

BÖLÜM 4

VARYANS ANALİZİ VE KARESEL FORMLAR

GİRİŞ

- Regresyon katsayı tahminleri, tahmin edilen ortalamalar ve artıklar, matris gösteriminde sunulmaktadır.
- Hepsi, orijinal \mathbf{Y} gözlemlerinin doğrusal fonksiyonlarıdır.
- Bu bölümde model, regresyon ve artık kareler toplamı ve doğrusal zıtlık karşılaştırması (kontrast) ya da doğrusal hipotez koleksiyonunun testi için kullanılan kareler toplamları, \mathbf{Y} 'nin karesel formları olarak gösterilmektedir.
- Bu, her kareler toplamının, $\mathbf{Y}' \mathbf{A} \mathbf{Y}$ şeklinde yazılabileceği anlamını taşımaktadır.
- Burada, \mathbf{A} tanımlayıcı matris olarak adlandırılan katsayılar matrisidir.
- $\mathbf{Y}' \mathbf{A} \mathbf{Y}$ 'nin karesel bir formu olarak ifade edilir.

KARESEL GÖSTERİMLERLE GİRİŞ

Öncelikle, önceki istatistiksel analiz derslerinden bildiğimiz doğrusal zıtlık karşılaştırmasına (kontrast) dayanan kareler toplamını ele alalım. Örneğin ilgilendiğiniz doğrusal zıtlık karşılaştırması aşağıdaki gibi olsun.

$$C_1^* = Y_1 + Y_2 - 2Y_3. \quad (4.1)$$

Bu zıtlık karşılaştırmasından kaynaklanan kareler toplamı

$$SS(C_1^*) = \frac{(C_1^*)^2}{6}. \quad (4.2)$$

şeklindedir. Formüldeki 6 böleni kontrastın katsayılarının kareler toplamıdır. Buradaki bölen, kareler toplamının beklenen değerindeki σ^2 'in katsayısını 1 yapmak için seçilmiştir. C_1^* 'i Y_i nin katsayısı $1/\sqrt{6}$ olacak şekilde yeniden tanımlarsak, kontrastıktan kaynaklanan kareler toplamı kontrastın karesidir. Böylece $C_1 = C_1^*/\sqrt{6}$, $\mathbf{a} = (1/\sqrt{6} \ 1/\sqrt{6} \ -2/\sqrt{6})'$ ve $\mathbf{Y} = (Y_1 \ Y_2 \ Y_3)'$ olacak şekilde tanımlayarak aşağıdaki matris gösterimindeki gibi yazılabilir.

KARESEL GÖSTERİMLERİN BEKLENEN DEĞERLERİ

- Y'nin varyans analizi de hesaplanan her karesel gösterim, model parametreleri fonksiyonlarını tahmin etmektedir.
- Eğer kareler toplamları ve onların ortalama kareleri doğru kullanılıyorsa, bu karesel formların beklenen değerleri mutlaka bilinmelidir.

KARESEL GÖSTERİMLERİN DAĞILIMI

Karesel gösterimlerin olasılık dağılımları parametrik anlamlılık testlerine temel oluşturur. Bu noktada ϵ_i 'ler üzerindeki normallik varsayımı devreye girer. Sonuçlar ϵ 'nin ve aynı zamanda Y 'nin normalliği varsayılarak özetlenir. Normallik sağlanmadığında, parametrik anlamlılık testleri yaklaşım olarak düşünülmelidir.

HİPOTEZ TESTLERİN GENEL GÖSTERİMİ

MS(Regr)'in MS(Res)'e oranı β_0 haricindeki tüm β_j 'ler aynı anda sıfırdan farklıdır sıfır hipotezini test eder. Hipotez testlerini kurarken bu prosedürlere izin verilenden daha çok esnekliğe ihtiyaç duyulur. Bu bölümde β 'nin fonksiyonlarını içeren bir hipotezi test etmek için kullanılan genel bir yöntem sunulmaktadır. Sıfır hipotezi tek bir doğrusal fonksiyon içerebilir, **basit hipotez**, ya da aynı zamanda birçok doğrusal hipotez içerebilir, **birleşik hipotez**.

GENEL DOĞRUSAL HİPOTEZLER

Genel doğrusal modeller

$$\begin{aligned}H_0 : \mathbf{K}'\boldsymbol{\beta} &= \mathbf{m} \\H_a : \mathbf{K}'\boldsymbol{\beta} &\neq \mathbf{m},\end{aligned}\tag{4.35}$$

şeklinde tanımlanır. Burada \mathbf{K}' test edilen β_j 'lerin k doğrusal fonksiyonunu tanımlayan $k \times p'$ boyutlu katsayılar matrisidir. \mathbf{K}' 'nin her satırı doğrusal bir fonksiyonun katsayılarını içerir. \mathbf{m} , $k \times 1$ boyutlu çoğunlukla sıfır olan sabitler vektörüdür. H_0 'daki k doğrusal denklem, doğrusal olarak bağımsız olmalıdır (fakat ortogonal olmak zorunda değildir).

GENEL GÖSTERİMLERİN ÖZEL DURUMLARI

Genel doğrusal hipotezlerin 3 özel durumu bizi ilgilendirmektedir.

Durum 1. Basit hipotez

β üzerinde basit hipotez test edildiğinde K' , tek bir satır vektörüdür ve böylece $[K'(X'X)^{-1}K]$ skaldır. Tersisi $1/[K'(X'X)^{-1}K]$ 'dir. Hipotezin kareler toplamı

$$Q = \frac{(K'\hat{\beta} - m)^2}{K'(X'X)^{-1}K} \quad (4.42)$$

şeklinde yazılır ve serbestlik derecesi 1'dir. Q 'nun payı $\hat{\beta}$ 'nin doğrusal fonksiyonunun karesidir ve paydası σ^2 haricindeki varyansdır. Böylece F -oranı

$$F = \frac{(K'\hat{\beta} - m)^2}{[K'(X'X)^{-1}K]s^2} \quad (4.43)$$

dir. Basit hipotezin F -testi iki kuyruklu (yanlı) t -testinin karesidir:

$$t = \frac{K'\hat{\beta} - m}{\{[K'(X'X)^{-1}K]s^2\}^{1/2}} \quad (4.44)$$

Payda, paydaki doğrusal fonksiyonun standart hatasıdır.

Durum 2. Belirli k tane β_j sifira eşittir

İlgilenilen hipotez belirli k tane regresyon katsayısının her birinin sifira eşit olduğudur. Bu durum için \mathbf{K}' , her satırda test edilen β_j 'yi tanımlayan giriş haricinde sıfırlar içeren $k \times p'$ boyutlu bir matristir, $\mathbf{m} = \mathbf{0}$. $[\mathbf{K}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{K}]$ matris çarpımı, buradaki \mathbf{K}' ile birlikte, test edilen β_j 'nin varyans-kovaryanları katsayılarını içeren $k \times k$ boyutlu $(\mathbf{X}'\mathbf{X})^{-1}$ alt matrisinden elde edilir. Örneğin, test edilen sıfır hipotezi, β_1, β_3 ve β_5 'in hepsi sifira eşittir hipotezi olsun. Q kareler toplamı

$$Q = (\hat{\beta}_1 \quad \hat{\beta}_3 \quad \hat{\beta}_5) \begin{bmatrix} c_{11} & c_{13} & c_{15} \\ c_{31} & c_{33} & c_{35} \\ c_{51} & c_{53} & c_{55} \end{bmatrix}^{-1} \begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_3 \\ \hat{\beta}_5 \end{pmatrix}, \quad (4.45)$$

formuna sahiptir. Eşitlikteki c_{ij} , $(\mathbf{X}'\mathbf{X})^{-1}$ matrisinin $(i+1)$ inci satırını ve $(j+1)$ inci sütunudur.

Hipotezin kareler toplamı, bu k tane bağımsız değişkenin başka bağımsız değişkenlerin de yer aldığı bir modele olan katkılarını ölçer. Buradaki kareler toplamı, bu k değişkenin modeldeki diğer bağımsız değişkenler için düzenlenmiş kareler toplamı olarak tanımlanır.

Durum 3. Bir β_j sıfıra eşittir; kısmi kareler toplamı

Üçüncü durum ilk iki durumun basitleştirilmesidir. Hipotez tek bir β_j 'nin sıfır olmasıdır, $H_0 : \beta_j = 0$. Bu hipotez için \mathbf{K}' , test edilen β_j 'ye ilişkin kolon haricinde diğer kolonları sıfır olan bir satır vektörüdür. Durum 2'de tanımlandığı gibi bu hipotezin kareler toplamı, modeldeki diğer bağımsız değişkenler için düzenlenen X_j 'nin modele olan katkısıdır. Bu kareler toplamına j inci bağımsız değişkeninin **kısmi kareler toplamı** denir.

Q 'daki $[\mathbf{K}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{K}]$ matris çarpımı $(\mathbf{X}'\mathbf{X})^{-1}$ 'den sadece $(j + 1)$ inci köşegen elemanı olan c_{jj} ile elde edilir. Bu $\hat{\beta}_j$ 'nin varyansının katsayısıdır. 1 serbestlik derecesindeki kareler toplamı

$$Q = \frac{\hat{\beta}_j^2}{c_{jj}}. \quad (4.46)$$

şeklinindedir. Bu herhangi bir bağımsız değişkenin kısmi kareler toplamının hesaplanmasını kolaylaştırır. Bu durum için iki kuyruklu t -testi

$$t = \frac{\hat{\beta}_j}{(c_{jj}s^2)^{1/2}}. \quad (4.47)$$

dir.

SAYISAL BİR ÖRNEK

Genel doğrusal hipotezin kullanımını örneklendirmek için N.C. State Üniversitesindeki fiziksel fitnes programından alınan veriler kullanılmıştır. (Veri A.C. Linnerud tarafından sağlanmıştır ve izin dahilinde kullanılmıştır.) Ölçümler 31 erkekten alınmıştır. Her kişinin yaş ve ağırlığına ek olarak, oksijen alımı (Y), koşu süresi (X_1), dinlenirkenki kalp atış hızı (X_2), koşarkenki kalp atış hızı (X_3) ve 1.5 mil koşarkenki maksimum kalp atış hızı (X_4) kaydedilmiştir. Veri, Tablo 4.3'te verilmiştir. Tartıştığımız sonuçlar oksijen alımı (Y) 'nin dört değişken (X_1 , X_2 , X_3 ve X_4) üzerindeki regresyon analizinin sonuçlarıdır.

TABLO 4.3. *North Carolina State Üniversitesi fiziksel fitnes programında yer alan 31 erkek üzerindeki fiziksel fitnes ölçümleri. (Ölçülen değişkenler yaş (yıl), ağırlık (kg), oksijen alım oranı (dakikada vücut ağırlığı aşına ml, 1.5 mil koşma zamanı (dakika), dinlenirken kalp artış hızı, koşarken kalp atış hızı (aynı zamanda oksijen alım ölçülmüştür) ve koşarken maksimum kalp artış oranı (A.C. Linnerud, N.C. State Üniversitesi, izniyle).*

| Yaş (yıl) | Ağırlık (kg) | O ₂ Alımı (ml/kg/dak.) | Zaman (dak.) | Kalp Atış Hızı | | |
|--------------|-----------------|--------------------------------------|-----------------|----------------|-------|----------|
| | | | | Dinlenme | Koşma | Maksimum |
| 44 | 89.47 | 44.609 | 11.37 | 62 | 178 | 182 |
| 40 | 75.07 | 45.313 | 10.07 | 62 | 185 | 185 |
| 44 | 85.84 | 54.297 | 8.65 | 45 | 156 | 184 |
| 42 | 68.15 | 59.571 | 8.17 | 40 | 166 | 172 |
| 38 | 89.02 | 49.874 | 9.22 | 55 | 178 | 180 |
| 47 | 77.45 | 44.811 | 11.63 | 58 | 176 | 176 |
| 40 | 75.98 | 45.681 | 11.95 | 70 | 176 | 180 |
| 43 | 81.19 | 49.091 | 10.85 | 64 | 162 | 170 |
| 44 | 81.42 | 39.442 | 13.08 | 63 | 174 | 176 |
| 38 | 81.87 | 60.055 | 8.63 | 48 | 170 | 186 |
| 44 | 73.03 | 50.541 | 10.13 | 45 | 168 | 168 |
| 45 | 87.66 | 37.388 | 14.03 | 56 | 186 | 192 |
| 45 | 66.45 | 44.754 | 11.12 | 51 | 176 | 176 |
| 47 | 79.15 | 47.273 | 10.60 | 47 | 162 | 164 |
| 54 | 83.12 | 51.855 | 10.33 | 50 | 166 | 170 |
| 49 | 81.42 | 49.156 | 8.95 | 44 | 180 | 185 |
| 51 | 69.63 | 40.836 | 10.95 | 57 | 168 | 172 |
| 51 | 77.91 | 46.672 | 10.00 | 48 | 162 | 168 |
| 48 | 91.63 | 46.774 | 10.25 | 48 | 162 | 164 |
| 49 | 73.37 | 50.388 | 10.08 | 67 | 168 | 168 |
| 57 | 73.37 | 39.407 | 12.63 | 58 | 174 | 176 |
| 54 | 79.38 | 46.080 | 11.17 | 62 | 156 | 176 |
| 52 | 76.32 | 45.441 | 9.63 | 48 | 164 | 166 |
| 50 | 70.87 | 54.625 | 8.92 | 48 | 146 | 186 |
| 51 | 67.25 | 45.118 | 11.08 | 48 | 172 | 172 |
| 54 | 91.63 | 39.203 | 12.88 | 44 | 168 | 172 |
| 51 | 73.71 | 45.790 | 10.47 | 59 | 186 | 188 |
| 57 | 59.08 | 50.545 | 9.93 | 49 | 148 | 160 |
| 49 | 76.32 | 48.673 | 9.40 | 56 | 186 | 188 |
| 48 | 61.24 | 47.920 | 11.50 | 52 | 170 | 176 |
| 52 | 82.78 | 47.467 | 10.50 | 53 | 170 | 172 |

TABLO 4.4. *Koşma zamanı, dinlenirkenki kalp atış hızı, koşarkenki kalp atış hızı ve maksimum kalp atış hızı üzerine oksijen alımının regresyonu için varyans analizinin özeti.*

| <i>Kaynak</i> | <i>s.d.</i> | <i>KT</i> | <i>OK</i> |
|---------------------------|-------------|-----------|----------------|
| Toplam _{bağımlı} | 30 | 851.3815 | |
| Regresyon | 4 | 658.2368 | 164.5659 |
| Artık | 26 | 193.1178 | 7.4276 = s^2 |

KARELER TOPLAMLARINI R-GÖSTERİMİ İLE ETİKETLENDİRİLMESİ

- Kısmi regresyon katsayılarının altkümesinin her birinin sıfır olduğu sıfır hipotezi, hem sıfır hipotezinde yer alan katsayılar altkümesine hem de modeldeki bütün katsayılar kümesine bağlıdır.
- Her durumda her ikisini de açıkça belirleyebilmek için daha güvenilir bir kareler toplamı gösterimine ihtiyaç duyulur.
- Bunun için sıkça kullanılan R-gösterimi ele alınmıştır.

TABLO 4.5. Oksijen alımı örneği için regresyon kareler toplamı, sıralı kareler toplamı ve artık kareler toplamı.

| | <i>Sıralı Kareler Toplamı</i> | <i>s.d.</i> | <i>F</i> |
|-----------|---|-------------|----------|
| KT(Regr)= | $R(\beta_1 \beta_3 \beta_2 \beta_4 \beta_0)=658.2638$ | 4 | 22.16 |
| | $R(\beta_1 \beta_0) = 632.9001$ | (1) | |
| | $R(\beta_3 \beta_0 \beta_1) = 15.3621$ | (1) | |
| | $R(\beta_2 \beta_0 \beta_1 \beta_3) = .4041$ | (1) | |
| | $R(\beta_4 \beta_0 \beta_1 \beta_3 \beta_2) = 9.5975$ | (1) | |
| KT (Hata) | $=193.1178$ | 26 | |

ÖRNEK: SIRALI VE KISMİ KARELER TOPLAMLARI

R -gösterimini, sıralı ve kısmi kareler toplamlarını örneklemek için Örnek 4.8'deki Oksijen alımı verisi kullanılmıştır. Tam modelin regresyon kareler toplamı dört serbestlik derecesi ile

$$SS(\text{Regr}) = R(\beta_1 \beta_2 \beta_3 \beta_4 | \beta_0) = 658.2638$$

dir (Tablo 4.4). X_1 , X_3 , X_2 ve X_4 sıralamasında uydurulan modelin sıralı kareler toplamı Tablo 4.5'te gösterilmektedir. Her sıralı kareler toplamı, bir bağımsız değişkenin modele eklenmesiyle modelde gerçekleşen aşamalı iyileşmeyi ölçer. Toplam regresyon kareler toplamına eklenen sıralı kareler toplamı $SS(\text{Regr}) = R(\beta_1 \beta_3 \beta_2 \beta_4 | \beta_0) = 658.2638$ 'dir ve bir ortogonal bölünmedir.

TABLO 4.6. Birikimli ardaşık kareler toplamı, sıfır hipotezi her bir birikimli kareler toplamı, oksijen örneğindeki sıfır hipotezinin F -testiyle sınanmaktadır.

| <i>Birikmiş Sıralı Kareler Toplamı</i> | <i>s.d.</i> | <i>Sıfır Hipotezi</i> | <i>F</i> |
|---|-------------|---|----------|
| $R(\beta_1 \beta_3 \beta_2 \beta_4 \beta_0) = 658.2638$ | 4 | $\beta_1 = \beta_3 = \beta_2 = \beta_4 = 0$ | 22.16 |
| $R(\beta_3 \beta_2 \beta_4 \beta_0 \beta_1) = 25.3637$ | 3 | $\beta_3 = \beta_2 = \beta_4 = 0$ | .67 |
| $R(\beta_2 \beta_4 \beta_0 \beta_1 \beta_3) = 10.0026$ | 2 | $\beta_2 = \beta_4 = 0$ | 1.14 |
| $R(\beta_4 \beta_0 \beta_1 \beta_3 \beta_2) = 9.5975$ | 1 | $\beta_4 = 0$ | 1.29 |
| SS(Error) = 193.1178 | 26 | | |

$$= 10.0016$$

$$= R(\beta_2 \beta_4 | \beta_0 \beta_1 \beta_3)$$

TABLO 4.7. Oksijen alımı örneği için kısmi kareler toplamı, her biri tarafından test edilen sıfır hipotezi ve sıfır hipotezinin F -testi.

| <i>Kısmi Kareler Toplamı</i> | <i>Sıfır Hipotezi</i> | F^a |
|--|-----------------------|-------|
| $R(\beta_1 \beta_0, \beta_2, \beta_3, \beta_4) = 397.8664$ | $\beta_1 = 0$ | 53.57 |
| $R(\beta_3 \beta_0, \beta_1, \beta_2, \beta_4) = 25.0917$ | $\beta_3 = 0$ | 3.38 |
| $R(\beta_2 \beta_0, \beta_1, \beta_3, \beta_4) = .0822$ | $\beta_2 = 0$ | .01 |
| $R(\beta_4 \beta_0, \beta_1, \beta_2, \beta_3) = 9.5975$ | $\beta_4 = 0$ | 1.29 |

^aTüm F -testleri tam modelden elde edilen kalıntı karelerle hesaplanmıştır.

TEK DEĞİŞKENLİ VE ORTAK GÜVEN BÖLGELERİ

- Parametrelerin güven aralıkları basit nokta tahminlerinin yapabileceğinden daha fazla bilgiyi okuyucuya iletir.
- Parametrelerin tek değişkenli güven aralıkları, parametre tahminleri arasındaki korelasyonu dikkate almaz.
- Buna ek olarak, bireysel güven aralıkları, ortak ifadelerdeki toplam güven derecelerini yansıtmazlar.
- Ortak güven bölgeleri bu iki noktayı ele alır.
- Ortak güven bölgelerine geçmeden önce tek değişkenli güven aralığı tahminlerinden kısaca bahsedilmiştir.

TEK DEĞİŞKENLİ GÜVEN ARALIKLARI

$$\hat{\beta}_j \pm t_{(\alpha/2, \nu)} s(\hat{\beta}_j), \quad j = 0, \dots, p, \quad (4.54)$$

dir. Burdaki $t_{(\alpha/2, \nu)}$, ν serbestlik derecesindeki student't dağılımıdır ve üst kuyruğa $\alpha/2$ olasılığını koyar. (Klasik çoklu regresyon modellerinde $\nu = (n-p')$ dir). $\hat{\beta}_j$ 'nin standard hatası $s(\hat{\beta}_j) = \sqrt{c_{jj}s^2}$ 'dir. Burada s^2 , ν serbestlik derecesiyle tahminlenmiştir ve c_{jj} , $(\mathbf{X}'\mathbf{X})^{-1}$ 'in $(j+1)$ 'inci köşegen elemanıdır.

Benzer şekilde, bağımsız değişkenlerin belirli değerleri için örneğin, $\mathbf{x}'_0 = (1 \ X_{01} \ \dots \ X_{0p})$, Y ortalamasının $\%(1-\alpha)100$ güven aralığı tahmini

$$\hat{Y}_0 \pm t_{(\alpha/2, \nu)} s(\hat{Y}_0), \quad (4.55)$$

dir. Burada, $\hat{Y}_0 = \mathbf{x}'_0 \hat{\boldsymbol{\beta}}$; $s(\hat{Y}_0) = \sqrt{\mathbf{x}'_0 (\mathbf{X}'\mathbf{X})^{-1} \mathbf{x}_0 s^2}$ ya da genel olarak $s(\hat{Y}_0) = \sqrt{v_{ii}s^2}$ dir. Buradaki \mathbf{x}'_0 , \mathbf{X} 'in i . satırını ve v_{ii} , \mathbf{P} 'nin i inci köşegen elemanını temsil eder. $t_{(\alpha/2, \nu)}$ Denklem 4.54'te tanımlanmıştır.

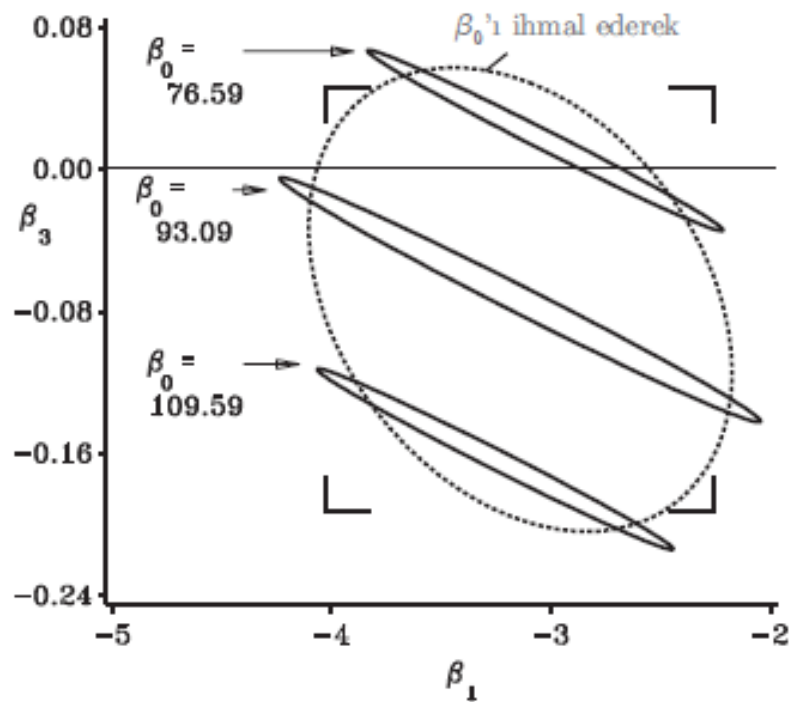
Bağımsız değişkenlerin belirli değerleri (örneğin, $\mathbf{x}'_0 = (1 \ X_{01} \ \dots \ X_{0p})$) için $Y_0 = \mathbf{x}'_0 \boldsymbol{\beta} + \epsilon$ 'in $\%(1-\alpha)100$ kestirim aralığı

$$\hat{Y}_0 \pm t_{(\alpha/2, \nu)} s(Y_0 - \hat{Y}_0), \quad (4.56)$$

EŞ ZAMANLI GÜVEN İFADELERİ

Klasik tek değişkenli güven aralıklarında her güven ifadesi için $(1 - \alpha) = .95$ güven katsayısı uygulanır. Eş zamanlı beş aralığın parametreleri içermelerine ilişkin güven düzeyi çok düşüktür. Beş aralık, istatistiksel olarak bağımsız olsalardı, ki değiller, bütün ya da ortak güven katsayısı $(1 - \alpha)^5 = .77$ olurdu.

Birçok eş zamanlı ifadenin ortak güven katsayısını önceden seçilen $(1 - \alpha)$ düzeyine yakın tutmak için iki prosedür vardır. En eski ve en basit prosedür, genelde **Bonferroni yöntemi** olarak adlandırılır. Bu yöntemle, Denklem 4.54 ve 4.55'teki gibi bireysel güven aralıkları oluşturulur; fakat, k eş zamanlı aralıklar ya da ifadeler için $\alpha^* = \alpha/k$ kullanılır. Yani, Denklem 4.54'teki $t_{(\alpha/2, \nu)}$, $t_{(\alpha/2k, \nu)}$ ile yer değiştirir. Bu prosedür, k eş zamanlı ifadenin gerçek ortak güven katsayısının en az $(1 - \alpha)$ olmasını garantiler.



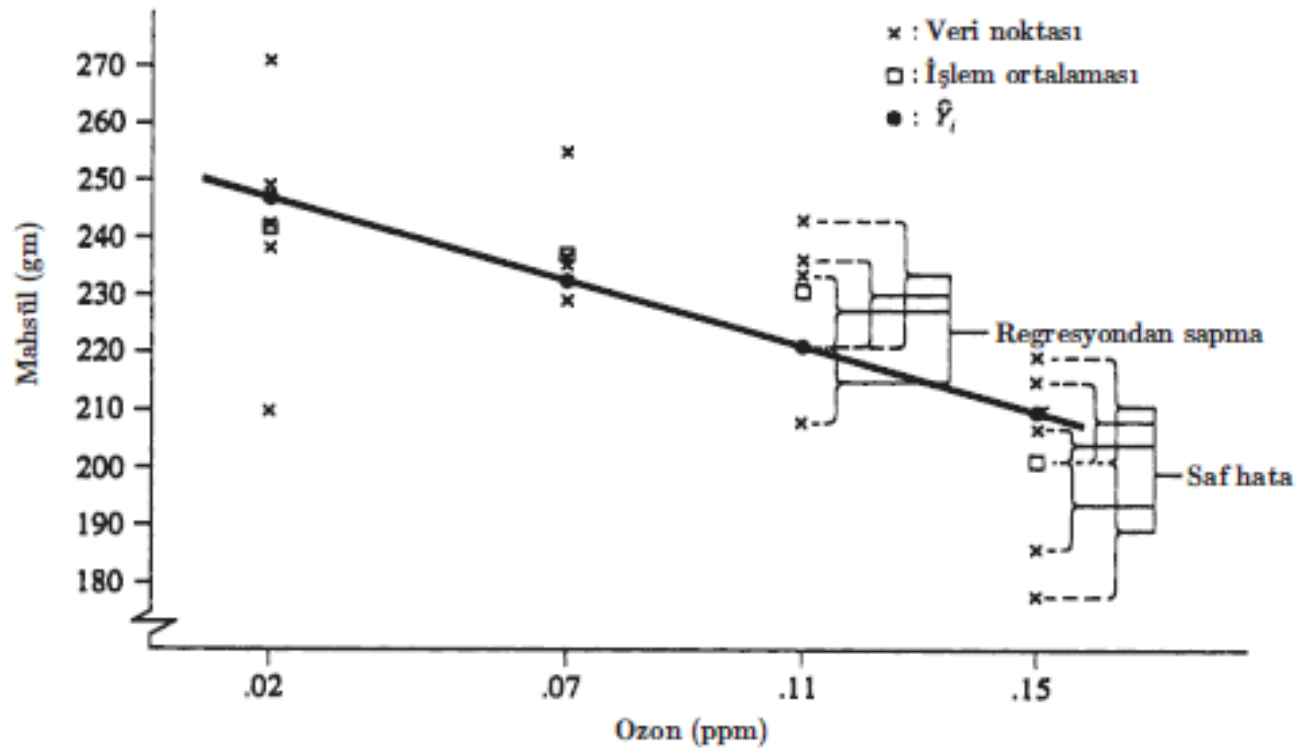
ŞEKİL 4.1. 1.5 mil koşma zamanı (X_1) ve koşarkenki kalp atış hızı (X_3) (kesintisiz elipsler) üzerine oksijen alımının regresyonu için ortak güven bölgesinin iki-boyutlu "dilimler"i ve β_0 ı göz ardı eden β_1 ve β_3 için iki boyutlu ortak güven bölgesi (kesikli elipsler). Bonferroni tek değişkenli güven aralıklarının kesişimi, kesişim tarafından oluşan dikdörtgenin köşeleri olarak gösterilir.

SAF HATA TAHMİNİ

Artık kare ortalaması bu zaman kadar σ^2 'in bir tahmini olarak kullanılmaktaydı. Bu prosedürün problemlerinden biri artık kare ortalamasının tahmin edilen modele bağımlı olmasıdır. Modeldeki bir yetersizlik, önemli bağımsız değişkenlerin göz ardı edilmesi ya da modelin yanlış bir formu, artık kare ortalamasının σ^2 'i aşırı tahminlemesine neden olacaktır. σ^2 'in bir tahmini, tahmin edilen modelin seçimine bağlı olmaması gerekir.

TABLO 4.8. Ozon kronik seviyelerine maruz kalan soya fasulyesinin tekrarlı mahsül verisi ve saf hatanın tahminleri (A.S. Heagle, North Carolina State Üniversitesi, izniyle).

| | Ozon seviyesi (ppm) | | | |
|-------------|---------------------|--------|--------|--------|
| | .02 | .07 | .11 | .15 |
| | 238.3 | 235.1 | 236.2 | 178.7 |
| | 270.7 | 228.9 | 208.0 | 186.0 |
| | 210.0 | 236.2 | 243.5 | 206.9 |
| | 248.7 | 255.0 | 233.0 | 215.3 |
| | 242.4 | 228.9 | 233.0 | 219.5 |
| \bar{Y}_i | 242.02 | 236.82 | 230.74 | 201.28 |
| s_i^2 | 476.61 | 114.83 | 179.99 | 325.86 |



ŞEKİL 4.2 Ozona karşı soya fasulyesi verisinin kullanarak "saf hata" ve "regresyondan sapmalar" karşılaştırılması .

TABLO 4.9. *Ozona karşılık soya fasulyesi mahsülü için tamamıyla rastgele deneysel tasarımın varyans analizi.*

| <i>Kaynak</i> | <i>s.d.</i> | <i>KT</i> | <i>OK</i> |
|--------------------|-------------|-----------|-----------|
| Toplam(bağlantılı) | 19 | 9366.61 | |
| İşlemler | 3 | 4977.47 | 1659.16 |
| Regresyon | 1 | 3956.31 | 3956.31 |
| Uydurma Eksikliği | 2 | 1021.16 | 510.58 |
| Saf Hata | 16 | 4389.14 | 274.32 |

Örnek 4.14'teki Ozon örneğinde, doğrusal modelin yeterlilik testi

$$F = \frac{\text{MS(Uyum eksikliği)}}{\text{MS(Saf hata)}} = \frac{510.58}{274.32} = 1.86,$$

dir ve model doğruysa, 2 ve 16 serbestlik derecesinde F dağılır. Kritik $F_{(.05,2,16)} = 3.63$ değeri ile karşılaştırıldığında hipotezin anlamlı olduğunu gösterir. Yani, bu veride ozona karşı soya fasulyesinin yanıtını temsil eden doğrusal modelin yetersiz olduğuna dair bir kanıt yoktur sonucuna varılır.

TABLO 4.10. Ozon verisinin varyans analizi.

| <i>Kaynak</i> | <i>s.d.</i> | <i>KT</i> | <i>OK</i> |
|--------------------------------|-------------|-----------|-----------|
| Toplam _(bağlantılı) | 19 | 9,366.61 | |
| Regresyon | 1 | 3,956.31 | 3,956.31 |
| Hata | 18 | 5,410.30 | 300.57 |
| Uydurma Eksikliği | 2 | 1,021.16 | 510.58 |
| Saf Hata | 16 | 4,389.14 | 274.32 |