BÖLÜM 3

ÇOKLU REGRESYON ANALİZİ: TAHMİN

KISIM 1
YATAY-KESİT VERİLER İLE REGRESYON ANALİZİ

İÇİNDEKİLER

Çoklu regresyon analizi, aynı zamanda bağımlı değişkeni etkileyen diğer birçok faktörü açıkça kontrol etmemize imkân sağladığından ceteris paribus analizine daha uygundur.

Deneysel olmayan veriye güvenmemiz gerektiğinde hem iktisat teorilerini test etme hem de politik etkileri değerlendirme için bu önemlidir.

1. ÇOKLU REGRESYON KURAMI

- İKİ BAĞIMSIZ DEĞİŞKENLİ MODEL
- K BAĞIMSIZ DEĞİŞKENLİ MODEL

Genel **çoklu doğrusal regresyon modeli** (aynı zamanda *çoklu regresyon modeli* de denilen) anakütlede

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \ldots + \beta_k x_k + u \qquad \boxed{3.6}$$

olarak yazılabilir. Burada

β_0, **kesim parametresi**dir.
β_1, x_1 ile ilişkili parametredir.
β_2, x_2 ile ilişkili parametredir ve bu şekilde devam eder.

Çoklu Regresyon Terminolojisi

y	$x_1, x_2, \ldots, x_k$
Bağımlı değişken	Bağımsız değişkenler
Açıklanan değişken	Açıklayıcı değişkenler
Tepki değişkeni	Kontrol değişkenleri
Öngörülen değişken	Öngören değişkenler
Bağlanan değişken	Açıklayıcılar

2. SIRADAN EN KÜÇÜK KARELERİN İŞLEYİŞİ VE YORUMU

- SEKK TAHMİNLERİNİN ELDE EDİLMESİ
- SEKK REGRESYON DENKLEMİNİN YORUMLANMASI
- ÇOKLU REGRESYONDA “DİĞER FAKTÖRLERİ SABİT TUTMA”NIN ANLAMI
- EŞ ANLI OLARAK BİRDEN FAZLA DEĞİŞKENİN DEĞİŞTİRİLMESİ
- SEKK’NİN TEORİK DEĞERLERİ VE ARTIKLAR
- ÇOKLU REGRESYONUN “KISMİ ARINDIRMA (PARTİALLİNG OUT)” YORUMU
- BASİT VE ÇOKLU REGRESYON TAHMİNLERİNİN KARŞILAŞTIRILMASI
- UYUM İYİLİĞİ

• ORİJİNDEN GEÇEN REGRESYON

Bazen bir iktisadi teori ya da öngörü, b_0'ın sıfır olması gerektiğini söylemektedir. Bu nedenle, kesim parametresi sıfır olduğunda SEKK tahmininden kısaca bahsetmemiz gerekir. Özellikle,

$$\tilde{y} 5 \tilde{b}_1 x_1 \; 1 \tilde{b}_2 x_2 \; 1 \ldots 1 \tilde{b}_k x_k \qquad (3.30)$$

şeklinde bir denklemi araştırmalıyız. Burada tahminler üzerindeki "~" sembolü, kesim parametresi ile birlikte elde edilen SEKK tahminlerinden onları ayırmak için kullanılmaktadır [(3.11)'deki gibi]. (3.30)'da $x_1 = 0$, $x_2 = 0$, …, $x_k = 0$ olduğunda öngörülen değer sıfırdır. Bu durumda $\tilde{b}_1, \ldots, \tilde{b}_k$, y'nin $x_1, x_2, \ldots, x_k$ ile orijinden geçen regresyonundan elde edilen SEKK tahminleri olarak ifade edilir.

3. SEKK TAHMİNCİLERİNİN BEKLENEN DEĞERİ

Varsayım ÇDR.1 (Parametrelerde Doğrusallık)

Anakütlede modeli,

$$y \ 5 \ b_0 \ 1 \ b_1x_1 \ 1 \ b_2x_2 \ 1 \ \ldots \ 1 \ b_kx_k \ 1 \ u \qquad \boxed{3.31}$$

olarak yazılabilmektedir. Burada b_0, b_1, …, b_k ilgilenilen bilinmeyen parametrelerdir (sabitler) ve u gözlemlenmeyen bir rassal hata ya da bozulma terimidir.

Varsayım ÇDR.3 (Tam Bağlılığın Olmaması)

Örneklemde (ve bu nedenle anakütlede), bağımsız değişkenlerin hiçbiri sabit değildir ve bağımsız değişkenler arasında kesin doğrusal ilişkiler yoktur.

Varsayım MLR.4 (Koşullu Sıfır Ortalama)

u hata teriminin, bağımsız değişkenlerin herhangi bir değeri verildiğinde beklenen değeri sıfıra eşittir. Diğer bir deyişle,

$$E(u|x_1, x_2, ..., x_k) = 0 \quad (3.36)$$

olacaktır.

Teorem 3.1 (SEKK'nin Sapmasızlığı)

ÇDR.1-ÇDR.4 varsayımları altında,

Anakütle parametresi b_j'nin her bir değeri için,

$$E(\hat{b}_j) \ 5 \ b_j, \ j \ 5 \ 0, 1, \ldots, k \qquad \boxed{3.37}$$

olacaktır. Diğer bir deyişle, SEKK tahmincileri anakütle parametrelerinin sapmasız tahmincileridir.

- **BİR REGRESYON MODELİNE İLGİSİZ DEĞİŞKENLERİN İLAVE EDİLMESİ**
- **İHMAL EDİLMİŞ DEĞİŞKEN SAPMASI: BASİT BİR DURUM**
- **İHMAL EDİLEN DEĞİŞKEN SAPMASI: DAHA GENEL DURUMLAR**

TABLO 3.2

Tahmin Edilen Denklemde x^2 İhmal Edildiğinde $\tilde{\beta}_1$'deki Sapmanın Özeti (3.40)

	$\mathrm{Corr}(x_1,x_2) > 0$	$\mathrm{Corr}(x_1,x_2) < 0$
$\beta_2 > 0$	Pozitif sapma	Negatif sapma
$\beta_2 < 0$	Negatif sapma	Pozitif sapma

4. SEKK TAHMİNCİLERİNİN VARYANSI

Varsayım ÇDR.5 (Sabit Varyans)

Açıklayıcı değişkenlerin herhangi bir değeri verildiğinde, hata terimi **u** aynı varyansa sahiptir. Diğer bir deyişle, Var(**uu**x_1, ..., x_k) 5 s^2'dir.

Teorem 3.2 (SEKK Eğim Parametresi Tahmincilerinin)

ÇDR.1-ÇDR.5 varsayımları altında, bağımsız değişkenlerin örnekleme değerleri üzerine belirlenen kısıt,

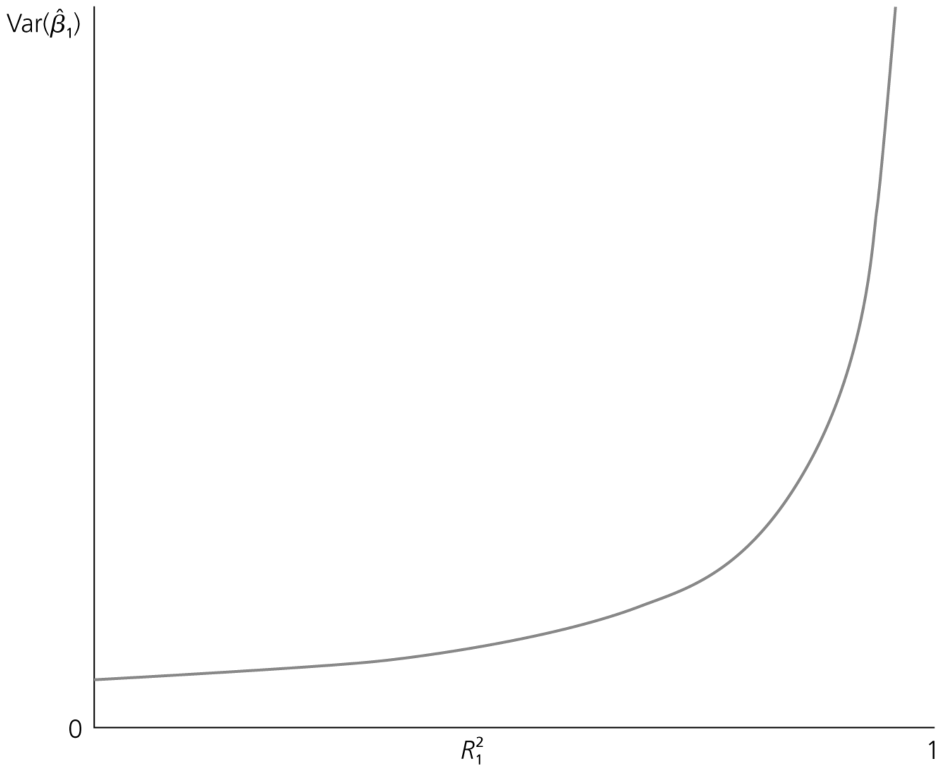
$$\mathrm{Var}(\hat{\beta}_j) = \frac{\sigma^2}{SST_j(1 - R_j^2)}, \tag{3.51}$$

olacaktır. $j = 1, 2, \ldots, k$ için, burada $SST_j = \sum_{i=1}^{n}(x_{ij} - \bar{x}_j)^2$'deki toplam örnekleme değişimidir ve R_j^2, x_j'nin diğer tüm bağımsız değişkenler üzerinden (ve bir kesme parametresi dâhil) regresyonundan elde edilen R-karedir.

- **SEKK VARYANSLARININ BİLEŞENLERİ: ÇOKLU DOĞRUSAL BAĞLILIK**
- **EKSİK BELİRLENMİŞ MODELLERDE VARYANSLAR**
- **σ^2'NİN TAHMİN EDİLMESİ: SEKK TAHMİNCİLERİNİN STANDART HATALARI**

ŞEKİL 3.1

R_1^2'nin fonksiyonu olarak Var($\hat{\beta}_1$)



5. SEKK'NİN ETKİNLİĞİ: GAUSS-MARKOV TEOREMİ

Teorem 3.4 (Gauss- Markov Teoremi)

ÇDR.1-ÇDR.5 varsayımları altında, $\hat{\beta}_0$, $\hat{\beta}_1$, …, $\hat{\beta}_k$ sırasıyla β_0, β_1, …, β_k'nin en iyi doğrusal sapmasız tahmincileridir (EİDST).