Öneri, C.1,S.5
Haziran, 1996, ss.49-54.

# REGRESYON ANALİZİNDE BAĞIMSIZ DEĞİŞKENLERİN BELİRLENMESİNİ SAĞLAYAN ALGORİTMALAR

*Dr. S. Ümit OKTAY FIRAT*

*Marmara Üniversitesi, AT Enstitüsü, Yardımcı Doçent*

***ABSTRACT:** In this paper, we discussed algorithms for selecting independent variables from a larger group of candidates in regression analyses.*

***We described the following four methods for composing predictor sets: i-Backward elimination of variables, ii- All possible regressions, iii-Forward selection of variables, iv-Stepwise regression selection. These methods can be applied by using computer.***

## I. GİRİŞ

Gerek sosyal bilimlerde, gerekse pozitif bilimlerde ilgilenilen herhangi bir sonuç (bağımlı değişken) veya durum, genellikle birden çok faktörün (bağımsız değişkenler) etkisiyle ortaya çıkmaktadır [1;42-43]. Y ile göstereceğimiz bir olayın veya sonucun, $X_1, X_2, ..., X_p$ gibi p-tane faktörün etkisiyle oluştuğunu varsayalım [2;45; 3;233]. Y bağımlı değişken, bağımsız değişkenlerden oluşan bir lineer regresyon denklemi ile ifade edilebilmektedir. Bu denklem bağımsız değişkenler setindeki elemanların herhangi bir kombinasyonu olabilir. İşte kurulacak bu regresyon denkleminde, bulunması gereken değişkenlerin seçimi nasıl olacaktır? Genellikle bilgisayar istatistik paket programları en iyi kombinasyonu seçerek bir denklem vermektedir. Ancak çoğu program, bu nihai denklemin belirlenmesinde kullanılacak yöntemin seçimini de kullanıcının tercihine bırakmaktadır. [4, 5] İşte bu noktada, bazı genel isimler altında sıralanan bu yöntemlerin özellikleri hakkında bilgi sahibi olmak önemli bir gereksinim olarak karşımıza çıkmaktadır.

Çalışmamızın amacı, regresyon denklemi kurulurken, denklemde yer alacak bağımsız değişkenlerin belirlenebilmesi amacıyla, istatistik paket programlarında sık kullanılan bazı yöntemlerin teorik esaslarını ve birbirlerine göre karşılaştırmalı üstünlüklerini incelemektir.

Genellikle regresyon denkleminin kurulmasında, birbirine zıt iki kriter gözönüne alınır[6;.207-223, 7;163-195 ]:

1. İyi bir tahmine olanak verecek bir denklem kurabilmek için, bağımlı değişkeni etkileyeceği düşünülen tüm bağımsız değişkenlerin ($X_i$'ler), regresyon analizinde kapsanması istenir.

2. Çok sayıda bağımsız değişkene ait bilgiyi elde etmenin maliyetinden ötürü de Y'yi en iyi açıklayan birkaç $X_i$ bağımsız değişken ile yetinmek istenir.

İşte bu iki ekstrem yaklaşım arasındaki uzlaşmayı sağlayarak, en iyi regresyon denkleminin seçiminde uygulanmak üzere önerilen ve bilgisayar paket programlarında kullanılan belli başlı yöntemler şunlardır [6,7]:

- Bütün mümkün regresyonlar yöntemi
- Geriye doğru eliminasyon yöntemi (Backward elimination, Step-down method)
- İleriye doğru seçim (Forward Selection, Step-up method)
- Stepwise regresyon

Bu yöntemler incelenecek ve sonra bir uygulamada elde edilen sonuçlar karşılaştırmalı olarak sunulacaktır.

## II. BÜTÜN MÜMKÜN REGRESYONLAR YÖNTEMİ

Yöntem ilk adım olarak, Y ve $X_1, X_2, ..., X_p$ gibi p-tane bağımsız değişkenin bütün mümkün kombinasyonlarını içeren tüm regresyon denklemlerinin oluşturulması ile başlamaktadır. Burada Xo ile ifade edeceğimiz bütün regresyon setlerinde bulunacak regresyon sabitine karşılık gelen bir kukla (dummy) değişken kullanmak yararlı olacaktır ve daima $X_0 =1$ dir. Bu durumda $X_0$ her denklemde bulunmak üzere, p elemanlı bir setin, $2^p$ kadar alt seti olacağından, örneğin p=4 için, ikinci adıma geçmeden önce $2^4=16$ denklemin incelenmesi gerekmektedir [8;443]. Burada, kurulabilecek regresyon denklemleri, m değişken içerenler ( m= 1, 2, ..., p) birer alt set oluşturmak üzere, p-tane alt sete bölünür ve her set kendi içinde bazı kriterlere göre derecelendirilir. Genellikle derecelendirme kriteri olarak determinasyon katsayısı ($R^2$) kullanılmaktadır

[8;444; 9;306-308]. Tüm mümkün regersyonlar yaklaşımında, n gözlem sayısının, potansiyel parametre sayısı k'nın maksimum değerinden daha büyük yani n >k olduğu varsayılmaktdır.

Yöntemi, 4-bağımsız değişkene ait 13 gözlem içeren Hald'ın datası [9;647] ve Draper and Smith'in [7;366] regresyon çözümü üzerinde incelemeyi tercih ediyoruz:

i- Kurulacak regresyon denklemleri, bulunabilecek bağımsız değişken sayısı esas alınarak beş alt sete bölünür:

Set I : Sabit terimden oluşan denklem,

$E(Y) = \beta_0$

Set II : Tek bağımsız değişkenli denklemler,

$E(Y) = \beta_0 + \beta_i X_i$ , (i=1, ..., p)

Set III: İki bağımsız değişkenli denklemler,

$E(Y) = \beta_0 + \beta_i X_i + \beta_j X_j$, ( i, j =1, ..., p)

Set IV:Üç bağımsız değişkenli denklemler,

$E(Y) = \beta_0 + \beta_i X_i + \beta_j X_j + \beta_k X_k$, (i, j, k = 1, ..., p)

Set V : Dört bağımsız değişkenli denklem,

$E(Y) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4$

ii- Her denklem için çoklu determinasyon katsayıları ($R^2$ ler) hesaplanır.

iii- Her setin içinden en yüksek $R^2$ değerine sahip olanlar seçilir. Yukarıda belirtilen data için elde edilen tüm regresyon setleri arasından aşağıda gösterilen en iyi denklemler seçilmiştir.

| Set | Denklemler | $R^2$ |
|---|---|---|
| II | $Y' = f(X_4)$ | % 67.5 |
| III | $Y' = f(X_1, X_2)$ | % 97.9 |
| IV | $Y' = f(X_1, X_2, X_3)$ | % 98.234 |
| V | $Y' = f(X_1, X_2, X_3, X_4)$ | % 98.237 |

Seçilen, regresyon denklemleri incelendiğinde, set II'den (tek bağımsız değişkenli denklemler) set III'e geçişte determinasyon katsayısı birden artış göstermekte, daha sonra değişken sayısının arttırılmasında ise determinasyondaki artışlar son derece az kalmaktadır. Determinasyon katsayısındaki bu çok küçük artışlar araştırmacı için çok önemli değilse, fazla sayıda değişken ile çalışmak yerine, maliyetleri azaltacağı için set III'den bir denklem seçmek daha akılcıdır. Hangi denklem seçilmelidir?

Sadece set III gözönüne alındığında, determinasyon katsayısı en yüksek olan $Y' = F(X_1, X_2)$ seçilebilir ancak bu denklemin seçimi set II ile tutarsızlık oluşturmaktadır. Çünkü tek değişkenli denklemlerde en yüksek korelasyon $Y' = f(X_4)$ denklemindedir. Bu durumda, $Y' = f(X_1, X_4)$ tercih edilmelidir.

Bütün mümkün regresyonların oluşturulması kesin bir çözüm getirmemekte ve bulunan alternatif çözümler arasından subjektif olarak bir denklem belirlenmektedir. Bu yönteme göre, elde edilen denklem

$$Y' = f(X_1, X_4) = 103.10 + 1.44\, X_1 - 0.61\, X_4$$

veya

$$Y' = f(X_1, X_2) = 52.58 + 1.47\, X_1 + 0.66\, X_2$$

olmalıdır.

## III. GERİYE DOĞRU ELİMİNASYON YÖNTEMİ

Bütün regresyonlar yönteminin biraz daha geliştirilmiş şeklidir. İleriye doğru seçim yönteminin ise karşıtıdır [8;458 ]. Yöntemin temel aşamaları şunlardır [6]:

1. Bütün değişkenleri içeren regresyon denklemi kurulur.

2. Regresyon denklemindeki değişkenler için F kısmi test ( veya t-test) değerleri hesaplanır.

3. Hesaplanan kısmi F değerleri arasında en düşük değer, önceden belirlenen anlamlılık seviyesindeki Fk kritik değeri ile karşılaştırılır. Eğer F<Fk ise, F'nin ait olduğu değişken regresyon denkleminden elimine edilir.

4. Model geriye kalan değişkenler için kurulur ve elimine edilecek aday değişken, 2. adımdaki yol ile yeniden tanımlanır. Proses tekrarlanır.

5. Proses, elimine edilecek bağımsız değişken kalmayıncaya kadar sürdürülür. Yani, bütün kısmi F değerleri, kritik Fk değerlerinden büyük veya eşitse (F≥Fk) regresyon denklemi olduğu gibi kabul edilir.

Yöntemi yine Hald'ın datasına dayanan örnek problem [7;395] üzerinde inceleyelim:

**i.** $Y' = f(X_1, X_2, X_3, X_4)$ denklemi kurulur. Kısmi F test = $t^2 = (b/\sigma_b)^2$ olmak üzere bütün değişkenler için kısmi F değerleri hesaplanır.

| Değ. | β katsayısı | Standart Hata | KısmiF |
|---|---|---|---|
| $X_4$ | -0.1440 | 0.7090 | 0.0412 |
| $X_3$ | 0.1019 | 0.7547 | 0.0182 |
| $X_2$ | 0.5101 | 0.7237 | 0.4968 |
| $X_1$ | 1.5511 | 0.7447 | 4.3375 |

$\alpha=0.10$ anlamlılık seviyesi için Fk değeri, $F_{(1,8;0.90)} = 3.46$ dır. En küçük kısmi F değeri $X_3$ değişkenine aittir. Kısmi F< Fk olduğundan $X_3$ red edilir, yani $X_3$ denklemden elimine edilir.

**ii.** Regresyon denklemi $Y' = f(X_1, X_2, X_4)$ haline gelmiştir.

**ANOVA:**

| Değişkenlik Kaynağı | Sapmaların Kareleri | Serbestlik Derecesi |
|---|---|---|
| Açıklanan | 2667.79 | k-1=3 |
| Açıklanamayan | 47.97 | n-k=9 |
| Toplam | 2715.76 | n-ı=12 |

F= 166.832

$\alpha=0.001$ için, F kritik değer = $F_{(3,9;0.999)} = 13.90$ ve $13.90<166.832$ olduğundan, bu üç bağımsız değişkene göre kurulan regresyon denklemi istatistiki açıdan anlamlıdır. Yani daha önceki adımda, X3'ü denklemden elimine etmiş olmak, denklemin anlamlılığını bozmamıştır. Şimdi bu aşamada, elenebilecek başka bir değişken var mıdır? Bunu belirleyebilmek için yeniden kısmi F testi uygulanır:

| Değ. | β katsayısı | Standart Hata | Kısmi F |
|---|---|---|---|
| $X_2$ | 0.4161 | 0.1856 | 5.0258 |
| $X_1$ | 1.4519 | 0.1169 | 154.0080 |
| $X_4$ | -0.2365 | 0.1732 | 1.8632 |

Bu denklemde en küçük kısmi F değeri $X_4$ değişkenine ait olup, Fk kritik değeri F (1,9; 0.90) = 1.8632 olduğundan $X_4$ değişkeni red edilir.

**iii.** Regresyon denklemi $Y' = f(X_1, X_2)$ halini alır. Bu denklemin anlamlı olup olmadığının denetimi gerekir.

**ANOVA:**

| Değişkenlik Kaynağı | Sapmaların Kareleri | Serbestlik Derecesi |
|---|---|---|
| Açıklanan | 2657.85 | k-1=2 |
| Açıklanamayan | 57.904 | n-k=10 |
| | F=229.504 | |

$\alpha=0.001$ için, F kritik değer = $F_{(2,10;0.999)} = 14.91 < 229.504$ olduğundan çoklu korelasyon katsayısının sıfırdan anlamlı şekilde farklı olduğu hipotezi kabul edilir.

Şimdi kısmi korelasyonların anlamlılığını inceleyerek denklemden indirgenmesi gereken değişken olup olmadığını araştıralım.

| Değ. | β katsayısı | Standart hata | Kısmi F |
|---|---|---|---|
| $X_2$ | 0.6622 | 0.458 | 208.582 |
| $X_1$ | 1.4683 | 0.1213 | 146.522 |

Burada kısmi F değerlerinin karşılaştırılacağı kritik F değeri , Fk = F(1, 10; 0.90) = 3.29 dur ve bu değer kısmi F değerlerinin ikisinden de küçük olduğundan $X_2$ ve $X_1$ değişkenleri anlamlıdır ve denklemde kalmalıdır. Eliminasyona devam edilmediğine göre yöntem durdurulur. Bu yöntemle ulaşılan sonuç:

$$Y' = f(X_1, X_2) = 52.58 + 1.47\,X_1 + 0.66\,X_2$$

denklemidir ve regresyon ilişkisi olarak kabul edilir.

## IV. İLERİYE DOĞRU SEÇİM YÖNTEMİ

Bu yöntem, geriye doğru eliminasyon yönteminin tam tersine, en az değişken içeren denklemle, yani tek değişkenli model ile başlamayı gerektirir. Stepwise yönteminin basitleştirilmiş bir versiyonudur [8;458].

Y bağımlı değişken ile en yüksek korelasyonu veren değişken için $Y' = f(X_i)$ başlangıç denklemi kurulur [6; 214]. Bundan sonraki aşamalarda da sırasıyla denkleme girmesi gereken değişkenler bazı incelemeler sonucu denkleme alınır. Her değişkenin denkleme alınmasından sonra aşağıdaki işlemler gerçekleştirilir:

1. Elde edilen denklemin, çoklu korelasyon katsayısının karesi ($R^2$) hesaplanarak, anlamlılığı test edilir.

2. Denkleme son alınan değişkenin bağımlı değişkeni açıklamak da anlamlı bir değişme sağlayıp sağlamadığını test edebilmek için, sözkonusu değişkenin kısmi F değeri hesaplanır. Eğer test sonucu, belli bir analamlılık düzeyi için değişken red edilirse, yöntem durdurulur [6;215].

Yöntemi yine Hald'ın datası [7;373] üzerinde inceleyelim:

**i.** $X_5$ = Y olmak üzere, $X_5$ ile bağımsız değişkenler arasındaki korelasyon katsayıları hesaplanır. Bunlar sırasıyla,

$r_{15} = 0.7307$   $r_{25} = 0.8162$   $r_{35} = 0.5346$   $r_{45} = 0.8213$

olarak bulunur. En yüksek korelasyon $X_4$ ile $X_5$ arasında olduğundan başlangıç denklemi $X_4$'e bağlı olarak kurulacaktır.

**ii.** Y' = f ($X_4$) denklemi kurulur. Denklemin anlamlı olup olmadığı F testi ile araştırılır. Bu denklem anlamlıdır.

**iii.** Regresyonda olmayan bütün değişkenlerin kısmi korelasyon katsayıları hesaplanarak, denkleme girmesi gereken yeni değişken araştırılır.

$X_1 : r_{15.4} = 0.91541$

$X_2 : r_{25.4} = 0.01696$

$X_3 : r_{35.4} = 0.80117$

En yüksek korelasyonu $X_1$ verdiği için, denkleme ilave edilecek değişken $X_1$ dir.

**iv.** Denklem Y' = f($X_4$ , $X_1$) şeklindedir.

Bu denklem için $R^2$ = 0. 972 dir. Bu katsayı için test yapılırsa , F=176.63 bulunur. $\alpha$=0.001 anlamlılık düzeyinde kritik değer Fk= $F_{(2,\ 10;\ 0.999)}$ =14.91 dir. F>Fk olduğundan $R^2$ anlamlıdır. Acaba $X_1$'in denkleme alınmış olması, Y'nin açıklanmasında yani açıklanan değişkenlikde bir artış sağlamış mıdır?

$X_1$ için kısmi F=108.22 olup kritik değer Fk= $F_{(1,\ 10;\ 0.999)}$ =21.04 ile karşılaştırısak, Fk<F olduğundan kısmi regresyon katsayısının sıfırdan anlamlı bir şekilde farklı olduğu kararı verilir, yani $X_1$'in denkleme alınması iyi bir yaklaşım sağlayacaktır.

**v.** Bu aşamada acaba hala denkleme alınması gereken değişken var mıdır? Regresyonda olmayan değişkenlerin kısmi korelasyon katsayıları hesaplanırsa

$r^2_{25.14} = 0.3583$   $r^2_{35.14} = 0.3200$

bulunduğundan denkleme alınması gereken değişken $X_2$ dir.

**vi.** Regresyon denklemi $X_2$'nin ilavesi ile Y' = f ($X_4$ , $X_1$ , $X_2$ ) dir. Bu denklem için çoklu korelasyon katsayısını karesi, $R^2$ = 0.9823 dür. $X_2$'nin ilavesinden önceki denklem için $R^2$ =0.972 olduğuna göre bağlı değişkendeki belirlilik artmıştır. $X_2$ için $\alpha$=0.10 anlamlılık düzeyinde kısmi F testi yapılırsa kritik F değeri F(1, 9; 0.90 )=3.36; kısmi F=5.03 olduğundan $X_2$'nin regersyona girmesi anlamlı olmuştur. Şimdi regresyonda olmayan $X_3$'ün denkleme alınıp alınmayacağının araştırılması gerekir. $X_3$ için $r^2_{35.124}$ =0.00227 olup oldukça küçük bir değerdir.

**vii.** Son değişken $X_3$'ü de regresyona aldığımızda Y' = f ($X_4$ , $X_1$ , $X_2$ , $X_3$ ) olacaktır.

$X_3$ için kısmi F=0.018; kritik F ise $F_{(1,\ 8;\ 0.90)}$ =3.46 olup, F<Fk olduğundan $X_3$ değişkeni red edilir, yani $X_3$ değişkeninin denkleme alınması açıklanan değişkenlikde artış sağlamayacaktır.

Şu halde yöntem sonucunda varılan regresyon denklemi Y'= F ($X_4$ , $X_1$ , $X_2$ ) şeklinde olacaktır. Regresyon denklemi

$$Y' = 71.65 - 0.24\ X_4 + 1.45\ X_1 + 0.42\ X_2 \quad \text{dir.}$$

## V. STEPWISE REGRESYON YÖNTEMİ

Bu yöntem farklı isim taşımasına rağmen daha önce incelene ileriye doğru seçim yönteminin geliştirilmiş şeklidir [6;217; 7;171; 11;495; 12;671], hatta bazı paket programlarda ileriye doğu stepwise (forward stepwise) olarak anılmaktadır [8;453]. İleriye doğru seçimden farklı olan aşaması, her adımda daha önceden denkleme sokulan değişkenlerin yeniden gözden geçirilmesidir. Çünkü bağımsız değişkenlerden biri tek başına en yüksek belirlik katsayısını vermesine rağmen, yeni bağımsız değişkenlerin ilavesinden sonra, bağlı değişkeni açıklamakda anlamsız hale gelebilmektedir. Böyle durumlarla karşılaşıldığında, gereksiz değişken bulundurmamak açısından yeniden gözden geçirmek yararlı olmaktadır. Bu da kısmi F testi ile gerçekleştirilen bir aşamadır. Anlamlılığı bozulan değişken regresyon denkleminden atılır. Halbuki ileriye doğru seçim yönteminde bir defa denkleme alınan değişken, o aşamadan sonra daima denklemde kalmaktadır[11;495; 12;672].

Yöntemi yine Hald'ın datası [10;647; 7;397-402] için inceleyelim:

Step 1. Yöntem basit bir korelasyon matrisinin kurulması ile başlar. Bu matriste en yüksek korelasyonu veren $X_i$ bağımsız değişkeni regresyona alınır. Bu örnek için, ileriye doğru seçim yönteminde olduğu gibi ilk değişken $X_4$ olacaktır.

Step 2. Denklem Y' = f (X4) dür. Kısmi korelasyon katsayılarını kullanarak. en yüksek korelasyonu veren $X_1$ değişkeni, denkleme girecek ikinci değişken olarak belirlenir.

Step 3. Bu adımda, denklem $Y' = f(X_4, X_1)$ haline gelir. Bu yöntem ileriye doğru seçim yönteminin yapmadığı bir test gerçekleştirir. Şöyle ki denklem $X_1$'e bağlı iken, ikinci giren değişken $X_4$ olursa anlamlı olabilir mi?

$X_4$ için kısmi F değeri 159.295 olup, $\alpha = 0.05$ anlamlılık düzeyinde istatistiki olarak önemli olduğundan, $X_4$ denklemde kalmalıdır. Bundan sonra denkleme alınacak değişken ise

$$r^2_{25.14} = 0.358$$

olduğundan $X_2$ değişkenidir.

Step 4. Denklem $Y' = f(X_4, X_1, X_2)$ şeklinde üç değişkene bağlı hale gelir. Burada. denklemdeki her değişken için, yine teker teker kısmi F testi yapmak gerekir. $\alpha = 0.10$ için F kritik değeri $Fk = F_{(1, 9; 0.90)} = 3.36$ dır.

$X_2$ için kısmi F =5.025>Fk ve $X_1$ için kısmi F = 154.07>Fk olduğundan $X_2$ ve $X_1$ değişkenlerinin denklemde kalmasının anlamlı olduğuna karar verilir.

$X_4$ için kısmi F = 1.8632<Fk olduğundan $X_4$ red edilir, yani denklemden çıkarılır. İşte burada da görüldüğü gibi, denkleme alınması anlamlı gibi görünen bir değişkenin, yeni bir değişken ilave edilmesi durumunda, anlamlılığı azalıyorsa denklemden çıkarılabilmektedir. Halbuki ileriye doğru seçim yönteminde bir değişken denkleme bir kez alındıktan sonra bir daha çıkarılması sözkonusu olmamaktadır.

Step 5. Kalan tek değişken $X_3$ dür. Kısmi F testi sonucu $X_3$ de red edilir ve yöntem durdurulur. Regresyon denklemi, $Y' = f(X_1, X_2)$ şeklindedir. Regresyon denklemi, tahmin edilen katsayılarla ifade edilirse,

$$Y' = 52.58 + 1.47 X_1 + 0.66 X_2 \text{ dir.}$$

## VI. İRDELEME VE SONUÇ

İncelenen dört yöntem de ancak bilgisayar aracılığı ile uygulanabilmektedir. Tüm regresyonlar yöntemi pratik yönü olmayan bir yoldur. Geriye doğru eliminasyon yöntemi hem bilgisayar zamanı hem de insan gücü açısından tüm regresyonlar yönteminden daha ekonomiktir. İleriye doğru seçim yöntemi her iki yönteme göre daha iyi bir yaklaşımdır. Fakat bu yöntemi tamamlayıcı aşamalarla geliştirilen stepwise yöntemi her zaman tercih edilir [6, 7, 8, 11, 12].

Örnek problem üzerinde, farklı yöntemler sonucu birbirinden biraz farklı denklemler elde edildi.

1. $Y' = 52.58 + 1.47X_1 + 0.66 X_2$

Bu denklem, tüm regresyonlar, geriye eliminasyon ve stepwise yöntemleri sonucu seçildi.

2. $Y' = 103.10 + 1.44 X_1 - 0.61 X_4$

denklemi tüm regresyonlar yöntemi ile ikinci seçenek olarak belirlendi.

3. $Y' = 71.65 + 1.45 X_1 + 0.42 X_2 - 0.24 X_4$ ileriye doğru seçim yöntemi ile belirlendi.

En iyi regresyon denklemini seçmek için tüm mümkün regresyonları oluşturup incelemek daha iyi sonuç vermektedir, ancak maliyetler gözönüne alındığında pek tercih edilememektedir. Bu yöntemler içinde her zaman ilk tercih, bilgisayar paket programlarında da sık rastlanan stepwise yöntemidir. İkinci tercih ise geriye doğru eliminasyon yöntemi olacaktır.

## KAYNAKLAR


[1] Serper, Ö., ve Gürsakal, N., *Araştırma Yöntemleri*, Filiz Kitapevi, İstanbul, 1989.

[2] Bağırkan, Ş., *İstatiksel Analiz*, Şampiyonlar Baskı ve Yayın Organizasyonu, İstanbul, 1978.

[3] Bağırkan. Ş., *İstatistiğe Giriş*, İ.İ.T.İ.A. Nihad Sayar Yayın ve Yardım Vakfı Yayınları No:337/570, İstanbul, 1980.

[4] *TSP User's Manual* Version 7.0, Quantitative Micro Software, Irvine, California.

[5] *SPSS$^X$ User's Guide*, Sec. Ed., SPSS Inc., Chicago.

[6] Morrison, D.F., *Applied Linear Statistical Methods*, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1983.

[7] Draper N. and Smith H., *Applied Regression Analysis*, John Wiley &Sons, New York, 1966.

[8] Neter, J., Wasserman, W., Kutner, M. H., *Applied Linear Statistical Modals*, Third Ed., Richard D. Irwin, Inc., Boston, 1990.

[9] Serper, Ö., *Uygulamalı İstatistik*-1, Filiz Kitapevi, İstanbul, 1985.

[10] Hald, A., *Statistical Theory with Engineering Applications*, John Wiley &Sons, Inc.,New York, 1962.

[11] Ostle, B. & Malone, L.C., *Statistics In Research: Basic Concepts and Techniques for Research Workers*, Fourth Ed. Iowa State University Press/Ames, Iowa, 1988.

[12] McClave, J.T., and Benson P.G., Statistics for Business and Economics, Fifth Ed., Maxwell Macmillian International Editions, Singapore, 1991.