Appendix B Probability Theory

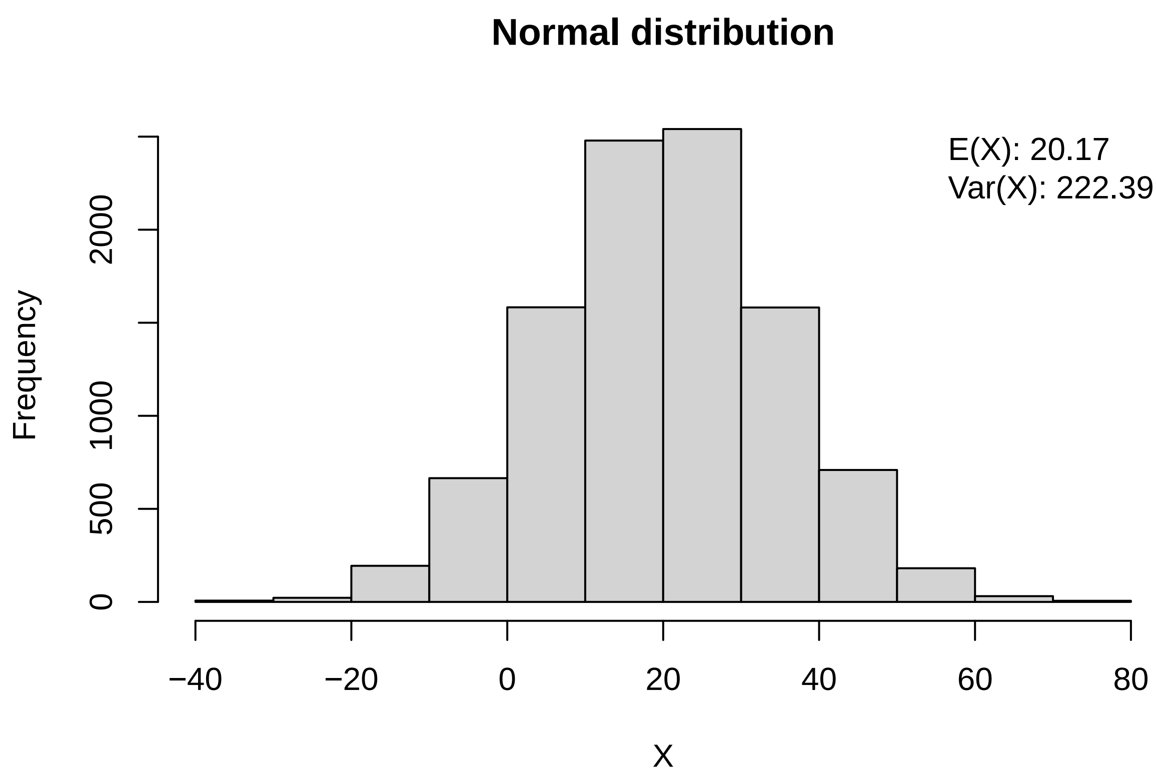
Susumu Shikano

Last compiled at 13. Juli 2022

Generate random numbers by using a normal distribution

We draw 10000 random numbers from a normal distribution whose mean is 20 and standard deviation is 15 (i.e. variance is $15^2 = 225$). The random numbers constitute a random variable which we denote by X.

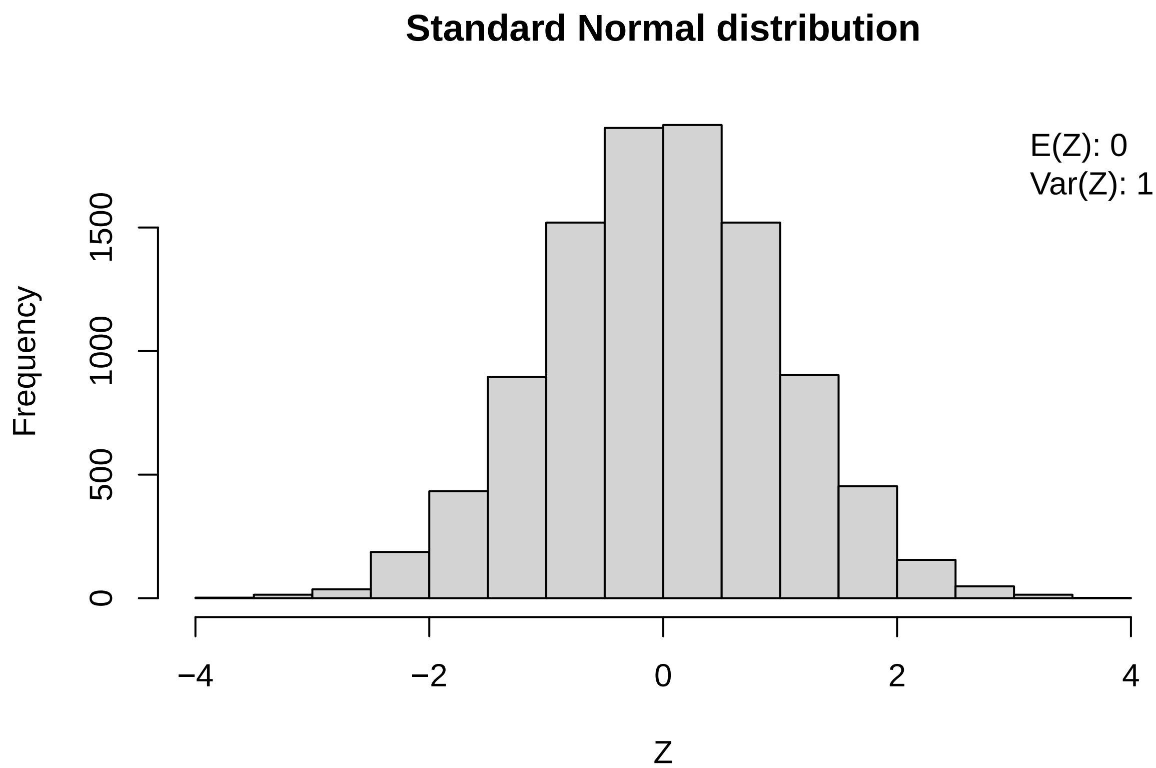
```
X <- rnorm(10000,mean=20,sd=15)
hist(X,main="Normal distribution")
legend("topright",c(paste("E(X):",round(mean(X),2)),
                    paste("Var(X):",round(mean(X^2)-mean(X)^2,2))),
       bty="n")
```



Z-Transformation

We subtract from X its mean and subsequently divide by its standard debiation. We call this transformation z-transformation and dnote the new variable by Z. Z is the standard normal distribution.

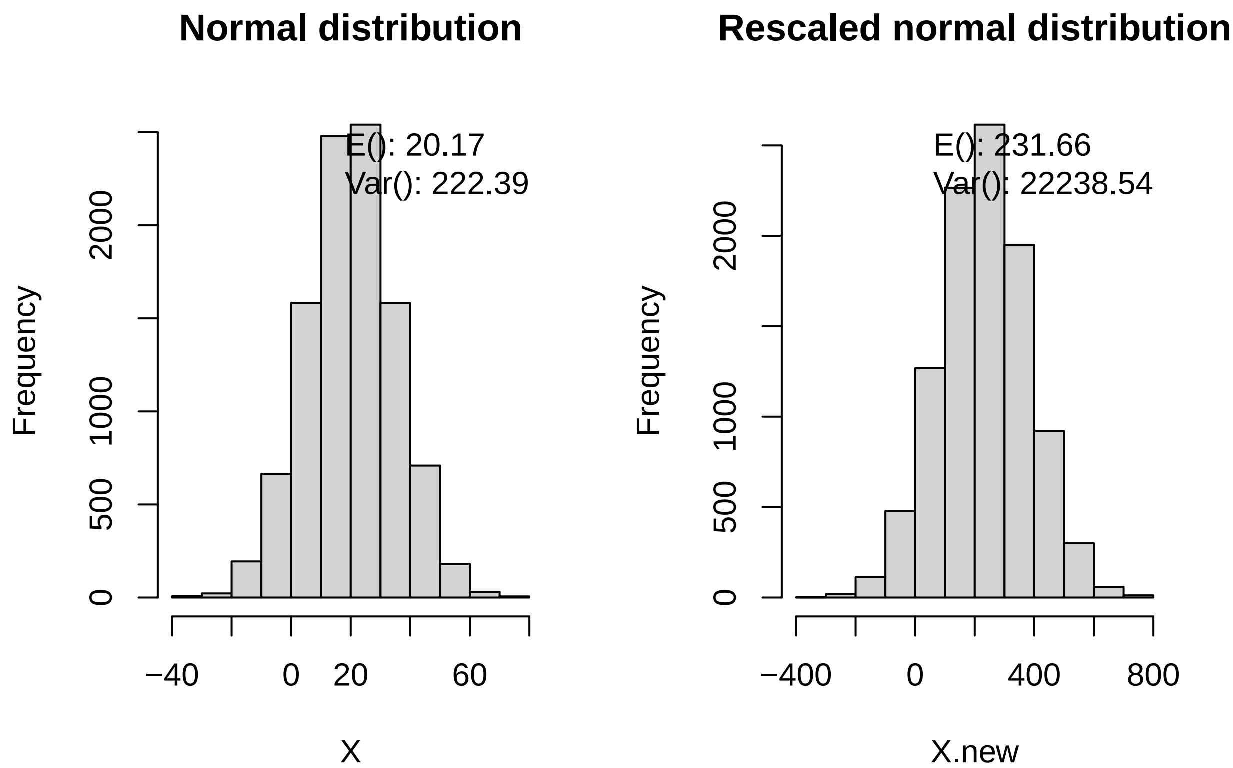
```r
Z <- (X - mean(X))/sqrt(mean(X^2)-mean(X)^2)
hist(Z,main="Standard Normal distribution")
legend("topright",c(paste("E(Z):",round(mean(Z),2)),
                    paste("Var(Z):",round(mean(Z^2)-mean(Z)^2,2))),
       bty="n")
```



Rescaling a normal distribution

We can divide/multiply and/or add/subtract an arbitrary constant value. Here, we construct a new random variably by multiplying X by 10 and subsequently add 30.

```r
X.new <- X * 10 + 30
par(mfrow=c(1,2))
hist(X,main="Normal distribution")
legend("topright",c(paste("E():",round(mean(X),2)),
                    paste("Var():",round(mean(X^2)-mean(X)^2,2))),
       bty="n")
hist(X.new,main="Rescaled normal distribution")
legend("topright",c(paste("E():",round(mean(X.new),2)),
                    paste("Var():",round(mean(X.new^2)-mean(X.new)^2,2))),
       bty="n")
```

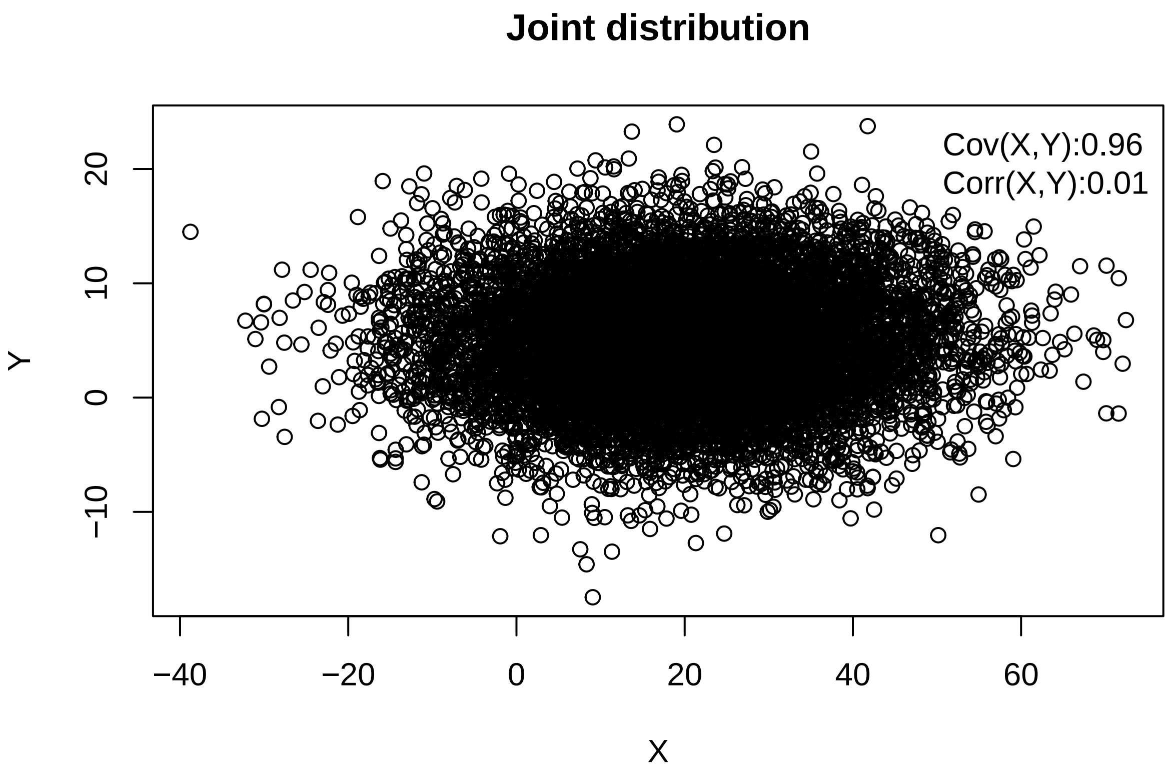


It is easily calculated What expected value and variance the new random variable should have (see Appsndix B).

Joint distribution of X and Y

You can generate another random variable Y from a normal distribution and observe its joint distribution with X.

```
Y <- rnorm(10000,mean=5,sd=5)
plot(X,Y,main="Joint distribution")
legend("topright",c(paste0("Cov(X,Y):",round(cov(X,Y),2)),
                    paste0("Corr(X,Y):",round(cor(X,Y),2))),
       bty="n")
```

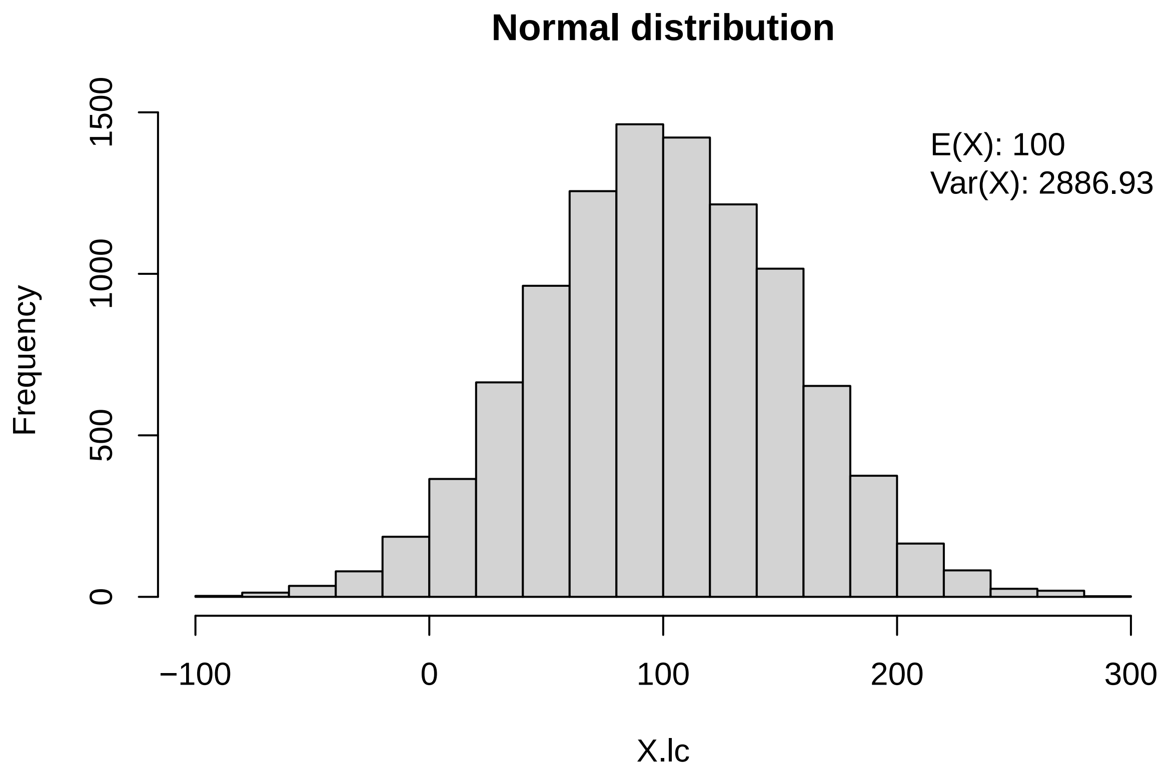


Since it is a result of random draws, the covariance and correlation of X and Y are nonzero. although X and Y are drawn independently. However, if you draw more random numbers they converge towards zero.

Linear combination of iid normal random variables

See Appendix B.

```
X2 <- rnorm(10000,mean=20,sd=15)
X.lc <- 3*X + 2*X2
hist(X.lc,main="Normal distribution")
legend("topright",c(paste("E(X):",round(mean(X.lc),2)),
                    paste("Var(X):",round(mean(X.lc^2)-mean(X.lc)^2,2))),
       bty="n")
```



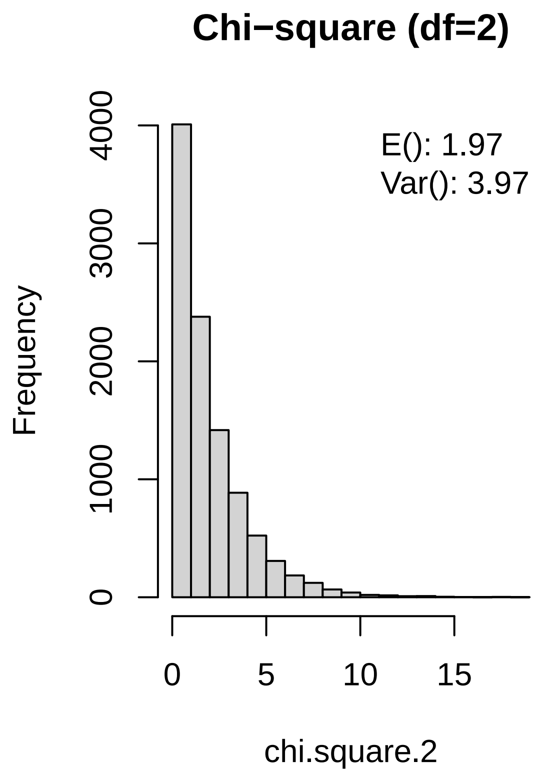
Chi-square distribution

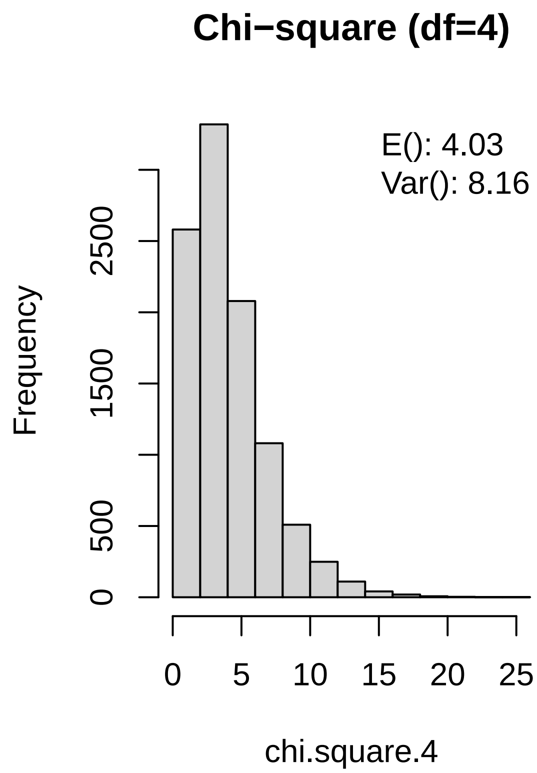
If one adds squared random values drawn from the standard normal distribution, the resulting random variable follows a chi-square distribution. Its degrees of freedom is determined by how many random variables are added. In the exapmles below, they are 2 and 6.

```
Z.square.1 <- rnorm(10000,mean=0,sd=1)^2
Z.square.2 <- rnorm(10000,mean=0,sd=1)^2
Z.square.3 <- rnorm(10000,mean=0,sd=1)^2
Z.square.4 <- rnorm(10000,mean=0,sd=1)^2
Z.square.5 <- rnorm(10000,mean=0,sd=1)^2
Z.square.6 <- rnorm(10000,mean=0,sd=1)^2

chi.square.2 <- Z.square.1 + Z.square.2
chi.square.4 <- Z.square.3 + Z.square.4 + Z.square.5 + Z.square.6

par(mfrow=c(1,2))
hist(chi.square.2,main="Chi-square (df=2)")
legend("topright",c(paste("E():",round(mean(chi.square.2),2)),
                    paste("Var():",round(mean(chi.square.2^2)-mean(chi.square.2)^2,2))),
       bty="n")
hist(chi.square.4,main="Chi-square (df=4)")
legend("topright",c(paste("E():",round(mean(chi.square.4),2)),
                    paste("Var():",round(mean(chi.square.4^2)-mean(chi.square.4)^2,2))),
       bty="n")
```



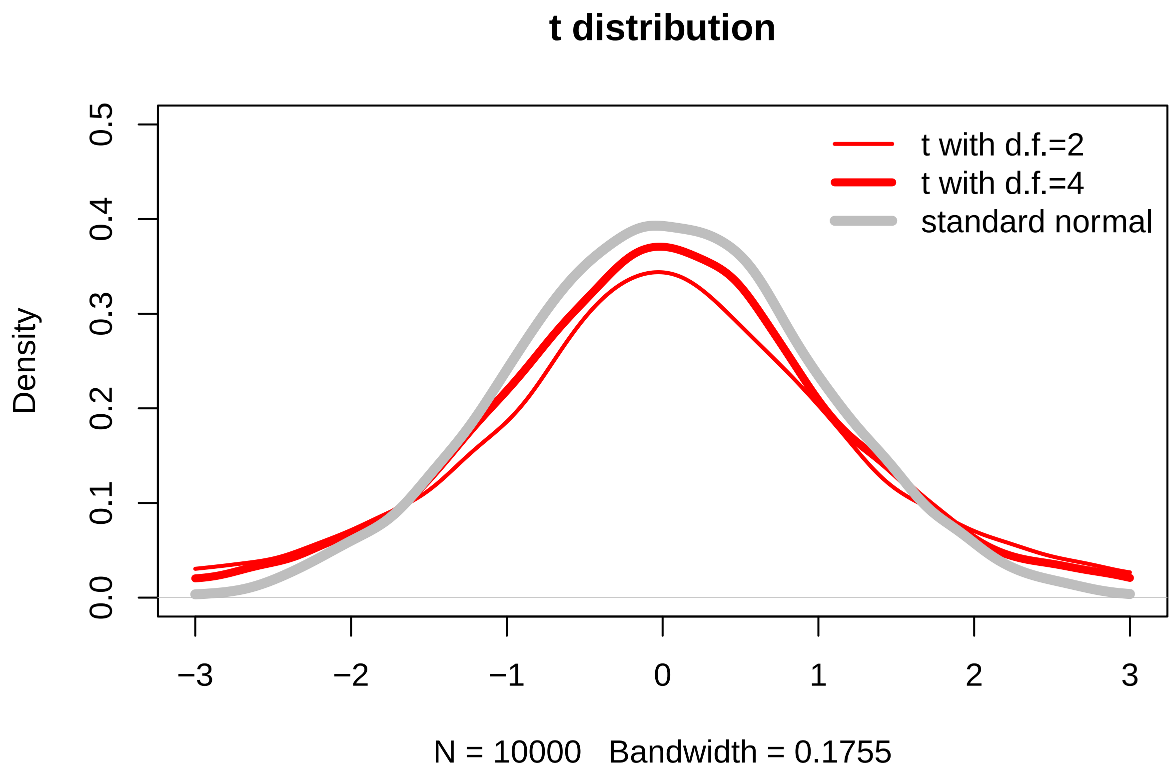


t distribution

A t-Distributed random variable can be generated through dividing a standard normal random variable by the square root of a chi-squared random variable divided by its degrees of freedom. Thus, the t-distributions have certain degrees of freedom, as well.

```
t.2 <- Z/sqrt(chi.square.2/2)
t.4 <- Z/sqrt(chi.square.4/4)

plot(density(t.2,from=-3,to=3),main="t distribution",ylim=c(0,0.5),lwd=2,col="red")
par(new=T)
plot(density(t.4,from=-3,to=3),ylim=c(0,0.5),ann=F,axes=F,lwd=4,col="red")
par(new=T)
plot(density(Z,from=-3,to=3),ylim=c(0,0.5),ann=F,axes=F,lwd=5,col="grey")
legend("topright",lwd=c(2,4,5),col=c("red","red","grey"),
       c("t with d.f.=2","t with d.f.=4","standard normal"),
       bty="n")
```

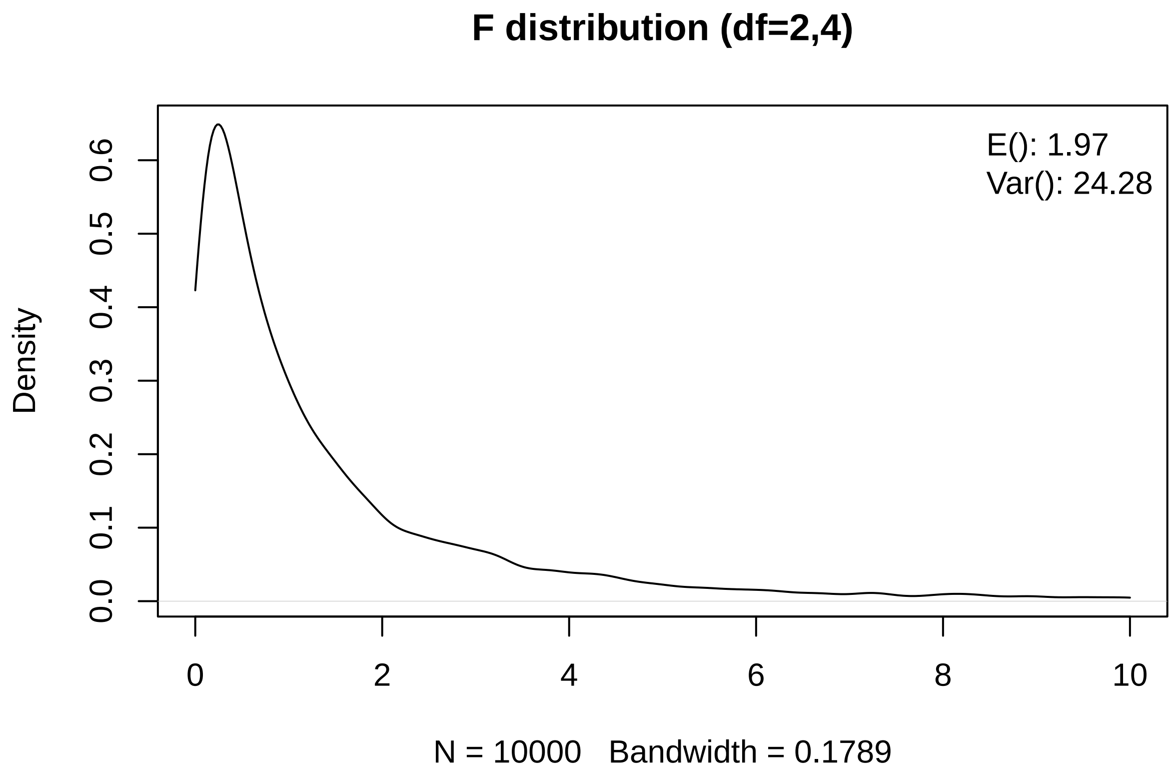


F distribution

A F-distributed random variable is the ratio of two chi-squared random variables divided by their own degrees of freedom. Thus, each F-distribution has a pair of degrees of freedom.

```
F.2.4 <- (chi.square.2 /2)/(chi.square.4 /4)

plot(density(F.2.4,from=0,to=10),main="F distribution (df=2,4)")
legend("topright",c(paste("E():",round(mean(F.2.4),2)),
                    paste("Var():",round(mean(F.2.4^2)-mean(F.2.4)^2,2))),
       bty="n")
```


F distribution (df=2,4)
E(): 1.97
Var(): 24.28
Density
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0
2
4
6
8
10
N = 10000   Bandwidth = 0.1789