BÖLÜM 4

4. ÇOK DEĞİŞKENLİ DOĞRUSAL REGRESYON VARYANS ANALİZİ VE KARESEL FORMLAR

Bölüm 2'de bir şans değişkeninin davranışını açıklamak amacıyla regresyon modelinin bir araç olarak kullanılabileceği belirtilmiş ve bir bağımsız değişkeni içeren model incelenmişti. Bununla birlikte tek bir bağımsız değişkeni içeren modeller genellikle şans değişkeninin davranışını yeterli miktarda açıklama yeteneğine sahip değildirler. Çünkü gerçek durum modelleri karmaşık bir yapıya sahiptirler. Gerçek durumu yansıtmaya çalışan modeller oluşturmak istendiğinde şans değişkeninin davranışı üzerinde etkisi olduğu düşünülen değişkenlerin en önemlilerinin modelde yer alması istenir. Sonuç olarak oluşturulan regresyon modeli, birden fazla bağımsız değişkeni içerir. Regresyon modelinde içerilen bağımsız değişken sayısı k ile belirtilecektir ve bu modeller çok değişkenli regresyon modeli olarak adlandırılacaktır. β_0 parametresine karşılık gelen ve tamamı 1 değerlerinden oluşan ilk sütun ise $\mathbf{x}_0$ sütun vektörü ile tanımlanacaktır. Sonuç olarak modeldeki parametre sayısı $k+1=p$ adettir. Konunun daha iyi anlaşılabilmesi için ilk olarak $k=2$ durumu incelenecek daha sonra genel model açıklanacaktır. Bu bölümden itibaren şans değişkenleri küçük italik harfler ile gösterilecektir.

4.1. İKİ BAĞIMSIZ DEĞİŞKENLİ REGRESYON MODELİ

Eşitlik (2.1) ile verilen modele ikinci bir açıklayıcı değişkenin eklenmesi ile

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \varepsilon_i \tag{4.1}$$

modeli elde edilir. Bu model için $E(\varepsilon_i)=0$ varsayımı ile beklenen yanıt

$$E(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 \tag{4.2}$$

elde edilir. Bu ifadeye verilen bir diğer isim de *regresyon fonksiyonudur*. Eşitlik (4.2) ile verilen regresyon fonksiyonu bir düzlem tanımlar bununla birlikte eşitlik (2.1) ile verilen basit regresyon fonksiyonu bir doğruyu belirtir. Çok değişkenli regresyonda, regresyon fonksiyonu *regresyon yüzeyi* (*surface*) ya da *yanıt yüzeyi* (*response surface*) olarak adlandırılır. Şekil 4.1'de örnek bir yanıt düzlemi

$$E(y) = 20 + 0.95x_1 + 0.50x_2$$

gösterilmiştir. Bu yanıt yüzeyi, basit bir düzlemdir. Bununla birlikte pek çok durum için yanıt yüzeyi daha karmaşık yapıda olabilir. Örnek yanıt düzlemindeki bir nokta bağımsız değişkenlerin verilen bir (x_1, x_2) değeri için ortalama yanıt $E(y)$ değerini verecektir.

Eşitlik (4.1) ile verilen regresyon modelinin parametrelerinden β_0 regresyon düzleminin Y eksenini kestiği noktayı tanımlar. Eğer modelin tanım aralığı $(x_1=0, x_2=0)$ noktasını içeriyor ise β_0 bu noktadaki ortalama yanıtı verir. Aksi durumda β_0 diğer parametrelerden farklı özel bir anlama sahip değildir.

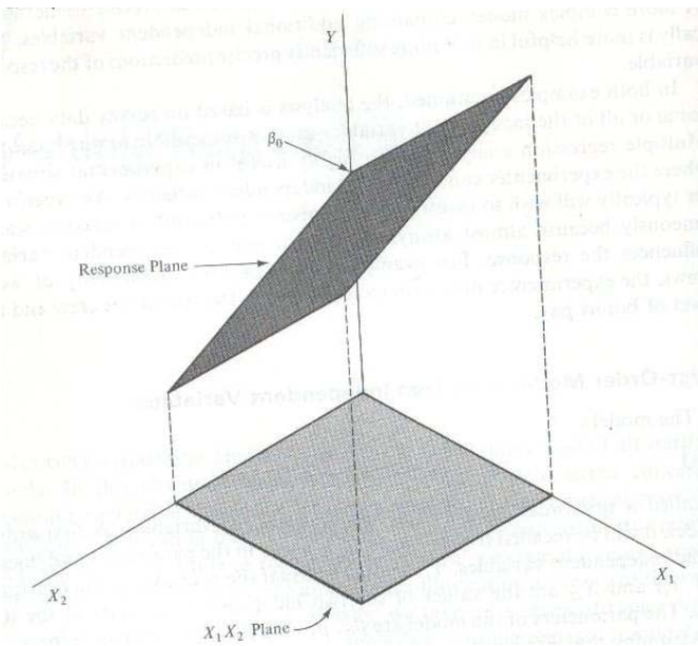
β_1 parametresi, x_2 değişkeninin değeri sabit tutulduğunda, x_1 değişkenindeki birim değişimin ortalama yanıtta oluşturduğu değişimi ifade eder. Benzer bir tanım β_2 için de yapılabilir. Verilen örnek fonksiyonunda $x_2=20$ alındığında, regresyon fonksiyonu,

$$E(y) = (20 - 10) + 0.95x_1$$

olup sağdaki ilk terim sabit olduğundan bir doğruyu tanımlar. x_2'nin sabitlenerek seviyenin değiştirilmesi kesişimi etkiler fakat eğimi etkilemez. Bu nedenle β_1 parametresi x_2 değeri sabit tutulduğunda x_1'deki birim değişimin ortalama yanıtta oluşturduğu değişikliği açıklamaktadır.

Ortalama yanıt üzerindeki x_1'in etkisinin x_2'nin değerlerine bağımlı olmadığı ve benzer şekilde x_2'nin etkisinin de x_1'in seviyesine bağılı olmadığı durumlarda bu iki bağımsız değişkenin etkileşimsiz olduğu ya da eklenebilir etkilere sahip olduğu söylenebilir.

β_1 ve β_2 parametreleri, modelde içerilen bir bağımsız değişken sabit iken diğer değişkenin kısmi etkisini tanımladıkları için kısmi regresyon parametreleri olarak da adlandırılırlar.



Şekil 4.1 *Regresyon düzlemi*

4.2 ÇOK DEĞİŞKENLİ DOĞRUSAL REGRESYON

Çok değişkenli regresyon modelinde bir y bağımlı değişkeni, k adet bağımsız değişkenin $\mathbf{x}_1, \ldots, \mathbf{x}_k$ doğrusal bir fonksiyonu olarak ifade edilmektedir. Burada $\mathbf{x}_j$ sütun vektörleri j-inci açıklayıcı değişkene ait gözlem verilerini içermektedir. Modele bir sabit terim eklenerek fonksiyon, i-inci gözlem için

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_k x_{ik} + \varepsilon_i \qquad (4.3)$$

şeklinde verilebilir. Eğer n adet gözlem mevcut ise model n adet denklemden oluşan bir set,

$$y_1 = \beta_0 + \beta_1 x_{11} + \beta_2 x_{12} + \ldots + \beta_k x_{1k} + \varepsilon_1$$

$$y_2 = \beta_0 + \beta_1 x_{21} + \beta_2 x_{22} + \ldots + \beta_k x_{2k} + \varepsilon_2$$

....

$$y_n = \beta_0 + \beta_1 x_{n1} + \beta_2 x_{n2} + \ldots + \beta_k x_{nk} + \varepsilon_n$$

şeklinde verilebilir. Bu n adet denklem matris ve vektörler formunda açık olarak,

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & X_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

ifade edilirler. Bu ifade daha kısa olarak,

$$\mathbf{y} = \mathbf{X\beta} + \boldsymbol{\varepsilon} \tag{4.4}$$

şeklinde verilebilir. Bağımlı değişken **y** ve hata terimleri **ε**, $n\times1$ boyutlu sütun vektörleridir. Katsayılar **β** ise k+1=p olmak üzere $p\times1$ boyutlu bir başka sütun vektörüdür. Bağımsız değişken matrisi **X** ise $n\times p$ boyutludur. Bu matrisin birinci sütunu bir sayılarından oluşmaktadır. Bu değerlerin β_0 sabit terimine karşılık gelen $\mathbf{x}_0$ değişkenini temsil ettikleri düşünülebilir. Eşitlik (4.4) matris notasyonunda çok değişkenli regresyon modelini tanımlar.

Eşitlik (4.4) ün sağ tarafında verilen $n\times p$ boyutlu bağımsız değişken **X** matrisi ile $p\times1$ boyutlu katsayılar **β** vektörünün çarpımı,

$$\mathbf{X\beta} = \begin{bmatrix} \beta_0 + \beta_1 x_{11} + \beta_2 x_{12} + \cdots + \beta_k x_{1k} \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \beta_0 + \beta_1 x_{n1} + \beta_2 x_{n2} + \cdots + \beta_k x_{nk} \end{bmatrix}$$

gerçekleştirilerek $n\times1$ boyutlu bir sütun vektörü elde edilir. Regresyon modeli için eğer (3.17) eşitliği, (4.4) eşitliğinin sağ tarafına uygulanırsa

$$\mathbf{X\beta} + \boldsymbol{\varepsilon} = \begin{bmatrix} \beta_0 + \beta_1 x_{11} + \beta_2 x_{12} + \cdots + \beta_k x_{1k} + \varepsilon_1 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \beta_0 + \beta_1 x_{n1} + \beta_2 x_{n2} + \cdots + \beta_k x_{nk} + \varepsilon_n \end{bmatrix}$$

$n\times1$ boyutlu bir sütun vektörü elde edilir. Bu vektörün elemanları boyutu $n\times1$ olan bağımlı değişken vektöründeki karşılık gelen elemanlara eşittir:

$$\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} \beta_0 + \beta_1 x_{11} + \beta_2 x_{12} + \cdots + \beta_k x_{1k} + \varepsilon_1 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \beta_0 + \beta_1 x_{n1} + \beta_2 x_{n2} + \cdots + \beta_k x_{nk} + \varepsilon_n \end{bmatrix} \tag{4.5}$$

Eşitlik (4.5)'den görüldüğü gibi **y** gözlem vektörü, **ε** hata terimi vektörü ile **X** sütunlarının doğrusal kombinasyonunun toplamıdır.

4.3 ÇOK DEĞİŞKENLİ REGRESYON İÇİN EKK TAHMİNLEME YÖNTEMİ

Eşitlik (4.4) ile tanımlanan genel doğrusal modeldeki bilinmeyen parametreler Bölüm 2 de açıklanan EKK tahminleme yöntemi kullanılarak elde edilebilir. Bu amaçla hata terimi,

$$\boldsymbol{\varepsilon} = \mathbf{y} - \mathbf{X\beta} \tag{4.6}$$

kareler toplamını,

$$\boldsymbol{\varepsilon}^T\boldsymbol{\varepsilon} = (\mathbf{y} - \mathbf{X\beta})^T(\mathbf{y} - \mathbf{X\beta}) \tag{4.7a}$$

$$\boldsymbol{\varepsilon}^T\boldsymbol{\varepsilon} = \mathbf{y}^T\mathbf{y} - \boldsymbol{\beta}^T\mathbf{X}^T\mathbf{y} - \mathbf{y}^T\mathbf{X\beta} + \boldsymbol{\beta}^T\mathbf{X}^T\mathbf{X\beta} \tag{4.7b}$$

minimize edecek parametre vektörü tahminlenmelidir. Eşitlik (4.7*a*) ile bulunan çarpım eşitlik (3.12) de tanımlanan bir iç çarpım,

$$\boldsymbol{\varepsilon}^T\boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 & \cdots & \varepsilon_n \end{bmatrix} \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix} = \varepsilon_1^2 + \ldots + \varepsilon_n^2 = \sum \varepsilon_i^2 \qquad (4.7c)$$

olması nedeniyle bir skalerdir. Hata kareler toplamını minimize edecek parametre tahminlerinin elde edilebilmesi için eşitlik (4.7*b*) nin, Kısım 3.4 de açıklandığı şekilde, $\boldsymbol{\beta}$ vektörüne göre türevi alınmalıdır. Eşitlik (4.7*b*) ile verilen tüm ifadeler birer skaler olduğundan transpozları kendilerine eşittir bu nedenle eşitliğin sağındaki ikinci ve üçüncü terimler için $(\mathbf{y}^T\mathbf{X}\boldsymbol{\beta})^T=\boldsymbol{\beta}^T\mathbf{X}^T\mathbf{y}$ olduğu görülebilir. Eşitlik (3.23) ve (3.24) den elde edilen iki sonuç eşitlik (4.7)'de tanımlanan hata terimi kareler toplamını minimize etmek amacıyla,

$$\frac{\partial\left(\boldsymbol{\varepsilon}^T\boldsymbol{\varepsilon}\right)}{\partial\boldsymbol{\beta}}$$

doğrudan kullanılabilir. Eşitlik (3.23) kullanılarak,

$$\frac{\partial(\mathbf{y}^T\mathbf{X}\boldsymbol{\beta})}{\partial\boldsymbol{\beta}} = \frac{\partial(\boldsymbol{\beta}^T\mathbf{X}^T\mathbf{y})}{\partial\boldsymbol{\beta}} = \mathbf{X}^T\mathbf{y}$$

bulunur. $\mathbf{X}^T\mathbf{y}$ tamamen bilinen k elemanlı bir vektördür. Eşitlik (3.24) kullanılarak,

$$\frac{\partial(\boldsymbol{\beta}^T\mathbf{X}^T\mathbf{X}\,\boldsymbol{\beta})}{\partial\boldsymbol{\beta}} = 2\;(\mathbf{X}^T\mathbf{X})\boldsymbol{\beta}$$

elde edilir. Bulunan iki sonuç birlikte ele alınarak,

$$\frac{\partial\left(\boldsymbol{\varepsilon}^T\boldsymbol{\varepsilon}\right)}{\partial\boldsymbol{\beta}} = -2\mathbf{X}^T\mathbf{y} + 2\mathbf{X}^T\mathbf{X}\boldsymbol{\beta}$$

eşitliği bulunur. Durağan bir nokta elde edebilmek için kısmi türevlerin her biri sıfıra eşit olmak zorunda olduğundan,

$$(\mathbf{X}^T\mathbf{X})\mathbf{b} = \mathbf{X}^T\;\mathbf{y} \qquad (4.8)$$

sonucu bulunur. Bulunan nokta bir minimuma karşılık gelir, bkz Alıştırma 4.1. Elde edilen eşitlik EKK için normal denklemleri verir. Bu denklem soldan $(\mathbf{X}^T\mathbf{X})^{-1}$ ile çarpılarak parametre tahminleyicileri vektörü,

$$\mathbf{b} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\;\mathbf{y} \qquad (4.9a)$$

bulunur. Eşitlik (4.9*a*) aynı zamanda parametre tahminlerinin doğrusal tahminleyiciler olduğunu ispatlamaktadır. Eşitlik (4.4), eşitlik (4.9*a*) da yerine konduğunda,

$$\mathbf{b} = \boldsymbol{\beta} + \left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\boldsymbol{\varepsilon} \qquad (4.9b)$$

EKK tahmincilerinin hataların doğrusal fonksiyonu oldukları bu nedenle dağılımlarının hataların dağılımına bağımlı olduğu görülmektedir. Bölüm 2 de verilen EKK tahminleme yöntemine ait varsayımların, hatalar ortalaması sıfır, sabit σ^2 varyanslı ve kovaryansları sıfır olan bir dağılama (normal) sahiptirler ifadesinin matris gösterimi,

$$E(\boldsymbol{\varepsilon})=\mathbf{0} \tag{4.10a}$$

$$\sigma^2(\boldsymbol{\varepsilon})=\sigma^2\mathbf{I} \tag{4.10b}$$

ya da kısaca:

$$\boldsymbol{\varepsilon} \sim N\left(\mathbf{0},\mathbf{I}\sigma^2\right) \tag{4.10c}$$

şeklindedir. Matris gösterimi basit regresyona da uygulanabilir: Tek değişkenli durum için $\mathbf{X}$ matrisi, ($\mathbf{X}^T\mathbf{X}$) matrisi, $\mathbf{X}^T\mathbf{y}$ vektörü ve $\mathbf{b}$ vektörü

$$\mathbf{X}=\begin{bmatrix} 1 & x_{11} \\ 1 & x_{12} \\ \vdots & \vdots \\ 1 & x_{1n} \end{bmatrix} \quad \mathbf{X}^T\mathbf{X}=\begin{bmatrix} n & \sum X_i \\ \sum X_i & \sum X_i^2 \end{bmatrix} \quad \mathbf{X}^T\mathbf{y}=\begin{bmatrix} \sum Y_i \\ \sum X_i Y_i \end{bmatrix} \quad \mathbf{b}=\begin{bmatrix} b_0 \\ b_1 \end{bmatrix}$$

şeklinde olup bu verilere göre eşitlik (4.8),

$$nb_0+b_1\sum X_i=\sum Y_i$$
$$b_0\sum X_i+b_1\sum X_i^2=\sum X_i Y_i$$

olarak elde edilebilir. Görüldüğü gibi bu denklemler normal denklemleri vermektedir.

4.4 ÇOK DEĞİŞKENLİ REGRESYON İÇİN EKK SONUÇLARI VE BEKLENEN DEĞERLER

Bu kısımda eşitlik (4.4) ile tanımlanan çok değişkenli regresyon modeli için kestirim değerleri ve artıklar matris gösterimi ile elde edilecek ve beklenen değerleri tanımlanacaktır. Parametre tahmin vektörü kullanılarak kestirilmiş değerler vektörü,

$$\hat{\mathbf{y}}=\mathbf{X}\mathbf{b} \tag{4.11}$$

Eşitlik (4.9*a*), eşitlik (4.10) da yerine konarak,

$$\hat{\mathbf{y}}=\left[\mathbf{X}\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\right]\mathbf{y} \tag{4.12}$$

köşeli parantez içindeki matris regresyon analizinde çok önemlidir ve *izdüşüm* matrisi ya da *şapka* (hat) matris olarak adlandırılır:

$$\mathbf{H}=\mathbf{X}\left(\mathbf{X}^\mathbf{T}\mathbf{X}\right)^{-1}\mathbf{X}^\mathbf{T} \tag{4.13a}$$

Eşitlik (4.13*a*) ile tanımlanan izdüşüm matrisi, simetrik ($\mathbf{H}^T$=$\mathbf{H}$) ve idempotent ($\mathbf{HH}$=$\mathbf{H}$) matristir. İzdüşüm matrisinin bir diğer önemli özelliği ise:

$$\mathbf{HX}=\mathbf{X}\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{X}=\mathbf{X} \tag{4.13b}$$

Sonuç olarak eşitlik (4.13*a*) eşitlik (4.12) de yerine konarak:

$$\hat{\mathbf{y}}=\mathbf{H}\mathbf{y} \tag{4.14}$$

elde edilir. Eşitlik (4.14) $\hat{\mathbf{y}}$ vektörünün, $\mathbf{y}$ şans vektörünün doğrusal bir fonksiyonu olduğunu ve bu fonksiyonun katsayılarının da $\mathbf{H}$ matrisi ile tanımlandığını belirtir. Örneğin $\mathbf{H}$ matrisinin birinci sırası $\hat{y}_1$ değerini veren tüm y_i değerlerinin doğrusal fonksiyonu için gerekli katsayıları verir.

Artık vektörü ise:

$$\mathbf{e} = \mathbf{y} - \hat{\mathbf{y}} \tag{4.15a}$$

$$\mathbf{e} = \mathbf{y} - \mathbf{Xb} \tag{4.15b}$$

$$\mathbf{e} = \mathbf{y} - \mathbf{Hy} = (\mathbf{I} - \mathbf{H})\mathbf{y} \tag{4.15c}$$

eşitliklerinden her hangi biri ile elde edilebilir. (**I-H**) matrisi de simetrik ve idempotent matristir. Hatalar ile artıklar arasındaki ilişkiyi belirlemek için eşitlik (4.15*c*) de $\mathbf{y}$ yerine eşitlik (4.4) konarak ve (4.13*b*) özelliği kullanılarak,

$$\mathbf{e} = (\mathbf{I} - \mathbf{H})\boldsymbol{\varepsilon} \tag{4.16}$$

bulunur.

Eşitlikler (4.9*a*), (4.14) ve (4.15c) den görüldüğü gibi regresyon sonuçları $\mathbf{b}$, $\hat{\mathbf{y}}$ ve $\mathbf{e}$ vektörleri şans değişkeni $\mathbf{y}$ vektörünün doğrusal fonksiyonlarıdır, bkz. Bölüm Ek 4. Bu nedenle kendileri de birer şans değişkenidir. Şans değişkenlerinin her biri, $\mathbf{y}$ vektörü normal dağılıma sahip olduğu için, normal dağılıma sahiptir. Normal dağılım gösteren şans değişkenlerinin beklenen değerleri aşağıda elde edilmiş olup varyansları Kısım 4.5 de incelenmiştir.

Model (4.4) için beklenen değer, bir şans vektörünün ya da matrisinin beklenen değeri için bkz. Bölüm Ek 4, $E(\mathbf{y})=E(\mathbf{X}\boldsymbol{\beta}+\boldsymbol{\varepsilon})= \mathbf{X}\boldsymbol{\beta}+E(\boldsymbol{\varepsilon})$ olup varsayımlar kullanılarak,

$$E(\mathbf{y})= \mathbf{X}\boldsymbol{\beta} \tag{4.17}$$

bulunur. Regresyon sonuçlarının beklenen değerleri ise parametre tahminleri için:

$$E(\mathbf{b}) = E\left[(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\ \mathbf{y}\right] = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T E\ (\mathbf{y})$$

$$E(\mathbf{b}) = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{X}\boldsymbol{\beta}$$

$$E(\mathbf{b}) = \boldsymbol{\beta} \tag{4.18}$$

Eşitlik (4.18) EKK tahmincilerinin sapmasızlığını ispatlamaktadır, minimum varyanslı olduklarının ispatı için bkz. Alıştırma 4.2. Bu durum sadece varsayılan modelin doğru olduğu durumlar için geçerlidir. Konu daha geniş anlamda ele alındığında sapmanın sadece kabul edilmiş doğru modele değil $\mathbf{X}$ değerlerine bağlı olduğu görülür. Başka bir deyişle tasarlanmış bir deney kullanıldığında sapma modele bağlı olduğu kadar deneysel tasarıma da bağlıdır. Bu konu kısım sonunda daha geniş olarak incelenecektir.

Kestirim için beklenen değer ise, eşitlik (4.11) ile,

$$E(\hat{\mathbf{y}}) = E(\mathbf{Xb}) = \mathbf{X}E(\mathbf{b})$$

ve eşitlik (4.18) ile,

$$E(\hat{\mathbf{y}}) = \mathbf{X}\boldsymbol{\beta} \tag{4.19}$$

bulunur. Artık vektörünün beklenen değeri, eşitlik (4.15*c*) ile

$$E(\mathbf{e}) = E\left[(\mathbf{I}-\mathbf{H})\mathbf{y}\right] = (\mathbf{I}-\mathbf{H})E(\mathbf{y})$$

$$E(\mathbf{e}) = (\mathbf{I}-\mathbf{H})\mathbf{X}\boldsymbol{\beta} = (\mathbf{X}-\mathbf{H}\mathbf{X})\boldsymbol{\beta} = (\mathbf{X}-\mathbf{X})\boldsymbol{\beta}$$

$$E(\mathbf{e}) = 0 \tag{4.20}$$

elde edilir. Başka bir deyişle artıklar sıfır ortalamalı şans değişkenleridir.

Regresyon tahminlerinde sapma; Eğer varsayılan model,

$$E(\mathbf{y}) = \mathbf{X}_1\boldsymbol{\beta} \tag{4.21}$$

ise EEK tahmini $\mathbf{b} = \left(\mathbf{X}_1^T\mathbf{X}_1\right)^{-1}\mathbf{X}_1^T\mathbf{y}$ olarak elde edilir ve $E(\mathbf{b}_1)=\boldsymbol{\beta}_1$ olup sapmasızdır. Fakat varsayılan model eşitlik (4.21) ile verilen model buna karşın gerçek model

$$E(\mathbf{y}) = \mathbf{X}_1\boldsymbol{\beta} + \mathbf{X}_2\boldsymbol{\beta}_2 \tag{4.22}$$

ise $\mathbf{X}_2\boldsymbol{\beta}_2$ terimi tahmin prosedürüne dahil değildir. Bu da **b** tahmininde sapmaya neden olur.

$$E(\mathbf{b}) = \left(\mathbf{X}_1^T\mathbf{X}_1\right)^{-1}\mathbf{X}_1^T E(\mathbf{y}) = \left(\mathbf{X}_1^T\mathbf{X}_1\right)^{-1}\mathbf{X}_1^T\left(\mathbf{X}_1\boldsymbol{\beta} + \mathbf{X}_2\boldsymbol{\beta}_2\right)$$

$$E(\mathbf{b}) = \left(\mathbf{X}_1^T\mathbf{X}_1\right)^{-1}\mathbf{X}_1^T\mathbf{X}_1\boldsymbol{\beta} + \left(\mathbf{X}_1^T\mathbf{X}_1\right)^{-1}\mathbf{X}_1^T\mathbf{X}_2\boldsymbol{\beta}_2$$

$$E(\mathbf{b}) = \boldsymbol{\beta} + \mathbf{A}\boldsymbol{\beta}_2 \tag{4.23a}$$

elde edilir. Burada $\mathbf{A} = \left(\mathbf{X}_1^T\mathbf{X}_1\right)^{-1}\mathbf{X}_1^T\mathbf{X}_2$ *sapma* ya da *eşyapı* matrisi olarak adlandırılır. Bu sapma sadece varsayılan model ile doğru modele değil, $\mathbf{X}_1$ ve $\mathbf{X}_2$ matrisleri arasındaki deneysel tasarıma bağlıdır. Tasarımın iyi seçilmesi daha az sapmalı tahminler elde edilmesine yardımcı olabilecektir.

Model (4.21) için kestirim değeri, $\hat{\mathbf{y}}=\mathbf{X}_1\mathbf{b}_1$ olup beklenen değeri,

$$E(\hat{\mathbf{y}}) = \mathbf{X}_1\boldsymbol{\beta} + \mathbf{X}_1\mathbf{A}\boldsymbol{\beta}_2 \tag{4.23b}$$

sonuç olarak kestirim değerinin de sapmalı olduğu görülür.

4.5 ÇOK DEĞİŞKENLİ REGRESYON İÇİN EKK SONUÇLARININ VARYANSLARI

Kısım 4.4 de şans vektörlerinin beklenen değerleri elde edilmiştir. Bu kısımda şans vektörlerinin varyans-kovaryans matrisleri elde edilmeye çalışılacaktır. İlk olarak Bölüm 2 de açıklanan şans değişkenlerinin doğrusal fonksiyonlarının varyansları, şans vektörleri için matris gösteriminde tekrar ele alınacaktır.

Varyans–kovaryans matrisi, bkz. Bölüm Ek 4, $V(\mathbf{y})$ olan $n\times1$ boyutlu bir şans vektörü **y** olsun. Bu şans vektörünün herhangi bir doğrusal fonksiyonu, bkz. Bölüm Ek 4, $\mathbf{w}= \mathbf{A}\mathbf{y}$ ile verilmiştir. Burada **A** boyutları $k\times n$ olan doğrusal fonksiyonunun katsayılarını belirten katsayılar matrisidir. Sonuç olarak yeni şans vektörü **w** için boyutlar $k\times1$ olarak tanımlanır.

Bir şans vektörünün doğrusal fonksiyonlarının varyansı matris işlemleri kullanılarak aşağıdaki gibi elde edilebilir. **y** şans vektörün varyans- kovaryans matrisi

$$V(\mathbf{y})=E\{[\mathbf{y}-E(\mathbf{y})]\,[\mathbf{y}-E(\mathbf{y})]^T\} \tag{4.24a}$$

şeklindedir. Eşitlik (4.24*a*)'daki çarpım $n \times n$ boyutlu bir matris oluşturur. Bu matrisin köşegen elemanları $[y_i - E(y_i)]^2$ ve köşegen dışı elemanları $[y_i - E(y_i)][y_j - E(y_j)]$ işlemleri ile elde edilir. Bu elemanların beklenen değerleri alınarak sırası ile varyans ve kovaryanslar bulunur. Varyans-kovaryans matrisinin bu tanımı, $\mathbf{w}=\mathbf{A}\mathbf{y}$ vektörünün varyansını elde etmek için kullanılabilir. $\mathbf{w}$ vektörü için varyans–kovaryans matrisi,

$$V(\mathbf{w})=E\{[\mathbf{w}-E(\mathbf{w})]\,[\mathbf{w}-E(\mathbf{w})]^T\}$$

olup $\mathbf{w}$ yerine $\mathbf{Ay}$ yazılıp, çarpım işlemleri yapılarak,

$$\begin{aligned} V(\mathbf{w}) &= E\{[\mathbf{Ay}-E(\mathbf{Ay})]\,[\mathbf{Ay}-E(\mathbf{Ay})]^T\} \\ &= E\{\mathbf{A}[\mathbf{y}-E(\mathbf{y})]\,[\mathbf{y}-E(\mathbf{y})]^T\mathbf{A}^T\} \\ &= \mathbf{A}E\{[\mathbf{y}-E(\mathbf{y})]\,[\mathbf{y}-E(\mathbf{y})]^T\}\mathbf{A}^T \end{aligned}$$

ve sonuç olarak şans değişkeni $\mathbf{w}$ 'nin varyansı $V(\mathbf{y})$' ye göre,

$$V(\mathbf{w}) = \mathbf{A}[V(\mathbf{y})]\mathbf{A}^T \qquad (4.24b)$$

olup eğer EEK'da varsayıldığı gibi $V(\mathbf{y}) = \mathbf{I}\sigma^2$ ise,

$$V(\mathbf{w}) = \mathbf{A}\mathbf{A}^T\sigma^2 \qquad (4.24c)$$

şeklinde elde edilir. $\mathbf{AA}^T$ matrisinin i-inci köşegen elemanı i-inci doğrusal fonksiyonun katsayılarının karelerinin toplamıdır. Bu elemanın σ^2 ile çarpılmasıyla i-inci doğrusal fonksiyonun varyansı elde edilebilir. Matrisin a_{ij} elemanı, i-inci ve j-inci doğrusal fonksiyonların katsayılarının çarpımlarının toplamını verir ve bu eleman σ^2 ile çarpılarak iki doğrusal fonksiyon arasındaki kovaryans elde edilir. Doğrusal fonksiyonların varyanslarının elde edilmesi iki örnek üzerinde açıklanmıştır, bkz Alıştırma 4.3 ve 4.4.

Doğrusal fonksiyonların varyansları için elde edilen bu genel sonuç $\mathbf{b}$, $\hat{\mathbf{y}}$ ve $\mathbf{e}$ şans değişkenlerinin varyans-kovaryans matrislerinin elde edilmesinde kullanılacaktır. $\mathbf{y}$'nin doğrusal bir fonksiyonu olan $\mathbf{b}$'nin katsayı matrisi $\mathbf{A}=(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T$ olduğu için tahminlenmiş regresyon parametrelerinin varyans-kovaryans matrisi,

$$V(\mathbf{b}) = \left[(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\right]V(\mathbf{y})\left[(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\right]^T$$

$$V(\mathbf{b}) = \left[(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\right]\left[(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\right]^T\sigma^2$$

$$V(\mathbf{b}) = (\mathbf{X}^T\mathbf{X})^{-1}\sigma^2 \qquad (4.25)$$

olarak elde edilir. Bu matrisin köşegen elamanları regresyon parametre tahminlerinin varyanslarını ve köşegen dışındaki elamanlar ise bu tahminler arasındaki kovaryansları verir.

$\hat{\mathbf{y}}$ için varyans-kovaryans matrisi eşitlik (4.11) veya (4.14) ilişkilerinden herhangi biri kullanılarak elde edilebilir. Doğrusal fonksiyonların varyansları için verilen kurallar eşitlik (4.14) e uygulanarak, burada $\mathbf{A}=\mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T=\mathbf{H}$,

$$V(\hat{\mathbf{y}}) = \mathbf{H}V(\mathbf{y})\mathbf{H}^T$$
$$V(\hat{\mathbf{y}}) = \mathbf{HH}^T\sigma^2$$

$$V(\hat{\mathbf{y}}) = \mathbf{H}\sigma^2 \tag{4.26a}$$

İkinci ilişki kullanılarak da yine aynı sonuç elde edilir, (**H** simetrik ve idempotent matristir). Uyumu yapılmış değerlerin herhangi bir alt setinin varyansı sadece karşılık gelen $\mathbf{X}_r$ sıraları kullanılarak ve eşitlik (4.11) ile

$$V(\hat{\mathbf{y}}_r) = \mathbf{X}_r V(\mathbf{b}) \mathbf{X}_r^T$$

$$V(\hat{\mathbf{y}}_r) = \mathbf{X}_r (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}_r^T \sigma^2 \tag{4.26b}$$

belirlenebilir. $\mathbf{H}\sigma^2$ ile verilen bu varyanslar, bağımsız değişkenlerin verilen seviyeleri için y ortalamalarını tahminlemekte kullanıldığında $\hat{\mathbf{y}}_i$ için uygun varyans değerlerini vermektedirler. Bağımsız değişkenlerin verilen $\mathbf{x}_0$ seviyelerinde, gelecekteki şans gözlemlerinin kestirim varyansı elde edilmek istendiğinde, kestirilecek değerin varyansına bireysel gözlemlerin varyansı olan σ^2 ilave edilmelidir. Bu durumda kestirim için varyans-kovaryans matrisi,

$$V(\hat{\mathbf{y}}_0) = (\mathbf{I} + \mathbf{H})\sigma^2 \tag{4.26c}$$

olarak elde edilir.

Artık vektörünün varyans-kovaryans matrisi, eşitlik (4.15*c*) kullanılarak, burada **A**=(**I**-**H**),

$$V(\mathbf{e}) = (\mathbf{I} - \mathbf{H})\sigma^2 \tag{4.27}$$

şeklinde elde edilir. Gözlenmiş artıkların varyansı σ^2 ye ve kovaryansları da sıfıra eşit değildir. Eşit varyanslılık ve sıfır kovaryans varsayımları e_i için değil ε_i için geçerlidir. Herhangi bir $\hat{y}_i$ varyansı karşılık gelen e_i değerinin varyansına eklenerek, $V(y_i)$ için,

$$\mathbf{H}\sigma^2 + (\mathbf{I} - \mathbf{H})\sigma^2 = \mathbf{I}\sigma^2 \tag{4.28a}$$

$$V(\hat{\mathbf{y}}) + V(\mathbf{e}) = V(\mathbf{y}) \tag{4.28b}$$

sonucu elde edilir. Varyanslar negatif olamayacağı için **H** matrisin her bir köşegen elamanı h_{ii} sıfır ile bir arasında olmak zorundadır. Buna uygun olarak herhangi i-inci gözlem için tahminlenmiş değerin varyansı daima bireysel gözlemlerin varyansından küçük olacaktır, diğer bir ifade ile $V(\hat{y}_i)<V(y_i)$.

Eşitlik (4.28)'den görüleceği gibi, $\hat{y}_i$ üzerinde büyük kesinliğe (düşük varyans) sahip veri noktaları e_i üzerinde düşük kesinliğe (büyük varyans) sahip olacaktır. Belsley, Kuh ve Welsch (1980), **H** matrisinin köşegen elamanlarının, veri noktalarının X uzayının merkezinden olan uzaklığının bir ölçümü olarak yorumlanabileceğini göstermiştir. X uzayının merkezinden uzak olan noktalar göreli olarak büyük h_{ii} değerine sahiptirler ve bu nedenle $\hat{y}_i$ üzerinde düşük, e_i üzerinde yüksek kesinliğe sahiptirler. Veri merkezinden uzak noktalar için daha küçük artık varyansı, uyumu yapılmış regresyon doğrusunun veya çıktı yüzeyinin, bu noktalar için, gözlenmiş değerlerin daha yakınına yönelmiş olduğunu belirtir.

Varyans ve kovaryansların σ^2 nin çarpanları olarak ifade edilebileceği ve katsayıların **X** matrisi ile belirtilebileceği eşitlikler (4.25), (4.26) ve (4.27) ile gösterilmişti. Bu sabitler matrisi uyumu yapılan modele ve çalışmadaki bağımsız değişkenlerin seviyelerine bağımlıdır. Tasarlanmış deneylerde, bağımsız değişkenlerin seviyeleri araştırmacıların kontrolü altındadır. Bu durumda σ^2 değeri hariç, deneyin kesinliği araştırmacının kontrolü altındadır ve deney başlamadan önce belirlenebilir. Alternatif deney tasarımlarının etkinliği, her bir tasarım için $(\mathbf{X}^T\mathbf{X})^{-1}$ ve $\mathbf{H}$ değerleri hesaplanıp, daha sonra karşılaştırılarak bulunabilir. Sonuç olarak ilgilenilen değerler için en küçük varyansı veren tasarım tercih edilebilecektir.

4.6 KARELER TOPLAMLARI

Eşitlik (4.15*a*),

$$\mathbf{y}=\hat{\mathbf{y}}+\mathbf{e} \tag{4.29}$$

olarak yeniden düzenlenebilir. Eşitlik (4.29) da belirtildiği gibi gözlenmiş bağımlı değişken vektörü $\mathbf{y}$, tahminlenmiş ortalama vektörü $\hat{\mathbf{y}}$ ve artık vektörü $\mathbf{e}$ olarak iki kısma ayrılabilmektedir. $\mathbf{y}$ vektörünün bu ayrışımı, bağımlı değişkenin toplam kareler toplamının bir ayrışımını elde etmek için

$$\mathbf{y}^T\mathbf{y}=\hat{\mathbf{y}}^T\hat{\mathbf{y}}+\mathbf{e}^T\mathbf{e} \tag{4.30}$$

kullanılabilir, bkz Alıştırma 4.5. Eşitlik (4.30) ile tanımlanan kareler toplamları ise, *Toplam kareler toplamı*,

$$KT(T)=\mathbf{y}^T\mathbf{y}=\sum y_i^2=y_1^2+\ldots+y_n^2 \tag{4.31a}$$

Model kareler toplamı,

$$KT(M)=\hat{\mathbf{y}}^T\hat{\mathbf{y}}=\mathbf{b}^T\mathbf{X}^T\mathbf{y}=\sum \hat{y}_i^2=\hat{y}_1^2+\ldots+\hat{y}_n^2 \tag{4.31b}$$

bkz. Alıştırma 4.6. *Artık kareler toplamı*,

$$KT(e)=\mathbf{e}^T\mathbf{e}=\mathbf{y}^T\mathbf{y}-\mathbf{b}^T\mathbf{X}^T\mathbf{y}=\sum e_i^2=e_1^2+\ldots+e_n^2 \tag{4.31c}$$

eşitlikleri ile tanımlanır. Eşitlik (4.30) $\mathbf{y}$ için toplam düzeltilmemiş kareler toplamını vermektedir. Düzeltilmiş kareler toplamı, $\mathbf{y}^T\mathbf{y}$ ve $\hat{\mathbf{y}}^T\hat{\mathbf{y}}$ değerlerinin her ikisinden de ortalamanın (sabit terimin) kareler toplamının (veya düzeltme faktörünün) çıkartılmasıyla elde edilir:

$$\mathbf{y}^T\mathbf{y}-KT(b_0)=\left[\hat{\mathbf{y}}^T\hat{\mathbf{y}}-KT(b_0)\right]+\mathbf{e}^T\mathbf{e} \tag{4.32a}$$

$$KT(Td)=KT(R)+KT(e) \tag{4.32b}$$

Ortalama için düzeltme faktörü kareler toplamı $KT(b_0)$, sadece sabit terim β_0 içeren bir model için kareler toplamıdır. Modelde sadece sabit terim mevcut olduğunda da model $\mathbf{y}=\mathbf{X}\boldsymbol{\beta}+\boldsymbol{\varepsilon}$ formunda yazılabilir. Bu modeldeki $\mathbf{X}$, sadece birlerden oluşan bir sütün vektörü ve $\boldsymbol{\beta}=\beta_0=\mu$ olup bir tek elemanı temsil eder. Model parametresinin tahmini,

$$\mathbf{b}=\left(\mathbf{1}^T\mathbf{1}\right)^{-1}\mathbf{1}^T\mathbf{y}=\left(\frac{1}{n}\right)\mathbf{1}^T\mathbf{y} \tag{4.33a}$$

ve eşitliğin en sonundaki vektörel çarpım gözlemlerin

$$\mathbf{1}^T\mathbf{y}=\sum y_i=y_1+\ldots+y_n=n\bar{y} \tag{4.33b}$$

toplamını verdiğinden,

$$\mathbf{b}=\bar{y} \tag{4.33c}$$

elde edilir. Model kareler toplamı, eşitlik (4.31*b*) ile $KT(M)=\mathbf{b}^T\mathbf{X}^T\mathbf{y}$ olup model bağımsız değişken içermediğinden,

$$KT(b_0) = \mathbf{b}^T(\mathbf{1}^T\mathbf{y}) \tag{4.34a}$$

eşitlik (4.33*a*) kullanılarak,

$$KT(b_0) = \left(\frac{1}{n}\right)\mathbf{y}^T\mathbf{1}(\mathbf{1}^T\mathbf{y}) = n\bar{y}^2 = \frac{\left(\sum y_i\right)^2}{n} \tag{4.34b}$$

bulunur. Eşitliğin en sağındaki değer düzeltme faktörü olarak adlandırılır. Elde edilen sonuçlar Kısım 4.9 da daha detaylı olarak ele alınacaktır.

4.7 σ² nin SAPMASIZ TAHMİNİ

Bu kısımda hata terimleri varyansının sapmasız bir tahmini elde edilecektir. $\boldsymbol{\varepsilon}$ değerleri doğrudan gözlenemediği için. σ^2 nin tahminin artık kareler toplamı $\mathbf{e}^T\mathbf{e}$ ile elde edilmesi uygun olabilecektir. Tek sorun sapmasız bir tahminci elde edebilmek için payda değerinin ne olması gerektiğidir. Eşitlik (4.16) kullanılarak,

$$\mathbf{e}^T\mathbf{e} = \boldsymbol{\varepsilon}^T(\mathbf{I}-\mathbf{H})\boldsymbol{\varepsilon} \tag{4.35}$$

elde edilir, her iki tarafın beklenen değeri alınarak,

$$\begin{aligned} E\left(\mathbf{e}^T\mathbf{e}\right) &= E\left[\boldsymbol{\varepsilon}^T(\mathbf{I}-\mathbf{H})\boldsymbol{\varepsilon}\right] \\ &= E\left\{tr\left[\boldsymbol{\varepsilon}^T(\mathbf{I}-\mathbf{H})\boldsymbol{\varepsilon}\right]\right\} \\ &= E\left\{tr\left[(\mathbf{I}-\mathbf{H})\boldsymbol{\varepsilon}^T\boldsymbol{\varepsilon}\right]\right\} \\ &= \sigma^2 tr(\mathbf{I}-\mathbf{H}) \end{aligned}$$

bulunur. Eşitliğin sağındaki iz terimi,

$$tr(\mathbf{I}-\mathbf{H}) = tr(\mathbf{I}_n) - tr\left[\mathbf{X}\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\right]$$

$$tr(\mathbf{I}-\mathbf{H}) = tr(\mathbf{I}_n) - tr\left[\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{X}\right]$$

$$tr(\mathbf{I}-\mathbf{H}) = n-p \tag{4.36a}$$

olduğundan eğer,

$$s^2 = \frac{\mathbf{e}^T\mathbf{e}}{n-p} \tag{4.36b}$$

tanımı yapılırsa,

$$E\left(s^2\right) = \sigma^2$$

şeklinde istenen sapmasız tahminci elde edilir. s^2 değerinin karekökü bu tahmincinin standart hatası olarak ifade edilir ve regresyon düzlemi çerçevesinde y değerlerinin standart sapması olarak değerlendirilebilir.

Yukarıda EKK tahminleme yönteminin matris cebri kullanılarak çok değişkenli regresyon yönteminde nasıl kullanılabileceği, parametre tahminlerinin kestirim değerleri ile artıkların nasıl elde edilebileceği açıklandı. Eşitlik (4.8), (4.9), (4.12), (4.13*a*) ve (4.15*c*) den görülebileceği gibi bu yöntem için ($\mathbf{X}^T\mathbf{X}$) matrisi ve onun ters matrisi önemli bir rol oynamaktadır. Bu matrisin varyans analizindeki önemi bundan sonraki kısımlarda, geometrik anlamı ise Bölüm 5 de açıklanacaktır.

4.8 KARESEL FORMLAR

Modelin uyumunu sağlamadaki amaç bağımsız değişkenlerde mevcut olan bilgiyi kullanarak, bağımlı değişkendeki değişkenliğin olabildiğince büyük bir kısmını açıklamaktır. Bağımsız değişkenlerin modele yaptığı katkı y bağımlı değişkeninin toplam kareler toplamının, bağımsız değişkenler tarafından açıklanabilen kısmı ile ölçülür. Kareler toplamının ayrıştırılan her bir bileşeni y değişkenindeki bir karesel formdur, bkz Kısım 3.12. Diğer bir deyişle kareler toplamlarının her biri

$$\theta=\mathbf{y}^{\mathrm{T}}\mathbf{A}\mathbf{y} \tag{4.37}$$

şeklinde yazılabilir. Burada $\mathbf{y}$ boyutu $n\times1$ olan şans değişkeni vektörü, $\mathbf{A}$ matrisi ise boyutu $n\times n$ olup katsayılar matrisi, *tanım matrisi* olarak adlandırılır. Elde edilen skaler $\mathbf{y}^{\mathrm{T}}\mathbf{A}\mathbf{y}$ bir şans değişkeni olup $\mathbf{y}$ vektörüne göre karesel formdur. Belirli bir kareler toplamı ile onun serbestlik derecesi ve farklı kareler toplamları ile arasındaki ortogonallik özelliği, karesel formdaki tanım matrisi ile belirlenir. Bu nedenle bağımlı değişken $\mathbf{y}$ için varyans analizi açıklanmadan önce karesel formlar ve bazı özellikleri basit bir örnek ile açıklanmaya çalışılacaktır. İlk olarak,

$$\tilde{c}_1 = y_1 + y_2 - 2y_3$$

şeklinde y şans değişkeninin bir tek doğrusal fonksiyonu ele alınsın. Doğrusal fonksiyonun katsayıları bir vektörü tanımlar.

$$\tilde{\mathbf{a}}^T = \begin{pmatrix} 1 & 1 & -2 \end{pmatrix}$$

Doğrusal fonksiyondan elde edilecek kareler toplamlarının beklenen değerindeki σ^2 nin katsayısını 1 e eşitlemek için bu vektörün her bir elemanı vektörün normuna bölünür:

$$\mathbf{a}^T = \begin{pmatrix} \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} \end{pmatrix}$$

Elde edilen ölçeklenmiş doğrusal fonksiyon:

$$c_1 = \tilde{c}_1 / \sqrt{6}$$

$$c_1 = \mathbf{a}^T\mathbf{y} = \frac{1}{\sqrt{6}}y_1 + \frac{1}{\sqrt{6}}y_2 - \frac{2}{\sqrt{6}}y_3$$

Bu fonksiyonun kareler toplamı, fonksiyonun karesini alınarak hesaplanır. Burada şans değişkeni vektörü, $\mathbf{y}^T=(y_1\ y_2\ y_3)$ şeklinde tanımlanmıştır. Matris gösteriminde c_1 için kareler toplamı,

$$KT(c_1) = c_1^2$$

$$KT(c_1) = \left(\mathbf{a}^T\mathbf{y}\right)^T\left(\mathbf{a}^T\mathbf{y}\right) = \mathbf{y}^T\mathbf{a}\mathbf{a}^T\mathbf{y}$$

$$KT(c_1) = \mathbf{y}^T\mathbf{A}\mathbf{y}$$

eşitliği elde edilir. Görülebileceği gibi $KT(c_1)$, **y**'nin karesel formu olarak ifade edilmiştir. $\mathbf{A}=\mathbf{aa}^T$ ise 3×3 boyutlu tanım matrisidir. Bu fonksiyon için tanım matrisi,

$$\mathbf{A}=\mathbf{aa}^T=\begin{bmatrix}1/\sqrt{6}\\1/\sqrt{6}\\-2/\sqrt{6}\end{bmatrix}\begin{bmatrix}1/\sqrt{6} & 1/\sqrt{6} & -2/\sqrt{6}\end{bmatrix}$$

$$=\begin{bmatrix}1/\sqrt{6} & 1/\sqrt{6} & -2/\sqrt{6}\\1/\sqrt{6} & 1/\sqrt{6} & -2/\sqrt{6}\\-2/\sqrt{6} & -2/\sqrt{6} & 4/\sqrt{6}\end{bmatrix} \tag{4.38}$$

olarak elde edilir. Karesel form çarpımı tanımlandığında,

$$\mathbf{y}^T\mathbf{Ay}=\begin{bmatrix}y_1 & y_2 & y_3\end{bmatrix}\begin{bmatrix}1/\sqrt{6} & 1/\sqrt{6} & -2/\sqrt{6}\\1/\sqrt{6} & 1/\sqrt{6} & -2/\sqrt{6}\\-2/\sqrt{6} & -2/\sqrt{6} & 4/\sqrt{6}\end{bmatrix}\begin{bmatrix}y_1\\y_2\\y_3\end{bmatrix}$$

$$=\frac{1}{6}y_1^2+\frac{1}{6}y_2^2+\frac{4}{6}y_3^2+\frac{2}{6}y_1y_2-\frac{4}{6}y_1y_3-\frac{4}{6}y_2y_3 \tag{4.39}$$

sonucu elde edilir. Eşitlik (4.38)'de verilen **A** matrisinin elemanları eşitlik (4.39) de verilen açılım ile karşılaştırıldığında, tanım matrisinin köşegen elemanlarının, karesel terimlerin katsayıları ve köşegen dışı simetrik elemanların toplamında çarpım terimlerinin katsayıları olduğu görülür. Bir karesel form için tanım matrisi daima simetriktir.

Doğrusal fonksiyon c_1 e ortogonal olan **y**'nin ikinci bir doğrusal fonksiyonu,

$$c_2=\mathbf{b}^T\mathbf{y}=\frac{1}{\sqrt{2}}y_1-\frac{1}{\sqrt{2}}y_2$$

olarak tanımlansın. Burada $\mathbf{b}^T=\left(1/\sqrt{2} \quad -1/\sqrt{2} \quad 0\right)$ şeklindedir. Bu fonksiyon için kareler toplamı,

$$KT(c_2)=\mathbf{y}^T\mathbf{By}$$

olup tanım matrisi,

$$\mathbf{B}=\mathbf{bb}^T=\begin{bmatrix}1/2 & -1/2 & 0\\-1/2 & 1/2 & 0\\0 & 0 & 0\end{bmatrix}$$

şeklindedir. Bu kareler toplamlarının her ikisi de 1 serbestlik derecesine sahiptir. Bunun nedeni her iki durumda da sadece bir tek fonksiyonun mevcut olmasıdır. Bir karesel form için serbestlik derecesi tanım matrisinin rankına eşittir. Eğer tanım matrisi, idempotent ise tanım matrisinin izi rankına eşit olacaktır. Bir karesel form için tanım matrisinin idempotent olması zorunlu değildir. Fakat ilgilenilen karesel formlar için tanım matrisleri idempotent olacaktır. Verilen her iki örnekteki **A** ve **B** tanım matrisleri idempotenttir. Eğer $1/\sqrt{6}$ ve $1/\sqrt{2}$ katsayılar vektöründe içerilmeseydi **A** ve **B** idempotent olmayacaktı. Her bir fonksiyon için $tr(\mathbf{A})=tr(\mathbf{B})=1$ serbestlik derecesi elde edilir.

A ve **B** tanım matrisleri ile tanımlanan karesel formlar her biri bir adet doğrusal fonksiyona ayrı ayrı uygulanmıştı. Bu nedenle her bir karesel form 1 serbestlik derecesine sahiptir. Bu iki fonksiyon birlikte ele alınabilir. Her iki fonksiyonun katsayıları 2×3 boyutlu $\mathbf{E}^T$ matrisi ile belirtilebilir.

$$\mathbf{c} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \mathbf{E}^T \mathbf{y} = \begin{bmatrix} 1/\sqrt{6} & 1/\sqrt{6} & -2/\sqrt{6} \\ 1/\sqrt{2} & -1/\sqrt{2} & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$$

c şans vektörünün tanımladığı karesel form,

$$KT(\mathbf{c}) = \mathbf{c}^T \mathbf{c} = \mathbf{y}^T \mathbf{E}\mathbf{E}^T \mathbf{y}$$

için tanım matrisi,

$$\mathbf{F} = \mathbf{E}\mathbf{E}^T = \begin{bmatrix} 2/3 & -1/3 & -1/3 \\ -1/3 & 2/3 & -1/3 \\ -1/3 & -1/3 & 2/3 \end{bmatrix}$$

olarak elde edilir. Örnekteki tanım matrisi idempotenttir ve izi 2 olup bu kareler toplamının serbestlik derecesine eşittir. İki orijinal fonksiyon birbirine ortogonal $\mathbf{a}^T\mathbf{b}$=0 olduğunda, tanımlanan karesel form idempotenttir. Eğer tanım matrislerinin çarpımı sıfır ise aynı **y** vektörünün iki karesel formu ortogonaldir. Örnekte verilen iki karesel formun ortogonalliği **A** ve **B** matrislerinin,

$$\mathbf{AB} = \begin{bmatrix} 1/\sqrt{6} & 1/\sqrt{6} & -2/\sqrt{6} \\ 1/\sqrt{6} & 1/\sqrt{6} & -2/\sqrt{6} \\ -2/\sqrt{6} & -2/\sqrt{6} & 4/\sqrt{6} \end{bmatrix} \begin{bmatrix} 1/2 & -1/2 & 0 \\ -1/2 & 1/2 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

çarpımı ile ispatlanabilir. **A** ve **B** simetrik olduğundan **BA** çarpımı da aynı sonucu verir. İki doğrusal fonksiyon ortogonal olduğunda bireysel olarak ele alınan iki fonksiyon için kareler toplamlarının toplamı, iki fonksiyonun birlikte ele alındıklarında oluşan kareler toplamına eşittir. Kareler toplamı için geçerli olan bu özellik serbestlik dereceleri içinde geçerlidir. Bu toplanabilirlik özelliğinin ikiden fazla fonksiyon için geçerli olabilmesi için tüm fonksiyon çiftlerinin ikişerli olarak ortogonal olması gereklidir. Karesel formların ortogonalliği, bireysel kareler toplamlarında içerilen bilgi parçalarının birbirinden bağımsız olduğunu belirtir. Karesel formlarla ilgili üç önemli özellik aşağıda verilmiştir.

1) Herhangi bir kareler toplamı $\mathbf{y}^T\mathbf{A}\mathbf{y}$ şeklinde yazılabilir. Burada **A** simetrik bir kare matristir.
2) Herhangi bir karesel formun serbestlik derecesi tanım matrisinin rankına eşittir. Tanım matrisi idempotent ise matrisin rankı izine eşittir.
3) Eğer sadece ve sadece tanım matrislerinin çarpımı boş matrisi (sıfır matris) sonucunu veriyorsa iki karesel form ortogonaldir.

4.9 VARYANS ANALİZİ

Gözlenmiş bağımlı değişken vektörü **y**, eşitlik (4.29) da belirtildiği gibi, tahmin vektörü ve artık vektörü olarak iki kısma ayrılabilmekteydi. **y** vektörünün bu ayrışımı, bağımlı değişkenin eşitlik (4.30) ile tanımlanan toplam kareler toplamının benzer bir ayrışımını elde etmek için kullanılabilir. Eşitlik (4.31*a*) ile verilen çarpım, **y** sütun vektöründeki elemanların toplam kareler toplamını vermektedir. Bu ifade tanım matrisi birim matrise eşit olan bir karesel formu belirtmektedir. Birim matris idempotent olduğundan izi matrisin boyutuna eşittir. Bu nedenle toplam (düzeltilmemiş) kareler toplamının $KT(T)$ serbestlik derecesi vektördeki eleman sayısına eşittir. İdempotent matrisler içinde sadece birim matris tam ranklıdır. Eşitlik (4.29) kullanılarak,

$$\mathbf{y}^T\mathbf{I}\mathbf{y} = (\hat{\mathbf{y}} + \mathbf{e})^T(\hat{\mathbf{y}} + \mathbf{e}) = \hat{\mathbf{y}}^T\hat{\mathbf{y}} + \hat{\mathbf{y}}^T\mathbf{e} + \mathbf{e}^T\hat{\mathbf{y}} + \mathbf{e}^T\mathbf{e}$$

elde edilir. Bu ifadede eşitlikler (4.14) ve (4.15*c*) dönüşümleri yerine konarak,

$$\mathbf{y}^T\mathbf{I}\mathbf{y} = (\mathbf{H}\mathbf{y})^T(\mathbf{H}\mathbf{y}) + (\mathbf{H}\mathbf{y})^T[(\mathbf{I}-\mathbf{H})\mathbf{y}] + [(\mathbf{I}-\mathbf{H})\mathbf{y}]^T(\mathbf{H}\mathbf{y}) + [(\mathbf{I}-\mathbf{H})\mathbf{y}]^T[(\mathbf{I}-\mathbf{H})\mathbf{y}]$$

H ve (**I-H**)matrisleri simetrik ve idempotent olduğundan,

$$\mathbf{y}^T\mathbf{I}\mathbf{y} = \mathbf{y}^T\mathbf{H}\mathbf{y} + \mathbf{y}^T\mathbf{H}(\mathbf{I}-\mathbf{H})\mathbf{y} + \mathbf{y}^T(\mathbf{I}-\mathbf{H})\mathbf{H}\mathbf{y} + \mathbf{y}^T(\mathbf{I}-\mathbf{H})\mathbf{y}$$

yazılabilir. Bu eşitlikteki ikinci ve üçüncü terimler **H**(**I-H**)=**0** olduğundan sıfırdır. Karesel formlar birbirine ortogonal olduğundan bileşenleri toplanabilecektir. Sonuç olarak,

$$KT(T) = KT(M) + KT(e) \tag{4.40a}$$

$$\mathbf{y}^T\mathbf{I}\mathbf{y} = \mathbf{y}^T\mathbf{H}\mathbf{y} + \mathbf{y}^T(\mathbf{I}-\mathbf{H})\mathbf{y} \tag{4.40b}$$

elde edilebilir. Eşitlik (4.40) dan görülebileceği gibi toplam düzeltilmemiş kareler toplamları **H** ve (**I-H**) tanım matrisleri kullanılarak iki karesel form şeklinde ayrıştırılabilir. $\hat{\mathbf{y}}^T\hat{\mathbf{y}}=\mathbf{y}^T\mathbf{H}\mathbf{y}$ kısmı uyumu yapılan model tarafından açıklanabilen bölümdür ve model kareler toplamı olarak adlandırılarak $KT(M)$ ile gösterilir. İkinci terim $\mathbf{e}^T\mathbf{e}= \mathbf{y}^T(\mathbf{I-H})\mathbf{y}$ ise model tarafından açıklanamayan bölüm olup artık kareler toplamı olarak isimlendirilir $KT(e)$ ile gösterilir.

Daha önce belirtildiği gibi karesel formların serbestlik dereceleri tanım matrislerinin rankına bağlıdır. **H** matrisinin rankı **X** matrisinin rankı ile belirlenir. Tam ranklı modeller için **X** matrisinin rankı sütun sayısına eşittir. Sütun sayısı aynı zamanda parametre sayısına eşittir. **H** idempotent olduğundan rankı izine eşittir. Eşitlik (4.36*a*) kullanılarak:

$$tr(\mathbf{H}) = tr\left[\mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\right] = tr\left[(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{X}\right]$$

$$tr(\mathbf{H}) = tr(\mathbf{I}_P) = p \tag{4.41}$$

Model tam ranklı olduğunda $KT(M)$ bileşeninin serbestlik derecesi p olacaktır. **I** matrisinin alt indisi p, birim matrisin derecesini belirtmektedir. $KT(e)$ için serbestlik derecesi ise eşitlik (4.36*a*) ile $(n\text{-}p)$ olarak bulunmuştur. Karesel formlar için uygun hesaplama formülleri de mevcuttur. Model kareler toplamı eşitlik (4.31*b*) den görülebileceği gibi,

$$\hat{\mathbf{y}}^T\hat{\mathbf{y}} = \mathbf{b}^T\mathbf{X}^T\mathbf{y}$$

eşitliğinden elde edilebilir. Bu formülasyonda, $\hat{\mathbf{y}}$ vektörünün hesaplanmasına gerek kalmamaktadır. Model kareler toplamı ve artık kareler toplamı eklenebilir olduğundan, $KT(e)$ için en basit hesaplama şekli, eşitlik (4.40*a*) kullanılarak,

$$KT(e) = KT(T) - KT(M)$$

$$\mathbf{e}^T\mathbf{e} = \mathbf{y}^T\mathbf{y} - \mathbf{b}^T\mathbf{X}^T\mathbf{y} \tag{4.42}$$

olarak verilebilir. Toplam düzeltilmemiş kareler toplamının model ve artık kareler toplamları şeklinde iki bileşene ayrılabileceği yukarıda açıklandı. Bununla birlikte **y**'nin orijine (sıfıra) göre değişkenliği yerine genellikle ortalaması çevresindeki değişkenliği ile ilgilenilir. Bu durumda bağımsız değişkenlerdeki bilgi ile bu değişkenliğin ne kadarının açıklanabildiği araştırılır. Çünkü regresyon analizinde araştırılan bağımsız değişkenler ile bağımlı değişken arasında bir regresyon olup olmadığıdır. Bağımsız değişkenlerden hiçbir bilgi elde edilemediği durumlarda, y'nin en iyi kestiricisi, anakütle ortalamasının en iyi tahmini olan aritmetik ortalamadır. Modelde bağımsız değişkenler mevcut olduğunda, bağımsız değişkenlerin şans değişkeni y'nin kestirimine (y'nin ortalaması hariç) ne miktar katkıda bulunduğu sorusu ile karşılaşılır.

Bağımsız değişkenler ile elde edilen bilginin ölçümü, modelde bağımsız değişken mevcut iken elde edilen $KT(M)$ değeri ile bağımsız değişken bulunmadığında elde edilen $KT(M)$ değeri arasındaki farktır. Model bağımsız değişken içermiyorsa sadece bir tek β_0 (ortalama) parametresi mevcuttur. Başka bir deyişle model $y_i=\beta_0+\varepsilon_i$ şeklinde olacaktır. Bu durumda $KT(M)$, $KT(b_0)$ ile belirtilecektir. $KT(b_0)$ düzeltme faktörü olarak adlandırılır. Bağımsız değişken(ler) tarafından kareler toplamına yapılan katkı regresyon kareler toplamı olarak adlandırılır ve $KT(M)$ ile belirtilir. Bu durumda

$$KT(R)=KT(M)-KT(b_0) \tag{4.43}$$

eşitliği yazılabilir. Bu eşitlikteki $KT(M)$, (bağımsız değişken içeren) model kareler toplamıdır.

$KT(b_0)$ için karesel formun tanım matrisi eşitlik (4.34b) de gösterilmiştir. Eşitlikte yer alan tanım matrisi $n\times n$ boyutlu birlerden oluşan ve $1/n$ skaleri ile çarpılan matristir. Bu matris **J** ile gösterilecektir:

$$\frac{1}{n}\left(\mathbf{1}\mathbf{1}^T\right)=\frac{1}{n}\begin{bmatrix}1 & 1 & \dots & 1\\ 1 & 1 & \dots & 1\\ \vdots & \vdots & \ddots & \vdots\\ 1 & 1 & \dots & 1\end{bmatrix}=\frac{1}{n}\mathbf{J} \tag{4.44}$$

$(1/n)\mathbf{J}$ matrisi, rankı izine $tr[(1/n)\mathbf{J}]=1$ eşit olan bir idempotent matristir. Matrisin rankı 1'e eşit olduğundan $KT(b_0)$, 1 serbestlik derecesine sahip olacaktır. $KT(M)$ için karesel form eşitlik (4.40b) de tanımlanmıştı. Sonuç olarak $KT(R)$ için karesel form eşitlik (4.43) kullanılarak,

$$KT(R)=\mathbf{y}^T\mathbf{H}\mathbf{y}-\mathbf{y}^T(\mathbf{J}/n)\mathbf{y}$$

$$KT(R)=\mathbf{y}^T(\mathbf{H}-\mathbf{J}/n)\mathbf{y} \tag{4.45}$$

ve $KT(R)$ için tanım matrisi $(\mathbf{H}\text{-}\mathbf{J}/n)$ elde edilir. Serbestlik derecesi ise $(p\text{-}1)=k$. Tanım matrisi $\mathbf{J}/n$, $(\mathbf{H}\text{-}\mathbf{J}/n)$ ve $(\mathbf{I}\text{-}\mathbf{H})$ matrislerine ortogonaldir. Bunun sonucunda toplam düzeltilmemiş kareler toplamı, üç ortogonal bileşene ayrıştırılabilir:

$$KT(T)= KT(b_0)+ KT(R)+ KT(e) \tag{4.46a}$$

$$\mathbf{y}^T\mathbf{I}\mathbf{y}=\mathbf{y}^T(\mathbf{J}/n)\mathbf{y}+\mathbf{y}^T(\mathbf{H}-\mathbf{J}/n)\mathbf{y}+\mathbf{y}^T(\mathbf{I}-\mathbf{H})\mathbf{y} \tag{4.46b}$$

Karesel formların serbestlik dereceleri sırası ile 1, k ve (n-p) şeklindedir. Varyans analizinde genellikle eşitlik (4.32*b*) ile verilen düzeltilmiş kareler toplamı kullanılır. Eşitlik (4.46) için düzenlenen varyans analiz tablosu Tablo 5.1`de verilmiştir.

Tablo 5.1 *Çok değişkenli regresyon için varyans analiz tablosu*

Değişkenlik Kaynağı	*Kareler Toplamları*	*Serbestlik Derecesi*	*Kareler Ortalaması*	*F-Testi*
Ortalamaya Bağlı Değişkenlik	$n\bar{y}^2$	1	$KO(b_0)=\dfrac{n\bar{Y}^2}{1}$	
Regresyona Bağlı Değişkenlik	$\mathbf{b}^T\mathbf{X}^T\mathbf{y}-n\bar{y}^2$	k	$KO(R)=\dfrac{\mathbf{b}^T\mathbf{X}^T\mathbf{y}-n\bar{y}^2}{k}$	$F=\dfrac{KO(R)}{KO(e)}$
Artığa Bağlı Değişkenlik	$\mathbf{y}^T\mathbf{y}-\mathbf{b}^T\mathbf{X}^T\mathbf{y}$	*n-p*	$KO(e)=\dfrac{\mathbf{y}^T\mathbf{y}-\mathbf{b}^T\mathbf{X}^T\mathbf{y}}{n-p}$	
Toplam Değişkenlik	$\mathbf{y}^T\mathbf{y}$	n		

Regresyon analizinde kullanılan tüm tanım matrisleri idempotenttir. Tanım matrislerinin hepsi birbirlerine çifterli olarak ortogonaldir. Bunun sonucu olarak düzeltilmemiş toplam kareler toplamı birbirine ortogonal kareler toplamlarına ayrılabilmektedir. Tablo 5.1 de verilen artık kareler ortalaması,

$$\mathbf{b}^T\mathbf{X}^T\mathbf{y}=\mathbf{y}^T\mathbf{X}\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{y}$$

özdeşliği kullanılarak,

$$s^2=\frac{\mathbf{y}^T\mathbf{y}-\mathbf{y}^T\mathbf{X}\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{y}}{n-p}$$

$$s^2=\frac{\mathbf{y}^T\left[\mathbf{I}-\mathbf{X}\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\right]\mathbf{y}}{n-p} \qquad (4.46c)$$

şeklinde ifade edilebilir.

4.9.1 Karesel Formların Beklenen Değeri ve Varyansı

y bağımlı değişkeninin varyans analizinde hesaplanan her bir karesel formu, model parametrelerinin tahminlenmiş bazı fonksiyonlarıdır. Diğer bir deyişle varyans analizindeki herhangi bir karesel form gözlemlerin karesel fonksiyonudur. Tablo 5.1 de verilen F-hesap değeri de bu karesel formların bir fonksiyonu olduğundan karesel formların dağılımlarının belirlenmesi oldukça önemlidir. Daha genel olarak çok değişkenli regresyon analizi için hipotez testleri ve aralık tahminlenmesinde de karesel formlar kullanılmaktadır. Bu aşamada ilk olarak karesel formların beklenen değerleri ve varyanslarının elde edilmesi ile ilgilenilecektir. y şans değişkeninin ortalaması $E(\mathbf{y})=\boldsymbol{\mu}$ ve varyans-kovaryans matrisi $V(\mathbf{y})=\mathbf{V}$ ile belirtilsin. $E(\mathbf{y}^T\mathbf{A}\mathbf{y})$ değerinin elde edilmesi için ilk olarak $\mathbf{V}=E[(\mathbf{y}-\boldsymbol{\mu})(\mathbf{y}-\boldsymbol{\mu})^T]=\mathrm{E}(\mathbf{y}\mathbf{y}^T)-\boldsymbol{\mu}\boldsymbol{\mu}^T$ yazılabilir. Bu ifade $\sigma^2=E(y^2)-\mu^2$ ifadesinin çok değişkenli durumunu belirtmektedir. Sonuç olarak

$E(\mathbf{yy}^T)$=$\mathbf{V}$+$\boldsymbol{\mu\mu}^T$ eşitliği verilebilecektir. $\mathbf{y}^T\mathbf{Ay}$ bir skaler olduğundan aynı zamanda izine de eşittir. Eşitlik (3.15) kullanılarak,

$$E\left(\mathbf{y}^T\mathbf{Ay}\right)=E\left[tr\left(\mathbf{y}^T\mathbf{Ay}\right)\right]=E\left[tr\left(\mathbf{Ayy}^T\right)\right]=tr\left[\mathbf{A}E\left(\mathbf{yy}^T\right)\right]$$

ve

$$E\left(\mathbf{y}^T\mathbf{Ay}\right)=tr\left[\mathbf{A}\left(\mathbf{V}+\boldsymbol{\mu\mu}^T\right)\right]=tr\left(\mathbf{AV}\right)+\boldsymbol{\mu}^T\mathbf{A}\boldsymbol{\mu} \qquad (4.47a)$$

elde edilir. Bu eşitliğin kullanılabilmesi için y değişkeninin normal dağılış göstermesi gerekli değildir. $\mathbf{y}^T\mathbf{Ay}$'nin varyansı ise

$$V\left(\mathbf{y}^T\mathbf{Ay}\right)=2tr\left(\mathbf{AV}\right)^2+4\boldsymbol{\mu}^T\mathbf{AVA}\boldsymbol{\mu} \qquad (4.47b)$$

şeklindedir. Eşitlik (4.47b) nin elde edilebilmesi için y şans değişkeninin normal dağılması gereklidir. Elde edilen (4.47a) ve (4.47b) sonuçları herhangi bir karesel form için genelleme yapılmasında kullanılır. Buna uygun olarak EKK varsayımları $E(\mathbf{y})$=$\mathbf{X}\boldsymbol{\beta}$ ve $V(\mathbf{y})$=$\mathbf{I}\sigma^2$ altında, karesel formun beklenen değeri,

$$E\left(\mathbf{y}^T\mathbf{Ay}\right)=\sigma^2 tr\left(\mathbf{A}\right)+\boldsymbol{\beta}^T\mathbf{X}^T\mathbf{AX}\boldsymbol{\beta} \qquad (4.48)$$

olarak verilebilir. Varyans analizinde karesel formların beklenen değeri $\mathbf{A}$ matrisi yerine uygun bir tanım matrisi kullanılarak elde edilir. $\mathbf{A}$ matrisi idempotent olduğunda σ^2 nin katsayısı karesel formun serbestlik derecesine eşit olacaktır. Aşağıda regresyon analizinde kullanılan karesel formların beklenen değerleri incelenmiştir:

Model kareler toplamının beklenen değeri, eşitlik (4.41) ve (4.13a) kullanılarak,

$$E\left[KT(M)\right]=E(\mathbf{y}^T\mathbf{Hy})=\sigma^2 tr(\mathbf{H})+\boldsymbol{\beta}^T\mathbf{X}^T\mathbf{HX}\boldsymbol{\beta}$$

$$=p\sigma^2+\boldsymbol{\beta}^T\mathbf{X}^T\mathbf{X}\boldsymbol{\beta} \qquad (4.49)$$

şeklindedir. Bu eşitlikteki ikinci terim (kesişim terimi β_0 ı da içeren) $\boldsymbol{\beta}$ vektörünün bir karesel formudur.

Regresyon kareler toplamının beklenen değeri;

$$E\left[KT(R)\right]=E\left[\mathbf{y}^T(\mathbf{H}-\mathbf{J}/n)\mathbf{y}\right]=\sigma^2 tr(\mathbf{H}-\mathbf{J}/n)+\boldsymbol{\beta}^T\mathbf{X}^T(\mathbf{H}-\mathbf{J}/n)\boldsymbol{\beta}\mathbf{X}$$

$$E\left[KT\left(R\right)\right]=k\sigma^2+\boldsymbol{\beta}^T\mathbf{X}^T(\mathbf{I}-\mathbf{J}/n)\mathbf{X}\boldsymbol{\beta} \qquad (4.50)$$

elde edilir. Burada $\mathbf{X}^T\mathbf{H}$=$\mathbf{X}^T$ olduğu hatırlanmalıdır. $\boldsymbol{\beta}$ vektörüne göre karesel bir form olan ikinci terim model kareler toplamındaki ifadeden farklıdır. Bu farkı oluşturan $\mathbf{X}^T(\mathbf{I}\text{-}\mathbf{J}/n)\mathbf{X}$ matrisidir. $\mathbf{X}$'in ilk sütunu bir sabit sütun olduğundan bu matrisin ilk satır ve sütunu sıfırlardan oluşacaktır. Bunun sonucu olarak karesel ifadede β_0 parametresi yer almaz Başka bir deyişle, regresyon kareler toplamının beklenen değerinde sadece bağımsız değişkenlerin regresyon katsayıları içerilmektedir.

Artık kareler toplamının beklenen değeri,

$$E\left[KT(e)\right]=E\left[\mathbf{y}^T(\mathbf{I}-\mathbf{H})\mathbf{y}\right]=\sigma^2 tr(\mathbf{I}-\mathbf{H})+\boldsymbol{\beta}^T\mathbf{X}^T(\mathbf{I}-\mathbf{H})\boldsymbol{\beta}\mathbf{X}$$

$$E\left[KT(e)\right]=(n-p)\sigma^2+\boldsymbol{\beta}^T\mathbf{X}^T(\mathbf{X}-\mathbf{X})\boldsymbol{\beta}$$

$$E\left[KT(e)\right]=(n-p)\sigma^2 \tag{4.51}$$

şeklindedir. Her üç beklenen değer ifadesinde σ^2 nin katsayısı kareler toplamlarının serbestlik derecesini belirtir. Kareler toplamlarını, karesel ortalamaya dönüştürmek için her bir beklenen değer kendi serbestlik derecesine bölünür. Bunun sonucu olarak tüm karesel formlardaki σ^2 katsayısı 1'e eşit olacaktır.

$$E\left[KO(M)\right]=\sigma^2+\left[\boldsymbol{\beta}^T\mathbf{X}^T\mathbf{H}\mathbf{X}\boldsymbol{\beta}/p\right] \tag{4.52}$$

$$E\left[KO(R)\right]=\sigma^2+\left[\boldsymbol{\beta}^T\mathbf{X}^T(\mathbf{I}-\mathbf{J}/n)\mathbf{X}\boldsymbol{\beta}/k\right] \tag{4.53}$$

$$E\left[KO(e)\right]=\sigma^2 \tag{4.54}$$

Eşitlik (4.54) artık kareler ortalamasının σ^2 nin sapmasız bir tahmini olduğunu gösterir. Regresyon kareler ortalaması β_0 hariç tüm β_j 'lerin karesel bir fonksiyonu artı σ^2 nin bir tahminidir. Buna göre $KO(R)$ ve $KO(e)$ ifadelerinin karşılaştırılması regresyon katsayılarının veya eşdeğer olarak bağımsız değişkenlerin anlamlılığının değerlendirilmesine temel oluşturur. $E[KO(R)]$ ifadesindeki ikinci terim $\boldsymbol{\beta}$'nın karesel bir fonksiyonu olduğu için negatif olamaz ve bağımsız değişkenlerin $\mathbf{y}$ vektörünün tahminlenmesine yaptığı katkıyı belirtir. Bu katkı büyüdükçe $KO(R)$ değeri ile $KO(e)$ değeri arasındaki fark da büyüyecektir. Sonuç olarak gözlenmiş $KO(R)$ değerinin gözlenmiş $KO(e)$ değerine oranı β_0 hariç tüm β_j değerlerinin sıfıra eşit olduğu,

H_0: $\beta_1=\beta_2=\ldots=\beta_k=0$

karmaşık hipotezin test edilmesini sağlar.

Tanımlanan beklenen değer ifadelerinin tümü varyans analizinde kullanılan modelin gerçekten doğru model olduğunu kabul eder. Eğer kullanılan model yanlış ise $E(\mathbf{y})=\mathbf{X}\boldsymbol{\beta}_1+\mathbf{X}_2\boldsymbol{\beta}_2\neq\mathbf{X}\boldsymbol{\beta}_1$ olacak ve eşitlik (4.51)'in ikinci terimi,

$$E\left[KT(e)\right]=\sigma^2 tr(\mathbf{I}-\mathbf{H})+\left[\mathbf{X}\boldsymbol{\beta}_1+\mathbf{X}_2\boldsymbol{\beta}_2\right]^T(\mathbf{I}-\mathbf{H})\left[\mathbf{X}\boldsymbol{\beta}_1+\mathbf{X}_2\boldsymbol{\beta}_2\right]$$

$$E\left[KT(e)\right]=\sigma^2(n-p)+\boldsymbol{\beta}_2^T\mathbf{X}_2^T(\mathbf{I}-\mathbf{H})\mathbf{X}_2\boldsymbol{\beta}_2$$

sıfıra yaklaşmayacaktır ve kareler ortalaması:

$$E\left[KO(e)\right]=\sigma^2+\boldsymbol{\beta}_2^T\mathbf{X}_2^T(\mathbf{I}-\mathbf{H})\mathbf{X}_2\boldsymbol{\beta}_2/(n-p) \tag{4.55}$$

Bu terimde, modele alınmaması sonucu spesifikasyon hatası ortaya çıkacak önemli bağımsız değişken(ler)in regresyon katsayısının karesel bir fonksiyonu mevcut olacaktır. Bu gibi durumlarda $KO(e)$, σ^2 nin pozitif sapmalı bir tahminin olacaktır.

4.9.2 Karesel Formların Dağılımı

Karesel formların olasılık dağılımları, bkz Bölüm Ek 4, parametrik anlamlılık testleri için bir temel oluşturur. Bu testler ve parametrelerle ilgili güven aralıklarının oluşturulabilmesi için hata teriminin normallik varsayımına ihtiyaç vardır. Hataların normal dağıldığının kabul edilmesi aynı zamanda $\mathbf{y}$'nin de normal dağıldığı anlamına gelmektedir.

Teorem 4.1 Eğer şans değişkeni $\mathbf{y} \sim N(\boldsymbol{\mu}; \mathbf{V})$ ve $\theta=\mathbf{y}^T\mathbf{A}\mathbf{y}$ ise**,** **AV** matrisinin idempotent bir matris olması koşulu ile, $\theta \sim \chi^2(r;\Omega)$ şeklinde *merkezi olmayan bir ki-kare* dağılımına sahiptir. Dağılımın parametrelerinden serbestlik derecesi **A** matrisinin rankı $r=\rho(\mathbf{A})$ ile merkezi olmayan parametresi ise $\Omega = \boldsymbol{\mu}^T \mathbf{A}\boldsymbol{\mu}/2$ şeklinde belirlenir.

Eğer $V(\mathbf{y})=\mathbf{V}=\mathbf{I}\sigma^2$ ise **AV** matrisinin idempotent olma şartı, $\mathbf{A}\mathbf{I}\sigma^2=\mathbf{A}\sigma^2$ olduğundan sadece **A** matrisinin idempotent olmasına indirgenmiş olur. **A** matrisinin önündeki sabit σ^2 çarpanı dikkate alındığında karesel form $\mathbf{y}^T\left(\mathbf{A}/\sigma^2\right)\mathbf{y}$ şekline dönüşür. Şans değişkenini bir sabit ile bölmek şans değişkeninin genel dağılımını değiştirmez fakat parametrelerini etkiler. Sonuç olarak $\mathbf{y}^T\left(\mathbf{A}/\sigma^2\right)\mathbf{y}$ karesel formu, serbestlik derecesi $r=\rho(\mathbf{A})$ ve merkezi olmayan parametresi olan bir $\Omega = (\boldsymbol{\mu}^T \mathbf{A}\boldsymbol{\mu})/2\sigma^2$ *merkezi olmayan ki-kare* dağılışı gösterir.

Bu durumda hataların normal dağıldığı varsayımı, σ^2 ile bölünen, kareler toplamlarının ki-kare değişkeni olduğunu belirtir.

Teorem 4.2 Merkezi olmayan ki-kare dağılımı gösteren bir karesel form, $\mathbf{y}^T\mathbf{A}\mathbf{y} \sim \chi^2\left(r;\boldsymbol{\mu}^T\mathbf{A}\boldsymbol{\mu}/2\right)$ ancak ve ancak $(\mathbf{y}-\boldsymbol{\mu})^T\mathbf{A}(\mathbf{y}-\boldsymbol{\mu}) \sim \chi^2(r;0)$ durumu için merkezi ki-kare dağılır.

Ki-kare dağılımı ve karesel formların arasındaki ortogonallik özelliği anlamlılık testleri için temel oluşturmaktadır. Örneğin boş hipotezin doğru olduğu durumlarda, *t*-istatistiği bir normal sapmanın, ölçeklenmiş bir bağımsız merkezi ki-kare değişkeninin kareköküne oranını verir, bkz Bölüm Ek 2. *F*-istatistiği de bir ölçeklenmiş merkezi olmayan ki-kare değişkeninin (eğer boş hipotez doğru ise merkezi ki-kare değişkeninin) ölçeklenmiş bağımsız merkezi ki-kare değişkenine oranını vermektedir. Ölçekleme ki-kare şans değişkeninin kendi serbestlik derecesine bölünmesiyle gerçekleştirilir. Bir merkezi ki-kare dağılımının merkezi olmayan parametresi sıfıra eşittir. *F*-testinin uygulanabilmesi için pay ve paydadaki karesel formların birbirinden bağımsız olması gereklidir. Aşağıdaki teorem bu varsayımın sağlanması için gerekli koşulları tanımlamaktadır:

Teorem 4.3 Eğer şans değişkeni $\mathbf{y} \sim N(\boldsymbol{\mu}; \mathbf{V})$ dağılımına sahip, $\theta_1=\mathbf{y}^T\mathbf{A}_1\mathbf{y}$ ve $\theta_2=\mathbf{y}^T\mathbf{A}_2\mathbf{y}$ ise θ_1 ile θ_2 ancak ve ancak $\mathbf{A}_1\mathbf{V}\mathbf{A}_2=\mathbf{0}$ koşulunun sağlandığı durumlarda birbirinden bağımsızdırlar.

Daha önce belirtildiği gibi eşitlik (4.46*b*) de yer alan tüm tanım matrisleri Teorem 4.3 ün şartını sağlamaktadır.

Merkezi olmayan parametre iki nedenle önemlidir. Birincisi, *F*-oranının payının merkezi olmayan parametresinin sıfıra eşit olması boş hipotezinin açık bir ifadesini sağlar. İkincisi, merkezi olmayan parametrenin büyüklüğü ile yanlış bir boş hipotezi belirlemenin ölçüsü olan testin gücü belirlenir. Eşitlik (4.51) den görülebileceği gibi $KT(e)/\sigma^2$ bir merkezi ki-kare değişkenidir, çünkü ikinci terimi sıfıra eşittir Eşitlik (4.50) ile tanımlanan $KT(R)/\sigma^2$ için merkezi olmayan parametre,

$$\Omega = \frac{\boldsymbol{\beta}^T\mathbf{X}^T(\mathbf{I}-\mathbf{J}/n)\mathbf{X}\boldsymbol{\beta}}{2\sigma^2} \tag{4.56}$$

şeklindedir. Bu ifade β_0 hariç tüm β_j parametrelerinin oluşturduğu bir karesel formudur. Bu durumda $KT(R)/\sigma^2$ nin merkezi bir ki-kare değişkeni olabilmesi için Ω=0 olması gereklidir. Bunun içinde $(\mathbf{I}-\mathbf{J}/n)\mathbf{X}\boldsymbol{\beta}=0$ eşitliği sağlanmalıdır. **X** matrisi tam sütun ranklı olduğundan Ω=0 eşitliği ancak ve ancak, $\beta_1=\beta_2=\ldots=\beta_k=0$ durumunda gerçekleşir. Sonuç olarak, F-oranının

$$F=\frac{KO(R)}{KO(e)} \tag{4.57a}$$

β_0 hariç tüm β_j parametrelerinin sıfıra eşit olduğu hipotezini test etmek için kullanılabileceği görülmektedir. Bu hipotez,

$$H_0: \boldsymbol{\beta}^*=0 \tag{4.57b}$$

$$H_1: \boldsymbol{\beta}^*\neq 0$$

şeklinde ifade edilebilir. Burada $\boldsymbol{\beta}^*$, β_0 hariç $k\times 1$ boyutlu regresyon parametreler vektörüdür. Gözlenmiş bir F-oranının 1'den yeterince büyük olması merkezi olmayan parametrenin sıfıra eşit olmadığını belirtir. Paydaki ki-kare değişkeninin merkezi olmayan parametresi büyüdükçe F-oranı da büyüyecektir. Bunun sonucunda da yanlış bir boş hipotezin belirlenmesi olasılığı da büyüyecektir. Bu olasılık testin gücü olarak adlandırılır. Bir F-testinin gücü her bir ki-kare değişkeninin, özellikle paydadaki ki-kare değişkeninin, serbestlik derecesi arttıkça artar. Deneye başlamadan önce bu merkezi olmayan parametredeki tüm değerler $\boldsymbol{\beta}$'lar hariç bilinmektedir. **X** değerleri araştırmacının kontrolü altındadır. Bunun sonucu olarak son tasarım belirlenmeden önce tüm farklı deney tasarımlarının göreceli güçleri karşılaştırılabilir.

Karesel formların beklenen değerleri modele bağımlıdır. Eğer model yanlış oluşturulmuş ise beklenen değerler de yanlış hesaplanacaktır. Bu durum özellikle $KO(e)$ için kritik bir durum oluşturur çünkü $KO(e)$ değeri σ^2'nin tahmini değeri olarak kullanılmaktadır. F-testinin pay ve paydasında yer alan kareler ortalamaları, boş hipotez doğru olduğunda aynı beklenen değere eşit olması, alternatif hipotez doğru olduğunda ise payın beklenen değerinin daha büyük olması gereklidir.

4.10 ÇOKLU BELİRLİLİK KATSAYISI

Çoklu belirlilik katsayısı, basit regresyonu benzer olarak,

$$R^2=\frac{KT(R)}{KT(Td)}=1-\frac{KT(e)}{KT(Td)} \tag{4.58}$$

tanımlanır ve bağımsız değişkenlerin bağımlı değişkeni açıklama yeteneğinin bir ölçüsünü verir. Basit regresyondaki r^2 ye benzer olarak:

$$0\leq R^2\leq 1$$

R^2 nin 0 değerini alabilmesi için tüm b_j=0, burada j=1,…,k, olması gereklidir. Uyumu yapılan tüm değerlerin regresyon yüzeyi üzerinde $y_i=\hat{y}_i$ olması durumunda ise R^2=1 olur.

Çoklu belirlilik katsayısı R^2, gözlenmiş y_i değerleri ile uyumu yapılmış $\hat{y}_i$ değerleri arasındaki basit belirlilik katsayısı r^2 olarak nitelendirilebilir.

Bazı araştırmalarda çoklu belirlilik katsayısına ait bir boş hipotezin test edilmesi gerekli olabilir. Testin gerçekleştirilebilmesi için R^2 nin örnekleme dağılımına ihtiyaç vardır. Bir istatistiksel teoreme göre,

$$R^2 \sim B\left(\frac{1}{2}v_1, \frac{1}{2}v_2\right) \tag{4.59}$$

Bununla birlikte beta dağılımına ait kritik değerleri veren tablolara ulaşmak her zaman mümkün olmadığı için kritik değerlerin elde edilmesinde aşağıda açıklanan yöntem tercih edilebilir. H_0 hipotezi altında F hesap değeri bir $F(v_1,v_2)$ dağılımına sahiptir. Eşitlik (4.58) ile verilen R^2 istatistiği yeniden düzenlenerek,

$$R^2 = \frac{KT(R)}{KT(R) + KT(e)} \tag{4.60a}$$

ve F test istatistiği,

$$F = \frac{KT(R)/v_1}{KT(e)/v_2} \tag{4.60b}$$

şeklinde olup, eşitlik (4.60*b*) den,

$$KT(e) = \frac{v_2 KT(R)}{v_1 F}$$

elde edilir ve eşitlik (4.60*a*) da yerine konarak,

$$R^2 = \frac{v_1 F}{v_1 F + v_2} \tag{4.60c}$$

elde edilir. Sonuç olarak R^2 kullanılarak H_0 hipotezine karşı H_1 hipotezinin testi gerçekleştirilebilir. Bu testin sonucu standart F-testinin sonucuna tamamen denk olacaktır. Kritik değer ise eşitlik (4.60*c*) de F değeri yerine tablodan bulunan $F(v_1,v_2)$ değeri konarak hesaplanır.

Büyük bir R^2 değeri kullanılan modelin kesinlikle faydalı bir model olduğunu göstermez. Örneğin az sayıda bağımsız değişken seviyesinin olduğu durum ele alınsın. Bu gibi durumlarda yüksek bir R^2, uyumu yapılan model pek çok kestirim için gözlem bölgesinin dışında ekstrapolasyon gerektirdiğinden faydalı olamayabilecektir. Modele daha fazla bağımsız değişken eklenmesi R^2 değerini arttırabilir fakat asla azaltmaz. R^2 değerinin pozitif karekökü ise,

$$R = \sqrt{R^2} \tag{4.61}$$

çoklu korelasyon katsayısı olarak adlandırılır.

Modeldeki bağımsız değişken sayısı arttıkça R^2 azalmadığından modeldeki bağımsız değişken sayısını dikkate alan, düzeltilmiş çoklu belirlilik katsayısının kullanılması tercih edilir. Düzeltilmiş çoklu belirlilik katsayısı, modele yeni bir bağımsız değişken eklendiğinde azalabilir. Bunun nedeni $KT(e)$ deki azalma, paydadaki serbestlik derecesindeki kayıptan daha fazla olabilir. R^2 formülündeki kareler toplamlarının serbestlik dereceleri dikkate alınarak yapılan bir düzeltme ile R_a^2 elde edilir.

$$R_a^2 = 1 - \frac{KT(e)/(n-p)}{KT(Tc)/(n-1)} = 1 - \left(1 - R^2\right)\left(\frac{n-1}{n-p}\right) \quad (4.62)$$

R_a^2 parametre sayıları farklı modellerin karşılaştırılmasında ve tamamen farklı iki ya da daha fazla veri setinden elde edilen modellerin karşılaştırılmasında kullanılır.

4.11 HİPOTEZ TESTLERİ

Bazı durumlarda araştırmacılar gerekli olandan daha genel modelleri kullanırlar. Gerçek model bilinmediğinden kullanılması gereken modellerle ilgili bazı varsayımları ve şüpheleri de mevcuttur. Örneğin uyumu yapılan model $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$ ve alternatif olarak doğru olabileceği düşünülen model $y = \beta_0 + \beta\left(x_1 - x_2\right) + \varepsilon$ ise modelin kontrolü nasıl gerçekleştirilecektir? Bu amaçla, β_1=-β_2=β ya da eş değer olarak, $\beta_1+\beta_2=0$ olup olmadığı sorusunun cevabı araştırılmalıdır. Yukarıdaki soru istatistikte, H_0:$\beta_1+\beta_2=0$ boş hipotezine karşılık, H_1:$\beta_1+\beta_2\neq0$ alternatif hipotezin testine karşılık gelir. H_0 hipotezinin β ların doğrusal bir kombinasyonu ile ilgili bir ifadeyi içermesi nedeniyle bu tip hipotezlere *doğrusal hipotez* adı verilir.

Regresyon kareler ortalamasının artık kareler ortalamasına oranı, β_0 hariç tüm parametrelerin sıfıra eşit olduğu sıfır hipotezinin test edilmesine imkan vermektedir. Fakat bazı durumlarda yukarıdaki örnekte tanımlandığı gibi daha esnek hipotez testlerinin oluşturulması gerekebilir. Bu kısımda parametrelerin doğrusal fonksiyonlarını içeren herhangi bir hipotezin testi için genel bir metot verilecektir. Eğer boş hipotez bir tek doğrusal fonksiyonu içeriyorsa basit hipotez, birden fazla fonksiyonu eş anlı olarak içeriyorsa karmaşık hipotez olarak adlandırılır.

4.11.1 Genel Doğrusal Hipotezler

Genel doğrusal hipotez

$$H_0 : \mathbf{K}^T\boldsymbol{\beta} = \mathbf{m} \quad (4.63)$$

$$H_1 : \mathbf{K}^T\boldsymbol{\beta} \neq \mathbf{m}$$

şeklinde tanımlanır. Burada $\mathbf{K}^T$, $r\times p$ boyutlu test edilecek parametrelerin r adet doğrusal fonksiyonunu tanımlayan katsayılar matrisidir. $\mathbf{K}^T$ matrisinin her bir sırası bir doğrusal fonksiyonun katsayılarını içerir. **m** ise genellikle sıfırlardan oluşan $r\times1$ boyutlu bir sabitler vektörüdür. H_0 daki r adet doğrusal denklem doğrusal bağımsız olmalıdır, fakat ortogonal olmaları gerekli değildir. Doğrusal bağımsızlık $\mathbf{K}^T$ matrisinin tam ranklı $\rho(\mathbf{K}^T)=r$ olduğu belirtir. Bunun sonucu olarak da H_0 daki denklemler **m**'nin her türlü seçimi için tutarlı olacaktır. H_0 daki doğrusal fonksiyon sayısı $\boldsymbol{\beta}$ daki parametre sayısını aşamaz, aştığı durumlarda $\mathbf{K}^T$ matrisi tam ranklı olmayacaktır. Örneğin parametre vektörünün $\boldsymbol{\beta}^T=(\beta_0\ \beta_1\ \beta_2\ \beta_3)$ olduğu ve $\beta_1=\beta_2$, $\beta_1+\beta_2=2\beta3$, $\beta_0=20$ hipotezlerinin test edilmek istendiği kabul edilsin, bu durumda H_0 hipotezi:

$$\begin{array}{rrrrcl} & \beta_1 & +\beta_2 & & = & 0 \\ & \beta_1 & +\beta_2 & -2\beta_3 & = & 0 \\ \beta_0 & & & & = & 20 \end{array}$$

şeklindedir. Bu üç doğrusal fonksiyon $\mathbf{K}^T\boldsymbol{\beta}=\mathbf{m}$ formunda

$$\mathbf{K}^T\boldsymbol{\beta}=\begin{bmatrix}0&1&1&0\\0&1&1&-2\\1&0&0&0\end{bmatrix}\begin{bmatrix}\beta_0\\\beta_1\\\beta_2\\\beta_3\end{bmatrix},\qquad \mathbf{m}=\begin{pmatrix}0\\0\\20\end{pmatrix}$$

yazılabilir. Alternatif hipotez H_1: $\mathbf{K}^T\boldsymbol{\beta}\neq\mathbf{m}$ şeklindedir. Boş hipotezin geçersiz olabilmesi için H_0 daki hipotezlerden herhangi birinin (veya daha fazlasının) yanlış olması gerekmektedir. $\mathbf{K}^T\boldsymbol{\beta}$-$\mathbf{m}$ için EKK tahminleri, $\boldsymbol{\beta}$ yerine $\mathbf{b}$ yazılarak $\mathbf{K}^T\mathbf{b}$-$\mathbf{m}$ şeklinde elde edilir. Normal dağılışı da kapsayan EKK varsayımları altında, $\mathbf{K}^T\mathbf{b}$-$\mathbf{m}$, ortalaması,

$$E(\mathbf{K}^T\mathbf{b}-\mathbf{m})=\mathbf{K}^T\boldsymbol{\beta}-\mathbf{m} \tag{4.64}$$

eğer boş hipotez doğru ise $\mathbf{K}^T\mathbf{b}$-$\mathbf{m}$=$\mathbf{0}$ olacaktır ve varyansı,

$$\begin{aligned}V\left(\mathbf{K}^T\mathbf{b}-\mathbf{m}\right)&=\mathbf{K}^T V\left(\mathbf{b}\right)\mathbf{K}\sigma^2\\&=\mathbf{K}^T\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{K}\sigma^2=\mathbf{V}\sigma^2\end{aligned} \tag{4.65}$$

olan bir normal dağılış gösterecektir. Çünkü şans vektörü $\mathbf{b}$ normal dağılmaktadır. Eşitlik (4.63) ile tanımlanan doğrusal hipotez için kareler toplamı,

$$\theta=\left(\mathbf{K}^T\mathbf{b}-\mathbf{m}\right)^T\left[\mathbf{K}^T\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{K}\right]^{-1}\left(\mathbf{K}^T\mathbf{b}-\mathbf{m}\right) \tag{4.66}$$

eşitliğinden hesaplanır. Karesel form, şans değişkeni $\mathbf{K}^T\mathbf{b}$-$\mathbf{m}$ nin karesinin varyansına oranı şeklinde elde edilmiştir. Bu nedenle tanım matrisi,

$$\mathbf{A}=\left[\mathbf{K}^T\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{K}\right]^{-1}=\mathbf{V}^{-1} \tag{4.67}$$

σ^2 sabit çarpanı hariç, şans vektörünün (H_0 hipotezindeki doğrusal fonksiyonların) varyans-kovaryans matrisinin tersidir. Buna göre karesel formun serbestlik derecesi, $\mathbf{AV}$=$\mathbf{I}$ idempotent matris olduğundan,

$$tr\left(\mathbf{AV}\right)=tr\left(\mathbf{I}_r\right)=r \tag{4.68}$$

burada $\rho(\mathbf{A})=\rho(\mathbf{K})=r$ dir ve beklenen değeri,

$$E\left(\theta\right)=r\sigma^2+\left(\mathbf{K}^T\boldsymbol{\beta}-\mathbf{m}\right)^T\left[\mathbf{K}^T\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{K}\right]^{-1}\left(\mathbf{K}^T\boldsymbol{\beta}-\mathbf{m}\right) \tag{4.69}$$

şeklinde bulunur. Normallik varsayımı kullanılarak, θ/σ^2 karesel formunun r serbestlik dereceli merkezi olmayan bir ki-kare dağılışı gösterdiği söylenebilir. Merkezi olmayan parametre ise,

$$\Omega=\frac{\left(\mathbf{K}^T\boldsymbol{\beta}-\mathbf{m}\right)^T\left[\mathbf{K}^T\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{K}\right]^{-1}\left(\mathbf{K}^T\boldsymbol{\beta}-\mathbf{m}\right)}{2\sigma^2} \tag{4.70}$$

şeklindedir. Eğer H_0 hipotezi doğru ise bu parametre sıfıra eşit olacaktır. Bu durumda boş hipotezi test etmek için kullanılacak F-istatistiğinin pay kısmı $KO(\theta)=\theta/r$ şeklinde verilen bir kareler ortalamasına sahip olacaktır. Payda kısmı ise σ^2 nin sapmasız bir tahmini olan $KO(e)$ den oluşacaktır.

$$F = \frac{\theta/\rho(\mathbf{K})}{KO(e)} = \frac{\theta/r}{s^2} \tag{4.71}$$

Verilen hipotezler için tüm kareler toplamları kullanılan modele bağlıdır. Bu nedenle bir bağımsız değişkenin modelden çıkarılması ya da ilave edilmesi her bir hipotez için kareler toplamlarını değiştirecektir.

4.11.2 Genel Formun Bazı Özel Durumları

Genel doğrusal hipotezlerin beş özel durumu ile ilgilenilecektir.

Durum 1. Basit Hipotez-Tek bir parametrenin testi

β vektörü üzerindeki bir basit hipotezde,

$$H_0 : \beta_i = 0, \tag{4.72}$$

$\mathbf{K}^T$ bir satır vektörüdür, $\mathbf{K}^T = \begin{bmatrix} 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \end{bmatrix}$. Burada 1 değeri β_j ye karşılık gelen pozisyondadır. Buna karşın $\mathbf{b}^T = \begin{bmatrix} b_0 & b_1 & b_2 & b_3 & \cdots & b_k \end{bmatrix}$ olduğundan $\mathbf{K}^T\mathbf{b}=b_j$ bir skalerdir ve $\mathbf{m}=0$. Benzer şekilde $\mathbf{K}^T\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{K}$ değeri ve onun tersi de skalerdir. Eşitlik (4.66) ile verilen hipotez kareler toplamı:

$$KT(H_0) = \theta = \frac{\left(\mathbf{K}^T\mathbf{b}-\mathbf{m}\right)^2}{\mathbf{K}^T\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{K}} = \frac{b_i^2}{c_{ii}} \tag{4.73}$$

Burada c_{jj} değeri, $(\mathbf{X}^T\mathbf{X})^{-1}$ matrisinin j+1-inci elemanıdır. Serbestlik derecesi 1'dir. θ nın payı b_j'nin karesi, paydası ise σ^2 hariç b_i'nin varyansına eşittir. Hesaplanan test değeri,

$$F = \frac{b_i^2}{c_{ii}s^2} \tag{4.74a}$$

olup tablo değeri $F(1,n\text{-}p)$. Eğer hipotez,

$$H_0 : \beta_i = \beta_{i0}$$

gibi sıfırdan farklı bir değerin testi ise, bu durumda $\mathbf{m}=\beta_{i0}$ olup, test istatistiği

$$F = \frac{\left(b_i - \beta_{i0}\right)^2}{c_{ii}s^2} \tag{4.74b}$$

şeklindedir. Bu hipotez için kareler toplamı, tüm diğer bağımsız değişkenler de modelde iken, x_i bağımsız değişkenin modele yaptığı katkıyı ölçer. Bu kareler toplamı i-inci bağımsız değişken için *kısmi kareler toplamı* olarak adlandırılır.

Basit hipotezin F-testi, iki yönlü bir t-testi olarak

$$t = \frac{\left(\mathbf{K}^T\mathbf{b}-\mathbf{m}\right)}{\left\{\left[\mathbf{K}^T\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{K}\right]s^2\right\}^{1/2}} = \frac{b_i}{s\sqrt{c_{ii}}} \tag{4.74c}$$

ifade edilebilir. Bu ifadede payda, paydaki doğrusal fonksiyonun standart hatasıdır.

Durum 2. Basit Hipotez-Parametrelerin bir doğrusal fonksiyonunun testi

Aşağıda örnekleri gösterilen β vektörü üzerinde tanımlanan tek bir doğrusal fonksiyonda basit bir hipotez olarak değerlendirilir:

$$H_0 : \beta_2 - \beta_3 = 0 \qquad (4.75a)$$

$\mathbf{K}^T = [0 \quad 0 \quad 1 \quad -1 \quad \cdots \quad 0]$, $\mathbf{b}^T = [b_0 \quad b_1 \quad b_2 \quad b_3 \quad \cdots \quad b_k]$ ve $\mathbf{m}$=0

$$H_0 : \beta_1 + \beta_2 = 1 \qquad (4.75b)$$

$\mathbf{K}^T = [0 \quad 1 \quad 1 \quad 0 \quad \cdots \quad 0]$, $\mathbf{b}^T = [b_0 \quad b_1 \quad b_2 \quad b_3 \quad \cdots \quad b_k]$ ve $\mathbf{m}$=1

Her iki örnek için de $\mathbf{K}^T$ bir satır vektörü olup $\mathbf{K}^T\mathbf{b}=b_i \pm b_j$ bir skalerdir. Eşitlik (4.75) ile verilen hipotezlerin kareler toplamı:

$$KT(H_0) = \theta = \frac{(\mathbf{K}^T\mathbf{b} - \mathbf{m})^2}{\mathbf{K}^T(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{K}} \qquad (4.76)$$

Paydadaki tanım matrisi bir skalerdir ve serbestlik derecesi 1'dir. Hesaplanan test değeri,

$$F = \frac{(\mathbf{K}^T\mathbf{b} - \mathbf{m})^2}{\mathbf{K}^T(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{K}s^2} \qquad (4.77a)$$

olup paydadaki ifade payda verilen doğrusal fonksiyonun varyansını tanımlamaktadır. Eşitlik (4.75) ile verilen hipotez basit bir hipotez olduğundan eşitlik (4.77*a*) daki *F*-testi iki yönlü bir *t*-testinin karesidir:

$$t = \frac{\mathbf{K}^T\mathbf{b} - \mathbf{m}}{\left[\mathbf{K}^T(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{K}s^2\right]^{1/2}} \qquad (4.77b)$$

Bu test için kritik değer (*n*-2) serbestlik dereceli bir *t*-dağılımından belirlenir.

Durum 3. Karmaşık Hipotez-Genel anlamlılık testi

Eşitlik (4.57*b*) ile tanımlanan hipotez,

$$H_0 : \begin{array}{l} \beta_1 = 0 \\ \beta_2 = 0 \\ \quad\vdots \\ \beta_k = 0 \end{array} \qquad (4.78)$$

şeklinde yeniden tanımlanabilir. Hipotez *k* adet birbirinden bağımsız doğrusal fonksiyon içermektedir. Bu hipoteze karşılık gelen katsayılar matrisi,

$$\mathbf{K}^T = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \qquad (4.79)$$

$k \times p$ boyutlu olup, $\rho(\mathbf{K})=k$ dır. $\mathbf{b}^T = [b_0 \quad b_1 \quad b_2 \quad b_3 \quad \cdots \quad b_k]$ $1 \times p$ boyutlu vektör ve $\mathbf{m}=\mathbf{0}$ ise $k \times 1$ boyutlu sıfırlardan oluşan vektördür. Eşitlik (4.78) ile tanımlanan doğrusal hipotez için kareler toplamı, eşitlik (4.79) da verilen tanım matrisi kullanılarak eşitlik (4.66) elde edilir. Hesaplanan test istatistiği,

$$F = \frac{\theta / \rho(\mathbf{K})}{KO(e)} = \frac{\theta / k}{s^2} \qquad (4.80)$$

ve kritik değer $F(k,n\text{-}p)$. Elde edilen sonuç eşitlik (4.57*a*) dan elde edilen sonuca eşit olacaktır.

Durum 4. Karmaşık Hipotez-Parametrelerin bir alt setinin sıfıra eşitliğinin testi

Bazı araştırmalarda baz alınan bir modele birkaç bağımsız değişkenin eklenmesinin anlamlı olup olmadığı araştırılmak istenir. Bu gibi durumlar için kullanılabilecek hipotez testi aşağıda açıklanmıştır. İlk olarak modele eklenmesi araştırılan r adet değişkene karşılık gelen parametreler için hipotez,

$$H_0: \quad \begin{aligned} \beta_{p-r} &= 0 \\ \beta_{p-r+1} &= 0 \\ &\vdots \\ \beta_p &= 0 \end{aligned} \qquad (4.81)$$

kurulur**.** Hipotez r adet birbirinden bağımsız doğrusal fonksiyon içermektedir. Hipotezden de anlaşılacağı gibi eklenmesi düşünülen değişkenler modeldeki son r adet değişkene karşılık gelmektedir. Hipotez son r adet parametrenin sıfıra eşitliğini ortak olarak test eder. Hipotezdeki doğrusal fonksiyonların katsayılar matrisi ise $r \times p$ boyutlu,

$$\mathbf{K}^T = [\mathbf{0} \quad \mathbf{I}_r] = \begin{bmatrix} 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \qquad (4.82)$$

Burada $\mathbf{0}$ boyutu $r \times (p\text{-}r)$ olan sıfır matrisidir. $\mathbf{I}_r$ ise boyutu $r \times r$ olan birim matristir. $\mathbf{m=0}$ ise r elemanlı sıfırlardan oluşan sütun vektörüdür. Hipotez kareler toplamını daha kolay elde edebilmek amacıyla $\mathbf{X}$ ve $\mathbf{b}$ yeniden düzenlenir:

$$\mathbf{y} = [\mathbf{X}_i \quad \mathbf{X}_r] \begin{bmatrix} \mathbf{b}_i \\ \mathbf{b}_r \end{bmatrix} + \mathbf{e} = \mathbf{X}_i \mathbf{b}_i + \mathbf{X}_r \mathbf{b}_r + \mathbf{e} \qquad (4.83)$$

Burada $\mathbf{X}_i$ boyutu $n \times (p\text{-}r)$ olup $\mathbf{X}$ deki ilk $p\text{-}r$ sütundan oluşur. $\mathbf{X}_r$ ise $n \times r$ boyutlu olup son r sütundan oluşur. $(\mathbf{X}^T\mathbf{X})$ matrisi ise,

$$\left(\mathbf{X}^T\mathbf{X}\right) = \begin{bmatrix} \mathbf{X}_i^T\mathbf{X}_i & \mathbf{X}_i^T\mathbf{X}_r \\ \mathbf{X}_r^T\mathbf{X}_i & \mathbf{X}_r^T\mathbf{X}_r \end{bmatrix} \qquad (4.84)$$

şeklinde bölümlenir. Hipotez kareler toplamı θ yı elde etmek için,

$$\mathbf{K}^T\mathbf{b} - \mathbf{m} = [\mathbf{0} \quad \mathbf{I}_r] \begin{bmatrix} \mathbf{b}_i \\ \mathbf{b}_r \end{bmatrix} - \mathbf{0} = \mathbf{b}_r$$

ve

$$\mathbf{K}^T\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{K} = [\mathbf{0} \quad \mathbf{I}_r] \begin{bmatrix} \left(\mathbf{X}_i^T\mathbf{X}_i\right)^{-1} & \left(\mathbf{X}_i^T\mathbf{X}_r\right)^{-1} \\ \left(\mathbf{X}_r^T\mathbf{X}_i\right)^{-1} & \left(\mathbf{X}_r^T\mathbf{X}_r\right)^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{0} \\ \mathbf{I}_r \end{bmatrix} = \left(\mathbf{X}_r^T\mathbf{X}_r\right)^{-1}$$

bulunur. $\left(\mathbf{X}_r^T\mathbf{X}_r\right)^{-1}$ matrisi ise, eşitlik (3.82) ya da (3.83) kullanılarak,

$$\left(\mathbf{X}_r^T\mathbf{X}_r\right)^{-1}=\left[\mathbf{X}_r^T\mathbf{X}_r-\mathbf{X}_r^T\mathbf{X}_i\left(\mathbf{X}_i^T\mathbf{X}_i\right)^{-1}\mathbf{X}_i^T\mathbf{X}_r\right]^{-1}=\left\{\mathbf{X}_r^T\left[\mathbf{I}-\mathbf{X}_i\left(\mathbf{X}_i^T\mathbf{X}_i\right)^{-1}\mathbf{X}_i^T\right]\mathbf{X}_r\right\}^{-1}$$

$$\left(\mathbf{X}_r^T\mathbf{X}_r\right)^{-1}=\left(\mathbf{X}_r^T\mathbf{M}_i\mathbf{X}_r\right)^{-1}$$

olup burada,

$$\mathbf{M}_i=\left[\mathbf{I}-\mathbf{X}_i\left(\mathbf{X}_i^T\mathbf{X}_i\right)^{-1}\mathbf{X}_i^T\right]. \tag{4.85}$$

Sonuç olarak hipotez kareler toplamı,

$$\theta=\mathbf{b}_r\left(\mathbf{X}_r^T\mathbf{M}_i\mathbf{X}_r\right)\mathbf{b}_r^T \tag{4.86}$$

Hesaplanan test istatistiği,

$$F=\frac{\theta/\rho(\mathbf{K})}{KO(e)}=\frac{\theta/r}{s^2} \tag{4.87}$$

ve kritik değer $F(r,n\text{-}p)$. Bu hipotez için kareler toplamı, tüm diğer bağımsız değişkenler de modelde iken, r adet bağımsız değişkenin modele yaptığı katkıyı ölçer.

Durum 5. Karmaşık Hipotez-Parametrelerin bir alt setinin sıfırdan farklı değerlere eşitliğinin testi

Bazı araştırmalarda baz alınan bir modeldeki parametrelerin teorik olarak bilinen bazı değerlere eşit olup olmadığı araştırılmak istenir. Bu gibi durumlar için kullanılabilecek hipotez testi aşağıda açıklanmıştır. Parametreler için hipotez,

$$\begin{aligned} H_0:\quad &\beta_0=\beta_{00}\\ &\beta_1=\beta_{10}\\ &\vdots\\ &\beta_k=\beta_{k0} \end{aligned} \tag{4.88}$$

kurulur**.** Hipotez p adet birbirinden bağımsız doğrusal fonksiyon içermektedir. Hipotezdeki doğrusal fonksiyonların katsayılar matrisi ise $p\times p$ boyutlu,

$$\mathbf{K}^T=\mathbf{I}_p=\begin{bmatrix}1&\cdots&0\\ \vdots&\vdots&\vdots\\ 0&\cdots&1\end{bmatrix} \tag{4.89}$$

birim matristir. Hipotezde test edilen değerler ise $\mathbf{m}=\boldsymbol{\beta}_0$ şeklinde r elemanlı sütun vektörü olarak ifade edilebilir. Sonuç olarak $\mathbf{K}^T\mathbf{b}-\mathbf{m}=(\mathbf{b}-\boldsymbol{\beta}_0)$ ve karesel formun tanım matrisi $[\mathbf{K}^T(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{K}]^{-1}=(\mathbf{X}^T\mathbf{X})$ olacaktır. Karesel form ise,

$$\theta=\left(\mathbf{b}-\boldsymbol{\beta}_0\right)^T\left(\mathbf{X}^T\mathbf{X}\right)\left(\mathbf{b}-\boldsymbol{\beta}_0\right) \tag{4.90}$$

olup, tanım matrisinin izi p olduğundan karesel formun serbestlik derecesi p olacaktır. Eşitlik (4.71) deki F-testi:

$$F=\frac{\left(\mathbf{b}-\boldsymbol{\beta}_0\right)^T\left(\mathbf{X}^T\mathbf{X}\right)\left(\mathbf{b}-\boldsymbol{\beta}_0\right)}{ps^2} \tag{4.91}$$

Kritik değer ise $F(p,n\text{-}p)$.

4.11.3 Ekstra Kareler Toplamları

Regresyon çalışmalarında oldukça sık ortaya çıkan sorulardan biri modelde yer alan değişkenlerin modele yapmış oldukları katkının yeterli faydayı sağlayıp sağlamadığıdır. Bu soru modelin regresyon kareler toplamı bileşenlere ayrıştırılarak araştırılabilir. Bu bileşenler *ekstra kareler toplamı* olarak adlandırılır. Ekstra kareler toplamlarından elde edilen kareler ortalamalarının istatistiksel olarak anlamlı olup olmadıkları s^2 ile test edilerek belirlenebilir. Eğer kareler ortalaması istatistiksel olarak anlamlı ise ilgili değişkenler modele dahil edilir. Sonuç anlamsız ise değişkenler modele dahil edilmez.

Ekstra kareler toplamlarına örnek olarak $KT(b_1 \backslash b_0)$ verilebilir. Bu ifade sabit terimli bir modele $\beta_1 x_1$ teriminin yaptığı katkıyı belirtir. Konuyu daha iyi açıklamak üzere aşağıdaki iki model ele alınsın:

$$y_i = \beta_0 + \beta_1 x_1 + \ldots + \beta_{q-1} x_{q-1} + \ldots + \beta_k x_k + \varepsilon_i \qquad \text{Model 1}$$

$$y_i = \beta_0 + \beta_1 x_1 + \ldots + \beta_{q-1} x_{q-1} + \varepsilon_i \qquad \text{Model 2}$$

Burada $q<k$ dır. Bir model için regresyon kareler toplamı $KT(R)$ ile gösterilmekte idi. Fakat bu gösterimde modelde yer alan terimlere ait bilgi olmadığından bu eksiği gidermek için genellikle R gösterimi tercih edilir. Yukarıdaki iki model için R notasyonunda regresyon kareler toplamları:

$$R_1(\beta_1,\ldots,\beta_k|\beta_0) = KT(R_1) \tag{4.92a}$$

$$R_2(\beta_1,\ldots,\beta_{q-1}|\beta_0) = KT(R_2) \tag{4.92b}$$

ve model kareler toplamları:

$$R_1(\beta_0,\beta_1,\ldots,\beta_k) = KT(M_1) \tag{4.92c}$$

$$R_2(\beta_0,\beta_1,\ldots,\beta_{q-1}) = KT(M_2) \tag{4.92d}$$

Bu iki modelin kareler toplamları arasındaki fark ise ekstra kareler toplamı olarak adlandırılır.

$$\theta = KT(R_1) - KT(R_2) \tag{4.93a}$$

Model 1'in içerdiği terim sayısı p adet Model 2'nin içerdiği terim sayısı q adet olduğu için $KT(R_1) \geq KT(R_2)$ olması beklenir. Eğer Model 1'in içerdiği $p\text{-}q$ adet terimin modele yaptığı katkı önemsiz ise $E[KT(R_1)]=E[KT(R_2)]$ olması beklenir. Bu durum $p\text{-}q$ adet parametrenin sıfıra eşit olması ile gerçekleşir. Kurulacak hipotez ve test yöntemi bir önceki kısımda Durum 4 ile açıklanmıştı. Eşitlik (4.93*a*) dan Durum 4 için elde edilen kareler toplamının, ekstra kareler toplamına eşit olduğu görülmektedir. İki model arasındaki terim farkı $p\text{-}q=r$ olarak alındığında eşitlik (4.87) nin ekstra kareler toplamı için uygulanabilecek bir anlamlılık testi olduğu görülmektedir.

Daha fazla terim içerdiği için Model 1 *tam model*, daha az terim içerdiği için Model 2 *indirgenmiş model* olarak adlandırılır.

Eşitlik (4.93*a*) ile verilen ekstra kareler toplamı artık kareler toplamı üzerinden de hesaplanabilir. Bununla birlikte $KT(e_2) \geq KT(e_1)$ olduğundan ekstra kareler toplamı,

$$\theta = KT(e_2) - KT(e_1) \tag{4.93b}$$

eşitliği ile tanımlanabilir.

Eşitlik (4.93a) R gösterimi ile de ifade edilebilir. Eşitlik (4.92) ve eşitlik (4.93a) kullanılarak,

$$\theta = R_1\left(\beta_1,\ldots,\beta_k \middle| \beta_0\right) - R_2\left(\beta_1,\ldots,\beta_{q-1} \middle| \beta_0\right)$$
$$= R\left(\beta_q,\ldots,\beta_k \middle| \beta_0,\ldots,\beta_{q-1}\right) \qquad (4.94)$$

elde edilir. Bu ifade, ($\beta_0,\ldots,\beta_{q-1}$) modelde iken ($\beta_q,\ldots,\beta_k$)'nın regresyon kareler toplamına yaptığı katkı olarak okunur. Tam model (Model 1) ile indirgenmiş model (Model 2) arasındaki farkı tanımlayan r adet parametrenin anlamlılık testi θ nın hesaplanmasında birbirine denk olan eşitlik (4.86), (4.93) ya da (4.94) kullanılarak, eşitlik (4.87) de tanımlanan F-testi ile gerçekleştirilir.

4.11.4 Ardışık Kareler Toplamları

Ekstra kareler toplamındaki yaklaşım kullanılarak her hangi bir model için; $R(\beta_0)$, $R(\beta_1\backslash\beta_0)$, $R(\beta_2\backslash\beta_0, \beta_1),\ldots, R(\beta_k\backslash\beta_0, \beta_1,\ldots, \beta_{k-1})$, şeklinde ardışık kareler toplamları elde edilebilir. Bu kareler toplamlarının tümü s^2'den bağımsız olup her biri 1 serbestlik derecesine sahip olduklarından kareler toplamları kareler ortalamalarına eşittir.

Ardışık kareler toplamları, değişkenlerin modele belirli bir sıra ile eklenmeleri durumunda her birinin kendinden önceki modelin kareler toplamına yaptığı katkıyı hesaplar. Örneğin iki değişkenli bir model ele alınsın değişkenlerin modele alınma sıraları da x_1 ve x_2 olsun. x_1 değişkeninin modele yaptığı katkı $R(\beta_1\backslash\beta_0)$ ve x_1 değişkeninin bulunduğu modelin kareler toplamına x_2 değişkeninin yaptığı katkı ise $R(\beta_2\backslash\beta_0, \beta_1)$ ile tanımlanır. Örneğin,

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4} + \varepsilon_i \qquad (4.95)$$

modeli için ardışık kareler toplamları,

$R(\beta_1\backslash\beta_0)$ $\quad$ $R(\beta_3\backslash\beta_0, \beta_1, \beta_2)$

$R(\beta_2\backslash\beta_0, \beta_1)$ $\quad$ $R(\beta_4\backslash\beta_0, \beta_1, \beta_2, \beta_3)$

şeklindedir. Ardışık kareler toplamları tam modelin regresyon kareler toplamını, birbiri ile eklenebilir bileşenlere ayırır. Eşitlik (4.95) modeli için:

$$KT(R) = R\left(\beta_1,\beta_2,\beta_3,\beta_4 \middle| \beta_0\right)$$
$$= R\left(\beta_1 \middle| \beta_0\right) + R\left(\beta_2 \middle| \beta_0,\beta_1\right) + R\left(\beta_3 \middle| \beta_0,\beta_1,\beta_2\right) + R\left(\beta_4 \middle| \beta_0,\beta_1,\beta_2,\beta_3\right) \qquad (4.96a)$$

Eşitlik (4.96a) p parametreli bir model için aşağıdaki,

$$KT(R) = R\left(\beta_1,\beta_2,\ldots,\beta_k \middle| \beta_0\right)$$
$$= R\left(\beta_1 \middle| \beta_0\right) + R\left(\beta_2 \middle| \beta_0,\beta_1\right) + \ldots + R\left(\beta_k \middle| \beta_0,\beta_1,\ldots,\beta_{k-1}\right) \qquad (4.96b)$$

şekilde genellenebilir. Ardışık kareler toplamlarının eklenebilirlik özelliği kullanılarak tam ve indirgenmiş modeller arasındaki ekstra kareler toplamı hesaplanabilir. Eşitlik (4.95) daki model ele alınsın. İndirgenmiş modelin $H_0{:}\beta_3{=}\beta_4{=}0$ boş hipotezine karşılık gelen model olduğu varsayılsın. Bu durumda test edilecek kareler toplamı:

$$\theta = R\left(\beta_1,\beta_2,\beta_3,\beta_4 \middle| \beta_0\right) - R\left(\beta_1,\beta_2 \middle| \beta_0\right)$$

eşitliğin sağındaki ikinci bileşene eşitlik (4.96b) uygulanarak,

$$\theta = R(\beta_1, \beta_2, \beta_3, \beta_4 | \beta_0) - [R(\beta_1|\beta_0) + R(\beta_2|\beta_0, \beta_1)]$$

olarak elde edilir. Eşitlik (4.96b) nin sağındaki bileşenler kullanılarak birbirinden bağımsız kareler toplamlarından türetilen testler ardışık F-testleri olarak adlandırılır. Tablo 4.1 bu amaçla kullanılan bir varyans analiz tablosunu göstermektedir.

Ardışık F-testleri modeldeki mevcut yapıya bir değişken eklenmesi durumunda bu değişken yaptığı katkı ile ilgili bilgi verir. Değişkenlerin modele dahil edilmesi sırası sonuçlar üzerinde önemli bir etkiye sahiptir. Örneğin (4.95) modeli için x_1, x_2 ve x_3 değişkenlerinin bulunduğu modele x_4 ün yaptığı katkı sadece x_1 in bulunduğu modele x_4 ün yaptığı katkıdan tamamen farklıdır. Diğer bir değişle bir değişkenin model için uygunluğu, genelde modelde hangi değişkenlerin bulunduğuna bağımlıdır. Bu nedenle ardışık F-testleri tam ölçekli bir değişken eleme yaklaşımında etkin olarak kullanılamaz.

Tablo 4.1 *Ardışık F-testleri için varyans analiz tablosu*

Değişkenlik Kaynağı	*Kareler Toplamları*	*Serbestlik Derecesi*	*Kareler Ortalaması*
Regresyon	$R(\beta_1, \beta_2, \ldots, \beta_k \| \beta_0)$	k	$R(\beta_1, \beta_2, \ldots, \beta_k \| \beta_0)/k$
β_1 *in katkısı*	$R(\beta_1 \| \beta_0)$	1	$R(\beta_1 \| \beta_0)$
β_2 *in katkısı*	$R(\beta_2 \| \beta_0, \beta_1)$	1	$R(\beta_2 \| \beta_0, \beta_1)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
β_k *in katkısı*	$R(\beta_k \| \beta_0, \beta_1, \ldots, \beta_{k-1})$	1	$R(\beta_k \| \beta_0, \beta_1, \ldots, \beta_{k-1})$
Artık	$(\mathbf{y}^T\mathbf{y} - n\bar{\mathbf{y}}^2) - R(\beta_1, \beta_2, \ldots, \beta_k \| \beta_0)$	n-p	s^2
Toplam düz. Değişkenlik	$\mathbf{y}^T\mathbf{y} - n\bar{\mathbf{y}}^2$	n-1	

4.11.5 Kısmi Kareler Toplamları

Bir araştırmada kullanılan bağımsız değişken kümesinin x_1, x_2,…, x_k olduğu varsayılsın. Herhangi bir x_j değişkeni için kısmi kareler toplamı,

$$R(\beta_j | \beta_0, \ldots, \beta_{j-1}, \beta_{j+1}, \ldots, \beta_k) \qquad (4.97)$$

şeklinde gösterilir. Eşitlik (4.97) den görüldüğü gibi kısmi kareler toplamı, bağımsız değişkenler kümesindeki tüm değişkenler modelde iken j-inci değişkenin oluşturduğu katkıyı belirtir. Kısmi kareler toplamlarının her biri bir serbestlik derecesine sahiptir. Uygun bir model kurulduktan sonra modelden bir değişkeni çıkarmak ya da modele bir değişken eklemek için kısmi F-testleri kullanılabilir. Kısmi kareler toplamları, regresyon kareler toplamını verecek şekilde,

$$R(\beta_1 | \beta_0, \beta_2, \ldots, \beta_k) + R(\beta_2 | \beta_0, \beta_1, \ldots, \beta_k) + \ldots + R(\beta_k | \beta_0, \beta_1, \ldots, \beta_{k-1}) \neq R(\beta_1, \beta_2, \ldots, \beta_k | \beta_0) \qquad (4.98)$$

eklenebilirlik özelliğine sahip değildir. Bu nedenle H_0:β_j=β_k=0 şeklinde karmaşık hipotezlerin testi için uygun değildirler. Bununla H_0:β_j=0 şeklindeki basit hipotezlerin testinde kullanılabilirler. Kısmi F-testleri için varyans analiz tablosu Tablo 4.2 de verilmiştir.

Tablo 4.2 *Kısmi F-testleri için varyans analiz tablosu*

Değişkenlik Kaynağı	*Kareler Toplamları*	*Serbestlik Derecesi*	*Kareler Ortalaması*	*Hipotez H_0*
Regresyon	$R(\beta_1,\beta_2,\ldots,\beta_k\|\beta_0)$	k	$R(\beta_1,\beta_2,\ldots,\beta_k\|\beta_0)/k$	
β_1 *in katkısı*	$R(\beta_1\|\beta_0,\beta_2,\ldots,\beta_k)$	1	$R(\beta_1\|\beta_0,\beta_2,\ldots,\beta_k)$	$\beta_1=0$
β_2 *in katkısı*	$R(\beta_2\|\beta_0,\beta_1,\ldots,\beta_k)$	1	$R(\beta_2\|\beta_0,\beta_1,\ldots,\beta_k)$	$\beta_2=0$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	
β_k *in katkısı*	$R(\beta_k\|\beta_0,\beta_1,\ldots,\beta_{k-1})$	1	$R(\beta_k\|\beta_0,\beta_1,\ldots,\beta_{k-1})$	$\beta_k=0$
Artık	$(\mathbf{y}^T\mathbf{y}-n\bar{\mathbf{y}}^2)-R(\beta_1,\beta_2,\ldots,\beta_k\|\beta_0)$	*n-p*	s^2	
Toplam düz. Değişkenlik	$\mathbf{y}^T\mathbf{y}-n\bar{\mathbf{y}}^2$	n-1		

4.12 X MATRİSİNDEKİ ORTOGONAL SÜTUNLAR

Kısmi kareler toplamları, regresyon kareler toplamını verecek şekilde, eklenebilirlik özelliğine sahip olmamasının temel nedeni **X** matrisindeki sütunların birbirine ortogonal olmamasıdır. Bir regresyon probleminin β_0, β_1, β_2 parametrelerini içerdiğini kabul edilsin Ardışık kareler toplamı prensibi kullanılarak aşağıda verilen değerleri hesaplanabilir.

$KT(\beta_2)=R(\beta_2)$ değeri $y=\beta_2x_2+\varepsilon$ modelinden,

$KT(\beta_2\backslash\beta_0)=\mathrm{R}(\beta_2\backslash\beta_0)$ değeri $y=\beta_0+\beta_2x_2+\varepsilon$ modelinden,

$KT(\beta_2\backslash\beta_0,\beta_1)=\mathrm{R}(\beta_2\backslash\beta_0,\beta_1)$ değeri $y=\beta_0+\beta_1x_1+\beta_2x_2+\varepsilon$ modelinden.

Bu değerler **X** matrisinin β_2 sütunun β_0 ve $\beta_1$sütunlarına ortogonal olmadığı durumlarda genellikle tamamen farklı değerler alırlar. Bu durum aşağıda detaylı olarak ele alınacaktır.

$\mathbf{y}=\mathbf{X\beta}+\boldsymbol{\varepsilon}$ modelinde **X** matrisinin t adet alt matristen oluşan bir seti ele alınsın, $\mathbf{X}=[\mathbf{X}_1\ \mathbf{X}_2\ \ldots\ \mathbf{X}_t]$. Bu sete karşılık gelen $\boldsymbol{\beta}$ vektörü, $\boldsymbol{\beta}^T=[\ \boldsymbol{\beta}_1\ \ \boldsymbol{\beta}_2\ \ldots\ \boldsymbol{\beta}_t]$ olarak verilirse model,

$$E(y)=\mathbf{X}_1\boldsymbol{\beta}_1+\mathbf{X}_2\boldsymbol{\beta}_2+\ldots+\mathbf{X}_t\boldsymbol{\beta}_t$$

şeklinde yazılır. Parametre tahmini **b** vektörü ise $\mathbf{b}^T=[\ \mathbf{b}_1\ \ \mathbf{b}_2\ \ldots\ \mathbf{b}_t]$ vektörü ile gösterilir. Bu model için **β**'nın tahmin vektörü $\mathbf{X}^T\mathbf{X}\mathbf{b}=\mathbf{X}^T\mathbf{y}$ normal eşitlikleri ile elde edilir. Eğer tüm i, j=1,…,t $i\neq j$ için $\mathbf{x}_i$ sütunları $\mathbf{x}_j$ sütunlarına ortogonal ise diğer bir deyişle $\mathbf{X}_i\mathbf{X}_j=0$ ise

$$KT(\mathbf{b})=KT(\mathbf{b}_1)+KT(\mathbf{b}_2)+\cdots+KT(\mathbf{b}_t)$$

$$=\mathbf{b}_1^T\mathbf{X}_1^T\mathbf{y}+\mathbf{b}_2^T\mathbf{X}_2^T\mathbf{y}+\cdots\mathbf{b}_t^T\mathbf{X}_t^T\mathbf{y}$$

eşitliği elde edilir. Ortogonallik özelliğinin sağlandığı durumlarda diğer terimler modelde olsun olmasın $\boldsymbol{\beta}_i$ için EKK tahmini daima $\mathbf{b}_i$ olup bu parametrelere bağlı regresyon kareler toplamı da daima aynı değeri, $KT(\mathbf{b}_i)=\mathbf{b}_i^T\mathbf{X}_i^T\mathbf{y}$ verir. Bu sonuç,

$$KT(\mathbf{b}_i)=KT(\mathbf{b}_i/\mathbf{b}_j,i\neq j)$$

şeklinde ifade edilebilir. Burada $\mathbf{X}_i$'nin sütunlarının kendi içinde birbirine ortogonal olması gerekli değildir. Sadece $\mathbf{X}_i$'nin tüm sütunlarının $\mathbf{X}$ matrisinin tüm diğer sütunlarına ortogonal olmalıdır. Konuyu daha iyi açıklamak üzere t=2 olduğu durum ele alınsın. $\mathbf{X}$=[$\mathbf{X}_1$ $\mathbf{X}_2$] olacaktır. Ortogonallik özelliği gereği $\mathbf{X}_1^T\mathbf{X}_2 = \mathbf{X}_2^T\mathbf{X}_1 = 0$ olup model

$$\mathbf{y} = \mathbf{X\beta} + \boldsymbol{\varepsilon} = \mathbf{X}_1\boldsymbol{\beta}_2 + \mathbf{X}_2\boldsymbol{\beta}_2 + \boldsymbol{\varepsilon}$$

olarak yazılır. Normal eşitliklerden,

$$\begin{bmatrix} \mathbf{X}_1^T\mathbf{X}_1 & \mathbf{X}_1^T\mathbf{X}_2 \\ \mathbf{X}_2^T\mathbf{X}_1 & \mathbf{X}_2^T\mathbf{X}_2 \end{bmatrix}\begin{bmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{X}_1^T\mathbf{y} \\ \mathbf{X}_2^T\mathbf{y} \end{bmatrix}$$

ve sol taraftaki matris ile vektör çarpılarak,

$$\mathbf{X}_1^T\mathbf{X}_1\mathbf{b}_1 = \mathbf{X}_1^T\mathbf{y}$$

$$\mathbf{X}_2^T\mathbf{X}_2\mathbf{b}_2 = \mathbf{X}_2^T\mathbf{y}$$

parametreler için EKK çözümleri

$$\mathbf{b}_1 = (\mathbf{X}_1^T\mathbf{X}_1)^{-1}\mathbf{X}_1^T\mathbf{y}$$

$$\mathbf{b}_2 = (\mathbf{X}_2^T\mathbf{X}_2)^{-1}\mathbf{X}_2^T\mathbf{y}$$

bulunur. Yukarıdaki sonuçlardan görülebileceği gibi $\boldsymbol{\beta}_2$ modelde bulunsada bulunmasada $\boldsymbol{\beta}_1$ in EKK tahmini daima $\mathbf{b}_1$ sonucunu verecektir. Aynı sonuç $\mathbf{b}_2$ içinde geçerlidir. Böylece

$$KT(\mathbf{b}_1) = \mathbf{b}_1^T\mathbf{X}_1^T\mathbf{y}$$

$$KT(\mathbf{b}_2) = \mathbf{b}_2^T\mathbf{X}_2^T\mathbf{y}$$

yazılabilir. Buna uygun olarak

$$\begin{aligned} KT(\mathbf{b}_1, \mathbf{b}_2) &= \mathbf{b}^T\mathbf{X}^T\mathbf{y} \\ &= (\mathbf{b}_1^T\mathbf{b}_2^T)\begin{pmatrix} \mathbf{X}_1^T & \mathbf{y} \\ \mathbf{X}_2^T & \mathbf{y} \end{pmatrix} \\ &= (\mathbf{b}_1^T\mathbf{X}_1^T\mathbf{y} + \mathbf{b}_2^T\mathbf{X}_2^T\mathbf{y}) \\ &= KT(\mathbf{b}_1) + KT(\mathbf{b}_2) \end{aligned}$$

elde edilir. Bunun sonucunda,

$$KT(\mathbf{b}_1 / \mathbf{b}_2) = KT(\mathbf{b}_1, \mathbf{b}_2) - KT(\mathbf{b}_2) = KT(\mathbf{b}_1)$$

ve benzer olarak

$$KT(\mathbf{b}_2 / \mathbf{b}_1) = KT(\mathbf{b}_2)$$

bulunur. Görüldüğü gibi bu durum sadece $\mathbf{X}_1$ ve $\mathbf{X}_2$ deki ortogonalliğe bağlıdır.

4.13 GÜVEN ARALIKLARI VE MÜŞTEREK GÜVEN BÖLGELERİ

Parametrelerin güven aralığı tahminleri araştırmacılara nokta tahminlerinden daha fazla bilgi verir. Bununla birlikte birkaç parametre için oluşturulan tek değişkenli güven aralıkları parametre tahmincileri arasındaki korelasyonu dikkate almaz. Ayrıca bireysel güven katsayıları müşterek ifadelerdeki kapsamlı güven derecesini belirtmezler. Klasik güven aralıkları için, güven katsayısı $(1-\alpha)=0{,}95$ her bir güven ifadesine uygulanır. Tüm p adet parametreyi de içeren eşanlı p güven aralığının güven seviyesi çok daha düşüktür. Eğer tüm aralık tahminleri istatistiksel olarak bağımsız iseler (ki değildirler) müşterek güven katsayısı $(1-\alpha)^p$ olacaktır. Müşterek güven bölgeleri ile ilgili bu iki problem Kısım 4.12.4 de ele alınacaktır.

EKK parametreleri $\boldsymbol{\beta}$, $E(\mathbf{y})$ ve $\hat{\mathbf{y}}$ birer vektördür. Güven aralıklarını oluşturulmasında Kısım 4.9.2 de açıklanan karesel formların dağılım bilgilerinin kullanılması durumunda bu vektörlerdeki her bir eleman için ayrı ayrı aralık tahminleri elde edilebileceği gibi elemanların tümü ya da bir alt seti içinde güven bölgeleri elde edilebilir. İlk olarak regresyon parametreleri incelenecektir.

4.13.1 β İçin Aralık Tahmini

Hipotez testlerinde eşitlik (4.66) ile genel olarak tanımlanan bir karesel form $\boldsymbol{\beta}$ parametrelerinin testi için Kısım 4.11.2 de verilen Durum 5 için eşitlik (4.90) ile tanımlanmıştır. Parametrelerin hipotez testlerinde boş hipotez $\mathbf{K}^T\mathbf{b}=\mathbf{m}$ ile belirtilmişti. Bu eşitlikte test edilen değerler yerine $\boldsymbol{\beta}$ parametre vektörünün yazılması ile $\mathbf{K}^T\boldsymbol{\beta}=\mathbf{m}$ elde edilir. EKK parametrelerinin sapmasızlık özelliği ile $E(\mathbf{K}^T\mathbf{b})=\mathbf{K}^T\boldsymbol{\beta}$ olduğundan,

$$\mathbf{K}^T(\mathbf{b}-\boldsymbol{\beta}) \sim N\left[0;\mathbf{K}^T\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{K}\right] \tag{4.99}$$

tanımlanabilir. Bu bilgiler kullanılarak ve tanım matrisinin rankı (karesel formun serbestlik derecesi) r kabul edilerek bir F-değişkeni,

$$F=\frac{\left[\mathbf{K}^T(\mathbf{b}-\boldsymbol{\beta})\right]^T\left[\mathbf{K}^T\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{K}\right]^{-1}\left[\mathbf{K}^T(\mathbf{b}-\boldsymbol{\beta})\right]\big/ r}{s^2} \sim F_{(r;n-p)} \tag{4.100a}$$

ya da

$$\left[\mathbf{K}^T(\mathbf{b}-\boldsymbol{\beta})\right]^T\left[\mathbf{K}^T\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{K}\right]^{-1}\left[\mathbf{K}^T(\mathbf{b}-\boldsymbol{\beta})\right]=s^2 rF_{(r;n-p)} \tag{4.100b}$$

elde edilir. Burada $F_{(r;n-p)}$ serbestlik dereceleri r ve $(n-p)$ olan, üst sınır olasılığı α ile belirtilen F-dağılımı değeridir. Eşitliğin sol tarafı $\boldsymbol{\beta}$ ya göre bir karesel formu belirtir. Çünkü tüm diğer değerler bilinmektedir. Bu eşitsizliğin sınırları için karesel formun çözülmesi ile r boyutlu bir elipsoid oluşturulur. Bu elipsoid modeldeki tüm parametreler için %100(1-α)'lük bir müşterek güven bölgesini tanımlar. Sırasıyla, eksenlerin eğimi ve elipsoidin dış merkezi, parametre tahminleri arasındaki korelasyonun yönünü ve şiddetini gösterir.

Eşitlik (4.100) r adet parametre için ortak güven bölgesini tanımlar. Hipotez testlerinde olduğu gibi $\mathbf{K}^T$ katsayı matrisinin farklı düzenlemeleri ile parametrelerin farklı alt setlerinin ortak güven bölgeleri oluşturulabilir. Eşitlik (4.100) kullanılarak iki özel durum tanımlanabilir. Bunlardan ilki tüm p adet parametre için ortak güven bölgesinin elde edildiği durumdur. İkincisi ise tek bir parametre için aralık tahminin oluşturulmasıdır.

Durum 1. Tüm parametrelerin ortak güven bölgesi.

Hipotez testlerinde belirtildiği gibi bu durum için $\mathbf{K}^T=\mathbf{I}_p$ alınır ve $\rho(\mathbf{X}^T\mathbf{X})=p$ olduğundan eşitlik (4.100*b*),

$$(\mathbf{b}-\boldsymbol{\beta})^T\left(\mathbf{X}^T\mathbf{X}\right)^{-1}(\mathbf{b}-\boldsymbol{\beta})=s^2pF_{(p;n-p)} \tag{4.101a}$$

şeklinde yeniden düzenlenir.

Durum 2. Tek bir parametrenin güven aralığı.

Her hangi bir β_j parametresinin güven aralığını elde etmek için $\mathbf{K}^T$ bir satır vektörüdür, $\mathbf{K}^T=\begin{bmatrix}0 & \cdots & 0 & 1 & 0 & \cdots & 0\end{bmatrix}$. Burada 1 değeri β_j ye karşılık gelen pozisyondadır. Bunun sonucu olarak $\mathbf{K}^T(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{K}$ bir skalerdir ve serbestlik derecesi birdir. Eşitlik (4.100*b*),

$$\frac{\left(b_j-\beta_j\right)^2}{s^2c_{jj}}=F_{(1;n-p)} \tag{4.101b}$$

Burada c_{jj} değeri, $(\mathbf{X}^T\mathbf{X})^{-1}$ matrisinin j+1-inci elemanıdır. $F_{(1;n-p)}=t^2_{(n-p)}$ olduğundan, β_j için aralık tahmini,

$$b_j \pm t_{(\alpha/2;n-p)}s\sqrt{c_{jj}} \tag{4.101c}$$

olarak belirlenir. Üç parametreli bir model için β_1 ve β_2 ortak güven bölgelerinin oluşturulması ile ilgili bir örnek aşağıda verilmiştir. Bu model için s^2=0.75, n=5 ve

$$\mathbf{b}=\begin{bmatrix}b_0\\2.5\\-1.5\end{bmatrix}$$

olsun. Örnek için katsayı matrisi:

$$\mathbf{K}^T=\begin{bmatrix}0 & 1 & 0\\0 & 0 & 1\end{bmatrix}$$

Sonuç olarak,

$$\mathbf{K}^T(\mathbf{b}-\boldsymbol{\beta})=\begin{bmatrix}2.5-\beta_1\\-1.5-\beta_2\end{bmatrix} \text{ ve } \left[\mathbf{K}^T\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{K}\right]^{-1}=\begin{bmatrix}10 & 6\\6 & 4\end{bmatrix}$$

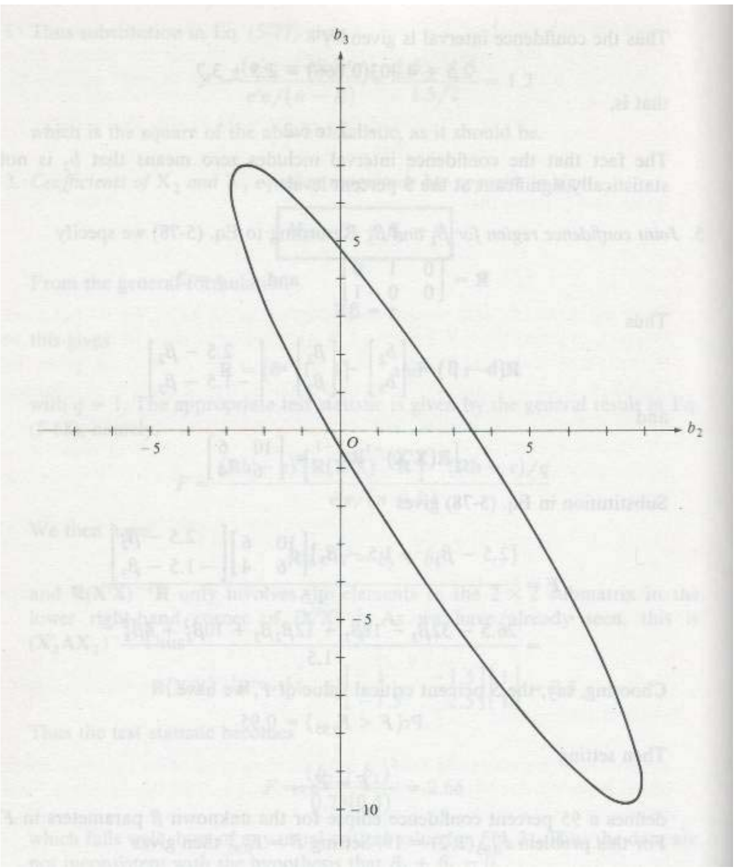
hesaplanır ve eşitlik (4.100*a*) da yerine konarak,

$$F=\frac{\begin{bmatrix}2.5-\beta_1 & -1.5-\beta_2\end{bmatrix}\begin{bmatrix}10 & 6\\6 & 4\end{bmatrix}\begin{bmatrix}2.5-\beta_1\\-1.5-\beta_2\end{bmatrix}\Big/2}{0.75}=\frac{26.5-32\beta_1-18\beta_2+12\beta_1\beta_2+10\beta_1^2+4\beta_2^2}{1.5}$$

α=0.05 için tablodan $F_{0.95}(2;2)=19$ ve sonuç olarak,

$$10\beta_1^2+12\beta_1\beta_2+4\beta_2^2-32\beta_2-18\beta_2-2=0$$

ortak güven bölgesi elde edilir. Bu ifade β_1 ve β_2 için %95 lik ortak güven elipsini tanımlar. Bu elips b_1=2.5, b_2=-1.5 noktasında merkezlenmiştir, bkz Şekil 4.2. Bu iki parametre tahmini arasında güçlü bir negatif korelasyon olduğu görülmektedir. Bu konu Kısım 4.12.5 de daha geniş olarak ele alınacaktır.



Şekil 4.2 johnstone sf 194

4.13.2 E(y) İçin Aralık Tahmini

Varsayılan modelin doğru olması durumunda $\hat{\mathbf{y}}$ vektörü $E(\mathbf{y})$ vektörünün sapmasız bir tahminleyicisidir. Kısım 4.2 de eşitlik (4.4) ve Kısım 4.4 de eşitlikler (4.11) ve (4.17) ile verilmiş olan ifadeler aralık tahmininde esneklik sağlamak amacıyla;

$$\mathbf{y} = \mathbf{K}^T\boldsymbol{\beta} + \boldsymbol{\varepsilon} \qquad (4.102a)$$

$$\hat{\mathbf{y}} = \mathbf{K}^T\mathbf{b} \qquad (4.102b)$$

$$E(\mathbf{y}) = \mathbf{K}^T\boldsymbol{\beta} \qquad (4.102c)$$

olarak yeniden tanımlanmıştır. Eşitlik (4.102b) ve (4.102c) kullanılarak,

$$E(\mathbf{y}) - \hat{\mathbf{y}} = -\mathbf{K}^T(\mathbf{b} - \boldsymbol{\beta}) \qquad (4.103)$$

Eşitlik (4.103) nin en sağındaki şans vektörü için karesel form eşitlik (4.100a) nın pay kısmında verilmiştir. Bu ifadede şans değişkeni $\mathbf{K}^T(\mathbf{b}\text{-}\boldsymbol{\beta})$ olup eşitlik (4.103) kullanılarak bu değişken yerine şans değişkeni olarak $E(\mathbf{y})$- $\hat{\mathbf{y}}$ alınarak,

$$[E(\mathbf{y}) - \hat{\mathbf{y}}]^T\left[\mathbf{K}^T\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{K}\right]^{-1}[E(\mathbf{y}) - \hat{\mathbf{y}}] = s^2 rF_{(r;n-p)} \qquad (4.104)$$

elde edilir. Eşitlik (4.104) r adet beklenen değer için ortak güven bölgesini tanımlar. Kısım 4.12.1 de olduğu gibi iki özel durum vardır.

Durum 1. Tüm n gözlem için beklenen değerlerin ortak güven bölgesi.

Katsayı matrisi $\mathbf{K}^T=\mathbf{X}$ olarak alındığında, $\rho(\mathbf{X})=p$ olduğundan, eşitlik (4.104) in yeni düzenlemesi,

$$[E(\mathbf{y})-\hat{\mathbf{y}}]^T\left[\mathbf{X}\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\right]^{-1}[E(\mathbf{y})-\hat{\mathbf{y}}]=s^2pF_{(p;n-p)} \qquad (4.105a)$$

olur.

Durum 2. Tek bir gözlem için beklenen değerin güven aralığı.

Her hangi bir $\mathbf{K}^T=\mathbf{x}_i^T=(1\ x_{i1}\ x_{i2}\ \ldots x_{ik})$ gözlemi için güven aralığı elde edilmek istendiğinde $\mathbf{x}_i^T\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{x}_i$ bir skaler olduğundan eşitlik (4.104),

$$\frac{[E(y_i)-\hat{y}_i]^2}{s^2\mathbf{x}_i^T\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{x}_i}=F_{(1;n-p)}$$

ve Kısım 4.12.1 e benzer işlemlerden sonra, $E(y_i)$ için güven aralığı,

$$\hat{y}_i\pm t_{(\alpha/2;n-p)}s\sqrt{\mathbf{x}_i^T\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{x}_i} \qquad (4.105b)$$

eşitliği ile tanımlanır. Burada,

$$\hat{y}_i=\mathbf{x}_i^T\mathbf{b}$$

olarak tanımlanır.

4.13.3 Bireysel Gözlemler y_i İçin Aralık Tahmini

Kısım 4.12.2 de verilen eşitlik (4.102*a*) ve (4.102*b*) kullanılarak,

$$\mathbf{y}-\hat{\mathbf{y}}=-\mathbf{K}^T(\mathbf{b}-\boldsymbol{\beta})+\boldsymbol{\varepsilon} \qquad (4.106)$$

bulunur. Elde edilen şans değişkeninin varyansı,

$$\begin{aligned}V\left[-\mathbf{K}^T(\mathbf{b}-\boldsymbol{\beta})+\boldsymbol{\varepsilon}\right]&=V\left(-\mathbf{K}^T\mathbf{b}\right)+V(\boldsymbol{\varepsilon})\\&=\mathbf{K}^T\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{K}\sigma^2+\sigma^2\end{aligned} \qquad (4.107)$$

ve sonuç olarak şans değişkenin dağılımı,

$$-\mathbf{K}^T(\mathbf{b}-\boldsymbol{\beta})+\boldsymbol{\varepsilon}\sim N\left[\mathbf{0};\mathbf{K}^T\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{K}\sigma^2+\sigma^2\right] \qquad (4.108)$$

bulunur. Eşitlik (4.108) den **y**-$\hat{\mathbf{y}}$ şans vektörünün karesel formu için tanım matrisi,

$$\mathbf{A}=\left[\mathbf{K}^T\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{K}+\mathbf{I}_k\right]^{-1}=\mathbf{V}^{-1}$$

bulunur. Sonuç olarak %100(1-α) güven bölgesini tanımlayan karesel form:

$$\frac{(\mathbf{y}-\hat{\mathbf{y}})^T\left[\mathbf{K}^T\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{K}+\mathbf{I}_k\right]^{-1}(\mathbf{y}-\hat{\mathbf{y}})/r}{s^2}=F_{(r;n-p)} \qquad (4.109)$$

İki özel durum aşağıda açıklanmıştır.

Durum 1. Tüm n gözlem için bireysel gözlemlerin ortak güven bölgesi.

Katsayı değerlerini veren bağımsız değişkenler kümesi boyutu $n\times p$ olan $\mathbf{K}^T=\mathbf{X}_0$ ise, $\rho(\mathbf{X})=p$ olduğundan, eşitlik (4.109):

$$(\mathbf{y}-\hat{\mathbf{y}})^T\left[\mathbf{K}^T\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{K}+\mathbf{I}_k\right]^{-1}(\mathbf{y}-\hat{\mathbf{y}})=rs^2F_{(r;n-p)} \tag{4.110a}$$

elde edilir.

Durum 2. Tek bir gözlem için bireysel gözlemlerin güven aralığı.

Her hangi bir $\mathbf{K}^T=\mathbf{x}_i^T=(1\,x_{i1}\;x_{i2}\ldots x_{ik})$ gözlemi için güven aralığı elde edilmek istendiğinde $\mathbf{x}_i^T\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{x}_i$ bir skaler olduğundan eşitlik (4.109),

$$\frac{[y_i-\hat{y}_i]^2}{s^2\left[1+\mathbf{x}_i^T\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{x}_i\right]}=F_{(1;n-p)}$$

ve Kısım 4.12.1 e benzer işlemlerden sonra, y_i için güven aralığı,

$$\hat{y}_i\pm t_{(n-p)}s\sqrt{\left[1+\mathbf{x}_i^T\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{x}_i\right]} \tag{4.110b}$$

eşitliği ile tanımlanır.

4.13.4 Ortak Güven Bölgeleri İçin Bonferroni ve Scheffe Yaklaşımları

Birkaç eşanlı aralık için ortak güven katsayısını önceden belirlenen bir (1-α) seviyesi civarında tutan iki yaklaşım mevcuttur. En eski ve en basit olanı Bonferroni metodu olarak bilinir. Bu metot ile eşitlik (4.101*c*) ve (4.105*b*) de verildiği şekilde bireysel güven aralıkları oluşturulur fakat bu eşitliklerde $\alpha^*=\alpha/r$ olarak alınır. Burada r eşanlı olarak oluşturulacak aralık sayısıdır. Bu nedenle $t_{(\alpha/2;n-p)}$ yerine $t_{(\alpha/2r;n-p)}$ kullanılır. Bu yaklaşım r adet eşanlı aralık için doğru müşterek güven katsayısının en az (1-α) değerine eşit olmasını sağlar. $\boldsymbol{\beta}$ parametre vektöründeki p adet parametre için Bonferroni eşanlı güven aralıkları:

$$b_j\pm t_{(\alpha/2p;n-p)}s_{b_j} \tag{4.111}$$

denklemi ile belirlenir. Bu metot özellikle önceden belirlenmiş r adet parametrenin aşanlı güven aralığının elde edilmesi için oldukça uygundur. Eğer r değeri küçükse tahminlenen aralıklar çok geniş olmayacaktır. Bununla birlikte r değeri büyüdükçe aralıklarda büyüyecektir. Sonuç olarak büyük r değerleri için Bonferroni yaklaşımı uygun değildir. Bir uç durum r değerinin sonsuz olmasıdır. Bu durumda $\alpha^*=\alpha/r$ sıfır değerini verecektir. Diğer bir deyişle birinci tip hata olasılığı$\alpha^*=0$ olacağı için Bonferroni aralıkları da sonsuz genişlikte olacaktır.

İkinci yaklaşım ise Scheffe (1953) metodu olarak bilinir. Bu metot p boyutlu parametre uzayının d boyutlu bir alt uzayındaki her hangi bir parametreler setinin tüm doğrusal kombinasyonları için eşanlı güven aralıklarının oluşturulmasını sağlar. $\boldsymbol{\beta}$ vektöründeki p adet parametre ve $E(\mathbf{y})$ için Scheffe'nin müşterek güven aralıkları (4.101*c*) ve (4.105*b*) eşitliklerdeki $t_{(\alpha/2;n-p)}$ ifadesi yerine $\left[pF_{(\alpha;p,n-p)}\right]^{1/2}$ yazılarak elde edilebilir. Eğer sadece d adet parametrenin doğrusal bağımsız bir alt seti ile ilgileniliyorsa $t_{(\alpha/2;n-p)}$ yerine $\left[dF_{(\alpha;d,n-p)}\right]^{1/2}$ yazılabilir. Yukarıda verilenlere uygun olarak,

$$b_j \pm \left[pF_{(\alpha;p,n-p)}\right]^{1/2} s\sqrt{c_{jj}} \qquad (4.112a)$$

$$\hat{y}_i \pm \left[pF_{(\alpha;p,n-p)}\right]^{1/2} s\sqrt{\mathbf{x}_i^T\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{x}_i} \qquad (4.112b)$$

eşitlikleri yazılabilir. Bonferroni aralıklarında olduğu gibi Scheffe aralıkları içinde müşterek güven katsayısı en az (1-α)'dır. Bu güven katsayısı β_j, $E(y_i)$ ve β_j'nin tüm ilgilenilen doğrusal kombinasyonlarına uygulanabilir. Bu durumda eşitlik (4.112*b*) ilgilenilen bölgedeki tüm bağımsız değişken değerleri için $E(y_i)$'nin Schefffe güven aralıklarının hesaplanmasıyla bütün regresyon yüzeyindeki bir güven bandının oluşturulması için kullanılabilir. Basit doğrusal regresyon durumu için güven bandı Working ve Hotelling (1929) tarafından geliştirilmiştir.

Bonferroni ve Scheffe metotlarının daha detaylı bilgisi için Miller (1981)'in çalısması tavsiye edilebilir. Parametreler setinin tüm doğrusal fonksiyonları için eşanlı güven aralıkları oluşturulduğundan Scheffe metodu ile elde edilen aralıklar Bonferroni aralıklarından daha geniş olabilecektir. Bu durum ile özellikle az sayıda eşanlı güven aralığının mevcut olduğu durumlarda karşılaşılır. Uygulamalarda daha dar aralıkları veren metot seçilebilecektir. Bonferroni ve Scheffe eşanlı güven aralıkları daima klasik tek değişkenli güven aralıklarından daha geniştir. Bunun nedeni de güven katsayının her bir aralığa uygulanmasıdır.

Bonferroni veya Scheffe metotları ile elde edilen eşanlı güven aralıklarını, tüm parametrelerin ortak güven bölgesi olarak düşünmek yanlıştır. Eşanlı aralıkların her biri bir parametrenin uygun değerlerinde tüm diğer parametrelerin ortalaması alınarak oluşturulmuş bir güven aralığıdır. Eşanlı aralıklar parametre tahminlerinin ortak dağılımlarını göz önünde bulundurmazlar. Bu nedenle onların kesişimlerini bir ortak güven bölgesi olarak yorumlamak hatalı olacaktır. Aralarındaki bu fark, müşterek güven bölgeleri tanımlandıktan sonra açıklanacaktır.

4.13.5 Ortak Güven Bölgelerinin Yorumu

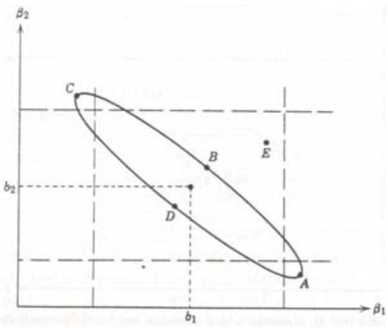
İki veya üçten daha fazla boyutlu bir elipsoit güven bölgesini yorumlamak zordur. $\boldsymbol{\beta}$ nın belirlenmiş değerlerinin, güven bölgesinin dışında ya da içinde olup olmadığı, bir bilgisayar yardımıyla, kontrol edilebilir. Böyle durumlarda eliptik bölgenin ana eksenlerinin uç noktalarının koordinatlarını bulmak bir çözüm olabilir. Şekil 4.3 de bu uç noktalar A, B, C ve D ile gösterilmişlerdir. Uç noktaları belirlemek için güven konturu bulunmalı ve kanonik yapıya indirgenmelidir. Bu amaçla kullanılabilecek yöntemler; ana bileşenler regresyonu ve latent kök regresyonudur.

Bununla birlikte bu çok boyutlu bölge ikişer veya en fazla üçer boyut bir arada olmak üzere incelenmelidir. Ortak güven bölgesi için görsel bir yaklaşımda iki parametrenin belirli değerleri için *p* boyutlu ortak güven bölgesinin değerlendirilmesidir. Belirlenmiş değerlerin her bir seti, çok boyutlu bölgenin iki boyutlu bir dilimi olan bir elips oluşturur. Bütün bölgenin resmini geliştirebilmek için bu iki boyutlu dilimler diğer parametrelerin seçilmiş birkaç değeri için grafikleri oluşturulabilir.

Tüm parametreler için *p* boyutlu müşterek güven bölgesinin bir alternatif kullanımı, parametreleri ikişerli olarak ele alıp diğer (*p*-2) parametre ihmal edilerek ortak güven bölgelerinin oluşturulmasıdır.

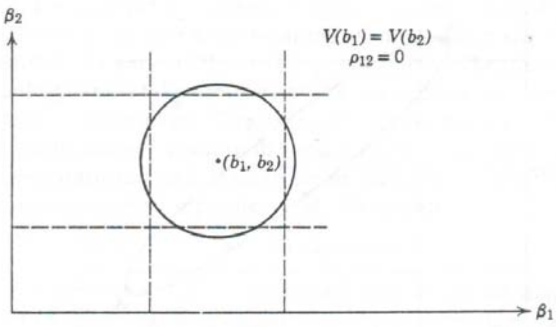
Kısım 4.12.1 de β_0 parametresi ihmal edilerek β_1 ve β_2 için bir ortak güven bölgesi elde edilmiştir. (1-α) güven katsayısı aynı anda ele alınan iki parametrenin ortak güven aralığına uygulanır. Bu yaklaşım, diğer parametrelerin değerlerini ihmal ederek β_j ve β_k'nın ortak dağılımını göz önünde bulundurur. Bu iki değişkenli ortak güven bölgesi, β_j ve β_k'nın diğer (p-2) parametre tahmini ile olan ortak dağılımını ihmal ettiğinden, tek değişkenli güven aralıklarının karşılaştığı kavramsal problemlerin aynısına sahiptir.

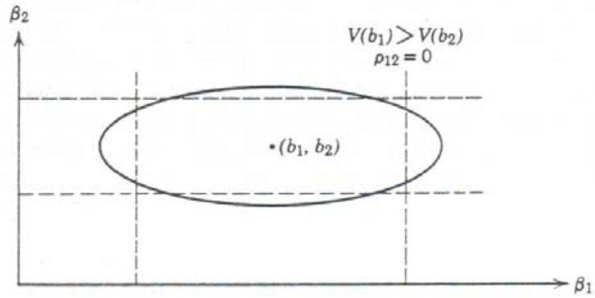
İki parametre dikkate alındığında ortaya çıkabilecek muhtemel durum Şekil 4.3'de gösterilmiştir. β_1 ve β_2 parametreleri için %95 güven katsayılı ortak güven bölgesi şekilden de görüldüğü gibi bir elips oluşturmaktadır ve parametrelerin ortak uygun değerlerini içermektedir. Bu güven aralıkları tek değişkenli olarak elde edildiklerinde oluşan dikdörtgenin içindeki E noktası β_1 ve β_2'nin müşterek değeri olarak görülebilir. Fakat görüldüğü şekilden gibi bu nokta müşterek güven bölgesinin dışındadır.



Şekil 4.3 Draper sf 95

Daha önce belirtildiği gibi ortakk güven bölgeleri parametre tahminleri arasındaki korelasyonu da dikkate almaktadır. İki parametre için bu durum ele alındığında $V(b_i)$, $V(b_k)$ ve $Cov(b_i, b_k)$ değerlerinin göreceli büyüklüğüne dikkat edilmelidir. b_i ve b_k farklı büyüklükteki varyanslara sahipse ve aradaki korelasyon küçük değil ise ve Şekil 4.3'deki durum ortaya çıkacaktır. Eğer $Cov(b_i, b_k)$ sıfıra yakınsa, bireysel güven aralıklarıyla tanımlanmış dikdörtgen bölge gerçek ortak güven bölgesine yaklaşacaktır. Bu durumda bölgenin uzunluğu $V(b_i)$ ve $V(b_k)$'nin göreceli büyüklüğüne bağımlı olacaktır. Birkaç örnek Şekil 4.4'de gösterilmiştir. Tüm parametreler için oluşturulan ortak güven bölgesi, iki değişkenli ortak güven ya da tek değişkenli güven aralığından çok daha kısıtlayıcıdır. , β_j ve β_k'nın mümkün kombinasyonları, diğer parametrelerin değerinin seçimine bağımlıdır. Tek değişkenli güven aralıklarının ve tüm parametreleri içermeyen ortak güven bölgelerinin aldatıcı olabileceği unutulmamalıdır.





Şekil 4.4 Draper sf 96

4.14 MATRİS GÖSTERİMİNDE EN YÜKSEK OLABİLİRLİK TAHMİNLEME YÖNTEMİ

Bölüm 2 de normal dağılmış hatalar için EKK ve EYOT yöntemleri açıklanmıştı. Doğrusal regresyon için hataların normal dağıldığı durumlarda EKK tahminleyicilerinin performansı, normal dağılışın gerçekleşmediği durumlara göre daha iyidir. Ayrıca normal dağılış durumunda EKK parametre tahminleyicileri EYOT a eşittir.

Genel doğrusal model eşitlik (4.4) ile

$$\mathbf{y} = \mathbf{X\beta} + \boldsymbol{\varepsilon}$$

tanımlanmıştır. Eğer hatalar için $\boldsymbol{\varepsilon}$~($\mathbf{0}$; $\sigma_2\mathbf{I}$) genel varsayımı yapılırsa ε_i için olasılık yoğunluk fonksiyonu,

$$f(\varepsilon_i) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{\varepsilon_i}{\sigma}\right)^2}$$

olup, olabilirlik fonksiyonu, $\varepsilon_1,\ldots,\varepsilon_n$, için

$$\prod f(\varepsilon_i) = \frac{1}{(2\pi)^{n/2}\sigma^n} e^{\left\{-\frac{1}{2\sigma^2}(\mathbf{y}-\mathbf{X\beta})^T(\mathbf{y}-\mathbf{X\beta})\right\}} \qquad (4.113a)$$

tanımlanır. Olabilirlik fonksiyonunun doğal logaritması ile çalışmak daha uygun olduğundan,

$$\log \prod f(\varepsilon_i) = -\frac{n}{2}\ln 2\pi - n\ln\sigma - \frac{1}{2\sigma^2}(\mathbf{y}-\mathbf{X\beta})^T(\mathbf{y}-\mathbf{X\beta}) \qquad (4.113b)$$

elde edilir ve log olabilirlik fonksiyonun parametrelere, $\boldsymbol{\beta}$ vektörüne, göre türevinin alınıp sıfıra eşitlenmesi,

$$\frac{\partial \log \prod f}{\partial \boldsymbol{\beta}} = \frac{2}{2\sigma^2}\left(\mathbf{X}^T\mathbf{y} - \mathbf{X}^T\mathbf{X\beta}\right) = 0$$

σ^2 ye göre türevinin sıfıra eşitlenmesi,

$$\frac{\partial \log \prod f}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{(\mathbf{y}-\mathbf{X\beta})^T(\mathbf{y}-\mathbf{X\beta})}{2\sigma^4} = 0$$

bu denklemlerin $\tilde{\mathbf{b}}$ ve $\tilde{s}^2$ için çözümleri,

$$\tilde{\mathbf{b}} = \left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{y} \qquad (4.114)$$

ve

$$\tilde{s}^2 = \frac{1}{n}\left(\mathbf{y}-\mathbf{X}\tilde{\mathbf{b}}\right)^T\left(\mathbf{y}-\mathbf{X}\tilde{\mathbf{b}}\right) = \frac{\tilde{\mathbf{e}}^T\tilde{\mathbf{e}}}{n} \tag{4.115}$$

bulunur. Eşitlik (4.114) ve (4.115) de verilen ~ sembolü EYOT yönteminden elde edilen tahminleri EKK yönteminden elde edilen tahminlerden ayırt etmek için kullanılır. Eşitlik (4.114) ve (4.9*a*) dan görüldüğü gibi EYOT ve EKK parametre tahminleyicileri denktir. Bu nedenle EYOT yönteminden elde edilen parametre tahminleri de sapmasızdır.

EYOT yönteminden elde edilen hata varyansının tahmini ise EKK tahminleyicisi s^2 den farklıdır ve sapmasızlığı araştırılmalıdır.

$$E\left[\tilde{s}^2\right] = \frac{1}{n}E\left[\left(\mathbf{y}-\mathbf{X}\tilde{\mathbf{b}}\right)^T\left(\mathbf{y}-\mathbf{X}\tilde{\mathbf{b}}\right)\right]$$

eşitlik (4.114) bu eşitlikte yerine konarak,

$$E\left[\tilde{s}^2\right] = \frac{1}{n}E\left[\mathbf{y}^T\left(\mathbf{I}-\mathbf{X}\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\right)^T\left(\mathbf{I}-\mathbf{X}\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\right)\mathbf{y}\right]$$

burada $\left(\mathbf{I}-\mathbf{X}\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\right) = \left(\mathbf{I}-\mathbf{H}\right)$ idempotent bir matris olduğundan,

$$E\left[\tilde{s}^2\right] = \frac{1}{n}E\left[\mathbf{y}^T\left(\mathbf{I}-\mathbf{X}\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\right)\mathbf{y}\right]$$

ve

$$E\left[\mathbf{y}^T\left(\mathbf{I}-\mathbf{X}\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\right)\mathbf{y}\right] = E\left[\boldsymbol{\varepsilon}^T\left(\mathbf{I}-\mathbf{X}\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\right)\boldsymbol{\varepsilon}\right]$$

Olduğundan eşitlik (4.48) kullanılarak,

$$E\left[\tilde{s}^2\right] = \frac{1}{n}\sigma^2 tr\left(\mathbf{I}-\mathbf{X}\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\right)$$

Elde edilir çünkü $\boldsymbol{\mu}=E(\boldsymbol{\varepsilon})=\mathbf{0}$ olduğundan ikinci terim sıfırdır. Eşitlik (4.36*a*) kullanılarak,

$$tr\left(\mathbf{I}-\mathbf{X}\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\right) = n-p$$

ve sonuç olarak,

$$E\left[\tilde{s}^2\right] = \frac{(n-p)}{n}\sigma^2 \tag{4.116}$$

bulunur. Eşitlik (4.116) EYOT yöntemi ile tahminlenen $\tilde{s}^2$ nin σ^2 nin sapmalı bir tahmini olduğunu göstermektedir.