HW\_3/9

Larin Anton

12/18/2020

## ИДЗ 3 Статан

## ИДЗ 9 Матпакеты

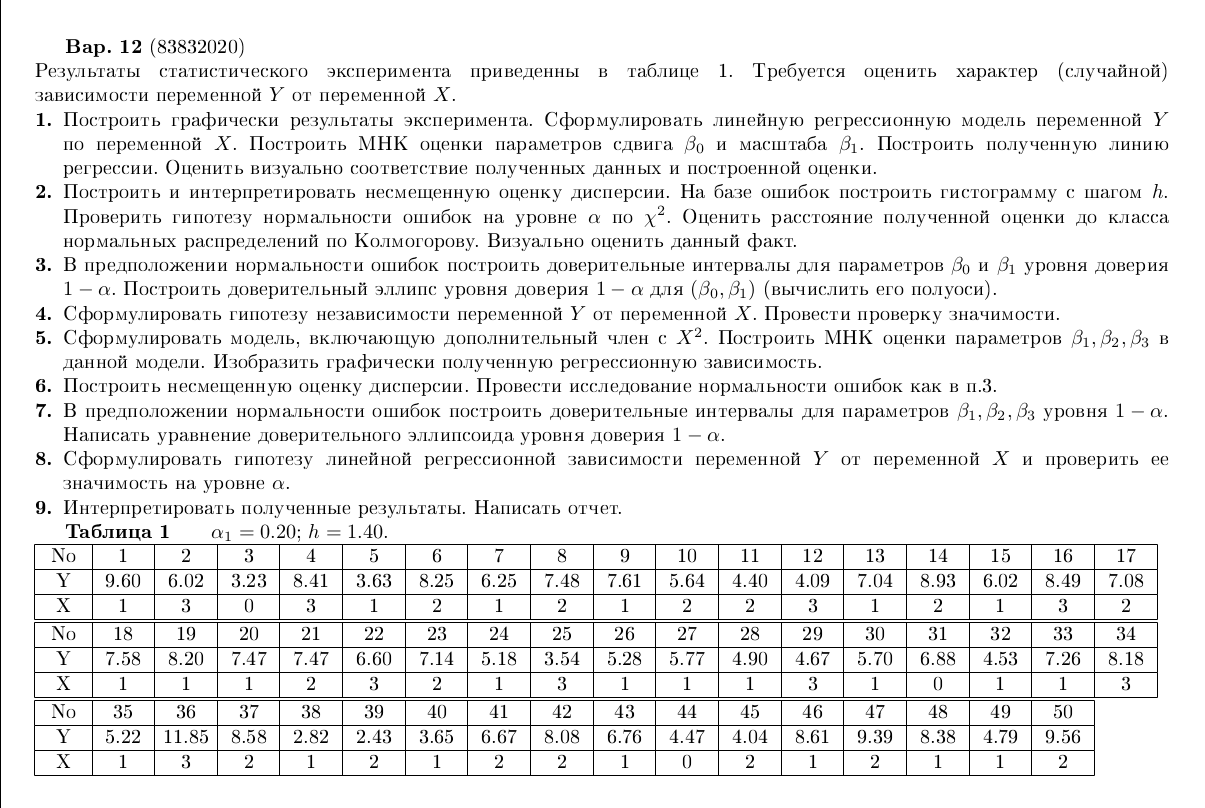
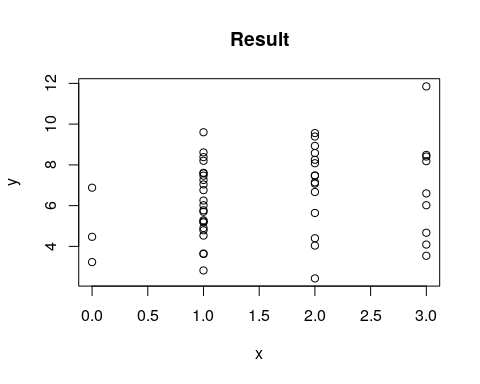


Image alt

#Ход работы ##Графические представление

plot(x,y,main="Result")



Линейная регрессионная модель:

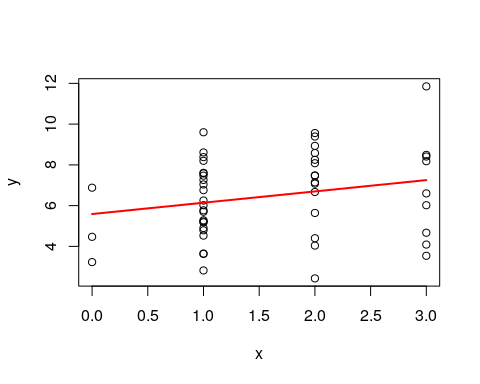
## 1

МНК оценка параметров сдвига и масштаба

n<-length(y)  
x0<-array(1,dim=n)  
X<-t(matrix(c(x0,x),nrow=n,ncol=2))  
Y<-as.matrix(y)  
S<-X%\*%t(X)  
S1<-solve(S)  
bhat<-S1%\*%X%\*%Y

#### Результат:

Нарисуем полученную регрессионную зависимость



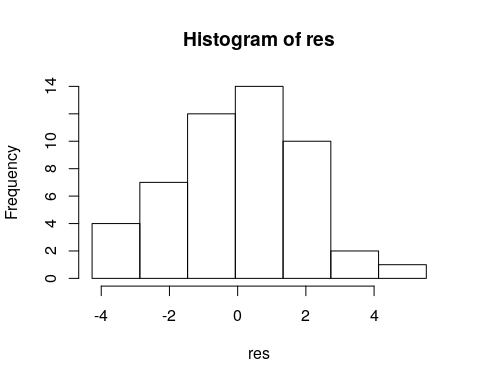
## 2

Построение несмещенной оценки дисперсии

res<-Y-t(X)%\*%as.matrix(bhat)  
SS<-sum(res^2)  
s2<-SS/(n-2)

Результат:

Построение гистограммы с шагом h = 1.4 на базе ошибок



Проверка гипотезы нормальности ошибок на уровне по

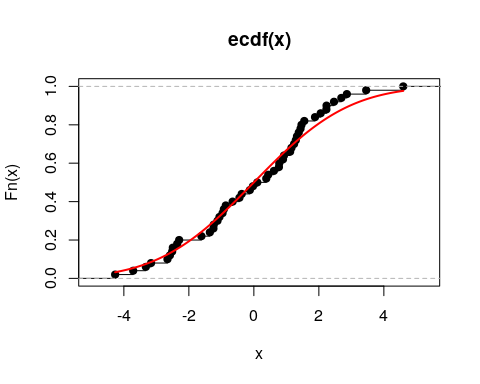
hh<-hist(res,breaks=brk,plot=FALSE)  
nu<-hh$counts  
breaks = hh$breaks;  
r = length(breaks) - 1  
l.b<-length(brk)  
csq0<-function(s){  
 if (s>0){  
 p<-pnorm(brk[2:l.b],0,s)-pnorm(brk[1:(l.b-1)],0,s)  
 return(sum((nu-n\*p)^2/n/p))  
 } else {  
 return(Inf)  
 }  
}  
csq.s<-nlm(csq0,p=sqrt(s2))$minimum  
pv<-pchisq(csq.s,r - 2,lower.tail=FALSE)

Результат: FALSE

Оценим расстояние оценки до класса нормальных распределений оп Колмогорову

kolm.stat<-function(s){  
 sres<-sort(res)  
 fdistr<-pnorm(sres,0,s)  
 max(abs(c(0:(n-1))/n-fdistr),abs(c(1:n)/n-fdistr))  
}  
ks.dist<-nlm(kolm.stat,p=sqrt(s2))$minimum

Полученое расстояние: 0.0708533



## 3

Построим ДИ для параметров с уровнем доверия

C<-diag(c(1,1))  
ph<-bhat  
V<-diag(S1)  
xa<-qt(1-al/2,n-2)  
s1<-sqrt(s2)  
d<-xa\*s1\*sqrt(V)  
CI<-data.frame(lw=ph-d,up=ph+d)

Для   
Для

Доверительный эллипс можно вычислить как

Где:

## [,1]  
## [1,] 5.5889333  
## [2,] 0.5546667

## [,1] [,2]  
## [1,] 0.09111111 -0.04444444  
## [2,] -0.04444444 0.02777778

## 4

Гипотеза независимости Y от X:   
Критерий:

Найдем статистику F-критерия и P-значение:

FST<-bhat[2]^2/V[2]/s2  
pv.f<-pf(FST,1,n-2,lower.tail=FALSE)

Получим 2.7210012 и 0.1055659  
Результат: FALSE

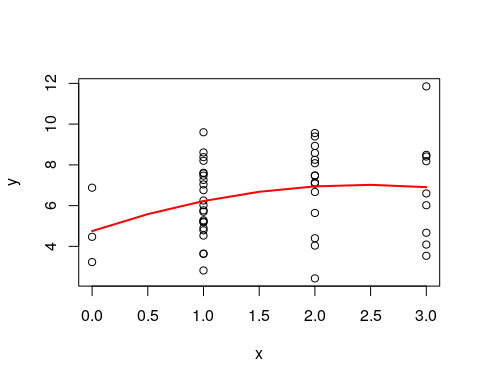
## 5

Добавим в модель член с :

Найдем МНК оценки

x0 <- array(1, dim=n)  
X <- t(matrix(c(x0, x, x^2), nrow=n, ncol=3))  
Y <- as.matrix(y)  
S <- X%\*%t(X)  
S1 <- solve(S)  
bht <- S1%\*%X%\*%Y

Нарисуем полученную регрессионную зависимость



## 6

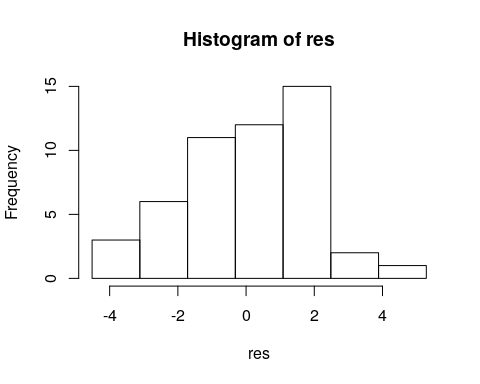
Несмещенная оценка дисперсии

res<-Y-t(X)%\*%as.matrix(bht)  
SS<-sum(res^2)  
s2<-SS/(n-2)

Результат: 3.9820608

Гистограмма на базе ошибок

brk<-seq(min(res), max(res) + h, by=h)  
hist(res,breaks=brk)



Проверка гипотезы нормальности ошибок на уровне по

l.b<-length(brk)  
brk[1]<- -Inf  
brk[l.b]<-Inf  
#r = length(breaks) - 1  
hh<-hist(res,breaks=brk,plot=FALSE)  
nu<-hh$counts  
breaks = hh$breaks;  
r = length(breaks) - 1  
  
l.b<-length(brk)  
csq0<-function(s){  
 if (s>0){  
 p<-pnorm(brk[2:l.b],0,s)-pnorm(brk[1:(l.b-1)],0,s)  
 return(sum((nu-n\*p)^2/n/p))  
 } else {  
 return(Inf)  
 }  
}  
csq.s<-nlm(csq0,p=sqrt(s2))$minimum

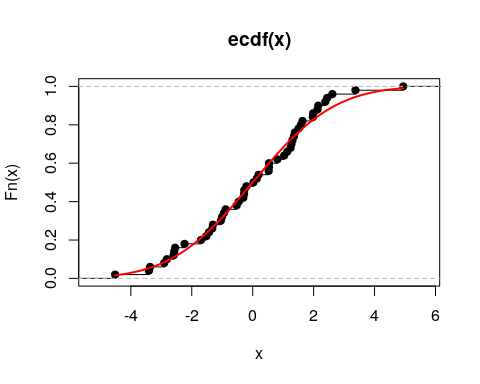
## Warning in nlm(csq0, p = sqrt(s2)): NA/Inf replaced by maximum positive value

pv<-pchisq(csq.s,r-3,lower.tail=FALSE)

Результат: FALSE

Оценка расстояния до нормального рапределения по Колмагорову

kolm.stat<-function(s){  
 sres<-sort(res)  
 fdistr<-pnorm(sres,0,s)  
 max(abs(c(0:(n-1))/n-fdistr),abs(c(1:n)/n-fdistr))  
}  
ks.dist<-nlm(kolm.stat,p=sqrt(s2))$minimum  
  
  
x2<-c(0:1000)\*(max(res)-min(res))/1000+min(res)  
y2<-pnorm(x2,0,nlm(kolm.stat,p=sqrt(s2))$estimate)



Полученое расстояние: 0.0677388

## 7

Построим ДИ для параметров

C<-diag(c(1,1,1))  
ph<-bht #t(C)%\*%bhat  
V<-diag(S1) # diag(C%\*%S1%\*%t(C))  
xa<-qt(1-al/2,n-2)  
s1<-sqrt(s2)  
d<-xa\*s1\*sqrt(V)  
CI<-data.frame(lw=ph-d,up=ph+d)

Полученые интервалы:

## lw up  
## 1 3.4410260 6.0647164  
## 2 0.1622953 3.5408451  
## 3 -0.8533866 0.0979088

Доверительный эллипсоид имеет форму

Где:

## [,1]  
## [1,] 4.7528712  
## [2,] 1.8515702  
## [3,] -0.3777389

## [,1] [,2] [,3]  
## [1,] 0.25594437 -0.3001346 0.07447286  
## [2,] -0.30013459 0.4244056 -0.11552266  
## [3,] 0.07447286 -0.1155227 0.03364738

## 8

Гипотеза линейной регрессионной зависимости Y от X:   
Критерий:

Найдем статистику F-критерия и P-значение:

FST<-bht[3]^2/S1[2,2]/s2  
pv.f<-pf(FST,1,n-2,lower.tail=FALSE)

Получим 0.0844295 и 0.7726338  
Результат: TRUE