SERİ SERIE B | CİLT TOME XXI | SAYI FASCICULE 1 | 1971

İSTANBUL ÜNİVERSİTESİ

ORMAN FAKÜLTESİ DERGİSİ

REVUE DE LA FACULTÉ DES SCIENCES FORESTIÈRES
DE L'UNIVERSITÉ D'ISTANBUL

EĞRİSEL İLİŞKİLERİN DOĞRUSAL HALE GETİRİLMESİ VE DOĞRUSAL ÇOĞUL REGRESYON ANALİZİ

Yazan :

Dr. Alptekin GÜNEL

«Basit Doğrusal Regresyon» adlı yazımızda (Orm. Fak. Drg. 1971 - B I) regresyon analizinin dayandığı esaslar verilmiş, bu esasların değişken sayısına bağlı olmadığı, değişken sayısının artmasının sadece hesap işlemlerinde bazı farklılıklar doğurduğu belirtilmişti. Ayrıca, regresyon analizinin dayandığı esasları açıklamak amacı ile, en basit olan, iki değişkenli durumun seçilmesinin, söz konusu esasların izahını kolaylaştıracağı zikredilerek, doğrusal olmayan ilişkilerin, dar aralıklarda, doğrusal olarak kabul edilebileceği gibi, bazı transformasyonlarla eğrisel ilişkilerin doğrusal hale dönüştürülebileceği nedenleri ile, açıklamada doğrusal model tercih edilmişti.

Bu yazımızda, sözü edilen transformasyonlar bir kaç örnekler tanıtıldıktan sonra, ikiden fazla değişkenli doğrusal regresyon, daha formel deyimi ile, doğrusal çoğul regresyon modeli üzerinde durulacaktır.

Doğrusal olmayan ilişkileri, doğrusal hale getirmede kullanılan transformasyonları iki grupta toplamak mümkündür :

1 — Logaritmik transformasyon,

2 — Tersine çevirme (invörs).

Bu iki transformasyon şekli ayrı ayrı kullanıldığı gibi, aynı denklem üzerinde beraberce de uygulanabilmektedir.

1 — Logaritmik transformasyon

Logaritma alma işleminin her iki değişken üzerinde de yapılıp yapılmadığına göre, iki kısma ayrılır :

a — Tam logaritmik transformasyon

Logaritma alma işlemi her iki değişken üzerinde uygulanır.

Bilindiği gibi, tek ağaç gövde hacmi

$$V = a\, d^{b} \tag{1}$$

formülünden tayin edilebilmektedir. Formülde,

V = Gövde Hacmi

a, b = Katsayılar,

d = Göğüs çapı'dır.

(1) eşitliği, logaritması alınacak olursa

$$\text{Log } V = \text{Log } a + b \text{ Log } d \tag{2}$$

şekline girer.

$$\text{Log } V = Y$$

$$\text{Log } a = p$$

$$\text{Log } d = X$$

konacak olursa, (2) eşitliğinden

$$Y = p + b X \tag{3}$$

doğru denklemi elde edilir. Serbest değişken olan çapa göre alınacak türevden de görülebileceği gibi, (1) eşitliği monoton artan bir eğri denklemidir. Tam logaritmik transformasyonla, eğri bir doğruya dönüştürülmüştür.

Aynı şekilde,

$$Y = a\, d^{-b} \tag{4}$$

eğrisi, tam logaritmik transformasyonla

$$\text{Log } Y = \text{Log } a - b \text{ Log } d \tag{5}$$

doğru denklemini verecektir. (4) eşitliğinden

$$Y\, d^{b} = a \tag{6}$$

bağıntısı elde edilir. a sabit bir sayıdır. Sabit iki eksene uzaklıkları çarpımı sabit olan noktaların geometrik yeri bir hiperbol olduğundan, yapılan transformasyon, hiperbol karakterindeki bir ilişkiyi doğrusal hale getirmiştir. Bilindiği gibi, seçme ormanlarında, geçiş müddetinin çapa göre değişimi bir hiperbol şeklindedir. *b* nin alacağı değerlere göre değişim seyri Şekil 1'de gösterilmiştir.

b — Yarı Logaritmik transformasyon

Logaritma alma işlemi, değişkenlerden yalnız biri üzerinde yapılır. (7) eşitliği bu çeşit transformasyona bir örnektir.

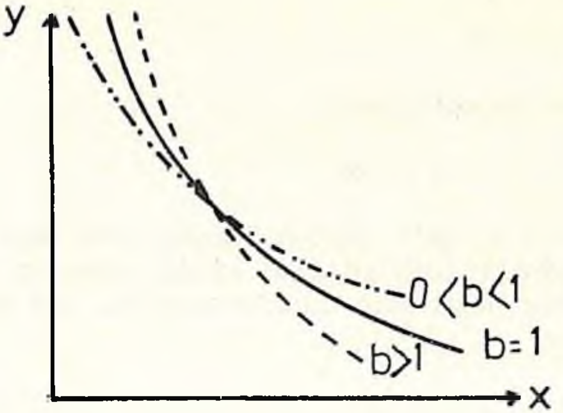
$$Y = A + B \operatorname{Log} X \tag{7}$$

Eşitlik (7) den

$$(Y - A) / B = \operatorname{Log} X$$

veya

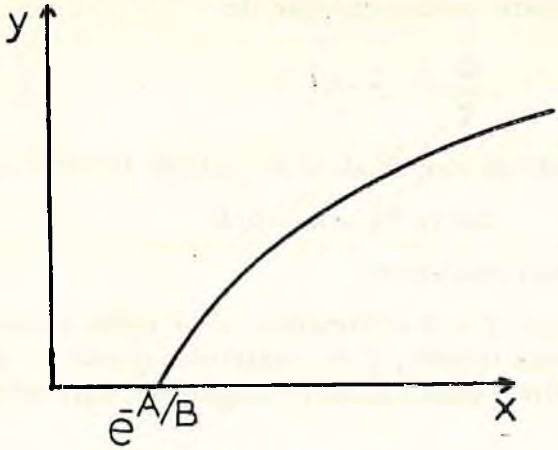
$$X = e^{(Y - A) / B} \tag{8}$$



Şekil 1. $a = d^b$ eğrisinin çeşitli *b* değerlerine göre değişimi

bulunur. (Logaritma bazı *e* sayısı alındığına göre).

A ve B sıfırdan büyük sayılar ise, Y = 0 için, $X_0 = e^{-A/B}$ olur. (8) fonksiyonunun türevi daima pozitif olduğundan, (8) eşitliği, X eksenini $e^{-A/B}$ de kesen ve moton artan bir eğridir. (Şekil 2).



Şekil 2. $X = e^{(Y-A)/B}$

2 — Tersine Çevirme (invörs)

Tek ağaçta, gövde eğrisini tayin için kullanılan formüllerden biri de

$$Y = \frac{X}{a + b\,X} \tag{9}$$

denklemidir. (9) eşitliği

$$\frac{X}{Y} = a + b\,X$$

şeklinde de yazılabilir. (X / Y) = Z konacak olursa

$$Z = a + b\,X \tag{10}$$

elde edilir ki, bu bir doğru denklemidir.

Tersine çevirme transformasyonu, genellikle bağımlı değişken Y'nin asimtotik bir değere sahip olduğu hallerde kullanılmaktadır. Yukardaki denklemde, $X = \infty$ için Y nin asimtotik değeri $Y_\infty = 1/b$ dir.

Her iki transformasyon çeşidinin birlikte ne şekilde kullanıldığı aşağıdaki örnek üzerinde açıklanmağa çalışılacaktır.

$$Y = \frac{1}{e^{a - b/X}} \tag{11}$$

Denklemde a, b ve X in yalnız pozitif değerler aldıklarını kabul edelim. (11) eşitliğinden, tersine çevirme ile

$$\frac{1}{Y} = e^{a - b/X} \tag{12}$$

eşitliği, bu eşitliğe uygulanacak logaritmik transformasyonla da

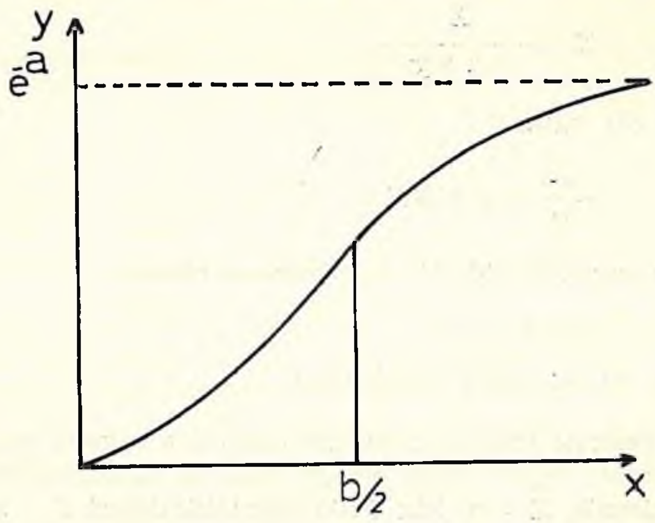
$$\mathrm{Log}\,(1/Y) = a - b/X \tag{12a}$$

doğru denklemi elde edilir.

X = 0 için, Y = 0 olduğundan (11) eğrisi orijinden geçer. X in sonsuza gitmesi halinde, Y nin asimtotik değeri e^{-a} dır. (11). eşitliğinin birinci türevi daima pozitif olduğundan, eğri monoton artar. İkinci türev

$$\frac{d^2Y}{d\,X^2} = e^{a - b/X}\,(b - 2X)$$

olup, X = $b/2$ noktasında eğrinin bir dönüm noktası mevcuttur. (Şekil 3) Şekil 3 den de görüldüğü gibi, (11) eşitliği tipik bir S — eğrisi vermektedir. Büyüme eğrilerinin bu formda oldukları bilinmektedir.



Şekil 3. $Y = \dfrac{1}{e^{a - b/X}}$

Yukarda verilen örneklerle, değişik karakterdeki eğrisel ilişkilerin nasıl doğrusal hale getirildikleri gösterilmek istenilmiştir. Bu çeşit transformasyonlarda dikkat edilecek husus, transformasyonlardan sonra, regresyon analizi esaslarının ihlâl edilmediğinden emin olmaktır. Bu amaçla geliştirilmiş, yeterli testler vardır.

Doğrusal olmayan ilişkilerde, ilişkiyi belirliyen katsayıların herhangi bir transformasyona baş vurmadan da tayini mümkündür. Meselâ,

$$V = a\, d^b$$

eşitliğinde, a ve b katsayıları, en küçük kareler prensibine göre elde edilecek aşağıdaki normal denklemlerden hesaplanabilir.

$$a \sum V b d^{b-1} - a^2 \sum b d^{2b-1} = 0$$
$$\sum V\, d^b - a \sum d^{2b} = 0$$

Ancak bu yol, transformasyon yapılması haline kıyasla, daha karışık olup, daha çok hesap işlemlerini gerektirmektedir. Amaç, var olan ilişkiyi, mümkün olan en küçük dereceden bir çok terimli ile ifade etmektir.

DOĞRUSAL(*) ÇOĞUL REGRESYON ANALİZİ

İkiden fazla değişkenler arasındaki istatistiksel ilişkiyi tayinde çoğul regresyon analizi kullanılır. Bu ilişki doğrusal veya eğrisel olabilir. Burada, doğrusal hal üzerinde durulacak ve önce üç değişkenli ilişki örnek olarak alınacaktır.

X_1 ve X_2 gibi iki bağımsız değişkenle Y bağımlı değişkeni arasındaki gerçek ilişki

$$Y_i = \alpha + \beta_1 (X_{i1} - U_1) + \beta_2 (X_{2i} - U_2) \qquad (13)$$

şeklinde olsun. Eşitlikte,

$\alpha;\ \beta_1;\ \beta_2$ = Denklemin parametreleri
$U_1 = X_1$ lerin gerçek ortalaması
$U_2 = X_2$ lerin gerçek ortalaması

dırlar. Denklem (13) deki gerçek değerleri tam olarak tayin etmek, genellikle mümkün değildir. Yapılacak ölçmeler yardımı ile, ilişkinin katsayıları, istatiksel anlamda hesaplanabilir.

(13) denkleminin, en küçük kareler prensibine göre bulunmuş regresyon denklemi

$$\hat{Y} = a + b_1 (X_{1i} - \bar{X}_1) + b_2 (X_{2i} - \bar{X}_2) \qquad (14)$$

olsun.

*) Doğrusallık deyimi, genellikle, denklemin parametreleri için kullanılır. Bu durum peşinen kabul edilmiş farz olunarak doğrusallık serbest değişkenin birinci dereceden olduğunu belirtmek için kullanılmıştır.

Denklem (14) üç boyutlu uzayda, düzlem denklemidir. Bu düzleme regresyon düzlemi denir.

En küçük kareler prensibine göre,

$$\Sigma (Y - \hat{Y})^2 = \Sigma [Y - a - b_1 (X_{1i} - \bar{X}_1) - b_2 (X_{2i} - \bar{X}_2)]^2$$
$$= \text{Minimum}$$

olmalıdır.

$$X_{1i} - \bar{X}_1 = x_i$$
$$X_{2i} - \bar{X}_2 = x_2$$
$$Y_i - \bar{Y} = y$$

koyarak, gerekli işlemler yapıldıktan sonra,

$$a = \bar{Y} = \Sigma Y_i/n$$
$$b_1 \Sigma x_1^2 + b_2 \Sigma x_1 x_2 = \Sigma x_1 y \qquad (15)$$
$$b_1 \Sigma x_1 x_2 + b_2 \Sigma x_2^2 = \Sigma x_2 y$$

normal denklemleri elde edilir. b_1 ve b_2 değerleri yukardaki denklemlerden bulunabilir. Meselâ, b_1

$$b_1 = \frac{\Sigma x_1 y \; \Sigma x_2^2 - \Sigma x_2 y \; \Sigma x_1 x_2}{\Sigma x_1^2 \; \Sigma x_2^2 - (\Sigma x_1 x_2)^2}$$

eşitliğinden hesaplanabilir. Ancak, değişken sayısı arttıkça, normal denklem sayısı da artacağından, yukardaki hesap şekli güç ve zaman alıcı olacaktır. Böyle çok değişkenli regresyon denklemlerinin katsayılarını hesaplamada matris işlemlerinin kullanılması büyük kolaylıklar sağlamaktadır. Bununla beraber, matris kavramının, hiç değilse Türk ormancıları arasında, yeteri kadar popüler olmaması nedeni ile, bu rada, elektrikli büro hesap makinesinin varlığı halinde, işlemlerin nispeten daha kolaylıkla yürütülebileceği Doolittle metodu tanıtılmaya çalışılacaktır *.

Metodu tanıtmadan evvel, çoğul regresyon analizi ile ilgili bazı kavramların açıklanması gerekir.

Bağımlı değişken Y ile bağımsız değişkenler (X_i, i = 1,2, ..) arasın-

(*) Doolittle metodu, esas itibariyle simetrik matrisin invörsünü bulma işleminden başka birşey değildir.

da bir korelasyon olduğu gibi, Y ile herhangi bir X_i arasında bir korelasyondan bahsedilebilir. Buna göre,

$$r_{12} = \frac{\Sigma x_1 \; x_2}{\sqrt{\Sigma x_1^2 \, \Sigma x_2^2}} \tag{15}$$

X_1 ile X_2 arasındaki basit korelasyon katsayısıdır. Aynı şekilde

$$r_{y1} = \frac{\Sigma x_1 y}{\sqrt{\Sigma y^2 \; \Sigma x_1^2}} \tag{16}$$ = Y ile X_1 arasındaki basit korelasyon katsayısı,

$$r_{y2} = \frac{\Sigma x_2 y}{\sqrt{\Sigma y^2 \; \Sigma x_2^2}} \tag{17}$$ = Y ile X_2 arasındaki basit korelasyon katsayısı,

$$r_{yj} = \frac{\Sigma x_j \, y}{\sqrt{\Sigma y^2 \; \Sigma x_j^2}} \tag{18}$$ = Y ile X_j arasındaki basit korelasyon katsayısı

dır. Çoğul regresyonda, basit korelasyon katsayılarına ilâveten, kısmi korelasyon katsayıları vardır. Kısmi korelasyon katsayıları yardımı ile Y ile X_k arasındaki ilişkinin diğer bağımsız değişkenlerden etkilenip etkilenmediği ve bu etkinin derecesi ölçülebilir. Üç değişkenli hal için,

$r_{y1\cdot 2}$ = X_2 nin sabit tutulması halinde, Y ile X_1 arasındaki kısmi korelasyon katsayısı,

$r_{y2\cdot 1}$ = X_1 in sabit tutulması halinde, Y ile X_2 arasındaki kısmi korelasyon katsayısı,

$r_{12\cdot y}$ = Y nin sabit tutulması halinde, X_1 ile X_2 arasındaki kısmi korelasyon katsayısı

nı ifade etmek üzere kullanılan işaretlerdir.

Y ile X_k arasındaki kısmi korelasyon katsayısını bulabilmek için, diğer serbest değişkenlerin etkilerinin yok edilmesi lâzımdır. Üç değişkenli halde, Y ile X_1 arasındaki kısmi korelasyon katsayısını bulmak ve bu amaçla X_2 nin Y ile X_1 üzerindeki etkisini yok etmek için, önce Y nin X_2 ye göre doğrusal regresyon denklemi bulunur. Bu denklem,

$$Y = a_{1\cdot 2} + b_{1\cdot 2} X_2 + u_i \tag{19}$$

olsun. Denklemde,

$$a_{1\cdot 2} = \bar{Y} - b_{1\cdot 2}\,\bar{X}_2$$

olduğundan

$$X_2 - \bar{X}_2 = x_2$$

$$Y_1 - \bar{Y} = y$$

yazarak,

$$y = b_{1\cdot 2}\,x_2 + u_i$$

şeklinde de yazılabilir.

Denklem (19a) dan, Y deki «açıklanmamış sapma» veya tesadüfi hata

$$u_i = y - b_{1\cdot 2}\,x_2 \tag{20}$$

bulunur.

Aynı şekilde, X_2 nin etkisi yok edildikten, yani, X_1 in X_2 ye göre doğrusal regresyon denklemi bulunduktan sonra, X_1 deki açıklanmamış sapma (veya tesadüfi hata)

$$v_i = x_1 - b_{2\cdot 2}x_2 \tag{21}$$

dir. Bu durumda, Y ile X_1 arasındaki kısmi korelasyon katsayısı

$$r_{y1\cdot 2} = \frac{\Sigma\, u_i\; v_i}{\sqrt{\Sigma\, u_i^2\; \Sigma\, v_i^2}}$$

$$= \frac{\Sigma\,(y - b_{1\cdot 2}x_2)\,(x_1 - b_{2\cdot 2}x_2)}{\sqrt{y^2\,(1 - r_{y2}^2)}\;\sqrt{x^2\,(1 - r_{12}^2)}}$$

Bazı kısaltmalardan sonra,

$$= \frac{r_{y1} - r_{y2}\; r_{12}}{\sqrt{1 - r_{y2}^2}\;\sqrt{1 - r_{12}^2}} \tag{22}$$

elde edilir. Benzer düşüncelerle hareket edilerek,

$$r_{y2\cdot 1} = \frac{r_{y2} - r_{y1}\; r_{12}}{\sqrt{1 - r_{y1}^2}\;\sqrt{1 - r_{12}^2}} \tag{23}$$

$$r_{12\cdot y} = \frac{r_{12} - r_{y1}\; r_{y2}}{\sqrt{1 - r_{y1}^2}\;\sqrt{1 - r_{y2}^2}} \tag{24}$$

yazılabilir.

(14) Denkleminin korelasyon katsayısı, yani çoğul korelasyon katsayısı

$$R_{y\cdot 12} = \sqrt{\frac{r_{y1}^2 + r_{y2}^2 - 2\, r_{y1}\, r_{y2}\, r_{12}}{1 - r_{12}^2}} \tag{25}$$

$$= \sqrt{1 - \frac{S_{y.12}^2}{S_1^2}} \tag{25a}$$

formülleri yardımı ile bulunur. Çoğul regresyon katsayısının karesi $R_{y.12}^2$ ye determinasyon katsayısı da denilmektedir.

Çoğul regresyona esas alınan değişkenlerin kendi varyansları bulunduğu gibi, Y bağımlı değişkeninin herhangi bir X_i bağımsız değişkenine göre bulunan regresyon denkleminin de bir varyansı olacaktır. Üç değişkenli hal için

Y nin varyansı =

$$S_y^2 = \frac{1}{n-1} \Sigma (Y_i - \overline{Y})^2 = \frac{1}{n-1} \left(\Sigma Y_i^2 - \frac{(\Sigma Y)^2}{n} \right) \tag{26}$$

X_1 in varyansı =

$$S_1^2 = \frac{1}{n-1} \Sigma (X_1 - \overline{X}_1)^2 = \frac{1}{n-1} \left(\Sigma X_1^2 - \frac{(\Sigma X_1)^2}{n} \right) \tag{27}$$

X_2 nin varyansı =

$$S_2^2 = \frac{1}{n-1} \Sigma (X_2 - \overline{X}_2)^2 = \frac{1}{n-1} \left(\Sigma X_2^2 - \frac{(\Sigma X_2)^2}{n} \right) \tag{28}$$

X_1 in X_2 ye göre bulunan regresyon denkleminin varyansı

$$S_{1.2}^2 = \frac{1}{n-1} \Sigma (x_1 - b_{2\cdot 2} x_2)^2 = S_1^2 \left(1 - r_{12}^2\right) \tag{29}$$

Y nin X_1 e göre bulunan regresyon denkleminin varyansı

$$S_{y.1}^2 = \frac{1}{n-1} \Sigma (y - b_{y1}\, x_1)^2 = S_y^2 \left(1 - r_{y1}^2\right) \tag{30}$$

Y nin X_2 ye göre bulunan regresyon denkleminin varyansı

$$S_{y.2}^2 = \frac{1}{n-1} \Sigma (y - b_{y2}\, x_2)^2 = S_y^2 \left(1 - r_{y2}^2\right) \tag{31}$$

X_2 nin X_1 e göre bulunan regresyon denkleminin varyansı

$$S_{2.1}^2 = \frac{1}{n-1} \Sigma(x_2 - b_{32}\, x_1)^2 = S_{12}^2 \left(1 - r_{12}^2\right) \tag{32}$$

Y nin X_1 ve X_2 ye göre bulunan regresyon denkleminin, yani çoğul regresyon denkleminin varyansı

$$S_{y.12}^2 = \frac{1}{n-2} \Sigma (y - b_1 x_1 - b_2 x_2)^2 \tag{33}$$

$$= \frac{1}{n-2} \Sigma (y_2 - \Sigma\, b_i\, \Sigma\, x_i\, y) \tag{33a}$$

$$= S_{y.2}^2\, (1 - r_{y2.1}^2) = S_{y.1}^2\, (1 - r_{y1.2}^2) \tag{33b}$$

$$= (\Sigma\, y^2 - R_{y.12}^2\, \Sigma\, y^2)/(n-2) = (1 - R_{y.12}^2)\; \Sigma\, y^2/(n-2) \tag{33c}$$

ile gösterilecek olursa, Denklem (14) ün katsayıları

$$b_1 = \frac{r_{y1} - r_{y2}\; r_{12}}{1 - r_{12}^2} \; \frac{S_y^2}{S_1^2} = r_{y1\cdot 2} \; \frac{S_{y\cdot 2}}{S_{1\cdot 2}} \tag{34}$$

$$b_2 = \frac{r_{y2} - r_{y1}\; r_{12}}{1 - r_{12}^2} \; \frac{S_y^2}{S_2^2} = r_{y2\cdot 1} \; \frac{S_{y\cdot 2}}{S_{2\cdot 1}} \tag{35}$$

denklemlerinden hesaplanabilirler. Yukarda verilen istatistikleri bulmada, denklemlerin en sağındaki eşitliklerin kullanılması, hesap işlemlerini daha kolaylaştırmaktadır.

Denklem (22), (23), (24), (25), 26), (27), (28), (29), (30), (31), (32) ve (33c) dikkate alınır ve gerekli kısaltmalar yapılırsa,

$$r_{y1.2}^2 = \frac{S_{y.2}^2 - S_{y.12}^2}{S_{y.2}^2} = \frac{S_y^2\,(1 - r_{y2}^2) - S_y^2\,(1 - R_{y.12}^2)}{S_y^2\,(1 - r_{y2}^2)} \tag{36}$$

$$= \frac{R_{y.12}^2 - r_{y2}^2}{1 - r_{y2}^2} \tag{36a}$$

benzer şekilde,

$$r^2_{y1.2} = \frac{R^2_{y.12} - r^2_{y2}}{1 - r^2_{y2}} \qquad (36)$$

$$r^2_{y2.1} = \frac{R^2_{y.12} - r^2_{y1}}{1 - r^2_{y2}} \qquad (37)$$

elde edilir. Bu sonuçlara göre,

$r^2_{y1.2}$, X_1 in ilâvesi ile Y de meydana gelen varyasyon oranını

$r^2_{y2.1}$, X_2 nin ilâvesi ile Y de meydana gelen varyasyon oranını ölçmektedirler.

Yukarıda, daha ziyade tatbikatçı yönünden, ana hatları ile açıklanmasına çalışılan üç değişkenli doğrusal çoğul regresyonla ilgili istatistiklerin bulunmasında takip edilecek işlem sıraları aşağıda özetlenmiştir :

1. Önce, (ΣY), (ΣX_1), (ΣX_2), (ΣY^2), (ΣX_1^2), (ΣX_2^2), $(\Sigma Y X_1)$, $(\Sigma Y X_2)$, $(\Sigma X_1 X_2)$ değerleri, bunlar yardımı ile $\bar{Y}$, $\bar{X}_1$, $\bar{X}_2$, $(\Sigma Y \Sigma X_1)$, $(\Sigma Y \Sigma X_2)$, $(\Sigma X_1 \Sigma X_2)$ dolayısiyle r_{y1}, r_{y2}, r_{12}, S_y^2, S_1^2, S_2^2 istatistikleri yukarda verilen ilgili formüller yardımı ile bulunur.

2. 1. şıkta bulunan istatistikler yardımı ile, gene ilgili denklemlerden

$$r_{y1.2} \; ; \; r_{y2.1} \; ; \; S^2_{y.1} \; ; \; S^2_{y.2} \; ; \; S^2_{1.2} \; ; \; S^2_{2.1}$$

değerleri bulunur.

3. 2. şıkta hesaplananlar yardımı ile Formül (34) ve (35) den b_1 ve b_2 katsayıları hesaplanır.

4. Aynı istatistikler yardımı ile, Formül (33a) dan $S_y^2{}_{.12}$ ve Formül (25) den çoğul regresyon katsayısı $R_{y.12}$ hesaplanır.

Yukarda izahına çalışılan çoğul regresyon denkleminin tayin metodu, değişkenlerin sayısının daha çok olması halleri için uygun değildir. Ayrıca, yukardaki hesap şekilleri, çoğul regresyonu ile ilgili varsayımların testinde gerekli istatistikleri vermemektedir. Bu nedenle, yukarda da belirtildiği gibi, daha genel durumlara kolaylıkla uygulanabilen Doolittle metodunun tanıtılmasında isabet görülmüştür.

Regresyon düzleminin katsayılarını tayin için en küçük kareler prensibinin uygulanması ile elde edilen normal denklemler

$$\begin{aligned} b_1 \Sigma x_1^2 + b_2 \Sigma x_1 x_2 &= \Sigma x_1 y \\ b_1 \Sigma x_1 x_2 + b_2 \Sigma x_2^2 &= \Sigma x_2 y \end{aligned} \qquad (15)$$

idi. b_1 ve b_2 tayini istenen elemanlar olduğundan, bunlara bilinmeyenler nazarı ile bakılırsa, bu bilinmeyenlerin normal denklemlerdeki katsayılarının simetrik olduğu görülür. Gerçekten, birinci denklemin sol tarafındaki ikinci bilinmeyenin katsayısı, ikinci denklemin sol yanındaki birinci bilinmeyenin katsayısına eşittir. Değişkenler daha çok sayıda olsaydı, bulunacak normal denklemler gene böyle simetrik bir durum göstereceklerdi. Yani, denklemin sol üst köşesini sağ alt köşesine birleştiren doğruya göre simetrik olan sayılar birbirlerine eşittirler. İşte, yukardaki durumda olduğu gibi, simetrik yapıdaki doğrusal denklemlerin çözümünde Doolittle metodu hesap işlemlerini büyük ölçüde azalttığı için, tercihen kullanılmaktadır.

Doolittle metodu ve bu metodun (b_i) katsayılarını tayin amacı ile kullanılan «öne doğru» çözüm şeklinde takip edilecek hareket tarzı şöyledir : (Metodun izahında dört değişkenli hal kullanılmıştır)

1. Değişkenler, Y, X_1, X_2, X_3 olsun.

$$Y - \bar{Y} = y \;;\; X_1 - \bar{X}_1 = x_1 \;;\; X_2 - \bar{X}_2 = x_2 \;;\; X_3 - \bar{X}_3 = x_3$$

yazarak,

Σx_i^2 ; $\Sigma x_i x_j$ ve $\Sigma x_i y$ değerleri bulunur. ($i = 1,2.3.4$)

Bundan sonra, x_1 in birinci çarpan olarak bulunduğu terimler birinci satıra, x_2 nin birinci çarpan olarak bulunduğu terimler ikinci satıra, x_3 ün birinci çarpan olarak bulunduğu terimler üçüncü satıra vb. yazılır. Yukardaki durum için, elde edilecek tablo şöyledir :

	$a_{11} = \Sigma x_1^2$	$a_{12} = \Sigma x_1 x_2$	$a_{13} = \Sigma x_1 x_3$	$a_{14} = \Sigma x_1 y$
I A		$a_{22} = \Sigma x_2^2$	$a^2{}_3 = \Sigma x_2 x_3$	$a_{24} = \Sigma x_2 y$
			$a_{33} = \Sigma x_3^2$	$a_{34} = \Sigma x_3 y$

2. I in birinci satırı, yukardaki tablonun altına aynen yazılarak, değer satırlardaki bütün terimler, birinci satırın ilk terimine bölünür; böylece yeni bir satır elde edilir:

	A_1	$A_{11} = \Sigma x_1^2$	$A_{12} = \Sigma x_1 x_2$	$A_{13} = \Sigma x_1 x_3$	$A_{14} = \Sigma x_1 y$
II	B_1	1	$B_{12} = \dfrac{\Sigma x_1 x_2}{\Sigma x_1^2}$	$B_{13} = \dfrac{\Sigma x_1 x_3}{\Sigma x_1^2}$	$B_{14} = \dfrac{\Sigma x_1 y}{\Sigma x_1^2}$

3. B_1 satırının ikinci terimi B_{12}, A_1 satırının ilk terimi hariç, diğer bütün terimleri ile ayrı ayrı çarpılarak, sonuçlar, I in ikinci satırındaki mütekabil terimlerden çıkarılır.

$$A_2 \qquad A_{22} = \Sigma x_2^2 - \frac{\Sigma x_1 x_2}{\Sigma x_1^2} \Sigma x_1 x_2 \qquad A_{23} = \Sigma x_2 x_3 - \frac{\Sigma x_1 x_2}{\Sigma x_1^2} \Sigma x_1 x_3$$

$$A_{24} = \Sigma\, x_2 y - \frac{\Sigma x_1 x_2}{\Sigma x_1^2} \Sigma x_1 y$$

A_2 satırının terimleri, bu satırının ilk terimi olan A_{22} bölünür ve B_2 satırı elde edilir

$$B_2 \qquad 1 \qquad B_{23} = A_{23} \,/\, A_{22} \qquad B_{24} = A_{24} \,/\, A_{22}$$

4. B_2 satırının ikinci terimi B_{23}, A_2 satırının ikinci terimi A_{23} ile çarpılır, I in üçüncü satırındaki ilk terimden çıkarılır. Aynı şekilde, B_{23} ile A_2 satırının üçüncü terimi A_{24} çarpılarak I in üçüncü satırının ikinci teriminden çıkarılır. Böylece A_3 satırı elde edilir.

$$A_3 \qquad A_{33} = a_{33} - B_{23}A_{23} \qquad A_{34} = a_{34} - B_{23}A_{24}$$

III A_3 satırının terimleri, bu satırın ilk terimi olan A_{33} ile bölünerek B_3 satırı elde edilir.

$$B_3 \qquad B_{33} = 1 \qquad B_{34} = A_{32} \,/\, A_{33}$$

Yukardaki işlemlerin toplu sonuçları Tablo 1 de verilmiştir. Tablo 1 den alınacak değerler yardımı ile (b_i) lerin hesabı şöyledir :

$b_3 = B_{34}$

$b_2 = B_{24} - B_{23}b_3$

$b_1 = B_{14} - B_{12}b_2 - B_{13}b_3$

Regresyon düzleminin katsayıları örnekten hesaplanmışlardır. Bu nedenle, bir varyasyon göstereceklerdir. (b_i) katsayısının varyansı

$$\text{Var}\,(b_i) = S^2_{y.123}\; c_{ii} \tag{38}$$

formülünden hesaplanır. Formülde,

$S^2_{y.123}$ = Dört değişkenli hal için, çoğul regresyon katsayısı

c_{ii} = Tablo 1 in I inci kısmındaki matrisin makûsu (invörs) dur.

c_{ij} katsayıları, Tablo 1 deki değerler yardımı ile bulunabilir. Yukarıdaki dört değişkenli hal için,

$$c_{33} = 1 / A_{33}$$
$$c_{23} = -c_{33}\ B_{23}$$
$$c_{13} = -c_{23}\ B_{12} - c_{33}\ B_{13}$$
$$c_{32} = c_{23}$$
$$c_{22} = (1 / A_{22}) - c_{23} B_{23}$$
$$c_{12} = -c_{22} B_{12} - c_{23}\ B_{13}$$
$$c_{11} = (1 / A_{11}) - c_{12} B_{12} - c_{13} B_{13}$$

eşitliklerinden yararlanılır. c_{ij} katsayılarının bu şekilde bulunmasına Doolittle metodunun «geriye doğru» çözümü denir. Bu hesaplardan sonra, herhangi bir regresyon düzlemi katsayısının varyansı kolaylıkla bulunabilir.

Yukarda izahına çalışılan esaslar kavranıldıktan sonra, metod istenilen sayıda değişkenli regresyon analizlerine uygulanabilir.

Tablo 1. Doolittle metodunun öne doğru çözüm şekli
Dört değişkenli hale örnek olarak verilmiştir.

I		a_{11}	a_{12}	a_{13}	a_{14}
			a_{22}	a_{23}	a_{24}
				a_{33}	a_{34}
II	A_1	$A_{11} = a_{11}$	$A_{12} = a_{12}$	$A_{13} = a_{13}$	$A_{14} = a_{14}$
	B_1	1	$B_{12} = \dfrac{A_{12}}{a_{11}}$	$B_{13} = \dfrac{A_{13}}{a_{11}}$	$B_{14} = \dfrac{A_{14}}{a_{11}}$
III	A_2		$A_{22} = a_{22} - a_{12}B_{12}$	$A_{23} = a_{23} - a_{13}B_{12}$	$A_{24} = a_{24} - a_{14}B_{12}$
	B_2		1	$B_{23} = A_{23} / A_{22}$	$B_{24} = A_{24} / A_{22}$
IV	A_3			$A_{33} = a_{33} - A_{23}B_{23}$	$A_{34} = a_{34} - A_{24}B_{23}$
	B_3			1	$B_{34} = A_{34} / A_{33}$

Çoğul Regresyonla ilgili varsayımların testi

a — X_i bağımsız değişkeninin ilâvesinin öneminin kontrolu

Üç boyutlu uzayda, regresyon düzlemimizin denklemi

$$Y = a + b_1 (X_1 - \bar{X}_1) + b_2 (X_2 - \bar{X}_2)$$

olsun. $y = Y - \bar{Y}$ olmak üzere, $(R^2_{y.12} - r^2_{y1}) \Sigma y^2$ ifadesi değişkenlerin regresyondaki önem derecelerini ölçmede kullanılabilir. Bu ifadede

$R^2_{y.12} \Sigma y^2$ terimi, hiç bir bağımsız değişkenin kullanılmaması halinde, Σy^2 nin miktarından, X_1 ve X_2 değişkenlerinin kullanılması ile meydana gelen azalmanın ölçüsüdür.

$r^2_{y1} \Sigma y^2$ terimi ise, hiç bir bağımsız değişken kullanılmaması halinde teşekkül edecek Σy^2 miktarından, X_1 in kullanılması ile meydana gelen azalmanın ölçüsüdür.

Böylece, $(R^2_{y.12} - r^2_{y1}) \Sigma y^2$ ifadesi, yalnız X_1 in kullanılmasına nazaran Σy^2 de, X_1 ve X_2 nin beraberce kullanılması sonucu meydana gelen azalma miktarını ölçecektir. Bu durum, denkleme X_1 in ilâvesinin önemli olup olmadığını kontrolda kullanılabilir.

Σy^2 nin serbestiyet derecesi 1 (bir) dir. X_1 in önem derecesini kontrolda kullanılacak istatistik

$$F_{(1,\ n-2)} = \frac{\dfrac{r^2_{y1} \Sigma y^2}{1}}{\dfrac{(1 - r^2_{y1}) \Sigma y^2}{n-2}}$$

$$= \frac{(n-2)\, r^2_{y1}}{1 - r^2_{y1}} \qquad (39)$$

Regresyon denklemine, X_1 den sonra X_2 yi eklemenin önemi

$$F_{(1,\ n-3)} = \frac{\dfrac{(R^2_{y.12} - r^2_{y1}) \Sigma y^2}{1}}{\dfrac{(1 - R^2_{y.12}) \Sigma y^2}{n-3}}$$

$$= \frac{(n-3)(R^2_{y.12} - r^2_{y1})}{(1 - R^2_{y.12})} \qquad (40)$$

X_3 gibi üçüncü bir bağımsız değişkenin ilâvesinin önemi söz konusu olsaydı, bu taktirde kullanılacak istatistik

$$F_{(1,\ n-4)} = \frac{(n-4)\,(R^2_{y.123} - R^2_{y.12})}{1 - R^2_{y.123}} \tag{41}$$

olacaktı.

Yukardaki formüllere göre bulunacak F istatistiklerinin, seçilen güvenirlik derecesi ve eşitliklerde gösterilen serbestiyet derecelerine göre F — tablosundan alınacak değerlerden küçük olması halinde, söz konusu bağımsız değişkenin ilâvesinin önemli olmadığı, aksi halde bu değişkenin regresyon denklemine dahil edilmesi gerektiği yargısına varılır.

b — Çoğul Regresyonun Katsayıları ile İlgili Varsayımlar

Çoğul regresyon denkleminde herhangi bir b_i katsayısının belirli bir B_i değerinden farklı olup olmadığını araştırmak amacı ile

$$t = \frac{b_i - B_i}{S_{b_i}} \tag{42}$$

istatistiği kullanılır. Eşitlikte,

B_i = b_i nin eşit olduğu varsayılan değer

S_{bi} = b_i nin vanyansının kare kökü, yani standard sapması olup (38) nolu formülden bulunabilir.

Bulunacak t değeri, seçilen güvenirlik derecesi ve $(n - v - 1)$ serbestiyet derecesine göre alınacak tablo değerinden büyükse varsayım red edilir. Burada,

n = örnek ünitesi sayısı (ölçme sayısı)

v = Bağımsız değişken sayısı

dırlar.

$B_i = 0$ varsayısı için (42) nolu eşitlik

$$t = b_i \,/\, S_{b_i} \tag{42a}$$

şekline girer. Bu varsayımın testi tek yönlüdür.

Değişkenler arasındaki korelasyonların tefsirinde dikkatli olmak gerekir. Meselâ, dört değişkenli — bir bağımlı, üç bağımsız — halde, $r_{y\cdot 123}$ kısmi korelasyon katsayısında, regresyona dahil edilebilecek diğer bağımsız değişkenler göz önüne alınmamıştır. Regresyona, dördüncü bir bağımsız değişkenin dahil edilmesi halinde durum ne olacaktır? Bu davranış, Y ile X_i arasındaki ilişkiyi değiştirecek midir? Bu soruya kısmen, çoğul korelasyon katsayısı ile cevap verilebilir. Zira, R^2, diğer değişkenlerin Y'de meydana getirdikleri varyasyonun bir ölçüsüdür. Böylece, meselâ, $R_y^2{\cdot}_{123}$ bir değerine çok yakınsa, X_1, X_2 ve X_3 bağımsız değişkenleri Y ile sıkı bir ilişki göstermektedirler; bu nedenle, yeni bir bağımsız değişkenin ilâvesi Y deki varyasyonu önemli ölçüde azaltmayacaktır; diğer bir deyişle, Y ile olan kısmı korelasyon katsayılarını etkilemeyecektir. Buna karşılık, $R_y^2{\cdot}_{123}$ birden farklı, fakat X_4 ün ilâvesi $R_y^2{\cdot}_{1234}$ bir değerine çok yaklaştırıyorsa, bu taktirde kısmi korelasyonların anlamları farklıdır.

Genel olarak, X_i ile X_j arasındaki *basit korelasyon* katsayısının, bu iki değişken arasındaki *kısmi korelasyon* katsayısından büyük olması, X_i ile X_j arasındaki basit korelasyonun, bu iki değişkenin diğer bir (veya daha fazla) değişkenle yakın ilgisi olması sonucu ortaya çıktığı düşüncesini uyandırır. Yukardakinin tersi bir durum varit yani, kısmi korelasyon katsayısı daha büyükse, X_i ile X_j aslında basit korelasyon katsayısının ifade ettiğinden daha kuvvetli bir korelasyona sahiptirler; ancak, bu korelasyon diğer değişkenlerin etkisi ile zayıflamış, adeta onlar tarafından gizlenmiştir.

FAYDALANILAN ESERLER

Anderson R.L., Bancroft T.A., Statistical Theory in Research. McGraw - Hill, N.Y.

Dixon W.J., F.J. Massey Jr. 1956. Introduction to Statistical Analysis, Mc - Graw - Hill, N.Y.

Draper N.L., H. Smith, 1966. Applied Repression Analysis, John Wiley and Sons. N.Y.

Fırat F., 1962, Dendrometri, Orm. Fak. Ya., İstanbul.

Fraser D.A.S., 1964, Statistics An Introduction John - Wiley and Sons, N.Y.

Kalıbsız A., 1968, Meşcere Hacım Artımı Tayininde Kullanılan Meyer Metodları ve Kritiği, Orm. Fak. Ya. İstanbul.

Li J.C.R., 1964, Statistical Inference Vol. I, II, Michigan.

Prodan M., 1966, Forest Biometrics Pergoman Press, London.

Snedecor G.W., 1953, Statistical Methods. Iowa State College Press.