YILDIZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

34715

BASİT LİNEER MODELLER VE GENELLEŞTİRİLMİŞ TERS MATRİSLERİN UYGULANMASI

İbrahim DEMİR

F.B.E. Matematik Anabilim Dalında
Hazırlanan

YÜKSEK LİSANS TEZİ

Tez Danışmanı: Y. Doç. Dr. Adnan MAZMANOĞLU

İSTANBUL, 1994

İÇİNDEKİLER

TEŞEKKÜR

Tez çalışmalarım sırasında yardımlarıyla ve eleştirileriyle beni daha iyi bir ödev hazırlamaya teşvik eden sayın hocam Yard. Doç. Dr. Adnan MAZMANOĞLU 'na ve çeşitli kaynak ve dökümlerin temininde bana rehberlik eden Üniversitemiz İstatistik Bölümü Başkanı Prof. Dr. Aziz BENER 'e kıymetli yardımlarından dolayı teşekkür ederim.

ÖZET

İstatistiğin büyük bir kısmı lineer modellerinoluşturulması ve regresyon analizinin incelenmesiyle uğraşmaktadır. Bizde bu çalışmamızda lineer modellerin genelleştirilmiş ters matrislerle çözümü ile ilgilendik.

1. Bölümde; lineer modelin tasarımı, en küçük kareler yöntemi ile söz konusu modelin lineer katsayıları olan ve nın elde edilmesi, varyans hesabı ve tahmini, linner model için güven aralığı tahmini ile bir uygulama bulunmaktadır.

2. Bölümde ise; genelleştirilmiş ters matrisin tanımı, özellikleri, elde edilmesi, lineer modellere adapte edilmesi yeralmaktadır.

Sonuç kısmında tezin konusu kısa ve öz olarak anlatılmıştır.

SUMMARY

Statistics mostly deals with structuring of the linear models and investigation of the regression analysis. In this research, we dealt with the linear models and the solution of this models by the generalized inverse matrix.

Chapter I contains project of the lineer model, the least-square process, obtaining of α and β coefficiants of the model, variance calculation and estimation and an application by ceonfidence interval for the linear model.

In chapter II, definition of generalized inverse matrix, it's properties obtaining and adaptation of its usage in the linear models are given.

In the last section, the subject of the thesis is written briefly.

GİRİŞ

İstatistikte, çoğu zaman x ve y gibi (tek veya çok boyutlu) iki büyüklük arasındaki ilişki ile ilgilenir. x ve y arasındaki ilişki değişik durumlar gösterir $y = \alpha x$ gibi bir ilişki varsa buna *fonksiyonel bağımsızlık*, *eğer aralarında bir ilişki yoksa bağımsızlık vardır* denir.

Bu iki uç durumdan, yani fonksiyonel bağımlılıktan farklı ara durumlar da vardır. Buna ise istatistiksel bağımlılık adı verilir.

İstatistiksel olarak bağımlı (x,y) çifti için iki farklı durum söz konusudur.

1-) x ve y nin ikiside rastgele değişken olup dağılım yasaları $F(x,y)$ ile verilir.

2-) (x,y) çiftindeki bir bileşene mesela x e bir değer verelim. O zaman y kendi ortalaması olan $E(y|x)$ etrafinda değişecektir. Buna göre ε ortalaması sıfır olan bir rastgele değişken olmak üzere

$$y = \varphi(x) + \varepsilon$$

yazılabilir. ε nın dağılımı x e bağlıdır ve y nin $\varphi(x)$ etrafindaki yaygınlığını gösterir. x i sabit tutarak tekrarlı denemeler yardımıyla $\varphi(x)$ e yaklaşabilir. Bazı hipotezler veya sağladığı kolaylıklardan dolayı $\varphi(x)$ fonksiyonu olarak $f(x;\beta_0,\beta_1,\beta_2)$ yi alırız.

Burada β_0,β_1 ve β_2 ler tahmin edilmesi gereken parametrelerdir. Yukarıdaki formüle regrasyon problemi denir.

Bir regrasyon problemi genel olarak

$$y = \varphi(x;\beta_0,\beta_1,\beta_2) + \varepsilon$$

ile ifade edilir. Burada

$$\varphi(x;\beta_0,\beta_1,\beta_2) = \beta_0 + \beta_1 x$$

olması halinde regrasyona lineer model adı verilir. Bizim burada inceleyeceğimiz model bu modeldir. O halde lineer modelimizin genel hali

$$y = \beta_0 + \beta_1 x + \varepsilon$$

olan denklemden elde edilen

$$(X'X)\hat{\beta} = X'Y$$

normal denkleminde, özel bir yapısı olan $X'X$ in rankı tam değildir. Dolayısıyla klasik anlamda ters matrisi yoktur. Bilinmeyen parametreler vektörü β için tek çözüm olmadığından birçok yazarlar "uygun koşullar","kaçınılmaz kısıtlamalar" denen yollara başvurarak çözümsüzlüğün üstesinden gelmişlerdir.

Son zamanlarda, rankı tam olmayan modeller yeni bir yaklaşımla ele alınmış ve bu yaklaşımda kullanılan yöntem; çok çözümlü lineer denklemlere uygulanılan "genelleştirilmiş ters matrislerle çözüm" yöntemidir. Dolayısıyla kısıtlamalara gerek kalmadan normal denklemler çözülebilmekte, kısıtlamaların bazı sakıncalarından kurtulunmakta, analiz yöntemleri daha genel bir bağlamda anlaşılabilmektedir.

Bu çalışmada güdülen amaç, son derece yararlı gördüğümüz bu yeni yaklaşımı dile getirmek, başka deyişle, nitel değişkenlerin söz konusu olduğu rank, tam olmayan lineer modellerin genelleştirilmiş ters matrislerle nasıl çözülebildiğini göstermektir.

BÖLÜM I
BASİT LİNEER MODEL

1. Lineer modele giriş

İstatistiğin büyük bir kısmı lineer model ya da regrasyon modellerin nasıl oluşturulacağını içerir. Bir lineer modelin oluşması ve lineer modele göre veri girişi analizi bu tezin konusu olacaktır.

Konuya bir tanım ile başlayalım. İlk olarak $(z_1,\ldots,z_N)$ bağımsız değişkenlerine fonksiyonel olarak bağlı olduğu düşünülen belli bir η cevap değişkeni. Yani $\eta=\alpha+\beta z$ yı göz önüne alalım. Örneğin belli bir üretim işleminde, bir ürünün verimi η, katalizin miktarı z_1, işlem süresince dikkate alınan (korunan) basınç z_2 ve işlemin başladığındaki sıcaklık z_3 'e bağlı olabilir .

Eğer

$$\eta=f(z_1,\ldots,z_N)=\sum_{j=1}^{k}\beta_j X_j(z_1,\ldots,z_N), \tag{1.1}$$

ise, denklem (1.1) in lineer modele uyduğunu söyleyebiliriz. Burada x_i ler sadece z_i nin fonksiyonudur. $\beta_1,\ldots,\beta_k$ nicelikleri (1.1) denklemine lineer olarak giren 1. dereceden genellikle bilinmeyen parametrelerdir.

(1.1) tipindeki modeller için bazı sonuçlar elde etmek için. İlk önce $s=1, k=2, \beta_1=\alpha, \beta_2=\beta$ ve $x_1(z)=1,\ x_2(z)=z\ , z_1=z_2$ alalım .Buna göre (1.1) modeli

$$\eta=\alpha+\beta z \tag{1.2}$$

dir. (1.2) deki denklemde $x=x_1(z)=z$ aldığımızda $\eta=\alpha+\beta x$ dır.

Diğer bir örnek, (1.2) genellemesi basit bir z değişkenindeki d. dereceden polinamal ilişkiyle gerçeklenir.

Yani;

$$\eta=\tau_0+\tau_1 z+\tau_2 z^2+\ldots+\tau_d z^d \tag{1.3}$$

dır. Burada $s=1$ ve $k=d+1$ olduğu için $\beta_j=\tau_{j-1}$, $x_j=x_j(z)=z^{j-1}$ ve $j=1,\ldots,d+1$ olarak yeniden yazdığımızda (1.3) denklemi yeni haliyle aşağıdaki gibi

$$\eta = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \ldots + \beta_{d+1} x_{d+1} \tag{1.3a}$$

olur. Dikkat edilirse β_j parametreleri (1.3) ve (1.3a) modeline lineer olarak katılırlar. polinomal olmayan şekildeki bir modelide, (1.1) denklemiyle verilen bir lineer model tipinde gösterebiliriz.

Şimdi,

$$\eta = \alpha + \beta \sin 2\pi z \tag{1.4}$$

modelini gözönüne alalım. α ve β, (1.4) denkleminde lineer olarak ifade edilebildiğinden dolayı bu model de bir lineer modeldir. Çünkü $s = 1$, $k = 2$, $x_1(z) = 1$, $x_2(z) = \sin 2\pi z$ ve $z_1 = z$, $\beta_1 = \alpha$, $\beta_2 = \beta$ olduğundan dolayı (1.4) denklemi, (1.1) modeli şeklindedir.

Yukarıdaki örneğin tersine nonlineer modelin bir örneğini, $\beta_2 > \beta_1$ ve $z > 0$ $\exp(a) = e^a$ ifadesini göz önünde bulundurarak

$$\eta = (\beta_2 - \beta_1)^{-1}\left[\exp(-\beta_1 z).\exp(-\beta_2 z)\right] \tag{1.5}$$

şeklinde ele alalım.

Daha açıkcası denklemde β_1 ve β_2 paremetreleri varolması halinde bile bu denklem, bir lineer model değildir.

Belirli nonlineer modelleri göz önüne alalım, bu parametrelerle nonlineer olan denklemler uygun dönüşümlerle lineer modellere dönüştürülebilir. Örneğin aşağıdaki cevap fonksiyonunu göz önünde bulunduralım.

$$\xi = \delta \exp(\gamma z) \tag{1.5a}$$

Açıkca (1.5a) de γ parametresi ile lineer değildir. Bununla birlikte e tabanına göre logaritmasının alınması ile

$$\eta = \beta_1 + \beta_2 x \tag{1.5b}$$

şeklinde olan

$$\ln \xi = \ln \delta + \gamma z \tag{1.5c}$$

denklemini elde ederiz. Burada $\beta_1 = \ln \delta$, $\beta_2 = \gamma$, $x_1(z) = 1$, $x_2(z) = z$ dir. Yani (1.5c) denklemi açıkca (1.1) formlu lineer bir modellidir. Belirli verileri baz almak süretiyle en küçük kareler analizini kullanarak (1.5c) yi elde etmek için bir yöntem takip edilebilir. Daha sonra (1.5c) belirli doğru kurallara uyarak izlenilen bir yolla (1.5a) çıkarımını elde edebiliriz. Bunun yanında (1.5) denkleminde olduğu gibi birçok nonlineer model dönüşümle lineer modele indirgenemeyebilir.

1.1 Basit Lineer Model ve Onun En Küçük Kareler Yöntemi İle Analizi

En küçük kareler analizini lineer modele uygulamak için (1.1) deki modelin analizini ele alacağız. Bu durum şimdilik matrisler veya n boyutlu geometrik argümanlar kullanılmaksızın yapılacaktır.

İşe somut bir örnekle başlayalım; sıcaklıktaki değişimlerin bir gazın hacmindeki değişmeleri nasıl etkilediği ile ilgilenelim. Bunu incelemek için $i = 1, \ldots, N$ kadar değerler alırken bağımsız değişken (sıcaklık buna x_i diyelim) önceden seçilmiş çeşitli durumlarda ölçülen bir silindirdeki gazın hacimlerinde (buna y_i diyelim) ki değişimleri ile ilgili deneyler yapılabilir. Farzedelimki deneysel hatalardan ayrı olarak verilen x e göre y nin matematik ümidini ifade edebilmek için y ve x arasındaki ilişkinin lineer olduğu kabul edilerek x e karşılık gelen y nın sözde regresyon fonksiyonu

$$E(y|x) = \eta_x = \alpha + \beta x \tag{1.6}$$

şeklinde olur.

Ayrıca, x ne olursa olsun gözlenen hacim , yani y ; bütün x değeri için

$$V(y|x) \equiv \sigma^2 \tag{1.7}$$

olarak ifade edilen aynı deneysel hatayı verir.

[(1.7) varsayımı literatürde homoscedestik varsayım olarak adlandırılır. Yani x ve y nin aynı anda değer almaları yada artmalardır]. Eğer (1.6) ve (1.7) denklemlerini ele alırsak iki yorum elde ederiz. Birincisi; Bağımsız değişkenin, verilen bir x değeri için, x değerindeki y den alınan gözlemler belirli bir D dağılımıyla uygun bir şekilde değişir. (Örneğin ,D, N normal dağılım olabilirdi) σ^2 ile x te üretilebilen y değerlerinin dağılım değişimlerini tanımlarız. $\alpha + \beta x$ doğrusu üzerinde bulunan x deki y nin beklenen değerini yani (1.6) denklemini;

$$y = D(\alpha + \beta x, \sigma^2) \tag{1.7a}$$

ifadesi ile gösterebiliriz. İkincisi; $x' \neq x$ bağımsız değişkenin diğer bir x' değeri için

(Örneğin ,D, N normal dağılım olabilirdi) σ^2 ile x te üretilebilen y değerlerinin dağılım değişimlerini tanımlarız. $\alpha+\beta x$ doğrusu üzerinde bulunan x deki y nin beklenen değerini yani (1.6) denklemini;

$$y = D(\alpha+\beta x, \sigma^2) \tag{1.7a}$$

ifadesi ile gösterebiliriz. İkincisi; $x' \neq x$ bağımsız değişkenin diğer bir x' değeri için

$$y = D(\alpha+\beta x', \sigma^2) \tag{1.7b}$$

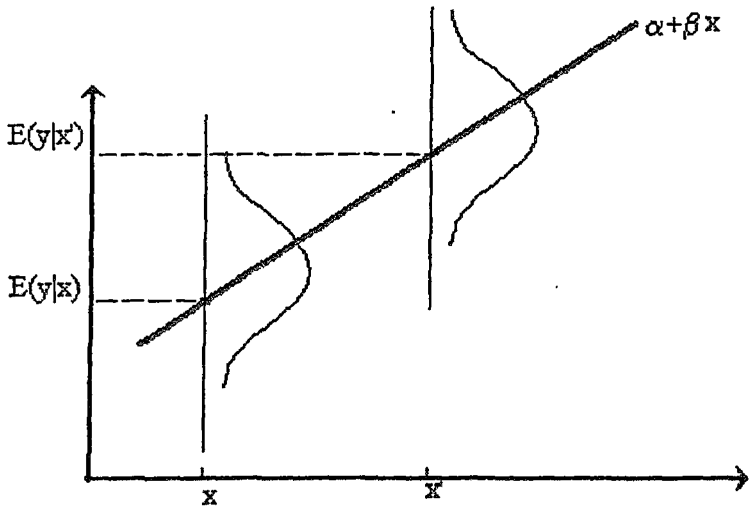
yi göz önünde bulundurarak özetleyebiliriz. Dağılımların ortalamaları ile ölçülen yerel değerlerden ayrı olarak x' nünde y ve x deki y nin dağılımları da ortalama bakımından farklılık gösterir. Bu durum şekil 1.1.1 de gösterilmiştir.

(1.7a) ve (1.7b) denklemlerini herhangi bir x için

$$y-(\alpha+\beta x)-\varepsilon=0 \tag{1.7c}$$

denklemini yazarak özetleyebiliriz. Burada rastgele hata ε, beklenen değeri sıfır ve varyansı σ^2 iken bütün x ler için aynı dağılıma sahiptir. x te y örneklendiğinde bazı zamanlar oluşum hatası (incurred error) olarak σ^2 dan bahsedilir. Yani anlatmak istediğimiz oluşum hatası ε dır. Burada ε sıfır ortalamalı ve σ^2 varyanslı dağılıma sahiptir. Daha ileri bir varsayımla x_i sıcaklığı kullanıldığında ortaya çıkan y_i gözlemi, x_j sıcaklığı kullanıldığında oluşan y_j den bağımsızdır. Yani x_i ile y_i ve x_j ile y_j birbirinden bağımsızdır.

(1.6) ve (1.7) denklemleri, istatistiksel bağımsızlık varsayımı deneyimiz için bir model oluşturur. İlişkinin x ile gerçekten lineer olduğunu veya x nin bilinen seçilen bölgelerde lineer olduğu bilindiğinden dolayı (1.6) denklemini ifade etmiştik. Şimdi adım olarak y ile x arasındaki fonksiyonel ilişki yapısının (1.6) ya bağlı olduğunu araştıralım istiyoruz. Önceden bahsedildiği gibi modelimizin alternatif bir tanımını kullanarak;



ŞEKİL 1.1.1

$$y_i = \alpha + \beta x_i + \varepsilon_i \tag{1.8}$$

denklemini yazabiliriz. Burada ε_i deneysel hataları gösterir. Ve

$$E(\varepsilon_i) = 0 \quad , V(\varepsilon_i) = \sigma^2$$

daki gibi dağılmıştır. ε_i diğer hatalardan bağımsızdır veya en azından karşılıklı olarak aralarında bir ilişkili yoktur. (x_i, y_i) $i = 1,\ldots,N$ verilerinin tipik cizimi şekil 1.1.2 deki gibidir.

Burada kullandığımız notasyonlar genel olarak kullanılan notasyonlardır. Birkaç x_i terimi aynı değere sahip olabilir. Fakat hiç değilse önceden enaz seçilmiş iki x in farklı olduğunu varsayıyoruz. Gerçekte deneye ait x lerin en iyi nasıl seçileceğinin sorusu burada aklımıza gelebilir. Bu noktada diğerleri arasında bir β eğimini tahmin etmek istediğimizden dolayı sadece x in farklı değerlerine karşı alınan enaz iki gözleme ihtiyaç duyduğumuzu söyleblirız.

Şimdi η fonksiyonel ilişkisindeki, bilinmeyen α ve β parametrelerinin ve deneysel hata σ^2 nin tahminlerini verecek bir metod göreceğiz. Bazı durumlarda, bilinmeyen α ve β yı $\alpha^{(g)}$ ve $\beta^{(g)}$ vasıtasıyla ifade edebiliriz. x 'e göre lineer kabul edilen bilinmeyen regresyon fonksiyonunu

$$\eta^{(g)} = \alpha^{(g)} + \beta^{(g)} x \tag{1.9}$$

olarak gösterelim. Burada $(x_i, \eta_i^{(g)})$ koordinatlarına sahip önceden seçilmiş x_i değerleri ile ilişkili doğru üzerindeki noktaları geçtiğimize dikkat ediniz. Burada η_i yi η_{x_i} yerine ve $\eta_i^{(g)}$ yi $\eta_{x_i}^{(g)}$ yerine yazabiliriz. Böylece (1.9) denklemi;

$$\eta_i^{(g)} = \alpha^{(g)} + \beta^{(g)} x_i \, , \qquad i = 1, \ldots, N \tag{1.9a}$$

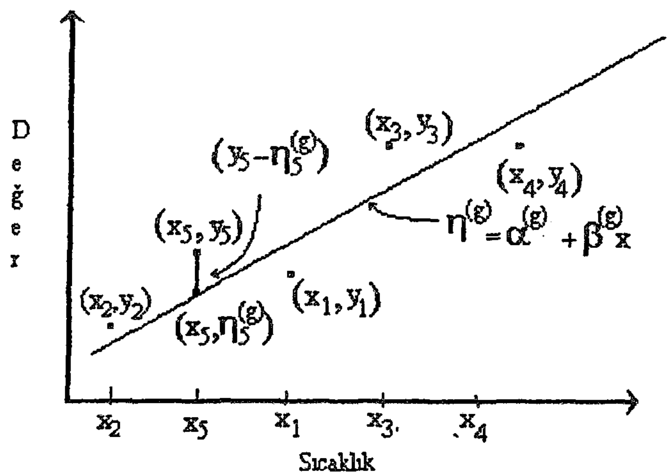
şeklinde olur. (1.9) tahmininin ne kadar iyi olduğunu ölçmenin bir yolu, "gözlenen (x_i, y_i) verilerinden (1.9) doğrusunun uzaklığı ne kadardır ? " sorusunu sormaktır. (1.10) ile ifade edilen bir kriter fonksiyonu gözlenen (x_i, y_i) noktalarının, tahmin edilen doğru üzerindeki $(x_i, \eta_i^{(g)})$ noktalarına olan düşey uzaklıklarının karelerinin toplamını ölçer. Yani;

$$\begin{aligned} Q_g &= \sum_{i=1}^{N} \left(y_i - \eta_i^{(g)} \right)^2 \\ &= \sum_{i=1}^{N} \left(y_i - \alpha^{(g)} - \beta^{(g)} x_i \right)^2 \end{aligned} \tag{1.10}$$

dir. Bu denklem bize en küçük kareler yönteminin denklemini verir.(Şekil 1.1.2 ye bakınız.) Şimdi; Eğer tahmin edilen iki doğrunun (yani $(\alpha^{(g)}, \beta^{(g)})$ ile (α, β)) iyiliğini karşılaştırmak istersek, -örneğin $(\alpha^{(g')}, \beta^{(g')})$ ifadesini ele alalım- aşağıdaki ifadelere göre sezgisel olarak $(\alpha^{(g)}, \beta^{(g)})$ ile $(\alpha^{(g')}, \beta^{(g')})$ arasında seçim yapılabiliriz.

$$\begin{aligned} &Q_g > Q_{g'} \text{ ise } \quad \left(\alpha^{(g')}, \beta^{(g')}\right) \text{ seçebiliriz} \\ &Q_g = Q_{g'} \text{ ise } \quad \left(\alpha^{(g')}, \beta^{(g')}\right) \text{ veya } \left(\alpha^{(g)}, \beta^{(g)}\right) \text{ yi seçebiliriz} \\ &Q_g < Q_{g'} \text{ ise } \quad \left(\alpha^{(g)}, \beta^{(g)}\right) \text{ yi seçebiliriz.} \end{aligned} \tag{1.11}$$

Q ile ölçüldüğü gibi veriye en yakın tahmin edilen doğruya yükselme verecek (α, β) tahmin çiftlerini seçebiliriz



ŞEKİL 1.1.2

Örneğin (α, β) tahminlerinin tüm münkün seçenekler üzerindeki veriye en yakın

$$\hat{\eta} = \hat{\alpha} + \hat{\beta} x \tag{1.12}$$

tahmin edilen (1.12) doğrusuna yükseltme verecekdir. Tekrarlanan haller için en küçük kareler yöntemi olarak adlandırılan en küçük kareler yöntemini kullanırız. (α, β) nın tahminleri olarak $(\hat{\alpha}, \hat{\beta})$ yı seçeriz . Burada

$$\min_{(\alpha,\beta)} Q = \min_{(\alpha,\beta)} \sum_{i=1}^{N} (y_i - \alpha - \beta x_i)^2 = \sum_{i=1}^{N} \left(y_i - \hat{\alpha} - \hat{\beta} x_i\right)^2 \tag{1.13}$$

dır.

$$Q = \sum_{i=1}^{N} \left(y_i - \hat{\alpha} - \hat{\beta} x_i \right)^2, \quad Q(\alpha, \beta) \tag{1.14}$$

(1.14) ü minimize eden $(\hat{\alpha}, \hat{\beta})$ değerlerini bulmanın bir çok yolu vardır.

Burada vereceğimiz ilk yöntemde hesaplama işlemleri kullanılır. (1.1) ifadesinde açıklamalarını verdiğimiz 2. yöntem cebirseldir. (1.14) de verilen $Q(\alpha, \beta)$ yı minimize eden değerlerini bulmak için α ve β ya göre kısmi türev alırız. Bu türevleri sıfıra eşitler ve yukarıda gösterildiği gibi $(\hat{\alpha}, \hat{\beta})$ ile çözümü göstererek (α, β) için çözüm elde ederiz. Bu işlemler aşağıdaki gibi yapılır.

$$\left.\frac{\partial Q(\alpha, \beta)}{\partial \alpha}\right|_{(\hat{\alpha}, \hat{\beta})} = 0 \tag{1.15}$$

$$\left.\frac{\partial Q(\alpha,\beta)}{\partial \beta}\right|_{(\hat{\alpha},\hat{\beta})} = 0$$

Bu sıra ile iki eşitlik üretir.

$$-2\sum_{i=1}^{N}\left(y_i - \hat{\alpha} - \hat{\beta}x_i\right) = 0 \qquad (1.15a)$$

$$-2\sum_{i=1}^{N}\left(y_i - \hat{\alpha} - \hat{\beta}x_i\right)x_i = 0$$

Bu denklemlerin her birini -2 ile bölerek, eşitliklerin sağ tarafını yalnız "y" terimleri içermesi için yeniden düzenleyerek ve toplamalar yaparak

$$(N)\hat{\alpha} + \left(\sum_{i=1}^{N} x_i\right)\hat{\beta} = \sum_{i=1}^{N} y_i \qquad (1.15b)$$

$$\left(\sum_{i=1}^{N} x_i\right)\hat{\alpha} + \left(\sum_{i=1}^{N} x_i^2\right)\hat{\beta} = \sum_{i=1}^{N} x_i y_i$$

elde ederiz.(1.15b) eşitlikleri α, β yı tahmin etmek için genellikle normal denklemler olarak adlandırılır. Kat sayılar determinantı

$$\begin{vmatrix} N & \sum_{i=1}^{N} x_i \\ \sum_{i=1}^{N} x_i & \sum_{i=1}^{N} x_i^2 \end{vmatrix} = \Delta(x_i,\ldots,x_N) \qquad (1.16)$$

sıfır olmadıkça (1.15b) normal denklemleri $(\hat{\alpha},\hat{\beta})$ ye göre bir tek çözüme sahip olacaktır. Şimdi bu determinant

$$\Delta(x_i,\ldots,x_N) = N\left[\sum x_i^2 - \frac{\left(\sum x_i\right)^2}{N}\right] \qquad (1.16a)$$

$$= N\sum_{i}^{N}(x_i - \bar{x})^2$$

sahip olur ve elbetteki sıfıra eşit değildir. Tüm x_i değerlerin hepsi birbirine eşittir. Bu da (1.8a) dan sonraki yaklaşımlara aykırıdır. (1.14) denklemini minimize etmenin bir sonucu olarak (1.15b) denklemleri bulunduğu için (1.15b) normal denklemleri, en küçük kareler yöntemi olarak tanımlanır. $\dot{x}$ ile $(x_i - \bar{x})$ kümesini ve

$$S(\dot{x},\dot{y}) = \sum_{1}^{N}(x_i - \bar{x})(y_i - \bar{y}) = \sum_{i=1}^{N} x_i y_i - N\bar{x}\bar{y} \qquad (1.17)$$

$$S(\dot{x}^2) = \sum_{1}^{N}(x_i - \bar{x})(x_i - \bar{x}) = \sum_{1}^{N} x_i^2 - N\bar{x}^2$$

göstererek $\hat{\alpha}$ ve $\hat{\beta}$ çözümleri,

$$\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x} \qquad (1.17a)$$

$$\hat{\beta} = \frac{\sum_1^N x_i y_i - N\bar{x}\bar{y}}{\sum_1^N x_i^2 - N\bar{x}^2} = \frac{\sum_1^N (x_i - \bar{x})(y_i - \bar{y})}{\sum_1^N (x_i - \bar{x})^2} = \frac{S(\dot{x},\dot{y})}{S(\dot{x}^2)}$$

şeklinde elde edebiliriz.

Daha önce gösterildiği gibi $(\hat{\alpha},\hat{\beta})$ kullanıldığında en küçük kareler yönteminin çözümü aşağıdaki gibidir. Q yu minimize eden $\alpha - \beta$ düzlemindeki noktaları araştırmada kullanılan (1.14) denklemi ile verilen $Q(\alpha,\beta)$ kareler toplamının diferansielinin bir sonucu olarak (1.15b) normal denklemlerini elde ederiz. Q yu minimize etmenin en basit yolu (1.15a) veya (1.15b) denklemlerini kullanmaktır. Buradan (α,β) nın tüm diğer seçenekleri üzerinde $Q(\alpha,\beta)$ değerlendirildiği zaman $Q(\hat{\alpha},\hat{\beta})$ en küçük olabilir. Bunu gösterebiliriz.(Ayrıca ek 1.1 de görülür.)

farzedelimki Q yu ;

$$Q = \sum_{i=1}^{N}\left[\left(y_i - \hat{\alpha} - \hat{\beta}x_i\right) + (\hat{\alpha} - \alpha) + \left(\hat{\beta} - \beta\right)x_i\right]^2 \qquad (1.17b)$$

olarak yazalım. Burada $(\hat{\alpha},\hat{\beta})$, (1.17) denklem sistemi ile bulunmuştu. Yani $(\hat{\alpha},\hat{\beta})$, (1.15a) denklem sistemini sağlar.Şimdi (1.17b) ifadesini genişleterek;

$$Q = \sum_{i=1}^{N}\left(y_i - \hat{\alpha} - \hat{\beta}x_i\right)^2 + \sum_{i=1}^{N}\left[(\hat{\alpha} - \alpha) + \left(\hat{\beta} - \beta\right)x_i\right]^2 \qquad (1.17c)$$
$$+2\sum_{i=1}^{N}\left(y_i - \hat{\alpha} - \hat{\beta}x_i\right)\left[(\hat{\alpha} - \alpha) + \left(\hat{\beta} - \beta\right)x_i\right]$$

ifadesini elde ederiz. (1.17c) denkleminin üçüncü teriminde verilen çapraz çarpımın sıfır olduğunu görürüz. Çünkü bu ifade (1.17d) ye

$$2\left[(\hat{\alpha} - \alpha)\sum_{i=1}^{N}\left(y_i - \hat{\alpha} - \hat{\beta}x_i\right) + \left(\hat{\beta} - \beta\right)\sum_{i=1}^{N}\left(y_i - \hat{\alpha} - \hat{\beta}x_i\right)x_i\right] \qquad (1.17d)$$

karşılıktır ve (1.15a) denklemini kullanarak (1.17d) denklemini sıfır bulabiliriz. Buradan;

$$Q=\sum_{i=1}^{N}\left(y_i-\hat{\alpha}-\hat{\beta}x_i\right)^2+\sum_{i=1}^{N}\left[(\hat{\alpha}-\alpha)+\left(\hat{\beta}-\beta\right)x_i\right]^2 \qquad (1.17e)$$

elde edilir. (α,β) nın üzerinde minimize edilen Q, minimize edilen (1.17e) denkleminin ikinci terimiyle minimize edileceği ve bu (α,β) ile $(\hat{\alpha},\hat{\beta})$ değiştirildiği zaman bu işlem daha kolay bir şekilde görülür.

Yukarıda gösterilen Q nun minumum değerini şöyle ifade edebiliriz.

$$Q\left(\hat{\alpha},\hat{\beta}\right)=\sum_{i=1}^{N}\left(y_i-\hat{\alpha}-\hat{\beta}x_i\right)^2=\min_{\alpha,\beta}Q(\alpha,\beta) \qquad (1.17f)$$

$(\hat{\alpha},\hat{\beta})$, (α,β) nın verilen tahminleridir. (1.6) denkleminin yani $\eta=\alpha+\beta x$ nın tahminlerine sahip oluruz. Bu ise (1.17) denklemi ile bulunan $(\hat{\alpha},\hat{\beta})$ nın en küçük kareler doğrusudur. Yani;

$$\hat{\eta}=\hat{\alpha}+\hat{\beta}x=\hat{\eta}_x \qquad (1.18)$$

dır. Eğer $x=\bar{x}$ ise o zaman

$$\hat{\eta}_{\bar{x}}=\hat{\alpha}+\hat{\beta}\bar{x}=\left(\bar{y}-\hat{\beta}\bar{x}\right)+\hat{\beta}\bar{x}=\bar{y} \qquad (1.18a)$$

denklemini elde ederiz. Burada en küçük kareler doğrusu olarak adlandırılan bu denklem $(\bar{x},\bar{y})$ verilerinin ağırlık merkezinden geçer. (1.18) de ulaştığımız neticeleri bütün işlemlere uygulayarak en küçük karelerle uygun bir doğru elde ederiz. Burada x de y nin regresyona tabi tutulması söz konusudur. Yalnız x nin önceden seçilmiş değerlerindeki en küçük kareler doğrusunun $(x_i,\hat{\eta}_i)$ noktaları olduğunu biliyoruz ve gerçekten (1.14),(1.15) ifadelerini ve diğerlerini göz önüne alarak;

$$Q\left(\hat{\alpha},\hat{\beta}\right)=\sum_{i}^{N}\left(y_i-\hat{\alpha}-\hat{\beta}x_i\right)^2=\sum_{i}^{N}\left(y_i-\eta_{x_i}\right)^2 \qquad (1.19)$$

$$=\min_{\alpha,\beta}Q(\alpha,\beta)$$

dır. Genellikle $Q(\hat{\alpha},\hat{\beta})$ ifadesi SS_e olarak gösterilir. - (α,β) nin tahminlerinin bütün mümkün seçilmişlerinin üzerindeki $Q(\alpha,\beta)$ nin minumum değerleri- ve SS_e , e_i sapmalarının karelerinin toplamları olarak adlandırılır. Burada e_i

$$e_i = y_i - \hat{\eta}_i = y_i - \left(\hat{\alpha} + \hat{\beta} x_i\right) \tag{1.20}$$

dir. e_i çokluğu, genellikle y_i nın kalanı olarak ifade edilir ve y_i nin görülen değerlerinden Yani $\eta_i = \alpha + \beta x_i$ nin ve $\hat{\eta}_i - \hat{\eta}$ tahmininin farkı olarak adlandırılır. (Bu şekil 1.1.3 de görülebilir.)

Kalanların toplamını genellikle sıfırdır. Yani;

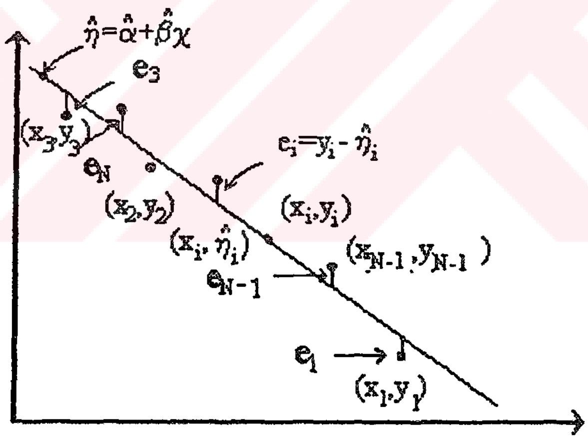
$$\sum_{i}^{N} e_i = \sum_{i=1}^{N}\left(y_i - \hat{\alpha} - \hat{\beta} x_i\right)$$

$$= \sum_{i=1}^{N}\left[(y_i - \bar{y}) - \hat{\beta}(x_i - \bar{x})\right] \tag{1.20a}$$

$$= \sum_{i=1}^{N}(y_i - \bar{y}) - \hat{\beta}\sum_{i=1}^{N}(x_i - \bar{x})$$

$$= 0$$

dır.



ŞEKİL 1.1.3

Bir en küçük kareler analizi yapıldığı zaman genelde bu özelliği dikkate almaya gerek yoktur. Fakat gerçek sabit terimle lineer bir model uydurulursa bu özellik göz önünde bulundurulmalıdır. Dikkat edilirse (1.20a) denklemi aşağıdaki gibi ifade edilebilir.

$$\sum\left(y_i - \hat{\eta}_i\right) = 0 \quad \text{veya} \quad \sum y_i = \sum \hat{\eta}_i \tag{1.20b}$$

Bunu ise;

$$Q(\hat{\alpha},\hat{\beta})=\sum_i^N (y_i-\hat{\eta}_{x_i})^2=\sum_i^N e_i^2 = SS_e \tag{1.20c}$$

şeklinde ifade edebiliriz. Kısım 1.2 de $E(\hat{\alpha})=\alpha$ ve $E(\hat{\beta})=\beta$ olduğunu göstereceğiz. Böylelikle e_i kalanların beklenen değerine sahip oluruz. ($E(e_i|\, x_1,\ldots,x_N)$yi $E(e_i)$ ile gösterebiliriz.)

$$\begin{aligned}E(e_i) &= E(y_i-\hat{\alpha}-\hat{\beta}x_i|x_i,\ldots,x_N)\\ &= E(y_i|x_i)-E(\hat{\alpha})-x_iE(\hat{\beta}) \qquad (1.20d)\\ &= \alpha+\beta x_i-\alpha-\beta x_i\\ &= 0\end{aligned}$$

dir. Bunun anlamı ise e_i terimleri tek başına hata hakkında bilgi verir demektir. Kısım 1.3 de görülecektir ki $\sum_i^N e_i^2 = SS_e$ yı, (1.6) ve (1.7) modelinin hata bileşeni σ^2 yi tahmin etmek için bir baz olarak kullanabiliriz.

Şimdi $\hat{\alpha}$, $\hat{\beta}$ ve SS_e , $\hat{\eta}_i$ ve e_i nicelikleri istatistiksel özelliklere sahip olduğunu bu kısımda araştıracağız. En küçük kareler analizinin açıklamalarına ise daha sonraki kısımlarda devam edeceğiz.

1.2. En Küçük Kareler Analizinin İstatistiksel Özellikleri

Bu kısımda buraya kadar hiç ihtiyaç duyulmayan kabullenmelerimizi hatırlayarak işe başlıyalım. Birinci kabullenme: Dağılımı D olan x'in önceden seçilmiş -sonlu sayıda x_i nin birinci ve ikinci momentleri vardır- x_i değerlerinde rastgele y değişkenlerini gözlüyoruz, .

$$E(y|x)=\alpha+\beta x \tag{1.21a}$$

$$V(y|x)=\sigma^2 \tag{1.21b}$$

ve ayrıca x_i de gözlenen N tane y_i gözlemleri istatistiksel olarak bağımsızdırlar. Şimdi (α,β) nın $(\hat{\alpha},\hat{\beta})$ tahminlerinin beklenen değerini ve varyansını araştıralım.

(1.17) denklemi yardımıyla $S(\dot{x},\dot{y})$ ifadesini basit cebirsel kurallarla

$$\sum_i^N (x_i-\bar{x})(y_i-\bar{y}) = \sum x_i y_i - N\bar{x}\bar{y} \tag{1.22}$$

$$= \sum (x_i-\bar{x})y_i$$

haline getirdikten sonra işlemlerimize başlamak için $\hat{\beta}$ 'yı

$$\hat{\beta} = \frac{\sum_i^N (x_i-\bar{x})y_i}{S(\dot{x}^2)} \tag{1.22a}$$

olarak yazabiliriz. Böylece (1.22)'den

$$\hat{\beta} = \sum_1^N c_i y_i \tag{1.22b}$$

elde ederiz. burada $c_i = \dfrac{(x_i-\bar{x})}{S(\dot{x}^2)}$ olduğunu görürüz.

Lineer kombinasyonun katsayıları olan c_i lerin $(i=1,\ldots,N)$ ağırlıkları sadece x_i nin önceden seçilmiş değerlerine bağlıdır ve bundan dolayı da sabittirler. Bunun anlamı $\hat{\beta}$ nın y_i lerin lineer kombinasyonu olduğudur. Bunlar şu özelliklere sahiptir.

i) $\quad \sum_1^N c_i = 0$

ii) $\quad \sum_1^N c_i^2 = \dfrac{1}{S(\dot{x}^2)} \qquad (1.22c)$

iii) $\quad \sum_1^N c_i x_i = 1$

(1.22b) denkleminden $E(\hat{\beta})$ yı yani beklenen değeri elde ederiz.

$$E(\hat{\beta}|x_1,\ldots,x_N) = \sum_1^N c_i E(y_i|x_1,\ldots,x_N)$$

$$= \sum_i^N c_i E(y_i|x_i) \tag{1.22d}$$

$$= \sum_i^N c_i \eta_i$$

Burada $\eta_i = E(y_i|x_i) = \alpha + \beta x_i$ dir. Böylece (1.22c) den

$$E(\hat{\beta}) = \sum_{1}^{N} c_i(\alpha + \beta x_i)$$

$$= \alpha \sum_{1}^{N} c_i + \beta \sum_{1}^{N} c_i x_i \qquad (1.22e)$$

$$= \beta$$

dır. Yani β nın en küçük kareler yöntemini $\hat{\beta}$, β için ön yargısızdır. İşlemlere devam ederek (1.22) ifadesinin her iki tarafının varyansını -$\hat{\beta}$ nın varyansı- alırız . Bunu işlemsel olarak gösterirsek;

$$V(\hat{\beta}|x_1,\ldots,x_N) = \sum_{i}^{N} c_i^2 V(y_i|x_1,\ldots,x_N) \qquad (1.23)$$

dır.

Çünkü y_i ler bağımsızdırlar, basit işlemlerde sonra

$$V(\hat{\beta}) = \sum_{1}^{N} c_i^2 V(y_i|x_i)$$

$$= \sigma^2 \sum_{1}^{N} c_i^2 \qquad (1.23a)$$

$$= \frac{\sigma^2}{S(\dot{x}^2)}$$

dır. [Şayet x_i ler $S(\dot{x}^2)$ nı maksimum yapacak şekilde mümkün olduğu kadar geniş aralıkta seçilirse $\hat{\beta}$ nın varyansı minumum olacaktır. Bu istenilen sonuçtur. Çünkü $\hat{\beta}$, β' nın bir tahminidir -bir eğimidir- ve bu eğimin iyi bir tahmini x lerin geniş bir aralığa dayandırılması ile ortaya çıkarılır.]

(1.22b) ve (1.17) ifadelerinin birinci satırını kullanarak $\hat{\alpha}$ yı y_i lerin lineer bir kombinasyonu olarak yazılabileceğini gösterebiliriz. Yani;

$$\hat{\alpha} = \sum_{i=1}^{N} \left(\tfrac{1}{N} - \bar{x} c_i\right) y_i \qquad (1.24)$$

dir. Bunu ispatlayabildiğimizden

$$E(\hat{\alpha}) = \alpha\,, \qquad V(\hat{\alpha}) = \sigma^2\left[\tfrac{1}{N} + \frac{\bar{x}^2}{S(\dot{x}^2)}\right] \tag{1.24a}$$

denklemlerini elde ederiz.

Ayrıca, $\eta_x = \alpha + \beta x$ -y nin x teki beklenen değeri- en az kareler yönteminin lineer bir fonksiyondur. Bu aşağıdaki gibi gösterilebilir.

$$\hat{\eta}_x = \hat{\alpha} + \hat{\beta} x = \sum_i^N \left[\tfrac{1}{N} + c_i(x - \bar{x})\right] y_i \tag{1.25}$$

Yukarıdaki sonuçlardan;

$$E(\hat{\eta}_x) = \eta_x\,, \qquad V(\hat{\eta}_x) = \sigma^2\left[\tfrac{1}{N} + \frac{(x-\bar{x})^2}{S(\dot{x}^2)}\right] \tag{1.25a}$$

denlemlerini sağlayabiliriz.

[(1.25a) denklemi; η_x tahmini için, $\bar{x}$ de veya x_i lerin merkezinde en küçük varyansı verdiğini gösterir.] Ayrıca (1.24) ve (1.24a), (1.25) ve (1.25a) denklemlerinin özel halini verir. Çünkü x=0 da uydurulan doğrunun kesim noktası $\hat{\alpha}$ dir. $\hat{\alpha}$ bilinmeyen $\eta_{x=0} = \alpha$ yı saptar , yani x=0 için

$$\hat{\eta}_{x=0} = \hat{\alpha}\,, \qquad E(\hat{\eta}_{x=0}) = \alpha\,, \qquad V(\hat{\eta}_{x=0}) = \sigma^2\left[\frac{1}{N} + \frac{(0-\bar{x})^2}{S(\dot{x}^2)}\right] \tag{1.25b}$$

yi elde ederiz.

Daha sonraki işlemler için $\hat{\alpha}$ ve $\hat{\beta}$ nın kovaryansını elde edelim. Yani ;

$$E\left[(\hat{\alpha}-\alpha)(\hat{\beta}-\beta)\right] = \operatorname{cov}(\hat{\alpha},\hat{\beta}) \tag{1.26}$$

(1.26) nın değerlerini belirlemek için (1.22) - (1.22e) den

$$(\hat{\beta}-\beta) = \sum_i^N c_i(y_i - \eta_i) = \sum_i^N c_i\left[y_i - E(y_i|x_i)\right] \tag{1.26a}$$

yı yazabiliriz ve benzer şekilde

$$(\hat{\alpha}-\alpha) = \sum_j^N \left(\tfrac{1}{N} - \bar{x} c_j\right)(y_j - \eta_j) \tag{1.26b}$$

yi görmek kolaydır. Böylece,

$$(\hat{\alpha}-\alpha)(\hat{\beta}-\beta)=\left[\sum_{1}^{N}\left(\frac{1}{N}-\bar{x}c_i\right)(y_i-\eta_i)\right]\left[\sum_{1}^{N}c_i(y_i-\eta_i)\right] \tag{1.27}$$

$$=\sum_{1}^{N}\left(\frac{1}{N}-\bar{x}c_i\right)c_i(y_i-\eta_i)^2$$

$$+\sum\sum_{i\neq j}\left(\frac{1}{N}-\bar{x}c_i\right)c_i(y_i-\eta_i)(y_j-\eta_j)$$

nin her iki tarafında beklenen değerleri alarak ve y_i nin $i \neq j$ için y_j den bağımsız olduğunu hatırlayarak;

$$\operatorname{cov}(\hat{\alpha},\hat{\beta})=\sigma^2\sum_{i=1}^{N}\left(\frac{1}{N}c_i+\bar{x}c_i^2\right)+0 \tag{1.27a}$$

$$=-\frac{\sigma^2\bar{x}}{S(x^2)}$$

yı elde edebiliriz.

Bu noktada şu soruyu sorabiliriz. "Bu durumda niçin en küçük kareler yöntemini kullanıyoruz ?". Çünkü; Gauss teoremi olarak adlandırılan teoreminin ifadesinde bu sonuç vardır. Bu teoremin bir yorumu burada tartışılacaktır. En basit şekilde ifade edersek Gauss teoremi; α veya β nın $\hat{\alpha}$ veya $\hat{\beta}$ en küçük kareler yönteminin, y_i lerde lineer olan herhangi bir başka tarafsız sapmalarının varyansından daha büyük olmayan varyansa sahip olduklarını iddia eder. Gerçekten bu $a_1\hat{\alpha}+a_2\hat{\beta}$, $a_1\alpha+a_2\beta$ ya ait bir minumum varyans tahminidir. Burada a_1 ve a_2 keyfi sabitlerdir. Buna göre aşağıdaki teoremi söyleyebiliriz.

Teorem: (x_i, y_i), $(i=1,\ldots,N)$ nin yukarıda bildirilen şartlar altında elde edilen gözlem noktaları olduklarını farz ediniz. Yani x de önceden seçilmiş değerleri $x_1,\ldots,x_N$, dir ve gözlenen $y_1,\ldots,y_N$ de x den bağımsızdır; Burada her i için $E(y_i|x_i)=\alpha+\beta x_i$ ve $V(y_i|x_i)=\sigma^2$ dir. (α,β) nın en küçük kareler yöntemini $(\hat{\alpha},\hat{\beta})$ ile gösterebiliriz.

$$\tau=a_1\alpha+a_2\beta \tag{1.28}$$

in tahmini ile ilgilendiğimizi farz ediniz .O zaman y_i ye göre lineer olan τ nun tüm tarafsız tahminleri arasında (en küçük kareler tahmini)

$$\hat{\tau} = a_1\hat{\alpha} + a_2\hat{\beta}$$

bir minimum varyansa sahiptir.

İspat: (1.22e) ve (1.24a) denklemlerinden $E(\hat{\tau}) = \tau$ olduğunu rahatlıkla görebiliriz. Şimdi τ dan bağımsız ve y ye göre lineer olan bir başka t tahminine sahip olduğumuzu farzedelim. O zaman

$$t = \sum_{i=1}^{N} d_i y_i \tag{2.29}$$

ve sapmasızlık (eğilimsizlik) bağımsızlık şartı (yani $E(t|x_1,\ldots,x_N) = E(t)$) şöyledir.

$$E(t) = \sum_{i=1}^{N} d_i E(y_i|x_i) = \sum_{i=1}^{N} d_i(\alpha + \beta x_i) = \tau$$

$$V(\hat{\tau}) = \sigma^2 \sum_{i}^{N} \left[\tfrac{a_1}{N} + (a_2 - a_1\bar{x})c_i\right]^2 \tag{1.30}$$

böylece tüm mümkün α ve β için

$$\alpha\sum_{i=1}^{N} d_i + \beta\sum_{i=1}^{N} d_i x_i = a_1\alpha + a_2\beta \tag{1.30a}$$

ve sonra

$$\sum_{i=1}^{N} d_i = a_1 \quad \text{ve} \quad \sum_{i=1}^{N} d_i x_i = a_2 \tag{1.30b}$$

denklemlerini elde ederiz. Şimdi, $\hat{\tau}$ en küçük kareler tahminini y_j nin bir lineer bir kombinasyonu olarak yazdığımızı farzedelim. Buradan

$$\hat{\tau} = a_1\hat{\alpha} + a_2\beta = \sum_{1}^{N} a_1\left(\tfrac{1}{N} - \bar{x}c_i\right)y_i + \sum_{1}^{N} a_2 c_i y_i \tag{1.31}$$

$$= \sum_{1}^{N} \left[\tfrac{a_1}{N} + (a_2 - a_1\bar{x})c_i\right] y_i$$

dir. Bu tarzda $\hat{\tau}$ nun varyansı -y_j lerin bağımsız olduklarını göz önünde tutarak-

$$V(\hat{\tau}) = \sigma^2 \sum_{i}^{N} \left[\tfrac{a_1}{N} + (a_2 - a_1\bar{x})c_i \right]^2 \tag{1.32}$$

dir. Ve (1.22 c) denklemini kullanarak

$$V(\hat{\tau}) = \sigma^2 \left[\frac{a_1^2}{N} + \frac{(a_2 - a_1 x)^2}{S(\dot{x}^2)} \right] \tag{1.33}$$

sonucunu elde ederiz. (1.31) deki beklenen değerden (1.30) daki beklenen değeri çıkartarak

$$\hat{\tau} - E(\hat{\tau}) = \hat{\tau} - \tau = \sum_{i}^{N} \left[\tfrac{a_1}{N} + (a_2 - a_1 x)c_i \right] (y_i - \eta_i) \tag{1.34}$$

ü elde ederiz ve benzer şekilde (1.29) dan (1.30) ifadesini çıkartırsak

$$t - E(t) = t - \tau = \sum_{i}^{N} d_i (y_i - \eta_i) \tag{1.35}$$

i elde ederiz. (1.27a) sonucunu elde ettiğimiz metodu kullanarak

$$\operatorname{cov}(\hat{\tau}, t) = \sigma^2 \sum_{i=1}^{N} \left[\tfrac{a_1}{N} + (a_2 - a_1 x)c_i \right] d_i \tag{1.35a}$$

yı elde ederiz ve (1.30b) yi kullanarak

$$\operatorname{cov}(\hat{\tau}, t) = \sigma^2 \left[\frac{a_1^2}{N} + (a_2 - a_1\bar{x}) \sum_{i}^{N} c_i d_i \right] \tag{1.35b}$$

yi elde ederiz. Fakat

$$\begin{aligned} \sum_{1}^{N} c_i d_i &= \sum_{1}^{N} \frac{(x_i - \bar{x}) d_i}{S(\dot{x}^2)} \\ &= \frac{\left[\sum_1^N x_i d_i - \bar{x} \sum_1^N d_i \right]}{S(\dot{x}^2)} \\ &= \frac{(a_2 - a_1\bar{x})}{S(\dot{x}^2)} \end{aligned} \tag{1.35c}$$

(1.35c) ifadesini (1.35b) denkleminde yerine koyarak

$$\operatorname{cov}(\hat{\tau}, t) = \sigma^2 \left[\frac{a_1^2}{N} + \frac{(a_2 - a_1\bar{x})^2}{S(\dot{x}^2)} \right] \tag{1.36}$$

yı elde ederiz .Bu ise bize (1.33) deki varyans ile (1.36) daki kovaryansın birbirine eşit olduğunu gösterir. Yani;

$$\mathrm{cov}(\hat{\tau},t)=V(\hat{\tau}) \tag{1.37}$$

dır. Farzedelimki $t=\hat{\tau}$ olsun. Buradan

$$0\leq V(t-\hat{\tau})=V(t)+V(\hat{\tau})-2\,\mathrm{cov}(\hat{\tau},t) \tag{1.38}$$

elde ederiz.(1.37) ve (1.38) de ispatlanmış olan

$$0\leq V(t)-V(\hat{\tau}) \tag{1.38a}$$

veya

$$V(\hat{\tau})\leq V(t) \tag{1.38b}$$

denklem sistemini elde ederiz. Burada Ancak ve ancak $t=\hat{\tau}$ olması durumunda eşitliğin meydana geldiğine dikkat edelim ki bu eşitlikte

$$t=\sum_{i=1}^{N}d_i y_i=\hat{\tau}=\sum_{i=1}^{N}\left[\tfrac{a_1}{N}+(a_2-a_1\bar{x})c_i\right]y_i$$

ve

$$d_i=\frac{a_1}{N}+(a_2-a_1\bar{x})c_i\ , \qquad i=1,\ldots,N \tag{1.39}$$

dir. Eğer eşitlik olmazsa

$$V(\hat{\tau})<V(t) \tag{1.39a}$$

dir. Burada yukarıdaki durumun çok genel ve özel olduğunu söyleyebiliriz ve buna göre aşağıdaki sonuçları elde edebiliriz.

Sonuç 1.2.1.1: $(a_1 = 0, a_2 = 1)$, y_i deki β nın bütün bağımsız tahminleri arasında, $\hat{\beta}$ en küçük kareler tahmini ile minumum varyansa sahiptir.

Sonuç 1.2.1.2: $(a_1 = 1, a_2 = 0)$, y_i de lineer olan α nın bütün bağımsız tahminleri arasında $\hat{\alpha}$ en küçük kareler tahmini ile minumum varyansa sahiptir.

Sonuç 1.2.1.3: $(a_1 = 1, a_2 = 1)$, y_i de lineer olan $\eta_x = \alpha + \beta x$ nın bütün bağımsız tahminleri arasında $\hat{\eta}_x = \hat{\alpha} + \hat{\beta}x$ en küçük kareler tahmini ile minumum varyansa sahiptir.

Bu bölümdeki bütün sonuçların zayıf varsayımları gözlemlerden elde edilmişlerdir. y_i gözlemleri birbiriyle ilişkisizdir. Gerçekte çok sıklıkla bu tahmin kullanılır. Elbette yapılan varsayımlar y_j lerle ilişkisiz ise ve normal şekilde dağılıyorsa elbette y_j ler birbirinden bağımsızdır.

1.3. σ^2 nin Tahmini

Daha önceden ifade ettiğimiz gibi uydurmakta olduğumuz model doğru ise [şayet $E(y|x) = \eta_x$, x e göre lineer ise ve biz bu verilere göre bir doğru uydurursak], o zaman $e_i = y_i - \hat{\eta}_{x_i}$ kalanları bize sadece hata hakkında yani sadece σ^2 hakkında bir bilgi verirler. Gerçekte $SS_e = \sum_1^N e_i^2$, σ^2 hakkındaki bilgiyi ifade eder. ((1.20c) yi göz önünde tutarak (1.44) ' e ifadesine bakınız.)

Bunu görmek için ilk olarak

$$SS_e = \sum_1^N (y_i - \hat{\alpha} - \hat{\beta}x_i)(y_i - \hat{\alpha} - \hat{\beta}x_i) \qquad (1.40)$$

$$= \sum_1^N (y_i - \hat{\alpha} - \hat{\beta}x_i)y_i - \hat{\alpha}\sum_1^N (y_i - \hat{\alpha} - \hat{\beta}x_i)$$

$$-\hat{\beta}\sum_1^N (y_i - \hat{\alpha} - \hat{\beta}x_i)x_i$$

yazabiliriz. [(1.20b) ye bakınız] Fakat yukarıdaki SS_e ifadesindeki son iki terimin (1.15a) denkleminden sıfır olduğu görülür. Böylece

$$SS_e = \sum_1^N y_i^2 - \left(\sum_1^N y_i + \hat{\beta} \sum_1^N x_i y_i \right) \tag{1.40a}$$

$$= SS_t - SS_r$$

yı elde ederiz. (1.40a) ifadesi hesaplama açısından önemlidir. Bu denklem bize kalanların karelerinin toplamının, gözlemlerinin karelerinin toplamından yani $\hat{\alpha}$ ve $\hat{\beta}$ tahminlerinin en küçük kareler yönteminin sırasıyla birinci ve ikinci normal denklemlerinin sağ taraflarıyla çarpımlarının toplamlarını çıkartarak hesaplanabildiğini ifade eder. Bu ikinci toplam SS_r ile gösterilir ve karelerin regresyon toplamı olarak ifade edilir. Normal denklemleri ise (1.15b) ile verilmişlerdir.

Şimdi (1.15b) normal denklemleri kullanılarak ,

$$SS_r = \hat{\alpha} \sum y_i + \hat{\beta} \sum x_i y_i = N\hat{\alpha}^2 + 2(\sum x_i)\hat{\alpha}\hat{\beta} + (\sum x_i^2)\hat{\beta}^2 \tag{1.40b}$$

ve de

$$SS_r = \sum \hat{\eta}_u^2 \tag{1.40c}$$

olduğu ispat edilebilir.

SS_r ye ait (1.40c) alternatif form bu noktada bizim için önemlidir. -Yani SS_r, $\hat{\eta}_u$ nun bir toplamı olduğundan bu karelerin regresyon toplamıdır; Burada $\hat{\eta}_u$, $\hat{\eta} = \hat{\alpha} + \hat{\beta} x$ regrasyon doğrusundaki x_u noktalarının ordinatlarıdır.

Ayrıca (1.40b) nın sağ tarafının şekli ilgi çekicidir. Çünkü $\hat{\alpha}^2$, $2\hat{\alpha}\hat{\beta}$ ve $\hat{\beta}^2$ terimlerinin kat sayıları (1.15b) normal denklemlerinin katsayılarından elde edilen elemanlardır. $\hat{\alpha}^2$ nin katsayıları için 1. normal denklemdeki 1. katsayıyı (yani 1. normal denklemdeki $\hat{\alpha}$ nın katsayısını kullanınız.); $\hat{\beta}^2$ nın katsayısı için 2. normal denklemdeki 2. katsayıyı (yani

2. normal denklemdeki $\hat{\beta}^2$ nın katsayısını kullanınız.); ve $2\hat{\alpha}\hat{\beta}$ nın katsayısı için 1. normal denklemde ortaya çıkan 2. katsayıyı yada 2. normal denklemde ortaya çıkan 1. katsayıyı kullanınız. (İkisinden birinin seçimi $\hat{\alpha}\hat{\beta}$' ın önünde olan 2.çarpanını hesaba katarız). Normal denklemlerin katsayıları herhangi bir (parametrelere göre) lineer model için çok önemli olduğu meydana çıkar

(1.40b) ve (1.40c) de zikredilen niceliklerden herhangi bir tanesine SS_r ile gösterilen karelerinin regrasyon toplamı olarak bakabiliriz. Yani

$$SS_t = N\hat{\alpha}^2 + 2\sum x_i\hat{\alpha}\hat{\beta} + \sum x_i^2\hat{\beta}^2 = \hat{\alpha}\sum y_i + \hat{\beta}\sum x_i y_i = \sum_{\mu=1}^{N}\hat{\eta}_u^2 \tag{1.40d}$$

dir. Şimdi onun beklenen değerini araştıralım. Biz sonsuz 1. ve 2. momentlerle herhangi bir rastgele değişkeni W için $\mu_W = E(W)$ ve $\sigma_W^2 = E\left[(W-\mu_W)^2\right]$, yerine koyarak

$$E(W^2) = \sigma_W^2 + \mu_W^2 \tag{1.41}$$

olduğunu ve benzer şekilde mevcut uygun momentlerle 2 rastgele W ve V değişkenleri için

$$E(WV) = \mathrm{cov}(W,V) + \mu_W\mu_V \quad ,\mathrm{cov}(W,V) = E\left[(W-\mu_W)(V-\mu_V)\right] \tag{1.41a}$$

olduğunu zaten biliyoruz. (1.40b) nin sağ tarafını kullanarak kısım 1.2 nin sonuçlarından

$$E(SS_r) = N\left\{\sigma^2\left[\frac{1}{N} + \frac{\bar{x}^2}{S(\dot{x}^2)}\right] + \alpha^2\right\} + 2N\bar{x}\left[-\frac{\sigma^2\bar{x}}{S(\dot{x}^2)} + \alpha\beta\right] \tag{1.42}$$

$$+(\sum x_i^2)\left[\frac{\sigma^2}{S(\dot{x}^2)} + \beta^2\right]$$

elde ederiz ve basit cebirsel işlemlerden sonra

$$E(SS_r) = 2\sigma^2 + \left[N\alpha^2 + 2(\sum x_i)\alpha\beta + (\sum x_i^2)\beta^2\right]$$

yı elde ederiz. [Esasen $S(\dot{x}^2)=\sum x_i^2 - N\bar{x}^2$ den olayı] (1.42a) hakkında iki uyarımız vardır. İlgilendiğimiz modelde iki tane α ve β parametresi vardır ve de (1.42a) da $2\sigma^2$ teriminde 2 çarpanını ortaya çıkartır. α ve β bilinmeyenlerindeki bir kuadratik ifade yardımıyla bu 2 çarpanı getirilir. Bu kuadratik formun katsayıları daha önce verilmişti ve (1.40b) ifadesinde tartışılmıştı.

$E(SS_e)$'yi hesaplamak kolaydır; (1.40a) ve (1.40b) denklemlerinden

$$SS_e = \sum_1^N y_i^2 - SS_r \tag{1.43}$$

elde ettiğimizden dolayı

$$E(SS_e) = \sum_1^N E(y_i^2) - E(SS_r) \tag{1.43a}$$

$$= \sum_1^N \left[\sigma^2 + (\alpha + \beta x_i)^2\right] - E(SS_r)$$

dır. Bu ifadeyi açarsak

$$E(SS_e) = N\sigma^2 + (N\alpha^2 + 2\alpha\beta\sum x_i + \beta^2 \sum x_i^2 - E(SS_r) \tag{1.43b}$$

yi elde ederiz. (1.42a) denklemini kullanarak

$$E(SS_e) = (N-2)\sigma^2 \tag{1.44}$$

elde ederiz. (1.44) ifadesinde görülen $N-2$ carpımına hataya ait serbestlik derecesi (DF) deriz. Şayet

$$S_{y.x}^2 = \frac{SS_e}{(N-2)} = \frac{\sum(y_i - \hat{\eta}_i)^2}{(N-2)} \tag{1.45}$$

ifadesini yazarsak

$$E(S_{y.x}^2) = \sigma^2 \tag{1.46}$$

yı elde ederiz. Yani $S_{y.x}^2$, σ^2 nın sapmasız bir tahminidir.

1.4. Varyans Analizi I

(1.43) den ilginç bir benzerliği

$$\sum_{1}^{N} y_i^2 = \sum_{i=1}^{N} (y_i - \hat{\eta}_{x_i})^2 + \left[N\alpha^2 + 2(\sum x_i)\alpha\beta + (\sum x_i^2)\beta^2\right] \tag{1.47}$$
$$= SS_e + SS_r$$

yı elde ederiz. (1.47) özdeşliği karelerin genel toplamını iki kısma bölebileceğimizi ifade eder. Biri SS_e hataları hakkında bilgi verir diğeri ise SS_r regresyon denkleminin parametreleri hakkında bilgi verir. (1.47) nın elemanlarını varyans analiz tablosu olarak isimlendirilen aşağıdaki tabloda gösterebiliriz. (Necla ÇÖMLEKÇİ,1988) kitabında yer almaktadır.

Tablo 1.4.1 : Modelin Varyans Analizi

$E(y|x) = \alpha + \beta x$, $V(y|x) = \sigma^2$

Kaynak	Serbestlik Dereceleri (DE)	Kareler Toplamı	Ortalama Kare**	Beklenen Ortalama Kare
Regresyon	2	SS_r	$SS_r/2$	$\sigma^2 + 1/2(N\alpha^2 + 2N\bar{x}\alpha\beta)$ $+1/2(\sum x_i^2)\beta^2$
Hata	$N-2$	SS_e	$SS_e/(N-2)$	σ^2
Toplam	N	$\sum y_i^2$		

**Ortalama Kare=Karelerin Toplamı/DF

Serbestlik derecesi hakkındaki yorumumuzu ileride yapacağız. Bu noktada bunlar şu şekilde değerlendirilebilirler. Verileri olmadan önce $y_1,\cdots,y_N$ veri noktası N boyutlu oklidyen uzayı R^N de herhangi bir yerde bulunabilir. (x_i, y_i), $i = 1,\cdots,N$ üzerineki α ve β yı hesaplamaya yani (1.15b) normal denklemlerini veya eşdeğer olarak (1.17) ile verilen çözümleri bulmak için iki kısıtlamaya gerek vardır. Yani $(\hat{\alpha},\hat{\beta})$, $y_1,\cdots,y_N$ nin N boyutluk bir örnek uzayın iki boyutlu örnek alt uzayında bulunmaya kısıtlandırılır. $e_i = y_i - \hat{\eta}_i$, N tane niceliğinin $(\hat{\alpha},\hat{\beta})$ yı içeren alt uzayda bulunması gerektiği ve onun boyutunun N-2 olduğu gösterilecektir. Burada e_i ler N sayıdadır. Fakat (1.15a) yı yani tekrar yazılmış hali ile

$$\sum e_i = \sum (y_i - \hat{\eta}_{x_i}) = 0 \tag{1.48}$$
$$\sum e_i x_i = \sum (y_i - \hat{\eta}_{x_i}) x_i = 0$$

denklemlerini sağlamak zorundadır.

Tablo 1.4.1 de beklenen ortalama kare sutununun incelenmesi ile "$\alpha = 0$ ve $\beta = 0$" bildirileri hakkındaki bilginin $SS_r/2$ regresyonuna ait ortalama kareyi $SS_e/(N-2)$ ile kıyaslayarak elde edilebildiğini gösterir. Şayet "$\alpha = 0$ ve $\beta = 0$" iddiası doğru ise o zaman bu ortalama karelerin her ikiside aynı σ^2 beklenilen değerine sahiptir. Bu ortalama karelerin tam olarak nasıl karşılaştırıldığı Kısım 1.6 da tartışılacaktır.

1.5 Varyans Analizi II

y yi x üzerinde regresyon ederekten ve kısım 1.1 deki model ile ilgilenirken biz "$\alpha = 0$ ve $\beta = 0$" durumu ile ilgilenmeyeceğiz. Parametrelerden sadece biri ile alakalı durum ile ilgilenneceğiz. (örneğin $\beta = 0$) O zaman bizim amacımız β hakkında bilgi veren karelerin toplam regresyon kısmını bulmak veya başka bir değişle α hakkında bilgi içermeyen karelerin regrasyon toplamı şeklinde ifade etmek olacaktır. Kullandığımız işlem yolu ortagonalleştirme olarak isimlendirilir. Bununla ilerideki konularda çok genel olarak inceleneceğiz. Burada anlatılmak istenen basit modeli aşağıdaki tarzda yazabiliriz.

$$\begin{aligned} E(y_i|x_i) &= \alpha + \beta x_i \\ &= (\alpha + \beta \bar{x}) + \beta (x_i - \bar{x}) \\ &= \phi + \beta (x_i - x) \end{aligned} \tag{1.49}$$

Burada $\bar{x}$ ler x'in N tane önceden seçilmiş x_i değerlerinin ortalamasıdır. $\sum_{i=1}^{N}(x - \bar{x}) = 0$ dır. Bundan başka şayet ϕ ve β değerlerini biliyorsak biz α 'yı otomatik olarak biliriz. Çünkü

$$\phi = \alpha + \beta \bar{x} \quad \text{veya} \quad \alpha = \phi - \beta \bar{x} \tag{1.49 a}$$

dır. Gerçekten (1.49), (ϕ ve β parametrelerine göre bir lineer modeldir ve $x_i - \bar{x}$ 'yi w_i ile gösterirsek

$$\eta_i = E(y_i|x_i) = \phi + \beta\, w_i \tag{1.50}$$

elde ederiz.

$$\sum (y_i - \eta_i)^2 = \sum (y_i - \phi - \beta\, w_i)^2 \tag{1.51}$$

i minimize etmekle ϕ ve β in en küçük kare tahminlerini bulabiliriz. [Tabii ki (1.51) denklemi Q gibi aynı niceliktedir]. Çünkü

$$\sum (y_i - \eta_i)^2 = \sum (y_i - \phi - \beta w_i)^2 = \sum (y_i - \alpha - \beta x_i)^2 .$$

ifadesinin ϕ ve β ya göre diferansiyelini alarak ve

$$\sum w_i = \sum (x_i - \bar{x}) = 0 \tag{1.51a}$$

olduğunu hatırlayarak (ϕ, β) nın en az kareler tahminlerine ait normal denklemlerini bulabiliriz. Örneğin $(\hat{\phi}, \hat{\beta})$

$$N\hat{\phi} = \sum y_i \tag{1.52}$$

$$\hat{\beta} \sum w_i^2 = \sum_{1}^{N} w_i y_i$$

dir. Bu denklemler kendilerini çözerler ve $w_i = x_i - \bar{x}$ olduğundan

$$\hat{\phi} = \bar{y} \tag{1.53}$$

$$\hat{\beta} = \frac{\sum (x_i - \bar{x}) y_i}{S(\dot{x}^2)} \tag{1.53a}$$

ü elde ederiz. Bundan başka Gauss teoremi $\tau = (1)\phi + (-\bar{x})\beta$ tüm (y_i göre) lineer tahminleri içinde

$$\hat{\tau} = (1)\hat{\phi} + (-\bar{x})\hat{\beta} \qquad (1.53b)$$

tahminin minumum varyansa sahip olduğunu ifade eder.

$$\hat{\tau} = (l)\bar{y} - \hat{\beta}\bar{x} = \hat{\alpha} \qquad (1.54)$$

olduğunu hatırlayalım ve gerçekten $\hat{\tau} = \hat{\alpha}$, $\tau = (1)\phi + (-\bar{x})\beta = \alpha$ nın en küçük kareler tahminidir. Yukarıdaki tahminlere ilave olarak uydurulan doğru önceden olduğu gibi

$$\hat{\eta}_i = \hat{\phi} + \hat{\beta} w_i = \hat{\phi} + \hat{\beta}(x_i - \bar{x}) = \hat{\alpha} + \hat{\beta} x_i \qquad (1.54a)$$

dir. ve (ϕ ve β) dan gelen karelerin toplam regresyonu (1.40b) denkleminden sonra tartışılan formasyon kuralına göre katsayıları (1.52) normal denkleminde elde edilen (ϕ,β) ya göre kuadratiktir. Böylece

$$SS_r = N\hat{\phi}^2 + 2(0)\hat{\phi}\,\hat{\beta} + (\sum w_i^2)\hat{\beta}^2 \qquad (1.55)$$

yi elde ederiz. Buradan

$$SS_r = N\hat{\phi}^2 + \sum (x_i - \bar{x})^2 \hat{\beta}^2 \qquad (1.55a)$$

$$= N\bar{y}^2 + \hat{\beta}^2 S(\dot{x}^2)$$

$$= \hat{\phi}\sum y_i + \hat{\beta} S(\dot{x},\dot{y}) \qquad (1.55b)$$

dir. Artık ifadenin tek başına $\hat{\phi}$ bağlı bir kısma ve tek başına $\hat{\beta}$ ya bağlı bir kısma ayrılmış oluduğunu görebiliriz. -(1.55a) denklemin tahminlerinin (1.52) normal denklemlerinin sağ tarafınına karşılık gelen tahminlerin çarpımlarının bir toplamı olduğunu ifade eder.- Gerçekten (1.43) denklemine bakarak

$$SS_e = \sum_1^N (y_i - \hat{\eta}_i)^2 = \sum_1^N y_i^2 - \left[N\bar{y}^2 + \hat{\beta}^2 S(\dot{x}^2)\right] \qquad (1.56)$$

$$= \sum_1^N (y_i - \bar{y})^2 - \hat{\beta}^2 S(\dot{x}^2)$$

yı elde ederiz. Bu (1.40a)'dan doğrudan ispat edilebilir. [Ek A.1.1. ın A.1.1.10 denklemine bakınız.] (1.56) 'denkleminin 1.satır elemanlarının varyans analizi tablosubir ile tablolaştırınız. [Tablo 1.5.1]

ϕ ve β dan gelen karelerin beklenen toplamı

$$2\sigma^2 + N(\alpha+\beta\overline{x})^2 + \beta^2 S(\dot{x}^2) \tag{1.57}$$

şeklinde ortaya çıkar. Bu biraz cebirsel işlemden sonra (1.42a) ile aynıdır.

Şimdi beklenen ortalama kare sutunun incelenmesi β 'dan geçen kareler regresyon toplamının hata ortalama karesi ile kıyaslanması "$\beta=0$" iddiasına bir yol gösterir. Şayet bu iddia doğru ise

$$E(\hat{\beta}^2 S(\dot{x}^2)) = E\left(SS_e/(N-2)\right) = \sigma^2 \tag{1.58}$$

dır. Tek başına β ile ilgilendiğimiz bu durumda varyans analizi tablosu tablo 1.5.1 'in bir iyileştirilmiş versiyonudur ve Tablo 1.5.1(a) daki gibi oluşturulur.

Tablo 1.5.1(a) esasen tablo 1.5.1 dır. Fakat ikinci tabloda birinci satır yoktur ve biz bu durumu birinci satırdaki elemanları Tablo 1.5.1 deki tüm satırdan çıkartarak telafi ederiz. Bunun beklenebildiğine dair bir ipucu karelerin kalan toplamı SS_e ye ait alternatif ifadede yatar. (1.40 a) dan

Tablo 1.5.1:β İle Elde Edilen Lineer Modelin Varyans Analizi

Kaynak	Serbestlik derecesi	Kareler Toplamı	Beklenen Ortalama Kare
ϕ	1	$N\hat{\phi}^2 = N\overline{y}^2 = \hat{\phi}\sum y_i$	$\sigma^2 + N\phi^2$
β	1	$\hat{\beta}^2 S(\dot{x}^2) = \hat{\beta} S(\dot{x},\dot{y})$	$\sigma^2 + \beta^2 S(\dot{x}^2)$
Hata	N-2	SS_e	σ^2
Toplam	N	$\sum y_i^2$	

Tablo 1.5.1(a)

Kaynak	Serbestlik Derecesi	Kareler Toplamı	Ortalama Kare	Beklenen ortalama Kare
β	1	$\hat{\beta}^2 S(\dot{x}^2)$	$\hat{\beta}^2 S(\dot{x}^2)/1$	$\sigma^2 + \beta^2 S(\dot{x}^2)$
Hata	*N-1*	SS_e	$SS_e/(N-2)$	σ^2
Toplam	*N-2*	$\sum y_i^2 - N\bar{y}^2 = \sum (y_i - \bar{y})^2$		

$$SS_e = \sum y_i^2 - (\hat{\alpha} N\bar{y} + \hat{\beta}\sum x_i y_i) \tag{1.59}$$
$$= \sum y_i^2 - \left[(\bar{y} - \hat{\beta}\bar{x})N\bar{y} + \hat{\beta}\sum x_i y_i\right]$$
$$= \sum y_i^2 - \left[N\bar{y}^2 + \hat{\beta}(\sum x_i y_i - N\bar{x}\bar{y})\right]$$

yazabiliriz ve (1.17a) yı kullanarak

$$SS_e = \sum y_i^2 - \left[N\bar{y}^2 + \hat{\beta}\,\mathrm{S}(\dot{x},\dot{y})\right]$$
$$= \sum (y_i - \bar{y})^2 - \hat{\beta}\,\mathrm{S}(\dot{x},\dot{y})$$
$$= \sum (y_i - \bar{y})^2 - \hat{\beta}^2 \mathrm{S}(\dot{x}^2)$$

veya bir başka tarzda

$$\sum (y_i - \bar{y})^2 = SS_e + SS_r(\beta) \tag{1.59a}$$

dır.Burada β dan gelen $SS_r(\beta)$ karelerin regresyon toplamı

$$SS_r(\beta) = \hat{\beta}\,\mathrm{S}(\dot{x},\dot{y}) = \hat{\beta}^2 \mathrm{S}(\dot{x}^2) \tag{1.59b}$$

olarak belirlenir. Tablo1.5.1 de ve tablo1.5.1(a) da SS_r, β nın σ^2 den ayrı olarak β hakkında bilgi verdiğini gördük. Çünkü β nın beklenen değeri $\sigma^2 + \beta^2 S(\dot{x}^2)$ dır. Bundan başka (1.44) ve (1.45) den $S_{y.x}^2$ nin σ^2 dan tarafsız olduğunu elde ederiz. (1.59a) nın elemanları Tablo1.5.1(a) varyans analizi için bir bazdır. Burada $\beta=0$ ifadesi ilgi odağımızı oluşturur.

Biz genellikle β ya bağlı karelerin extra regrasyon toplamı olarak $SS_r(\beta)$ yı sunarız. Gerçekten $\eta=\alpha+\beta x$ olduğu durumda bazen SS_e 'i, $SS_e(\alpha,\beta)$ olarak yazabiliriz. Böylece (1.59a) dan

$$SS_r(\beta)=\sum(y_i-\bar{y})^2-SS_e(\alpha,\beta) \qquad (1.59c)$$

yi elde ederiz. Aşağıda gösterildiği gibi (1.59c) denkleminin ilginç bir yanı vardır. Bağımsız y_i leri $\eta_i=E(y_i)=\alpha$ ve $V(y_i)=\sigma^2$ şartlarını içeren durumlarda olduğunu farzedelim. O zaman çok iyi bilindiği gibi eğer en küçük kareleri uygularsak ($\sum(y_i-\alpha)^2$ kriterini minimize ederek $\eta=\alpha$ doğrusunun ayarlanması)

$$\hat{\alpha}=\bar{y} \qquad (1.59d)$$

çözümünü elde ederiz ve bu model için $SS_e(\alpha)$ kalanlarının karelerin toplamı , σ^2 ile ilgili bilginin

$$SS_e(\alpha)=\sum(y_i-\bar{y})^2=\min_{\alpha}\sum(y_i-\alpha)^2 \qquad (1.59e)$$

gibi olduğunu ortaya koyar. Gerçekte eğer bütün i ler için $E(y_i)=\alpha$ ve $V(y_i)=\sigma^2$ ise $E(\hat{\alpha})=E(\bar{y})=\alpha$ ve $E\left[\sum(y_i-\bar{y})^2\right]=(N-1)\sigma^2$ olur. Böylece (1.59e) denklemini

$$SS_r(\beta)=SS_e(\alpha)-SS_e(\alpha,\beta) \qquad (1.59f)$$

şeklinde yeniden yazabiliriz. Şimdi "β 'ya bağlı karelerin ek regresyon toplamı" ifadesi modelde β nın (yani βx) dahil edilmesi ile elde edilen karelerin regresyon toplamının parçası olarak gözönünde tutulacaktır.

Denklem (1.59a) nın gözönüne alınmasıyla yukarıda sözü edilen bütün nicelikler pozitif olduğundan dolayı

$$\sum(y_i-\bar{y})^2>SS_e \qquad (1.59g)$$

yi elde ederiz. Gerçekten de bu

$$\frac{SS_e}{\sum(y_i-\bar{y})^2}=1-\frac{\hat{\beta}^2S(\dot{x}^2)}{\sum(y_i-\bar{y})^2} \qquad (1.59h)$$

dır. Sağ taraftaki son terim genellikle R^2 ile temsil edilir. Bu durumda (1.59g) ve (1.59h) sonuçlarıyla

$$SS_e = \sum (y_i - \bar{y})^2 (1 - R^2) \qquad , 0 \le R^2 \le 1 \tag{1.59i}$$

ifadesini elde ederiz. (1.59i) yi gözönüne alarak R^2 yi bir korelasyon katsayısı olarak kullanabiliriz. (Eğer $R^2 = 0$ ise $SS_e = \sum (y_i - \bar{y})^2$ olur.) Veriler $E(y_i) = \alpha$ modeline uymuş olacağını ifade eder ve x_i 'e karşılık gelen y_i leri regresyona tabii tutmada pasiftir. Çünkü eğer $V(y_i) = \sigma^2$ bütün i ler için (y_i, y_j) birbirinden bağımsız olduğu zaman $E(y_i) = \alpha$ ye "en küçük kareler yöntemini" uyguladığımızda elde edeceğimiz şey eninde sonunda karelerin hata toplamıdır. Çünkü bununla birlikte R^2, 1 değerine yaklaşırken SS_e azalır ve sıfıra yaklaşır. Veriler bize bu durumu $E(y_i | x_i) = \alpha + \beta x_i$ modeli ile iyi bir şekilde açıklar ve x_i üzerinde y_i regresyonun yapılması gerektiğini ifade eder.

R^2 için alternatif bir form aşağıdaki gibi bulunabilir:

$$R^2 = \frac{\hat{\beta}^2 S(\dot{x}^2)}{\sum (y_i - \bar{y})^2} \tag{1.59j}$$

$$= \frac{[S(\dot{x}, \dot{y})]^2}{S(\dot{x}) S(\dot{y})}$$

veya

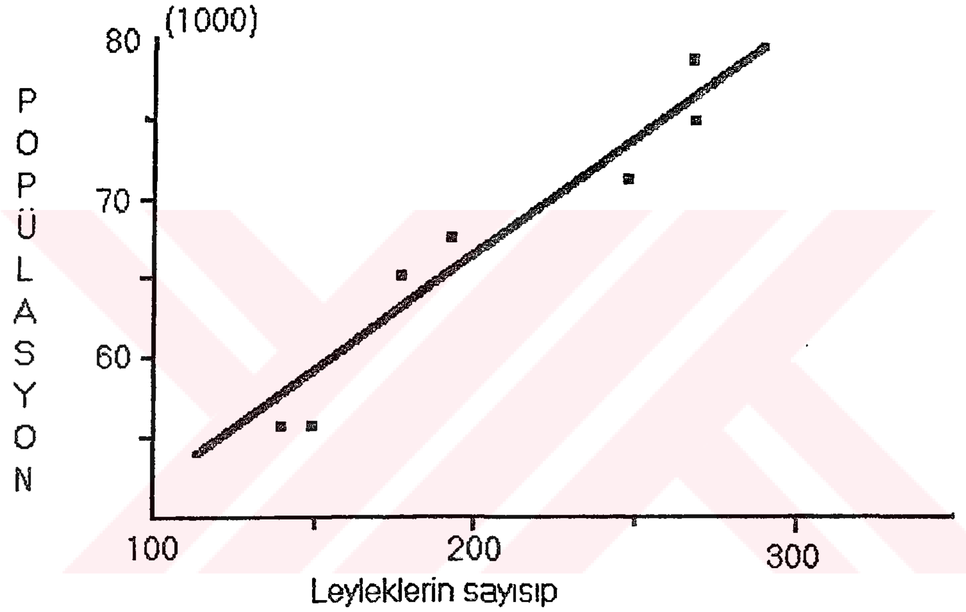
$$R = \frac{S(\dot{x}, \dot{y})}{S(\dot{x}^2) S(\dot{y}^2)} = \hat{\beta} [S(\dot{x}^2) / S(\dot{y}^2)]^{1/2} \tag{1.59k}$$

dir.(1.59i) den

$$-1 \le R \le 1 \tag{1.59l}$$

yi elde edebiliriz. R genellikle x ve y ler arasındaki korelasyon katsayısıdır. Dikkat edilirse yukarıda (1.59l) formülü ile açıklanan duyarlılıkta doğrusal ayarın bir ölçüsüdür.

Maalesef birçok kişi R nin "x, y ye sebep olsun " olarak atılan 1 civarındaki değerini kanıt olarak alır. Bu yanlış algılama genellikle x ile y arasında yüksek bir korelasyon olduğunu ortaya koyar. Genellikle başka sonuç yada sonuçlara neden olur. Herkez tarafindan bilinen bir örnek Norveçteki bir kasabanın nüfus büyüklüğü ve leyleklerin sayısı arasındaki ilişki ile ilgili örnektir. Şekil 1.5.1 de verilen çizim yani toplanan veri bir düz doğru uydurulmasını gerekli kılmıştır ve burada R nin yüksek bir değeri ortaya çıkar. Tabiiki leylekler doğumun sebebi olmadığından buradaki nokta leylek ve insanların nüfusunun zamanla artan fonksiyonlar olmasıdır.



ŞEKİL 1.5.1

(1.59j) den ve (1.59f) yi kullanarak aşağıdaki ilginç sonucu elde ederiz.,

$$R^2 = \frac{SS_r(\beta)}{\sum(y_i - \bar{y})^2} \tag{1.59m}$$

$$= \frac{SS_e(\alpha) - SS_e(\alpha,\beta)}{SS_e(\alpha)}$$

Bu bağıntı bir lineer karşılık fonksiyonunda birden fazla bağımsız değişken içerdiği zaman genelleşir. Çoklu korelasyon katsayısının bir tanımını verir.

1.6. İhtiyaç Duyulan Dağılım Teorisi: Testler ve Güven aralıkları:

Önceki kısımlarda

$$E(y|x) = \alpha + \beta x \qquad ,V(y|x) = \sigma^2 \tag{1.60}$$

Lineer modelimizin paremetreleri hakkındaki çeşitli iddialarımızı incelemek için değişik istatistiklerin nasıl kullanılabileceğinden emin değildik. Bunu yapmanın temeli önce x_i lerle verilen y dağılımın normal olduğunu farz etmek ve sonra bu kabul üzerine dayandırılmış bazı işlem yollarına baş vurmak ve en nihayetinde bu işlem yollarının normal olmayan dağılımlar için ne kadar iyi olduğunu araştırmaktır. Eğer y_i ve y_j iki rastgele değişkenleri ikideğişkenli normal dağılımı bitişik olarak dağıtılırsa ve aralarında bir ilişki kurulamazsa o zaman y_i ve y_j bağımsızdırlar. Böylece normallik kabulüne ilave olarak modeli aşağıdaki gibi

$$y_i = \mathrm{ID}N(\alpha + \beta x_i, \sigma^2) \qquad ,i = 1,\ldots,N \tag{1.61}$$

özetleyebiliriz. Burada ID notasyon gösterimi bağımsız olarak dağılanı temsil eder. IID notasyonunu ileride kullanma imkanımız olacaktır. Çünkü $x_i \neq x_j$ için

$$E(y_i|x_i) \neq E(y_j|x_j)$$

dir. Normallik kabulü geçerli olduğu zaman -önceki notasyonlarda *D=N dir-.* aşağıdaki teoremi söyleyebiliriz.

Teorem 1.6.1. (1.61) denkleminin kabulü altında (1.17) ile verilen $(\hat{\alpha},\hat{\beta})$ en az kare tahminleri normal ikideğişkenli dağılımından kaynaklanan bir birleşik dağılıma sahiptirler. Bu dağılımın yoğunluğu

$$\frac{\left[NS(x^2)\right]^{1/2}}{2\pi\sigma^2}\exp\left\{-\frac{1}{2\sigma^2}\left[N(\hat{\alpha}-\alpha)^2 + 2(\sum x_i)(\hat{\alpha}-\alpha)(\hat{\beta}-\beta) + \sum x_i^2(\hat{\beta}-\beta)^2\right]\right\} \tag{1.62}$$

dir. Bu

$$E(\hat{\alpha}) = \alpha \quad , E(\hat{\beta}) = \beta \quad , Cov(\alpha,\hat{\beta}) = -\frac{\bar{x}\sigma^2}{S(\dot{x}^2)} \tag{1.62a}$$

$$V(\hat{\alpha}) = \frac{\sigma^2 \sum x_i^2}{NS(\dot{x}^2)} = \sigma^2\left[\frac{1}{N} + \frac{\bar{x}^2}{S(\dot{x}^2)}\right]$$

$$V(\hat{\beta}) = \frac{\sigma^2}{S(\dot{x}^2)}$$

olduğunu ifade eder. $(\hat{\alpha},\hat{\beta})$ SS_e nin bağımsız dağılımlarıdır; Burada SS_e ifadesi

$$SS_e = \sigma^2 X_{N-2}^2 \tag{1.62b}$$

şeklindedir. Burada bu teoremi ispatsız bir şekilde kullanacağız.

Sonuç 1.6.1.1.:

$$N(\hat{\alpha}-\alpha)^2 + 2(\sum x_i)(\hat{\alpha}-\alpha)(\hat{\beta}-\beta) + (\sum x_i^2)(\hat{\beta}-\beta)^2 = T(\alpha,\beta) \tag{1.63}$$

rastgele değişken dağılımı

$$T = \sigma^2 X_2^2 \tag{1.63a}$$

şeklindedir ki T, $SS_e = \sigma^2 X_{N-2}^2$ den bağımsızdır.

Sonuç 1.6.1.1. birçok yollarda kullanabilirz. 1.Yol: (α,β) nın $(\hat{\alpha},\hat{\beta})$ tahminine ulaşıldıktan sonra

$$H_0 : \alpha = \alpha_0 \quad \text{ve} \quad \beta = \beta_0 \tag{1.64}$$

olarak gösterilen ve bir boş hipotez olarak bakılan H_0 ifadesini araştıracağımızı farzedelim. Her iki parametrede kullanılmıştır. Böyle bir hipoteze ekseriya ikisi aynı anda gerçekleşen bir hipotez olarak adlandırılır ve $x = x_i$ de örneklendirilmiş olan dağılımın $\alpha_0 + \beta_0 x_i$ ortalamasına sahip olduğnu ifade eder. σ^2 değeri hakkında herhangi bir şey söylenemediğinden dolayı dağılım H_0 ile belirsiz olarak bırakılmış olur. Böylece H_0 bir karışık hipotezdir.

Şimdi H_0 doğru ve

$$T = T(\alpha_0, \beta_0) \tag{1.65}$$
$$= N(\hat{\alpha} - \alpha_0)^2 + 2(\sum x_i)(\hat{\alpha} - \alpha_0)(\hat{\beta} - \beta_0) + (\sum x_i)(\hat{\beta} - \beta_0)^2$$

ise $\sigma^2 X_2^2$ ifadesi $SS_e = \sigma^2 x_N^2$ den bağımsız olarak dağıldığını kabul edelim. Burada (α, β) nın gerçek değerlerinin neler olabileceği önemli değildir. Böylece H_0 ı incelemek üzere bir istatistik test olarak

$$F = \frac{T(\alpha_0, \beta_0)/2}{SS_e/(N-2)} \tag{1.65a}$$
$$= \frac{T(\alpha_0, \beta_0)/2}{S_{y.x}^2}$$

yı kullanabiliriz. Çünkü H_0 doğru ise sonuç 1.6.1.1 den

$$F = F_{2,N-2} \tag{1.66}$$

olduğunu görmek çok kolaydır. Burada genel olarak F_{m_1,m_2}, (m_1, m_2) serbestlik dereceli Snedecor-Fisher rastgele değişkenini gösterir. γ önem seviyesinin red bölgesi

$$\text{Red : } H_0 : \alpha = \alpha_0, \beta = \beta_0 \qquad F > F_{2,N-1:\gamma} \text{ ise } \gamma \text{ seviyesinde} \tag{1.67}$$

ile belirtilir. (m_1, m_2) serbest dereceli Snedecor-Fisher dağılımını kullandığımız zaman $F_{m_1,m_2;\delta}$, δ olasılığı ile geçilen noktayı gösterir.]

Tablo 1.4.1 'i (varyans analizi tablosu) $\alpha_0 = 0$, $\beta_0 = 0$ geçerli olduğu zaman yukarıdakiler için hazırlarız. Çünki $T(0,0)/2$, $SS_r/2$, $SS_e/(N-2)$ i kıyaslayarak ve onların oranlarını kullanarak yukarıdakinin özel bir halini elde ederiz.

H_0: $\alpha = 0, \beta = 0$ test etmek için , Şayet

$$F = \frac{T(0,0)/2}{S_{y.x}^2} > F_{2,N-2:\gamma} \tag{1.68}$$

ise γ seviyesinde red ederiz. Tablo 1.4.1 deki beklenen ortalama kare kolonuna bakınız.

(1.68) de verilen işlem yolu "$\alpha=0$ ve $\beta=0$" iddiasını test etmek için $SS_r/2$; $SS_e/(N-2)$ ile kıyasladığımızı bildirir. Burada SS_r (1.40c) ile daha önce verilmişti. $\alpha=\alpha_0$, $\beta=\beta_0$ ve $(\alpha_0,\beta_0)\neq(0,0)$ ı test etmek için kullanılan işlem yolu ile (Burada (1.65a) da tanımlanmıştır.) aynı formdadır. Yani $T(\alpha_0,\beta_0)$ bulunan karelerin bir regresyon toplamıdır.

$$E(y_i|x_i)=\alpha+\beta\,x_i \tag{1.68a}$$

yı elde ederiz. $y_i=\alpha_0-\beta_0 x_i$ i y_i^* ile göstererek

$$E(y_i^*|x_i)=\alpha^*+\beta^*x_i \tag{1.68b}$$

yi elde ederiz. Burada

$$y_i^*=y_i-\alpha_0-\beta_0 x_i \quad , \ \alpha^*=\alpha-\alpha_0 \quad ,\beta^*=\beta-\beta_0 \tag{1.68c}$$

dır. Şayet y_i^* ı x_i üzerine regresyon uygularsak bu durumdaki karelerin regresyon toplamının

$$SS_r^*=n\hat{\alpha}^{*2}+2\sum x_i\hat{\alpha}^*\hat{\beta}^*+(\sum x_i^2)\hat{\beta}^{*2} \tag{1.68d}$$

olduğunu ve karelerin hata toplamının

$$SS_e^*=SS_e \tag{1.68e}$$

olduğunu buluruz.

Ayrıca "$\alpha=\alpha_0$ ve $\beta=\beta_0$" durumunun "$\alpha^*=0$, $\beta^*=0$" a eşdeğer olduğunu görürüz. Son olarak

$$\hat{\alpha}^*=\hat{\alpha}-\alpha_0, \qquad \hat{\beta}^*=\hat{\beta}-\beta_0 \tag{1.68f}$$

olduğu ve de

$$SS_e^* = T(\alpha_0,\beta_0) \tag{1.68g}$$
$$= N(\hat{\alpha}-\alpha_0)^2 + 2(\sum x_i)(\hat{\alpha}-\alpha_0)(\hat{\beta}-\beta_0) + (\sum x_i^2)(\hat{\beta}-\beta_0)^2$$

olduğu gösterilebilir. Tüm bu sonuçlardan sonra "bir kesmenin α_0 değerine sahip olduğu ve de bir eğimin β_0 değerine sahip olduğu" probleminden "belirli bir kesmenin sıfır değerine ve belirli bir eğiminde sıfır değerine sahip olduğu" problemini test etme hipotezine dönebiliriz. Bunu y_i den y_i^* a uygun dönüşümlerle ve y_i^* ı x_i üzerine regresyon etmekle ve bir regresyon ortalama karesini bir hata ortalaması ile yaparız.

Sonuç 1.6.1.1. in ilginç bir başka kullanımı, α ve β nokta tahminlerini bir eşzamanlı (α,β) güven aralığı ile tanımlamak istediğimiz zaman ortaya çıkar. Sonuç 1.6.1.1. i kullanarak

$$\frac{T(\alpha,\beta)/2}{S_{y.x}^2} = F_{2,N-2,\gamma} \tag{1.69}$$

ifadesini elde ederiz. Böylece

$$\Pr\left[T(\alpha,\beta) + 2S_{y.x}^2 F_{2.N-2,\gamma}\right] + \gamma \tag{1.70}$$

i elde ederiz. Burada P_r olasılığı temsil eder.) Çünkü (1.70) ifadesinin sol tarafını ve (1.69) u kullanılarak

$$\Pr\left[F_{2.N-2} + F_{2.N-2,\gamma}\right] + \Pr\left[F_{2.N-2} + F_{2.N-2,\gamma}\right] \tag{1.70a}$$

elde edilir ve böylece (α,β) ya ait $100(1-\gamma)\%$ birleşik veya eşzamanlı güven bölgesi α-β düzlemindeki (α,β) noktalarının bir kümesidir, bu (1.70) in sol tarafındaki parantezlerdeki eşitsizliği sağlar. Böyle bir küme

$$\left\{(\alpha,\beta)\middle| T(\alpha,\beta) \le 2S_{y.x}^2 F_{2.N-2,\gamma}\right\} \tag{1.70b}$$
$$= \left\{(\alpha,\beta)\middle| N(\alpha-\hat{\alpha})^2 + 2(\sum x_i)(\alpha-\hat{\alpha})(\beta-\hat{\beta}) + (\sum x_i^2)(\beta-\hat{\beta})^2 - 2S_{y.x}^2 F_{2.N-2,\gamma}\right\}$$

ile gösterilebilir. Bu $(\hat{\alpha},\hat{\beta})$ kümeside merkezlenmiş bir elipsin içi ve sınırıdır.

Çok sıkça kullanılan bir parametre ile örneğin α veya β nın herhangi birisiyle veya (α,β) nın $\tau = a_1\alpha + a_2\beta$ gibi bir lineer kombinasyonu ile ilgilenilir. Güven aralıkları ile uğraşmak için ve/veya böyle durumlara ait istatistikleri test etmek için aşağıdaki adımları kullanırız.

Sonuç 1.6.1.2. (1.61) kabullenmeleri altında

$$\hat{\tau} = a_1\hat{\alpha} + a_2\hat{\beta} \tag{1.71}$$

rastgele değişkeni normal olarak dağılır. Özel olarak

$$\hat{\tau} = N\left\{\tau, \sigma^2\left[\frac{a_1^2}{N} + \frac{(a_1 - a_2\bar{x})^2}{S(\dot{x}^2)}\right]\right\} \tag{1.71a}$$

ve $\hat{\tau}$, $SS_e = \sigma^2 X_{N-2}^2$ den bağımsız ve dolayısıyla

$$\hat{\tau}, \quad S_{y.x}^2 = SS_e/(N-2) = \sigma^2 X_{N-2}^2/(N-2)$$

den de bağımsızdır.

Bu sonuçtan

$$\frac{\hat{\tau} - \tau}{S_{y.x}\left[a_1^2/N + (a_2 - a_1\bar{x})^2/S(\dot{x}^2)\right]^{½}} = t_{N-2} \tag{1.72}$$

yi veya

$$\frac{(\hat{\tau} - \tau)^2}{S_{v=x}^2\left[a_1^2/N + (a_2 - a_1\bar{x})^2/S(\dot{x}^2)\right]} = F_{1.N-2} \tag{1.72a}$$

ifadesini kolaylıkla elde edebiliriz ve değişik durumlarda (1.72) ve (1.72a) yı kullanabiliriz. ilk önce $\eta = \alpha + \beta x$ regresyon dogrusunun eğimi olan β ile ilgilendiğimizi farzedelim. -β değişim oranıdır-. Yukarıdaki sonuçlarda $a_1 = 0, a_2 = 1$ koyarak

$$\tau = (0)\alpha + (1)\beta - \beta, \quad \hat{\tau} = \hat{\beta} \tag{1.73}$$

ve

$$\frac{\hat{\beta} - \beta}{\left[S_{y.x}^2/S(\dot{x}^2)\right]^{1/2}} = t_{N-2} \tag{1.73a}$$

yı elde ederiz. Bu denklem sadece β ile alakalı sıfır hipotezine ait bir test istatistiğini teşkil etmek için veya β ya ait güven aralıklarını bulmak için kullanılabilir.

Özel olarak β' nın en küçük kareler tahmini olan $\hat{\beta}$ 'ya ulaştığımızı farzedelim.

$$H_0^{(1)}: \beta = \beta_0 \tag{1.74}$$

ile tariflenen ifadeyi test etmek amacıyla kullanalım. [Herhangi bir sıfır H_0 hipotezi için, Eğer bildirilen sıfır hipotez doğru değilse diğer ifadenin doğru olduğunu kabul ederiz. Örneğin eğer (1.74) denklemi geçerli değilse o zaman β nın β_0 'a eşit olmayan herhangi bir pozitif reel değere sahip olabileceğini yani bunu alternatifi $H_1^1: \beta \neq \beta_0$ hipotezinin doğru olduğunu kabul ederiz.] Şimdi (1.73a) daki formulü kullanark $H_0^{(1)}$ i incelemeye ait bir test istatistiğini kullanalım

$$t = \frac{\hat{\beta} - \beta_0}{\left[S_{y.x}^2 / S(\dot{x}^2)\right]^{1/2}} \tag{1.75}$$

ve şayet t nin gözlenmiş bir değeri aşağıdaki gibi ise γ seviyesindeki $H_0^{(1)}$ 'i ret ederiz

$$|t| = t_{N-2.2} \tag{1.75a}$$

Burada genellikle $t_{m\delta}$, m serbestlik dereceli Student t dağılımını kullanırken δ olasılığı ile geçilen noktayı gösterir. Bu yukarıdaki işlem yolu ile eşdeğerdir.

Şayet gözlenen değer (1.75b) ise γ seviyesinde $H_0^{(1)}$ red edilir

$$F = t^2 = \frac{(\hat{\beta} - \beta_0)^2 S(\dot{x}^2)}{S_{y.x}^2} > F_{1.N-2;\gamma} \tag{1.75b}$$

$t_m^2 = F_{t,m}, F_{t,m,\delta} = t_{m,\delta/2}^2$ olduğunu hatırlayacak olursak. $\beta_0 = 0$ özel halinde tablo 1.51(a) nın elemanları ile ilgileniriz ve beklenen ortalama kare sütununa bakarak $H_0^{(1)} = \beta$ altında

$SS_r(\beta)$ ve SS_e yi onların ortalama karelerinin oranını kullanarak kıyaslarız; bu oran $\beta_0 = 0$, (1.75b) daki denklem tarafından tanımlanan F istatistiğidir. (1.73a) denkleminden

$$\left\{\beta \middle| \hat{\beta} - \left[\frac{S_{y.x}^2}{S(\dot{x}^2)}\right]^{1/2} t_{N-2;\gamma/2} \le \beta \le \hat{\beta} + \left[\frac{S_{y.x}^2}{S(\dot{x}^2)}\right]^{1/2} t_{N-2;\gamma/2}\right\} \tag{1.76}$$

aralığının (β değerlerinin kümesi) β ya ait bir $100(1-\gamma)\%$ güven aralığı olduğunu rahatlıkla görebiliriz.

Şimdi (1.72) denkleminde $a_1 = 0$ ve $a_2 = x$ değerleri yerine konulduğunda

$$\frac{\hat{\eta}_x - \eta_x}{S_{y.x}\left[1/N + (x-\bar{x})^2/S(\dot{x}^2)\right]^{1/2}} = t_{N-2} \tag{1.77}$$

denklemini elde ederiz. Burada $\eta_x = \alpha + \beta x$ ve $\hat{\eta}_x = \hat{\alpha} + \hat{\beta} x$ dir. (1.77) deki denklemi bir istatistik test olarak η_x e ait bir güven aralığını türetmek için kullanabiliriz.sonra $(1-\gamma)$ güven seviyesi için η_x e ait bir güven aralığı

$$\left\{\eta_x \middle| \hat{\eta}_x - S_{y.x}\left[\frac{1}{N} + \frac{(x-\bar{x})^2}{S(\dot{x}^2)}\right]^{1/2} t_{N-2;\gamma/2} \le \eta_x \le \hat{\eta}_x + S_{y.x}\left[\frac{1}{N} + \frac{(x-\bar{x})^2}{S(\dot{x}^2)}\right]^{1/2} t_{N-2;\gamma/2}\right\} \tag{1.78}$$

formunda elde edilir. Eğer x in değişmesine izin verirsek bir güven kuşağı elde ederiz. Üst ve alt şeritleri x e göre çizildiği zaman

$$\eta = \hat{\eta}_x + S_{y.x}\left[\frac{1}{N} + \frac{(x-\bar{x})^2}{S(\dot{x}^2)}\right]^{1/2} t_{N-2;\gamma/2} \tag{1.79}$$

denklemlerini elde ederiz. Bu denklem bize η_x in değerini verir ve bu kuşak $x = x_1$ de en az genişlemeye sahiptir.

(1.77)-(1.79) denklem silsilesinin özel hali $x=0$ olduğu haldir. O zaman

$$\eta_x = \eta_0 = \alpha \quad , \hat{\eta}_0 = \hat{\alpha} \tag{1.80}$$

ve (1.77) denkleminden

$$\frac{\hat{\alpha}-\alpha}{S_{y.x}\left[1/N+\bar{x}^2/S(\dot{x}^2)\right]^{1/2}}=t_{N-2} \tag{1.81}$$

i elde ederiz. Bu α için $100(1-\gamma)\%$ güven aralığını verir.

$$\left\{\alpha\middle|\hat{\alpha}-S_{y.x}\left[\frac{1}{N}+\frac{\bar{x}^2}{S(\dot{x}^2)}\right]^{1/2}t_{N-2;\gamma/2}\le\alpha\le\hat{\alpha}+S_{y.x}\left[\frac{1}{N}+\frac{\bar{x}^2}{S(\dot{x}^2)}\right]^{1/2}t_{N-2;\gamma/2}\right\} \tag{1.82}$$

verilen bir veri kümesi için (α,β) nın postula değerleri Örneğin (α_0,β_0) postula değerleri (1.82)-(1.76) aralıkları içinde sırasıyla çok güzel bir şekilde yer alabilirler α_0 ve β_0 değeri ise (1.70b) denklemi ile verilen (α,β) ya ait güven aralığının dışında yer alır. Yani (α_0,β_0) postula değerleri marjinal olarak kabul edilebilirler yani bu değerler kendi güven aralıklarına aittirler. Eğer (α_0,β_0) noktası (α,β) ya ait bitiş güven aralığının dışında yer alırsa eşzamanlı olarak gözlendikleri zaman kabul edilemezler. Burada verilerin büyük bir dikkatle incelenmesi gerektiği ve deneyin amacının akıldan çıkarmak gerektiğini söylemek yersizdir. Örneğin eğer deney β daki ilgiden dolayı yapılacak ise -β x' e göre karşılığın değişim oranıdır. O zaman (1.76) denklemi önem kazanır. Bununla birlikte eğer deney (α,β) ya ait bilgiyi elde etmek için yapılacaksa o zaman (1.70b) denklemini kullanmak amaca daha uygun olur.

1.7. Basit Bir Modeli Test Etme (Uyarlamanın iyiliği)

Şimdiye kadar bu bölümde veriye x in bir lineer fonksiyonu uydurmak için ortaya koyduğumuz faraziyeleri kullandık. Gerçekten η karşılık x değişkeninin lineer bir fonksiyonu olduğu yada iyi yaklaşımlarda (emprik) temellerde yaygın şekilde kullanılan x değerleri için karşılık fonksiyonunun, x e göre lineer olduğu birçok durumlar vardır. Bununla beraber lineerlik kabulünün kontrol edilmesinin önemli olduğu birçok örnekler vardır ve biz artık bu durum için uyumsuzluk olarak adlandırılan böyle bir metodu inceleyeceğiz.

Örnek olarak bir doğru uydurabiliriz. Fakat bu doğrunun uymadığını kabul edelim. Çünkü karşılık fonksiyonu x e göre kuadratiktir veya x in nonlineer bir fonksiyonudur ve biz ortaya çıkan bu uyumsuzluğu tesbit etmeyi arzuluyoruz. (Fikri AKDENİZ, 1985)' in kitabındada değindiği gibi şayet bir deneyi aşağıdaki şekilde ele alırsak yukarıdaki problemle uğraşmak için kendimize kolayca bir yöntem bulabiliriz. $x_0, x_1, x_2, \ldots, x_m$ in m tane değerini önceden seçtiğimizi ve herbir x_i de birbirinden istatistik olarak bağımsız n_j 'i gözlediğimizi farzediniz. Toplam numunenin büyüklüğünü $N = \sum_{j=1}^{m} n_j$ ile gösterelim ve x_j ler de üretilen n_j gözlemlerini $(y_{1j}, y_{2j}, \ldots y_{nj})$ ile gösterelim. Verilerin uyarlanan grafiği (x_{ij}, y_{ij}), $i = 1,2,\ldots,n_j$ ve $i = 1,2,\ldots,m$ noktaları tarafından oluşturulmuş olarak düşünebiliriz. Herbir , j *için* $n_j \geq 1$ olduğunu farzediyor ve $j - j'$ dan enaz birinin olduğunu farzediyoruz. Bunun için $n_j > 1$ dir. Burada $m \geq 3$ olduğunu kabul edeceğiz. (aşağıdaki (1.88a) ya bakınız). Bunun yanında $x_1 < x_2 < \cdots < x_m$ olduğunuda kabul diyoruz.

Bir diğer nokta en küçük kareler uyarlandığı zaman

$$\hat{\eta} = \hat{\alpha} + \hat{\beta} x \tag{1.83}$$

en küçük kareler eğrisini elde ederiz.Burada $\hat{\alpha}$ ve $\hat{\beta}$

$$\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x} \quad , \quad \bar{y} = \frac{1}{N}\sum_{j=1}^{m}\sum_{i=1}^{n_j} y_{ij} = \frac{1}{N}\sum_{j=1}^{m} n_j \bar{y}_{ij} \tag{1.84a}$$

ile

$$\bar{y}_{ij} = n_{ij}^{-1}\sum_{i=1}^{n_j} y_{ij} \tag{1.84b}$$

ve

$$\hat{\beta} = \frac{\sum_{j=1}^{m}\sum_{i=1}^{n_j}(x_j - \bar{x})(y_{ij} - \bar{y})}{S(\dot{x}^2)} = \frac{\sum_{j=1}^{m} n_j (x_j - \bar{x})(\bar{y}_{ij} - \bar{y})}{S(\dot{x}^2)} \tag{1.84c}$$

ile

$$S(\dot{x}^2) = \sum_{j=1}^{m} n_j(x_j - \bar{x})^2 \ , \quad x = \frac{\sum_{j=1}^{m} n_j x_j}{N} \tag{1.84d}$$

elde edilir. Tabii ki sapmaların karelerin minimum toplamı kalanların karelerinin toplamıdır. Burada

$$\hat{\eta}_{ij} = \hat{\alpha} + \hat{\beta} x_j \quad , \qquad i = 1,\ldots,n_j \tag{1.85a}$$

denklemi gerçekleştiğinde

$$SS_e = \sum_{j=1}^{m} \sum_{i=1}^{n_j} e_{ij}^2 = \sum\sum (y_{ij} - \hat{\eta}_{ij})^2 \tag{1.85}$$

denklemi elde edilir. Eğer bir başka fonksiyon uyarlanması gerekseydi bir lineer fonksiyondan ziyade kalanlar hatadan başka bilgiyi de içermeliydiler. Çünkü y_{ij} bize gerçek fonksiyon hakkında bilgi verir. $\hat{\eta}_{ij}$ sadece lineer bileşenler hakkında bilgi vermek üzere dizayn edilirler. Bu meyanda SS_e nin $\sigma^2 = V(y_{ij}|x_j)$ $(i = 1,2,\ldots,n_j$ ve $j = 1,2,\ldots,m)$ hakkındaki bilgiye ilaveten doğru fonksiyonu olarak kabul edilen lineer fonksiyonundan farklı bilgiyide içereceği beklenir ve böylece SS_e nin tek başına σ^2 hakkındaki bilgiden daha büyük olacağı beklenecektir. Bu eğer σ^2 hatasının bir tahminini elde edebilirsek -bu herhangi bir uydurmaya bağlı değildir.- bunu kontrol edebiliriz; veriyi elde ettiğimiz yol bunu yapmamıza müsaade eder.

İlk adım olarak yukarıdaki veri için " tek yol varyans analizi" yöntemini veriyi analiz etmek için kullanılabiliriz. Burada n_j gözlemleri ile m tane grup vardır $j = 1,2,\ldots,m$. Burada grupların ortalamalarının

$$E(y_{ij}|x_i) = \alpha + \beta x_j \ , \qquad j = 1,\ldots,m \tag{1.86}$$

olacağı hipotez edilir. Fakat (1.86) nın doğru olup olmadığından ve de x_j $, j = 1,2,\mathrm{K}\ ,m$ de populasyonlarının ortalamaları lineer bir fonkksiyon olmayacağından -x in daha yüksek mertebeden bir fonksiyonu örneğin kuadratik, kübik, exponansiyel veya sinüsoidal bir fonksiyon gibi x in herhangi birbaşka nonlineer fonksiyonu olabileceğinden- tek yol varyans analizi için hata terimi karelerin toplamı içindedir. Gerçekten m gruplarına ait tek yol varyans analizi tablosu Tablo 1.7.1. de gösterildiği gibi oluşturulabilir.

Tablo 1.7.1: Uygunsuz araştırma için m guplarına ait tek yol varyans tablosu analizi

Kaynak	Serbestlik Derecesi	Kareler Toplamı
Arası	*m-1*	$SS_\beta = \sum_{j=1}^{m} n_j (y_{ij} - \bar{y})^2$
İçi	$N - m = \sum_{j=1}^{m} (n_j - 1)$	$SS_W = \sum_{j=1}^{m} \sum_{i=1}^{n_j} (y_{ij} - \bar{y}_{ij})^2$
Toplam	$N - 1 = \sum n_j - 1$	$SS_L = \sum_{j=1}^{m} \sum_{i=1}^{n_j} (y_{ij} - \bar{y})^2$

"Karelerin örnek toplamı içinde" olarak adlandırılan SS_W tek başına σ^2 hatası hakkındaki bilgiyi verir. Standart elemanter teoriden $y_{1j},\ldots,y_{ij},\ldots,y_{nj}$,n_j tane bağımsız ve özdeş olarak dağıtılmış gözlemler -dağılım $E(y_{ij}|x_j) = \sigma^2$ ortalamasına $i = 1,2,\ldots,n_j$ için $V(y_{ij}|x_j) = \sigma^2$ varyansına sahiptir-. olduğundan

$$E\left[\sum_{i=1}^{n_j} (y_{ij} - \bar{y}_i)^2\right] = (n_j - 1)\sigma^2 \tag{1.87}$$

veya

$$E\left[\sum_{j=1}^{m}\sum_{i=1}^{n_j} (y_{ij} - \bar{y}_i)^2\right] = \sigma^2 \sum_{j=1}^{m} (n_j - 1) = \sigma^2 (N - m) \tag{1.87a}$$

yı elde ederiz.

Şimdi biz düzgün bir doğru uydurmakta olduğumuzdan *m-1* serbestlik dereceli karelerin toplamı arasını; $SS_r(\beta) = \hat{\beta}^2(\dot{x}^2)$ -bir serbestlik dereceli uyarlanan doğrunun eğiminden gelen karelerin regrasyon toplamına- ve (uyarlama bileşeninin eksikliğinin karelerinin toplamını temsil eden) *m-2* serbestlik dereceli ve SS_t ile gösterdiğimiz bir niceliğe ayırabiliriz. Burada SS_t

$$SS_L = SS_\beta - SS_r(\beta) \tag{1.88}$$

dır. Şayet lineerlik kabulü doğru ise o zaman

$$E(SS_L) = (m-2)\sigma^2 \tag{1.88a}$$

olduğu gösterilebilir. Halbuki doğru değilse

$$E(SS_L) = (m-2)\sigma^2 + A^2 \tag{1.88b}$$

denklemini yazabiliriz. Burada A^2 gerçek fonksiyonun lineerlikten ne kadar uzaklaştığına bağlıdır ve $\eta = \alpha + \beta x$ ise $A^2 = 0$ şeklindedir. Burada eğer $\eta \neq \alpha + \beta x$ ise o zaman $\eta = \gamma_0 + \gamma_1 x + \gamma_2 x^2$ olduğu farzedilir]. Buna mukabil varyans ile ilgili bir iki aşamalı varyans analizini tablolaştırabiliriz. (Tablo 1.7.2.)

Uyumsuzluğu test etmek için sadece $SS_L/(m-2)$ yi, $SS_W/(N-m)$ ile kıyaslayalım. Gerçekten eğer biz y_{ij} dağılımına ait normalliği ispatsız bir şekilde kabul edersek

$$F_L = \frac{SS_L/(m-2)}{SS_W/(N-m)} \tag{1.89}$$

eğer

$$\eta = \alpha + \beta x \tag{1.90}$$

ve (1.89) hipotezi doğru ise $F_{m-2,N-m}$ dağılımına sahip olduğu gösterilebilir. Yani uydurulan bir doğru uyumsuzluk göstermez ve γ önem seviyesini (Tablo 1.7.2. deki beklenen ortama kare sütununa bakan işlem yolu) aşağıdaki gibidir.

Şayet $F_L > F_{m-2,N-m}$; γ ise lineerliği red ediniz. Aksi halde kabul ediniz. (1.91)

dir.

Tablo 1.7.2 : Bir Doğrunun Uyarlanmasında Uygunsuz varyans Analizi

Kaynak	Sebestlik Derecesi	Kareler Toplamı	Beklenen Kare Ortalaması
Arasında	Regresyon 1 *m-1*	$SS_r(\beta)=\hat{\beta}^2 SS(\dot{x}^2)$	$\sigma^2+\beta^2 S(\dot{x}^2)$ $+g(A^2)^a$
İçinde	Uygunsuzluk $N-m$	$SS_L=SS_\beta-\hat{\beta}^2 SS(\dot{x}^2)$ $SS_W=\sum\sum(y_{ij}-y_{i.})^2$	$\sigma^2+A^a/(m-2)$ σ^2
Toplam	$N-1$	$SS_T=\sum\sum(y_{ij}-\bar{y})^2$	

SS_e nin yani kalanların kareleri toplamının bir uygunsuzluğunu göstermek için Tablo.1.7.2. nin şeklini değiştirebiliriz.(1.59a) denklemini kullanarak

$$SS_e = SS_T - \hat{\beta}^2 S(\dot{x}^2) \quad , \quad SS_T = \sum\sum(y_{ij}-\bar{y})^2 \tag{1.92}$$

denklemini elde ederiz. Ayrıca Tablo.1.7.2. den

$$SS_T - \hat{\beta}^2 S(\dot{x}^2) = SS_L + SS_W \tag{1.92a}$$

ifadesinin gerçekleştiğini görürüz. Yani (1.92) ve (1.92a) denklemlerini bir arada düzenleyerek

$$SS_e = SS_L + SS_W \tag{1.93}$$

denklemini elde ederiz. Buna göre Tablo 1.7.2. nin yeni şekli Tablo 1.7.3. deki gibi olur.

Kareler sütununun toplamındaki elemanlar için $\hat{\beta}$ en küçük kareler tahmini ifadenin bir kare bulunduğu zaman işlemlere aşağıdaki gibi devam edebiliriz. Önce

$$SS_T = \sum\sum y_{ij}^2 - \frac{(\sum\sum y_{ij})^2}{N} \tag{1.94}$$

bağıntısını kullanarak $SS_r(\beta)$ yı buluruz ve SS_T yi hesaplarız. Sonra

$$SS_e = SS_T - SS_r(\beta) \tag{1.94a}$$

denklemini kullanarak SS_e yi buluruz.

$$SS_W = \sum_{j=1}^{m}\left[\sum_{i=1}^{n_j} y_{ij}^2 - \frac{\left(\sum_{i=1}^{n_j} y_{ij}\right)^2}{n_j}\right] \tag{1.94b}$$

yı yani

$$SS_W = \sum_{j=1}^{m}\sum_{i=1}^{n_j} y_{ij}^2 - \sum_{j=1}^{m}\left[\frac{\left(\sum_{i=1}^{n_j} y_{ij}\right)^2}{n_j}\right] \tag{1.94c}$$

yi hesaplayarak karelerin örnek toplamı içerisinde gerçek hata terimi $SS_{W'}$ yı elde ederiz.

Tablo 1.7.3 : Bir Doğru Uyarlandığı Zaman Varyansın Uygunsuzluk Analizi

Kaynak	Serbestlik Derecesi	Kareler Toplamı	Beklenen Ortalama Kare
Regresyon	*1*	$SS_r(\beta) = \hat{\beta} S(\dot{x}^2)$	$\sigma^2 + \hat{\beta}^2 S(\dot{x}^2) + g(A^2)$
	m-2	$SS_L = SS_e - SS_W$	$\sigma^2 + A^2/(m-2)$
Kalanlar	*N-2*	$SS_{W'} = \sum\sum(y_{ij} - y_{.j})^2$	σ^2
	N-m		
Toplam	*N-1*	$SS_T = \sum\sum(y_{ij} - y)^2$	

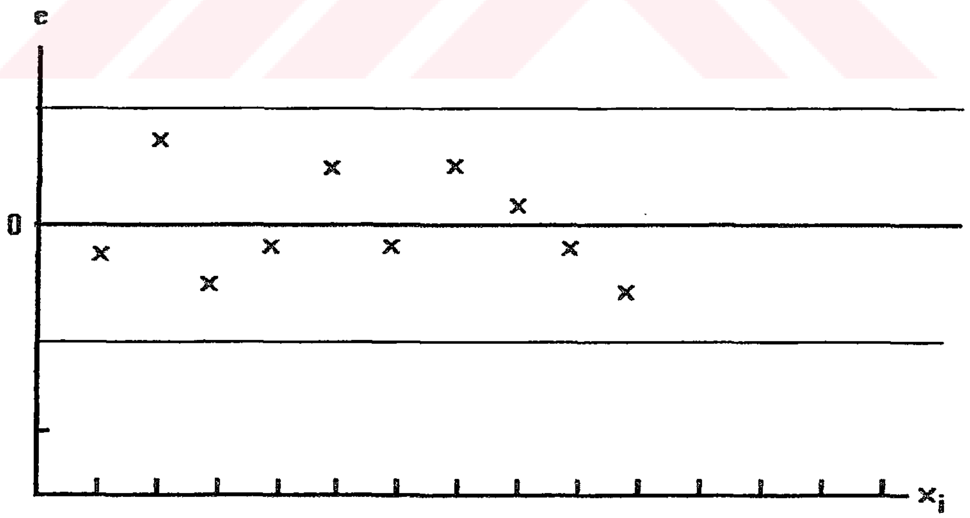
(1.94b) bağıntısı SS_W nın $y_{ij} - \bar{y}_{.j}$ sapmalarının karelerinin toplamını yani herbir $j = 1,\ldots,m$ için onların örnek ortalaması $\bar{y}_{ij}$ den j nci gruptaki gözlemlerin sapmalarını ve sonra bunları toplayarak (1.94c) ile tüm m gruplarının düzeltme faktörlerinin toplamını yani tüm $(\sum_{i=1}^{n_j} y_{ij})^2 / n_j$ niceliklerinin toplamını çıkartarak ve bütün y_{ij} gözlemlerin karelerin toplamını bularak hesaplanabildiğini ifade eder.

Eğer lineerlik hipotezi reel edilirse veriler η nün x in herhangi bir nonlineer bir fonksiyonu -sinüsoidal exponansiyel vs.- olabileceği kabülünün uyumsuzluğunu destekler. Şimdi η ya x in bir polinomial bir fonksiyonu olduğunu veya x in bir polinomial fonksiyonu ile iyi bir şekilde yaklaşıldığını biliyoruz. O zaman eğer lineerlik hipotezi red edilirse verilere kuadratik bir denklem uyarlayabilir ve yeni bir uygunsuzluk testi, kuadratiklik hipotezin karşılık fonksiyonunun kübikliğine karşı incelemek için yapılabilir. Bu tarzda veri gösterimini açmak için veriye en iyi şekilde uyan polinomial modeli bulabiliriz.

Eğer (1.91) testi lineerliği kabul ederse $\beta \neq 0$ hipotezine karşı $\beta = 0$ hipotezini (karşılık fonksiyonu x in sabit bir fonksiyonudur.) incelemekle meşgul olabiliriz. Bunu yapmak için $SS_t(\beta)$ ve $SS_e/(N-2)$ yi karşılaştırabilirz. Çünkü eğer $A=0$ [ve biz burada datanın (1.91) in kullanım vasıtası ile bunu gösterdiğini farzediyoruz.] ve $\beta = 0$ ise

$$F_\beta = \frac{SS_r(\beta)}{SS_e/(N-2)} \qquad (1.94d)$$

nin $F_{1,N-2}$ olarak dağıldığı gösterilebilir.



Şekil 1.7.1

Yukarıdaki tartışma $m \geq 3$ bağımsız değişkenlerin değerlerinde tekrarlama gözlemlerinin olduğu kabulü ile oluşturulmuştur. Bununla birlikte deney tekrar gözlemleri olmaksızın ilişkilendirilmiştir. Fakat uygunsuzluk hakkındaki gereksinimler hala büyük öneme sahiptir. Bu durumda sıksık yapılan şey $e_i = y_i - \hat{\eta}_i$ kalanlarını incelemektir. Bunun için teorik bir baz vardır. Fakat kalanları incelemek e_i değerlerini değişik şekillerde çizerek de yapılabilir;

i) i ye karşı e_i

ii) x_i *ye karşı* e_i

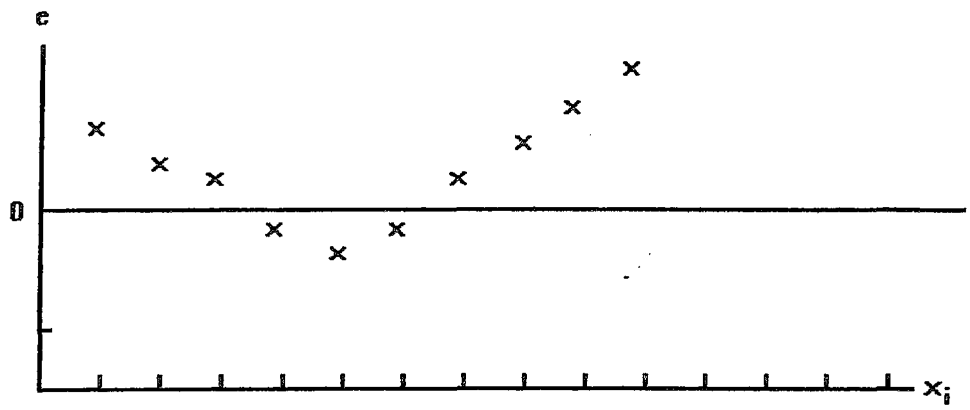
iii) $\hat{\eta}_{x_i}$ *ye karşı* e_i

Eğer uygunsuzluk yok ise ne bu üç çizim ne de Şekil 1.7.1. deki gibi $e = 0$ çizgisinin üstündeki ve altındaki noktaları verecektir. Burada e_i değerleri $e = 0$ civarındaki bir banda düşeceği görülür.

Bununla birlikte eğer kalanlar Şekil 1.7.1. deki gibi bir bant içine düşmez ise gösterilen örnek kabullerden biri veya daha fazlasının geçerli olmadığı gösterilebilir. Örneğin eğer x_i ye karşı çizilen e_i değerleri Şekil 1.7.2. deki davranışı gösterirse (burada y_i ler $\alpha + \beta x_i$ üzerine regresyon yapılır.) o zaman lineerlik kabulünden şüpheleneceğiz ve bir kuadratiği $(\gamma_0 + \gamma_1 x_i + \gamma_2 x_i^2)$ uyduracağız.

1.8. Tahmin

Çok sık olarak yapılan deneysel bir incelemenin amacı, bağımsız x değişkeninin gözlenmiş bir y değerinin sonucuna göre tahmin yapmasını sağlayacak örnek veri toplamaktır. Bu durumda deneysel verinin uydurulacakğı modele güvenmeyi isteriz ve buradaki tartışma için önceki bir bilginin veya uyumsuzluğu değerlendirmelerinin karşılığının x e göre lineer olduğunu farzederiz. Yani



Şekil 1.7.2.

$$E(y|x) = \alpha + \beta x = \eta_x \tag{1.95}$$

dir. (x_i, y_i), $i = 1, \ldots N$ verisinin uydurulan en küçük kareler doğrusuna götürdüğünü biliyoruz ve gerçekte bu uyumsuzluğu test etmek için kullanılabilirdi. Uyumsuzluğun bulunamadığını farzedelim; burada $(\hat{\alpha}, \hat{\beta})$,(α, β) nın en küçük kareler tahminidir. Şimdi y ye göre tahminde bulunalım. Bu $x = x_0$ kullanıldığı zaman karşılığı y in alabileceği değeridir; burada bu değer

$$\hat{\eta}_x = \hat{\alpha} + \hat{\beta} x \tag{1.96}$$

ya götüren veriden bağımsız olarak gözlenecektir. Tahminde bulunulan değeri $\bar{y}$ ile gösterelim (burada y rastgele bir değişken olup bir parametre olmadığından dolayı y harfi üzerinde ^ sembolünü kullanamayız.) Şimdi soru şudur: "$x - x_0$ da gözlenecek olan bağımsız bir gözlenmiş y karşılığının $[(x_i, y_i), i = 1, \ldots, N]$ örnek bilgisine dayandırılan uygun bir $\hat{y}_{x_0}$ tahmini nedir?

Sezgiyle $(x_i, y_i), i = 1, \ldots, N$ datasına dayanan x_0 da gözlenecek bir y tahminininde bulunmak için

$$\tilde{y}_{x_0} = \hat{\alpha} + \hat{\beta} x_{x_0} \tag{1.97}$$

tahmincisini kullanmak makul gözükür. Çünkü (1.97) ile verilen $\hat{y}_{x_0}$ tahmini her iki tahminde η_{x_0} a eşit olduğundan

$$E(\tilde{y}_{x_0} | (x_i, y_i), i = 1, \ldots, N) = E(y | x_0) \tag{1.98}$$

şeklindedir. (1.97) nin seçimini şu şekilde doğrulayacağız. y nin herhangi bir $\tilde{y}$ tahmincisi için $\tilde{y}_{x_0}$ tahmini değerlerinin x_0 daki gerçek y değerlerinden sapmalarının minimum olmasını istiyoruz ve bu sapmaların genişliğinin bir ölçüsü $x = x_0$ için

$$\begin{aligned} &E((\tilde{y}_{x_0} - y) | (x_i, y_i), i = 1, \ldots, N) \\ &\quad = E((\tilde{y}_{x_0} - \eta_{x_0})^2 | (x_i, y_i), i = 1, \ldots, N) + E((y - \eta_{x_0})^2 | x_0) \end{aligned} \tag{1.99}$$

çapraz çarpımı terimi ile ortadan kalkar. Çünkü y ile $\tilde{y}_{x_0}$ arasındaki kovaryans y nin x_0 da gözleneceği ve $(x_i, y_i), i = 1, \ldots, N$ verisinden bağımsız olacağı kabulüne göre sıfırdır. Böylece eğer biz kendimizi y ye göre lineer olan $\tilde{y}$ tahmin fonksiyonlarına kısıtlarsak Gauss teoreminin sonuç 1.2.1.3. undan (1.99) denkeleminin sağ tarafındaki birinci teriminin (1.97) yi kullanarak minimum yapıldığını görürüz ve bu işlemlere devam ettiğimiz takdirde (1.99) denklemi ile verilen tahmincinin kareli hatasını minimum yaparız.Konuyu özetlersek eğer y_i nin bağımsız değişkeni y

$$ID(\alpha + \beta x_i, \sigma^2) \quad , y_i, i = 1, \ldots, N \tag{1.100}$$

ise (1.97) ile verilen y nin $\tilde{y}$ tahmincisi [(1.99) denklemini ve (1.100) denklemini, (1.25a) ifadesi ile birlikte kullanarak] y_i lere göre lineer olan tüm tahminler arasında ortalama kare minimum tahminidir.

$$E\left[(\tilde{y}_{x_0} - y)^2 | (x_i, y_i)\right] = \sigma^2 \left[1 + \frac{1}{N} + \frac{(x_0 - \bar{x})^2}{S(\dot{x}^2)}\right] \tag{1.101}$$

Eğer (1.100) denkleminde $D=N$ alırsak yani denklemin doğrusal olduğunu kabul edersek o zaman hemen

$$t_{N-2}=\frac{\tilde{y}_{x_0}-y}{[SS_e/(N-2)]^{1/2}\left[1+1/N+(x_0-\bar{x})^2/S(\dot{x}^2)\right]^{1/2}} \tag{1.102}$$
$$=\frac{\tilde{y}_{x_0}-y}{S_{y.x}\left[1+1/N+(x_0-\bar{x})^2/S(\dot{x}^2)\right]^{1/2}}$$

yi elde ederiz. Böylece x_0 daki y nin bir uç noktalarına sahip $(1-\gamma)$ aralığı tahmin aralığı

$$\tilde{y}_{x_0}+S_{y.x}\left[1+\frac{1}{N}+\frac{(x_0-\bar{x})^2}{S(\dot{x}^2)}\right]^{1/2} t_{N-2;\gamma/2} \tag{1.103}$$

dır. Burada $\tilde{y}_{x_0}$ nokta tahmini (1.97) ile verilmişti. Eğer x_0 ı değişken olarak alırsak η_x e ait güven aralığını çok daha geniş bir tahmin aralığı olarak elde ederiz.[(1.97a) bakınız.]

1.9 UYGULAMALAR

Örnek 1: Rastgele seçilmiş 30 birey üzerinde farklı yiyecek ile yapılmış bir diyette bir riboflavin çıkarmanın (karışımın mililitresi başına miligramlar) günlük değişim miktarları (x) ile ve bunların sistotik kan yoğunluğu (y) gösterilmek koşulu ile elde edilen veriler tablo 1.9.1 de gösterilmiştir.

Bu tablo değerlerinden yararlınarak basitçe yapılan hesaplamalardan sonra elde edilen sonuçlar şöyledir.

$$N=30;\quad \sum x_i=1354;\quad \sum x_i y_i=199.576;\quad \sum y_i=4276 \tag{1.104}$$
$$\sum x_i^2=67.894;\quad \sum y_i^2=624.260$$

Bu deney kalp uzmanlarının kontrolü altında yapılan muayenelerle ortaya çıkmıştır. x ile y arasındaki ilişki genellikle lineerdir. Gerçekte kullanılan x in değerlerinin sonuçları kaydedilmiştir. Birçok hasta aynı şeylerle diyete devam etmesinden bıkmışlardı. Bu yüzden eksik bir test meydana çıkabilirdi.

$$= \frac{\bar{y}_{x_0} - y}{S_{y.x}\left[1 + 1/N + (x_0 - \bar{x})^2 / S(\dot{x}^2)\right]^{1/2}}$$

yi elde ederiz. Böylece x_0 daki y nin bir uç noktalarına sahip $(1-\gamma)$ aralığı tahmin aralığı

$$\bar{y}_{x_0} + S_{y.x}\left[1 + \frac{1}{N} + \frac{(x_0 - \bar{x})^2}{S(\dot{x}^2)}\right]^{1/2} t_{N-2;\gamma/2} \qquad (1.103)$$

dır. Burada $\bar{y}_{x_0}$ nokta tahmini (1.97) ile verilmişti. Eğer x_0 ı değişken olarak alırsak η_x e ait güven aralığının çok daha geniş bir tahmin aralığı olarak elde ederiz.[(1.97a) bakınız.]

1.9 Uygulamalar

Örnek 1: Rastgele seçilmiş 30 birey üzerinde farklı yiyecek ile yapılmış bir diyette bir riboflavin çıkarmanın (karışımın mililitresi başına miligramlar) günlük değişim miktarları (x) ile ve bunların sistotik kan yoğunluğu (y) gösterilmek koşulu ile elde edilen veriler tablo 1.9.1 de gösterilmiştir.

Bu tablo değerlerinden yararlınarak basitçe yapılan hesaplamalardan sonra elde edilen sonuçlar şöyledir.

$$N = 30; \quad \sum x_i = 1354; \quad \sum x_i y_i = 199.576; \quad \sum y_i = 4276 \qquad (1.104)$$
$$\sum x_i^2 = 67.894; \quad \sum y_i^2 = 624.260$$

Bu deney kalp uzmanlarının kontrolü altında yapılan muayenelerle ortaya çıkmıştır. x ile y arasındaki ilişki genellikle lineerdir. Gerçekte kullanılan x in değerlerinin sonuçları kaydedilmiştir. Birçok hasta aynı şeylerle diyete devam etmesinden bıkmışlardı. Bu yüzden eksik bir test meydana çıkabilirdi.

Hasta Sayısı	x	y	Hasta Sayısı	x	y	Hasta Sayısı	x	y
1	39	144	11	64	162	21	36	136
2	47	220	12	56	150	22	50	142
3	45	138	13	59	140	23	39	120
4	47	145	14	34	110	24	21	120
5	65	162	15	42	128	25	44	160
6	46	142	16	48	130	26	53	158
7	67	170	17	45	135	27	63	144
8	42	124	18	17	114	28	29	130
9	67	158	19	20	116	29	25	125
10	56	154	20	19	124	30	69	175

Uyumsuzluk sorusunun araştırıldığı bu noktayı elde etmel için ilk önce (1.104) deki değerleri kullanarak y i, x üzerinde regresyona tabii tutarak aşağıdaki normal denklemler elde edilir.

$$30\hat{\alpha}+1354\hat{\beta}=4276 \qquad (1.105)$$

$$1354\hat{\alpha}+67.894\hat{\beta}=199.576$$

bunu çözdüğümüzde aşağıdaki sonuçları elde ederiz.

$$\hat{\alpha}=98.71\,, \qquad \hat{\beta}=0.97087 \qquad (1.105a)$$

dir.(1.104) daki değerler ile (1.105a) de elde ettiğimiz değerleri kullanılarak

$$S(\dot{x}^2)=\sum(x_i-\bar{x})^2=\sum x_i^2-\frac{(\sum x_i)^2}{30}=6783.47 \qquad (1.106)$$

değerini elde ederiz. Böylece β dan dolayı kareler toplamının extra regrasyonu

$$SS_r(\beta)=\hat{\beta}^2 S(\dot{x}^2)=6394.02 \qquad (1.106a)$$

dır. Ayrıca buradan

$$S(\dot{y}^2) = \sum y_i^2 - \frac{(\sum y_i)^2}{30} = 14.787.47 \quad (1.107)$$

sonucunuda elde ederiz. Buradan çıkan sonuçlarla

$$SS_e = S(\dot{y}^2) - SS_r(\beta) = 8393.45 \quad (1.108)$$

yı buluruz.

Böylece bir uygunsuzluk testi için elementler -arzu edilen gerçek hata terimi- SS_{μ}. çıkarılmasıyla hesaplanır. Tablo 1.9.2. de görüleceği gibi *x=39,42,45,47,56,67* olmak üzere altı tane x değeri için y in iki gözlem değeri bulunmuştur, gerçekte yapılacak x in diğer tüm değerleri yalnız bir gözlem değeri alınmıştır. Bu nedenle m değeri *24* olur.

n_j gözlemlerinde y_{ij} " koşul j " in kullanılmasıyla bulunur. Bundan sonra kareler toplamı içerisinde şüphesiz sonuca gidilir.

$$\sum_{i=1}^{n_j} (y_{ij} - \bar{y}_{.j})^2, \qquad \bar{y}_{.i} = n_j^{-1} \sum_{i=1}^{n_j} y_{ij} \quad (1.109)$$

Tablo 1.9.2

x	39	42	45	47	56	67
y	144;120	124;128	138;135	220;145	150;154	170;158

$n_j = 2$ olduğu zaman bizim yapacağımız

$$\sum_{i=1}^{2} (y_{ij} - \bar{y}_{.i})^2 = \frac{(y_{1j} - y_{2j})^2}{30} \quad (1.109a)$$

dır. Buradan hızlıca görülebileceği gibi verilerin kümesi ile gerçek hata terimi SS_{μ}. yi aşağıdaki gibi elde ederiz

$$SS_{H'} = \left[(144-120)^2+\cdots+(170-158)^2\right]/2 = 3193 \tag{1.109b}$$

Şimdi biz burada Tablo1.9.3. ün yardımıyla varyansın uygun analizlerini yapma konumundayız.

" uygunsuzluk" un kareleri toplamı şüphesiz çıkarılma ile bulunur. Böylece bu

$$SS_T = SS_L - SS_r(\beta) - SS_{H'} \tag{1.110}$$

dır. SS_T için girilen değerler sonucu bulunan sonuçlar Tablo 1.9.3. de verilmiştir. Buradan

$$F_L = \frac{5200.45/22}{3193/6} = \frac{236.38}{532.17} = 0.44 \qquad , F_{22.6;0.05} = 3.855 \tag{1.111}$$

hipotezini elde ederiz.Böylece bu hipotezi kabul ederiz. Çünkü $\alpha+\beta x$ doğru modeline uyarlandığında "uygunsuzluk" bulunamaz, yani " uygunsuzluk " olmaz.

Tablo 1.9.3 : Tablo 1.9.1 deki dataların varyans analizi

Kaynak	Serbestlik Derecesi	Kareler Toplamı
β	1	$SS_r(\beta) = 6394.02$
Uyumsuzluk	22	$SS_L = 5200.45 \quad , F_L = 0.44$
Has Hata	6	$SS_{H'} = 3193$
Toplam	29	$SS_T = S(\dot{y}^2)$ $= 14.787.47$

β ile ilgili bir test yapalım. Örneğin %5 seviyesinde $\beta = 0$ iddiasını kullandığımızda açıkça görülebileceği gibi hipotez reddedilir. , buradan (1.94d) ü kullanarak F_β bulunuruz.

$$F_\beta = \frac{6394.02}{8393.45/28} \qquad F_{1,28;0.05} = 4.196 \tag{1.112}$$

Eğer x *in* y ile ilişkisinin gidişi (lineer ilişkisi) tam ise $\beta = 1$ iddiasını test edelim. Buna göre testimizi

$$F_\beta^* = \frac{(\hat{\beta} - \beta_0)^2 S(\dot{x}^2)}{SS_e / 28} \tag{1.113}$$

dağılımı ile yaparız. Bu durum için F_β^* nin gözlenmiş değeri

$$F_\beta^* = \frac{(0.971 - 1)^2 (6783.47)}{299.77} \tag{1.113a}$$

dir ve gözlenmiş $F_\beta^* \leq F_{1,28;0.05} = 4.196$ olduğundan dolayı $\beta = 1$ için hipotezi kabul ederiz.

Şimdi lineer ilişkiyi gösterelim. Sonuç çıkarmak için herzamanki değerleri kullanacağız. Örneğin, β için %95 güven aralığı olan $I_{0.95}(\beta)$ nin bulunması

$$I_{0.95}(\beta) = \left[\hat{\beta} + (SS_e/28)^{1/2} \left(1/S(\dot{x}^2)\right)^{1/2} t_{28,0.025}\right] \tag{1.114}$$

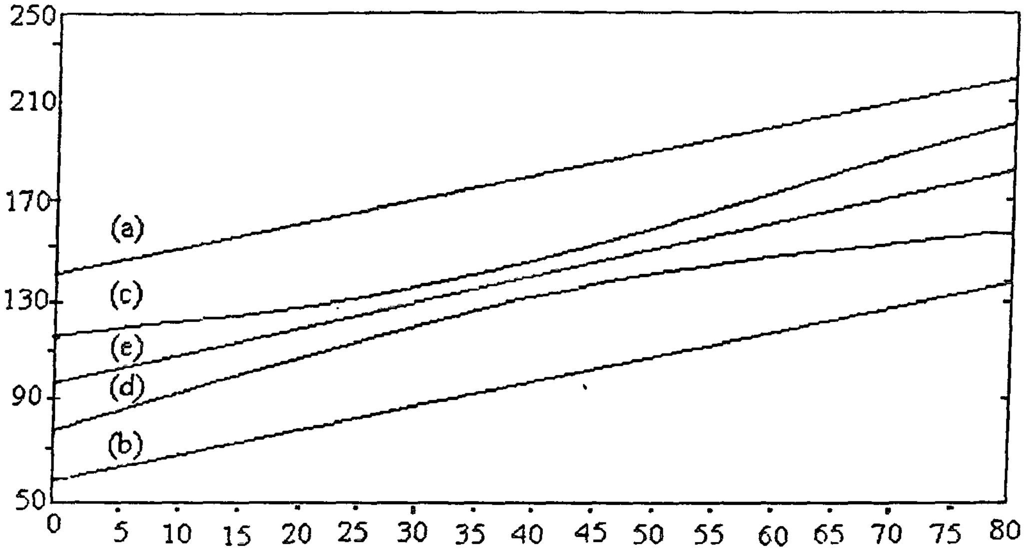
$$= [0.971 + 0.4305]$$

şeklindedir. Buradan

$$I_{0.95}(\beta) = [0.5405, 1.4015] \tag{1.114a}$$

elde ederiz. Yapılacak yaklaşımlarla η için %95 güven aralığında, limitlerden bulunması kolayca görülebilir.

$$\hat{\alpha} + \hat{\beta}x + \left[\frac{1}{30} + \frac{(x - 45.13)^2}{6783.5}\right]^{1/2} \left(\frac{SS_e}{28}\right)^{1/2} t_{28,0.025} \tag{1.115}$$



ŞEKİL 1.9.1

ve x in değişimi için, bu aralık Şekil 1.9.1. de grafiksel olarak gösterilmiştir. İlerideki y gözlemi için önceden elde edilen 0.95 seviyesindeki tahmin aralığı, x in karşısında bağımsızca gözlenir. Ayrıca Şekil 1.9.1. dè grafikleştirilmiştir ve limitler tarafindan genelleştirilir.

$$\hat{\alpha}+\hat{\beta}x+\left[1+\frac{1}{30}+\frac{(x-45.13)^2}{6783.5}\right]^{1/2}\left(\frac{SS_e}{28}\right)^{1/2} t_{28,0.025} \tag{1.115a}$$

Sonuçta, (α,β) için %95 lik güven bölgesi meydana getirebiliriz. Bu $C_{0.95}$ şöyle bulunabilir, (α,β) düzlemi içerisinde noktalar kümesi tarafindan verilir.

$$C_{0.95}=\left\{(\alpha,\beta)\left|\begin{array}{c}30(\alpha-98.71)^2+2(1354)(\alpha-98.7)(\beta-0.971)\\ +67.894(\beta-0.971)^2\leq 2(SS_e/28)F_{2,28,0.05}\end{array}\right.\right\} \tag{1.116}$$

Bu (1.116) ifadesinde işaret edilen denkliğin kullanımıyla bulunan denklemle elde edilen aralık bölgesi elipsdir . Bu elipsin merkezi (α,β) dır.

Örnek 2: Büyük bir okuldaki öğrenci kayıt burosunda çalışan yetkili bir yönetici bir sonraki yılda okula kayıt olacakların sayısı içinde bulunan öğretim yılı boyunca bilgi almak için başvuranların sayısını esas alarak tahmin edebileceğini düşünmektedir. Geçmiş 9 yıllık veriler aşağıda verilmiştir.

Başvuranların Sayısı (x/1000)	2.5	2.7	2.9	2.6	2.8	3	2.9	3.1	3.3
Kayıtlar (y/100)	18.3	18.5	19.4	19.6	19.7	20.1	20.6	20.7	21

$\hat{\alpha}$ ve $\hat{\beta}$ değerinin bulunması

Her x değeri için $\hat{y}_i$ değerinin bulunması

$y_i - \hat{y}_i$ farklarının hesaplanması

$\hat{\sigma}^2$ nin hesaplanması

$\alpha = 0.10$ alarak lineer modelin eğiminin sıfırdan önemli derecede farklı olup olmadığı hipotezinin incelenmesi.

α, β ve $\mu_{y \cdot x=3.5}$ için % 90 lık güven aralığının bulunması işlemleri yapılacaktır.

Ayrıca bilgi almak için başvuranların sayısı 3500 olduğuna göre kayıtlar üzerinde gelecekteki gözlemler için % 90 lık güven aralığını bulalım.

Bizim için gerekli olacak veriler aşağıdaki gibi bulunmuştur.

x_i	y_i	$y_i x_i$	x_i^2	$(x_i - \bar{x})^2$	$\hat{y}_i$	$y_i - \hat{y}_i$	$(y_i - \hat{y}_i)^2$
2.5	18.3	45.75	6.25	0.137	18.6	-0.3	0.09
2.7	18.5	49.45	7.29	0.029	19.2	-0.7	0.49
2.9	19.4	56.26	8.41	0.001	19.9	-0.5	0.25
2.6	19.6	50.96	6.76	0.073	18.9	0.7	0.49
2.8	19.7	55.16	7.84	0.005	19.6	0.1	0.01
3.0	20.1	60.30	9.00	0.017	20.2	-0.1	0.01
2.9	20.6	59.74	8.41	0.001	19.9	0.7	0.49
3.1	20.7	64.17	9.61	0.053	20.5	0.2	0.04
3.3	21.0	69.30	10.89	0.185	21.2	-0.2	0.04
Toplam							
25.8	177.9	511.59	74.46	0.5	178	-0.1	1.91

$\hat{\alpha}$ ve $\hat{\beta}$ değerlerini

$$N\hat{\alpha} + \hat{\beta}\sum x_i = \sum y_i$$

$$\hat{\alpha}\sum x_i + \hat{\beta}\sum x_i^2 = \sum y_i x$$

normal denkleminde yerine koyarsak

$$9\hat{\alpha}+2.87\hat{\beta}=19.77$$
$$2.87\hat{\alpha}+74.44\hat{\beta}=511.6$$

dan $\hat{\alpha}=10.536$ ve $\hat{\beta}=3.22$ elde edilir. O halde regresyon denklemi

$$\hat{y}=10.536+3.22x$$

dir. Burada x_i değerlerini vererek $\hat{y}_i$ değerlerini tablodaki gibi buluruz. Yine tabloda $y_i-\hat{y}_i$ nın değerleri gösterilmiştir. Şimdi elimizdeki değerlerle σ^2 yi tahmin edebiliriz. Bu ise

$$\sigma^2=\frac{\sum(y_i-\hat{y}_i)^2}{N-2}$$

denklemi ile elde edilir. Yani

$$\sigma^2=\frac{1.91}{7}=0.273$$

tir. Burada standart sapma ise $\sigma=0.522$dir.

Şimdi de

$$H:\beta=0$$
$$H:\beta\neq 0$$

Hipotezlerinin doğruluğunu $\alpha=0.10$ da araştıralım.

$$t_{\alpha/2;N-2}=t_{0.05;7}=-1.895$$
$$t_{1-\alpha/2;N-1}=t_{0.95;7}=1.95$$

dir. t-istatistiğinin değeri

$$t=\frac{\hat{\beta}-\beta}{\sqrt{\frac{\sigma^2}{S(\dot{x}^2)}}}=\frac{\hat{\beta}\sqrt{S(\dot{x}^2)}}{\sigma}=\frac{3.22}{0.7434}=4.33$$

dir hesaplanan t değeri -1.895 ile 1.95 arasında olmadığından H_0 hipotezi red edilir. Yani β nın değeri sıfırdan farklıdır.

Şimdi de α için güven aralığını hesaplayalım.

$$(\hat{\alpha}-t_{1-\alpha/2}\sigma\frac{\sqrt{\sum x_i^2}}{\sqrt{NS(\dot{x}^2)}},\hat{\alpha}+t_{1-\alpha/2}\sigma\frac{\sqrt{\sum x_i^2}}{\sqrt{NS(\dot{x}^2)}})$$

dır. O halde

$$P[10.536-1.895/0.214<\alpha<10.536+1.895/0.214]=0.90$$

dır.Böylece α nın %90 lık güven aralığı

$$10.130<\alpha<10.942$$

dır.

Şimdi de β için güven aralığını hesaplayalım.

$$(\hat{\beta}-t_{1-\alpha/2}\frac{\sigma}{\sqrt{S(\dot{x}^2)}},\hat{\beta}+t_{1-\alpha/2}\frac{\sigma}{\sqrt{S(\dot{x}^2)}})$$

dır. O halde

$$P[3.22-1.895/0.744<\beta<3.22+1.895/0.744]=0.90$$

dir. Böylece β için %90 lık Güven aralığı

$$1.810<\beta<4.630$$

dır. Şimdi $E(y|x=3.5)$ 'nun tahmini bulalım.

$$\hat{y}_{x=3.5}=E(y|x=3.5)=10.536+3.22(3.5)=21.806$$

dir. Şimdi $\hat{y}_{x_0}$ nin varyansını hesaplayalım.

$$\sigma^2_{\hat{y}_{x_0}}=\sigma^2(1/N+\frac{(x_0-\bar{x})^2}{S(\dot{x}^2)})$$

dır. O halde

$$\sigma^2_{\hat{y}_{x_0}}=0.277(1/9+\frac{(3.5-2.87)^2}{0.5})=0.251$$

dir. Böylece beklenen değerin güven aralığı

$$(21.806-1.895/0.5;21.806+1.895/0.5)$$

dir.

x=3500 olduğuna göre gelecekteki kayıtların tahmini

$$\mu=10.536+3.22(3.5)=21.806$$

dır. Varyansının tahmini ise

$$\sigma^2(1+1/N+\frac{(x_0-\bar{x})^2}{S(\dot{x}^2)})=0.277(1+1/9+\frac{(3.5-2.87)^2}{0.5})=0.528$$

dır. Standart sapması =0,726 dır. Güven aralığı ise

$$(21.806-1.895(0.726) < y(x=3.5) < 21.806+1.895(0.726))$$
$$(20.430 < y < 23.182)$$

dır.

BÖLÜM II
GENELLEŞTİRİLMİŞ TERS MATRİSLER

2. Genelleştirilmiş ters matrise giriş

$AX=Y$ denklem sisteminin bir çözümü olabilmesi için A matrisinin nxn tekil olmayan bir matris olması lazımdır. O zaman bu denklemin çözümü tektir ve $X = A^{-1}Y$ ile verilir. Eğer A matrisi kare matris değilse bir çok çözümü olabilir. A matrisinin kare matris omadığı durumlarda çözümü araştırmak için genelleştirilmiş ters matris yöntemini kullanacağız.

2.1.1. Genelleştirilmiş Ters Matrisin Tanımı, Özellikleri

Genelleştirilmiş ters matris gerçekte yeni bir kavram değildir. Bu konuda ilk yapıtın Moore'ye ait olduğu söylenir. 1920 de yayınlanan bu makaleden sonra 1955 e kadar bu konuda önemli sayılabilecek araştırmalar görülmez. Bu tarihte Penrose'un bir makalesi yayınlanır. Söz konusu yapıtta ister kare ister dikdötrgen olsun bir A matrisinin aşağıdaki özellikleri yerine getiren bir tek genelleştirilmiş ters matrisi bulunduğunu kanıtlamıştır. A matrisinin genelleştirilmiş ters matrisini A^- ile göstereceğiz. Matris cebirinde genelleştirilmiş ters matrisi " g-inverse " olarak adlandırılır.

2.1.1. g- inverse tanımı

1) $A A^- A = A$
2) $A^- A A^- = A^-$
3) $(A A^-)' = A A^-$
4) $(A^- A)' = A^- A$

Bu özellikleri sağlayan matrise inverse matris denir. Burada A nın tekil olmaması veya kare matris olmaması inverse nin bulunmasını etkilemez ve bu inverse matris $AX=Y$ denklem sistemi için bir çözüm oluşturur ve bu çözüm tektir.

Bu çalışmada, Hausman-Searle denklemini temel alarak

$$AGA=A \tag{2.1}$$

koşulunu (Penrose'nin 1. koşulu) sağlayan herhangi bir G matrisine A matrisinin genelleştirilmişidir diyeceğiz.

İstatistik uygulamalarda lineer modellerin çözümlenmesinde tekbir g-inverse in olmaması herhangi bir sakınca oluşturmamış hangi G bulunursa bulunsun bazı sonuçların değişmediği ortaya çıkmış ve önemli bir yaklaşıma yol açmıştır.

2.1.2.g-inverse'in Özellikleri

$H=GA$ diyelim ($G=A^{-1}$ için $H=I$ olacağı aşikardır.)

1) $r(H)=r(A)=r$ dır.
2) $r(G)\geq r(A)$ dır.
3) $H^2=H$ dır.
4) G,A nın herhangi bir g-inverse ise G',A' nın herhangi bir g-inverse dir: $A'G'A'=A'$

2.1.3. G nin hesabı

2.1.3.1. Genel Yöntem

Herhangi bir A matrisini $PAQ \doteq \Delta$ gibi köşegen biçime indirgeyebileceğimizi biliyoruz.

$$PAQ=\Delta=\begin{bmatrix} D_r & 0 \\ 0 & 0 \end{bmatrix}$$

Burada PAQ nun mertebesi A ile aynı r(A)=r, D_r ,r elemanları sıfır olmayan bir köşegen matrisi, sıfırlar ise uygun mertebeden sıfır matrisleri, P ile Q satır ve sütunlara ilişkin elemanter işlemleri gösterirler. Burada Δ nın inversini aşağıdaki gibi tanımlayalım.

$$\Delta^{-1} = \begin{bmatrix} D_r^{-1} & 0 \\ 0 & 0 \end{bmatrix}$$

$$G = Q\Delta^{-}P \tag{2.2}$$

Bu şekilde tanımlanan bir matris A nın bir genelleştirilmiş tersidir. Başka bir deyişle G, $AGA=A$ koşulunu gerçekleştirir. Gerçektende $\Delta\Delta^{-} = \Delta^{-}\Delta$ yazabiliriz. Dolayısıyla

$$\Delta\Delta^{-}\Delta = \Delta, \qquad \Delta^{-}\Delta\Delta^{-} = \Delta^{-} \tag{2.3}$$

olur. Öte yandan P ile Q elemanter işlemcilerin çarpımları olduklarından tekil değillerdir. Klasik anlamda tersleri olacağından $PAQ = \Delta$ eşitligine dayanarak

$$A = P^{-1}\Delta Q^{-1} \tag{2.4}$$

yazabiliriz. A nın bu değeri ile G nin $G = Q\Delta^{-}P$ değerini AGA da yerine koyarsak

$$AGA = (P^{-1}\Delta Q^{-1})(Q\Delta^{-1}P)(P^{-1}\Delta Q^{-1}) \tag{2.5}$$

olur. (2.3) ve (2.4) denklemleriyle

$$AGA = P^{-1}\Delta\Delta^{-1}\Delta Q^{-1} = P^{-1}\Delta Q^{-1} = A$$

sonucuna varırız. Böylece A matrisinin bir g-inverse ni elde etmiş oluruz.

Diğer yandan P ve Q bulunmadan da A nın satır ve sütunlarında yapılacak işlemler ile Δ doğrudan elde edilebilir. Fakat çoğu kez Q ya ve P ye gereksinim olduğundan bu ikisi bulunmalıdır.

Yukarıdaki yöntemle çözümün iki ilginç özelliğini belirtebiliriz.

1) $GAG=G$

2) A,(n,q) boyutlu bir matris $H=GA$ olduğuna göre $(H-I)$nın rankı q-r dir.r$(H-I)$ = q-r

Bu iki özellik ile, kısım 2.1.2. deki özellikler $AX=Y$ sisteminin çözümünde önemli rol oynar.

2.1.3.2. Bir Başka Yöntem

g-inverse ler başka yollardan da bulunabilir. Bu bir A matrisini düzgün kare matrislere ayırmak ile mümkündür. Yani bir A matrisi,

$$A=\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

biçiminde öyle kısımlara ayırabiliriz ki, A_{11} rxr boyutlu A matrisi içinde tekil olmayan en geniş alt matris olur. r nin maksimum olabilmesi için A nın satır ve sütunlarında değişiklik yapabiliriz. Ayrıca A_{11} in tekil olması da gerekmez.

$$G=\begin{bmatrix} A_{11}^{-1} & 0 \\ 0 & 0 \end{bmatrix} \tag{2.6}$$

matrisi A matrisinin bir g-inverse dir. AG ile GA hesaplandığında $AGA=A$, $GAG=G$ olduğu görülür. Burada eğer A matrisi simetrik ise çözüm daha kolay bulunabilir. Bunu bir prosedür ile gösterecek olursak aşağıdaki işlemleri takip edebiliriz.

a) Rankı r ve simetrik olan bir A matrisinde r. mertebeden tekil olmayan bir minör bulunur. Buna M diyelim.
b) M^{-1} oluşturulur.
c) M^{-1} in transpozesi $(M^{-1})'$ oluşturulur.
d) A matrisi içinde M yi oluşturan her elemanın bulunduğu yerde, $(M^{-1})'$ de karşılık gelen elemanlara yer verilir. Öteki elemanların yerine sıfır konur.
e) Bulunan matris, A nın genelleştirilmiş ters matrisi G dir.

2.1.4. Bir Uygulama

Rankı tam olmayan ($|A|=0$) aşağıdaki A matrisinin Genelleştirilmiş Tersini bulalım

$$A=\begin{bmatrix} 1 & 2 & 1 \\ 2 & 3 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

$$A = \begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & -2 & -1 \\ 0 & -1 & -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{array}$$

3. sutun yerine 3. + 2. sutun x (-1) işlemini yaparsak

$$A = \begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & -2 & 1 \\ 0 & -1 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{array}$$

demek ki $Q = \begin{bmatrix} 1 & -2 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}$

dir. Şimdi Δ yı elde edelim.

$$PA = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 1 & -1 & 1 \end{bmatrix} . \begin{bmatrix} 1 & 2 & 1 \\ 2 & 3 & 1 \\ 1 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 1 \\ 0 & -1 & -1 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\Delta = PAQ = \begin{bmatrix} 1 & 2 & 1 \\ 0 & -1 & -1 \\ 0 & 0 & 0 \end{bmatrix} . \begin{bmatrix} 1 & -2 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\Delta^{-1} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$G = Q\Delta^{-1}P$$

den

$$G = \begin{bmatrix} 1 & -2 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} . \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix} . \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 1 & -1 & 1 \end{bmatrix}$$

$$G=\begin{bmatrix}1 & -4 & 0\\ 0 & -1 & 0\\ 0 & 0 & 0\end{bmatrix}$$

dir

2.2. LİNEER DENKLEMLERİN ÇÖZÜMÜ

Genelleştirilmiş ters matrisleri tanımladığımıza göre burada g-inverse ile genel çözümün nasıl elde edilebileceğini açıklayacağız.

2.2.1. Genel Çözüm

G yi bildiğimize göre genel çözüme geçmeden önce konuyla ilgili teoremleri verelim

Teorem 2.1. $AX=Y$ denklem sisteminin bir çözümü olması için gerek ve yeter koşul herhangi bir G için $AGY=Y$ olmasıdır.

Gerçekten de $AX=Y$ nin X^* ile gösterebileceğimiz bir çözümü olduğunu kabul edelim. $AX^*=Y$ olacaktır. G herhangi bir g-inverse olduğuna göre $(AGA\text{-}A)$,$AX^*=Y$ nın her iki yanında AG ile çarparsak

$$AGAX^*=AGY$$
$$AX^*=AGY=Y$$

elde ederiz. Demek ki X^* bir çözüm ise, $AGY=Y$ dir. Böylece ispat tamamlanmış olur. Burada dikkat edilmesi gereken husus $X^*=GY$ olmasıdır.

Teorem 2.2. Ancak ve ancak $AGA=A$ ise $AX=Y$ fonksiyonun çözümü $X^*=GY$ dir.Gerçekten de GY bir çözüm ise $AGY=Y$

$$AGAX=Y$$
$$AX=Y$$

olur. Demek ki G bir çözümdür. Diğer taraftan $AGA=A$ ise $AGAX^*=AX^*$ yazılabilir. Ancak $AX^*=Y$ den $AGY=Y$ yi elde edebiliriz. Bu ise bize GY nin bir çözüm olduğunu gösterir. Böylece ispat tamamlanır.

Teorem 2.3. $AGA=A$ ve $H=GA$ ise $AX=Y$ denklem sisteminin bir çözümü

$$X_0 = GY + (H - I)Z \tag{2.7}$$

dır. Burada Z, q. mertebeden herhangi bir vektördür.

Gerçekten de (2.7) denkleminin her iki yanını A ile çarparsak

$$AX_0 = AGY + A(H - I)Z$$

ve $AH=AGA=A$ olduğundan (daha önceden gösterildi.), $AX_0 = AGY$ yi elde ederiz. Ancak bu Teorem 2.2 den GY bir çözüm olduğuna göre $AX_0 = AX^*$ özdeşliğine varırız.

Demek ki her (q,1) boyutlu Z vektörü için X_0 vektörü bir çözümdür. g-inverse matrisi tek olmayacağı gibi Z te istenildiği gibi seçilebilir. Burada denklemin sonsuz çözümü olacağı ortadadır.

2.2.2. Lineer Olarak Bağımsız çözüm Sayısı

Denklem sisteminin sonsuz çözümünün olduğunu ayrıca G nin tek olmadığını biliyoruz. önemli olan nokta bunun ne kadarının lineer olarak bağımsız olduğunu ve aralarında ne gibi bir ilişki bulunduğunu bilmektir. Lineer denkleme ilişkin lineer vektör sayısı, her X çözümü q. mertebeden olacağına göre q dan fazla olamaz. Bunu aşağıdaki teoremle dile getirelim.

Teorem 2.4. A, rankı r olan q sutunlu bir matris ise Y de bir sıfır vektörü değilse AX=Y sisteminin Lineer olarak bağımsız çözüm sayısı q-r+1 dir. Bunun ispatı Merih İpek,1980 in kitabında gösterilmiştir.

Bazı bileşimler hangi X_0 çözümü kullanılırsa kullanılsın değişmezler. Lineer modellerin çözümlenmesinde önemli bir rol oynadığı için bu özelliğin üzerinde biraz duralım.

Rao 'dan kayanklanan bir teoremle şöyle dile getirilir.

Teorem 2.5. $AGA=A$, $H=GA$ olmak üzere $AX=Y$ denklem sisteminin bir X_0 çözümünün öğelerinin $k'X_0$ gibi bir lineer birleşimi, ancak ve ancak $k'H=k'$ ise hangi çözüm kullanılırsa kullanılsın çözüm değişmez.

Yani hangi Z vektörü seçilirse seçilsin $k'X_0$, Z den bağımsızdır. Bu ise farklı Z vektörleri için elde edilen X_0 çözümlerinin k' ile çarpımının aynı sabite eşit olduğunu gösterir.

2.3. Rankı Tam Olmayan Lineer Modeller

Burada ele alacağımız konu rankı tam olmayan lineer modeldir. Bunun için önce normal lineer denklemimizi ele alarak bunun üzerinde tartışmaya devam edelim.

$$Y = X\beta + \varepsilon$$

denkleminde Y, nx1 boyutlu gözlem matrisi, X, nxq boyutlu katsayılar matrisi, β, qx1 boyutlu bilinmeyen paremetreler matrisi ε nx1 boyutlu gözlenemeyen hatalar matrisidir.

Burada dikkat edilirse tek farklılık X matrisinin özel yapısından doğar. Çoğu istatistikçiler ele aldığımız modele, regresyon modelinin özel bir hali olarak bakarlar.

Denklemdeki bilinmeyen $\hat{\beta}$ parametre matrisinin tahmini için " en küçük kareler " yöntemine başvurulur.Bu

$$(X'X)\hat{\beta}=X'Y \qquad (2.9)$$

normal denklemlerini verir. Matrisin eğer rankı tam ise çözüm

$$\hat{\beta}=(X'X)^{-1}X'Y$$

olacak şekilde tekdir. Oysa yukarıdaki denklemde X in rankının tam olmadığını ve ($r(X)=r(X'X)$) matris özelliklerinden $X'X$ inde rankının tam olmadığını biliyoruz.

Buna göre (2.8) denkleminin $\hat{\beta}$ vektörleri yoktur yada sonsuz sayıda $\hat{\beta}$ vektörleri vardır. $\hat{\beta}$ vektörlerinin bulunamıyacağı ihtimalini göz önünde bulundurmuyoruz. Çünkü matrisin genelleştirilmiş g-inverse alındığında bu olasılık ortadan kalkar.

Öyleyse $\hat{\beta}$ için sonsuz çözüm söz konusudur. İlk bakışta bu çok kötü bir sonuçtur. Çünkü farklı kişiler, aynı modeli kullansalar da, değişik $\hat{\beta}$ vektörü bulabilirler. Dolayısıyla β için tek bir tahmin vektöründen söz edilemez. Bazı yazarlar bu durumda çözümü tek yapma yoluna gitmişlerdir. Uygulanan yöntem şöyledir.

$$Y = \beta X + \varepsilon$$

Denklem sistemini

$$Y = \beta^* X^* + \varepsilon$$

biçimine sokulur. Burada β^* ortalama değeri β vektöründen çıkartılan q-1x1 boyutlu matris X^* ise 1. sütunun ortadan kaldırılmasıyla elde edilen (nxq-1) boyutlu bir kat sayılar matrisidir. Böylece β^* in rankı tam olacağından çözüm de tek olacaktır. X^* matrisinin rankı tam olaağından $X'X$ in de rankı tam olacaktır.

Bu yönteme "dönüştürme ve yeniden parametreleme" adı verilir.

2.3.1. Rankı Tam Olmayan Modellerin q-inverse İle Çözümleri

Lineer denklemlerin genelleştirilmiş ters matrislerle nasıl çözüldüğünü daha önce gösterdik. Benzer şekilde Teorem 2.2 ye dayanarak (2.9) denklem sisteminin bir çözümünü

$$\beta_0 = GX'Y \qquad (2.10)$$

olacağını söyleyebiliriz. Burada g-inverse nin tanımı olarak

$$X'XGX'X = X'X \qquad (2.11)$$

koşulunu sağlayan, ayrıca Penrosenin 2 koşulunu da gerçekleştirir. Yani

$$GX'XG=G$$

dir. Demekki G matrisi $X'X$ matrisinin bir genelleştirilmiş ters matrisidir. Ayrıca G tek olmayacaktır. Bundan dolayı $\hat{\beta}$ tahmincisi yerine β_0 simgesini kullanacağız. Çünkü çözümlerden birini göstermektedir.

2.3.1.1. Simetrik Matrislerin Genelleştirilmiş Terslerinin Özellikleri.

Simetrik $X'X$ matrisinin birçok özellikleri vardır. Bunu bir teoremle belirtelim.
Teorem.3.1. Eğer G, $X'X$ in $X'XGX'X=X'X$ koşulunu sağlayan bir g-inverse ve $H=GX'X$ ise

a-) G, $X'X$ in bir g-inverse idir.
b-) $X'GX'X=X$ yada $XH=X$ dır.
c-) G ne olursa olsun XGX' değişmez.
d-) G simetrik olsun olmasın XGX' simetriktir.
e-) $(XGX')^2 = XGX'$
f-) $(I - XGX')^2 = (I - XGX')$

2.3.1.2. Genel Çözüm:

$X'XGX'X=X'X$ ve $H=GX'X$ ise teorem 2.2. ye göre normal denklemlerin genel çözümü

$$\beta_0 = GX'Y + (H - I)Z \tag{2.12}$$

eşitliği ile verilir. Z, (qx1) boyutlu herhangi bir vektör olacağına göre β_0, bir çözüm kümesi sağlar.

Varyans analizinde genellikle faktörlerin etkisinin olup olmadığı aranır. Aynı şekilde lineer modellerde faktörlerin etkisi aranırken σ^2 nın tahminine gerek duyulur. Burada β_0 ın varyansı

$$V(\beta_0) = G\sigma^2$$

dir.Bunun ispatı Merih İpek in kitabında yapılmıştır.

Gerek varyans gerekse regresyon analizinde önemli bir yeri olan temel özellik şöyledir.

$$SS_{W} = SS_{L} + SS_{e}$$

Daha açıkça yazılacak olursa

$$\sum\sum(y_{ij} - \bar{y})^2 = \sum\sum(y_{ij} - \hat{y}_{ij})^2 + \sum\sum(\hat{y}_{ij} - \bar{y})^2$$

dir.

Şimdi β_0 bir çözüm olduğuna göre genel denklemi bu çözüme göre

$$Y = X\beta_0 + \varepsilon$$

şeklinde yazabiliriz. Buradan $\hat{Y} = X\beta_0$ olmak üzere kalanlar

$$Y - \hat{Y} = \varepsilon$$

ya da

$$Y - \hat{Y} = Y - X\beta_0$$

dır. (2.11) denklemi ise $XH{-}X$ özelliğinden yararlanılarak

$$\begin{aligned} Y - \hat{Y} &= Y - X(GX'Y + (H - I)Z) \\ &= Y - XGX'Y - X(H - I)Z \\ &= Y - XGX'Y \\ &= (I - XGX')Y \end{aligned}$$

sonucuna varırız. Buna dayanarak SS_{W} yi bulalım. Matrisyel yazılışla

$$\begin{aligned} SS_{W} &= (Y - \hat{Y})'(Y - \hat{Y}) \\ &= ((I - XGX')Y)'((I - XGX')Y) \end{aligned}$$

ancak teorem 3.1. deki (d)) ve (f) özelliğinden

$$SS_{W} = Y'(I - XGX')Y \tag{2.13}$$

dir. Bu önemli bir sonuçtur. Çünkü teorem 3.1. deki (d) özelliğinden hangi ters matris kullanılırsa kullanılsın SS_W değişmez.

Öte yandan (2.13) ifadesini kullanarak

$$SS_W = Y'Y - XGX'Y$$

dir. Burada $\beta_0 = GX'Y$ olduğundan

dır. $Y'X\beta_0 = X'Y\beta_0$ yazılabileceğinden

$SS_W = Y'Y - \beta_0 X'Y$

dir. Bu regresyon modellerinde varılan sonucun kendisidir. Bu demektir ki hangi çözüm kullanılırsa kullanılsın kalanların karelerinin toplamı değişmeyecektir.

σ^2 tahmini ise

$$\sigma^2 = \frac{SS_W}{n-r}$$

dir.Yalnız burada σ^2, X in rankına ($r=r(X)$) bağlıdır.

Aynı şekilde SS_e ve SS_L ninde herhangi bir çözümü için değişmediğini görebiliriz.

Demekki; β nın bileşenlerinin sapmasız kestirimi yapılamamakla birlikte β_0 ne olursa olsun bazı değişmez sonuçlar alınabiliyor.

2.4. Bir Uygulama

Her hangi konuda üç ayrı eğitim yöntemine başvuruluyor. Bunların öğrencilerin başarısında değişik etkilerinin olup olmadığını araştırmak istiyelim. Rastgele seçilen 9 öğrenciden üçüne birinci yöntemi ,dördüne ikinci yöntemi , diğer ikisinede üçüncü yöntemi uygulayalım. Sonuçlar ise aşağınaki gibidir.

Biz burada sadece β_0 ının bir çözümü ile ilgileneceğiz. Ayrıca varyansı ve kalanların kareler toplamının tahminini bulacağız. Eğitim yöntemlerinin öğrencilerin başarısına etkisi olup olmadığını şimdilik araştırmayacağız. Bu söylenen işlemleri ise genelleştirilmiş ters matrislerle çözeceğiz.

EĞİTİM YÖNTEMLERİ

	I	II	III
	9	14	5
PUANLAR	13	16	7
	11	14	
		8	

Bu deneye ilişkin modeli

$$y_{ij} = \mu + \alpha_j + \varepsilon_{ij}$$

biçiminde yazabiliriz. Bu bir lineer modeldir. Burada y_{ij} üç bileşenin toplamı, μ ana kütle ortalaması, α_j faktörlerin j. katagorisi, ε_{ij}, oluşan hatalardır. Yukarıdaki modeli düzenleyerek

$$Y = \beta X + \varepsilon$$

yi elde ederiz. Burada da Y gözlemler vektörü, X katsayılar matrisi, β, bilinmeyen parametreler vektörü, ε ise gözlenemeyen hatalar vektörüdür.

Şimdi denklem sistemimizi yazalım.

$$\begin{bmatrix} 9 \\ 13 \\ 11 \\ 14 \\ 16 \\ 14 \\ 8 \\ 5 \\ 7 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix} . \begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} + \begin{bmatrix} \varepsilon_{11} \\ \varepsilon_{21} \\ \varepsilon_{31} \\ \varepsilon_{12} \\ \varepsilon_{22} \\ \varepsilon_{32} \\ \varepsilon_{13} \\ \varepsilon_{23} \\ \varepsilon_{33} \end{bmatrix}$$

$$(X'X)\hat{\beta} = X'Y$$

den

$$\hat{\beta} = (X'X)^{-1} X'Y$$

yi hesaplayabileceğimizi biliyoruz. Buna göre önce $X'X$ i ve $X'Y$ yi hesaplayalım. Sonra da $(X'X)^{-1}$ i hesaplarız.

$$X'X = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} . \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 9 & 3 & 4 & 2 \\ 3 & 3 & 0 & 0 \\ 4 & 0 & 4 & 0 \\ 2 & 0 & 0 & 2 \end{bmatrix}$$

$$X'Y=\begin{bmatrix}1&1&1&1&1&1&1&1&1\\1&1&1&0&0&0&0&0&0\\0&0&0&1&1&1&1&0&0\\0&0&0&0&0&0&0&1&1\end{bmatrix}.\begin{bmatrix}9\\13\\11\\14\\16\\14\\8\\5\\7\end{bmatrix}=\begin{bmatrix}97\\33\\52\\12\end{bmatrix}$$

Şimdi $(X'X)^{-1}$ yi hesaplamak için daha önce gösterdiğimiz g-inverse bulma yöntemini kullanacağız. Burada P ve Q yu bulma işlemini göstermeden kısaca yazalım $(X'X)^{-1}$ matrisi simetrik olduğundan Q–P' dür.

$$P=\begin{bmatrix}1&0&0&0\\-0.33&1&0&0\\-0.66&0.66&1&0\\-1&1&1&1\end{bmatrix},\quad Q=\begin{bmatrix}1&-0.33&-0.66&-1\\0&1&0.66&1\\0&0&1&1\\0&0&0&1\end{bmatrix}$$

$$\Delta=PAQ=\begin{bmatrix}9&0&0&0\\0&2&0&0\\0&0&1.33&0\\0&0&0&0\end{bmatrix}$$

$$\Delta^{-}=\begin{bmatrix}0.16&0&0&0\\0&0.5&0&0\\0&0&0.75&0\\0&0&0&0\end{bmatrix}$$

$$Q\Delta^{-}=\begin{bmatrix}0.11 & -0.18 & -0.5 & 0\\ 0 & 0.5 & 0.5 & 0\\ 0 & 0 & 0.75 & 0\\ 0 & 0 & 0 & 0\end{bmatrix}$$

$$G=Q\Delta^{-}P=\begin{bmatrix}0.5 & -0.5 & -0.5 & 0\\ -0.5 & 0.81 & 0.5 & 0\\ -0.5 & 0.5 & 0.75 & 0\\ 0 & 0 & 0 & 0\end{bmatrix}$$

Burada $(X'X)^{-1}=G$ dir. Şimdi $(X'X)^{-1}X'Y$ yani $GX'Y$ yi hesaplayalım. Çünkü $\beta=GX'Y$ dir.

$$\beta_{01}=GX'Y=\begin{bmatrix}0.5 & -0.5 & -0.5 & 0\\ -0.5 & 0.81 & 0.5 & 0\\ -0.5 & 0.5 & 0.75 & 0\\ 0 & 0 & 0 & 0\end{bmatrix}.\begin{bmatrix}93\\ 33\\ 52\\ 12\end{bmatrix}=\begin{bmatrix}0\\ 11\\ 13\\ 6\end{bmatrix}$$

dır. Bu sonuçtan genel çözümü şöyle yazabiliriz.

$$\beta_0=\begin{bmatrix}0\\ 11\\ 13\\ 6\end{bmatrix}+\left\{\begin{bmatrix}0 & 0 & 0 & 0\\ 1 & 1 & 0 & 0\\ 1 & 0 & 1 & 0\\ 1 & 0 & 0 & 1\end{bmatrix}-\begin{bmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{bmatrix}\right\}.\begin{bmatrix}z_1\\ z_2\\ z_3\\ z_4\end{bmatrix}$$

Basit cebirsel işlemlerden sonra

$$\beta_0 = \begin{bmatrix} 0 \\ 11 \\ 13 \\ 6 \end{bmatrix} + \begin{bmatrix} -1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} . \begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \end{bmatrix} = \begin{bmatrix} -z_1 \\ 11+z_2 \\ 13+z_3 \\ 6+z_4 \end{bmatrix}$$

dır. $Z' = \begin{pmatrix} 2 & 1 & 2 & 1 \end{pmatrix}$ alalım. Buna göre

$$\beta_{02} = \begin{bmatrix} 0 \\ 11 \\ 13 \\ 6 \end{bmatrix} + \left\{ \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \right\} . \begin{bmatrix} 2 \\ 1 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} -2 \\ 12 \\ 15 \\ 7 \end{bmatrix}$$

Şimdi β_{01} ve β_{02} çözümleri için kalanların kareler toplamının değişmediğini gösterelim.

$$\beta'_{01}.X'Y = \begin{pmatrix} 0 & 11 & 13 & 6 \end{pmatrix} . \begin{bmatrix} 97 \\ 33 \\ 52 \\ 12 \end{bmatrix} = 1111$$

$$\beta'_{02}.X'Y' = \begin{pmatrix} -2 & 12 & 15 & 7 \end{pmatrix} . \begin{bmatrix} 97 \\ 33 \\ 52 \\ 12 \end{bmatrix} = 1111$$

Böylece iki ayrı çözüm için $\beta_0 X'Y$ nin değişmelerini gördük. SS_e yi hesaplayalım.
$Y'Y$= 1157 den
SS_e=1157 - 1111=46
dır. σ^2 tahmini ise $\sigma^2 = SS_e/(9-3) = 46/6 = 7.67$ dir.

SONUÇ

İstatistiğin en önemli konularından biri olan ve istatistiği merak eden bir kişi istatistiğin bazı temel konularını öğrendikten sonra olaylar arasında ilişki kurmak, onları değerlendirmek ve çeşitli sonuçlara ulaşmak için öğrenmesi gereken en önemli konulardan biride regresyon analizinin özel bir durumu olan lineer modeller konusudur.

Bende bunun için lineer model konusunu en ince detayına kadar bu tezimde inceleyip, dünyada her ne kadar temeli 1945 lere dayanıyorsada son zamanların en popüler konusu olan genelleştirilmiş ters matrisin çok az kısmını lineer modele uygulayarak β tahmincisini, varyansını ve kalanların kareler toplamını bulmak için kullandım.

Söz konusu olan bu tez iki bölümden oluşmaktadır. İlk bölüm lineer modellerle ilgili tüm gerekli açıklamalarverilmiştir. Bölümün sonunda ise konuyla ilgili iki tane örneğe yer verilmiştir. İkinci bölüümde ise genelleştirilmiş ters matrislere kısaca değinilmiş ve bu ters matrislerin lineer modellere uygulanması tanım ve teoremlerle anlatılmış, özellikle rankı tam olmayan lineer modellere uygulanması üzerinde durulmuş β tahmincisinin elde edilmesi, varyansın tahmini ve kalanların kareler toplamının bulunan farklı β_0 değerleri için değişmediği gösterilmiştir. Ayrıca konu bir örnekle desteklenmiştir.

Söz konusu olan bu tezde birinci bölümün içeriğini biraz daha açalım: Lineer modelin oluşumunu inceledikten sonra en küçük kareler yöntemi ile modelin bilinmeyen parametrelerin tahmin değeri kareler toplamını minimum yapan değerini bulduk. Parametrelerin tahmincisi için önemli olan varyansın tahminini, gözlenen değerler ile elde edilen sonuçlar arasında ilişki olup olmadığını araştırmak için kullanılan kovaryansı ve farklı olaylar arasındaki ilişki kurulduktan sonra bu ilişkinin güvenirliğini araştıran korelasyon katsayısı ile bu katsayının güvenirliğini araştıran Fisher testinin kullanılmasını inceledik. Sonra $\hat{\alpha}$ ve $\hat{\beta}$ nın güven aralıklarının belirlenmesinde en önemli rolü oynayan α ve β tahmincilerinin bulunması ve bunların güven aralıkları ile x in verilmiş bir değeri için y in beklenen değerini tahmin etme ve ayrıca x in verilmiş bir değeri için y in özel bir değerinin t-istatistiği yada student istatistiği tahmin edilmesi, en son olarak da lineer modelin doğrusallığının Fisher testi yer almaktadır.

İkinci bölümde ise genelleştirilmiş ters matrislerin kısa bir tarihçesi, tanımı, elde edilmesi ve özellikleri kısaca ele alınmış ayrıca bir örnekle herhangi bir genelleştirilmiş g-inverse matrisi bulunmuştur. Sonra genelleştirilmiş ters matrislerin lineer modellere nasıl uygulandığı anlatılmış ve rankı tam olmayan matrislerin genelleştirilmiş ters matrislerle nasıl inversinin bulunduğu ele alınmıştır. Ayrıca β tahmincisi için gerekli olan varyansın tahmini ile kalanların kareler toplamının tahmini bulunmuştur. Son olarak rankı tam olmayan bir matris örneği ile konu özetlenmiştir.

Bu işlemleri yaparken model olarak tek faktörlü varyans analizi modelini temel olarak aldık.

Bu çalışmamızdan cıkaracağımız sonuçlar şunlardır:

Genelleştirilmiş ters martrislerle, hiçbir kısıtlamaya başka bir deyişle zorlamaya gerek kalmadan çözüm yapılabilmektedir. Bu ise hangi g-inverse yada hangi β_0 çözümü kullanılırsa kullanılsın çözüm için bazı değerlerin değişmediğini görürüz.

Diğer yandan rankı tam olmayan modeller, rankı tam olan modellerin genelleştirilmiş bir halidir. Graybill gibi bazı yazarların, X matrisinin özel bir yapısı nedeniyle, rankı tam olmayan lineer modelleri diğerlerinin özel bir biçimi olarak gösterdiklerini belirtmiştik. Ama bunun terside doğrudur. Şöyleki rankı tam olmayan modellerde $\beta_0 = \hat{\beta}$ olursa, tam ranklı modellere ulaşılır. Demekki rankı tam olan modeller, rankı tam olmayanların özel bir şeklidir.

Son olarak şunu belirtelim: Lineer denklemler Penrose un yalnızca birinci koşulunu sağlayan bir g-inversle çözülebiliyorsa, lineer modellerinde yalnızca $X'XGX'X-X'X$ koşulunu sağlayan bir g-inversle çözülebileceğini görüyoruz.

KAYNAKLAR

1-) Necla ÇÖMLEKÇİ, Deney Tasarımı ve Çözümlemesi, Sf. 25-145, Anadolu Üniversitesi Basım Evi, 1988, Eskişehir.

2-) Merih İPEK, Genelleştirilmiş Ters Matrisler ve Rankı Tam Omayan Modellere Uygulama, Sf. 20-70, Doçentlik Tezi, İstanbul Üniversitesi İktisat Fakültesi, 1980, İstanbul.

3-) Aziz BENER, Matematiksel İstatistik, Sf. 405-412, Basılmamış Ders Notları, İstanbul.

4-) Fikri AKDENİZ, İstatistik ve Olasılık, Sf. 417-456, Ankara Üniversitesi, 1985, Ankara.

5-) R,L,SEARLE The Linear Models and Applications, Sf. 1-61.

6-) R.A.Penrose, A Generalized İnverse and Marix Theory, Sf. 105-145.

7-) Köksal B. ALOBA, İstatistik Analiz Metotları,Sf. 100-200, Çağlayan Basım Evi, 1985, İstanbul.

8-) Özden KILIÇAY, Turbo Pascal 6.0 for Wındows, Sf. 51-63, 1990, Ankara.

9-) Ekrem ULUSOY, Ters Matrisler, sf.7-28, Karadeniz Üniversitesi Basımevi, 1984, Trabzon

10-) Mark L. Berenson, David M. Levine, Matthew Goldstein, Intermediate Statıcsal Methods and Aplications a Computer Package Approach, sf. 203-242, Prentice-Hall, Inc, Enlewood Cliffs,New Jersey

11-) Adnan MAZMANOĞLU, " Etkileşimsiz 2 - Faftörlü çapraz sınıflandırma Modellerinin Matrislerle Çözümü", Doktora Tezi, İstanbul Üniversitesi İktisat Fakültesi,1984 , İstanbul

EKLER

EK A1.I:

Bu bölümde, y nin x üzerindeki regresyon doğrusu için bir formül geliştirilmiştir. Yapılan hesaplamalar cebirsel yapılabilir.

İlk olarak, N tane bağımsız gözlemler $(w_1,\ldots,w_N)$ olarak farzedelim ve $V(k)$ niceliği miktarını k ya riayet ederek minimize edilir. Böylece şunu yazabiliriz;

$$V(k)=\sum_{i=1}^{N}(w_i-k)^2 \qquad \text{(A1.1.1.)}$$

Sonra, $k=\overline{w}$ değerinde k üstüne V minimize edilir. Böylece

$$V(k)=\sum_{i=1}^{N}\left[(w_i-\overline{w})+(\overline{w}-k)\right]^2 \qquad \text{(A.1.1.2.)}$$

$$=\sum_{i=1}^{N}\left[(w-\overline{w})^2+(\overline{w}-k)^2+2(\overline{w}-k)(w_i-\overline{w})\right]$$

denklemini yazabiliriz.Burada

$$\sum_{i=1}^{N}(w_i-\overline{w})=\sum w_i-N\overline{w}=N\overline{w}-N\overline{w}=0$$

olduğundadan

$$=\sum_{i=1}^{N}(w_i-\overline{w})^2+N(\overline{w}-k)^2$$

dir. Şimdi $k=\overline{w}$ karşısında sabit $(w_1,\ldots,w_N)$ için $V(k)$, negatif olmayan iki niceliğinin toplamıdır ve açıkça görülebileceği gibi minimumdur.

Regresyon problemi minimum $Q(\alpha,\beta)$ yı verecek α ve β değerlerini ister.(1.14)de tanımlandığı gibi

$$Q(\alpha,\beta)=\sum_{i=1}^{N}(y_i-\alpha-\beta x_i)^2 \qquad \text{(A1.1.3.)}$$

Burada $x_1,\ldots,x_N$ eşit değildirler. Eğer biz $y_i = \beta x_i = w_i$ alırsak $\bar{y} - \beta\bar{x} = \bar{w}$ yı elde ederiz ve (A1.1.2.) yı kullanarak, şöyle yazarbiliriz.

$$Q(\alpha,\beta) = \sum_{i=1}^{N}(w_i - \alpha)^2 = \sum_{i=1}^{N}(w_i - \bar{w})^2 + N(\bar{w} - \alpha)^2 \qquad \text{(A1.1.4.)}$$

Şüphesiz, yapılacak seçim β nın herhangi bir sabit değeri için minimine edilir.

$$\alpha = w = \bar{y} - \beta\bar{x} \qquad \text{(A1.1.5.)}$$

Ayrıca α nın $(y - \beta x)$ olarak alınması ve herhangi β değeri ile $Q(\alpha,\beta)$ minimize edilir. Sonuçta $Q(\alpha,\beta)$ değeri

$$Q(\bar{y} - \beta\bar{x},\beta) = \sum_{i-1}^{N}\left[y_i - (\bar{y} - \beta\bar{x}) - \beta x_i\right]^2 \qquad \text{(A1.1.6.)}$$

$$= \sum_{i-1}^{N}\left[(y_i - \bar{y}) - \beta(x_i - \bar{x})\right]^2$$

Bu yalnız β nın bir fonksiyonudur. Bunu $Q(\beta)$ ile gösterebiliriz. Birbirlerine eşit olmayan $x_1,\ldots,x_N$ sabit değerleri için $Q(\beta)$ yi minimum yapan β değeri bulunur. Burada

$$A = \sum_{i=1}^{N}(x_i - \bar{x})^2 > 0 \qquad ,B = \sum(x_i - \bar{x})(y_i - \bar{y}) \quad \text{ve } C = \sum(y_i - \bar{y})^2$$

olduğunda

$$Q(\beta) = A\beta^2 - 2B\beta + C \qquad \text{(A1.1.7.)}$$

$$= A\left[\left(\beta - \frac{B}{A}\right)^2\right] + \frac{AC - B^2}{A}$$

dır. Böylece eğer β yı aşağıdaki gibi seçersek, seçilen β değeri ile $Q(\beta)$ değeri minimize edilebilir.

$$\beta = \frac{B}{A} = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum(x_i - \bar{x})^2} \qquad \text{(A1.1.8.)}$$

Yani seçilen α ve β değerlerini sırasıyla $\hat{\alpha}$ ve $\hat{\beta}$ gibi seçersek $Q(\alpha,\beta)$ minimize edilir.

$\hat{\alpha}$ ve $\hat{\beta}$ aşağıdaki gibidir.

$$\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x} \qquad \text{A1.1.9.)}$$

$$\hat{\beta} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2}$$

Biz burada (1.56) daki noktayı direkmen elde edebiliriz. Sonra $Q(\alpha,\beta)$ nın minimumu olan $Q(\hat{\alpha},\hat{\beta})$ ve (A1.1.7.) den yararlanarak Şöyle yazılabilir;

$$Q(\hat{\alpha},\hat{\beta}) = C - \frac{B^2}{A} \qquad \text{(A1.1.10)}$$

$$= \sum (y_i - \bar{y})^2 - \frac{\left[\sum (x_i - \bar{x})(y_i - \bar{y})\right]^2}{\sum (x_i - \bar{x})^2}$$

$$= \sum (y_i - \bar{y})^2 - \hat{\beta}^2 \sum (x_i - \bar{x})^2$$

denklemini elde ederiz. (1.56) dan aldığımız notta $Q(\alpha,\beta) = SS_e$ dir.

EK A1.2

Genelleştirilmiş Ters Matrislerin Paskal Dili ile çözümü verilmiştir tezdeki örnek için ise test edilmiştir.

```
program g_inverse;
uses dos,crt;
type
d=array[1..20,1..20] of real;
var
sat,sut,islem,boyut:integer;
n:integer;
i,j,s,k:integer;
bolum,sutis,pivot,f:real;
del,a,b,m1,m2,i1,i2,g,yaz:array[1..20,1..20] of real;
procedure yazim;
begin
    for i:=1 to n do begin
        for j:=1 to n do begin
        gotoxy(j*9,i+1);write(yaz[i,j]:4:3,'          ');
        end;
        writeln;
        end;
end;

procedure carpim;
begin
for i:=1 to n do
    for j:=1 to n do
        for k:=1 to n do
            del[i,j]:=del[i,j]+m1[i,k]*m2[k,j];
end;
procedure sifir;
begin
for i:=1 to 20 do
    for j:=1 to 20 do
            del[i,j]:=0;
end;
procedure invers;
begin
for k:=1 to n-1 do begin
    for i:=1 to n-1 do
        for j:=1 to n-1 do
            if (i=k) or (j=k) then
            {}
```

```
            else del[i,j]:=del[i,j]-del[i,k]*del[k,j]/del[k,k];
  del[k,k]:=-1/del[k,k];
                  for i:=1 to n-1 do
            if i=k then
            {}
            else begin
                del[i,k]:=del[i,k]*del[k,k];
                del[k,i]:=del[k,i]*del[k,k];
           end;
     for i:=1 to n-1 do
        for j:=1 to n-1 do
        del[i,j]:=-del[i,j];
  end;
  end;
 begin
 clrscr;
 writeln('matrisin boyutu');readln(n);
 for i:=1 to n do
    for j:=1 to n do
       begin
           a[i,j]:=0;
           i1[i,j]:=0;
           i2[i,j]:=0 ;
           i1[i,i]:=1;
           i2[i,i]:=1;
      end;
for i:=1 to n do
   for j:=1 to n do begin
      readln(a[i,j]);
    . b[i,j]:=a[i,j];
      end;
                 for j:=1 to n-1 do
                    begin
                        pivot:=a[j,j];
                                   for i:=j+1 to n do
                                   begin
                                       f:=a[i,j]/pivot ;
                                       for s:=1 to n do begin
                                       a[i,s]:=a[i,s]-f*a[j,s];
                                       i1[i,s]:=i1[i,s]-f*i1[j,s];
                                       end
                                   end
                        end;
.for i:=1 to n do
   for j:=1 to n do
       yaz[i,j]:=i1[i,j];
```

```
clrscr;
writeln('P matrisi');
writeln;
yazim;

for sut:=1 to n-1 do begin
        sutis:=a[sut,sut];
                  for sat:=sut+1 to n  do begin
                     bolum:=a[sut,sat]/sutis;
                     a[sut,sat]:=a[sut,sat]-bolum*a[sut,sut];
                     for islem:=1 to n do
                  i2[islem,sat]:=i2[islem,sat]-bolum*i2[islem,sut];
                  end
        end;

writeln('Q matrisi');
for i:=1 to  n do  begin
   for j:=1 to n do begin
      write(i2[i,j]:3:3,'            ');
      end;
writeln;
end;

{ P matrisi ile A matrisinin carpimi}

for i:=1 to n do
   for j:=1 to n do begin
      m1[i,j]:=i1[i,j];
      m2[i,j]:=b[i,j];
      end;
sifir;
carpim;

{ PA matrisi ile Q matrisinin carpimi }

for i:=1 to n do
   for j:=1 to n do begin
      m1[i,j]:=del[i,j];
      m2[i,j]:=i2[i,j];
      end;
sifir;
carpim;

Writeln(' delta matrisi');
for i:=1 to  n do  begin
   for j:=1 to n do begin
```

```
        write(del[i,j]:3:3,'          ');
        end;
        writeln;
        end;

 { Delta matrisinin inversinin alinmasi }

 invers;
 writeln(' invers matris');
        for i:=1 to  n do  begin
           for j:=1 to n do begin
           write(del[i,j]:3:3,'          ');
           end;
           writeln;
        end;

 { Q ile deltanin inversinin carpimi }

 for i:=1 to n do
    for j:=1 to n do begin
       m1[i,j]:=i2[i,j];
       m2[i,j]:=del[i,j];
       end;
       sifir;
 carpim;

 { G invers matrisin elde edilmesi }

 for i:=1 to n do
    for j:=1 to n do begin
       m1[i,j]:=i2[i,j];
       m2[i,j]:=del[i,j];
       end;
sifir;
carpim;

Writeln(' g invers matrisi');
for i:=1 to  n do  begin
   for j:=1 to n do begin
      write(del[i,j]:3:3,'          ');
      end;
      writeln;
   end;
end.
```

ÖZGEÇMİŞ

Adı Soyadı	: İbrahim Demir
Doğum Tarihi	: 27 Mayıs 1970
Doğum Yeri	:Akkuş/ORDU
İlk Öğrenimi	:1982 yılında Sakarya İlkokulundan mezun oldu.
Orta Öğrenimi	:1988 yılında İzmit İmam Hatip Lisesinden Mezun oldu.
Yüksek Öğrenimi	:1988 yılında Yıldız Üniversitesi Fen-Edebiyat Fakültesi Matematik Bölümünde başladığı eğitimini 1992 yılında tamamladı.
Görevi	:Y.T.Üniversitesi Fen-Edebiyat Fakültesi Matematik Bölümü Uygulamalı Matematik anabilim dalı Araştırma Görevlisi