen(n = 1000, mu = c(0.5, 0.3, 10), sigma = sig, factorization = "Spectral")
Y = 71 - 0.28 *mv$X1 + 0.05*mv$X2 - 0.007*mv$X3+error
data2 <- cbind(mv,Y)

model<-lm(Y ~ X1 +X2 + X3 ,data = data2)

model%>%
  tidy() %>%
  kable()

```

term	estimate	std.error	statistic	p.value
(Intercept)	71.0014400	0.0192700	3684.55010	0
X1	-0.2790970	0.0065873	-42.36882	0
X2	0.0517965	0.0026530	19.52409	0
X3	-0.0069816	0.0004300	-16.23521	0