Rastgele Etkili Model

Rastgele Etkili Model

Örnek

Bir tekstil atölyesinin çok sayıda dokuma tezgahı vardır. Her bir tezgahın dakikada aynı kumaş çıktısı sağladığı varsayılıyor. Bu varsayımı araştırmak için 5 tezgah rastgele seçiliyor ve çıktıları farklı zamanlarda ölçülüyor. Aşağıdaki veriler elde ediliyor.

```
##            çıktılar çıktılar çıktılar çıktılar çıktılar
## 1. Tezgah     1.80     1.77     1.90     1.60     1.72
## 2. Tezgah     1.90     1.72     1.91     1.72     1.60
## 3. Tezgah     1.91     1.77     1.90     1.80     1.77
## 4. Tezgah     1.80     1.80     1.80     1.77     1.72
## 5. Tezgah     1.90     1.80     1.77     1.68     1.80
```

Bu veri için

a) ANOVA tablosunu oluşturarak $H_0 : \sigma_\tau^2 = 0$ $H_1 : \sigma_\tau^2 \neq 0$ hipotezlerini test ederek yorumlayınız.

b) σ^2 ve σ_τ^2 için tahmin edicileri bulunuz.

c) σ^2, $\sigma_\tau^2/(\sigma^2 + \sigma_\tau^2)$ ve μ için %95 lik güzen aralıkları oluşturunuz.

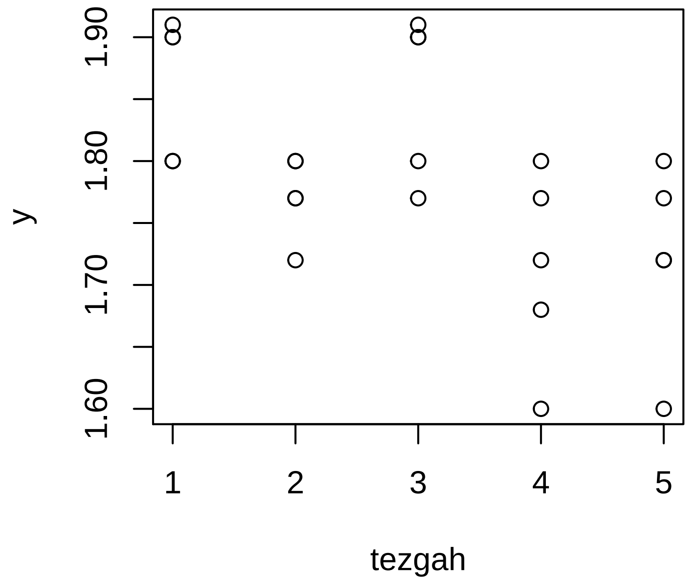
d) ANOVA'nın varsayımlarını kontrol ediniz.

ÇÖZÜM

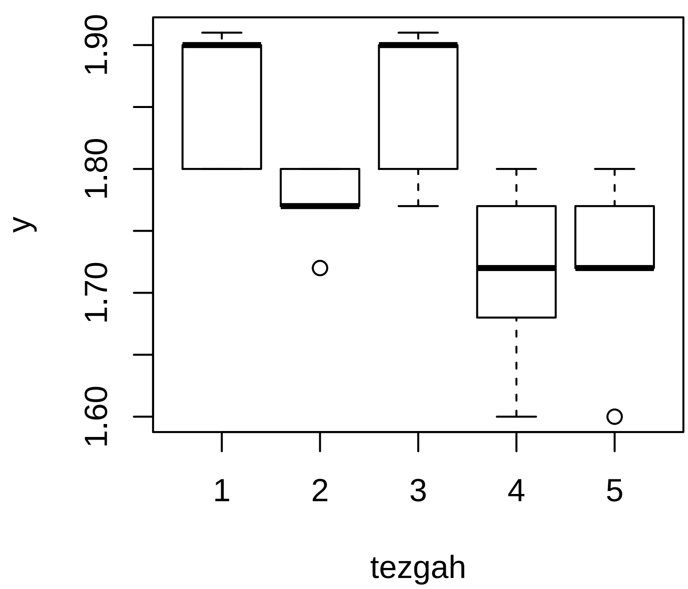
```
y<- c(1.80,1.90,1.91,1.80,1.90,1.77,1.72,1.77,1.80,1.80,1.90,1.91,1.90,1.80,1.77,1.60,1.72,1.80,1.77,1.
tezgah<- factor(rep(1:5, each= 5))
data<- data.frame(y,tezgah)
str(data)
```

```
## 'data.frame':    25 obs. of  2 variables:
##  $ y     : num  1.8 1.9 1.91 1.8 1.9 1.77 1.72 1.77 1.8 1.8 ...
##  $ tezgah: Factor w/ 5 levels "1","2","3","4",..: 1 1 1 1 1 2 2 2 2 2 ...
```

```
stripchart(y ~ tezgah, vertical = TRUE, pc=1, xlab = "tezgah")
```




```
boxplot(y ~ tezgah)
```



a) ANOVA tablosu sabit etkili modelde olduğu gibi oluşturulur.

```r
anova<-aov(y ~ tezgah)
summary(anova)
```

```
##             Df  Sum Sq  Mean Sq F value  Pr(>F)
## tezgah       4 0.10074 0.025186   6.107 0.00222 **
## Residuals   20 0.08248 0.004124
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

ANOVA tablosuna göre $p - value = 0.00222 < 0.05$ olduğundan $H_0 : \sigma_\tau^2 = 0$ hipotezi red edilir. Böylece, tezgahların kumaş çıktıları arasında anlamlı bir farklılık vardır.

b)

Rastgele etki modeli için "lme4" paketindeki "lmer" fonksiyonunu kullanacağız.

```r
library(lme4)
random_anova <- lmer(y ~ (1 | tezgah), data = data)
summary(random_anova)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: y ~ (1 | tezgah)
##    Data: data
##
## REML criterion at convergence: -53.2
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.0609 -0.7110 -0.0648  0.7875  1.1576
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  tezgah   (Intercept) 0.004212 0.06490
##  Residual             0.004124 0.06422
## Number of obs: 25, groups:  tezgah, 5
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept)  1.78520    0.03174   56.24
```

Rastgele etki modeli olduğu için lmer de "(1 | tezgah)" kullanırız. Farklı modeller için (mixed effect gibi) bu fonksiyon kullanılabilir.

Bu sonuca göre varyanslar için tahmin ediciler $\widehat{\sigma}_\tau^2 = 0.004212$ ve $\widehat{\sigma}^2 = 0.004124$ bulunur.

Ayrıca, $\widehat{\sigma}_\tau^2/(\widehat{\sigma}^2 + \widehat{\sigma}_\tau^2) = 0.004212/(0.004212 + 0.004124) = 0.5052783$ bulunur.

Bu oran bize tezgah türlerindeki farklılığın ürün çıktısındaki farklılığın ne kadarını açıkladığını söyler.

Böylece, kumaş çıktısındaki farklılığın %50,5 i tezgah türündeki farklılıktan kaynaklanmaktadır.

c)

confint(random_anova) ile tam olarak hesaplayamadığımız σ_τ^2 için yaklaşık güven aralığı bulunur.

```r
confint(random_anova)
```

```
##                  2.5 %    97.5 %
## .sig01      0.02181814 0.1353173
## .sigma      0.04848987 0.0907209
## (Intercept) 1.71694536 1.8534546
```

Sonuçtaki ilk satır σ_τ^2 için %95 lik yaklaşık güven aralığıdır ve $0.02181814 \leq \sigma_\tau^2 \leq 0.1353173$ bulunur.

Ayrıca, son satırdan (lineer modeldeki eğim katsayısı gibidir) μ için %95 lik güven aralığı $1.71694536 \leq \mu \leq 1.8534546$ olarak bulunur.

d)

Normallik varsayımı için artıkları kullanırız. Bu modeldeki artıklarımız aşağıdaki gibidir.

```r
residuals(anova)
```

```
##      1      2      3      4      5      6      7      8      9     10     11
## -0.062  0.038  0.048 -0.062  0.038 -0.002 -0.052 -0.002  0.028  0.028  0.044
##     12     13     14     15     16     17     18     19     20     21     22
##  0.054  0.044 -0.056 -0.086 -0.114  0.006  0.086  0.056 -0.034 -0.002 -0.122
##     23     24     25
##  0.048 -0.002  0.078
```

Aşağıda normallik için 5 farklı test uygulanmıştır. Kolmogorov-Smirnov, Shapiro Wilk, Liiliefor, Anderson-Darling ve Cramer-Von Mises testleri.

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

```
##
##  One-sample Kolmogorov-Smirnov test
##
## data:  residuals(anova)
## D = 0.16639, p-value = 0.4931
## alternative hypothesis: two-sided
```

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

```
##
##  Shapiro-Wilk normality test
##
## data:  residuals(anova)
## W = 0.92925, p-value = 0.08352
```

```r
library(nortest)
lillie.test(residuals(anova))
```

```
##
##  Lilliefors (Kolmogorov-Smirnov) normality test
##
## data:  residuals(anova)
## D = 0.16639, p-value = 0.07235
```

```r
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 5 groups
##  Null hypothesis: Normal distribution
##  with parameters mean = -3.60930903220424e-18, sd = 0.058623089876487
##  Parameters assumed to have been estimated from data
##
## data:  residuals(anova)
## Anmax = 1.3574, p-value = 0.7
```

```r
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 5 groups
##  Null hypothesis: Normal distribution
##  with parameters mean = -3.60930903220424e-18, sd = 0.058623089876487
##  Parameters assumed to have been estimated from data
##
## data:  residuals(anova)
## omega2max = 0.17298, p-value = 0.8688
```

Bu sonuçlara göre normallik varsayımı sağlanır.

Varyansların homejenliğini Bartlett ve Levene testleri ile kontrol edelim.

```r
bartlett.test(y ~ tezgah)
```

```
##
##  Bartlett test of homogeneity of variances
##
## data:  y by tezgah
## Bartlett's K-squared = 2.9051, df = 4, p-value = 0.5738
```

```r
library(car)
leveneTest(y, tezgah) #medyana göre
```

```
## Levene's Test for Homogeneity of Variance (center = median)
##       Df F value Pr(>F)
## group  4  0.3872 0.8152
##       20
```

```r
leveneTest(y, tezgah,mean) #ortalamaya göre
```

```
## Levene's Test for Homogeneity of Variance (center = mean)
##       Df F value Pr(>F)
## group  4  1.0158  0.423
##       20
```

Bu sonuçlara göre homojen varyanslılık varsayımı da sağlanmış olur.