REGRESYON ÖDEV-1

BERKE DİLMEN

21723746

KONU

Bir ülkenin atletizm federasyonu, sporcularının 30 km maratonunu koşma süreleriyle ilgili bir araştırma yapıyor. Bu araştırma için 100 milli sporcu seçiliyor. Sporcuların günlük antrenman süresi(x1), bir günde telefon kullanma süresi(x2), hava sıcaklığı(x3) ve sporcuların yaş aralıkları(x4 [18-22 ->1, 22-26->2, 26+ ->3]) olarak verilmiştir. Bu araştırma ile ilgili regresyon analizi yapınız.

TEMEL İSTATİSTİKLER

**basicStats(odev)**

y x1 x2 x3 x4

nobs 100.000000 100.000000 100.000000 100.000000 100.000000

NAs 0.000000 0.000000 0.000000 0.000000 0.000000

Minimum 11.995795 1.074151 -1.063915 -1.571245 1.000000

Maximum 161.615522 6.156524 4.606860 3.000806 3.000000

1. Quartile 51.133082 3.475043 1.428921 0.297231 1.750000

3. Quartile 78.053326 4.783061 2.558154 1.673506 3.000000

Mean 67.275078 4.049988 2.032580 1.009539 2.150000

Median 64.819051 4.138544 2.025288 1.092905 2.000000

Sum 6727.507782 404.998820 203.258044 100.953893 215.000000

SE Mean 2.568870 0.099120 0.098325 0.091812 0.079614

LCL Mean 62.177882 3.853313 1.837481 0.827364 1.992029

UCL Mean 72.372274 4.246663 2.227680 1.191714 2.307971

Variance 659.909443 0.982472 0.966790 0.842948 0.633838

Stdev 25.688703 0.991197 0.983255 0.918122 0.796140

Skewness 1.118331 -0.285603 0.109344 -0.058128 -0.269012

Kurtosis 2.221259 -0.214254 0.450691 -0.396550 -1.391464

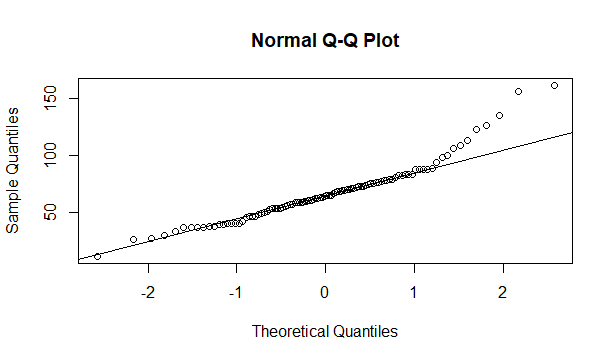
1-NORMALLİK TESTİ

Normallik Testi incelemesinde P değeri,α=0.05 ten büyük ise normallik varsayımına uyduğu kabul edilir.

**attach(odev)**

**qqnorm(odev$y)**

**qqline(odev$y)**



shapiro.test(odev$y)

Shapiro-Wilk normality test

data: odev$y

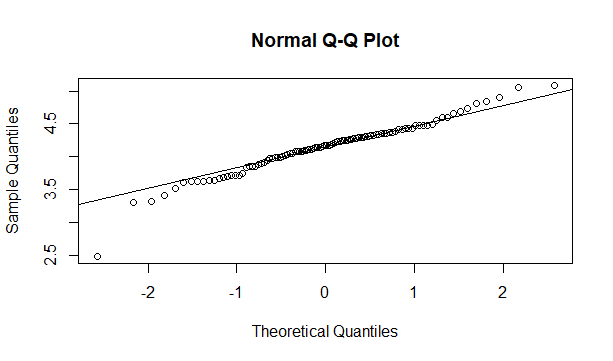
W = 0.92975, p-value = 4.82e-05

**p<α=0.05 olduğundan ln dönüşümü uygulanacak.**

**lny<- log(odev$y)**

**qqnorm(lny)**

**qqline(lny)**



shapiro.test(lny)

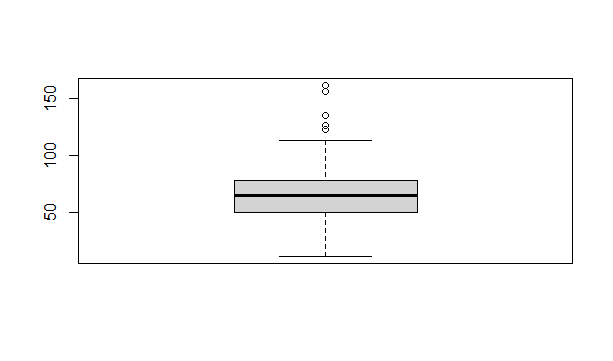
Shapiro-Wilk normality test

data: lny

W = 0.96225, p-value = 0.00584

P=0.00584<α=0.05 olduğundan bu sefer box plot kullanılacak.

**boxplot(odev$y)**



**which(odev$y%in%boxplot(odev$y)$out)**

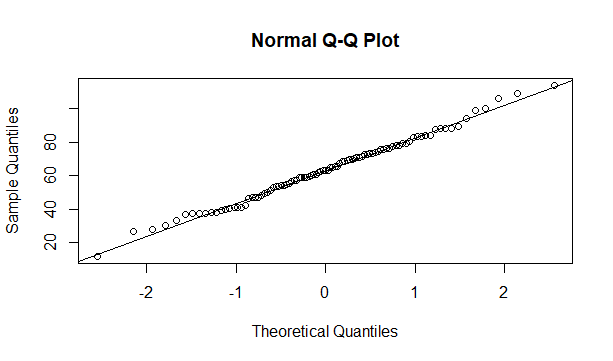
[1] 2 21 47 73 90

Veri setimizden bu değerleri sildirip tekrar test edeceğiz.

**odevy<- odev[-c(2,21,47,73,90),]**

**qqnorm(odevy$y)**

**qqline(odevy$y)**



Grafikten de görüldüğü üzere çok daha doğrusal bir sonuç elde ettik.

**shapiro.test(odevy$y)**

Shapiro-Wilk normality test

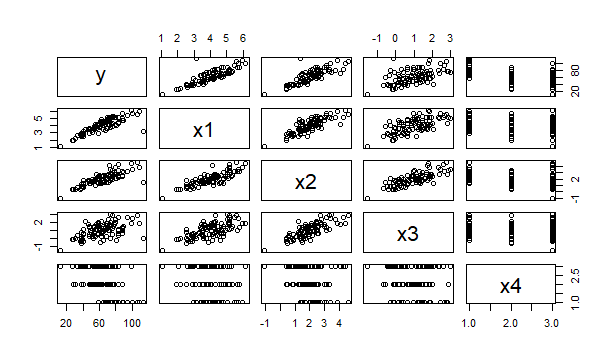
data: odevy$y

W = 0.99321, p-value = 0.9152

**P=0.9152> α=0.05 olduğundan normallik varsayımını sağladı.**

2-DOĞRUSALLIK

pairs(odevy)



**Grafiğe bakarsak en kuvvetli doğrusal bağlantı [y,x1] arasında görülmektedir.**

**x4 ile hiçbir değişkenin doğrusal bağlantısı olduğu söylenemez.**

3-REGRESYON ANALİZİ

**sonuc<- lm(odevy$y~odevy$x1+odevy$x2+odevy$x3)**

**summary(sonuc)**

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 4.134 4.626 0.894 0.373859

odevy$x1 10.510 1.723 6.101 2.53e-08 \*\*\*

odevy$x2 11.382 2.117 5.376 5.85e-07 \*\*\*

odevy$x3 -6.031 1.670 -3.612 0.000497 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 9.298 on 91 degrees of freedom

Multiple R-squared: 0.7801, Adjusted R-squared: 0.7728

F-statistic: 107.6 on 3 and 91 DF, p-value: < 2.2e-16

R^2= 0.7801 => %78

**\*Maraton koşu süresinin %78’i, antrenman süres,telefon kullanımı ve hava sıcaklığı ile açıklanabilir.**

**\*%22 lik etkisi ile bu modellemeyle açıklanamamaktadır.**

Eğer;

Ho: ß=0 -> Model anlamlı değildir.

Ho: ß≠0 -> Model anlamlıdır.

P=2.2e-16 < α=0.05 -> Ho Red. Model anlamlıdır.

**^yi= 4.134+10.510(xi1)+11.382(xi2)-6.031(xi3) ± 9.298**

**(4.626) (1.723) (2.117) (1.670)**

Katsayı Yorumları

**\*bo=4.134**

>Maraton süresinin 1 dakika artması, modeli 4.134 kat değiştirir.

**\*b1=10.510**

**>x2 ve x3 sabit tutulduğunda, günlük antrenman süresinin 1 saat artması, maraton koşu süresini 10.510 dakika etkilemektedir. Güven değeri p=0.000 < α =0.05 olduğundan katsayı istatistiksel olarak anlamlıdır.**

**\*b2=11.382**

**>x1 ve x3 sabit tutulduğunda, telefon kullanımının 1 saat artması artması, maraton koşma dakikasını 11.382 birim arttırır. Güven değeri p=0.000 < α =0.05 olduğundan katsayı istatistiksel olarak anlamlıdır.**

**\*b3=-6.031**

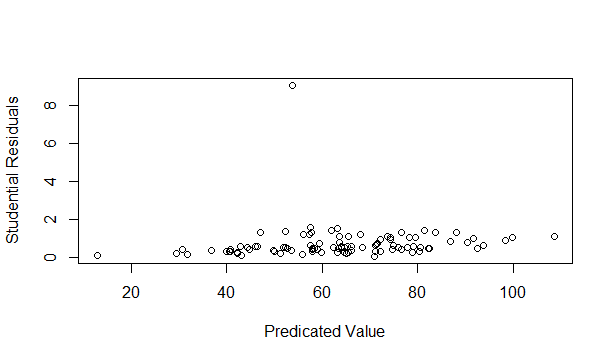
**>x1 ve x2 sabit tutulduğunda, hava durumunın 1 derece artması, maraton süresini 6.031 dakika azaltır.**

4-ARTIK İNCELEMESİ

**inf<- ls.diag(sonuc)**

**influence.measures(sonuc)**

**par(mfrow= c(1,1))**

**plot(predict(sonuc),abs(inf$stud.res),ylab= "Studential Residuals", xlab= "Predicated Value")**

**Grafikte 50-60 arasındaki bir değerin aykırı değer olduğu izlenimini yapabiliriz. Detaylı bir şekilde incelenecek.**

5-AYKIRI DEĞER İNCELEMESİ

5.1-COOK UZAKLIĞI İLE

**n<- 95**

**k<- 4 (x4 ü saymadık)**

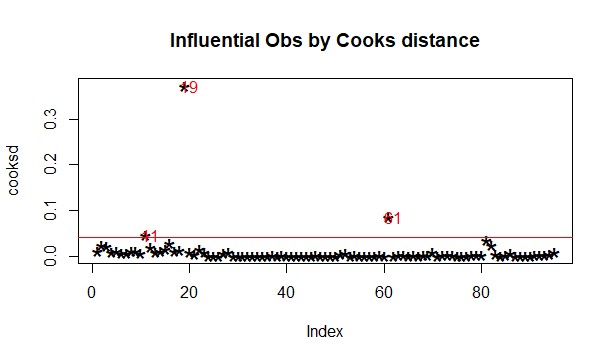
**cooksd<- cooks.distance(sonuc)**

**plot(cooksd, pch="\*", cex=2, main="Influential Obs by Cooks distance")**

**abline(h = if (n>50) 4/n else 4/(n-k-1) , col="red")**

**text(x=1:length(cooksd)+1, y=cooksd, labels=ifelse(cooksd>if (n>50) 4/n else 4/(n-k-1),names(cooksd),""), col="red")**

**boxplot(cooksd)$out**



11, 19 ve 61 değerlerinin aykırı olduğunu gördük. Boxplot ile çizdirip ortak olanları veriden sildireceğiz.

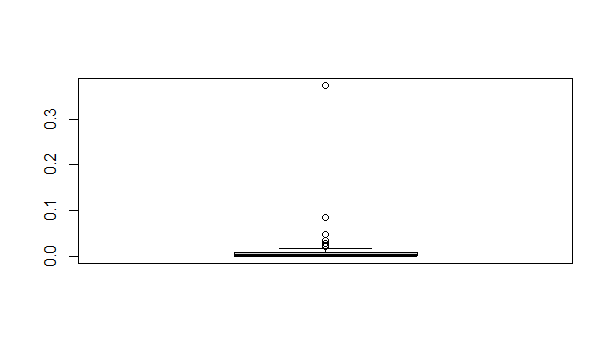
**boxplot(cooksd)$out**

2 3 11 12 16 19 61

0.02435299 0.02218216 0.04756723 0.02081272 0.02842395 0.37368824 0.08591348

81 82

0.03466535 0.02419899



Her ikisinde de ortak olan 11,19 ve 61 değerleri olduğundan sildirme işlemi uyguluyoruz.

c**ooksdd<- cooksd[-c(11,19,61)]**

**\*Normallik varsayımının bozulup bozulmadığını incelemek için Shapiro Wilk testi yapıyoruz.**

**shapiro.test(cooksdd)**

Shapiro-Wilk normality test

data: cooksdd

W = 0.71586, p-value = 4.734e-12

**\*p<0.05 çıktı ln dönüşümünden yardım alınır.**

**lncook<- log(cooksdd)**

**shapiro.test(lncook)**

**Shapiro-Wilk normality test**

data: lncook

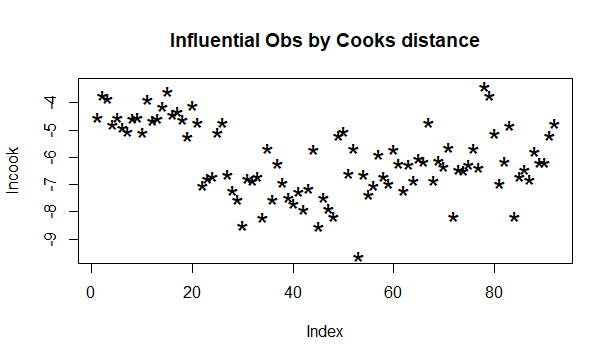
W = 0.98196, p-value = 0.2338

**\*p>0.05, nomalliği sağladı. Son olarak tekrar artık değerimiz var mı test edelim.**

**plot(lncook, pch="\*", cex=2, main="Influential Obs by Cooks distance")**

**abline(h = if (n>50) 4/n else 4/(n-k-1) , col="red")**

**text(x=1:length(lncook)+1, y=lncook, labels=ifelse(lncook>if (n>50) 4/n else 4/(n-k-1),names(lncook),""), col="red")**

**Herhangi bir aykırı değere rastlanmadığından sıradaki adımlara geçebiliriz.**

5.2-UÇ DEĞER İLE

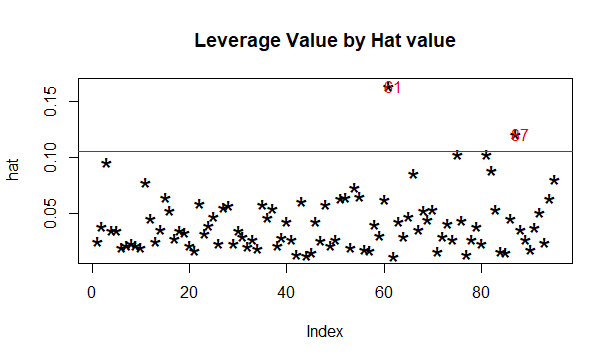
**library(zoo)**

**hat<- inf$hat**

**plot(hat, pch="\*", cex=2, main="Leverage Value by Hat value")**

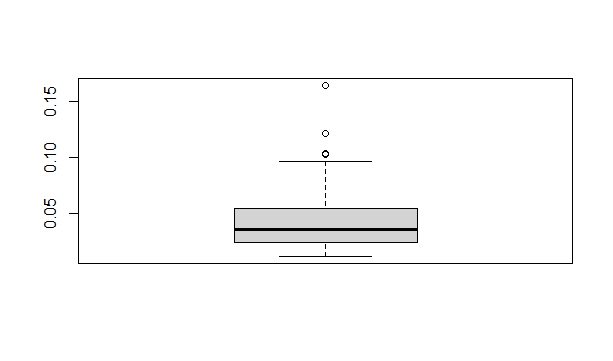
**abline(h = 2\*(k+1)/n , col="red")**

**text(x=1:length(hat)+1, y=hat, labels=ifelse(hat>2\*(k+1)/n,index(hat),""), col="red")**



Grafikte 61 ve 87 değerlerinin aykırı değer olduğunu gördük. Aynı şekilde boxplot ile karşılaştırma yapacağız.

boxplot(hat)$out

[1] 0.1641402 0.1035062 0.1026703 0.120968

**which(hat%in%boxplot(hat)$out)**

[1] 61 75 81 87

**61 ve 87 ortak aykırılık içerdiği için ikisi silinecek.**

**hatt<- hat[-c(61,87)]**

shapiro.test(hatt)

Shapiro-Wilk normality test

data: hatt

W = 0.90893, p-value = 7.653e-06

**#p<0.05, normallik varsayımını sağlamadı, dönüşüm yapılır.**

**lnhat<-log(hatt)**

**shapiro.test(lnhat)**

Shapiro-Wilk normality test

data: lnhat

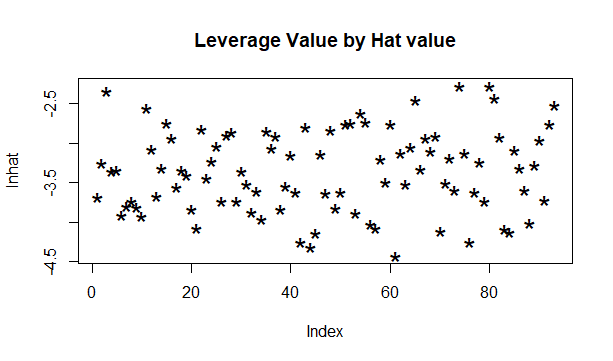
W = 0.98357, p-value = 0.2947

#p>0.05, normallik varsayımını sağladı.

**plot(lnhat, pch="\*", cex=2, main="Leverage Value by Hat value")**

**abline(h = 2\*(k+1)/n , col="red")**

**text(x=1:length(lnhat)+1, y=lnhat, labels=ifelse(lnhat>2\*(k+1)/n,index(lnhat),""), col="red")**



**Tekrar kontrol ettik ve aykırılığa rastlamadık. Kabul edip devam edeceğiz.**

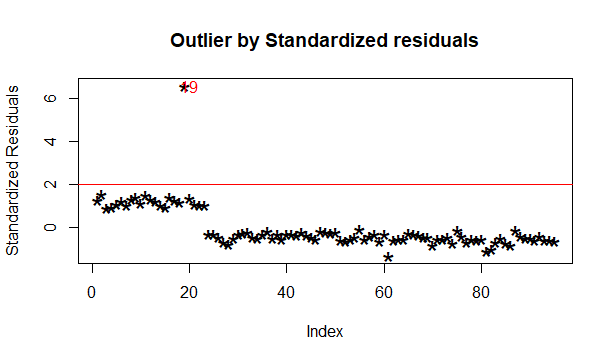
6-STANDARTLAŞTIRILMIŞ ARTIKLAR

**std<-inf$std.res**

**plot(std, pch="\*", cex=2, main="Outlier by Standardized residuals",ylab="Standardized Residuals", xlab="Index")**

**abline(h = c(-2,2) ,col="red")**

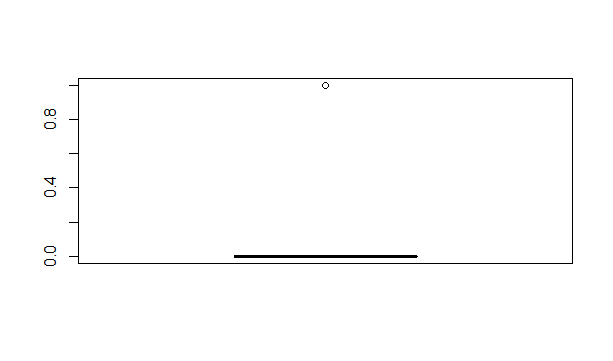
**text(x=1:length(std)+1, y=std, labels=ifelse(std<-2 & std>2,index(std),""), col="red")**



(-2,2) dışında kalan tek artık gözlem 19 çıktı. Box plot ile teyit edeceğiz.

boxplot(std)$out

[1] 1



Görüldüğü üzere tek artık var. Hangisi olduğunu sorgulayalım.

**which(std%in%boxplot(std)$out)**

[1] 19

19. değeri sildireceğiz.

**stdd<- std[-c(19)]**

**shapiro.test(stdd)**

Shapiro-Wilk normality test

data: stdd

W = 0.78868, p-value = 2.694e-10

**\*p değeri 0.05 ten küçük. Dönüşüm yapılır.**

**lnstdd<-log(stdd)**

**shapiro.test(lnstdd)**

**Shapiro-Wilk normality test**

data: lnstdd

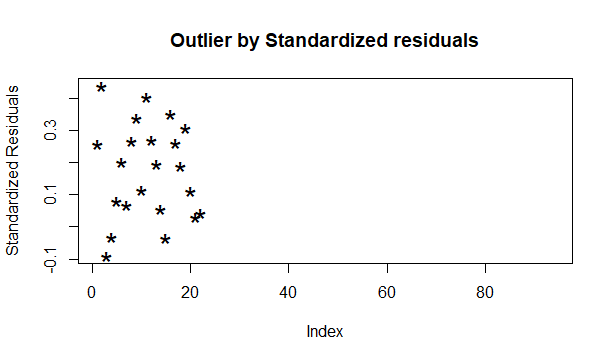
W = 0.97023, p-value = 0.7164

**\*p>0.05. Normallik varsyaımını sağladı.**

**plot(lnstdd, pch="\*", cex=2, main="Outlier by Standardized residuals",ylab="Standardized Residuals", xlab="Index")**

**abline(h = c(-2,2) ,col="red")**

**text(x=1:length(lnstdd)+1, y=lnstdd, labels=ifelse(lnstdd<-2 & lnstdd>2,index(lnstdd),""), col="red")**



Artık bulunmadığından diğer işleme geçiyoruz.

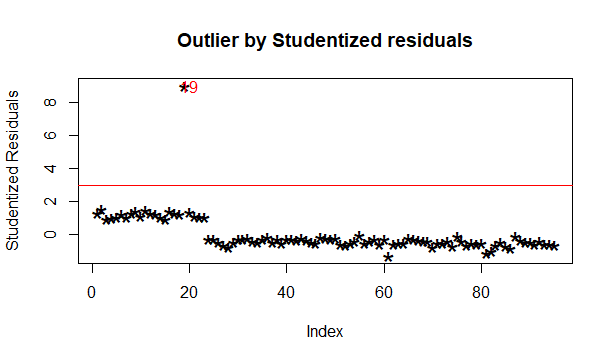
7-STUDENT TÜRÜ ARTIKLAR

**stud<- inf$stud.res**

**plot(stud, pch="\*", cex=2, main="Outlier by Studentized residuals",ylab="Studentized Residuals", xlab="Index")**

**abline(h = c(-3,3) ,col="red")**

**text(x=1:length(stud)+1, y=stud, labels=ifelse(stud<-3 & stud>3,index(stud),""), col="red")**



**boxplot(stud)$out**

[1] 1

**which(stud%in%boxplot(std)$out)**

[1] 19

**Aynı şekilde her ikisinde de 19 aykırı çıktığı için sildiriyoruz.**

**studd<- stud[-c(19)]**

**shapiro.test(studd)**

Shapiro-Wilk normality test

data: studd

W = 0.78906, p-value = 2.764e-10

**\*p<0.05 çıktı. Dönüşüm uygulanır.**

**lnstudd<-log(studd)**

**shapiro.test(lnstudd)**

**Shapiro-Wilk normality test**

data: lnstudd

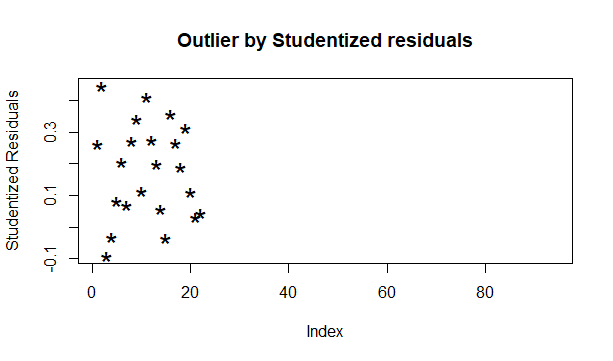
W = 0.9703, p-value = 0.718

**\*p>0.05, normallik varsayımını sağladı.**

**plot(lnstudd, pch="\*", cex=2, main="Outlier by Studentized residuals",ylab="Studentized Residuals", xlab="Index")**

**abline(h = c(-3,3) ,col="red")**

**text(x=1:length(lnstudd)+1, y=lnstudd, labels=ifelse(lnstudd<-3 & lnstudd>3,index(lnstudd),""), col="red")**



**Herhangi bir aykırılık olmadığından olmadığından kabul ediyoruz. Artık değer kalmadığından regresyon analizine geçebiliriz.**

8-GÜVEN ARALIĞI

**confint(sonuc, level = .99) #güven aralığı**

0.5 % 99.5 %

(Intercept) -8.036313 16.30415

odevy$x1 5.977651 15.04286

odevy$x2 5.811663 16.95276

odevy$x3 -10.424261 -1.63837

**\*İlk değer negatif olduğundan yorum yapılmaz.**

**\*x2 ve x3 sabit tutulduğunda, günlük antrenman süresinin 1 saat artmasıi maraton süresini 5.977651 dakika ile 15.04286 dakika arası etkilediğini %99 güven düzeyinde söyleyebiliriz.**

**\*x1 ve x3 sabit tutulduğunda, telefon kullanımının 1 saat artması, maraton koşusunun süresini 5.811663 dakika ile 16.95276 dakika arttırdığını, %99 güven düzeyinde söyleyebiliriz.**

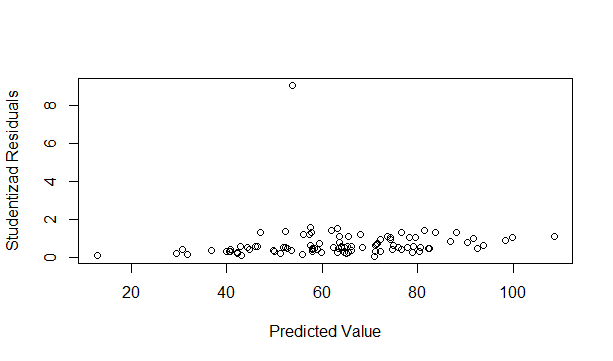
**\*x3 değişkeninde sonuç negatif çıktığından yorum yapılamaz.**

9-DEĞİŞEN VARYANSLILIK SORUNU

**summary(lm(abs(residuals(sonuc)) ~ fitted(sonuc)))**

**par(mfrow=c(1,1))**

**plot(predict(sonuc), abs(inf$stud.res), ylab="Studentizad Residuals", xlab="Predicted Value")**



9.1-BREUSH-PAGAN TESTİ

Breush Pagan Testi ile değişen varyanslılık var mı yok mu tes edeceğiz.

**library(lmtest)**

**bptest(sonuc)**

studentized Breusch-Pagan test

data: sonuc

BP = 2.098, df = 3, p-value = 0.5523

**#p-value 0.05 ten büyükse değişen varyanslılık sorunu yok demektir. Ho reddedilemez.**

9.2 ÖZ İLİŞKİ SORUNU

Durbin-Watson Testi ile analiz etmeye çalışacağız.

**dwtest(sonuc) #Durbin Watson testi**

Durbin-Watson test

data: sonuc

DW = 0.68075, p-value = 9.204e-14

**#DW sonucu 0<d<2 aralığında olduğundan ve 0<p<0.5 olduğundan, pozitif öz ilişki vardır diyoruz.**

10-ÇOKLU BAĞLANTI SORUNU

**library(car)**

**vif(sonuc) #Varyans şişme değerini verir.**

odevy$x1 odevy$x2 odevy$x3

3.267807 4.838279 2.557297

**#Eğer değer 10 dan büyük ise güçlü çoklu bağlantı olduğunu gösterir.**

**#Bizim değerlerimiz ise 10 dan küçük, çoklu bağlantı yoktur denebilir.**

**library(perturb)**

**colldiag(sonuc)** #Koşul sayıları ve varyans bozulum oranı

Condition

Index Variance Decomposition Proportions

intercept odevy$x1 odevy$x2 odevy$x3

1 1.000 0.003 0.001 0.003 0.011

2 3.533 0.045 0.005 0.001 0.366

3 8.555 0.210 0.004 0.461 0.614

4 17.763 0.742 0.990 0.535 0.008

**#Değerler 30 dan büyük ise güçlü çoklu bağlantı var demektir.**

**#Bizim değerlerimiz 30 dan küçük, bu yüzden çoklu bağlantı yoktur.**

11-ÖZDEĞER VE ÖZVEKTÖRLERİN BULUNMASI

sonuc

Call:

**lm(formula = odevy$y ~ odevy$x1 + odevy$x2 + odevy$x3)**

Coefficients:

(Intercept) odevy$x1 odevy$x2 odevy$x3

4.134 10.510 11.382 -6.031

**ort1<-mean(odevy$x1)**

**kt1<-sum((odevy$x1-ort1)^2)**

**skx1<-(odevy$x1-ort1)/(kt1^0.5)**

**ort2<-mean(odevy$x2)**

**kt2<-sum((odevy$x2-ort2)^2)**

**skx2<-(odevy$x2-ort2)/(kt2^0.5)**

**ort3<-mean(odevy$x3)**

**kt3<-sum((odevy$x3-ort3)^2)**

**skx3<-(odevy$x3-ort3)/(kt3^0.5)**

**x<-cbind(skx1,skx2,skx3)**

**sm<- eigen (t(x)%\*%x)**

**signif(sm$values,3)**

[1] 2.510 0.355 0.134

**signif(sm$vectors,3)**

[,1] [,2] [,3]

[1,] -0.572 -0.637 0.518

[2,] -0.603 -0.102 -0.791

[3,] -0.557 0.764 0.325

**#0’a yakın değer olmadığından güçlü çoklu bağlantı vardır yorumunu yapamıyoruz.**

**#İki ya da daha fazla değişken arasında değerler çok yakın olmadığından, ilişki vardır diyemiyoruz.**

**V<-sm$vectors**

**t(V)%\*%V**

[,1] [,2] [,3]

[1,] 1.000000e+00 -1.665335e-16 1.665335e-16

[2,] -1.665335e-16 1.000000e+00 -1.942890e-16

[3,] 1.665335e-16 -1.942890e-16 1.000000e+00

12.1-İLERİYE DOĞRU SEÇİM

**library(stats)**

**lm.null<- lm(odevy$y~1)**

**forward <- step(lm.null,odevy$y~odevy$x1+odevy$x2+odevy$x3, direction =**

**"forward")**

Start: AIC=565.46

odevy$y ~ 1

Df Sum of Sq RSS AIC

+ odevy$x1 1 25370.4 10404 450.13

+ odevy$x2 1 23548.4 12226 465.46

+ odevy$x3 1 9726.4 26048 537.31

<none> 35775 565.46

Step: AIC=450.13

odevy$y ~ odevy$x1

Df Sum of Sq RSS AIC

+ odevy$x2 1 1409.02 8995.3 438.31

<none> 10404.3 450.13

+ odevy$x3 1 38.64 10365.7 451.78

Step: AIC=438.31

odevy$y ~ odevy$x1 + odevy$x2

Df Sum of Sq RSS AIC

+ odevy$x3 1 1128 7867.3 427.58

<none> 8995.3 438.31

Step: AIC=427.58

odevy$y ~ odevy$x1 + odevy$x2 + odevy$x3

forward

Call:

lm(formula = odevy$y ~ odevy$x1 + odevy$x2 + odevy$x3)

Coefficients:

(Intercept) odevy$x1 odevy$x2 odevy$x3

4.134 10.510 11.382 -6.031

**summary(forward)**

Call:

lm(formula = odevy$y ~ odevy$x1 + odevy$x2 + odevy$x3)

Residuals:

Min 1Q Median 3Q Max

-11.245 -4.899 -3.306 -0.673 60.167

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 4.134 4.626 0.894 0.373859

odevy$x1 10.510 1.723 6.101 2.53e-08 \*\*\*

odevy$x2 11.382 2.117 5.376 5.85e-07 \*\*\*

odevy$x3 -6.031 1.670 -3.612 0.000497 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 9.298 on 91 degrees of freedom

Multiple R-squared: 0.7801, Adjusted R-squared: 0.7728

F-statistic: 107.6 on 3 and 91 DF, p-value: < 2.2e-16

**En iyi model: yi = 4.134 + 10.510x1i+11.382x2i-6.031x3i ±9.298**

**\*Pr değerlerine baktığımızda 3 değişken de 0.05 ten küçük olduğundan, anlamlı olduğu yorumunu yapabiliriz.**

**\*Sabit terim 0.05 ten büyük olduğundan anlamlı olduğu söylenemez.**

**\*p-value<2.2e-16 değeri de, modelin genel olarak anlamlı olduğunu belirtir.**

**\*y bağımlı değişkeni x1,x2 ve x3 tarafından %78 oranında açıklanabilmektedir.**

12.2-GERİYE DOĞRU SEÇİM

backward<-step(sonuc,direction="backward")

Start: AIC=427.58

odevy$y ~ odevy$x1 + odevy$x2 + odevy$x3

Df Sum of Sq RSS AIC

<none> 7867.3 427.58

- odevy$x3 1 1128.0 8995.3 438.31

- odevy$x2 1 2498.4 10365.7 451.78

- odevy$x1 1 3217.6 11084.9 458.15

> summary(backward)

Call:

lm(formula = odevy$y ~ odevy$x1 + odevy$x2 + odevy$x3)

Residuals:

Min 1Q Median 3Q Max

-11.245 -4.899 -3.306 -0.673 60.167

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 4.134 4.626 0.894 0.373859

odevy$x1 10.510 1.723 6.101 2.53e-08 \*\*\*

odevy$x2 11.382 2.117 5.376 5.85e-07 \*\*\*

odevy$x3 -6.031 1.670 -3.612 0.000497 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 9.298 on 91 degrees of freedom

Multiple R-squared: 0.7801, Adjusted R-squared: 0.7728

F-statistic: 107.6 on 3 and 91 DF, p-value: < 2.2e-16

**\*Pr değerlerine baktığımızda 3 değişken de 0.05 ten küçük olduğundan, anlamlı olduğu yorumunu yapabiliriz.**

**\*Sabit terim 0.05 ten büyük olduğundan anlamlı olduğu söylenemez.**

**\*p-value<2.2e-16 değeri de, modelin genel olarak anlamlı olduğunu belirtir.**

**\*y bağımlı değişkeni x1,x2 ve x3 tarafından %78 oranında açıklanabilmektedir.**

12.3-ADIMSAL SEÇİM YÖNTEMİ

**library(MASS)**

**step.model <- stepAIC(sonuc, direction = "both", trace = FALSE)**

**summary(step.model**)

Call:

lm(formula = odevy$y ~ odevy$x1 + odevy$x2 + odevy$x3)

Residuals:

Min 1Q Median 3Q Max

-11.245 -4.899 -3.306 -0.673 60.167

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 4.134 4.626 0.894 0.373859

odevy$x1 10.510 1.723 6.101 2.53e-08 \*\*\*

odevy$x2 11.382 2.117 5.376 5.85e-07 \*\*\*

odevy$x3 -6.031 1.670 -3.612 0.000497 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 9.298 on 91 degrees of freedom

Multiple R-squared: 0.7801, Adjusted R-squared: 0.7728

F-statistic: 107.6 on 3 and 91 DF, p-value: < 2.2e-16

**\*Pr değerlerine baktığımızda 3 değişken de 0.05 ten küçük olduğundan, anlamlı olduğu yorumunu yapabiliriz.**

**\*Sabit terim 0.05 ten büyük olduğundan anlamlı olduğu söylenemez.**

**\*p-value<2.2e-16 değeri de, modelin genel olarak anlamlı olduğunu belirtir.**

**\*y bağımlı değişkeni x1,x2 ve x3 tarafından %78 oranında açıklanabilmektedir.**

13-UYUM KESTİRİMİ VE %95 GÜVEN DÜZEYİ

**predict(sonuc, data.frame(x1=3.684, x2=2.147, x3=1.580 ), interval = 'prediction')**

fit lwr upr

1 4.811803 -28.78428 38.40789

**predict(sonuc, data.frame(x1=3.684, x2=2.147, x3=1.580 ), interval = 'confidence',level = .95)**

fit lwr upr

1 4.811803 -26.79295 36.41656

14-ÖN KESTİRİM VE %95 GÜVEN DÜZEYİ

**predict(sonuc, data.frame(x1=2.556, x2=4.628, x3=9.325 ), interval = 'confidence',level = .95)**

fit lwr upr

1 -7.21136 -48.57553 34.15281

**predict(sonuc, data.frame(x1=2.556, x2=4.628, x3=9.325), interval = 'prediction')**

fit lwr upr

1 -7.21136 -50.11627 35.69355