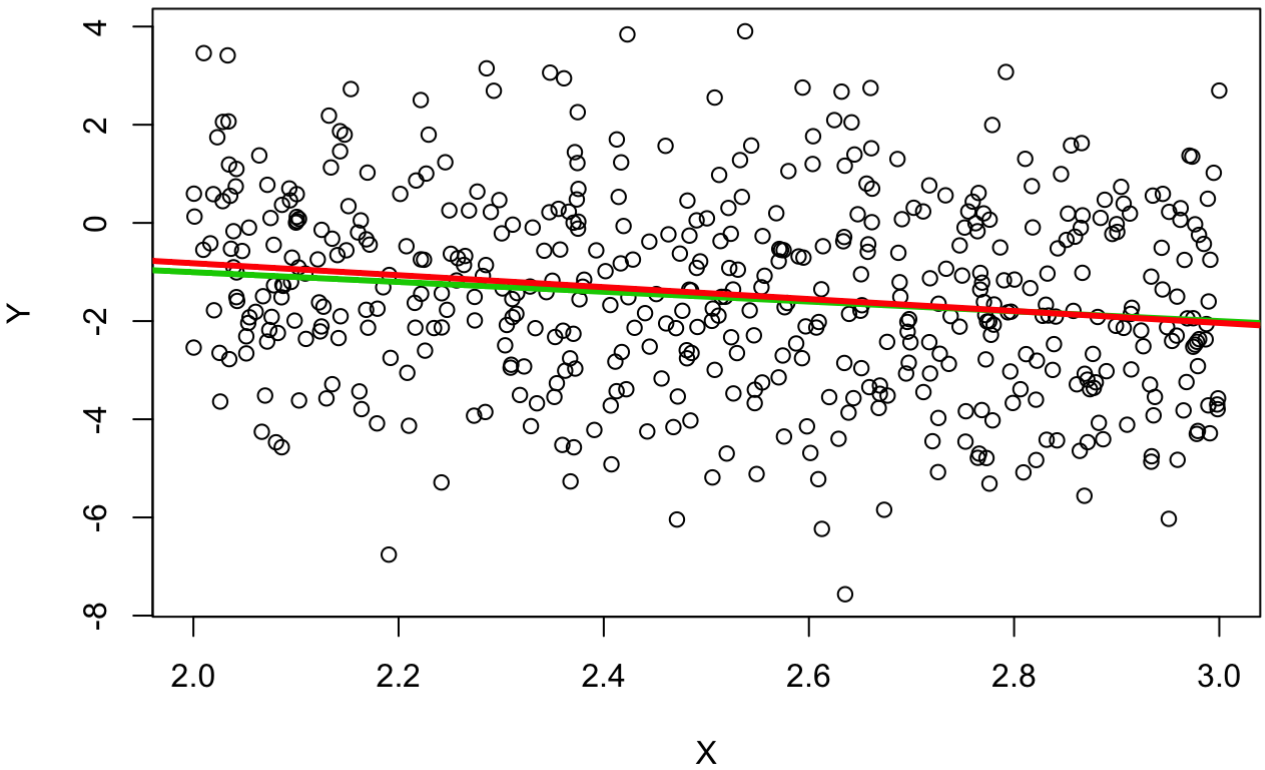
RTSM_SLR.R

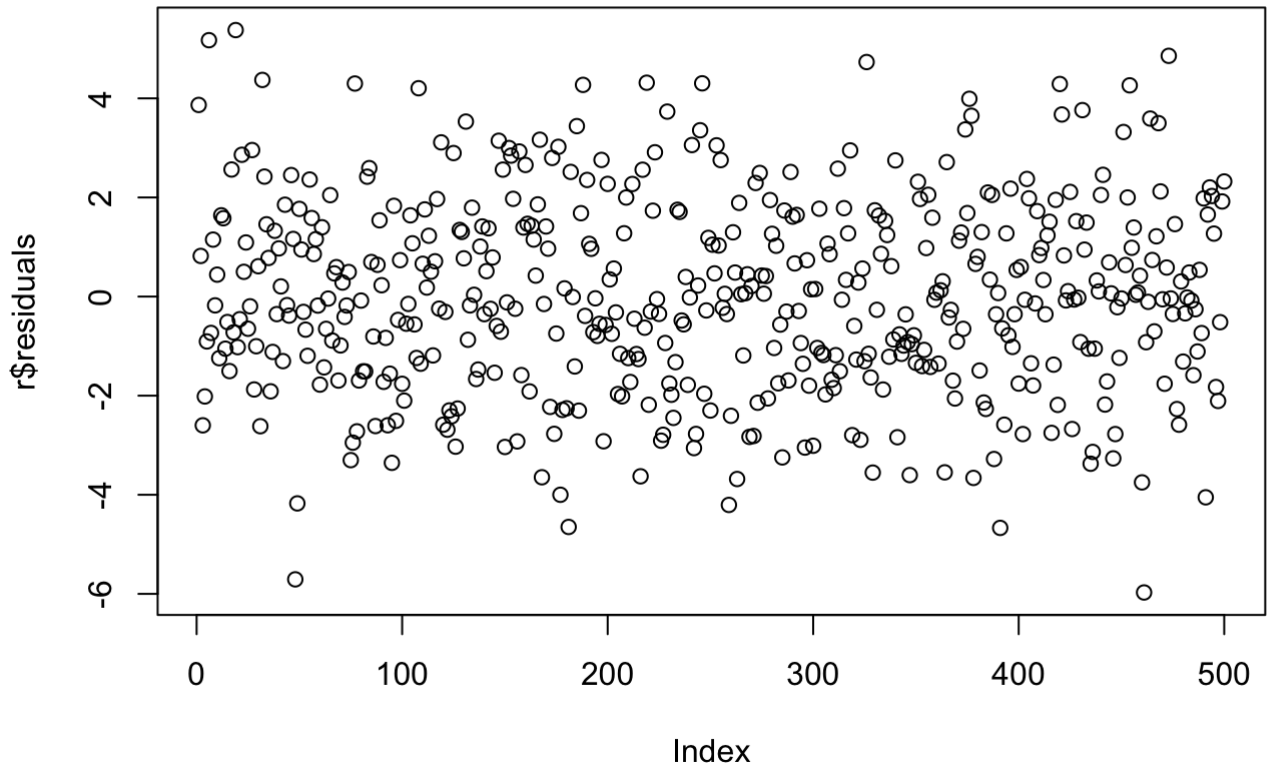
buddhananda

Tue Sep 22 22:18:36 2020

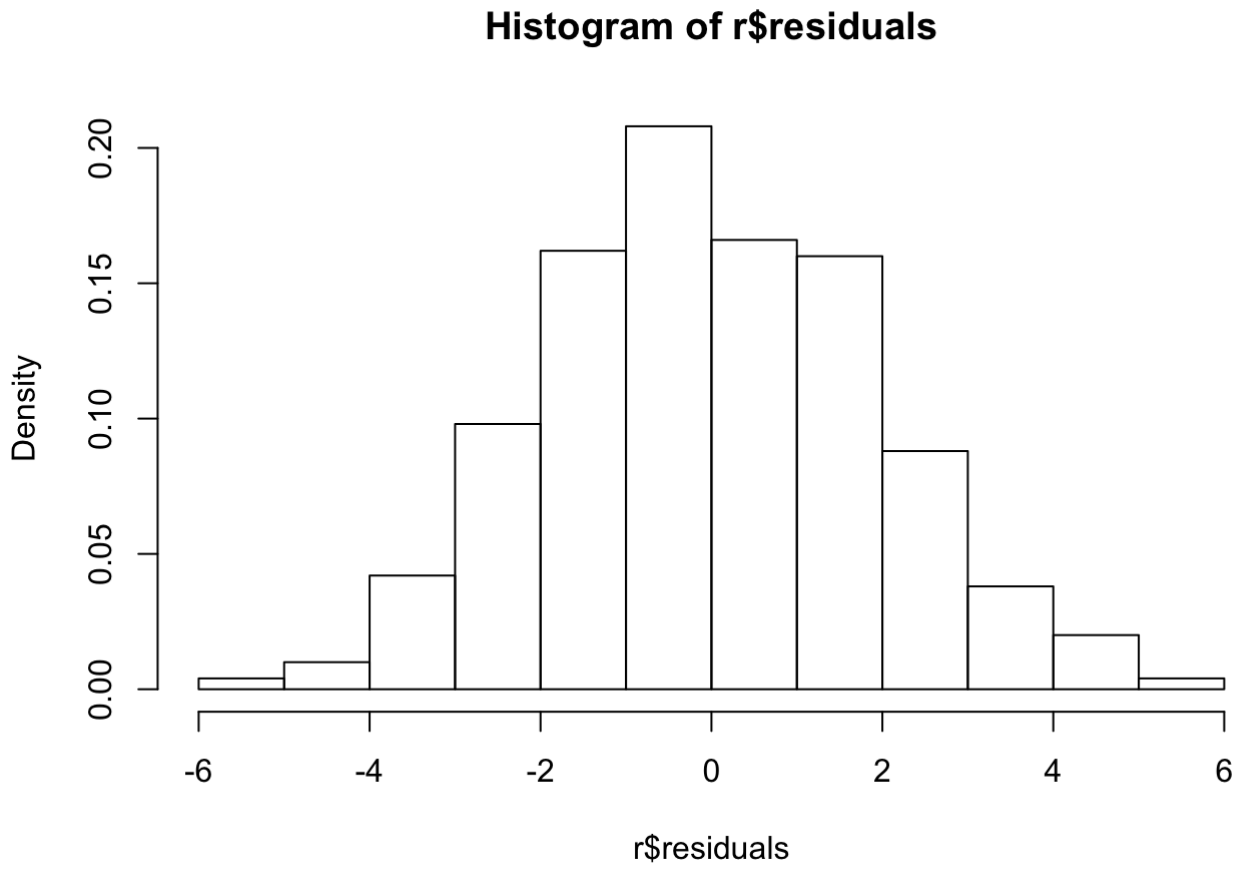
```
# Simple linear regression
####################################
N <- 500  # sample size
a <- 1   # intercept
b <- -1  # slope

epsilon <- rnorm(N, 0, 2) # random error
X <- runif(N,2,3)  # X values
Y <- a + b * X + epsilon # Y values

plot(X,Y)   # scatter plot
abline(a,b, lty=1, col=3, lwd=3)  # Oreginal line in GREEN
r<-lm(Y~X)  # filling model

abline(r$coefficients[1],r$coefficients[2], col="red", lwd=3) # fitted line
```




```
yh<-r$coefficients[1]+r$coefficients[2]*X  # estimated values of y
plot(r$residuals) # residul plot  ############
```




```
hist(r$residuals,probability = T) # residual histogram ###############
```



```
cor(r$residuals,yh) # correlation betewen y_hat and estimated error
```

```
## [1] 6.057997e-16
```

```
print(r)  # estimated parameter values
```

```
##
## Call:
## lm(formula = Y ~ X)
##
## Coefficients:
## (Intercept)            X
##       1.619       -1.219
```

```
cat("Estimated sigma square=",sum((r$residuals)^2)/(r$df.residual),"\n")
```

```
## Estimated sigma square= 3.795471
```

```
itrn<-10000
ah<-array(0,dim=c(itrn))
bh<-array(0,dim=c(itrn))
varh<-array(0,dim=c(itrn))

for(i in 1 : itrn){
  epsilon <- rnorm(N, 0,2)
  Y <- a + b * X + epsilon
  r<-lm(Y~X)
 # print(r)

  ah[i]<-r$coefficients[1]
  bh[i]<-r$coefficients[2]
 # abline(r$coefficients[1],r$coefficients[2], col="red", lwd=1, lty=4) #####
  varh[i]<-sum((r$residuals)^2)/(r$df.residual)
}

cat("MEAN(ah)=", mean(ah), "SD(ah)=",sd(ah),"\n")
```

```
## MEAN(ah)= 0.999453 SD(ah)= 0.7588602
```

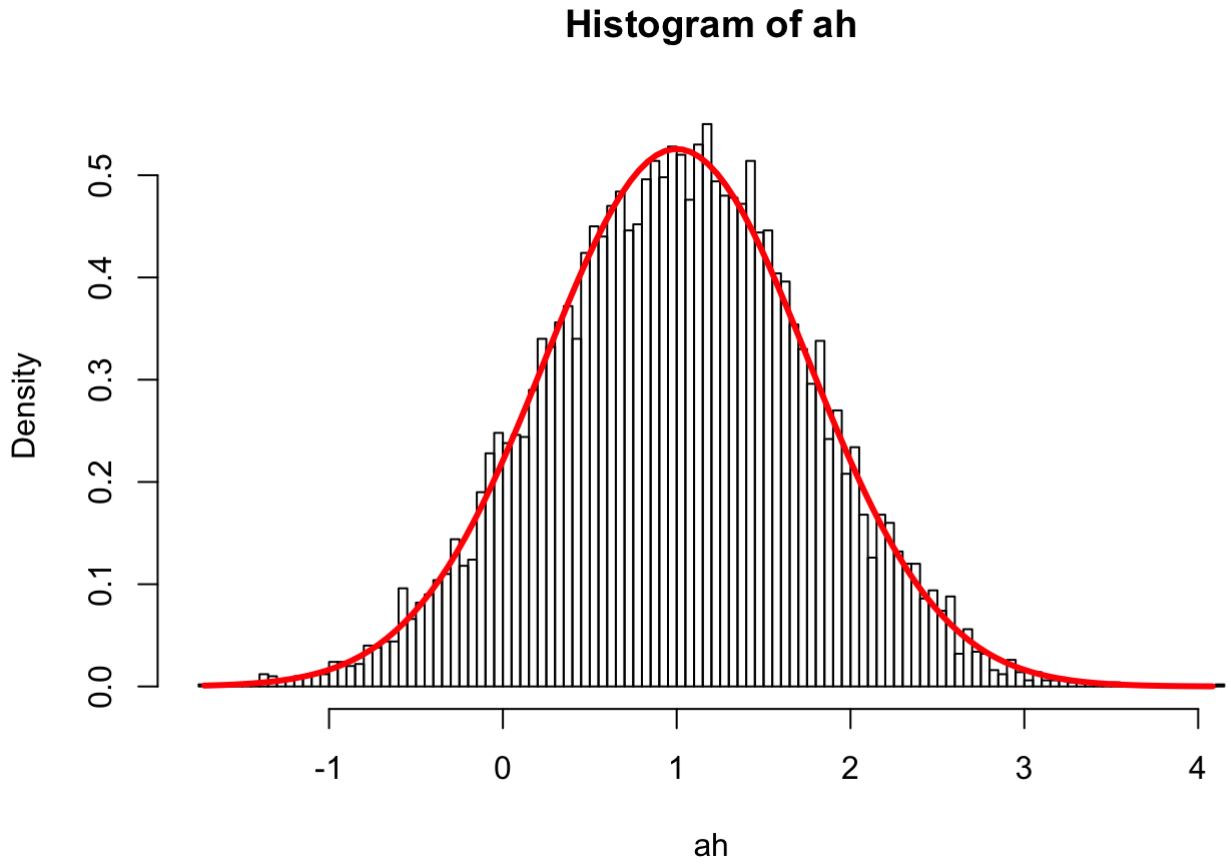
```
cat("MEAN(bh)=", mean(bh), "SD(bh)=",sd(bh),"\n")
```

```
## MEAN(bh)= -0.9995818 SD(bh)= 0.2997363
```

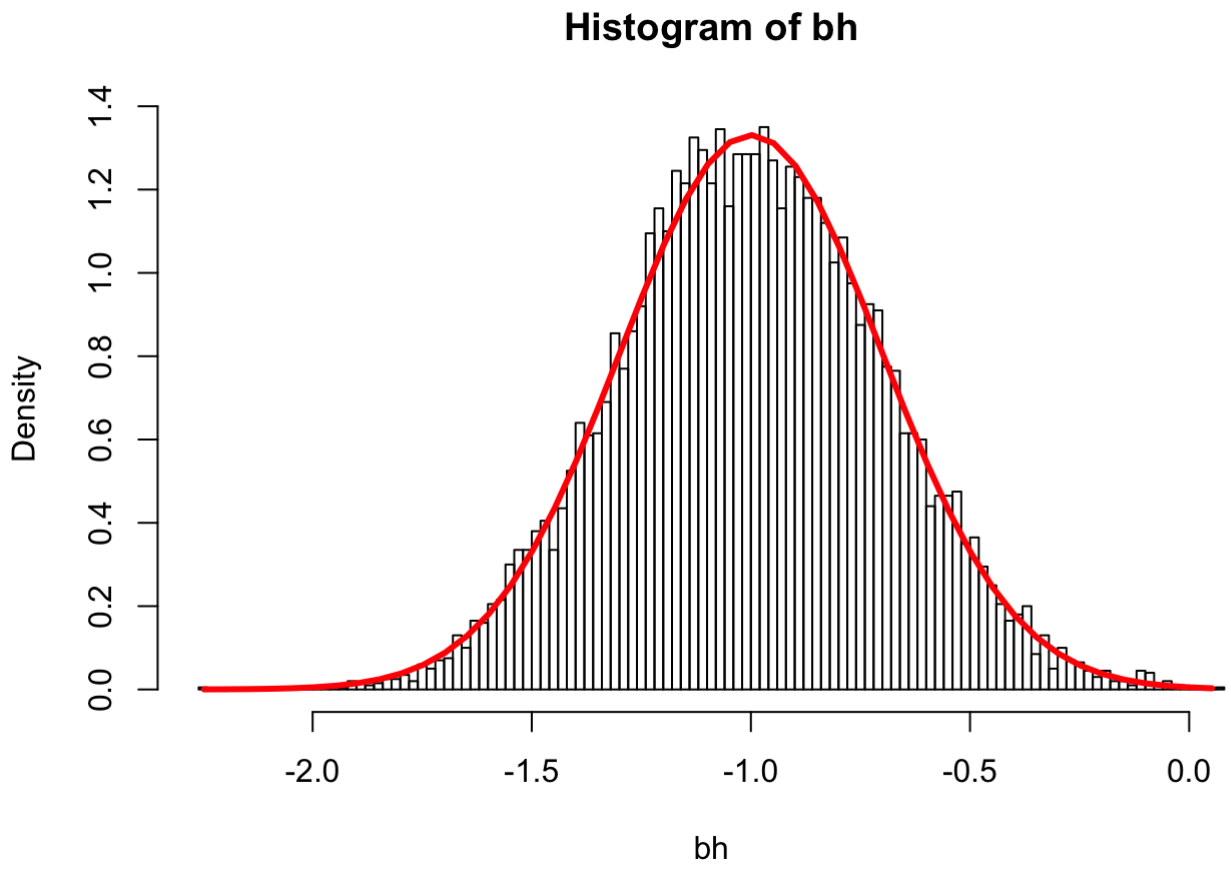
```
cat("MEAN(varh)=", mean(varh), "\n")
```

```
## MEAN(varh)= 3.997597
```

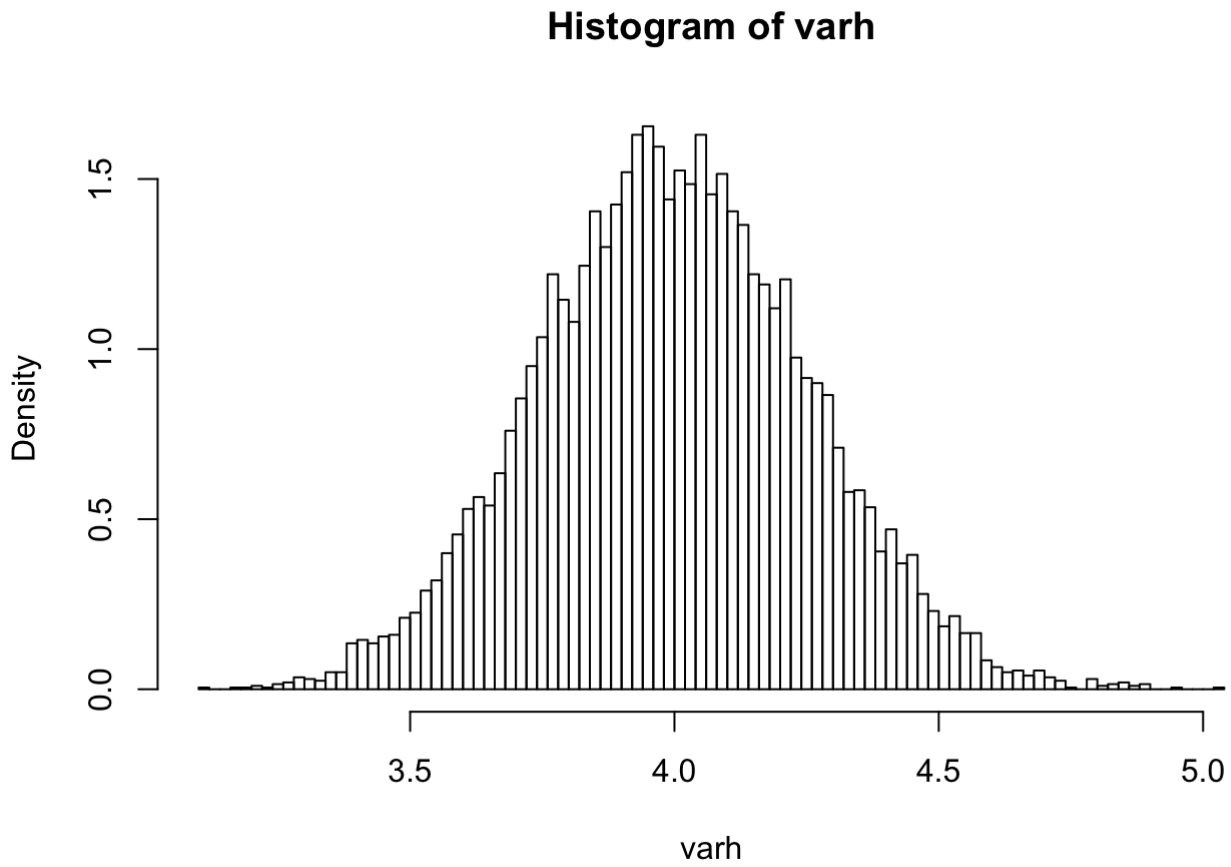
```
hist(ah, probability = T, breaks = 100)
s<-seq(min(ah),max(ah),by=0.05)
lines(dnorm(s, mean=mean(ah), sd=sd(ah))~s, col=2, lwd=3)
```




```
hist(bh,probability = T, breaks = 100)
s<-seq(min(bh),max(bh),by=0.05)
lines(dnorm(s, mean=mean(bh), sd=sd(bh))~s, col=2, lwd=3)
```




```
hist(varh,probability = T, breaks = 100)
```



```
v<-(r$df.residual)*varh/4
hist(v,probability = T, breaks = 100,xlim = c(0,max(v)))
s<-seq(0,max(v),by=0.05)
lines(dchisq(s, df=(r$df.residual))~s, col=2, lwd=3)
```