

**ÇOKLU REGRESYON MODELİNDE HATA TERİMLERİNİN İKİ  
FARKLI DAĞILIM VARSAYIMI ALTINDA ANALİZİ**

**Demet HAN**


**YÜKSEK LİSANS TEZİ  
(İSTATİSTİK)**

**GAZİ ÜNİVERSİTESİ  
FEN BİLİMLERİ ENSTİTÜSÜ**




**Haziran 2004**

**ANKARA**

Demet HAN tarafından hazırlanan ÇOKLU REGRESYON MODELİNDE HATA TERİMLERİNİN İKİ FARKLI DAĞILIM VARSAYIMI ALTINDA ANALİZİ adlı bu tezin Yüksek Lisans tezi olarak uygun olduğunu onaylarım.

  
Prof. Dr. Müslim EKNİ  
Tez Yöneticisi

Bu çalışma, jürimiz tarafından İstatistik Anabilim Dalında Yüksek Lisans tezi olarak kabul edilmiştir.

Başkan : Prof. Dr. Hülya Bayrak   
Üye : Prof. Dr. Fikri Öztürk   
Üye : Prof. Dr. Müslim Eknî   
Üye : \_\_\_\_\_  
Üye : \_\_\_\_\_

Bu tez, Gazi Üniversitesi Fen Bilimleri Enstitüsü tez yazım kurallarına uygundur.



## İÇİNDEKİLER

	Sayfa
ÖZET.....	i
ABSTRACT.....	ii
TEŞEKKÜR.....	iii
ŞEKİLLERİN LİSTESİ.....	iv
EKLERİN LİSTESİ.....	v
SİMGELER VE KISALTMALAR.....	vi
1. GİRİŞ.....	1
2. BAYES ANALİZ.....	3
2.1. Bilgi Verici Olmayan Önsel Dağılımlar.....	7
3. BAZI ÖNEMLİ DAĞILIMLAR VE KAVRAMLAR.....	15
3.1. Çok Değişkenli Student-t Dağılımı.....	15
3.2. Ters Gamma Dağılımı.....	16
3.3. Bilinen en Küçük Kareler Tahmin Edicisi ve Özellikleri.....	17
3.4. Gauss Markov Teoreminin Özel bir Durumu.....	19
4. ÇOK DEĞİŞKENLİ STUDENT-t HATA TERİMLİ DOĞRUSAL REGRESYON MODELİNİN BAYES VE KLASİK YÖNTEMLE ANALİZİ.....	23
4.1. Olabilirlik Fonksiyonu ve Örneklem Teorisi.....	24
4.2. Çok Değişkenli Student-t Hata Terimli Regresyon Modeli.....	26
4.3. Bayes Analizi.....	34
5. DOĞRUSAL REGRESYON MODELİNDEKİ HATA TERİMLERİNİN BAYES ANALİZİ.....	39
5.1. Regresyon Modeli ve Parametrelerin EÇOB Tahmin Edicileri.....	39
5.2. Parametrelerin Bayes Analizi.....	42

6. HATA TERİMLERİNİN İKİ FARKLI DAĞILIM VARSAYIMI ALTINDA ANALİZ SONUÇLARININ ÖZETİ.....	44
7. SONUÇ VE ÖNERİLER.....	48
KAYNAKLAR.....	50
EKLER.....	52
ÖZGEÇMİŞ.....	57



**ÇOKLU REGRESYON MODELİNDE HATA TERİMLERİNİN İKİ FARKLI  
DAĞILIM VARSAYIMI ALTINDA ANALİZİ  
(Yüksek Lisans Tezi)**

**Demet HAN**

**GAZİ ÜNİVERSİTESİ  
FEN BİLİMLERİ ENSTİTÜSÜ  
Haziran 2004**

**ÖZET**

Bu çalışmada, Bayes analiz yöntemi, çok değişkenli Student-t dağılımı, ters Gamma dağılımı, en küçük kareler tahmin edicisi ve özellikleri tanıtılmıştır. İki farklı dağılım varsayımına dayanan çoklu regresyon modelinin Bayes analizleri yapılmış, parametrelerin en çok olabilirlik tahmin edicileri bulunmuş ve analiz sonuçlarının yorumları verilmiştir.

**Bilim Kodu : 406.01.00**  
**Anahtar Kelimeler : Bayes Analizi, Bilgi İçermeyen Önsel Dağılım, Çoklu Regresyon Modeli, Ters Gamma Dağılımı, Çok Değişkenli Student-t Dağılımı.**  
**Sayfa Adedi : 57**  
**Tez Yöneticisi : Prof.Dr. Müslim EKNİ**

**THE ANALYSIS OF ERROR TERMS BASED ON TWO DIFFERENT  
DISTRIBUTIONS IN THE MULTIPLE REGRESSION MODEL**

**(M.Sc. Thesis)**

**Demet HAN**

**GAZI UNIVERSITY**

**INSTITUTE OF SCIENCE AND TECHNOLOGY**

**June 2004**

**ABSTRACT**

**In this study, Bayesian analysis method, multivariate Student-t distribution, inverted Gamma distribution, least square estimator and their properties are introduced. The Bayesian analysis of multiple regression which is based on two different distribution hypothesis, are examined. Finally, maximum likelihood estimators of parameters are found and the interpretations of the results of the analysis are given.**

**Science Code : 406.01.00**

**Key Words : Bayesian Analysis, Noninformative Prior Distribution, Multiple Regression Model, Inverted Gamma Distribution, Multivariate Student-t Distribution.**

**Page Number: 57**

**Adviser : Prof.Dr.Müslim EKNİ**

## TEŐEKKÜR

Çalıőmalarım boyunca deęerli yardım ve katkılarıyla beni yönlendiren Hocam Prof. Dr. Müslim EKNİ'ye teşekkür ederim.



## ŞEKİLLERİN LİSTESİ

Şekil	Sayfa
Şekil 2.1. Üç farklı veri setine bağlı olabilirlik fonksiyonu.....	8
Şekil 2.2. Üç farklı olabilirlik fonksiyonunun $\alpha$ yükseklik değeri.....	9
Şekil 2.3. Üç farklı olabilirlik fonksiyonunun $\alpha$ yükseklik değeri.....	10



**EKLERİN LİSTESİ**

<b>Ek</b>	<b>Sayfa</b>
Ek-1. Bölüm Dörtteki Bazı Eşitliklerin İspatları.....	53
Ek-2. İnternet Kaynakları .....	58



## SİMGELER VE KISALTMALAR

Bu çalışmada kullanılmış bazı simgeler ve kısaltmalar, açıklamaları ile birlikte aşağıda sunulmuştur.

<b>Simgeler</b>	<b>Açıklama</b>
<b><math>r(X)</math></b>	X matrisinin rankı
<b>oyf</b>	Olasılık Yoğunluk Fonksiyonu
<b>Eş.</b>	Eşitlik
<b>Bkz.</b>	Bakınız

<b>Kısaltmalar</b>	<b>Açıklama</b>
<b>EÇOB</b>	En Çok Olabilirlik
<b>EKK</b>	En Küçük Kareler

## 1. GİRİŞ

Bağımsız değişkenler ile bağımlı değişken arasındaki doğrusal ilişkiyi açıklayan regresyon modeli genelde belirli bir örneklemden yararlanılarak elde edilir. Burada asıl amaç doğrusal regresyon modelinin parametrelerini elde etmektir. Bu noktada iki yöntem izlenebilir. Birincisi sadece örneklemdaki bilgiyi kullanan klasik analiz yöntemi, ikincisi ise parametrelere ilişkin önsel dağılımlar konusunda bazı varsayımlarda bulunarak bunu örneklem bilgisiyle birleştirip parametrelerin sonsal dağılımını belirleyen Bayes analiz yöntemidir.

Hata terimlerindeki normallik, bağımsızlık ve eşvaryanslılıktan ortaya çıkan değişiklikler genellikle çeşitli yöntemlerle belirlenebilir. Daha önce yapılan çalışmalar bu doğrultuda olmuştur. Zellner ve Tiao (1) bağımsızlık koşulu sağlanmayan regresyon modelinin Bayes analizini yapmıştır. Mouchart (2), çalışmasında, bilinmeyen ortalamaları tahmin ederek hata terimlerini elde etmiştir. Uygulamada, hata terimleri Normallik varsayımını sağlamayabilir. Birçok araştırmacı bunun farkına varmıştır. Bunların içinde Newcomb (3), Jeffreys (4), Fama ve Roll (5), Box ve Tiao (6), Praetz (7) gösterilebilir. Newcomb ve Jeffreys çalışmalarında, dağılımların kesikli karışımlarını kullanmışlardır. Box ve Tiao güç dağılımlarını analiz ettiğinde Fama ve Roll durağan Paretian dağılımlar sınıfını dikkate alır. Praetz tarafından dikkate alınan Normal dağılımla sürekli dağılımların karışımı, bağımsız her biri aynı tek değişkenli Student-t dağılımına sahip hata terimlerine götürür. Ayrıca Jeffreys (4) bağımsız, her biri aynı tek değişkenli Student-t dağılımına sahip hata terimli regresyon modelinin analiz sonuçlarını vermiştir.

Bu çalışma hata terimlerinin iki farklı dağılım varsayımına göre çoklu regresyon modelinin Bayes analizine dayanmaktadır.

İkinci bölümde Bayes analizinin genel yapısı üzerinde durulacaktır. Bilgi içermeyen önsel dağılım ve sonsal dağılım hakkında bilgiler verilecektir. Bayes analizi istatistiksel çıkarsama ve karar vermede klasik yaklaşımdan örneklem bilgisini

kullanması ile farklı olan bir yaklaşımdır. Örneklem bilgisi elde edilir ve öteki bilgilerle birleştirilir. Bilgiyi birleştirme yöntemi Bayes teoremidir. Sonuçta elde edilen bilgi birleşimi, sonuç çıkarsama ve karar verme için bir taban oluşturur. Üçüncü bölüm, çoklu regresyon modelinin analizinde geçen bazı dağılımlar ve kavramlardan Çok Değişkenli Student-t dağılımı, Ters Gamma dağılımı, OLS tahmin edicileri ve Gauss Markov teoreminin tanıtılmasını içermektedir. Dördüncü bölümde Çok Değişkenli Student-t dağılımına sahip hata terimli doğrusal regresyon modelinin Bayes ve klasik yöntemle analizi verilecektir. Beşinci ve altıncı bölümlerde, Çok Değişkenli Normal dağılıma sahip hata terimli doğrusal regresyon modelinin Bayes analizi ve bölüm dört ve bölüm beşteki analiz sonuçları verilecektir. Yedinci bölüm olan sonuç bölümünde ise bu iki farklı varsayıma dayanan regresyon modelinin Bayes analiz sonuçları üzerinde durulacaktır.

## 2. BAYES ANALİZİ

Bayes teoremi istatistiksel sonuç çıkarsama yöntemidir. Sonuç, istatistiksel analizin yalnızca bir aşaması olduğundan tekrar modele ve analize dönülür. Bu yüzden araştırmacı araştırmacının adımlarında bu sonucu kullanır.

İstatistiksel sonuç yöntemleri arasındaki fark, Bayes teoreminden elde edilen sonuçtan kaynaklanır. Bilimsel araştırmalar için Bayes analiz kuralı verilerek özellikleri üzerinde durulmuş olunur.

Bayes analizi verideki bütün bilginin kullanılmasıyla gerçekleştirildiğinden verilen model otomatik olarak analiz edilmiş olur. Sonuçların uygunsuzluğu yöntemin yetersiz olmasından çok yaklaşımların uygunsuz olmasından kaynaklanır. Bu sebeple önsel dağılım, modelin bütün bölümlerini içerdiğinden uygun sonuç ortaya çıkar. Bu yöntem herhangi bir olasılık modeline kolayca uygulanabilir. Üzerinde durulan modelin matematiksel uygunluğunun çok az önemi olur bu yüzden modelin bilimsel değeri daha önemlidir. Örnekleme teorisinde karşılaşılan uyumsuz problemlerden dolayı tahmin edicilerin seçimi ve onların güven aralıklarının oluşumu mümkün olmayabilir.

Bayes analizinde, bilgi içeren yada bilgi içermeyen önsel dağılımla ilgili yaklaşımları açık bir şekilde ortaya koymanın yolu yeterli olma koşuludur. Bazı önsel dağılımlar hakkındaki bilgi bütün sistemlerde uygulandığından kullanılabilir. Örneğin veri Normal dağılımı bir önsel dağılıma sahip, her bir gözlemde bağımsız, aynı varyansa sahipse örnekleme teorisi yöntemi ve Bayes analiz yöntemi uygulanabilir. Fakat artıklar üzerindeki bir çalışmadan sonra modelin yetersizliği söz konusu olduğunda ilk modele az miktarda kısıtlama getirilerek verinin tekrar analizi istenebilir. Eğer verinin Normal dağılmadığı tahmin ediliyorsa analizi örnekleme teorisiyle sonuçlandırmak zordur. Bu yüzden Bayes analizi ile bilinmeyen parametre tahminleri kolayca yapılabilir.

Bayes analiz yönteminde klasik analiz yöntemden farklı olarak bilinmeyen  $\theta$  parametresinin  $f(\theta)$  ile gösterilen önsel dağılım ek bilgisine sahip olduğu kabul edilir. Örneklem bilgisi,  $f(\theta|x)$  ile gösterilen  $\theta$  parametresinin sonsal dağılımını bulmak için kullanılır. Bu dağılım  $f(x|\theta)$ ,  $x$  verilerinin oylu olmak üzere

$$f(\theta|x) = \frac{f(x|\theta)f(\theta)}{\int_{-\infty}^{\infty} f(x|\theta)f(\theta)d\theta} \quad [2.1]$$

eşitliği ile bulunabilir. Eş. 2.1 ile verilen ifade

$$f(x) = \int_{-\infty}^{\infty} f(x|\theta)f(\theta)d\theta$$

olmak üzere,

$$f(\theta|x) = \frac{f(x|\theta)f(\theta)}{f(x)} \quad [2.2]$$

biçiminde yazılabilir. Eş. 2.2'de yer alan  $f(x)$  çarpım sabiti olduğundan sonsal dağılımın şekline etki etmez. Bu nedenle

$$f(\theta|x) \propto f(x|\theta)f(\theta) \quad [2.3]$$

olur.  $f(x|\theta)$  ifadesinde  $\theta$  bilinmeyen parametredir. Bu yüzden  $f(x|\theta)$  olabirlik fonksiyonu olarak da adlandırılır ve  $L(\theta|x)$  ile gösterilir. Eş. 2.3'deki Bayes formülü,

$$f(\theta|x) \propto L(\theta|x)f(\theta) \quad [2.4]$$

biçiminde yazılabilir. Bu durumda sonsal dağılımın önsel dağılım ve olabilirlik fonksiyonu ile orantılı olduğu söylenebilir. Olabilirlik fonksiyonu çarpanlarına ayrılarak

$$L(\theta | x) = L'(\theta, S(x))g(x)$$

olarak yazılabilir. Yukarıdaki ifadede yer alan  $S(x)$  örneklemden alınan  $x$  verisinin bir fonksiyonudur ve yeterli istatistik olarak adlandırılır.  $\theta$  parametresiyle ilgili bütün bilgi, bu istatistik ile sağlanır.  $g(x)$  ifadesi de parametreden bağımsız bir sabittir.  $\theta$  parametresinin sonsal dağılımı  $L(\theta | x) = L'(\theta, S(x))g(x)$  olmak üzere,

$$f(\theta | x) = \frac{L'(\theta, S(x))g(x)f(\theta)}{\int L'(\theta, S(x))g(x)f(\theta)d\theta} \quad [2.5]$$

biçiminde yazılabilir.

Yeterli istatistiğin olmadığı durumlarda  $\theta$  parametresiyle ilgili bilgi, olabilirlik fonksiyonunun çarpanları arasında yer almayan Bayes istatistikleri ile sağlanır. Ayrıca örneklemden alınan farklı bir  $x$  verisi sonsal dağılımı sağlarsa önsel dağılım gibi ele alınır ve yeni olabilirlik fonksiyonu  $L(\theta | x)$  ile yeni sonsal dağılımı hesaplamak için kullanılır. Fazla veri orijinal önsel dağılım etkisini azaltırken olabilirlik fonksiyonunu baskın kılar.

Örnek :

$\theta$  parametresi Normal dağılımın bilinmeyen kitle ortalaması ve  $\sigma^2$  parametresi bilinen kitle varyansı olmak üzere  $\theta$  parametresinin önsel dağılımı,

$$f(\theta) = \frac{1}{c}, \quad -c/2 \leq \theta \leq c/2$$

olsun.  $x_1, x_2, \dots, x_n$  değerlerinin olabilirlik fonksiyonu,

$$L(\theta | x) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -1/2 \left[ \frac{x_i - \theta}{\sigma} \right]^2 \right\}$$

$$\propto \exp \left\{ -1/2 \sum_{i=1}^n \left[ \frac{x_i - \theta}{\sigma} \right]^2 \right\}$$

olarak yazılabilir. Yine  $x_i - \theta = (x_i - \bar{x}) + (\bar{x} - \theta)$  eşitliği kullanılarak

$$\sum_{i=1}^n (x_i - \theta)^2 = \sum_{i=1}^n (x_i - \bar{x})^2 + n(\bar{x} - \theta)^2$$

elde edilen ifadede sadece son terim  $\theta$  parametresine dayandığından

$$L(\theta | x) \propto \exp \left\{ -1/2 \left[ \frac{\bar{x} - \theta}{\sigma/\sqrt{n}} \right]^2 \right\}$$

olur.  $L(\theta | x)$  olabilirlik fonksiyonunun dağılımı  $N[\theta, \sigma^2/n]$  ile gösterilen Normal dağılım olur. Bayes analizinden  $\theta$  parametresinin sonsal dağılımı

$$f(\theta | x) \propto \exp \left\{ -1/2 \left[ \frac{\theta - \bar{x}}{\sigma/\sqrt{n}} \right]^2 \right\} \frac{1}{c}, \quad -c/2 \leq \theta \leq c/2$$

olarak bulunur. Yukarıdaki sonsal dağılım budanmış Normal dağılımdır. Çok büyük  $c$  değerleri için  $f(\theta | x) \sim N[\bar{x}, \sigma^2/n]$  olur.

## 2.1. Bilgi Verici Olmayan Önsel Dağılımlar

Bayes analizinde örneklem bilgisini elde etmeden önce, parametreyle ilgili önsel dağılım bilgisinin mevcut olduğu kabul edilir. Çoğu zaman bu bilgi yetersiz olduğundan dolayı aşağıda verilen yöntemlerden biri uygulanır.

- i. Laplace metodu uygulanır ve bütün  $\theta$  parametrelerinin tekdüze dağıldığı kabul edilir.
- ii.  $\theta$  parametresinin sonsal dağılımını elde etmek için olabilirlik fonksiyonu aracılığıyla veriyi sağlayan önsel dağılım seçilir. Bu tür önsel dağılımlar bilgi içermeyen yada belirsiz (vague) önsel dağılımlar olarak adlandırılırlar.

Mad. i'de  $\theta$  parametresi ne olursa olsun önsel dağılımın tekdüze dağılım olduğu kabul edilir. Fakat her zaman bu uygun olmayabilir. Bu durumda bilgi içermeyen önsel dağılım yaklaşımı kullanılır.

$\theta$  parametresinin belirsiz değeriyle ilgili önsel dağılımının tekdüze olduğu kabul edilsin. Bu durumda  $\theta$  parametresinin önsel dağılımı

$$f(\theta) \equiv \text{tekdüze}$$

ile gösterilir.

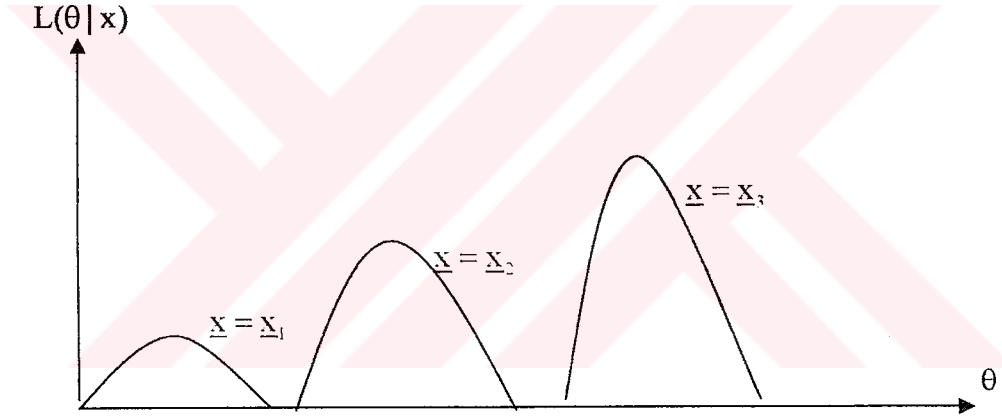
$\beta$ ,  $\theta$  parametresinin herhangi bir fonksiyonu olsun. Eğer  $\theta$  parametresinin önsel dağılımı tekdüze ise genelde  $\beta$  fonksiyonunun önsel dağılımı tekdüze değildir. Bu durumda  $\theta$  parametresinin önsel dağılımının tekdüze olduğu gerçeği görmezlikten gelinir ve başka bir dağılımdan değişken değiştirme tekniği kullanılarak

$$f_{\beta}(\beta) = f_{\theta}(\theta) \left| \frac{d\beta}{d\theta} \right|^{-1}$$

$$= f_{\theta}(f^{-1}(\beta)) \left| \frac{d\beta}{d\theta} \right|^{-1} \quad [2.6]$$

ifadesiyle verilen  $\beta$  fonksiyonu için bir önsel dağılım seçilir.

Mad. ii ile tanımlanan önsel dağılım bir tanedir. Olabilirlik fonksiyonu sonsal dağılım üzerinde baskındır ve sonsal dağılım belirsiz önsel dağılım ve olabilirlik fonksiyonu ile orantılıdır. Bu nedenle olabilirlik fonksiyonunun incelenmesi gerekir. Olabilirlik fonksiyonunun Şekil 2.1 ile gösterilen farklı üç tane veri kümesine dayandığı kabul edilsin.



Şekil 2.1. Üç farklı veri setine bağlı olabilirlik fonksiyonu

Bilgi içermeyen önsel dağılım yaklaşımı için, olabilirlik fonksiyonu ile orantılı olarak alınan  $\theta$  değeriyle ilgili tamamen tarafsız olunması gerekir. Bu nedenle  $\underline{x} = \underline{x}_1$  kadar  $\underline{x} = \underline{x}_2$  yada  $\underline{x} = \underline{x}_3$  ile gösterilen farklı örneklemelerin  $\theta$  değeriyle ilgili tüm bilgiyi içermesi anlamına gelir. Bu nedenle örneklemden alınan olabilirlik fonksiyonlarında tarafsız olunmalıdır. O halde  $\theta$  parametresine ilişkin önsel dağılımın güvenilirliği hakkında ne söylenebilir?

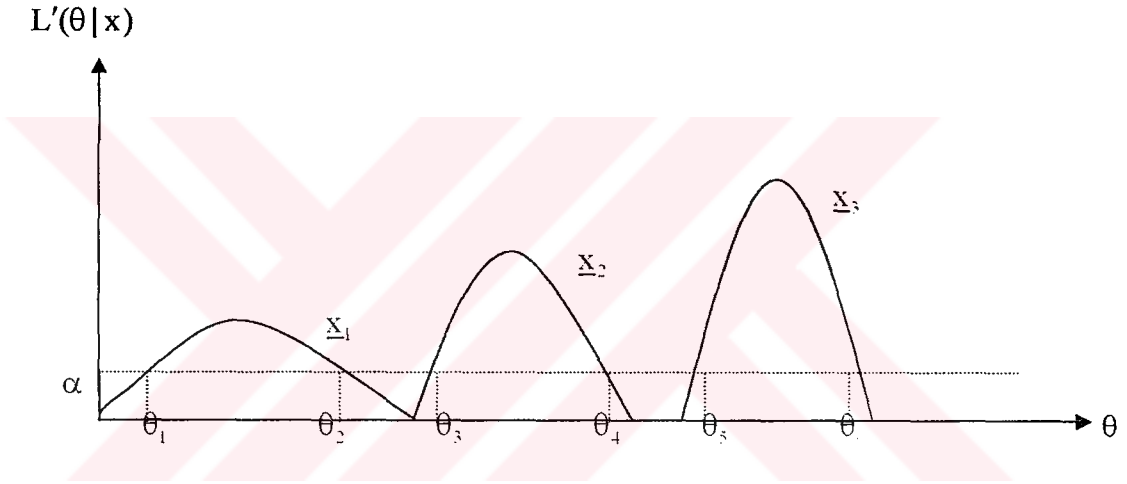
Yukarıdaki grafikte  $L(\theta | \underline{x}_3)$ 'ün yayılması  $L(\theta | \underline{x}_2)$ 'den daha azdır.  $L(\theta | \underline{x}_2)$ 'nin yayılması da  $L(\theta | \underline{x}_1)$ 'den daha azdır. Bu olabilirliklerin her biri için  $\theta$  parametresi  $-\infty < \theta < \infty$  aralığında değer alır.

Ayrıca  $\alpha$  bu üç olabilirlik fonksiyonunun yüksekliklerine ilişkin her hangi bir küçük değer olmak üzere

$$L(\theta | \mathbf{x}) \geq \alpha \text{ için} \quad L'(\theta | \mathbf{x}) = L(\theta | \mathbf{x})$$

$$L(\theta | \mathbf{x}) < \alpha \text{ için} \quad L'(\theta | \mathbf{x}) = 0$$

olarak tanımlansın. Örnek olarak Şekil 2.2



Şekil 2.2. Üç farklı olabilirlik fonksiyonunun  $\alpha$  yükseklik deęeri

alınabilir. Eđer  $\alpha$  yeteri kadar küçükse  $L'(\theta | \mathbf{x}) \approx L(\theta | \mathbf{x})$  olur. Artık önsel dağılım, olabilirlikler arasında tarafsız olduğundan gerçek  $\theta$  deęerinin,  $\theta_1$  ve  $\theta_2$  arasında olduğu kadar  $\theta_3$  ile  $\theta_4$ ,  $\theta_5$  ile  $\theta_6$  arasında olması önsel dağılım olarak söz konusudur.

$$P(\theta_1 \leq \theta \leq \theta_2) = P(\theta_3 \leq \theta \leq \theta_4) = P(\theta_5 \leq \theta \leq \theta_6)$$

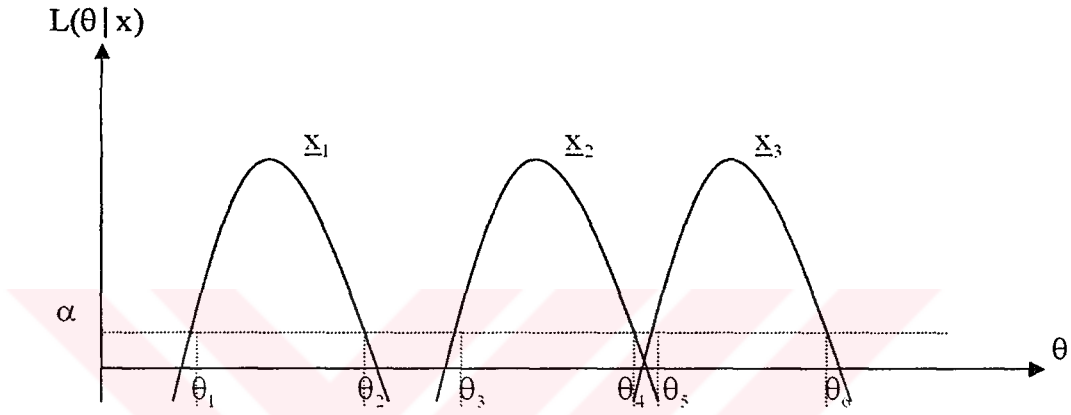
örnekten

$$\theta_2 - \theta_1 \neq \theta_4 - \theta_3 \neq \theta_6 - \theta_5$$

[2.7]

olduğundan  $\theta$  parametresi tekdüze dağılımından gelmez. Olabilirlik fonksiyonu bakımından  $f(\theta)$  ile gösterilen  $\theta$  parametresinin belirsiz önsel dağılımıyla ilgili istenen tekdüze dağılmamasıdır.

Olabilirlik fonksiyonları aşağıdaki gibi olduğu kabul edilsin.



Şekil 2.3. Üç farklı olabilirlik fonksiyonunun  $\alpha$  yükseklik değeri

Eğer olabilirlik fonksiyonu bu şekilde olursa örneklemden gelen bilginin etkisi sadece, olabilirlik fonksiyonunun konumunu değiştirmektir. Bu durumda olabilirlik fonksiyonunun değişen veri olduğu söylenebilir. Eş. 2.7'deki eşitsizlik tekrar ele alındığında bu eşitsizlik

$$\theta_2 - \theta_1 = \theta_4 - \theta_3 = \theta_6 - \theta_5$$

biçiminde eşitliğe dönüşür ve bundan dolayı  $\theta$  parametresinin önsel dağılımı tekdüze olur. Bu bilgiler yaklaşık olarak bilgi içermeyen önsel dağılım yaklaşımını açıklar. Bilgi içermeyen önsel dağılımın bilinen kullanımı aşağıdaki gibi verilir:

$\theta$  parametresinin  $g(\theta)$  ile gösterilen parametre dönüşümünün bulunduğu varsayalım.  $g(\theta)$  için olabilirlik fonksiyonu değişen veridir ve önsel dağılımı ise burada  $c$  bir sabit olmak üzere

$$f_g(g(\theta)) = \text{tekdüze}$$

$$f_g(g(\theta)) = c$$

olarak yazılır.  $\theta$  parametresinin bilgi içermeyen önsel dağılımı Eş. 2.6'daki ifadenin kullanılması

$$f(\theta) = c \left| \frac{dg(\theta)}{d\theta} \right|$$

yada

$$f(\theta) \propto \left| \frac{dg(\theta)}{d\theta} \right| \quad [2.8]$$

ile şeklinde elde edilir.

Örnek :

$\theta$  parametresi Normal dağılımın bilinmeyen kitle ortalaması ve  $\sigma^2$  parametresi bilinen kitle varyansı olmak üzere  $\theta$  parametresinin olabilirlik fonksiyonu,

$$L(\theta | \underline{x}) \propto \exp \left\{ -\frac{1}{2} \left[ \frac{\bar{x} - \theta}{\sigma / \sqrt{n}} \right]^2 \right\}$$

olarak yazılır. Bu fonksiyon zaten değişen veridir ( $\bar{x}$ 'nin farklı değerleri için başka bir değişme olmadan  $L(\theta | \underline{x})$  olabirlik fonksiyonunun konumu değişecek). Bundan dolayı, aranılan parametre dönüşümü

$$g(\theta) = \theta$$

olur. Bilgi içermeyen önsel dağılımdan  $c$  bir sabit olmak üzere  $f(\theta) = c$  olur.

Örnek :

$\mu$  parametresi Normal dağılımın bilinen kitle ortalaması ve  $\theta^2$  parametresi de bilinmeyen kitle varyansı olmak üzere  $\theta$  parametresinin olabirlik fonksiyonu,

$$s^2 = \sum_{i=1}^n \frac{(x_i - \mu)^2}{n} \text{ olmak üzere}$$

$$L(\theta | \underline{x}) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\theta^2}} \exp \left\{ -1/2 \left[ \frac{x_i - \mu}{\theta} \right]^2 \right\}$$

$$\propto \theta^{-n} \exp \left[ -\frac{1}{2\theta^2} \sum (x_i - \mu)^2 \right]$$

$$L(\theta | \underline{x}) \propto \theta^{-n} \exp \left[ -\frac{ns^2}{2\theta^2} \right]$$

olarak yazılabilir. Burada  $L(\theta | \underline{x})$  olabirlik fonksiyonu değişen veri olmadığından  $\theta$  parametresinin bilgi içermeyen önsel dağılımı da tekdüze değildir. Bundan dolayı  $L(\theta | \underline{x})$  olabirlik fonksiyonu  $s^n$  ile  $L(\theta | \underline{x})$  çarpılarak

$$L(\theta | \underline{x}) \propto \left[ \frac{s}{\theta} \right]^n \exp \left[ -\frac{ns^2}{2\theta^2} \right]$$

yazılır. Ayrıca olabilirlik fonksiyonu  $\ln \theta$  ile tekrar oluşturularak

$$L(\ln \theta | \underline{x}) \exp \left\{ n(\ln s - \ln \theta) - \frac{n}{2} \exp[2(\ln s - \ln \theta)] \right\}$$

ifadesi elde edilir. Burada  $\ln \theta$  parametre dönüşümüdür ve bilgi içermeyen önsel dağılımı tekdüzedir. O halde olabilirlik fonksiyonu artık değişen veridir. Böylece  $\theta$  parametresinin bilgi içermeyen önsel dağılımı

$$f(\theta) \propto \left| \frac{d \ln \theta}{d \theta} \right| = \frac{1}{\theta}$$

olup,  $\theta$  parametresinin sonsal dağılımı

$$f(\theta | \underline{x}) \propto L(\theta | \underline{x}) f(\theta) = \theta^{-n+1} \exp \left[ -\frac{ns^2}{2\theta^2} \right]$$

elde edilir. Yine  $f(\theta | \underline{x})$  ile verilen  $\theta$  parametresinin sonsal dağılımının oyf olabilmesi için gerekli sabit,

$$\int_0^{\infty} \theta^{-n+1} \exp \left[ -\frac{ns^2}{2\theta^2} \right] d\theta = \frac{\Gamma(n/2) \left( \frac{ns^2}{2} \right)^{-n/2}}{2} \\ = \frac{2^{n/2-1} \Gamma(n/2)}{(ns^2)^{n/2}}$$

olarak bulunur. Buradan  $\theta$  parametresinin sonsal dağılımı

$$f(\theta | \underline{x}) = \frac{(ns^2)^{n/2}}{2^{n/2-1} \Gamma(n/2)} \theta^{-n+1} \exp \left[ -\frac{ns^2}{2\theta^2} \right]$$

olur. Burada  $\frac{ns^2}{\theta^2}$  ifadesi yerine  $\chi^2$  yazılırsa elde edilen sonuç n serbestlik dereceli ki-kare dağılımıdır. Bu ise bilinen örnekleme sonucudur.

Burada anlatılan yöntem, bölüm dört ve beşte Bayes analizlerinde parametrelerin bilgi içermeyen önsel dağılımlarının elde edilmesinde kullanılmıştır.



### 3. BAZI DAĞILIMLAR VE KAVRAMLAR

Bu bölümde, dördüncü ve beşinci bölümlerde çoklu regresyon modelinin Bayes analizleri içerisinde yer alan bazı dağılımlar ve kavramlardan Çok Değişkenli Student-t dağılımı, Ters Gamma dağılımı, bilinen en küçük kareler tahmin edicileri ve Gauss Markov teoremi üzerinde durulacaktır.

#### 3.1. Çok Değişkenli Student-t Dağılımı

Merkezi olmayan Student-t dağılımı, serbestlik derecesi  $\nu$  ve merkezi olmama parametresi  $\delta$  ile,

$$t_{\nu}(\delta) = \frac{U + \delta}{\chi_{\nu} / \sqrt{\nu}}$$

olarak tanımlanır. Burada  $U$  tek değişkenli Normal dağılıma,  $\chi_{\nu}$  ise  $\nu$  serbestlik dereceli ki-kare dağılımına sahip rassal değişkenlerdir. Eğer  $\delta = 0$  ise  $t_{\nu}$ ,  $\nu$  serbestlik dereceli merkezi Student-t dağılımı olur.

Çok Değişkenli Merkezi Student-t dağılımının oyf,

$$P_Y(y) = \frac{\Gamma\left(\frac{1}{2}(\nu + m)\right)}{(\Pi\nu)^{1/2+m} \Gamma\left(\frac{1}{2}\nu\right) |R|^{1/2}} (1 + \nu^{-1}y'R^{-1}y)^{-\frac{1}{2}(\nu+m)} \quad [3.1]$$

eşitliği ile verilir. Bu oyf

$$Y_j = \frac{X_j}{S/\sqrt{\nu}} \quad j = 1, 2, \dots, m$$

eşitliğinde  $m$  sayıda bağımsız Student-t değişkeninin ortak oylur. Burada  $X_j$  rassal değişkenlerinin ortak dağılımı  $N[\underline{0}, R]$  ( $R$  varyans kovaryans matrisi) ile gösterilen Çok Değişkenli Normal dağılımdır.  $S$  rassal değişkeni ise  $\nu$  serbestlik dereceli ki-kare dağılımına sahiptir. Ayrıca  $X_j$  ve  $S$  rassal değişkenleri birbirinden bağımsızdır. Yine  $X_j$  rassal değişkenlerinin karşılıklı olarak bağımsız olması durumunda Eş. 3.1'de verilen oylur

$$P_Y(y) = \frac{\Gamma\left(\frac{1}{2}(\nu+m)\right)}{(\Pi\nu)^{1/2+m} \Gamma\left(\frac{1}{2}\nu\right)} \left(1 + \nu^{-1} \sum_{j=1}^m y_j^2\right)^{-\frac{1}{2}(\nu+m)} \quad [3.2]$$

eşitliğine dönüşür. Bu eşitlik  $\nu=1$  için Çok Değişkenli Cauchy dağılımı olur. Çok Değişkenli Student-t dağılımı monoton azalan bir fonksiyon ve orijine göre simetrik bir dağılımdır (8).

### 3.2. Ters Gamma Dağılımı

$\alpha$  ve  $\beta$  parametrelili Ters Gamma dağılımı  $IG(\alpha, \beta)$  ile gösterilir ve dağılım özellikleri, Gamma dağılımının özelliklerinden yararlanılarak kolayca bulunabilir.  $G(\alpha, \beta)$  ile gösterilen  $\alpha$  ve  $\beta$  parametrelili Gamma dağılımına sahip  $X$  rassal değişkenine

$$Y = \sqrt{\frac{1}{X}}$$

dönüşümü uygulanırsa  $Y$  rassal değişkeni Ters Gamma dağılımına sahip olur. bu dağılımın oylur,

$$f(y) = \frac{(\beta/2)^{\alpha-2}}{\Gamma(\alpha/2)} y^{2-(\alpha+2)/2} \exp\left\{-\beta/(2y^2)\right\}, \quad 0 \leq y < \infty, \quad \alpha > 0 \text{ ve } \beta > 0$$

dir. Ters Gamma dağılımının modu

$$\frac{\alpha}{\beta+2}$$

eşitliği ile, ortalaması ise

$$E(y) = \frac{\alpha}{\beta-2}, \quad \beta > 2$$

dir. Ayrıca bu dağılımın ortalaması modundan daha büyük olduğundan sağa çarpık dağılımdır. Yine varyansı ise,

$$\text{Var}(y) = \frac{2\alpha^2}{(\beta-2)^2(\beta-4)}, \quad \beta > 4$$

ifadesidir. Ters Gamma dağılımının uygulamadaki en iyi bilinen özelliği, Bayes analizinde Normal dağılımın bilinmeyen standart sapmasının önsel dağılımı olarak alınmasıdır (9).

### 3.3. Bilinen En Küçük Kareler Tahmin Edicisi ve Özellikleri

Bilinen en küçük kareler tahmin edicisinin anlamı gözlenen ve beklenen bağımsız y rassal değişkenleri arasındaki karesel uzaklığın minimum olmasıdır. Üzerinde çalışılan regresyon modeli,

$$y = X\beta + \varepsilon, \quad r(X_{n \times k}) = k < n$$

olsun. Bu modelde geçen bazı terimlerin tanımları

$n$  : gözlem sayısı

- $y$  :  $(n \times 1)$  boyutlu gözlem vektörü  
 $X$  :  $(n \times k)$  boyutlu bağımsız değişkenlere ilişkin matris  
 $\beta$  :  $(k \times 1)$  boyutlu bilinmeyen regresyon parametreleri vektörü  
 $\varepsilon$  :  $(n \times 1)$  boyutlu hata vektörü  
 $X'X$  : pozitif tanımlı bir matris

olarak verilir. Bu terimlerin özellikleri:

- i.  $E(\varepsilon) = 0$
- ii.  $\varepsilon$  ve  $X$  bağımsızdır. Yani  $Cov(X, \varepsilon) = 0$  olur.
- iii.  $E(\varepsilon\varepsilon') = \sigma^2 I_n$
- v.  $E(\varepsilon\varepsilon') = \sigma^2 \Omega$ ,  $\Omega = \text{diag}(\sigma_1^2, \sigma_2^2, \dots, \sigma_n^2) \neq I_n$  (otokorelasyon)
- vi.  $\varepsilon \sim N[0, \Sigma]$  (10).

olarak sıralanabilir.  $\beta$  ve  $\sigma^2$  parametrelerinin EKK tahmin edicileri

$$\hat{\beta} = (X'X)^{-1}X'y, \quad s^2 = \hat{\sigma}^2 = \frac{\hat{\varepsilon}'\hat{\varepsilon}}{n-k} \quad \text{yada} \quad s^2 = \frac{(y - X\hat{\beta})'(y - X\hat{\beta})}{n-k}$$

eşitlikleridir. Yine EKK tahmin edicilerinin bazı özellikleri

- i.  $\beta$  parametresinin EKK tahmin edicisi yansızdır.

$$E(\hat{\beta}) = \beta$$

- ii.  $\beta$  parametresinin EKK tahmin edicisi minimum varyansa sahip en iyi yansız tahmin dicitir. Ayrıca varyansı,

$$V(\hat{\beta}) = \sigma^2 (X'X)^{-1}$$

dir.

- iii. EKK tahmin edicisi

$$\hat{\beta} \sim N[\beta, \sigma^2 (X'X)^{-1}]$$

ile Normal dağılıma sahiptir.

- iv.  $\hat{\beta}$  ve  $\hat{\sigma}^2$  tahmin edicileri birbirinden bağımsızdır.

- v.  $\hat{\sigma}^2$  tahmin edicisinin dağılımı  $(n-k)s^2 / (\sigma^2) \sim \chi_{n-k}^2$  ile gösterilen n-k serbestlik dereceli ki-kare dağılımıdır (10).

### 3.4. Gauss Markov Teoreminin Özel bir Durumu

Üzerinde çalışılan regresyon modeli,

$$y_i = \alpha + \beta x_i + \varepsilon_i$$

olsun. Burada  $\beta$  parametresi için Gauss Markov teoreminin ispatı verilecektir. Gauss Markov teoremi olarak ifade edilen, EKK tahmin edicilerinin en küçük varyansa sahip olma özelliği üzerinde durulacaktır. Gauss Markov teoremine göre doğrusal yansız tahmin ediciler sınıfı içinde EKK tahmin edicileri, en küçük varyansa sahip en iyi doğrusal yansız tahmin edicilerdir. Ayrıca bu özellik BLUE (Best Linear Unbiased Estimator) olarak da ifade edilir.

İspat

$\beta$  parametresinin herhangi bir tahmin edicisi  $\hat{\beta}$  olsun. Ayrıca  $y$  rassal değişkeninin doğrusal bir fonksiyonu,  $c_i$  ifadeleri sabit olmak üzere

$$\hat{\beta} = \sum c_i (y_i - \bar{y}) \quad [3.5]$$

alınabilir. Yine  $\hat{\beta}$  tahmin edicisinin beklenen değeri

$$\begin{aligned} E[\hat{\beta}] &= \sum c_i E[y_i - \bar{y}] \\ &= \sum c_i E[\alpha + \beta x_i - (\alpha + \beta \bar{x})] \\ &= \sum c_i E[\alpha + \beta x_i - \alpha - \beta \bar{x}] \\ &= \sum c_i E[\beta x_i - \beta \bar{x}] \\ &= \sum c_i \beta (x_i - \bar{x}) \end{aligned}$$

$$E[\hat{\beta}] = \beta \sum c_i (x_i - \bar{x})$$

olup  $\hat{\beta}$  tahmin edicisinin yansız olması için

$$\sum c_i (x_i - \bar{x}) = 1 \quad [3.6]$$

eşitliği sağlanmalıdır. Diğer taraftan  $\hat{\beta}$  tahmin edicisinin varyansı

$$V[\hat{\beta}] = V[\sum c_i (y_i - \bar{y})]$$

$$= \sum c_i^2 V(y_i - \bar{y})$$

$$V[\hat{\beta}] = (\sum c_i^2) \sigma^2$$

olarak yazılır. Burada  $\sum c_i^2$  toplamının minimum olması gerekir. Bu nedenle Eş. 3.6 ifadesine Lagrange formülü uygulanırsa

$$L(c_i, \lambda) = \sum c_i^2 - \lambda [\sum c_i (x_i - \bar{x}) - 1]$$

yazılır. Bu eşitliğin sırasıyla  $c_i$  ve  $\lambda$  değişkenlerine göre kısmi türevi alınırsa

$$\frac{\partial L(c_i, \lambda)}{\partial c_i} = 2c_i - \lambda(x_i - \bar{x}) = 0$$

$$\frac{\partial L(c_i, \lambda)}{\partial \lambda} = \sum c_i (x_i - \bar{x}) - 1 = 0 \quad [3.7]$$

elde edilir. Böylece bu denklem sisteminin çözümünden

$$c_i = \left( \frac{\lambda}{2} \right) (x_i - \bar{x}) \quad [3.8]$$

olur. Elde edilen  $c_i$  ifadesi Eş. 3.7'de yerine yazılırsa

$$\sum \left\{ \frac{\lambda}{2} (x_i - \bar{x})(x_i - \bar{x}) \right\} = 1$$

$$\lambda = \frac{2}{\left(\sum (x_i - \bar{x})^2\right)} \quad [3.9]$$

olarak yazılır. Bu  $\lambda$  ifadesi Eş. 3.8'de yazılırsa,

$$c_i = \frac{(x_i - \bar{x})}{\sum (x_i - \bar{x})^2}$$

elde edilir. Böylece  $\hat{\beta}$  tahmin edicisi

$$\begin{aligned} \hat{\beta} &= \sum c_i (y_i - \bar{y}) \\ &= \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2} \end{aligned}$$

olur. Bu  $\hat{\beta}$  tahmin edicisi  $\beta$  parametresinin EKK tahmin edicisidir. Buna göre EKK tahmin edicisi, minimum varyansa sahip doğrusal yansız bir tahmin edicidir. İspatı bu şekilde verilen Gauss Markov teoremi

i.  $y_i = \alpha + \beta x_i + \varepsilon_i$

ii.  $E[\varepsilon_i] = 0$

iii.  $V(\varepsilon_i) = \sigma^2$

iv.  $\text{Cov}(\varepsilon_i, \varepsilon_j) = 0$ ,  $i \neq j$  olmak üzere

ile verilen koşullar altında sağlanır (11).

#### 4. ÇOK DEĞİŞKENLİ STUDENT-t HATA TERİMLİ DOĞRUSAL REGRESYON MODELİNİN BAYES VE KLASİK YÖNTEMLE ANALİZİ

Doğrusal regresyon analizinde kullanılacak modelin bazı varsayımları sağlaması gerekir.

i. Bağımsız değişkenlerin aldığı değerler sabit sayılardır. Araştırmacı bağımsız değişkenleri kendisi belirler. Bağımsız değişkenin aldığı değere bir bağımlı değişkenin rassal olarak seçilmiş birimlerdeki değeri karşılık gelir. Bu nedenle hata terimi ile bağımsız değişken birbirinden bağımsızdır.

ii. Hata teriminin ortalaması 0 ve varyansı  $\sigma^2$ 'dir. Bu varyans bağımsız değişkenin aldığı değere göre değişmez yani sabittir. Hata teriminin varyansı bağımlı değişkenin varyansıdır.

iii. Hata terimleri birbirinden bağımsızdır. Yani aralarında otokorelasyon yoktur.

$$\text{Cov}(\varepsilon_i, \varepsilon_j) = 0$$

iv. Hata terimi  $\varepsilon \sim N[0, \sigma^2 I]$  Normal dağılıma sahiptir (12).

Tek değişkenli doğrusal regresyon modeli uygulamalarında hata terimlerinin yukarıda da ifade edildiği gibi her birinin  $N[0, \sigma^2]$  dağıldığı kabul edilir. Bu bölümde ise hata terimleri 0 ortalama ve  $v_0 \sigma^2 / (v_0 - 2) I_n$  varyans-kovaryans matrisi ile Çok Değişkenli Student-t dağılımına sahiptir ve her bir hata teriminin marjinal dağılımı ise tek değişkenli Student-t dağılımıdır. Ayrıca bağımlı olmalarına rağmen ilişkisizdirler.

#### 4.1. Olabilirlik Fonksiyonu ve Örneklem Teorisi

Bu bölümde kullanılan çoklu doğrusal regresyon modeli,

$$y = X\beta + \varepsilon, \quad r(X_{n \times k}) = k < n, \quad [4.1]$$

olsun. Bu modeldeki terimlerin tanımları

- $n$  : gözlem sayısı  
 $y$  :  $(n \times 1)$  boyutlu gözlem vektörü  
 $X$  :  $(n \times k)$  boyutlu bağımsız değişkenlere ilişkin matris  
 $\beta$  :  $(k \times 1)$  boyutlu bilinmeyen regresyon parametreleri vektörü  
 $\varepsilon$  :  $(n \times 1)$  boyutlu hata vektörü

olarak verilir.  $\varepsilon$  ile gösterilen rassal hata vektörünün dağılımı,

$$p(\varepsilon | \cdot) = \int_0^{\infty} p_N(\varepsilon | \tau) p(\tau | \cdot) d\tau, \quad \tau, \sigma, \nu_0 > 0 \quad [4.2]$$

eşitliği ile verilen dağılımlar sınıfının elemanı olan Çok Değişkenli Student-t dağılımıdır. Bu dağılım Çok Değişkenli Normal dağılımın, Ters Gamma dağılımı ile ağırlıklandırılması ile elde edilir. Burada  $p_N(\varepsilon | \tau)$  ile gösterilen

$$p_N(\varepsilon | \tau) = \frac{1}{(2\pi\tau^2)^{n/2}} \exp\left\{-\frac{\varepsilon'\varepsilon}{2\tau^2}\right\} \quad -\infty < \varepsilon_i < \infty, \quad i = 1, 2, \dots, n \quad [4.3]$$

eşitlik, Çok Değişkenli Normal dağılım oyf ile oluşturulan hata terimleri vektörünün dağılımıdır.  $p(\tau | \cdot)$  ifadesi ise

$$p(\tau | v_0, \sigma^2) = \left[ \frac{2}{\Gamma(v_0/2)} \right] \left( \frac{v_0 \sigma^2}{2} \right)^{v_0/2} \tau^{-(v_0+1)} \exp \left\{ -\frac{v_0 \sigma^2}{2\tau^2} \right\} \quad [4.4]$$

olan  $v_0$  ve  $\sigma^2$  parametrelili Ters Gamma (13) dağılımıdır.

Eş. 4.2'deki integralin sonucu

$$\begin{aligned} p(\varepsilon | \cdot) &= \int_0^{\infty} p_N(\varepsilon | \tau) p(\tau | \cdot) d\tau \\ &= \int_0^{\infty} (2\pi\tau^2)^{-n/2} \exp \left\{ -\frac{\varepsilon' \varepsilon}{2\tau^2} \right\} \frac{2}{\Gamma(v_0/2)} \left( \frac{v_0 \sigma^2}{2} \right)^{v_0/2} \tau^{-(v_0+1)} \exp \left\{ -\frac{v_0 \sigma^2}{2\tau^2} \right\} d\tau \\ &= \frac{2}{\Gamma(v_0/2)} \left( \frac{v_0 \sigma^2}{2} \right)^{v_0/2} (2\pi)^{-n/2} \int_0^{\infty} \tau^{-(n+v_0+1)} \exp \left\{ -\frac{\varepsilon' \varepsilon + v_0 \sigma^2}{2\tau^2} \right\} d\tau \\ &= \frac{2}{\Gamma(v_0/2)} \left( \frac{v_0 \sigma^2}{2} \right)^{v_0/2} (2\pi)^{-n/2} \frac{\Gamma \left( \frac{n+v_0}{2} \right)}{2 \left\{ \frac{\varepsilon' \varepsilon + v_0 \sigma^2}{2} \right\}^{(n+v_0)/2}} \\ &= \frac{(v_0 \sigma^2)^{v_0/2}}{\Gamma(v_0/2)} 2^{-(n+v_0)/2} \pi^{-(n/2)} \frac{\Gamma \left( \frac{n+v_0}{2} \right)}{\left\{ v_0 + \frac{\varepsilon' \varepsilon}{\sigma^2} \right\}^{(n+v_0)/2}} 2^{(n+v_0)/2} (\sigma^2)^{-(n+v_0)/2} \\ &= \frac{(v_0)^{v_0/2}}{\Gamma(v_0/2)} \frac{(\sigma^2)^{-n/2}}{\pi^{n/2}} \Gamma \left( \frac{n+v_0}{2} \right) \left\{ v_0 + \frac{\varepsilon' \varepsilon}{\sigma^2} \right\}^{-(n+v_0)/2} \\ p(\varepsilon | \cdot) &= \frac{(v_0)^{v_0/2} \Gamma \left( \frac{n+v_0}{2} \right)}{\Gamma(v_0/2) \pi^{n/2} (\sigma^2)^{n/2}} \left\{ v_0 + \frac{\varepsilon' \varepsilon}{\sigma^2} \right\}^{-(n+v_0)/2} \end{aligned}$$

olur. Bu sonuç

$$p(\varepsilon | v_0, \sigma) = \frac{g(v_0)}{(\sigma^2)^{n/2}} \left\{ v_0 + \frac{\varepsilon' \varepsilon}{\sigma^2} \right\}^{-(n+v_0)/2} \quad [4.5]$$

$$\sigma, v_0 > 0, \quad -\infty < \varepsilon_i < \infty, \quad i = 1, 2, \dots, n$$

olarak düzenlenebilir. Burada  $g(v_0)$  ifadesi,

$$g(v_0) = \frac{v_0^{v_0/2} \Gamma((n+v_0)/2)}{\Gamma(v_0/2) \pi^{n/2}}$$

eşitliğidir.  $\varepsilon$  hata terimleri vektörünün dağılımının ortalaması

$$E(\varepsilon) = 0, \quad v_0 > 1$$

olup, kovaryans matrisi ise

$$E(\varepsilon \varepsilon') = \frac{v_0 \sigma^2}{(v_0 - 2)} I_n, \quad v_0 > 2 \quad [4.6]$$

ifadesidir. Burada  $I_n$ ,  $n \times n$  boyutlu birim matristir.

## 4.2. Çok Değişkenli Student-t Hata Terimli Regresyon Modeli

Daha önce ifade edildiği gibi hata terimleri ilişkisiz olmalarına rağmen, bağımsız değildir ve  $v_0$  serbestlik derecesi yeterince büyük olduğunda hata vektörünün dağılımı  $N[0, \sigma^2 I_n]$  ile Çok Değişkenli Normal Dağılıma yaklaşır.  $v_0$  serbestlik derecesinin sonlu bir değeri için  $\varepsilon_i / \sigma$  rassal değişkeni,  $v_0$  serbestlik dereceli tek

değişkenli Student-t dağılımına sahip olur. Serbestlik derecesinin  $v_0 = 1$  değeri için, hata vektörünün dağılımı

$$P(\varepsilon | \sigma) = \frac{\Gamma((1+n)/2)}{\pi^{(n+1)/2} (\sigma^2)^{n/2}} \left\{ 1 + \frac{\varepsilon' \varepsilon}{\sigma^2} \right\}^{-(n+1)/2}$$

eşitliği ile verilen Çok Değişkenli Cauchy dağılımı olur.

$\varepsilon$  hata terimleri vektörü, Çok Değişkenli Student-t dağılımına sahip bu doğrusal regresyon modeli için olabirlik fonksiyonu (14),

$$p(y | \beta, v_0, \sigma) = \left[ \frac{g(v_0)}{(\sigma^2)^{n/2}} \right] \left\{ v_0 + \left[ v s^2 + (\beta - \hat{\beta})' X' X (\beta - \hat{\beta}) \right] / \sigma^2 \right\}^{-n+v_0/2} \quad [4.7]$$

dir. Burada,  $\hat{\beta}$  ve  $v s^2$  ifadeleri

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

$$v s^2 = (y - X\hat{\beta})' (y - X\hat{\beta}) \text{ ve } v = n - k \quad [4.8]$$

olarak verilir. Ayrıca  $\hat{\beta}$  ve  $s^2$  ifadeleri yeterli istatistiklerdir. Eş. 4.7 ile verilen olabirlik fonksiyonu,  $\sigma^2$  ve  $\beta$  parametrelerinin EÇOB tahmin edicilerinin elde edilmesinde kullanılır. Yine sözü edilen  $p(y | \beta, v_0, \sigma)$  olabirlik fonksiyonunda

$$v s^2 = A; \quad (\beta - \hat{\beta})' X' X (\beta - \hat{\beta}) = B$$

ve ayrıca  $P_x = X(X'X)^{-1}X'$

olsun. Bu  $P_x$  ifadesi

i.  $P_x = P_x^2$  eşgüçlü bir matristir.

ii.  $P_x = P_x'$  simetriktir.

iii.  $r(P_x) = r(X)$

özelliklerini sağlayan matris olmak üzere, B ifadesi

$$\begin{aligned}
 B &= (\beta - \hat{\beta})' X'X (\beta - \hat{\beta}) \\
 &= (\beta - (X'X)^{-1} X'y)' X'X (\beta - (X'X)^{-1} X'y) \\
 &= \beta' X'X \beta - \beta' X' P_x y - y' P_x X \beta + y' P_x P_x y \\
 &= \beta' X'X \beta - \beta' X' P_x y - y' P_x X \beta + y' P_x y
 \end{aligned}$$

olarak düzenlenir. Şimdi  $vs^2$  eşitliğinde  $\hat{\beta}$  tahmin edicisi yerine eşiti olan  $(X'X)^{-1} X'y$  matrisi yazılırsa

$$\begin{aligned}
 A = vs^2 &= (y - P_x y)' (y - P_x y) \\
 &= y'y - y' P_x y - y' P_x y + y' P_x P_x y
 \end{aligned}$$

$$= y'y - y'P_x y - y'P_x y + y'P_x y$$

elde edilir. A ve B eşitlikleri taraf tarafa toplanırsa

$$A + B = \beta'X'X\beta - \beta'X'P_x y - y'P_x X\beta + y'P_x y + y'y - y'P_x y - y'P_x y + y'P_x y$$

$$= y'y + \beta'X'X\beta - \beta'X'P_x y - y'P_x X\beta$$

$$= (y - P_x X\beta)'(y - P_x X\beta)$$

$$= (y - X\beta)'(y - X\beta)$$

bulunur. Buradan Eş. 4.7'deki  $p(y | \beta, v_0, \sigma)$  olasılık fonksiyonu

$$P(y | \beta, v_0, \sigma) = \left[ \frac{g(v_0)}{(\sigma^2)^{n/2}} \right] \left\{ v_0 + (y - X\beta)'(y - X\beta) / \sigma^2 \right\}^{-(n+v_0)/2}$$

olup, bu eşitliğin her iki tarafının ln fonksiyonu altındaki görüntüsü alınırsa

$$L(\beta, \sigma^2) = \ln g(v_0) - \frac{n}{2} \ln \sigma^2 - \left( \frac{n+v_0}{2} \right) \ln \left( v_0 \sigma^2 + (y - X\beta)'(y - X\beta) \right) + \left( \frac{n+v_0}{2} \right) \ln \sigma^2$$

elde edilir. Buradan  $\beta$  parametresinin EÇOB tahmin edicisi

$$\frac{\partial L(\beta, \sigma^2)}{\partial \beta} = \frac{2(X'X)\beta - 2X'y}{v_0 \sigma^2 + (y - X\beta)'(y - X\beta)} = 0$$

$$2(X'X)\beta - 2X'y = 0$$

$$(X'X)\beta = X'y$$

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

olur. Bu sonuç  $\beta$  parametresinin EKK tahmin edicisinin aynı zamanda EÇOB tahmin edicisi olduğunu gösterir.  $\hat{\beta}$  tahmin edicisinin beklenen değeri alınırsa

$$E(\hat{\beta}) = E[(X'X)^{-1}X'y]$$

$$= (X'X)^{-1}X'E(y)$$

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

bulunur. Buradan  $\hat{\beta}$  tahmin edicisi  $\beta$  parametresinin yansız bir tahmin edicisi olduğu görülür. Yine  $\hat{\beta}$  tahmin edicisinin varyansı

$$V(\hat{\beta}) = E(\hat{\beta} - E(\hat{\beta}))(\hat{\beta} - E(\hat{\beta}))' \quad [4.9]$$

$$= E(\hat{\beta} - \beta)(\hat{\beta} - \beta)'$$

$$= E(\beta + (X'X)^{-1}X'\epsilon - \beta)(\beta + (X'X)^{-1}X'\epsilon - \beta)'$$

$$= E((X'X)^{-1}X'\epsilon)((X'X)^{-1}X'\epsilon)'$$

$$= E((X'X)^{-1}X'\epsilon\epsilon'X(X'X)^{-1})$$

$$= (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{X} \mathbf{E}(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}') \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}$$

Eş. 4.6 kullanılarak

$$V(\hat{\boldsymbol{\beta}}) = (\mathbf{X}'\mathbf{X})^{-1} \sigma_{\varepsilon}^2 = (\mathbf{X}'\mathbf{X})^{-1} v_0 \sigma^2 / (v_0 - 2)$$

elde edilir.  $v_0$  serbestlik derecesinin yeterince büyük olması durumunda bu eşitlik  $(\mathbf{X}'\mathbf{X})^{-1} \sigma^2$  varyans kovaryans matrisine yaklaşır.

Yine aynı şekilde  $\sigma^2$  parametresinin EÇOB tahmin edicisi

$$\frac{\partial L(\boldsymbol{\beta}, \sigma^2)}{\partial \sigma^2} = -\frac{n-1}{2\sigma^2} - \left(\frac{n+v_0}{2}\right) \frac{v_0}{v_0\sigma^2 + (\mathbf{y}-\mathbf{X}\boldsymbol{\beta})'(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})} + \left(\frac{n+v_0}{2}\right) \frac{1}{\sigma^2} = 0$$

$$\frac{v_0}{2\sigma^2} = \frac{n+v_0}{2} \frac{v_0}{v_0\sigma^2 + (\mathbf{y}-\mathbf{X}\boldsymbol{\beta})'(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})}$$

$$n\sigma^2 + v_0\sigma^2 = v_0\sigma^2 + (\mathbf{y}-\mathbf{X}\boldsymbol{\beta})'(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})$$

$$n\sigma^2 = (\mathbf{y}-\mathbf{X}\boldsymbol{\beta})'(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})$$

$$\hat{\sigma}^2 = (\mathbf{y}-\mathbf{X}\hat{\boldsymbol{\beta}})'(\mathbf{y}-\mathbf{X}\hat{\boldsymbol{\beta}}) / n$$

olarak elde edilir.  $\hat{\sigma}^2$  tahmin edicisinin beklenen değeri

$$E(\hat{\sigma}^2) = E\boldsymbol{\varepsilon}'(\mathbf{I}-\mathbf{P}_x)\boldsymbol{\varepsilon} / n = 1/n(\text{tr}(\mathbf{I}-\mathbf{P}_x)\sigma^2\mathbf{I}_n + (\mathbf{E}(\boldsymbol{\varepsilon}))'(\mathbf{I}-\mathbf{P}_x)\mathbf{E}(\boldsymbol{\varepsilon}))$$

$$E(\hat{\sigma}^2) = (n-k)\sigma_{\varepsilon}^2 / n$$

olur. Buradan  $\hat{\varepsilon}'\hat{\varepsilon}/(n-k)$  eşitliği ile de gösterilen  $\hat{\sigma}^2$  tahmin edicisinin,  $\sigma_e^2$  parametresi için yansız bir tahmin edicisi olduğu görülür. Yine  $\sigma_e^2 = \frac{v_0 \sigma^2}{(v_0 - 2)} I_n$  olduğundan,  $(v_0 - 2)\hat{\varepsilon}'\hat{\varepsilon}/v_0(n-k)$  tahmin edicisi de  $\sigma^2$  parametresi için yansız bir tahmin edicisi olur.  $q\hat{\varepsilon}'\hat{\varepsilon}$  ile gösterilen tahmin ediciler sınıfında  $q$  pozitif bir skaler olmak üzere  $\sigma_e^2$  parametresinin EKK tahmin edicisi

$$\tilde{\sigma}_e^2 = (v_0 - 4)\hat{\varepsilon}'\hat{\varepsilon}/(v_0 - 2)(v + 2)$$

olup  $\sigma^2$  parametresinin EKK tahmin edicisi ise

$$\tilde{\sigma}^2 = (v_0 - 4)\hat{\varepsilon}'\hat{\varepsilon}/v_0(v + 2)$$

olur ve bütün olabilirlik fonksiyonları için  $\hat{\varepsilon}'\hat{\varepsilon} = (y - X\hat{\beta})'(y - X\hat{\beta})$  eşitliği içerisindeki  $\hat{\beta}$ ,  $\beta$  parametresinin EÇOB tahmin edicisidir. Ayrıca  $v_0$  serbestlik derecesinin  $v_0 > 2$  koşulu altında Eş.4.6 sağlandığında Gauss Markov teoreminin şartları yerine getirilmiş olur. Bu nedenle  $\hat{\beta}$  tahmin edicisi, en küçük varyansa sahip doğrusal yansız bir tahmin edici olur.

Hata terimlerinin bağımsız olmamasına ve Eş. 4.5 ile verilen Çok Değişkenli Student-t dağılımına sahip olmasına rağmen parametrelere ilişkin hipotez testleri ve güven aralıkları  $t$  ve  $F$  istatistiklerine dayanmaktadır. Örneğin,

$$t = \frac{\hat{\beta}_i - \beta_i}{s(m^{ii})^{1/2}}$$

olsun. Burada

$m^{ii}$  ifadesi  $(X'X)^{-1}$  matrisinin  $(i,i)$ 'inci elemanı,  $s^2$  ise  $\hat{\varepsilon}'\hat{\varepsilon}/(n-k)$  eşitliğidir.

$\varepsilon$  hata vektörünün dağılımı,  $\tau$  parametresi biliniyorken, Eş. 4.3'deki Normal dağılımdır. Buradan  $t$  istatistiğinin dağılımı,  $(n-k)$  serbestlik dereceli tek değişkenli Student- $t$ 'dir.  $t$  istatistiği ve  $\tau$  parametresinin ortak oyları  $p(t, \tau) = p(t \setminus \tau)p(\tau)$  eşitliği ile bulunur. Bu eşitlik bağımsız olmaları durumunda ise  $p(t \setminus \tau) = p(t)$  olur. Benzer olarak herhangi bir  $F$  istatistiği ile  $\tau$  parametresinin ortak oyları  $p(F, \tau) = p(F \setminus \tau)p(\tau)$  eşitliği ile bulunduğundan,  $F$  istatistiğinin marjinal dağılımı ( $\tau$  biliniyorken)  $F_{v_1, v_2}$  oyları olur. Diğer taraftan  $p(F \setminus \tau)$  koşullu olasılığı,  $\tau$  parametresine bağlı olmadığından  $F$  istatistiği ve  $\tau$  parametresi bağımsız olur. Bu nedenle  $F$  istatistiği koşulsuz  $F_{v_1, v_2}$  oylarına sahip olur. Bu durum sadece  $\tau$  parametresinin dağılımının ters Gamma olmasından kaynaklanmaz, Eş. 4.2'deki ifadeyi sağlayan  $\tau$  parametresinin herhangi bir özel oyları için de sağlanır.

$\sigma^2$  yada  $\sigma_\varepsilon^2$  parametrelerine ait güven aralıkları oluşturmak ve hipotez testleri yapmak için, Eş. 4.3 ile verilen ( $\tau$  biliniyorken) Normal dağılıma sahip  $\varepsilon' = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)$  hata vektörü kullanılır. Bu durumda,

$$\frac{\hat{\varepsilon}'\hat{\varepsilon}}{\tau^2} \equiv \frac{\omega}{v\sigma^2} \sim \chi_v^2$$

olur. Örneğin  $s^2 = \frac{\hat{\varepsilon}'\hat{\varepsilon}}{v}$  olmak üzere

$$x = \frac{\omega}{v\sigma^2} = \frac{s^2}{\sigma^2}$$

olsun. Eş. 4.4 ile verilen Ters Gamma dağılımından

$$\sigma^2 / \tau^2 \sim (\chi_{v_0}^2 / v_0)^{-1}$$

yazılır. Yine  $x = \frac{(\omega/2\sigma^2)}{(\sigma^2/\tau^2)}$  eşitliğinden

$$x = s^2/\sigma^2 \sim F_{v,v_0}$$

dır. Bu sonuç  $\sigma^2$  ve  $\sigma_e^2$  parametreleri ile ilgili hipotezlerin testi ve güven aralıklarının oluşumunda kullanılabilir.

Son olarak Eş. 4.a'deki olabilirlik fonksiyonunu maksimum yapan  $\beta$  ve  $\sigma^2$  parametrelerinin EÇOB tahmin edicileri varken,  $v_0$  serbestlik derecesinin EÇOB tahmin edicisi yoktur. Fakat tekrarlanan gözlemler için  $v_0$  değeri veriden tahmin edilebilir. Bu sonuçları kullanmadan önce hata terimlerinin dağılım özelliklerini ortaya koymak,  $v_0$  değerine karar vermek açısından önemlidir. Hata terimlerinin Normal dağılım varsayımı ile  $v_0$  değerini büyük almak arasında hiçbir fark yoktur.  $v_0$  serbestlik derecesinin küçük değerleri için ise bu sonuçlar elde edilir.

### 4.3. Bayes Analizi

Bu kısımda, Eş. 4.7'deki olabilirlik fonksiyonu ve bilgi içermeyen önsel dağılım ile  $\beta$  ve  $\sigma^2$  parametrelerinin Bayes analizleri yapılacaktır.  $\beta$  ve  $\sigma^2$  parametrelerinin bilgi içermeyen önsel dağılımları

$$p(\beta, \sigma^2) \propto \frac{1}{\sigma^2} \quad -\infty < \beta_i < \infty, \quad 0 < \sigma^2 < \infty, \quad i = 1, 2, \dots, k \quad [4.10]$$

olup, bu parametrelerin sonsal dağılımları

$$A(\beta) = (y - X\beta)'(y - X\beta),$$

$$\bar{\sigma}^2 = (v_0\sigma^2 + vs^2)/v_1, \quad v_1 = v + v_0, \quad v = n - k,$$

$$\hat{\beta} = (X'X)^{-1}X'y, \quad vs^2 = (y - X\hat{\beta})'(y - X\hat{\beta})$$

olmak üzere,

$$p(\beta, \sigma^2 | y, v_0) \propto \frac{\{v_0 + (y - X\beta)'(y - X\beta)/\sigma^2\}^{-(v_1+v_0)/2}}{(\sigma^2)^{n+2}} \quad [4.11]$$

$$\propto \frac{(\sigma^2)^{v_0/2-1}}{(\bar{\sigma}^2)^{n+v_0/2}} \left\{ v_1 + (\beta - \hat{\beta})' X'X (\beta - \hat{\beta}) / \bar{\sigma}^2 \right\}^{-(v_1+k)/2} \quad [4.12]$$

$$\propto \{A(\beta)\}^{-n/2} \left\{ \frac{[v_0\sigma^2 / A(\beta)]^{v_0/2-1}}{A(\beta)[1 + v_0\sigma^2 / A(\beta)]^{n+v_0/2}} \right\} \quad [4.13]$$

şeklindedir. Eş. 4.12 ( $\sigma^2$  ve  $v_0$  biliniyorken),  $\beta$  parametresinin koşullu sonsal dağılımı Çok Değişkenli Student-t'dir. Dağılımın ortalaması

$\hat{\beta}$  tahmin edicisi olup, kovaryans matrisi,

$$V(\beta | y, \sigma^2, v_0) = \frac{(X'X)^{-1}v_1\bar{\sigma}^2}{(v_1-2)}, \quad v_1 > 2$$

$$= \frac{(X'X)^{-1}(v_0\sigma^2 + vs^2)}{(v_0 + v - 2)} \quad [4.14]$$

eşitliğidir.  $v_0$  serbestlik derecesinin yeterince büyük değerleri için ( $\sigma^2$  biliniyorken),  $\beta$  parametresinin koşullu sonsal dağılımı  $\beta \sim N[\hat{\beta}, (X'X)^{-1}\sigma^2]$  dağılımına yaklaşır.

Eş. 4.13 ile verilen ifade de  $z = n\sigma^2 / A(\beta)$  ( $0 < z < \infty$ ) olsun. Bu durumda  $\beta$  ve  $z$  parametrelerinin ortak sonsal dağılımları,

$$p(\beta, z | y, v_0) \propto \frac{\{A(\beta)\}^{-n/2} z^{(v_0/2)-1}}{(1+v_0/n)^{(n+v_0/2)}} \quad [4.15]$$

ve

$$z = n\sigma^2 / A(\beta) \sim F_{v_0, n}$$

olur.  $\beta$  ve  $z$  parametreleri bağımsız olduğundan  $\beta$  parametresinin marjinal sonsal dağılımı,

$$p(\beta | y, v_0) \propto \{A(\beta)\}^{-n/2}$$

$$\propto \left\{ v s^2 + (\beta - \hat{\beta})' X'X (\beta - \hat{\beta}) \right\}^{-(v+k)/2} \quad [4.16]$$

eşitliği ile verilen  $k$ -boyutlu Student-t dağılımı olur. Bu sonuç,  $\beta$  ve  $\sigma^2$  parametrelerinin bilgi içermeyen önsel dağılımı ile genel regresyon modelinin Bayes analizi sonucudur. Kısım 4.1'de açıklanan örnekleme teorisi sonuçlarının Bayes karşılığıdır.

$\sigma^2$  parametresinin marjinal sonsal dağılımı, Eş. 4.12'deki Çok Değişkenli Student-t dağılımının  $\beta$  parametresine göre integralenmesiyle elde edilir. Bu sonuç

$$p(\sigma^2 | y, v_0) \propto \frac{(\sigma^2 / s^2)^{(v_0/2)-1}}{(1+(v_0\sigma^2 / v s^2))^{(v_0+v_0/2)}} \quad [4.17]$$

olup,  $\sigma^2 / s^2$  ifadesinin marjinal sonsal dağılımı  $F_{v_0, v}$  ile gösterilen F dağılımıdır. Bu da Kısım 4.1'de  $\sigma^2$  parametresi ile ilgili anlatılan örnekleme teorisi sonucuna

paraleldir.  $F_{v_0, v}$  dağılımının modu  $\frac{v(v_0-2)}{v_0(v+2)}$  ve ortalaması  $\frac{v}{(v-2)}$  eşitlikleri ile

verilir.  $v_0$  serbestlik derecesinin yeterince büyük olması durumunda  $vs^2/\sigma^2$  ifadesinin sonsal dağılımı  $v$  serbestlik dereceli ki-kare dağılımına yaklaşır.

$z = \frac{n\sigma^2}{(y - X\beta)'(y - X\beta)}$  eşitliğinin sonsal dağılımı  $F_{v_0, n}$  ile gösterilen F dağılımıdır.

Bu sonuçlar, Eş. 4.10'deki  $\beta$  ve  $\sigma^2$  parametrelerinin bilgi içermeyen önsel dağılımı ile doğrusal regresyon modelinin Bayes analizi sonuçlarıdır.

Şimdi Çok değişkenli Student-t dağılımına sahip hata terimli doğrusal regresyon modelinin Bayes analizinin ikinci kısmı olan bilinmeyen parametrelerin eşlenik önsel dağılımları ile Bayes analizleri anlatılacaktır. Eş. 4.7 ile verilen olabilirlik fonksiyonu,

$v_1 = v + v_0$  ve  $\bar{\sigma}^2 = \frac{(v_0\sigma^2 + vs^2)}{v_1}$  olmak üzere

$$p(y | \beta, v_0, \sigma^2) \propto \left\{ \frac{(v_0\sigma^2 / vs^2)^{v_0/2}}{(1 + v_0\sigma^2 / vs^2)^{(v_0 + v)/2}} \right\} \left\{ \frac{(\bar{\sigma}^2)^{-k/2}}{\left[ v_1 + (\beta - \hat{\beta})' X'X (\beta - \hat{\beta}) / \bar{\sigma}^2 \right]^{v_1 - k/2}} \right\}$$

[4.18]

olarak yazılabilir.  $\beta$  ve  $\sigma^2$  parametrelerinin eşlenik önsel dağılımları, bilgi içermeyen önsel dağılım ile Bayes analizi sonuçlarındaki  $p_F(\sigma^2 | \cdot)$  eşitliği ile verilen  $\sigma^2$  parametresinin marjinal sonsal dağılımı ve  $p_S(\beta | \sigma^2, \cdot)$  eşitliği ile verilen  $\beta$  parametresinin koşullu sonsal dağılımının çarpımı olarak

$$p(\beta, \sigma^2) = p_F(\sigma^2 | \cdot) p_S(\beta | \sigma^2, \cdot) \quad [4.19]$$

yazılabilir. Yukarıda adı geçen  $p_F(\sigma^2 | \cdot)$  ve  $p_S(\beta | \sigma^2, \cdot)$  ifadeleri  $\bar{\beta}$   $\beta$  parametresinin eşlenik önsel dağılım ortalaması,  $\bar{v}_a = v_0 + v_a$  ve  $\bar{\sigma}_a^2 = (v_a S_a^2 + v_0 \sigma^2) / \bar{v}_a$  olmak üzere

$$p_F(\sigma^2 | s_a^2, v_a, v_0) \propto \frac{(v_0 \sigma^2 / v_a s_a^2)^{(v_0-2)/2}}{(1 + v_0 \sigma^2 / v_a s_a^2)^{(v_0+v_a)/2}}, \quad 0 < \sigma < \infty$$

$$v_a, s_a > 0 \quad [4.20]$$

ve

$$p_S(\beta | \sigma^2, \bar{\beta}, A, v_a) \propto (\bar{\sigma}_a^2)^{-k/2} \left\{ \frac{\bar{v}_a + (\beta - \bar{\beta})' A (\beta - \bar{\beta})}{\bar{\sigma}_a^2} \right\}^{-(\bar{v}_0+k)/2}$$

$$-\infty < \beta_i < \infty, \quad i = 1, 2, \dots, k \quad [4.21]$$

eşitlikleri ile verilebilir. Bu analiz sonuçları Ek-1'de verilmiştir (Bkz. Ek. 1).

Eş. 4.19 ile verilen eşlenik önsel dağılımda  $\beta$  ve  $\sigma^2$  parametreleri bağımsız değildir. Eşlenik önsel dağılımın yeterli önsel dağılım bilgisi içerdiği düşünülürse, sonsal dağılımı elde etmek için kullanılabilir. Zaten Eş. 4.19'daki eşlenik önsel dağılımlar, bilgi içermeyen önsel dağılım ile olabilirlik fonksiyonunun birleşimi ile elde edilen parametrelerin sonsal dağılımlarıdır.

Son olarak basit regresyon modelinde olduğu gibi, parametre sayısı çok değilse sonsal dağılımın analizde kullanılan sayısal integrasyon teknikleri, önsel dağılımların seçimine oldukça esneklik verir. Bu da önsel dağılım bilgileriyle basit modellerin analizinde kullanılır (14).

## 5. DOĞRUSAL REGRESYON MODELİNDEKİ HATA TERİMLERİNİN BAYES ANALİZİ

Bu bölümde çalışmanın ikinci aşaması olan Normal dağılıma sahip hata terimli çoklu regresyon modelinin Bayes analizi anlatılacaktır. Bu modelde regresyon varsayımlarından olan, dördüncü bölümdeki modelin sağlamadığı hata terimlerinin Normal dağılım ve bağımsızlık koşulları sağlanmaktadır. Hata terimlerinin Normal dağılım varsayımı klasik yöntemde hipotez testleri ve güven aralıklarının oluşturulması aşamasında yani parametrik testler için istenmektedir. Normal dağılım varsayımı olmadan, parametrik testlerin yapılması pek mümkün değildir. Bu yüzden çalışmalarda Normallik varsayımı aranır.

### 5.1. Regresyon Modeli ve Parametrelerin EÇOB Tahmin Edicileri

Burada kullanılan çoklu regresyon modeli,

$$y = X\beta + \varepsilon, \quad r(X_{n \times k}) = k < n, \quad \varepsilon \sim N(0, \sigma^2 I_n) \quad [5.1]$$

olsun. Bu modeldeki terimlerin tanımları

- $n$  : gözlem sayısı
- $y$  :  $(n \times 1)$  boyutlu gözlem vektörü
- $X$  :  $(n \times k)$  boyutlu bağımsız değişkenlere ilişkin matris
- $\beta$  :  $(k \times 1)$  boyutlu bilinmeyen regresyon katsayıları vektörü
- $\varepsilon$  :  $(n \times 1)$  boyutlu hata vektörü

olarak verilir.  $\varepsilon$  ile gösterilen rassal hata vektörünün dağılımı,

$$P_N(\varepsilon | \sigma) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left\{-\frac{\varepsilon'\varepsilon}{2\sigma^2}\right\}$$

dir.  $y$  gözlem vektörünün dağılımı ise  $y \sim N[X\beta, \sigma^2 I_n]$ 'dir.

Eş. 5.1'deki doğrusal regresyon modeli için olabilirlik fonksiyonu

$$P(y | \beta, \sigma) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left\{-\frac{(y - X\beta)'(y - X\beta)}{2\sigma^2}\right\} \quad [5.2]$$

olup,  $\sigma$  ve  $\beta$  parametrelerinin EÇOB tahmin edicilerinin elde edilmesinde kullanılır.

Bu eşitliğin her iki tarafının ln fonksiyonu altındaki görüntüsü alınırsa

$$L(\beta, \sigma^2) = \ln P(y | \beta, \sigma^2) = -\frac{n}{2} \ln \sigma^2 - \frac{(y - X\beta)'(y - X\beta)}{2\sigma^2}$$

elde edilir. Buradan  $\beta$  parametresinin EÇOB tahmin edicisi

$$\frac{\partial L(\beta, \sigma^2)}{\partial \beta} = \frac{[2(X'X)\beta - 2X'y]2\sigma^2}{(2\sigma^2)^2} = 0$$

$$2(X'X)\beta - 2X'y = 0$$

$$(X'X)\beta = X'y$$

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

olur. Bu sonuç  $\beta$  parametresinin EKK tahmin edicisinin aynı zamanda EÇOB tahmin edicisi olduğunu gösterir.  $\hat{\beta}$  tahmin edicisinin beklenen değeri alınırsa

$$E(\hat{\beta}) = E[(X'X)^{-1} X'y]$$

$$= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{E}(\mathbf{y})$$

$$\mathbf{E}(\hat{\boldsymbol{\beta}}) = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}\boldsymbol{\beta} = \boldsymbol{\beta}$$

bulunur. Buradan  $\hat{\boldsymbol{\beta}}$  tahmin edicisi  $\boldsymbol{\beta}$  parametresinin yansız bir tahmin edicisi olduğu görülür. Yine  $\hat{\boldsymbol{\beta}}$  tahmin edicisinin varyansı

$$\mathbf{V}(\hat{\boldsymbol{\beta}}) = \mathbf{E}(\hat{\boldsymbol{\beta}} - \mathbf{E}(\hat{\boldsymbol{\beta}}))(\hat{\boldsymbol{\beta}} - \mathbf{E}(\hat{\boldsymbol{\beta}}))'$$

$$= \mathbf{E}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})'$$

$$= \mathbf{E}(\boldsymbol{\beta} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\varepsilon} - \boldsymbol{\beta})(\boldsymbol{\beta} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\varepsilon} - \boldsymbol{\beta})'$$

$$= \mathbf{E}((\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\varepsilon})(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\varepsilon}'$$

$$= \mathbf{E}((\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1})$$

$$= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}\mathbf{E}(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}')\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}$$

$$\mathbf{V}(\hat{\boldsymbol{\beta}}) = (\mathbf{X}'\mathbf{X})^{-1}\sigma^2$$

eşitliği olarak bulunur.

Yine aynı şekilde  $\sigma^2$  parametresinin EÇOB tahmin edicisi

$$\frac{\partial \mathcal{L}(\boldsymbol{\beta}, \sigma^2)}{\partial \sigma^2} = -\frac{n-1}{2\sigma^2} - \left(\frac{n+v_0}{2}\right) \frac{v_0}{v_0\sigma^2 + (\mathbf{y}-\mathbf{X}\boldsymbol{\beta})'(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})} + \left(\frac{n+v_0}{2}\right) \frac{1}{\sigma^2} = 0$$

$$\frac{\partial \mathcal{L}(\beta, \sigma^2)}{\partial \sigma^2} = -\frac{n}{2} \ln \sigma^2 + \frac{(y - X\beta)'(y - X\beta)}{2(\sigma^2)^2} = 0$$

$$\frac{n}{2\sigma^2} = \frac{(y - X\beta)'(y - X\beta)}{2(\sigma^2)^2}$$

$$n\sigma^2 = (y - X\beta)'(y - X\beta)$$

$$\hat{\sigma}^2 = (y - X\hat{\beta})'(y - X\hat{\beta})/n$$

olarak elde edilir.  $\hat{\sigma}^2$  tahmin edicisinin beklenen değeri

$$E(\hat{\sigma}^2) = E\epsilon'(I - P_x)\epsilon/n = 1/n(\text{tr}(I - P_x)\sigma^2 I_n + (E(\epsilon))'(I - P_x)E(\epsilon))$$

$$E(\hat{\sigma}^2) = (n - k)\sigma^2/n$$

olur. Buradan  $\hat{\sigma}^2/(n - k)$  eşitliği ile de gösterilen  $\hat{\sigma}^2$  tahmin edicisinin,  $\sigma^2$  parametresi için yansız bir tahmin edici olduğu görülür.

## 5.2. Parametrelerin Bayes Analizi

$\sigma$  ve  $\beta$  parametrelerinin bilgi içermeyen önsel dağılımı

$$p(\beta, \sigma) \propto \frac{1}{\sigma}, \quad -\infty < \beta_i < \infty, \quad i = 1, 2, \dots, k \quad 0 < \sigma < \infty \quad [5.3]$$

olup, Olabilirlik fonksiyonu Eş. 5.2'dir. Bu parametrelerin sonsal dağılımları

$$\hat{\beta} = (X'X)^{-1}X'y, \quad v\hat{\sigma}^2 = (y - X\hat{\beta})'(y - X\hat{\beta}) \quad \text{ve } v = n - k \quad \text{olmak üzere}$$

$$p(\beta, \sigma | y) \propto p(\beta, \sigma)p(y | \beta, \sigma)$$

$$\begin{aligned} &\propto \frac{1}{\sigma^{n+1}} \exp\left\{-\frac{1}{2\sigma^2}(y - X\beta)'(y - X\beta)\right\} \\ &\propto \frac{1}{\sigma^{n+1}} \exp\left\{-\frac{1}{2\sigma^2}\left[vs^2 + (\beta - \hat{\beta})'X'X(\beta - \hat{\beta})\right]\right\} \end{aligned} \quad [5.4]$$

eşitliklerinden kolayca hesaplanabilir.  $\sigma$  parametresinin marjinal sonsal dağılımı,

$$p(\sigma | y) \propto \frac{1}{\sigma^{v+1}} \exp\left\{-\frac{vs^2}{2\sigma^2}\right\} \quad v > 0 \quad [5.5]$$

eşitliği ile verilen  $v$  ve  $s^2$  parametrelili Ters Gamma dağılımıdır.  $\beta$  parametresinin marjinal sonsal dağılımı

$$p(\beta | y) \propto \left\{vs^2 + (\beta - \hat{\beta})'X'X(\beta - \hat{\beta})\right\}^{-(v+k)/2} \quad v > 0 \quad [5.6]$$

şeklinde yazılan, bu oyf  $k$ -boyutlu Çok Değişkenli Student-t dağılımıdır. Ayrıca Ortalaması

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

olup, kovaryans matrisi

$$\frac{(X'X)^{-1}vs^2}{(v-2)} \quad v > 2$$

eşitliğidir (15).

## 6. HATA TERİMLERİNİN İKİ FARKLI DAĞILIM VARSAYIMI ALTINDA ANALİZ SONUÇLARININ ÖZETİ

Bu bölümde Çok Değişkenli Student-t dağılımına sahip hata terimli çoklu regresyon modeli ile Çok Değişkenli Normal dağılıma sahip çoklu regresyon modeli analiz sonuçları özetlenmektedir.

### Model 1

$$y = X\beta + \varepsilon, \quad r(X_{n \times k}) = k < n$$

Burada,

- $y$  : gözlem vektörü  
 $\beta$  :  $(k \times 1)$  boyulu regresyon katsayıları vektörü  
 $\varepsilon$  :  $(n \times 1)$  boyutlu hata vektörüdür. Hata vektörünün dağılımı, 0 ortalama vektörü ve  $v_0 \sigma^2 / (v_0 - 2) I_n$  varyans-kovaryans matrisi ile Çok Değişkenli Student-t dağılımdır.

olup  $\varepsilon$  hata vektörünün dağılımı Çok Değişkenli Normal dağılımın Ters Gamma dağılımı ile ağırlıklandırılması ile

$$p(\varepsilon | \cdot) = \int_0^{\infty} p_N(\varepsilon | \tau) p(\tau | \cdot) d\tau \quad [6.1]$$

şeklinde elde edilir. Eş. 6.1'de adı geçen  $p_N(\varepsilon | \tau)$  ve  $p(\tau | \cdot)$  ifadeleri sırasıyla

$$p_N(\varepsilon | \tau) = \frac{1}{(2\pi\tau^2)^{n/2}} \exp\left\{-\frac{\varepsilon'\varepsilon}{2\tau^2}\right\}$$

$$p(\tau | v_0, \sigma) = \left[ \frac{2}{\Gamma(v_0/2)} \right] \left( \frac{v_0 \sigma^2}{2} \right)^{v_0/2} \tau^{-v_0+1} \exp \left\{ -\frac{v_0 \sigma^2}{2} \right\}$$

eşitlikleridir. Eş. 6.1'deki integralin sonucu

$$p(\varepsilon | \cdot) = \frac{(v_0)^{v_0/2} \Gamma((n+v_0)/2)}{\Gamma(v_0/2) \pi^{n/2} (\sigma^2)^{n/2}} \left\{ v_0 + \frac{\varepsilon' \varepsilon}{\sigma^2} \right\}^{-(n+v_0)/2}$$

olup, bu da hata vektörünün dağılımı olan Çok Değişkenli Student-t'dir. Parametrelerin EÇOB tahmin edicileri sırasıyla

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

$$\hat{\sigma}^2 = (y - X\hat{\beta})'(y - X\hat{\beta})/n$$

dir.  $\beta$  ve  $\sigma^2$  parametrelerinin bilgi içermeyen önsel dağılımları

$$p(\beta, \sigma^2) \propto \frac{1}{\sigma^2},$$

olup,  $\beta$  parametresinin sonsal dağılımı,

$$p(\beta | y, v_0) \propto \left\{ v_0 \sigma^2 + (\beta - \hat{\beta})' X'X (\beta - \hat{\beta}) \right\}^{-(v_0+k)/2}$$

olur. Bu dağılım  $\hat{\beta}$  ortalama vektörü ve  $\frac{(X'X)^{-1} v_0 \sigma^2}{(v_0 - 2)}$  varyans-kovaryans matrisi ile verilen Çok Değişkenli Student-t'dir.

$\sigma^2$  parametresinin sonsal dağılımı

$$p(\sigma^2 | y, v_0) \propto \left(\frac{\sigma^2}{s^2}\right)^{(v_0+2)} \left[1 + (v_0 \sigma^2 / v s^2)\right]^{-(v_0+v+2)}$$

oyf ile F dağılımıdır.

### Model 2

$$y = X\beta + \varepsilon, \quad r(X_{n \times k}) = k < n$$

Burada

- y : gözlem vektörü
- $\beta$  : (kx1) boyulu regresyon katsayıları vektörü
- $\varepsilon$  : (nx1) boyutlu hata vektörüdür.

olup  $\varepsilon$  hata vektörünün dağılımı

$$P(\varepsilon | \sigma) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left\{-\frac{\varepsilon'\varepsilon}{2\sigma^2}\right\}$$

ile verilen Çok Değişkenli Normal dağılımdır. y gözlem vektörünün dağılımı ise,

$y \sim N[X\beta, \sigma^2 I_n]$  olur. Parametrelerin EÇOB tahmin edicileri ise

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

$$\hat{\sigma}^2 = (y - X\hat{\beta})'(y - X\hat{\beta})/n$$

eşitlikleridir.  $\beta$  ve  $\sigma$  parametrelerinin bilgi içermeyen önsel dağılımları

$$p(\beta, \sigma) \propto \frac{1}{\sigma}$$

olup,  $\beta$  parametresinin sonsal dağılımı

$$p(\beta | y) \propto \left\{ vs^2 + (\beta - \hat{\beta})' X'X(\beta - \hat{\beta}) \right\}^{-(v+k)/2}$$

ile verilen,  $\hat{\beta}$  ortalama vektörü ve  $\frac{(X'X)^{-1} vs^2}{v-2}$  varyans-kovaryans matrisi ile Çok

Değişkenli Student-t dağılımıdır.  $\sigma^2$  parametresinin sonsal dağılımı ise

$$p(\sigma | y) \propto \frac{1}{\sigma^{v+1}} \exp \left\{ -\frac{vs^2}{2\sigma^2} \right\}$$

olan,  $v$  ve  $s^2$  parametrelili Ters Gamma dağılımıdır.

## 7. SONUÇ VE ÖNERİLER

Bu çalışmada, hata terimlerinin iki farklı dağılım varsayımı altında çoklu regresyon modeli Bayes analizleri yapılmıştır. Bölüm dörtte verilen 0 ortalama vektörü ve

$\frac{v_0 \sigma^2}{(v_0 - 2)} I_n$  varyans-kovaryans matrisi ile, Çok Değişkenli Student-t dağılımına sahip

hata terimli çoklu regresyon modelinin klasik yöntem ve Bayes yöntemiyle analizi yapılmıştır. Her iki analiz yönteminde elde edilen sonuçlar bakımından birebir paralellik gözlenmiştir.

Bu modelde hata terimleri ilişkisiz olmalarına rağmen bağımsız değildir. Regresyon katsayıları vektörünün EKK tahmin edicisi, EÇOB tahmin edicisidir. Aynı zamanda en küçük varyanslı doğrusal yansız tahmin edicidir. Hata terimlerinin Normal dağılıma sahip olmamasına rağmen ilgili momentler t ve F istatistiklerine dayanmaktadır. Bununla birlikte,  $\sigma^2$  parametresi ile ilgili sonuçlarda, ki-kare dağılımı yerine F dağılımı kullanılmıştır.

Parametrelerin bilgi içermeyen önsel dağılım ile modelin Bayes analizinde ise,  $\beta$  regresyon katsayılarının vektörünün sonsal dağılımı Çok Değişkenli Student-t'dir. EKK tahmin edicisi (regresyon katsayılarının), bu sonsal dağılımın ortalamasıdır.  $\sigma^2$  parametresinin sonsal dağılımı ise F dağılımıdır. Serbestlik derecesi büyüdüğünde hata terimlerinin dağılımı Normal dağılıma,  $vs^2/\sigma^2$ 'nin sonsal dağılımı ise, ki-kare dağılımına yaklaşır. Bu da  $\sigma^2$  parametresi ile ilgili sonuçların serbestlik derecesinin büyüklüğüne bağlı olarak değiştiğini gösterir. Ayrıca parametrelerin eşlenik önsel dağılım ile Bayes analiz sonuçları da verilmiştir.

Bölüm beşte verilen  $N[\underline{0}, \sigma^2 I_n]$  ile gösterilen Çok Değişkenli Normal dağılıma sahip çoklu regresyon modelinin analiz sonuçlarında ise  $\beta$  regresyon katsayıları vektörünün EKK tahmin edicisi, EÇOB tahmin edicisidir. Tahmin edicilerinin varyansları ise bölüm dörtteki modele göre daha küçüktür. ...

Bilgi içermeyen önsel dağılım ile modelin Bayes analizinde ise  $\beta$  regresyon katsayılarının sonsal dağılımı, Çok Değişkenli Student-t'dir. EKK ( $\beta$  parametresinin) tahmin edicisi burada da sonsal dağılımın ortalamasıdır.

$\sigma^2$  parametresi ile ilgili sonuçlarda ise birinci modele göre tamamen farklılıklar vardır. Sonsal dağılımı Ters Gamma dağılımıdır.

Kısaca belirtmek gerekirse her iki modelin analiz sonuçları benzerlik göstermektedir. Elde edilen tahmin ediciler aynı olmasına rağmen tahmin edicilerin varyansları açısından farklılıklar vardır. Birinci modeldeki tahmin edicilerin varyansları daha büyüktür .

Son olarak Çok Değişkenli Student-t dağılımına sahip hata terimleri ile ilgili elde edilen sonuçların, basit regresyon modeli içinde, geçerli olduğu görülecektir. Regresyon varsayımlarının (normallik, bağımsızlık gibi) sağlanmadığı durumlarda koşulları yerine getirmek yerine Çok Değişkenli Student-t hata terimli regresyon modeli önerilebilir.

## KAYNAKLAR

- 1- Zellner, A., Tiao, G. C., "Bayesian analysis of the regression model with autocorrelated errors", *Journal of American Statistical Association*, (59): 763-78 (1965).
- 2- Mouchart, M., "A regression model with prior information on the disturbances", *University of London*, CORE Discussion Paper No. 7116 (1971).
- 3- Newcomb, S., "A generalized theory of the combination of observations so as to obtain the best results", *American Journal of Mathematics*, (8): 394-66 (1986).
- 4- Jeffreys, H., "Scientific inference", 2nd ed., *Cambridge University Press Cambridge*, England, 1-190 (1957).
- 5- Fama, E.F. and Roll, R., "Parameter estimates for symmetric stable distributions", *Journal of the American Statistical Association*, (66): 331-8 (1971).
- 6- Box, G., Tiao, G., "A further look at robustness via Bayes theorem", *Biometrika*, (49): 419-33 (1962).
- 7- Praetz, P.D., "The distribution of share price changes", *Journal of Business*, (45): 49-55 (1972).
- 8- Johnson, N. L., Kotz, S., "Multivariate t-distributions", *Continuous Multivariate Distributions*, *John Wiley & Sons, Inc.*, New York, 132-136 (1972).
- 9- O'Hagan, A., "The linear model", *Kendall's Advanced Theory of Statistics*, 2B, *John Wiley & Sons, Inc.*, New York, 244-245 (1994).
- 10- Schmidheiny K., "OLS regression",  
<http://www.hec.unil.ch/schmidheiny/sea2/ols2up.pdf> (5.11.2003)
- 11- Dahl, G., "Proof of the Gauss-Markov theorem",  
<http://www.courses.rochester.edu/dahl/ECO231/Proof%20of%20Gauss%20Markov%20Theorem.pdf> (2002).
- 12- Ersoy N., Erbaş S., "Temel istatistik ve olasılığa giriş", *Özkan Matbaacılık*, Ankara, 346 (1996).
- 13- Billah, B., Saleh, E., "Performance of three large sample tests in the actual formation of the pretest estimators for regression model with t-error", *Journal of Applied Statistical Sciences*, 9 (3): 237-51 (2000).

- 14- Zellner, A., "Bayesian and non-bayesian analysis of the regression model with multivariate Student-t error terms", *Journal of American Statistical Association*, 71 (354): 400-404 (1976).
- 15- Zellner, A., "Bayesian analysis of regression error terms", *Journal of American Statistical Association*, 70 (349): 138-144 (1975).





**EKLER**

**Ek-1**

Bu ekte, bölüm dördte Eş. 4.19'daki eşlenik önsel dağılım ile  $\sigma^2$  ve  $\beta$  parametrelerinin marjinal sonsal dağılımlarının elde edilmeleri gösterilecektir. Bu çalışmada,

$$p(y, \tau | \beta, \sigma^2) \propto \left[ \frac{1}{\tau^n} \exp \left\{ -\frac{1}{2\tau^2} (y - X\beta)'(y - X\beta) \right\} \right] \left[ \frac{(\sigma^2)^{v_0-2}}{\tau^{v_0+1}} \exp \left\{ -\frac{v_0 \sigma^2}{2\tau^2} \right\} \right]$$

[EK.1]

ile yazılan ortak oyf kullanılmıştır. Yukarıdaki dağılım  $\tau$  ( $0 < \tau < \infty$ ) üzerinden integrallenirse Eş. 4.7'deki olabilirlik fonksiyonu elde edilir. Eş. 4.19 ile verilen önsel dağılımın analizi için  $p_F(\sigma^2 | \cdot)$  ifadesi, Eş. 4.20 ve  $\Theta$  yardımcı parametre olmak üzere

$$p(\beta, \sigma^2, \Theta | \cdot) \propto p_F(\sigma^2 | \cdot) \left[ \frac{1}{\Theta^k} \exp \left\{ -\frac{1}{2\Theta^2} (\beta - \bar{\beta})' A (\beta - \bar{\beta}) \right\} \right]$$

$$\left[ \frac{(\bar{\sigma}_a^2)^{\bar{v}_a-2}}{\Theta^{\bar{v}_a+1}} \exp \left\{ -\frac{\bar{v}_a \bar{\sigma}_a^2}{2\Theta^2} \right\} \right]$$

[EK.2]

olarak yazılabilir. Eş. EK.2'deki oyf  $\Theta$  parametresine göre integrallenirse, Eş. 4.19'da verilen  $\beta$  ve  $\sigma^2$  parametrelerinin marjinal önsel dağılımları elde edilir.

Eş. EK.1 ve Eş. EK.2 ile verilen ifadelerin çarpımı,  $\lambda = \tau^2 / \Theta^2$  olmak üzere

$$p_F(\sigma^2 | \cdot) (\sigma^2)^{v_0-2} (\bar{\sigma}_a^2)^{\bar{v}_a-2} \left[ \frac{1}{\tau^{n+v_0+1}} \exp \left\{ -\frac{vs^2 + v_0 \sigma^2}{2\tau^2} \right\} \right] \left[ \frac{1}{\Theta^{k+\bar{v}_a+1}} \exp \left\{ -\frac{\bar{v}_a \bar{\sigma}_a^2}{2\Theta^2} \right\} \right]$$

$$\cdot \exp \left\{ -\frac{1}{2\tau^2} \left[ (\beta - \hat{\beta})' X' X (\beta - \hat{\beta}) + (\tau^2 / \Theta^2) (\beta - \hat{\beta})' A (\beta - \hat{\beta}) \right] \right\} \quad [\text{EK.3}]$$

olarak yazılır. Bu eşitlik  $\sigma^2$  ve  $\beta$  parametrelerinin marjinal sonsal dağılımlarını bulmak için kullanılır. Eş. EK.3'deki son üstel ifade  $A = X'X$ ,  $M = X'X + \lambda A$  ve  $\bar{\beta}_\lambda = M^{-1}(X'X\hat{\beta} + \lambda A\bar{\beta})$  olmak üzere

$$\beta - \hat{\beta})' X' X (\beta - \hat{\beta}) + \lambda (\beta - \hat{\beta})' A (\beta - \hat{\beta}) = (\beta - \bar{\beta}_\lambda)' M (\beta - \bar{\beta}_\lambda) + \hat{\beta}' X' X \hat{\beta} + \lambda \bar{\beta}' A \bar{\beta} - \bar{\beta}_\lambda' M \bar{\beta}_\lambda \quad [\text{EK.4}]$$

olarak yazılabilir. Eş. EK.4'deki ifade, Eş. EK.3'de yerine yazılıp  $\beta$  parametresine göre integrellenirse elde edilen sonuç

$$\frac{p_F(\sigma^2 | \cdot) (\sigma^2)^{v_0 - 2} (\bar{\sigma}_a^2)^{\bar{v}_a - 2}}{\tau^{n+v_0+1}} \exp \left\{ -\frac{vs^2 + v_0\sigma^2}{2\tau^2} \right\} \frac{1}{\Theta^{k+\bar{v}_a+1}} \exp \left\{ -\frac{\bar{v}_a \bar{\sigma}_a^2}{2\Theta^2} \right\} \\ \left| \frac{M}{\tau^2} \right|^{-1,2} \exp \left\{ -\frac{1}{2\tau^2} \left[ \hat{\beta}' X' X \hat{\beta} + \lambda \bar{\beta}' A \bar{\beta} - \bar{\beta}_\lambda' M \bar{\beta}_\lambda \right] \right\} \quad [\text{EK.5}]$$

olur. Eş. EK.5 ifadesinde  $\sigma^2$ ,  $\tau$  ve  $\Theta$  parametrelerini  $\sigma^2$ ,  $\tau$  ve  $\lambda = \tau^2 / \Theta^2$  parametreleri olarak değiştirilirse bu dönüşümün jakobiyeni  $\lambda^{1/2}$  ile orantılıdır. Böylece Eş. EK.5  $\sigma^2$ ,  $\tau$  ve  $\lambda$  terimleriyle

$$\frac{p_F(\sigma^2 | \cdot) (\sigma^2)^{v_0 - 2} (\bar{\sigma}_a^2)^{\bar{v}_a - 2}}{\tau^{n+v_0+\bar{v}_a+3}} |M|^{-1,2} \lambda^{(k+\bar{v}_a+1)/2} \\ \cdot \exp \left\{ -\frac{1}{2\tau^2} \left[ vs^2 + v_0\sigma^2 + \bar{v}_a \bar{\sigma}_a^2 \lambda + \hat{\beta}' X' X \hat{\beta} + \lambda \bar{\beta}' A \bar{\beta} - \bar{\beta}_\lambda' M \bar{\beta}_\lambda \right] \right\} \quad [\text{EK.6}]$$

olarak yazılır.  $\lambda$  ve  $\sigma^2$  parametrelerinin marjinal sonsal dağılımlarını bulmak için

Eş. EK.6  $\tau$  parametresine göre integrallenirse,  $\bar{\sigma}_a^2 = \frac{(v_a s_a^2 + v_0 \sigma^2)}{\bar{v}_a}$  ve  $\bar{v}_a = v_0 + v_a$

olmak üzere

$$\frac{p_F(\sigma^2 | \cdot) \sigma^{2v_0-2} (\bar{\sigma}_a^2)^{\bar{v}_a-2} |M|^{-1/2} \lambda^{(k+\bar{v}_a+4)/2}}{\left\{ v s^2 + v_0 \sigma^2 + \bar{v}_a \bar{\sigma}_a^2 \lambda + \hat{\beta}' X' X \hat{\beta} + \lambda \bar{\beta}' A \bar{\beta} - \bar{\beta}'_z M \bar{\beta}_z \right\}^{(n+v_0+\bar{v}_a+2)/2}} \quad [\text{EK.7}]$$

elde edilir. Eş. 4.20 ile Eş. EK.7,

$$p(\sigma^2, \lambda | y, \cdot) = p_1(\sigma^2 | \lambda, y, \cdot) p_2(\lambda | y, \cdot) \quad [\text{EK.8}]$$

olarak yazılabilir. Yukarıdaki eşitlikte adı geçen ifadeler

$g(\lambda) = \lambda(v_a s_a^2 + \bar{\beta}' A \bar{\beta}) - \bar{\beta}'_z M \bar{\beta}_z + \hat{\beta}' X' X \hat{\beta}$  olmak üzere

$$p_1(\sigma^2 | \lambda, y, \cdot) \propto \left[ \frac{(1+\lambda)v_0 \sigma^2}{v s^2 + g(\lambda)} \right]^{2v_0-2} \left[ 1 + \frac{(1+\lambda)v_0 \sigma^2}{v s^2 + g(\lambda)} \right]^{-1/2} \quad [\text{EK.9}]$$

ve

$$p_2(\lambda | y, \cdot) \propto \frac{\lambda^{(k+\bar{v}_a-4)/2}}{|M|^{-1/2} (1+\lambda)^{v_0-2} [v s^2 + g(\lambda)]^{(n+\bar{v}_a+4)/2}} \quad [\text{EK.10}]$$

olarak yazılabilir. Eş. Ek.9 ifadesinden  $\sigma^2$  parametresinin koşullu sonsal dağılımının F oyl olduğu görülür.

$\beta$  parametresinin marjinal sonsal dağılımını  $c_z = \hat{\beta}' X' X \hat{\beta} + \bar{\beta}' A \bar{\beta} - \bar{\beta}'_z M \bar{\beta}_z$  olmak üzere

$$\frac{(\sigma^2)^{v_0-1}}{\Theta^{k+\bar{v}_i+1}\tau^{n+v_0+1}} \exp\left\{-\frac{1}{2\tau^2}\left[vs^2 + v_0\sigma^2 + (\tau^2/\Theta^2)(v_a s_a^2 + v_0\sigma^2)\right]\right\}$$

$$\exp\left\{-\frac{1}{2\tau^2}\left[c_{\lambda} + (\beta - \bar{\beta}_{\lambda})'M(\beta - \bar{\beta}_{\lambda})\right]\right\} \quad [\text{EK.11}]$$

olarak yazılan ifadedeki  $\sigma^2$ ,  $\tau$  ve  $\Theta$  parametreleri yerine  $\sigma^2$ ,  $\tau$  ve  $\lambda = \tau^2/\Theta^2$  olarak değiştirilip  $\sigma^2$  ve  $\tau$  parametresine göre integral alınır, elde edilen sonuç

$$p(\beta, \lambda | y, \cdot) \propto \frac{\lambda^{(k+v_0+4)/2} (1+\lambda)^{v_0}}{\left\{vs^2 + \lambda v_a s_a^2 + c_{\lambda} + (\beta - \bar{\beta}_{\lambda})'M(\beta - \bar{\beta}_{\lambda})\right\}^{(n+k+v_0+1)/2}} \quad [\text{EK.12}]$$

olup, bu oyf'da  $\lambda$  biliniyorken  $\beta$  parametresinin koşullu sonsal dağılımının Çok Değişkenli Student-t olduğu görülür. Bu oyf'ndan herhangi bir  $\beta_i$  ve  $\lambda$  parametrelerinin ortak sonsal dağılımları bulunabildiğinden,  $\beta_i$ 'nin marjinal sonsal dağılımında bulmak mümkündür (14).

## ÖZGEÇMİŞ

Demet HAN 25.04.1978 tarihinde Yozgat'ta doğdu. İlk, orta ve lise öğrenimini Ankara'da tamamladı. 2000 yılında Ankara Üniversitesi Fen Fakültesi İstatistik Bölümünden mezun oldu.



## OLS Regression

### 1 The Econometric Model

The multiple linear Regression model is of the following form.

A single observation follows

$$y_i = \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_K x_{iK} + \varepsilon_i \quad i = 1, \dots, N$$

$$y_i = \begin{bmatrix} x_{i1} & \dots & x_{iK} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_K \end{bmatrix} + \varepsilon_i$$

$$1 \times 1 \quad 1 \times K \quad K \times 1 \quad 1 \times 1$$

and the whole sample follows

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1K} \\ x_{21} & x_{22} & \dots & x_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ x_{N1} & x_{N2} & \dots & x_{NK} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_K \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_N \end{bmatrix}$$

$$N \times 1 \quad N \times K \quad X \quad K \times 1 \quad \beta \quad N \times 1 \quad \varepsilon$$

$$= \begin{bmatrix} x_1 & \dots & x_K \end{bmatrix} \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_K \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_N \end{bmatrix} = \begin{bmatrix} x'_1 \\ \vdots \\ x'_N \end{bmatrix} \cdot \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_K \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_N \end{bmatrix}$$

where a set of assumptions is an integral part of model.

### 2 Assumptions

There are several sets of assumptions that complete the description of the statistical model. Several of the following assumptions are therefore formulated in different alternatives. Table 1 links different sets assumptions with the properties of the OLS estimator.

- A1: Correct functional form (linearity):  
 $y = X\beta + \varepsilon$ , where  $E(\varepsilon) = 0$
- A2: Data behaviour:
- $\text{rank}(X) = K < N$  (full rank)
  - $\text{plim}(\frac{1}{N}X'X) = Q$  and  $\text{plim}(\frac{1}{N}X'\Omega X)$  are positive definite
- A3: Exogeneity of independent variables
- $X$  is fixed (non-random)
  - $\varepsilon \perp X$  (independent)
  - $E(\varepsilon|X) = 0$  (mean independent)
  - $\text{cov}(X, \varepsilon) = 0$  (uncorrelated)
- A4: Spherical or non-spherical disturbances
- $E(\varepsilon'\varepsilon|X) = \sigma^2 I_N$  (homoscedastic and no autocorrelation)
  - $E(\varepsilon'\varepsilon|X) = \sigma^2 \Omega$ ,  $\Omega = \text{diag}(\sigma_1^2, \dots, \sigma_N^2) \neq I_N$  (heteroscedastic)
  - $E(\varepsilon'\varepsilon|X) = \sigma^2 \Omega$ ,  $\Omega \neq I_N$ , (autocorrelation)
  - "some sort of independence across observations"
- A5: Normality:  
 $\varepsilon|X \sim N(0, \Sigma)$

### 3 Estimation with OLS

Ordinary least squares (OLS) means minimizing the squared distance between the observed and the predicted dependent variable  $y$ :

$$S(\alpha, \beta) = \sum_{i=1}^N e_i^2 = e'e = (y - X\beta)'(y - X\beta) \rightarrow \min_{\beta, \sigma^2}$$

The resulting OLS estimator of  $\beta$  and  $\sigma$  is:

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

$$s^2 = \hat{\sigma}^2 = \frac{e'e}{N-K}$$

where  $e = y - \hat{y} = y - X\hat{\beta}$ .

### 4 Small Sample Properties of the OLS Estimator

The following small sample properties hold under the classical assumptions A1, A2a, A3c, A4a and A5:

- The OLS estimator of  $\beta$  is unbiased:

$$E(\hat{\beta}|X) = \beta$$

- Gauß-Markov-Theorem:  $\hat{\beta}$  is BLUE (best linear unbiased)
- The variance of the OLS estimator is given by:
 
$$V(\hat{\beta}|X) = \sigma^2 (X'X)^{-1}$$
- The OLS estimator of  $\sigma^2$  is unbiased:

$$E(\hat{\sigma}^2|X) = E s^2|X = \sigma^2$$

- The OLS estimator is normally distributed:

$$\hat{\beta}|X \sim N(\beta, \hat{\sigma}^2 (X'X)^{-1})$$

- $\hat{\beta}$  and  $\hat{\sigma}^2$  are independent

- The distribution of  $\hat{\sigma}^2$  follows

$$(T-K) s^2 / \sigma^2 | X \sim \chi_{T-K}^2$$

### 5 Tests in Small Samples

Assume again A1, A2a, A3c, A4a and A5, then the  $t$  statistic of a particular  $\beta_k$  follows a  $t$  distribution with  $N - K$  degrees of freedom:

$$\frac{\hat{\beta}_k - \beta_k}{\sqrt{s^2 (X'X)^{-1}_{kk}}} \sim t_{N-K}$$

where  $(X'X)^{-1}_{kk}$  is the scalar in the  $k$ -th row and  $k$ -th column of  $(X'X)^{-1}$ .

The  $F$  statistic to simultaneously test  $J$  linear restrictions  $R\beta = q$  follows an  $F$  distribution with  $J$  and  $N - K$  degrees of freedom:

$$\frac{(R\hat{\beta} - q)' [R s^2 (X'X)^{-1} R']^{-1} (R\hat{\beta} - q)}{J} \sim F_{J, N-K}$$

### 6 Asymptotic Properties of the OLS Estimator

The following large sample properties hold under the assumptions A1, A2a, A2b, A3d, A4a and A4d:

- The OLS estimator is consistent:

$$\text{plim } \hat{\beta} = \beta$$

- The OLS estimator is asymptotically normally distributed:

$$\sqrt{N}(\hat{\beta} - \beta) \xrightarrow{d} N(0, \sigma^2 Q^{-1})$$

- The OLS estimator is therefore approximately normally distributed samples in large, but finite samples:

$$\hat{\beta} \overset{\Delta}{\sim} N\left(\beta, \frac{\sigma^2}{N} Q^{-1}\right)$$

where  $\sigma^2$  can be replaced by the consistent estimator  $s^2 = e'e/(N-K)$  and  $Q^{-1}/N$  by  $(X'X)^{-1}$ .

### 7 Asymptotic Tests

Several asymptotic test can be performed under the assumptions A1, A2a, A2b, A3d, A4a and A4d. Asymptotic tests do not require the normality assumption A5. However, they are only approximately valid and require "large" (which is difficult to define) samples.

The t statistic of a particular  $\beta_k$  follows *approximately* the standard normal distribution:

$$t_k = \frac{\hat{\beta}_k - \beta_k}{\sqrt{s^2 (X'X)^{-1}_{kk}}} \overset{\Delta}{\sim} N(0, 1)$$

The Wald statistic to simultaneously test  $J$  linear restrictions  $R\beta = q$  follows approximately an  $\chi^2$  distribution with  $J$  degrees of freedom:

$$W = J F = (R\hat{\beta} - q)' [R^2 (X'X)^{-1} R']^{-1} (R\hat{\beta} - q) \overset{\Delta}{\sim} \chi^2$$

### References

Greene, William H. (2003), *Econometric Analysis*, 5th ed., Prentice Hall, chapters 2-5.

Table 1: Properties of the OLS estimator.

case	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]	[13]	[14]	[15]
<i>Assumptions</i>															
A1: linearity	✓	✓	✓	✓	✓	✓	✓	✓	×	✓	✓	✓	✓	✓	✓
A2a: full rank	✓	✓	✓	✓	✓	✓	✓	✓	✓	×	✓	✓	✓	✓	✓
A2b: plim				✓	✓	✓	✓	×	✓	✓	✓	✓	✓	✓	✓
A3a: fixed	✓						×				×				
A3b: independent	✓	✓					×				×				
A3c: mean indep.	✓	✓	✓	✓	✓		×	✓			×	✓	✓	✓	✓
A3d: uncorrelated	✓	✓	✓	✓	✓	✓	✓	✓			×		✓	✓	✓
A4a: homoscedastic	✓	✓	✓	✓	✓	✓	✓	✓				×	×	×	×
A4b: heteroscedastic	×	×	×	×	×	×	×	×				✓	✓	✓	✓
A4c: autocorrelated	×	×	×	×	×	×	×	×				×	×	×	×
A4d: independent obs.				✓	✓	✓	✓	✓				✓	✓	×	×
A5: normality	✓	✓	✓	✓	×							✓	×	✓	×
<i>Small sample properties of <math>\hat{\beta}</math></i>															
BL: best linear	✓	✓	✓	✓	✓		×	✓	×	-	×	×	×	×	×
U: $\hat{\beta}$ unbiased	✓	✓	✓	✓	✓		×	✓	×	-	×	✓	✓	✓	✓
N: $\hat{\beta}$ normally distributed	✓	✓	✓	✓	×					-	-	✓	×	✓	×
<i>Large sample properties of <math>\hat{\beta}</math></i>															
$\hat{\beta}$ asymptotically efficient				✓	×	×	×	×	×	-	×	×	×	×	×
C: $\hat{\beta}$ consistent				✓	✓	✓	✓	×	×	-	×	✓	✓	✓	✓
AN: $\hat{\beta}$ asymptotically normal				✓	✓	✓	✓	×	×	-	-	✓	✓	✓	✓

Notes: empty = not required, ✓ = fulfilled, × = violated, - = inexistent, \* = needs more (technical) assumptions.

## ECONOMETRICS

Economics 231W  
University of Rochester

Gordon Dahl  
Spring 2002

### Proof of the Gauss Markov Theorem:

The Gauss Markov Theorem says: "Within the class of linear unbiased estimators, the least squares estimator has minimum variance." Another way of saying this is that the least squares estimator is the Best Linear Unbiased Estimator (BLUE).

Consider the regression equation

$$y_i = \alpha + \beta x_i + \varepsilon_i$$

We will do the proof of the Gauss Markov theorem for the slope coefficient.

#### Proof:

Consider a linear estimator for  $\beta$ . Since we are interested in the slope coefficient only, all variables will be written as deviations from means. If  $\hat{\beta}$  is a linear function of the random variable  $y$ , it must be of the form

$$(1) \quad \hat{\beta} = \sum c_i (y_i - \bar{y}) \quad \text{for constants } c_i.$$

Our theorem concerns unbiased estimators, so we require that  $E[\hat{\beta}] = \beta$ .

$$\begin{aligned} E[\hat{\beta}] &= \sum c_i E[y_i - \bar{y}] \\ &= \sum c_i (\beta(x_i - \bar{x})) \\ &= \beta \sum c_i (x_i - \bar{x}) \\ &= \beta \end{aligned}$$

So for our linear estimator to be unbiased, the following condition must hold:

$$(2) \quad \sum c_i (x_i - \bar{x}) = 1$$

The variance of  $\hat{\beta}$  is

$$\begin{aligned} \text{Var}(\hat{\beta}) &= \text{Var}(\sum c_i (y_i - \bar{y})) \\ &= \sum c_i^2 \text{Var}(y_i - \bar{y}) \\ &= (\sum c_i^2) \sigma^2 \end{aligned}$$

Since  $\sigma^2$  is fixed we need only minimize  $\sum c_i^2$ .

Thus, we need to minimize  $\sum c_i^2$  subject to our unbiasedness restriction (2). Forming the Lagrangian:

$$(3) \quad L = \sum c_i^2 - \lambda (\sum c_i (x_i - \bar{x}) - 1)$$

The first order conditions (FOC) for a minimum are:

$$(4) \quad \frac{\partial L}{\partial c_i} = 2c_i - \lambda(x_i - \bar{x}) = 0$$

$$(5) \quad \frac{\partial L}{\partial \lambda} = \sum c_i (x_i - \bar{x}) - 1 = 0$$

We must solve these equations to get an expression for the  $c_i$ 's.

Equation (4) implies:

$$(6) \quad c_i = (\lambda/2)(x_i - \bar{x}) \quad \forall i$$

To solve for  $\lambda$  substitute (6) into (5):

$$(7) \quad \begin{aligned} \sum \{(\lambda/2)(x_i - \bar{x})(x_i - \bar{x})\} &= 1 \\ \lambda &= 2/(\sum(x_i - \bar{x})^2) \end{aligned}$$

Substituting (7) into (6):

$$\begin{aligned} c_i &= (1/2)(2/(\sum(x_i - \bar{x})^2))(x_i - \bar{x}) \\ &= (x_i - \bar{x})/\sum(x_i - \bar{x})^2 \end{aligned}$$

But notice that this gives the OLS estimator!

$$\begin{aligned} \hat{\beta} &= \sum c_i (y_i - \bar{y}) \\ &= \sum (x_i - \bar{x})(y_i - \bar{y}) / \sum (x_i - \bar{x})^2 \end{aligned}$$

Hence, we have shown that the least squares estimator is BLUE (best linear unbiased estimator).

