

Rastgele Blok Tasarımı 2. Uygulama

Fatih Kızılaslan

01 05 2020

Rastgele Blok Tasarım: Etkileşimli Model ve Rastgele Etkili Model

Örnek 1: (Rastgele Etkili Model)

Bir hastalığın tedavisinde kullanılan 4 farklı ilaç ve bu tedavide kullanılan 5 farklı dozu vardır. Rastgele seçilen 3 farklı doz (günde 50mg, 400 mg ve 150 mg gibi) için bu 4 ilaç farklı hastalara uygulanmış ve iyileşme süreleri (saat) aşağıdaki gibi elde edilmiştir.

| ## | | 1. Doz | 2. Doz | 3. Doz |
|------|-------|--------|--------|--------|
| ## A | İlacı | 27 | 14 | 13 |
| ## B | İlacı | 22 | 5 | 9 |
| ## C | İlacı | 34 | 11 | 18 |
| ## D | İlacı | 15 | 7 | 10 |

$\alpha = 0.05$ olmak üzere bu veri için

- Rastgele tam blok tasarımı kullanarak oluşturduğunuz modeli yazınız.
- Bu model için ANOVA tablosunu oluşturunuz.
- Kullanılan ilaçların hastaların iyileşme süreleri üzerindeki etkileri arasında fark olup olmadığını test ediniz.
- İlaç dozlarının iyileşme süresi üzerindeki etkileri arasında fark olup olmadığını test ediniz.
- Modelin varsayımlarını kontrol ediniz.
- Varyans bileşenlerinin tahmin edicilerini bulunuz.

ÇÖZÜM:

a) Bu veride 4 farklı ilacın iyileşme süresi üzerindeki etkisi araştırılmak isteniyor. Bu nedenle ana faktörümüz ilaç türleri ve ilacın uygulama dozları ise blok faktörümüz olur. Bloklar rastgele belirlendiği için rastgele etkili ve faktör düzeyleri ise sabit etkilidir. Böylece modelimizi

$$y_{ij} = \mu + \tau_i + \beta_j + \epsilon_{ij}, \quad i = 1, \dots, a, \quad j = 1, \dots, b$$

biçiminde yazarız. Burada, $\epsilon_{ij} \sim N(0, \sigma^2)$ ve $\beta_j \sim N(0, \sigma_\beta^2)$ dir.

Veriyi faktörlere uygun olarak aşağıdaki gibi tanımlarız.

Not: Tanımlama işleminden sonra verinizin faktör atamaları ile orjinal verideki blok ve düzeyleri karşılaştırarak kontrol ediniz.

```
y<-c(27,22,34,15,14,5,11,7,13,9,18,10)
```

```
ilac<-factor(rep(1:4, times = 3)) # İlaç türleri  
blok<- factor(rep(1:3, each = 4)) # 3 blok: ilaç dozları  
data<- data.frame(y, ilac,blok)  
print(data)
```

```
##      y ilac blok  
## 1  27    1    1  
## 2  22    2    1  
## 3  34    3    1  
## 4  15    4    1  
## 5  14    1    2  
## 6   5    2    2  
## 7  11    3    2  
## 8   7    4    2  
## 9  13    1    3  
## 10  9    2    3  
## 11 18    3    3  
## 12 10    4    3
```

```
tapply(data$y,data$ilac,mean)
```

```
##      1      2      3      4  
## 18.00000 12.00000 21.00000 10.66667
```

```
tapply(data$y,data$blok,mean)
```

```
##      1      2      3  
## 24.50  9.25 12.50
```

b) ANOVA tablosu aşağıdaki gibi elde edilir.

```
anova<-aov(y~ ilac + blok, data = data)  
summary(anova)
```

```
##           Df Sum Sq Mean Sq F value Pr(>F)  
## ilac       3  216.2   72.08   5.805 0.03306 *  
## blok       2  516.2  258.08  20.785 0.00201 **  
## Residuals  6   74.5   12.42  
## ---  
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

c) İlaçların hastaların iyileşme süreleri üzerinde etkileri arasında karşılaştırma yapmak için

$H_0 : \tau_1 = \tau_2 = \tau_3 = \tau_4 = 0$ ve $H_1 : \text{En az bir } \tau_i \neq 0$

hipotezlerini test ederiz.

ANOVA tablosundan $p - \text{value} = 0.03306 < 0.05$ olduğundan H_0 hipotezi red edilir. İlaçların etkileri arasında anlamlı bir farklılık vardır.

d) İlaç dozları arasındaki farklılığın iyileşme süresi üzerindeki etkilerini karşılaştırmak için

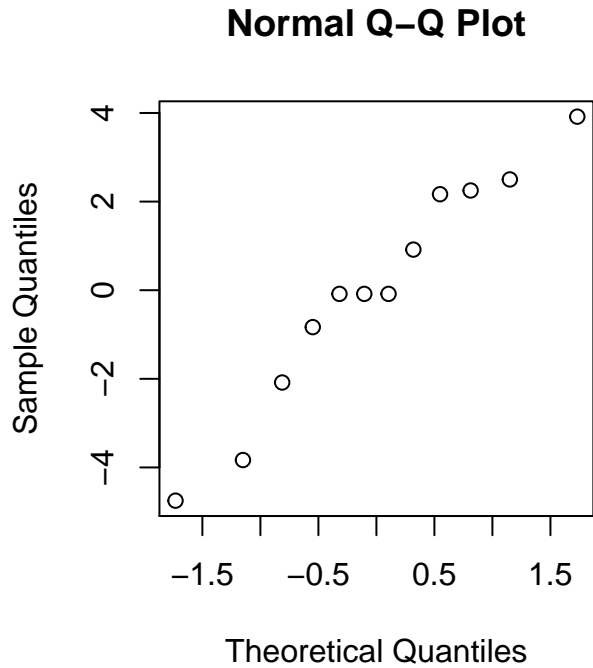
$$H_0 : \sigma_{\beta}^2 = 0 \text{ ve } H_1 : \sigma_{\beta}^2 > 0$$

hipotezlerini test ederiz.

ANOVA tablosundan $p\text{-value} = 0.00201 < 0.05$ olduğundan H_0 hipotezi red edilir. Bloklar yani ilaç dozları arasında anlamlı farklılık vardır.

e) Varsayım kontrolü: Aşağıdaki sonuçlardan varsayımların sağlandığı görülür.

```
qqnorm(residuals(anova))
```



```
shapiro.test(residuals(anova))
```

```
##  
## Shapiro-Wilk normality test  
##  
## data: residuals(anova)  
## W = 0.95527, p-value = 0.7148
```

```
ks.test(residuals(anova), "pnorm", mean(residuals(anova)), sd(residuals(anova)))
```

```
##  
## One-sample Kolmogorov-Smirnov test  
##  
## data: residuals(anova)  
## D = 0.15389, p-value = 0.8981  
## alternative hypothesis: two-sided
```

```

library(goftest)
ad.test(residuals(anova), "pnorm", mean=mean(residuals(anova)), sd=sd(residuals(anova)), estimated=TRUE)

##
## Anderson-Darling test of goodness-of-fit
## Braun's adjustment using 3 groups
## Null hypothesis: Normal distribution
## with parameters mean = -6.01732205729455e-17, sd = 2.60244640150902
## Parameters assumed to have been estimated from data
##
## data: residuals(anova)
## Anmax = 2.532, p-value = 0.1445

cvm.test(residuals(anova), "pnorm", mean=mean(residuals(anova)), sd=sd(residuals(anova)), estimated=TRUE)

##
## Cramer-von Mises test of goodness-of-fit
## Braun's adjustment using 3 groups
## Null hypothesis: Normal distribution
## with parameters mean = -6.01732205729455e-17, sd = 2.60244640150902
## Parameters assumed to have been estimated from data
##
## data: residuals(anova)
## omega2max = 0.11807, p-value = 0.8909

bartlett.test(y ~ data$ilac) # ilac

##
## Bartlett test of homogeneity of variances
##
## data: y by data$ilac
## Bartlett's K-squared = 1.6539, df = 3, p-value = 0.6472

library(car)

## Loading required package: carData

leveneTest(y, data$ilac) #medyana göre

## Levene's Test for Homogeneity of Variance (center = median)
##      Df F value Pr(>F)
## group 3  0.3108 0.8173
##      8

leveneTest(y, data$ilac, mean) #ortalamaya göre

## Levene's Test for Homogeneity of Variance (center = mean)
##      Df F value Pr(>F)
## group 3  1.3922 0.3139
##      8

```

f) Rastgele etki modeli için “lme4” paketindeki “lmer” fonksiyonunu kullanarak varyans bileşenleri σ^2 ve σ_β^2 tahmin edicileri buluruz.

```
library(lme4)
random_anova <- lmer(y ~ ilac+(1 | blok), data = data)
summary(random_anova)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: y ~ ilac + (1 | blok)
## Data: data
##
## REML criterion at convergence: 53.3
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -1.22398 -0.39828  0.01845  0.54769  1.23553
##
## Random effects:
## Groups   Name      Variance Std.Dev.
## blok    (Intercept) 61.42    7.837
## Residual                12.42    3.524
## Number of obs: 12, groups: blok, 3
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept)   18.000     4.961    3.628
## ilac2         -6.000     2.877   -2.085
## ilac3          3.000     2.877    1.043
## ilac4        -7.333     2.877   -2.549
##
## Correlation of Fixed Effects:
##      (Intr) ilac2  ilac3
## ilac2 -0.290
## ilac3 -0.290  0.500
## ilac4 -0.290  0.500  0.500
```

Buradan, $\hat{\sigma}^2 = 12.42 = MS_E$ ve $\hat{\sigma}_\beta^2 = 61.42 = (MS_{Block} - MS_E)/a = (258.08 - 12.42)/4$ bulunur. Böylece, modeldeki toplam varyansın büyük bir kısmının bloklardaki farklılıklardan kaynaklandığı görülmektedir.

Örnek 2: (Eksik veri durumu)

Örnek 1'deki veri için

a) 2. ilacın 1. dozu ile ilgili veri yani y_{21} kayıp olsun. Bu durumda ANOVA tablosunu oluşturarak düzeyler ve bloklar için ilgili hipotezleri test ediniz.

b) 1. ilacın 3. dozu ile ilgili veri yani y_{13} kayıp olsun. Bu durumda ANOVA tablosunu oluşturarak düzeyler ve bloklar için ilgili hipotezleri test ediniz. (ÖDEV)

ÇÖZÜM: Bir $y_{ij} = x$ gözlemimiz eksik ise

$$\hat{x} = \frac{ay_i^* + by_j^* - y_{..}^*}{(a-1)(b-1)}$$

biçiminde verilen EKK tahmin edicisini kullanarak ve hatanın serbestlik derecesini 1 azaltarak analizimizi yapabiliriz.

a) $y_{21} = x_1$ olmak üzere verimiz aşağıdaki gibi olur.

```
##          1. Doz 2. Doz 3. Doz
## A İlacı "27"  "14"  "13"
## B İlacı "x1"  "5"   "9"
## C İlacı "34"  "11"  "18"
## D İlacı "15"  "7"   "10"
```

Bu durumda kayıp gözlemimiz için EKK tahmin edicisi

$$\widehat{x}_1 = \frac{ay_2^* + by_1^* - y_3^*}{(a-1)(b-1)} = (4 * (14) + 3 * (76) - 163) / (3 * (2)) = 20.16667$$

olarak bulunur. Böylece verimiz bu tahmin değerini kullanarak aşağıdaki gibi olur.

```
y1<-c(27,20.16667,34,15,14,5,11,7,13,9,18,10)

ilac<-factor(rep(1:4, times = 3)) # İlaç türleri
blok<- factor(rep(1:3, each = 4)) # 3 blok: ilaç dozları
data1<- data.frame(y1, ilac,blok)
print(data1)
```

```
##          y1 ilac blok
## 1  27.00000    1    1
## 2  20.16667    2    1
## 3  34.00000    3    1
## 4  15.00000    4    1
## 5  14.00000    1    2
## 6   5.00000    2    2
## 7  11.00000    3    2
## 8   7.00000    4    2
## 9  13.00000    1    3
## 10  9.00000    2    3
## 11 18.00000    3    3
## 12 10.00000    4    3
```

Bu durumda, y1 verisi için ANOVA sonuçları aşağıdaki gibi olur.

```
anova1<-aov(y1~ ilac + blok, data = data1)
summary(anova1)
```

```
##          Df Sum Sq Mean Sq F value Pr(>F)
## ilac      3  229.6   76.54   6.307 0.02763 *
## blok      2  483.4  241.71  19.916 0.00224 **
## Residuals  6   72.8   12.14
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Ancak, kayıp gözlem için tahmin edici kullandığımızdan hatanın (Residuals) serbestlik derecesi $6-1=5$ olacak.

Bu sonuçlara göre $SS_{Deneme} = 229.6$, $SS_{Blok} = 483.4$ ve $SS_E = 72.8$ olduğundan $MS_E = 72.8/5 = 14.56$,

$$F_{ilac} = MS_{ilac}/MS_E = 76.54/14.56 = 5.256868 \text{ ve}$$

$$F_{blok} = MS_{blok}/MS_E = 241.71/14.56 = 16.60096$$

olarak bulunur.

$F_{(a-1), N-a-b, 0.05} = F_{3,5,0.05} = qf(1 - 0.05, 3, 5) = 5.409451$ olduğundan $F_{ilac} = 5.256868 < 5.409451 = F_{3,5,0.05}$ ve $p - value = 1 - pf(5.256868, 3, 5) = 0.05268142 > 0.05$ bulunur.

Böylece, $H_0 : \tau_1 = \tau_2 = \tau_3 = \tau_4 = 0$ hipotezi kabul edilir.

Yani, faktör düzeyleri olarak kullanılan 4 ilacın hastaların iyileşme süreleri üzerindeki etkileri arasında anlamlı bir farklılık yoktur.

$F_{(b-1), N-a-b, 0.05} = F_{b,5,0.05} = qf(1 - 0.05, 2, 5) = 5.786135$ olduğundan $F_{blok} = 16.60096 > 5.786135 = F_{2,5,0.05}$ ve $p - value = 1 - pf(16.60096, 2, 5) = 0.03954868 < 0.05$ bulunur.

Böylece, $H_0 : \sigma_\beta^2 = 0$ hipotezi red edilir.

Böylece, bloklar ilac dozları arasında anlamlı bir farklılık vardır. Yani, iyileşme süresi üzerinde dozların etkisi önemlidir biçiminde de yorumlanabilir.

b) Yukarıdakilere benzer olarak çözüünüz. **ÖDEV**

Örnek 3: (Kaggle'dan sağlık verisi)

Kaggle'da <https://www.kaggle.com/ronitf/heart-disease-uci> web sayfasına paylaşılan kalp hastalığı ile ilgili veriyi kullanacağız.

Bu veride toplam 303 gözlem vardır. Kişilerin cinsiyet, yaş, göğüs ağrısı şiddeti, dinlenme durumundaki tansiyon değeri (trestbps), kolesterol değeri (chol) gibi veriler bulunmaktadır.

```
data2<- read.csv("heart.csv")
colnames(data2)[1]<-"age"
head(data2,10)
```

```
##      age sex cp trestbps chol fbs restecg thalach exang oldpeak slope ca thal
## 1    63  1  3     145  233   1      0     150     0     2.3    0  0    1
## 2    37  1  2     130  250   0      1     187     0     3.5    0  0    2
## 3    41  0  1     130  204   0      0     172     0     1.4    2  0    2
## 4    56  1  1     120  236   0      1     178     0     0.8    2  0    2
## 5    57  0  0     120  354   0      1     163     1     0.6    2  0    2
## 6    57  1  0     140  192   0      1     148     0     0.4    1  0    1
## 7    56  0  1     140  294   0      0     153     0     1.3    1  0    2
## 8    44  1  1     120  263   0      1     173     0     0.0    2  0    3
## 9    52  1  2     172  199   1      1     162     0     0.5    2  0    3
## 10   57  1  2     150  168   0      1     174     0     1.6    2  0    2
##      target
## 1         1
## 2         1
## 3         1
## 4         1
## 5         1
## 6         1
## 7         1
## 8         1
## 9         1
## 10        1
```

Kolesterol değeri ile göğüs ağrısı ve cinsiyet arasındaki ilişkiyi inceleyelim. Faktörleri aşağıdaki gibi tanımlarız.

```
chest_pain<-as.factor(data2$cp) #Faktör düzeyleri ağrı seviyeleri "0,1,2,3"  
levels(chest_pain)
```

```
## [1] "0" "1" "2" "3"
```

```
sex<-as.factor(data2$sex) #cinsiyet bloklama faktörü "0: Kadın, 1: Erkek"
```

```
tapply(data2$chol,data2$cp,mean) # göğüs ağrısına göre kolesterol ortalamaları
```

```
##          0          1          2          3  
## 250.1329 244.7800 243.1724 237.1304
```

Göğüs ağrısı şiddetlerinin kolesterol değerlerinin ortalamalarını karşılaştıralım.

```
anova1<-aov(chol~ chest_pain, data = data2) #tek yönlü ANOVA  
summary(anova1)
```

```
##                Df Sum Sq Mean Sq F value Pr(>F)  
## chest_pain      3   5001    1667   0.618  0.604  
## Residuals     299 806300    2697
```

Görüldüğü gibi ağrı şiddetlerinin kolesterol değerlerinin ortalamaları arasında anlamlı bir fark yoktur.

Şimdi, göğüs ağrısı şiddeti yerine cinsiyeti kullanalım. Bu durumda

```
tapply(data2$chol,data2$sex,mean)
```

```
##          0          1  
## 261.3021 239.2899
```

```
anova2<-aov(chol~ sex, data = data2) #tek yönlü ANOVA  
summary(anova2)
```

```
##                Df Sum Sq Mean Sq F value Pr(>F)  
## sex              1  31778    31778  12.27 0.00053 ***  
## Residuals     301 779523    2590  
## ---  
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Bu sonuçlara göre kadın ve erkeklerin kolesterol değerleri arasında anlamlı bir fark vardır.

Kullandığımız bu iki faktörü birleştirelim. Yani, cinsiyeti blok faktörü olarak rastgele blok tasarımı oluşturalım. Bu durumda

```
anova3<-aov(chol~ chest_pain+sex, data = data2) #Blok tasarım  
summary(anova3)
```

```
##                Df Sum Sq Mean Sq F value Pr(>F)  
## chest_pain      3   5001    1667   0.643 0.588070  
## sex              1  33443    33443  12.895 0.000385 ***  
## Residuals     298 772857    2593  
## ---  
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```


Yukarıdakilere benzer sonuçlar alırız. Bu sonuçlara göre göğüs ağrısı şiddetlerine göre kolesterol değerlerinin ortalamaları arasında anlamlı bir fark yoktur. Bloklar arasında anlamlı bir fark vardır.

Bu verideki yaş değerlerini de bir faktör olarak kullanabiliriz. Görüldüğü gibi en küçük yaş 29 ve en büyük ise 77'dir.

```
summary(data2$age)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      29.00  47.50   55.00   54.37  61.00   77.00
```

29-77 arasındaki yaşları 4 farklı gruba ayırarak yaş faktörümüzü aşağıdaki gibi oluşturalım. Bu durumda veri aşağıdaki gibi olur.

```
data3<- data2
```

```
data3$age[which(data3$age>=29& data3$age<41)]=0
data3$age[which(data3$age>=41& data3$age<53)]=1
data3$age[which(data3$age>=53& data3$age<65)]=2
data3$age[which(data3$age>=65& data3$age<78)]=3
```

```
age1<-as.factor(data3$age)
```

```
head(data3,10)
```

```
##      age sex  cp  trestbps  chol  fbs  restecg  thalach  exang  oldpeak  slope  ca  thal
## 1     2   1   3     145    233   1         0     150     0     2.3    0  0    1
## 2     0   1   2     130    250   0         1     187     0     3.5    0  0    2
## 3     1   0   1     130    204   0         0     172     0     1.4    2  0    2
## 4     2   1   1     120    236   0         1     178     0     0.8    2  0    2
## 5     2   0   0     120    354   0         1     163     1     0.6    2  0    2
## 6     2   1   0     140    192   0         1     148     0     0.4    1  0    1
## 7     2   0   1     140    294   0         0     153     0     1.3    1  0    2
## 8     1   1   1     120    263   0         1     173     0     0.0    2  0    3
## 9     1   1   2     172    199   1         1     162     0     0.5    2  0    3
## 10    2   1   2     150    168   0         1     174     0     1.6    2  0    2
##      target
## 1         1
## 2         1
## 3         1
## 4         1
## 5         1
## 6         1
## 7         1
## 8         1
## 9         1
## 10        1
```

Böylece orjinal veriyi **data2** de tutarak, **data3** de **age1** 4 faktör düzeyine sahip bir değişken olarak tanımlanmış oldu.

data3 verisi için diğer faktörleri tanımlayalım.

```
chest_pain1<-as.factor(data3$cp) #Faktör düzeyleri ağırlık seviyeleri "0,1,2,3"
sex1<-as.factor(data3$sex) #cinsiyet bloklama faktörü "0: Kadın, 1: Erkek"
```

Yaşa göre kolesterol değerleri ortalamaları için ANOVA.

```
anova4<-aov(chol~ age1, data = data3)
summary(anova4)
```

```
##              Df Sum Sq Mean Sq F value Pr(>F)
## age1          3  42000    14000   5.441 0.00118 **
## Residuals    299 769301     2573
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Yaş ve cinsiyet faktörlerine göre kolesterol değerlerinin ortalamaları için ANOVA.

```
anova5<-aov(chol~ age1+sex1, data = data3) #Blok tasarım
summary(anova5)
```

```
##              Df Sum Sq Mean Sq F value Pr(>F)
## age1          3  42000    14000   5.612 0.000936 ***
## sex1          1  25937    25937  10.398 0.001402 **
## Residuals    298 743364     2495
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Bu sonuca göre yaş faktörünün düzeylerinin kolesterol değerleri arasında anlamlı bir fark vardır. Son iki anova sonuçlarında (anova4 ve anova5) cinsiyet ile bloklama yaparak hata kareler ortalaması MS_E azalmıştır. Ayrıca, anova5 sonucunda age1 için p-değeri anova4 den daha küçüktür. Böylece bloklama faktörünün analize olumlu katkısı olduğunu söyleyebiliriz.

Şimdi, yaş ile cinsiyet faktörünün etkileşimini de ekleyelim.

```
anova6<-aov(chol~ age1+sex1+age1:sex1, data = data3) #Blok tasarım etkileşimli
summary(anova6)
```

```
##              Df Sum Sq Mean Sq F value Pr(>F)
## age1          3  42000    14000   5.645 0.000897 ***
## sex1          1  25937    25937  10.459 0.001359 **
## age1:sex1     3  11798     3933   1.586 0.192874
## Residuals    295 731566     2480
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Bu sonuca göre yaş ile cinsiyetin etkileşimlerinin kolesterol değerleri üzerindeki etkisi önemsizdir.

anova6 modeli için varsayımları kontrol edelim.

```
shapiro.test(residuals(anova6))
```

```
##  
## Shapiro-Wilk normality test  
##  
## data: residuals(anova6)  
## W = 0.96512, p-value = 1.118e-06
```

```
ks.test(residuals(anova6), "pnorm", mean(residuals(anova6)), sd(residuals(anova6)))
```

```
## Warning in ks.test(residuals(anova6), "pnorm", mean(residuals(anova6)), : ties  
## should not be present for the Kolmogorov-Smirnov test
```

```
##  
## One-sample Kolmogorov-Smirnov test  
##  
## data: residuals(anova6)  
## D = 0.052081, p-value = 0.3837  
## alternative hypothesis: two-sided
```

```
library(goftest)
```

```
ad.test(residuals(anova6), "pnorm", mean=mean(residuals(anova6)), sd=sd(residuals(anova6)), estimated=TRUE)
```

```
##  
## Anderson-Darling test of goodness-of-fit  
## Braun's adjustment using 17 groups  
## Null hypothesis: Normal distribution  
## with parameters mean = -5.4873368408737e-16, sd = 49.2179206022896  
## Parameters assumed to have been estimated from data  
##  
## data: residuals(anova6)  
## Anmax = 3.2253, p-value = 0.3093
```

```
cvm.test(residuals(anova6), "pnorm", mean=mean(residuals(anova6)), sd=sd(residuals(anova6)), estimated=TRUE)
```

```
##  
## Cramer-von Mises test of goodness-of-fit  
## Braun's adjustment using 17 groups  
## Null hypothesis: Normal distribution  
## with parameters mean = -5.4873368408737e-16, sd = 49.2179206022896  
## Parameters assumed to have been estimated from data  
##  
## data: residuals(anova6)  
## omega2max = 0.33724, p-value = 0.8505
```

Shapiro-Wilk testine göre hatalar normal dağılıma sahip değildir. Ancak, diğer testlere göre normallik varsayımı sağlanır.

```
bartlett.test(data3$chol ~ age1) # ilac
```

```
##  
## Bartlett test of homogeneity of variances  
##  
## data: data3$chol by age1  
## Bartlett's K-squared = 16.943, df = 3, p-value = 0.0007262
```

```
library(car)  
leveneTest(data3$chol,age1) #medyana göre
```

```
## Levene's Test for Homogeneity of Variance (center = median)  
##      Df F value Pr(>F)  
## group  3  1.7682 0.1533  
##      299
```

```
leveneTest(data3$chol,age1,mean) #ortalamaya göre
```

```
## Levene's Test for Homogeneity of Variance (center = mean)  
##      Df F value Pr(>F)  
## group  3  1.9075 0.1284  
##      299
```

Levene testine göre homojen varyanslılık varsayımı sağlanır.

ÖDEV: Sizde yukarıda tanımlanan göğüs ağrısı şiddeti, cinsiyet ve yaş faktörlerini kullanarak kolestrol değeri yerine dinlenme durumundaki tansiyon değeri **trestbps** veya maksimum kalp atışı değeri **thalach** kullanarak varyans analizleri yapınız.