

**AYKIRI GÖZLEMLERİN GRUPLAR HALİNDE VARLIĞININ
ÇOK DEĞİŞKENLİ ÇOKLU REGRESYON ANALİZİNDE
BAYESGİL YAKLAŞIM İLE İRDELENMESİ**

UFUK EKİZ

**DOKTORA TEZİ
(İSTATİSTİK)**

133419

**GAZİ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

**Mayıs 2003
ANKARA**

**TC YÜKSEKÖĞRETİM KURULU
BİLİMSEL ARAŞTIRMA ENSTİTÜSÜ**

133419

Ufuk EKİZ tarafından hazırlanan AYKIRI GÖZLEMLERİN GRUPLAR HALİNDE VARLIĞININ ÇOK DEĞİŞKENLİ ÇOKLU REGRESYON ANALİZİNDE BAYESGİL YAKLAŞIM İLE İRDELENMESİ adlı bu tezin Doktora tezi olarak uygun olduğunu onaylarım.

Prof. Dr. Müslim EKİNİ.....

Tez Yöneticisi

Bu çalışma, jürimiz tarafından.....*İstatistik*.....Anabilim Dalında
...*Doktora*.....tezi olarak kabul edilmiştir.

Başkan : Prof. Dr. Süleyman GÜNAY

Üye : Prof. Dr. Alptekin ESİN

Üye : Prof. Dr. Fikri Öztürk

Üye : Prof. Dr. Müslim EKİNİ

Üye : Prof. Dr. Hamza GANİGAM

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

İÇİNDEKİLER

	Sayfa
ÖZET.....	i
ABSTRACT.....	ii
TEŞEKKÜR.....	iii
ŞEKİLLERİN LİSTESİ.....	iv
1. GİRİŞ.....	1
2. AYKIRI GÖZLEMLERLE İLGİLİ İSTATİSTİKSEL YÖNTEMLER.....	5
2.1. Aykırı Gözlemleri Uygun Hale Getirme Görüşü	5
2.2. Uyuşmazlık Testi Görüşü.....	10
2.2.1. Hipotez testi prosedürüne bağlı uyumsuzluk testi görüşü.....	12
2.2.2. Sezgisel uyumsuzluk testi görüşü.....	21
3. ÇOK DEĞİŞKENLİ VERİDE AYKIRI GÖZLEM	31
4. AYKIRI GÖZLEMLERİN TEK TEK VARLIĞININ İRDELENMESİNE ÇOK DEĞİŞKENLİ ÇOKLU REGRESYON ANALİZİNDE BAYESGİL YAKLAŞIM.....	36
4.1. Çok Değişkenli-Çoklu Regresyon Analizinde Aykırı Gözlemlerin Uyuşmazlıklarının Tek Tek İrdelendiği Bayesgil Bir Yaklaşım.....	36
4.2. Gerçekleşmiş Hatanın Karesel Formu R_i 'nin Sonsal Dağılımı.....	40
4.3. Önsel Dağılım.....	40
4.4. Jeffrey Kuralı.....	41
4.5. Olabilirlik Fonksiyonu.....	45
4.6. $p(\Theta / \Sigma, Y)$ Üzerinden i.nci Gözleme İlişkin Gerçekleşmiş Hatanın Sonsal Dağılımı.....	47
4.7. i.nci Gözleme İlişkin Gerçekleşmiş Hatanın Karesel Miktarının Sonsal Dağılımı.....	50

5. GİZLEME (Masking) VE YANILGIYA-DÜŞÜRME (Swamping) PROBLEMLERİ.....	52
6. AYKIRI GÖZLEMLERİN GRUPLAR HALİNDE VARLIĞININ İRDELENMESİNE ÇOK DEĞİŞKENLİ ÇOKLU REGRESYON ANALİZİNDE BAYESGİL YAKLAŞIM.....	55
7. UYGULAMA.....	63
8. SONUÇ VE ÖNERİLER.....	66
KAYNAKLAR.....	68
EKLER.....	71
ÖZGEÇMİŞ.....	85



**AYKIRI GÖZLEMLERİN GRUPLAR HALİNDE VARLIĞININ
ÇOK DEĞİŞKENLİ ÇOKLU REGRESYON ANALİZİNDE
BAYESGİL YAKLAŞIM İLE İRDELENMESİ
(Doktora Tezi)**

Ufuk EKİZ

**GAZİ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

Mayıs 2003

ÖZET

Çok değişkenli-çoklu regresyon analizinde, tesadüfi örneğin herhangi bir gerçekleşişinin aykırı gözlem içerip içermediğine ilişkin pek çok yöntem literatürde yer almaktadır. Bu tür gözlemlerin belirlenmesini sağlayacak bayesgil yaklaşımlar kabaca üç başlık altında toplanmaktadır. Bu başlıklardan birine uygun olarak, gerçekleşmiş ancak gözlenemeyen hatanın karesel formunun sonsal dağılımı üzerinden aykırı gözlemlerin tek tek irdelenmesi fikrine yönelik bayesgil bir yaklaşım, Varbanov(1998) tarafından ortaya atılmıştır. Tesadüfi örneğin herhangi bir gerçekleşişinde yer alan gözlemlerin tek tek aykırı olup olmadıklarının irdelenmesi, gizleme ve yanılgıya düşürme problemlerinin incelenmesi açısından yeterli değildir. Bu problemlerin varlığını ortaya koyabilmek, gözlemlerin gruplar halinde aykırı olup olmadıklarının irdelenmesi ile mümkün olabilmektedir. Bu amaçla Varbanov tarafından ileri sürülen yöntem, gruplar halinde gözlemlerin aykırı olduklarını irdememize imkan verecek şekilde genişletilmiştir.

Bilim Kodu : 406.01.01

Anahtar Kelimeler : Gerçekleşmiş ancak gözlenemeyen hata, Gizleme ve Yanılgıya-Düşürme Problemleri, Aykırı Gözlem

Sayfa Adedi : 85

Tez Yöneticisi : Prof.Dr.Müslim EKNİ

**A BAYESIAN METHOD TO INVESTIGATE GROUP OF OBSERVATIONS
WHICH IS AN OUTLIER IN MULTIVARIATE-MULTIPLE LINEAR
REGRESSION ANALYSIS**

(Dr.Thesis)

Ufuk EKİZ

**GAZİ UNIVERSITY
INSTITUTE OF SCIENCE AND TECHNOLOGY**

May 2003

ABSTRACT

In multivariate-multiple regression analysis, there are many methods in the literature to search whether there are outliers or not in the sample. Bayesian methods are used to and summarized in three groups. Varbanov(1998) proposed a bayesian method to investigate outlier observation individually. He proposed a posterior distribution for the mean squared error from of unobserved but realized error. To handle observations individually it is not satisfactory investigate masking and swamping problems. To show the presence of this problems, it is necessary to handle the observations as group to see whether the groups are outliers or not. Here, Varbanov's method is expanded to the multivariate-multiple regression and the expanded method is applied to the masking and swamping problems.

Science Code : 406.01.01

**Key Words : Realized but unobserved error, Masking and Swamping
Problems, outlier observation**

Page number : 85

Adviser : Prof.Dr.Müslim EKNİ

TEŞEKKÜR

Tez izleme komitemde yer alan danışmanım sayın Prof.Dr.Müslim EKNİ, sayın Prof.Dr.Süleyman GÜNAY, sayın Prof.Dr.Hamza GAMGAM ve jürimde yer alan sayın hocalarıma, ayrıca sayın hocam Yrd.Doç.Dr.Celal AYDIN başta olmak üzere tüm hocalarım, arkadaşlarım ve aileme çalışmalarında yapmış oldukları değerli katkılarından ve manevi desteklerinden dolayı teşekkür ederim.



ŞEKİLLERİN LİSTESİ

Şekil	Sayfa
Şekil 4.1 φ metriğindeki olabilirlik fonksiyonları.....	42
Şekil 4.2 ϕ metriğindeki olabilirlik fonksiyonları.....	43
Şekil 5.1 Gizleme.....	52
Şekil 5.2 Yanılgıya-Düşürme.....	53



1.GİRİŞ

İstatistiksel bir çalışmada, tesadüfi örneğin herhangi bir gerçekleşişinde yer alan gözlemlerden bir veya bir kaç araştırmanın beklentileri üzerinde şüphe doğurabilir. Araştırmacının beklentisi ya da başlangıçtaki varsayımı, gözlemlerin tamamının aynı yapıya ait olduğudur. Aynı yapıya ait olmadığı düşünölen, diğerleri ile tutarsız olan, sıra dışı gözlem yada gözlemler verinin kaynağı hakkındaki bilgiyi düşürmekte veya çarpıtmaktadır (veya tesadüfi değişken için öngördüğümüz dağılımı tekrar gözden geçirmemizi gerektirmektedir). Pek çok çalışmada bu tarz gözlemlerin varlığını sezebilmek çok zor bazen de imkansız olabilir. Bu sebepten dolayı, tesadüfi örneğin herhangi bir gerçekleşişinin, aykırı gözlem (outlier) olarak isimlendirilen bu gözlem ya da gözlemleri içerip içermediğini araştırmak, içeriyorsa bu gözlemleri tesadüfi örneğin mevcut gerçekleşişinden çıkararak veriyi varsaydığımız kaynağına uygun hale getirmek ya da istatistiksel analizler üzerindeki etkisini azaltacak yöntemler uygulamak kaçınılmaz olmaktadır.

“Aynı yapıya ait olmayan gözlem”, “diğerlerinden farklı olduğu düşünölen gözlem” şeklindeki genel tanımların dışında, çok daha özel bir tanımlamayı aykırı gözlem için yapabilmek her probleme ve her araştırmacıya göre değişmektedir. Probleme çözüm getirinceye kadarki süreçte karşılaşılan sorulara verilecek yanıtlar aykırı gözlem için yeni bir tanımlama olacaktır.

Cevap aranması gereken ilk soru diğerlerinden farklı olduğu veya onlarla aynı yapıya ait olmadığı düşünölen gözlemin karakteristiğinin ne olduğudur. Yani aykırı olduğu düşünölen gözlemin karakteristiği tesadüfi mi yoksa deterministik (her olayın başka olayların kaçınılmaz sonucu olarak ortaya çıktığı düşüncesi) midir? Aykırı gözlemin karakteristiğinin deterministik olup olmadığını incelemek çoğu zaman mümkün olmayabilir. Mümkün olduğu durumlarda gözlemi düzeltmek, çıkarmak ya da deneyi tekrarlamak yoluyla problemi giderebiliriz. Mümkün olmadığı durumlarda ise, gerçekte deterministik bir karakteristiğin elemanı olan aykırı gözlemi karakteristiği tesadüfiymiş gibi değerlendirmek zorunda kalabiliriz. Aykırı gözlemin karakteristiğinin tesadüfi olduğunun düşünölmesi problemin istatistiksel açıdan

değerlendirilmesini gerektirmektedir. Tüm gözlemlerin aynı yapının elemanları olması tanımı, istatistikte gözlemlerin her birinin tesadüfi değişken için ön görülen model tarafından üretilmiş olmasına veya modele uygunluk göstermesine karşılık gelir. Öngörülen modele uymadığı düşünülen gözlem gerçekten aykırıdır, yoksa öngörülen model tesadüfi değişken için yanlış bir seçim midir? İstatistikçinin, gözlemin öngörülen modele uygun olup olmadığı sorusuna cevap ararken kullandığı araç bu amaca ne denli hizmet etmektedir? Eğer kullanılan aracın amaca iyi hizmet ettiği ve öngörülen modelin tesadüfi değişkene uygun olduğu düşünülüyor ise, aykırı olduğu düşünülen gözlem hakkında ne karar verilmelidir. Yani gözlem yok mu varsayılmalı yoksa yığın hakkında bize negatif katkı getiriyor düşüncesinden hareketle bu katkıyı en aza indirecek Robust (sağlam) yöntemlere benzer yöntemlerin kullanılması mı tercih edilmelidir?

Öngörülen model tesadüfi değişken için yanlış bir seçim değilse, tesadüfi karakteristiğe sahip bir gözleme aykırıdır diyebilmek için kullanılan yöntemler, gözlemin öngörülen model için çok uç olduğu veya öngörülen model dışında bir modele uyduğu görüşüne dayanmaktadırlar.

Sadece öngörülen modelden hareketle aykırı gözlemi belirlemeye çalışmak görüşü, uç(extreme) gözlemleri dikkate almaktadırlar. Her uç gözlemin aykırı olduğunu söylemek doğru olmaz. Dolayısıyla tesadüfi değişken için bir kritik değerin belirlenmesi ve bu kritik değere göre uç gözlemin aykırı olup olmadığının analiz edilmesi gerekmektedir. Yani tüm gözlemler öngörülen model tarafından üretilmiştir ancak uç gözlemlerden bazılarının belli bir ölçü üzerinden diğerlerinden daha da uç olduğu tespit edilebilir. Belirli bir sınırdan daha uçta yer alan bu uç gözlemlerde aykırı olarak tanımlanmaktadırlar.

Öngörülen model dışında bir başka modele daha uygun olan gözlem kirletici(contaminant) gözlem olarak tanımlanmaktadır. Yani tüm gözlemler için öngörülen model F olmak üzere, eğer herhangi bir gözlem problemin kendine has özelliklerine uygun olarak belirlenmiş bir G alternatif modeline daha çok uygunluk gösteriyor ise bu gözlem kirletici olarak isimlendirilmektedir. Kirletici gözlem uç bir

gözlem olabileceği gibi merkezde yer alan bir gözlem de olabilir. Dolayısıyla aykırı gözlem için alternatif bir modelin tanımlanması görüşüne göre, her kirletici gözlem aykırıdır diyemeyiz. Aykırı gözlem için alternatif bir modelin tanımlandığı yöntemlerin bu ayrımı ortaya koyması gerekmektedir. Yani bu görüşe göre bir gözleme aykırıdır diyebilmek için gözlemin hem uç hem de kirletici olması gerekmektedir.

Aykırı gözlemlerle ilgili yukarda kısaca verilen görüşlerden hangisini benimsiyor olursak olalım, ilgilenilen istatistiksel analiz ne denli karmaşıklşırsa aykırı gözlemi ve onu ortaya çıkaracak yöntemleri belirlemek daha da zorlaşmaktadır. Örneğin regresyon analizinde aykırı gözlemi belirlemek için önerilecek yöntemlerin ortaya atılması, işlem karmaşası bir tarafa, uygulamacının regresyon analizi yapmaktaki asıl amacımızın ne olduğuna da çok bağlı olarak değişmektedir. Amacımız kestirim yapmaksa regresyon modelini destekleyen bir uç gözlemi aykırıdır diye yok saymak(sadece uç gözlem olup olmaması kriteri ile hareket edilirse) veya birtakım sağlam yöntemlerle bu gözlemin sağladığı bilgiyi düşürmek istemeyiz. Böyle bir durumda regresyon doğrusunun parametre tahminleri üzerinde büyük değişmeye sebep olacak uç bir gözlemi aykırı diye niteleyecek bir yöntemi önermek daha mantıklı olacaktır.

Aykırı gözlemin öngörülen model tarafından üretilmiş aşırı uç olan bir gözlem olması tanımından hareketle, çok değişkenli tesadüfi örneğin herhangi bir gerçekleşişinin aykırı gözlem içermediği araştırılırken ortaya çıkabilecek zorluklardan biri uç gözlemin tespiti için gözlemlerin sıralanmasının gerekliliğidir. Her bir değişken üzerinden sıralama yaparak uç gözlemleri tespit etmek yanlış olacaktır. Çünkü tek bir değişken bakımından uç olan bir gözlem diğer değişkenler bakımından uç olmayabilir. Öyleyse çok değişkenli tesadüfi örneğin herhangi bir gerçekleşişinin aykırı gözlem içermediği araştırılırken uç gözlemlerin belirlenmesine olanak verecek ortak bir sıralama ölçüsünün tanımlanması gerekmektedir. Buna göre kullanılacak sıralama ölçüsünün tanımına bağlı olarak uç gözlem teşhisleri de değişecektir.

Regresyon analizi yapmaktaki amacımıza, analizin çok deęişkenli oluşuna, gözlemlerin ilişkili olup olmamalarına ve bunlar gibi pek çok sebebe baęlı olarak aykırı gözlem tanımlamalarımız ve yöntemlerimiz deęişmektedir.

Dięer yandan aykırı gözlemlerle ilgili tanımlama ve yöntemlerin deęişiklik göstermesindeki başlıca sebeplerden biri de istatistikçinin baęlı olduęu istatistik okuludur. Yani tesadüfi örneğin herhangi bir gerçekleşişinde aykırı gözlemin yer alıp almadığını analiz etmekte veya olası aykırı gözlemlerin etkisini azaltmakta kullanılacak yöntemler ve aykırı gözleme ilişkin tanımlar Klasik (sampling theory) yaklaşımın ya da Bayesgil (Bayesian)yaklaşımın benimsenmesine göre de deęişiklik gösterecektir.

Çok deęişkenli-çoklu doğrusal regresyon (Multivariate Multiple Linear Regression) analizinde gözlemlerin aykırı olup olmadıklarını, dięerleri ile uyuşup uyuşmadıklarını test etmekte kullanılacak Bayesgil bir yöntemin tanımlamasını ve literatürde yer alan dięer yöntemlerle karşılaştırıldığında bu yöntemin hangi grupta yer aldığını daha rahat anlayabilmek açısından, yöntemlerin belli bir sistematik içinde kısaca özetlenmesi uygun olacaktır.

Literatürde aykırı gözlemlerle ilgili yer alan tüm farklı yöntemler, aykırı gözlemin sebep olduęu problemleri gidermek adına benimsenen görüşe uygun olarak kabaca iki grup altında toplanabilirler. Benimsenen bu görüşlerden ilki, aykırı gözlemleri uygun hale getirmek yani onları örnekte yer alan dięer gözlemlerle benzeştirmektir, ikincisi de, gözlem ya da gözlemlerin örnekte yer alan dięer gözlemlerle uyuşup uyuşmadığını test etmektir.

2. AYKIRI GÖZLEMLERLE İLGİLİ İSTATİSTİKSEL YÖNTEMLER

Aykırı gözlemin sebep olduğu problemi gidermek adına benimsenen görüşleri kısaca, “aykırı gözlemleri uygun hale getirme” ve “uyuşmazlık testleri” olarak isimlendirebiliriz. Aykırı gözlemleri uygun hale getirmek görüşü altında, tesadüfi örneğin herhangi bir gerçekleşişinde analizin sonuçları üzerinde negatif etki yaratabilecek gözlem ya da gözlemlerin var olabileceği olasılığını göz önünde bulundurarak bu tür gözlemlerden en az şekilde etkilenecek yöntemin belirlenmesi amacı yatmaktadır. Uyuşmazlık testlerinde ise tesadüfi örneğin herhangi bir gerçekleşişinde aykırı gözlemin yer alıp almadığının test edilmesi ve testin sonucuna göre de gözlemin yok sayılması(reddedilmesi) görüşü yatmaktadır. Bu iki genel görüş altında incelenecek yöntemlerde, klasik ya da bayesgil yaklaşımın benimsenmesine, ilgilenilen istatistiksel analizin karmaşıklığına ve hangi gözlemin aykırı sayılacağına ilişkin olarak yapılacak tanımın kapsamına göre çeşitlilik gösterecektir.

2.1. Aykırı Gözlemleri Uygun Hale Getirme Görüşü

Tesadüfi örneğin herhangi bir gerçekleşişinden yığın hakkında çıkarsama yapmak ve bu çıkarsamayı yaparken aykırı gözlemin varlığından çok fazla etkilenecek istatistiksel yöntemler, aykırı gözlemleri uygun hale getiren yöntemler olarak isimlendirilmektedir. Bu yöntemler aykırı gözlemlerin varlığına karşı sağlamdır(robust) diyebiliriz.

Sağlamlık istatistiksel çıkarsama yapmakta çok önemli bir kavramdır. Ancak sadece aykırı gözlem çalışmalarına özgü bir anlam taşımamaktadır. Son yıllarda, veriyi ürettiği düşünülen yapı(önerilmiş bir istatistiğin kullanılabilmesi için sağlanması gereken varsayımların tümü) hakkında çeşitli belirsizliklerin olması durumunda bu belirsizliklerden etkilenecek bir ölçü sağlayan istatistiksel yöntemler elde etmek için önemli bir çaba sarf edilmektedir. Veriyi ürettiği düşünülen yapının bir elemanı olan öngörülen dağılımın, parametrelerinin tahmini ve hipotezlerinin testi gibi problemlerle ilgilenirken, bu yapı içinde yer alan diğer elemanlardaki

belirsizliklerden etkilenmeyecek şekilde önerilen yöntemler sağlam olarak isimlendirilmektedir. Gözlemlerden bir veya birkaçının öngörülen model dışında bir başka modelden kaynaklandığı düşünülüyor ise, bu belirsizlikten dolayı sağlam yöntemler aykırı gözlem problemi içerisinde yer alırlar. Yani sağlamlık ve aykırı gözlem çalışmaları biri olmazsa diğeri olmaz türden kavramlar değildir. Öngörülen modelin dışında başka bir modelin varlığının olabileceğinin getirdiği belirsizliğin dışında, veriyi ürettiği düşünülen yapı içerisinde başka belirsizliklerde yer alabilir ve bunların çözümlerine de sağlam yaklaşımlarla gidilebilir.

Tesadüfi örneğin herhangi bir gerçekleşmesi içerisinde, aykırı gözlemle karşılaşıldığında veya böyle bir beklentinin olması durumunda, bu gözlemlerin etkilerinden korunmak için sağlam istatistiksel yöntemler araştırmak istenebilir. Örneğin, tesadüfi örneğin herhangi bir gerçekleşmesi içerisinde yer alan uç bir gözlemin bu gerçekleşiş üzerinden hesaplanacak varyansa ilişkin tahmin değerini büyük ölçüde etkilemesi kaçınılmazdır. Eğer tesadüfi değişken için öngörülen dağılımın parametre(lerinin) tahmini ile ilgileniyor ve bu tahminlerin aykırı gözlemlerden etkilenmesini istemiyorsak, aykırı gözlemlere karşı çok duyarlı olmayan tahmin edicileri kullanmayı isteyebiliriz. Basit bir örnek olarak, tesadüfi değişken için ön görülen dağılımın merkezi parametresini tahmin etmekte aykırı gözlemlerden çok etkilenmeyecek bir tahmin edici olan ortanca(medyan) istatistiğinin kullanılması verilebilir.

Tesadüfi değişken için öngörülen dağılım gerçeği yansıtmakta mıdır? Sorusu açısından aykırı gözlem problemine bakıldığında, sağlam istatistiksel yöntemlerin, uyumsuzluk testi yöntemlerine göre daha uygun tahminler vereceği düşünülmektedir. Tesadüfi değişken için öngörülen model üzerinden bir gözleme aykırıdır denebilir. Ancak gerçekte tesadüfi değişkenin sahip olduğu model bizim öngördüğümüz modelden farklı olabilir. Örneğin normal dağılım öngörmüşken, tesadüfi değişken gerçekte daha yüksek kuyukları olan bir dağılıma sahip olabilir (student t dağılımı gibi). Bu taktirde aykırı gözlemlerden korunmak için uç gözlemleri yok saymak, atmak, onları reddetmek dağılımın parametrelerini tahmin etmekte çok fazla bilgi kaybına sebep olacak bir tercih olur. Oysa ki uç gözlemlerin etkisini azaltacak

sağlam yöntemlerin kullanılması sonucu elde edilen tahminler, gözlemlerin çıkarılması sonucu elde edilen tahminlerden daha çok gerçeği yansıtacaktır.

Aykırı gözlemlerin varlığına rağmen parametre tahmini ve hipotez testleri gibi öncelikli konularla ilgilenebilmek aykırı gözlem analizlerindeki sağlam yaklaşımların temelini teşkil etmektedir.

Bayesgil yaklaşımla, aykırı gözlemlerin varlığının etkisinin uygun hale getirilmesi, önerilen önsel(prior) dağılım ile çok ilişkilidir. Önerilen önsel dağılım fonksiyonunun sonsal(posterior) dağılım fonksiyonu üzerinde olabilirlik(likelihood) fonksiyonundan daha baskın(daha çok bilgi veriyor) olması, başka bir deyişle güçlü bir önsel dağılım fonksiyonunun kullanılması uç gözlemlerin dolayısı ile aykırı gözlemlerin istatistiksel çıkarsamalar üzerindeki muhtemel negatif etkilerini azaltacaktır. Bayesgil yaklaşımla aykırı gözlemlerin uygun hale getirilmesi aşağıda verilecek örnekle daha kolay anlaşılabilir.

Konum parametresi θ ve ölçüm parametresi σ^2 olan X tesadüfi değişkenine ilişkin herhangi bir gerçekleşişin x olduğunu ve θ 'nın da konum parametresi θ_0 ve ölçüm parametresi σ_0^2 olan önsel dağılıma sahip olduğunu varsayalım. $X=x$ değerinin çok büyük olduğunu düşünelim. x değerinin çok büyük olması, θ 'ya ilişkin sonsal dağılımın beklenen değerini nasıl etkileyecektir? X 'in büyük değerli bir gerçekleşişi, önsel dağılımın konum parametresi θ_0 'ın, θ 'nın sonsal dağılımının beklenen değeri üzerindeki etkisini azaltacak ve onu kendisine doğru çekecektir. Yani sonsal dağılımın beklenen değeri, x 'e hızla yaklaşacaktır. Diğer yandan, eğer X tesadüfi değişkeninin dağılımı ve θ 'ya ilişkin önsel dağılım normal ise, θ 'ya ilişkin sonsal dağılımın beklenen değeri,

$$E(\theta/x) = x - (x - \theta_0)\sigma^2 / (\sigma^2 + \sigma_0^2)$$

'in x arttıkça x 'e yaklaşma hızı yavaşlayacaktır. Lindley(29) θ 'ya ilişkin önsel dağılımın normal değil de, student t dağılımına sahip olması halinde sonsal dağılımın beklenen değerinin x değeri arttıkça x değerine daha hızlı yaklaştığını göstermiştir.

Dawid(11) önsel dağılımın sonsal dağılım üzerinde az baskın yada çok baskın olmasına göre, x değerinin çok büyük olması durumunda sonsal dağılımın beklenen değerinin sırasıyla,

$$E(\theta / x) \rightarrow x, \quad x \rightarrow \infty \text{ için, (önsel dağılım az baskın)}$$

$$E(\theta / x) \rightarrow \theta_0, \quad x \rightarrow \infty \text{ için, (önsel dağılım çok baskın)}$$

olduğunu belirtmektedir. Sonsal dağılımın beklenen değerindeki, büyük x değerine bağlı bu istenmeyen değişme önsel dağılımın türü ile son derece ilgili görülmektedir.

O'Hagan(30) dağılımları, aykırı gözleme olan duyarlılıkları bakımından iki sınıfa ayırmaktadır. Bunlar, aykırı gözleme-yatkın(outlier-prone) ve aykırı gözleme-dirençli(outlier-resistant) dağılımlardır. Aykırı gözleme-yatkın ve dirençli dağılımlar, konum parametresinin sonsal dağılımı üzerinden tanımlanmışlardır.

Buna göre, x_n çok büyük bir değer iken, tüm $c, x_1, x_2, \dots, x_{n-1}$ ve θ 'ya ilişkin herhangi bir önsel dağılım için, sonsal olasılık,

$$p(\theta \leq c / x_1, x_2, \dots, x_n) \rightarrow p(\theta \leq c / x_1, x_2, \dots, x_{n-1})$$

oluyorsa, sonsal dağılımın aykırı gözleme-yatkın olduğu söylenmektedir. O'Hagan aykırı gözleme-yatkın dağılımlar içinde en az kuyruk yüksekliğine sahip dağılım olarak student t dağılımını göstermektedir.

Eğer, x_n çok büyük bir değer iken

$$p(\theta \leq c / x_1, x_2, \dots, x_n)$$

sonsal olasılığı, tüm $c, x_1, x_2, \dots, x_n$ ve θ 'ya ilişkin herhangi bir önsel dağılım için, x_n 'den etkilenen parametre tahminlerinin azalan bir fonksiyonu ise dağılım aykırı gözleme-dirençli olarak tanımlanmaktadır. Normal dağılım aykırı gözleme-dirençli bir dağılım olarak ifade edilmiştir.

O'Hagan'ın bu tanımlamaları yapmaktaki amacı, θ parametresine ilişkin bir tahmin elde etmekte, aykırı gözlemin varlığının irdelenip irdelenmeyeceğinin dağılımın aykırı gözleme-dirençli ya da yatkın olmasına bağlı olarak değişeceğidir. Yani tesadüfi değişkene ilişkin öngörülen dağılım aykırı gözleme-yatkın bir dağılım ise, parametrelere ilişkin tahminler aykırı sayılabilecek aşırı uç gözlemlerden etkilenmeyeceklerdir. Tesadüfi değişken için aykırı gözleme-dirençli bir dağılım öngörülmüş ise, aşırı uç gözlemlerin olması durumunda, bu gözlemlerden parametre tahminlerinin etkileneceği açıktır. Amaç, aynı aykırı gözleme-dirençli dağılıma sahip bağımsız tesadüfi değişkenlerden oluşan tesadüfi örneğin herhangi bir gerçekleşişi üzerinden, aykırı gözlemin olması riskine rağmen bayesgil yaklaşımla parametreye ilişkin bir tahmin elde etmek ise, elde edilen tahminin aykırı gözlemlerden etkilenme olasılığını en aza indirmek için aykırı gözleme-dirençli bir dağılımı önsel fonksiyon olarak kullanmak mantıklı olacaktır.

Uygun hale getirme görüşünden farklı olarak bayesgil yaklaşımda amaç, tesadüfi örneğin herhangi bir gerçekleşişinde aykırı olabilecek gözlemleri belirlemek ve bunları yok varsaymak(reddetmek)ise, bu durumda gözlemlerin geldikleri varsayılan (öngörülen dağılım) dağılıma uyduklarının test edilmesinde aykırı gözleme-yatkın bir dağılım önsel dağılım olarak belirlenmelidir. Hatta sonsal dağılım üzerinde en az etkili olacak önsel fonksiyonun kullanılması daha doğru olacaktır. Çünkü merak edilen gözlenmiş değer, tesadüfi değişken için öngörülen dağılıma uygun olup

olmadığıdır. Bunun cevabı ise, sonsal fonksiyonun iki ayağından biri olan olabilirlik fonksiyonu içinde yatmaktadır.

Özetlenecek olursa bayesgil yaklaşımlarla aykırı gözlemleri uygun hale getirmeye ilişkin yöntemlerde aykırı gözleme-dirençli, uyumsuzluk testlerinde ise aykırı gözlemlere-yatkın (hatta bilgi-vermeyen(non-informative)) önsel fonksiyonların kullanılması uygun olacaktır.

2.2. Uyuşmazlık Testi Görüşü

Aykırı gözlemin sebep olduğu problemi gidermek adına benimsenen görüşlerden ikincisi “uyuşmazlık testi görüşü” olarak Bölüm 2 'de isimlendirilmişti. Bu görüş, tesadüfi örneğin herhangi bir gerçekleşişinde yer alabilecek aykırı gözlemin, test etmek yoluyla reddedilmesini ya da önemle üzerinde durulması gereken bir durum olduğunun tanımlanmasını içermektedir.

Tesadüfi örneğin herhangi bir gerçekleşiş $x_1, x_2, \dots, x_n$ küçükten büyüğe doğru sıralandığında en küçük $x_{(1)}$ ve en büyük $x_{(n)}$ değerlerinin birer uç gözlem olduğunu düşünelim. Eğer bu uç gözlemlerden herhangi biri (örneğin $x_{(n)}$) tesadüfi değişken için öngördüğümüz dağılım bakımından aşırı uç bir gözlem ise, bu gözlemin aykırı olduğu düşünülebilir. Yani aşırı uç olan bir gözlemin tesadüfi değişken için ön gördüğümüz dağılımdan gelmediğine kanaat getirebiliriz. Öyleyse bu tür bir gözlem öngördüğümüz dağılımın dışında başka bir dağılımdan gelmiştir.

Tesadüfi örneği oluşturan her bir tesadüfi değişkenin aynı dağılıma sahip olduğu varsayılın. $x_{(n)}$ gözleminin yalnızca uç bir gözlem değil aynı zaman da istatistiksel açıdan mantıklı bir sonuç olup olmadığının incelenmesi için uyumsuzluk testi olarak isimlendirilen istatistiksel bir test uygulansın. Yani $x_{(n)}$ 'in $X_{(n)}$ tesadüfi değişkeninin dağılımı bakımından mantıklı olmayacak derecede büyük bir gözlem olmadığının testi yapılıyor olsun. Eğer istatistiksel açıdan $x_{(n)}$ gözlemi mantıklı

olmayacak derecede büyük görülüyorsa, testin anlamlılık düzeyine göre $x_{(n)}$ uyuşmaz (üst aykırı) gözlem olarak isimlendirilir. Tabii ki bu aşama sadece üst aykırı gözlemler için sınırlı değildir. $x_{(1)}$ 'in uyuşmaz (alt aykırı) bir gözlem olup olmadığı da incelenebilir.

$x_{(n)}$ gözleminin belirli bir anlamlılık düzeyinde uyuşmaz olduğunun gösterilmesi, bu gözlemin $X_{(n)}$ tesadüfi değişkeni için ön görülen model tarafından üretildiği düşüncesinin mantıksız olduğu anlamını taşımaktadır. Sadece öngörülen model üzerinden belirli bir anlamlılık düzeyinde herhangi bir uç gözlemin uyuşmaz olduğu kanaatine varmak, bu uç gözlemin aykırı olduğu anlamına gelmektedir. Bu aşamada söylenen tek şey, gözlemin tesadüfi değişken için öngörülen model tarafından üretilip üretilmediğidir. Yani, ön görülen modele uymadığı düşünülen gözlem için ön görülen modele alternatif oluşturacak başka bir modelden bahsedilmemektedir.

Eğer tesadüfi örneğin herhangi bir gerçekleşişindeki gözlemlerden birinin ön görülen modele alternatif olacak başka bir model tarafından üretildiği ve diğerlerinin öngörülen model tarafından üretildiği düşüncesi ile bir gözlemin uyuşmazlığı araştırmak istenirse, alternatif modelin varlığından dolayı, bu amaca uygun bir test istatistiği ile çalışmak mümkün olur. Bu testin sonucuna göre de gözleme aykırıdır ya da değildir denir.

Görüldüğü gibi, uyuşmazlık testi görüşü kendi içinde iki farklı görüşü barındırmaktadır. Bu görüşlerden ilki, bir gözleme aykırıdır diyebilmek için sadece öngörülen model üzerinden tanımlama (belirli bir anlamlılık düzeyine karşılık gelen kritik değerden daha büyük olan gözlenmiş değer in aykırı sayılması) yapmanın yeterli olduğudur. İkincisi de, aykırılığın şüphe edilen gözlemin üreticisi olarak öngörülen modele alternatif bir modelin tanımlanması ile elde edilebilecek bir test istatistiği üzerinden, gözlemin uyuşmazlığının test edilebileceği görüşüdür.

İkinci görüşte bir zorluk mevcuttur. Bu zorluk, aykırı olup olmadığı test edilmek istenen gözlem için nasıl bir alternatif modelin ortaya koyulacağıdır. Tesadüfi

değişken için öngörülen model tarafından üretilmediği düşünülen gözlemin hangi model tarafından üretilmiş olabileceğini bir test istatistiği üretebilmek için ortaya koymak gereklidir. Bu da ilgilenilen problemin özellikleri ve bizim sezgilerimizle çok alakalıdır. Aykırı gözlem için öngörülmüş olan alternatif modelin türü(çeşidi) ortaya çıkacak test istatistiğinin özellikleri ve başka test istatistiklerine göre avantajları bakımından önemlidir. Aykırı gözlemin üreticisi olarak ne tür alternatif modellerin kullanıldığını, uyumsuzluk testi görüşlerinden ikincisini daha ayrıntılı irdelerken vermek uygun olacaktır.

2.2.1. Hipotez testi prosedürüne bağlı uyumsuzluk testi görüşü

Bu görüş daha önce de ifade edildiği üzere, bizim tesadüfi örnek için öngördüğümüz modele alternatif olarak, aykırı gözlemi de tanımladığı düşünülen bir başka modelin ortaya koyulmasını gerektirmektedir. Tesadüfi örneğe ilişkin olarak, aykırı gözlemi de tanımladığı düşünülen alternatif bir modelin ileri sürülebilmesi, öngörülen modelin doğruluğunu test etmemizi sağlayacaktır.

Bir gözlemin uyumsuzluğuna kanaat getirerek ona aykındır diyebilmek için, bize mantıklı gelen herhangi bir istatistiği kullanabiliriz. Bu istatistiğin dağılımının elde edilmesi ve belirli bir anlamlılık düzeyine karşılık gelen kritik değerin belirlenerek, istatistiğin örnekten hesaplanan değerinin bu kritik değerden büyük ya da küçük olmasına göre gözlemin aykırı sayılabilmesi, yalnızca tesadüfi örnek için öngörülen modelin varlığına dayalıdır. Fakat testin gücünü incelemek ya da istatistiksel bazı özellikleri sağlaması istenen testler (test istatistikleri) ortaya konmak isteniyorsa, bu durumda alternatif bir modelin tanımlanması gerekmektedir. Literatürde, aykırı gözlemi de tanımladığı düşünülen pek çok farklı alternatif model tanımlaması yer almaktadır. Aşağıda tesadüfi örnek için öngörülen modele ve bu alternatif modellerden bazılarına yer verilecektir.

2.2.1.1. Öngörülen model

Bir deneyin n kez tekrarlanması sonucunda elde edilen veri kümesinin nasıl bir yapı tarafından üretilmiş olabileceğine dair birtakım öngörülerimiz mevcuttur. Deneye ilişkin n tane sonuç elde edilmeden önce, cevabı merak edilen soruya dair bir tahmin ortaya koyabilmek adına, deneyin her bir tekrarının aynı şartlarda ve birbirinden bağımsız, birbirlerinin sonuçlarını etkilemeyecek şekilde yapılmasını hedefleriz. Böylece deneyin sonuçları arasında görülecek farklılığın tek kaynağı, ilgilenilen soruya cevap olacağını düşündüğümüz inceleme altında tutulan özellik bakımından meydana gelen tesadüfi değişimler olacaktır. Eğer ilgilenilen soruya cevap olacağını düşündüğümüz tesadüfi karakteristiğe sahip özelliğe ilişkin olarak deneyin sonucunda karşımıza çıkan değer, belli bir olasılık yapısına göre ortaya çıktığını öngörüyorsak, bu öngörünün deneylerin bağımsız ve aynı şartlarda yapılıyor olmasından dolayı, deneyin her bir tekrarında aynı kalmasını bekleriz. Öyleyse bir deneyin n kez tekrarı sonucunda karşımıza çıkacak veri kümesi, deneyin her bir tekrarında “aynı olasılık yapısına”(deneyin her bir tekrarının bağımsız ve aynı şartlarda yapılıyor olması nedeniyle) göre sonuç vereceğini öngördüğümüz tesadüfi karakteristiğe sahip özelliğin, mümkün sonuçlarından biri olacaktır. Deneyin sonucunda tesadüfi karakteristiğe sahip özelliğe ilişkin n tane değerin nasıl bir yapı tarafından üretilmiş olabileceğine dair ileri sürülen varsayımlar bütünü(deneylerin sonuçları arasındaki farklılığın sadece tesadüfi karakteristiğe sahip özellikten ileri gelmesini sağlayacak) öngörülen model olarak tanımlayabiliriz.

Öngörülen model tarafından üretilmiş olduğunu düşündüğümüz tesadüfi örneğe ilişkin herhangi bir gerçekleşişte, aykırı bir gözlemin var olup olmadığını irdelemek isteyebiliriz. Eğer aykırı gözlemi, öngörülen model içinde yer alan aynı olasılık yapısına ait olma öngörüsüyle zıtlığa düşen gözlem olarak tanımlıyorsak, tesadüfi örneğin mevcut gerçekleşişindeki değerlerin tümünün modelde öngörüldüğü gibi aynı olasılık yapısına uygun olarak gerçekleştiklerini test etmek gerekir. Tesadüfi örneğin herhangi bir gerçekleşişinin aykırı gözlem içermediği anlamına gelen ve öngörülen model içerisinde yer alan aynı olasılık yapısına ait olma öngörüsü böyle bir test için gerekli yokluk hipotezini teşkil edecektir. Bu yokluk hipotezini,

$x_1, x_2, \dots, x_n$, $X_1, X_2, \dots, X_n$ tesadüfi örneğinin herhangi bir gerçekleşişi ve F de bu tesadüfi örnekte yer alan karşılıklı bağımsız tesadüfi değişkenlerin her biri için öngörülen herhangi bir dağılım fonksiyonu olmak üzere,

$$H_0 ; x_i \in F \quad ; \quad i=1,2,\dots,n$$

şeklinde de tanımlayabiliriz.

Aykırı gözlemi, farklı bir dağılımdan gelen gözlemdir şeklinde tanımlayan uyumsuzluk testi görüşüne göre, böyle bir gözlemin olmadığı anlamına gelen yokluk hipotezini test etmek için kullanılacak herhangi bir test istatistiği, gözlemin nasıl bir dağılımdan geldiğine dair önerilen alternatif hipotezin yapısına bağlı olacaktır.

2.2.1.2. Alternatif modeller

2.2.1.2.1. Aykırılığın şüphe edilen gözlemi yadırgamayan (inherent) alternatif model (A.Ş.E.G.Y)

Yokluk hipotezinde belirtilen, tesadüfi örneğin herhangi bir gerçekleşişindeki gözlemlerin aynı dağılımdan geldikleriydi. Tesadüfi değişken için özel bir dağılım adı kullanılmadı. Genel olarak dağılımların hepsi üzerinden konuşuldu. Yalnızca aynı dağılım tarafından üretilip üretilmediğine göre aykırı gözlemin tanımlanabileceğinden bahsedildi. Oysa ki gözden kaçırılmaması gereken bir başka konu, uygulamada tesadüfi değişken için özel bir dağılımın öngörüleceğidir. Tesadüfi örneğin herhangi bir gerçekleşişi öngördüğümüz bu özel dağılım için sürpriz sayılabilecek bir gözlemi içeriyor olabilir. Yani öngördüğümüz bu özel dağılım tesadüfi değişken için doğru bir tercih midir? Tesadüfi değişken için başka bir dağılım öngörülseydi, daha önceden üzerinde durulan gözleme yine sürpriz olarak bakılabilecek miydi? Örneğin, tesadüfi örneğin herhangi bir gerçekleşişinde yer alan gözlemlerin her birinin normal dağılımdan geldiğini öngörmüşken, bu gözlemlerden bir veya birkaçının aykırılığın şüphe ediyor olabiliriz. Belki de tesadüfi örneğe

ilişkin bu gerçekleşişte yer alan gözlemlerin her biri kuyrukları daha yüksek olan başka bir dağılımdan gelmiş olabilir. Öyleyse bu aşamada cevap verilmesi gereken soru, tesadüfi örneğe ilişkin herhangi bir gerçekleşişteki gözlemlerin tamamının tesadüfi deęişken için öngörülen dağılım tarafından mı, yoksa öngörülen dağılım bakımından aykırı gibi görünen gözlemleri de tanımlayacağı düşünölen başka bir dağılım tarafından mı üretildiğidir. Bu soru, yokluk hipotezi

$$H_0 ; x_i \in F \quad ; i=1,2,\dots,n \text{ (tesadüfi örneğin herhangi bir}$$

gerçekleşişinde yer alan gözlemlerin
tamamı F dağılımından gelmektedir)

ve alternatif hipotez

$$H_1 ; x_i \in G \quad ; i=1,2,\dots,n \text{ (tesadüfi örneğin herhangi bir}$$

gerçekleşişinde yer alan gözlemlerin
tamamı G dağılımından gelmektedir)

aracılığıyla da ifade edilebilir.

F ve G birbirinden tamamen farklı dağılımlar olabileceğı gibi, herhangi bir dağılım ailesinin farklı parametre deęerlerine sahip birer üyesi de olabilirler. Burada ve daha sonra tanımlanacak dięer alternatif modellerin hepsinde dikkat edilmesi gereken husus, G dağılımının F 'den farklı olarak aykırı gözlemi üretebilme kabiliyetine sahip olması zorunluluğudur. Yani F dağılımı altında aykırı olduğı düşünölen gözlemin G dağılımı altında böyle bir düşönceyi meydana getirmemesi gerekmektedir. Örneğın; G, F dağılımının merkezi parametresinin sağ ya da sol tarafa kaydırılması(shift) ile tanımlanmış bir dağılım olabilir. G dağılımı için dięer tanımlar, G 'nin F 'den daha büyük yayılma (dispersion) ya da belli bir yönde F 'den daha yüksek kuyruğa(çarpık bir dağılım) sahip olması şeklinde verilebilir. Dikkat edilirse G 'nin aykırı gözlemi de tanımlayan bir dağılım olabilmesi için, merkezi parametresinin F 'den farklı ve(veya) yayılım(ölçüm) parametresinin F 'den daha büyük olması zorunluluğı mevcuttur.

Shapiro ve Wilk(36,37) normal ve üstel dağılımlar tarafından üretildiği düşünülen veri kümeleri için farklı Aykırılığın Şüphe Edilen Gözlemi Yadırgamayan alternatif hipotezler tanımlamışlardır.

2.2.1.2.2. Karma(mixture) alternatif model

Karma model, öngörülen ve Aykırılığın Şüphe Edilen Gözlemi Yadırgamayan(A.Ş.E.G.Y.) modellerden farklı olarak, tesadüfi örneğin herhangi bir gerçekleşişinin iki farklı dağılımdan gelen gözlemler tarafından oluştuğunu ifade etmektedir. Yalnız, bu gözlemlerden büyük çoğunluğunun öngörülen modele uygun olarak gerçekleştiği, geriye kalan çok az sayıda gözlemin ise başka bir dağılım tarafından üretildiği varsayılmaktadır. Bu şekilde bir varsayıma, sadece aynı tür birimlerin bulunduğu bir yığının tanımlanamamasından(farklı türleri dışarıda bırakmamızı sağlayacak bir çerçeve oluşturamamamızdan) dolayı ihtiyaç duyulmaktadır. Yani yığın, iki farklı türe ait birimlerden oluşmaktadır. İlgilenilen türdeki birimler üzerinden istatistiksel bir çıkarsamayı bu karma yığın üzerinden yapabilmek, tesadüfi örneğin herhangi bir gerçekleşişinden elde edilecek bilginin ne kadarının ilgilenilen tür tarafından sağlandığı ile alakalıdır. Örneğin, tek bir cinse ait olduğunu varsaydığımız bir fosil yığınının, tesadüfi örneğe ilişkin gerçekleşişlerden herhangi biri ile ilgilendiğimizi düşünelim. Tesadüfi örneğe ilişkin bu mevcut gerçekleşişte, varsayılanın aksine farklı tesadüfi karakteristiğe sahip farklı cinsten bir ya da iki fosilin yer alması muhtemel olabilir. İki farklı cinsi de içinde bulundurmasının muhtemel olduğu bir yığın üzerinden ilgilenilen cinse ait bir çıkarsama yapmakla ilgileniliyor ise, ilgilenilmeyen cinse ait gözlemleri belirlemek adına, tesadüfi örneğe ilişkin mevcut gerçekleşişteki birimlerin üreticisi varsayılan ve yokluk hipotezinde ortaya koyulan yapıya alternatif olarak,

$$H_1 ; x_i \in (1-\lambda)F + \lambda G , (i=1,2,...,n)$$

karma hipotezi önerilebilir. Burada λ , F 'den çok farklı olan G dağılımı tarafından üretilmiş ve böyle olmasından dolayı da aykırı diye nitelenecek olan gözlemin, G

dağılımı tarafından üretilmiş olması ihtimalidir. G dağılımına ilişkin olası formların hangi kriterlere bağlı olarak tanımlanması gerektiğine bir önceki alternatif modelde yer verilmişti.

Box ve Tiao(7), Guttman(18) aykırı gözlem çalışmalarında karma modelin kullanılabileceğini tartışmaktadırlar. Box ve Tiao(7) tarafından, eğer tesadüfi örneğin mevcut gerçekleşişinde yer alan gözlem sayısı en fazla 20 ve λ 'da küçük ise karma modelin aykırı gözlem çalışmalarında kullanılabileceği vurgulanmaktadır. Ayrıca, λ 'nın en fazla 0.10 olarak alınabileceği, mümkünse 0.05 'in de altında seçilmesi önerilmektedir. Aksi taktirde, tesadüfi örneğin mevcut gerçekleşişinde, az olmasını ümit ettiğimiz aykırı gözlem sayısından daha fazla gözlemin G dağılımı altında ortaya çıktığı gibi durumlarla karşılaşabiliriz. Guttman(18) tesadüfi örneğe ilişkin gerçekleşişlerin herhangi birinde en fazla bir tane aykırı gözlem olabileceğini varsaymaktadır. Böyle bir varsayımda bulunmak yani tesadüfi örneğe ilişkin gerçekleşişlerden herhangi birinin içerebileceği aykırı gözlem sayısını kesin bir rakama bağlamak, bir kısır döngü oluşturmaktadır. Çünkü tesadüfi örneğe ilişkin gerçekleşişlerden herhangi bir tanesinin içerebileceğine dair öngördüğümüz aykırı gözlem sayısına uygun olarak karma alternatif hipotez anlam kazanacaktır. Eğer aykırı gözlem sayısına ilişkin bir rakam belirlemişken(λ 'ya bağlı olarak), belirlediğimizden daha fazla sayıda aykırı gözlemin tesadüfi örneğe ilişkin mevcut gerçekleşişte yer alması , karma alternatif modeli anlamsız kılacaktır.

Karma modellerin sadece bazı formları aykırı gözlemlerle ilişkilidir. Aşağıda daha ayrıntılı olarak verilecek olan ölçek-değiştirme(scale-shift)(alt ve üst aykırı gözlemleri belirlemek için kullanılır) ve merkez-değiştirme(location-shift)(değiştirmenin hangi yönde olduğuna bağlı olarak alt ve üst aykırı gözlemleri belirlemekte kullanılır) modelleri karma alternatif modelde yer alan G dağılımı için en çok tercih edilen dağılımı tanımlamaktadırlar.

2.2.1.2.3. Kaydırmalı (Slippage) alternatif model

Bundan önce tanımlanmış olan alternatif modeller üzerinden elde edilecek test istatistiklerinin sonuçlarına göre, eğer öngörülen model (yokluk hipotezi) reddediliyorsa, söylenecek tek şey tesadüfi örneğin mevcut gerçekleşişinin aykırı gözlem içerdiği. Hangi gözlemin aykırı olduğuna dair bir çıkarsamanın doğrudan yapılması mümkün olmamaktadır. Ancak tesadüfi örneğin mevcut gerçekleşişinin aykırı gözlem içerdiği kanısına varılıyor ise, aykırı gözlemin tesadüfi örneğin mevcut gerçekleşişinde yer alan uç gözlem (en büyük ya da en küçük) olması gerektiği düşüncesi ile bir karar verilebilmektedir. $(1-\lambda)F + \lambda G$ karma modeli ile aykırı gözlem belirleme amaçlı çalışmalarda, G dağılımı için kullanılabilecek formlardan bazılarının merkezi-değiştirilmiş ya da ölçeği-değiştirilmiş modeller olduğu söylenmişti. Merkezi-değiştirilmiş veya Ölçeği-değiştirilmiş modeller karma modelde G dağılımı olarak kullanılabileceği gibi, tek başlarına bu modeller öngörülen modele alternatif olarak kullanılabilirler. Bu modellerin alternatif olarak kullanılması ile yalnızca tesadüfi örneğin mevcut gerçekleşişinin aykırı gözlem içermediğini test etmekle kalmaz, aynı zamanda uç gözleminde aykırı olmadığını doğrudan test etmiş oluruz. Genel olarak, Kaydırmalı alternatif model olarak isimlendirilen bu alternatif modellerin en sık kullanılan formları, tesadüfi örneğin herhangi bir gerçekleşişindeki gözlemlerden pek çoğunun bağımsız ve parametreleri μ ve σ^2 olan F dağılımından, geri kalan gözlemlerin bağımsız ancak F 'nin parametreleri kaydırılmış bir versiyonundan geldiği şeklindedir. Yani,

$$H_0 ; x_i \in F \text{ (tüm } i\text{'ler için) öngörülen model, (yokluk hipotezi)}$$

olmak üzere, kaydırmalı alternatif model

$$H_1 ; x_1, x_2, \dots, x_{i-1}, x_{i+1}, \dots, x_n \in F$$

$$x_i \in G$$

şeklinde tanımlanmaktadır. Burada G, F dağılımının merkezi yada ölçek parametresi değiştirilmiş bir versiyonudur. G 'nin aykırı gözlemi tanımlayabilmesi açısından F 'nin merkezi parametresi her iki yönde de değiştirilebilecekken, ölçek parametresi sadece büyük olacak şekilde değiştirilebilir.

Tesadüfi örneğin herhangi bir gerçekleşişindeki aykırı gözlemleri belirlemekte, aykırı gözlem ya da gözlemleri tanımladığı düşünülen H_1 kaydırmalı alternatif modelinin kullanılması sonucu En çok Olabilirlik Oranından elde edilecek test istatistiğinin yeterince büyük değeri bize yalnızca tesadüfi örneğin herhangi bir gerçekleşişinin aykırı gözlem içerdiğini gösterecektir. Ancak hangisinin yada hangilerinin olduğunu göstermeyecektir. Bu testin sonucunda (karma modelde olduğu gibi) öngörülen model reddediliyorsa, uç gözlemler aykırı olarak değerlendirileceklerdir. Oysaki H_1 'in alternatif model olarak kullanılması sonucu elde edilen test istatistiği uç gözlemin aykırı olduğunu söylemeye yeterli değildir. Bu ancak bir yorum olur. Etiketlenmiş-kaydırma (Labelled Slippage) olarak isimlendirilen,

$$H_2 : x_{(1)}, x_{(2)}, \dots, x_{(n-1)} \in F$$

$$x_{(n)} \in G \quad (x_{(i)}, i.\text{nci sıradaki gözlem } (i=1,2,\dots,n))$$

alternatif hipotezin kullanılması sonucu elde edilecek test istatistiği üzerinden, tesadüfi örneğin herhangi bir gerçekleşişinin aykırı gözlem içermediğini, içeriyorsa da bunun bir uç gözlem olacağını test etmiş oluruz.

Tesadüfi örneğin herhangi bir gerçekleşişinin iki veya daha fazla aykırı gözlem içermediği test edilirken kullanılan etiketlenmiş-kaydırma alternatif modelleri

$$\bar{H} : x_{(i)}, x_{(j)} \in G, (\text{diğer } (n-2) \text{ sıralı gözlem } F\text{'nin elemanı})$$

yada

$$\bar{H}; x_{(i)} \in G_1, x_{(j)} \in G_2, (\text{diğer } (n-2)\text{ sıralı gözlem } F\text{'nin elemanı})$$

şeklinde tanımlanabilmektedir.

Karma alternatif modelin kullanılması sonucu elde edilen test istatistiği üzerinden, tesadüfi örneğin herhangi bir gerçekleşişinin yalnızca aykırı gözlem içerip içermediği söylenebilirken, etiketlenmiş-kaydırmalı alternatif modellerin kullanılması sonucu elde edilen test istatistiği üzerinden tesadüfi örneğin herhangi bir gerçekleşişindeki en büyük ya da en küçük gözlemin (uç gözlemlerin) aykırı olup olmadığı söylenebilmektedir. Yani, yalnızca $x_{(n)}$ gözleminin aykırı olmadığını, başka bir deyişle $x_{(n)}$ gözleminin merkezi ya da ölçeği-kaydırılmış bir dağılımdan gelmediğini test etmiş oluyoruz. Oysa ki esas test edilmek istenen, tesadüfi örneğin herhangi bir gerçekleşişi içinde yer alan herhangi bir gözlemin öngörülendekinden farklı bir dağılımdan gelmediğidir.

Tesadüfi örneğin herhangi bir gerçekleşişinde yer alan gözlemlerden her birinin öngörülen model dışında başka bir model tarafından üretilmediğini test etmekte kullanılabilecek bir alternatif model Değişebilir (exchangable) alternatif modeldir.

2.2.1.2.4. Değişebilir (exchangable) alternatif model

Tesadüfi örneğin herhangi bir gerçekleşişinin tek bir aykırı gözlem içerdiği düşünülüyor olsun. Bu alternatif modelde, $x_1, x_2, \dots, x_{i-1}, x_{i+1}, \dots, x_n$ gözlemlerinin bağımsız olarak F dağılımı tarafından, x_i gözleminin de G dağılımı tarafından üretildikleri varsayılmaktadır. Ayrıca, n tane gözlemden herhangi birinin (i.nci) G dağılımı tarafından üretilmiş olması ihtimalinin eşit olduğu varsayılmaktadır. Bu modelde, $X_1, X_2, \dots, X_n$ tesadüfi değişkenlerinin bağımsız olmadıkları ancak değişebilir oldukları düşünülmektedir (31)

Tesadüfi örneğin herhangi bir gerçekleşişinde, G dağılımı tarafından üretilmiş tek bir gözlemin olduğu düşünülüyor ise, değişebilir alternatif model altında olabilirlik fonksiyonu,

$$\ell(x_1, x_2, \dots, x_n / F, G) = \frac{1}{n} \sum_{i=1}^n g(x_i) \prod_{j=1}^n f(x_j)$$

olarak ifade edilmektedir. Burada $f(x)$, $g(x)$ sırasıyla F ve G dağılımlarının olasılık (yoğunluk) fonksiyonlarıdır. F ve G farklı dağılım ailelerinden olabilecekleri gibi aynı ailenin farklı parametrelere sahip üyeleri de olabilirler.

2.2.2. Sezgisel uyumsuzluk testi görüşü

Hipotez Testi Prosedürüne Bağlı Uyuşmazlık Testi Görüşünde, öngörülen ve alternatif modelin yapılarına uygun olarak Klasik ya da Bayesgil hipotez testi prosedürlerinin uygulanması sonucunda ortaya çıkan test istatistikleri üzerinden gözlemlerin aykırı olmadıklarının test edilmek istendiği vurgulanmaya çalışılmıştır. Belli bir hipotez testi prosedürüne bağlı kalarak elde edilmiş test istatistiği sahip olduğu özelliklerden dolayı tercih edilebilir. Ancak belirli bir hipotez testi prosedürüne bağlı kalmaksızın sadece sezgisel olarak önerilmiş bir takım istatistiklerin öngörülen ya da alternatif model altındaki dağılımlarından faydalanarak gözlemlerin aykırı olup olmadıklarının irdelenmesi ile ilgili literatürde pek çok çalışma yer almaktadır.

Barnett ve Lewis((3) sayfa94) sezgisel olarak önerilmiş bu test istatistiklerini yedi farklı sınıfa ayırmıştır. Bu sınıflamada yer alan istatistiklerden kimilerinin öngörülene alternatif bazı modeller altında En Çok Olabilirlik Oran Testinden elde edilebildiği de yine aynı kaynakta yer almaktadır.

Bayesgil yaklaşımla aykırı gözlemlerin belirlenmesine yönelik öne sürülmüş yöntemler Barnett ve Lewis(3), Pena ve Guttman(21), Justel ve Pena(27) tarafından aynı biçimde sınıflandırılmışlardır. Bu sınıflamalarda yer alan bayesgil aykırı gözlem

belirleme yöntemleri (bayesgil uyumsuzluk testi yöntemleri) belirli bir hipotez testi prosedürüne dayanmadıkları için bu bölümde ele alınacaklardır. Sadece öngörülen ya da alternatif modelin doğruluğu koşulu altındaki uyumsuzluk testi yöntemleri olarak da isimlendirebileceğimiz bu bölümde yer alan bayesgil yöntemler, iki guruba sınıflandırılmışlardır. Bu guruplardan ilki, tesadüfi örneğin herhangi bir gerçekleşişinin üreticisi olarak bölüm (2.2.1) 'de tanımlanmış olan öngörülen modelin benimsendiği yöntemlerden, ikinci gurup ise tesadüfi örneğin herhangi bir gerçekleşişinin üreticisi olarak bölüm (2.2.1.2)'de tanımlanan alternatif modellerden herhangi birinin benimsendiği yöntemlerden oluşmaktadır.

Bu iki grupta yer alan bayesgil uyumsuzluk testi yöntemlerine, çoklu doğrusal regresyon analizi üzerinden kısaca değinilecektir.

2.2.2.1. Öngörülen modelin doğruluğu altında , bayesgil uyumsuzluk testi yöntemleri

Çoklu doğrusal regresyon modeli ,

$$y = X\theta + \varepsilon$$

y , $(n \times 1)$ boyutlu normal dağılıma sahip tesadüfi değişkenler vektörü , X $(n \times p)$ boyutlu tam kolon ranklı $(p < n)$ bilinen bir matris , θ $(p \times 1)$ boyutlu parametreler vektörü ve ε_i $(n \times 1)$ boyutlu ortalaması sıfır , kovaryans matrisi $\sigma^2 I_n$ olan normal dağılıma sahip tesadüfi değişkenler vektörü , şeklinde tanımlandığında “öngörülen model” olarak isimlendirilmektedir.

Öngörülen modelin doğruluğu koşulu altında , aykırı gözlemleri tanımlamak için kullanılan Bayesgil yöntemler aşağıda kısaca özetlenmektedir. Çalışmalarda θ ve σ^2 parametrelerine ilişkin bileşik bilgi-vermeyen(non-informative) önsel(prior) fonksiyon, $p(\theta, \sigma^2) \propto (\sigma^2)^{-1}$ olarak kullanılmaktadır.

2.2.2.1.1. Kestirici dağılım (Predictive Distribution)

Box (6) , Geisser (13 , 14 , 15 , 16 , 17) , Pettit & Smith (34) ve Pena & Tiao (33) aykırı gözlemlerin belirlenmesine yönelik yöntemler geliştirmekte kestirici dağılımın kullanılmasını önermektedirler. Buradaki temel fikir kestirici yoğunluk,

$$p(y_1 / y_{(1)}) = \int p(y_1 / \theta, \sigma^2) p(\theta, \sigma^2 / y_{(1)}) d\theta d\sigma^2$$

nın hesaplanmasıdır. Burada Y_1 , k sayıda gözlemden oluşan bir vektör ve $Y_{(1)}$ 'da k sayıda gözlemin dışında kalan , üzerinde θ parametresine ilişkin sonsalın oluşturulduğu bağımlı değişkene ait gözlenmiş değerlerdir.

Öngörülen modelin doğruluğu koşulu altında, bilgi-vermeyen önsel için Geisser(12) kestirici yoğunluğu(predictive density) ,

$$K = \frac{\Gamma((n-p)/2)}{\Gamma(\frac{1}{2})^k \Gamma((n-p-k)/2)(n-p-k)^{k/2}}$$

ve

$$Q_1 = \frac{(y_1 - X_1 \hat{\theta}_{(1)})'(I - H_1)(y_1 - X_1 \hat{\theta}_{(1)})}{(n-p-k)s_{(1)}^2}$$

olmak üzere,

$$p(y_1 / y_{(1)}) = K(s_{(1)}^2)^{-k/2} |I - H_1|^{1/2} (1 - Q_1)^{-(n-p)/2}$$

olarak ifade etmektedir. Burada $\hat{\theta}_{(1)}$ ve $s_{(1)}^2$, $(X_{(1)}, y_{(1)})$ üzerinden θ ve σ^2 'nin tahmin edicileri ve H_1 'da H 'nin elemanlarından oluşan $(k \times k)$ 'lık izdüşüm

matrisidir. $H_1 = X_1(X_1'X_1)^{-1}X_1'$ şeklinde ifade edilmekte olup, y_1 gözlemlerinin kaldıraç-gücü (leverage) ölçüsü olarak da isimlendirilmektedir. Dikkat edilirse, herhangi bir gözlem ya da gözlem grubu için kestirici ordinatların küçük çıkması kaldıraç-gücü ölçüsünün büyük çıkması ile aynı anlamı taşımaktadır. Öyleyse tesadüfi örneğin herhangi bir gerçekleşişinde, aykırı gözlem ya da gözlemlerin yer almadığını incelemekte kestirici ordinatlar fikir verici olabilirler.

2.2.2.1.2. Gözlenemeyen bir takım düzensizliklerin sonsal

Chaloner ve Brant (10) tarafından önerilmiş olan bu yöntem ,

$$p_i = \text{pr}(|\varepsilon_i| > q\sigma / y, X)$$

sonsals olasılığına dayandırılmaktadır. Burada ε_i 'ler $\varepsilon_i = y_i - x_i'\beta$ ile ifade edilen gerçekleşmiş ancak gözlenemeyen hatalardır. Chaloner ve Brant bu çalışmalarında uygun q 'nun seçimini ve p_i sonsal olasılığı büyük olan y_i gözleminin aykırılığını tartışmaktadırlar.

2.2.2.1.3. Tesadüfi örneğin herhangi bir gerçekleşişinden bir gözlem kümesi çıkartıldığında, gelecekteki bir gözlem için kestirici yoğunlukta meydana gelen değişimin ölçüsü

İlk olarak Johnston ve Geisser (25) tarafından ortaya atılan bu yaklaşımı kısaca aşağıdaki gibi özetleyebiliriz.

Eğer amaç, $Z = W\theta + \varepsilon^*$, ($\varepsilon^* \sim N(0, \sigma^2 I)$ ve ε dan bağımsız ; W , kestirim alınmak istenen bağımsız değişken vektörü ya da matrisi) modelinden geleceğe ilişkin bir gözleme dair kestirim yapmak ise, y koşulu altında Z 'nin kestirici yoğunluğu,

$$p(z / W, X, y) = \int p(z / W, \theta, \sigma^2) p(\theta, \sigma^2 / X, y) d\theta d\sigma^2$$

ve $y_{(1)}$ koşulu altında Z 'nin kestirici yoğunluğu,

$$p(z / W, X_{(1)}, y_{(1)}) = \int p(z / W, \theta, \sigma^2) p(\theta, \sigma^2 / X_{(1)}, y_{(1)}) d\theta d\sigma^2$$

arasında oluşacak büyük farklılık y_1 gözlemlerinin aykırılığının bir ölçüsü olarak kullanılabilir. Bu kestirici yoğunluklar arasındaki farklılığın ölçüsünü Johnston ve Geisser(25) Kestirici Etki Fonksiyonu(Predictive Influence Function(PIF)) adı altında tanımlamaktadır.

Johnston ve Geisser(26) ve Guttman ve Pena(20) tesadüfi örneğin herhangi bir gerçekleşişinden bir gözlem seti çıkarıldığında sonsalda meydana gelen değişimle ilgili aykırı gözlem tanımlama yöntemi geliştirmişlerdir. Son olarak Guttman ve Pena(21) bu değişikliklerin ortaya çıkışının , yokluk modeli dışında başka bir modelden gelen bir grup gözlemin olasılıkları ile ilişkili olduğunu göstermişlerdir.

2.2.2.2. Alternatif modelin doğruluğu koşulu altında , bayesgil uyumsuzluk testi yöntemleri

Öngörülen modelden farklı olarak alternatif bir modellerin kullanılmasındaki amaç , aykırı olabilecek gözlem veya gözlemlerin diğerlerinden farklı bir dağılımdan geliyor olmasının sonsal olasılığını elde etmektir. Bu sonsal olasılığın büyük olması o gözlem veya gözlemlerin aykırı olabileceği anlamını taşımaktadır. Kullanılan alternatif modellerden bazıları bölüm(2.2.1.2)'de de tanımlanmış olan , Varyansı-Değiştirilmiş(Variance-Inflation yada Variance-Shift), Karma ve Ortalaması-Değiştirilmiş(Mean-Shift) modellerdir.

Varyansı-Değiştirilmiş Alternatif Model

Box ve Tiao(7) tarafından önerilen varyansı-değiştirilmiş alternatif model ,

$$y = X\theta + \varepsilon$$

çoklu doğrusal regresyon modelinde , hatalardan her birinin bağımsız olarak F dağılımından($f(\varepsilon/\xi_1)$) yada G dağılımından($g(\varepsilon/\xi_2)$) geldiği varsayımı ile tanımlanmıştır.

$a_{(k)}$, n tane ε 'dan k tanesinin $g(\varepsilon/\xi_2)$ dağılımından ve $s=n-k$ tanesinin de $f(\varepsilon/\xi_1)$ dağılımından gelmesi olayı olarak tanımlanırsa, y verilmişken $(\theta, \xi_1, \xi_2, a_{(k)})$ 'nın olabilirlik fonksiyonu

$$\begin{aligned} \ell(\theta, \xi_1, \xi_2, a_{(k)}) &= \dot{f}(\varepsilon_{(s)}/\xi_1) \dot{g}(\varepsilon_{(k)}/\xi_2) \\ &= \dot{f}(y_{(s)} - X_{(s)}\theta/\xi_1) \dot{g}(y_{(k)} - X_{(k)}\theta/\xi_2) \end{aligned}$$

şeklinde yazılmaktadır. Burada $\dot{f}$ ve $\dot{g}$ sırasıyla $\varepsilon_{(s)}$ ve $\varepsilon_{(k)}$ 'nin elemanlarının yoğunluk fonksiyonlarının çarpımıdır. Dolayısı ile, tam olabilirlik fonksiyonu ε 'lardan kaç tanesinin $g(\varepsilon/\xi_2)$ dağılımından geldiğine bağlı olarak oluşacak 2^n tane mümkün olabilirlik fonksiyonunu içerecektir.

$p^{(k)}$, $a_{(k)}$ olayının ortaya çıkmasının önsel olasılığı olarak düşünülürse

$$p^{(k)} \geq 0 \quad , \quad \sum_k p^{(k)} = 1$$

dir ve $p(\theta, \xi_1, \xi_2)$ 'de parametrelerin önsel dağılımı olmak üzere, y verilmişken $(\theta, \xi_1, \xi_2, a_{(k)})$ 'nın sonsal dağılımı Box ve Tiao tarafından ifade edilmiştir. Burada

esas üzerinde durulan, 2^n tane $p(\theta/a_{(k)}, y)$ marjinal koşullu sonsal dağılım üzerinden θ 'nın ağırlıklandırılmış marjinal sonsal dağılımı,

$$p(\theta/y) = \sum_k p(a_{(k)}/y)p(\theta/a_{(k)}, y)$$

nın elde edilmek istenmesidir. $p(a_{(k)}/y)$ ağırlıkları tesadüfi örneğin herhangi bir gerçekleşişinde k tane hatanın $g(\varepsilon/\xi_2)$ dağılımından gelmesinin sonsal olasılığıdır. Yukarıdaki toplamı açık şekli ile yazmak gerekirse, $p(a_{(k)}/y) = w_{(k)}$ şeklinde ifade edilmek üzere,

$$p(\theta/y) = w_{(0)}p(\theta/a_{(0)}, y) + \sum_{i=1}^n w_{(i)}p(\theta/a_{(i)}, y) + \\ + \sum_{i < j} w_{(ij)}p(\theta/a_{(ij)}, y) + \sum_{i < j < t} w_{(ijt)}p(\theta/a_{(ijt)}, y) + \dots$$

2^n tane toplam üzerinden ifade edilir. $w_{(i)}$ ağırlığı sadece i .nci gözlemin, $w_{(ij)}$ ağırlığı sadece i .nci ve j .nci gözlemlerin,... $g(\varepsilon/\xi_2)$ dağılımından gelmesinin sonsal olasılıklarıdır.

Herhangi bir hatanın $f(\varepsilon/\xi_1) = N(0; \sigma^2)$ ve $g(\varepsilon/\xi_2) = N(0; \delta^2 \sigma^2)$ dağılımlarından birinden gelmesinin önsel olasılıkları sırasıyla $(1-\alpha)$ ve α , θ ve $\log \sigma^2$ 'de bölgesel bağımsız(locally independent) tekdüze(uniform) önsel olmak üzere $(p^{(k)} = \alpha^k (1-\alpha)^{n-k} ; p(\theta, \xi_1, \xi_2) = p(\theta, \sigma^2) = \frac{1}{\sigma^2})$

$$w_{(k)} = p(a_{(k)}/y) = C \left(\frac{\alpha}{1-\alpha} \right)^k \delta^{-k} \frac{|X'X|^{1/2}}{|X'X - \phi X'_{(k)} X_{(k)}|^{1/2}} \left\{ \frac{s_{(k)}^2}{s^2} \right\}^{-1/2(n-p)}$$

olarak elde edilmiştir. Burada $\phi = 1 - \delta^{-2}$ ve C 'de k tane gözlemin aykırı olmasının sonsal olasılığı olan $w_{(k)}$ 'ların toplamının bir olmasını sağlayacak sabittir. Büyük $w_{(k)}$ değeri k sayıda gözlemin aykırı olduğunu nitelemektedir.

Karma Alternatif Model

Pena ve Tiao(33) ve Box ve Tiao(7) hataların dağılımı için

$$\varepsilon_i \sim (1 - \alpha)N(0; \sigma^2) + \alpha N(0; \delta^2 \sigma^2) \quad , i=1,2,\dots,n$$

karma alternatif (karma dağılım için varyansı-değiştirilmiş model kullanılmıştır) modelinin benimsenmesi durumunda $p(\theta / \delta, \alpha, y)$ sonsal olasılığının varyansı-değiştirilmiş model altında elde edilen $p(\theta / y)$ sonsal olasılığı ile aynı olduğunu göstermişlerdir.

Ortalaması-Kaydırılmış Model

$\{1,2,\dots,n\}$ kümesinden k tane farklı tamsayının seçilmesi ile oluşan küme, I ile ifade edilsin. Bu durumda y vektörü $y' = (y'_1, y'_{(I)})$ şeklinde ayrıştırılabilir. Burada (I) notasyonu I 'nın tümleyenini ifade etmektedir. Benzer şekilde X matrisi , $X' = (X'_1, X'_{(I)})$ şeklinde parçalanabilir. Bu notasyonlara bağlı olarak , $y' = (y'_1, y'_{(I)})$ gözlemlerine ilişkin Ortalaması-Kaydırılmış model ,

$$y_1 = X_1 \theta + a + \varepsilon_1 \quad , \quad y_{(I)} = X_{(I)} \theta + \varepsilon_{(I)}$$

olarak tanımlanmaktadır. Burada a , (kx1) boyutlu ortalaması-kaydırılmış sabitler vektörüdür. $\varepsilon_1 \sim N(0, \sigma^2 I_k)$, $\varepsilon_{(I)} \sim N(0, \sigma^2 I_{n-k})$ 'dan bağımsızdır. Bu model Guttman(18) tarafından kullanılmış ve sonra Guttman,Dutter ve Freeman(19) tarafından faydalanılmıştır. Guttman, Dutter ve Freeman k 'yı sabit tutmuşlar ve bu

durumda oluşabilecek $\binom{n}{k}$ kombinasyonu kadar farklı I kümesi üzerinden, y verilmişken a, θ ve σ^2 'ye ilişkin olabilirlik fonksiyonunu,

$$\ell(a, \theta, \sigma^2 / y) = \sum_{\binom{n}{k}} \binom{n}{k}^{-1} (\sigma^2)^{-(n-k)/2} \exp\left\{-\frac{1}{2\sigma^2} (y_{(I)} - X_{(I)}\theta)'(y_{(I)} - X_{(I)}\theta)\right\} \cdot \left[(\sigma^2)^{-k/2} \exp\left\{-\frac{1}{2\sigma^2} (y_I - X_I\theta - a)'(y_I - X_I\theta - a)\right\} \right]$$

olarak ifade etmişlerdir. Burada k sayıda gözlemden oluşan herhangi bir I kümesinde yer alan gözlemlerin ortalaması-kaydırılmış model tarafından üretilmiş

olmasının önsel olasılığı $\frac{1}{\binom{n}{k}}$ ve (a, θ, σ^2) parametrelerine ilişkin önsel olasılıkta

$\frac{1}{\sigma^2}$ varsayılmaktadır. Olabilirlik fonksiyonu ve önsel fonksiyonlardan hareketle tanımlanan $p(a, \theta, \sigma^2 / y)$ sonsal fonksiyonundan θ 'ya ilişkin marjinal sonsal dağılım,

$$p(\theta / y) = \sum_{\binom{n}{k}} w_{(k)} p(\theta / I, y)$$

olarak ifade edilmiştir. Buradaki $\binom{n}{k}$ sayıdaki $w_{(k)}$ ağırlıkları, k sayıda gözlemin ortalaması-kaydırılmış model tarafından üretilmiş olmasının sonsal olasılıklarıdır ve

$$w_{(k)} = \frac{\left\{ s_{(I)}^{-(n-k)} \left[s_{(I)}^p | (X'_{(I)} X_{(I)})^{-1} | \right]^{1/2} \right\}}{\sum \binom{n}{k} \left\{ s_{(I)}^{-(n-k)} \left[s_{(I)}^p | (X'_{(I)} X_{(I)})^{-1} | \right]^{1/2} \right\}},$$

$$s_{(I)} = (y_{(I)} - X_{(I)} (X'_{(I)} X_{(I)})^{-1} X'_{(I)} y_{(I)})' (y_{(I)} - X_{(I)} (X'_{(I)} X_{(I)})^{-1} X_{(I)} y_{(I)})$$

şeklinde ifade edilmiştir. Bu yaklaşımda da büyük $w_{(k)}$ sonsal olasılığı I kümesinde yer alan k sayıda gözlemin aykırı olduğu anlamına gelmektedir.



3. ÇOK DEĞİŞKENLİ VERİDE AYKIRI GÖZLEM

İlgilenilen istatistiksel analiz ne denli karmaşıklaşırsa aykırı gözlemi ve onu ortaya çıkaracak yöntemleri belirlemenin zorlaşacağına giriş bölümünde kısaca değinilmişti. Çok değişkenli-çoklu (Multivariate-multiple) doğrusal regresyon analizinde aykırı gözlemleri belirlemek üzere önerilmiş bayesgil bir yöntemle ilgili bu tez çalışmasında, çoklu doğrusal regresyon analizindeki aykırı gözlemleri belirlemekten daha zor bir durum ile karşılaşılmaktadır. Bu ek zorluk bağımlı değişken sayısının birden fazla olmasıdır. Başka bir deyişle çok değişkenli bir analiz tek değişkenli bir analize nazaran, aykırı gözlemi belirlemekte üstesinden gelinmesi gereken bir problemi daha beraberinde getirmektedir. Bu problem bir gözlemin diğerlerinden farklılığını ortaya koymanın matematiksel karşılığı olduğunu düşündüğümüz uzaklık kavramının, çok değişkenli analizde birden fazla tesadüfi değişkenin olması sebebi ile sekteye uğramasıdır. Tek değişkenli analizde sıralanmış gözlemlerden en uçta yer alanlar kuşku uyandırmakta ve en uçta yer alan bu gözlemlerin aykırılığının istatistiksel açıdan mantıklı olup olmadığı, uygun bir uzaklık ölçüsünün ön görülen yada alternatif model altındaki dağılımı üzerinden değerlendirilmektedir. Tek bir tesadüfi değişkene ilişkin tanım cümlesi üzerinde gözlemlerin büyüklüklerine göre sıralanması mümkünken, birden fazla tesadüfi değişkene ilişkin tanım cümlesi üzerinde gözlemlerin büyüklüklerine göre sıralanması mümkün değildir. Örneğin bir grup insanın boylarını ölçüp gerçek sayı doğrusu üzerinde işaretlersek gözlemlerin küçükten büyüğe doğru sıralanışlarından bahsetmek mümkün olacaktır. Ancak hem boy hem de kiloları ölçüldüğünde bu iki tesadüfi değişkenin tanım cümlesi olan düzlem üzerinde yer alan gözlemlerin büyüklüklerine göre bir sıralanıştan bahsetmek mümkün değildir. Düzlem üzerinde hangi gözlemin daha büyük yada küçük olduğunu yani uç gözlem olduğunu bilmemizi sağlayacak, düzlemde yer alan tüm noktalar için tanımlı doğal bir sıralamanın mümkün olmayışlığı, düzlemde yer alan gözlemlerin belli bir ölçü bakımından aynılık gösterenlerinden oluşan alt-gruplara ayrılmasını ve bu alt-gruplar üzerinden bir sıralamanın yapılabileceği fikrini güçlü kılmaktadır. Çoklu düzlemde belli bir ölçü bakımından aynılık gösteren gözlemlerden oluşan alt-grupların kullanılan ölçü üzerinden sıralanması alt-sıralama (sub-ordering) olarak isimlendirilmektedir. Bir ölçü üzerinden aynı olduğu düşünülen

gözlemlerden meydana gelen alt-gruplardan yararlanarak sıralama yapmak fikrinin sahip olduğu dezavantajlar, kullanılacak tek ve değişmez bir ölçünün olmayışlığı ve aynı grupta yer alan gözlemlerin tek bir değerle ifade ediliyor olmasıdır. Aynı grupta yer alan gözlemlerin tek tek irdelenmesi mümkün olmamaktadır. Bu dezavantajlara rağmen çok değişkenli veride uç gözlemleri ifade edebilmek açısından gerekli alt-sıralamayı sağlayacak herhangi bir ölçünün kullanılması kaçınılmaz olmaktadır. Barnett(4) alt-sıralama ölçülerini özelliklerine göre dört farklı bölüme ayırmaktadır. Bu bölümler marjinal, indirgenmiş (reduced), kısmi (partial) ve koşullu (conditional) alt-sıralama olarak adlandırılmaktadırlar. İndirgenmiş-alt-sıralama ölçüsü neredeyse kullanılan tek ölçü olarak karşımıza çıkmaktadır.

İndirgenmiş-alt-sıralama ile (p boyutlu) çok değişkenli herhangi bir x gözlemi tek bir rakam olan $R(x)$ miktarına dönüştürülmüş olur. Böylece tesadüfi örneğin herhangi bir gerçekleşmesi olan $x_1, x_2, \dots, x_n$, $R_i = R(x_i)$, ($i=1,2,\dots,n$) değerleri bakımından sıralanabilir. En büyük R_i değerine karşılık gelen x_i gözlemi aykırı olmaya aday(kuşkulu) gözlem olarak değerlendirilir. Tesadüfi örneğin herhangi bir gerçekleşişinde yer alan en büyük gözlemin aykırı olmadığı test edilmek isteniyorsa, öngörülen model üzerinden $R_{(n)}$ 'in dağılımı ile ilişkili olarak en büyük gözleme karşılık gelen $R_{(n)}$ değerine bağlı, $x_{(n)}$ 'in uyumsuzluğuna karar verilebilir.

İndirgeme ölçüsü $R(x)$ 'in seçimi nasıl yapılacaktır? Sezgisel (kişisel) bir seçim yapmanın tehlikesi, çok değişkenli verinin yapısının ilgisiz(alakasız) olabilecek bir indirgeme ölçüsü üzerinden ortaya konulamayabileceğidir. $R(x)$ için uygun bir seçimin, öngörülen ve alternatif modeller üzerinden herhangi bir hipotez testi prosedürüne bağlı kalarak ortaya çıkmasını isteriz. Ancak bu uygun seçim nadiren ortaya çıkmaktadır. Hatta hipotez testi prosedürüne bağlı kalarak bir $R(x)$ elde edilmiş olsa da, kimi dağılımlar için $x_{(n)}$ uç gözleminin uyumsuzluğunu test etmekte kullanılacak $R_{(n)}(x)$ istatistiğinin öngörülen model altındaki dağılımının elde edilmesi çok kolay olmayabilir.

Örneğin çok boyutlu bir x gözlemini, sezgisel olarak tek boyutta

$$R(x; x_0, \Gamma) = (x - x_0)' \Gamma^{-1} (x - x_0) \quad [3.1]$$

uzaklık ölçüsü ile ifade etmeyi seçmiş olalım. Burada x_0 öngörülen dağılımın merkezi parametre vektörünü veya onun tahminini ve Γ 'da varyans kovaryans matrisini veya onun tahminini ifade etmektedir. Eğer tesadüfi değişken için öngörülen dağılım çok değişkenli normal dağılım ($N \sim (\theta, \Sigma)$) ise, [3.1]'de ki indirgeme ölçüsü bu dağılıma göre,

$$R(x; \theta, \Sigma) = (x - \theta)' \Sigma^{-1} (x - \theta)$$

formunda ifade edilecektir. Sezgisel olarak önerilen bu indirgeme ölçüsü aynı zamanda tesadüfi değişken için çok değişkenli normal dağılımın öngörüldüğü durum için, klasik yaklaşımda en çok olabilirlik oran testinden elde edilmiş bir test istatistiği olarak karşımıza çıkmaktadır. Diğer dağılımlar için $R(x; x_0, \Gamma)$ indirgeme ölçüsü uygun olmayabilir. Tesadüfi değişken için çok değişkenli normal dağılımın öngörülmesi durumunda, klasik yaklaşımda en çok olabilirlik oran testinden indirgeme ölçüsünün elde edilmesine geçmeden önce, çok değişkenli aykırı gözlemi belirlemekte Barnett(5) tarafından önerilmiş iki temel kriteri vermek mantıklı olacaktır.

A Kriteri; Tesadüfi örneğin herhangi bir gerçekleşmesi olan $x_1, x_2, \dots, x_n$ 'den çıkarılması durumunda, öngörülen dağılım altında geriye kalan gözlemler üzerinden en büyüklenmiş(maksimize edilmiş) olabilirlikle hiç gözlem çıkarılmadan en büyüklenmiş olabilirlik arasında en büyük farkı meydana getirecek olan gözlem, en uç gözlemdir. Eğer bu en büyük artış önemli derecede büyük ise bu uç gözlem aykırı olarak tanımlanır.

Bu kriter sadece öngörülen modelin tanımlanmasını gerektirir. Eğer aykırı gözlemi de tanımladığını düşündüğümüz alternatif bir modelin varlığı kabul ediliyorsa daha karmaşık bir kriterin tanımlanması gerekmektedir.

B Kriteri; Tesadüfi örneğin herhangi bir gerçekleşmesi olan $x_1, x_2, \dots, x_n$ 'in öngörülen ve alternatif modeller altındaki logaritmik olabilirlik fonksiyonları arasındaki farkı en büyük yapan gözlem en uç gözlemdir. Eğer bu en büyük fark önemli derecede büyük ise bu uç gözlem aykırı olarak tanımlanır.

A ve B kriterinin tesadüfi örneğin herhangi bir gerçekleşmesinde birden fazla aykırı gözlemin olması durumuna da genişletilebileceği söylenmektedir.

Öngörülen model $x_1, x_2, \dots, x_n$ 'in parametreleri θ ve Σ olan çok değişkenli normal dağılıma sahip bağımsız tesadüfi değişkenlerden oluşan $X_1, X_2, \dots, X_n$ tesadüfi örneğinin herhangi bir gerçekleşmesi olduğu şeklinde tanımlanıyor olsun. Alternatif model de, tesadüfi örneğin herhangi bir gerçekleşmesinde yer alan gözlemlerden sadece bir tanesinin öngörülenden farklı bir dağılım tarafından üretildiği şeklinde tanımlanıyor ise Barnett(5)'in ileri sürdüğü A kriterine göre, Σ 'nın biliniyor olması durumu için indirgenmiş-alt-sıralama ölçüsü Barnett ve Lewis(3) tarafından aşağıdaki gibi ifade edilmektedir.

Öngörülen model altında, tesadüfi örneğin herhangi bir gerçekleşmesi olan $x_1, x_2, \dots, x_n$ 'e ilişkin olabilirlik fonksiyonu,

$$\ell(x_1, x_2, \dots, x_n / \Sigma) \propto \frac{1}{|\Sigma|^{n/2}} \exp \left\{ -\frac{1}{2} \sum_{j=1}^n (x_j - \theta)' \Sigma^{-1} (x_j - \theta) \right\}$$

olmak üzere, en büyüklenmiş logaritmik olabilirlik fonksiyonu (sabit terim hariç)

$$L(x_1, x_2, \dots, x_n / \Sigma) \propto -\frac{1}{2} \sum_{j=1}^n (x_j - \bar{x})' \Sigma^{-1} (x_j - \bar{x})$$

şeklinde ifade edilmektedir. i.nci gözlemin çıkarılması durumunda, öngörülen model altında en büyüklenmiş logaritmik olabilirlik fonksiyonu ise,

$$L(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n / \Sigma) \propto -\frac{1}{2} \sum_{j \neq i} (x_j - \bar{\bar{x}})' \Sigma^{-1} (x_j - \bar{\bar{x}})$$

olarak verilmektedir. Bu ifadedeki $\bar{\bar{x}}$, x_i hariç diğer (n-1) gözlem üzerinden elde edilen ortalamadır. A kriterine göre bu iki en büyüklenmiş olabilirlik fonksiyonu arasındaki fark,

$$R_i(x_i, \bar{x}, \Sigma) = (x_i - \bar{x})' \Sigma^{-1} (x_i - \bar{x})$$

şeklinde elde edilmektedir.

Tesadüfi örneğin herhangi bir gerçekleşmesi olan $x_1, x_2, \dots, x_n$ de yer alan herhangi bir gözlem için elde edilen $R_i(x_i, \bar{x}, \Sigma)$ değerinin, $R_i(x_i, \bar{x}, \Sigma)$ istatistiğinin öngörülen model altındaki dağılımı üzerinden belli bir anlamlılık düzeyine karşılık gelecek şekilde tanımlanmış kritik değerden büyük olması, öngörülen modelin reddedilmesini gerektirir. Bu ise mevcut gözlemlerden herhangi birinin farklı bir dağılım tarafından üretildiği anlamını taşımaktadır. Bu istatistik üzerinden hangi gözlemin farklı dağılım tarafından üretildiğini test edemeyiz. Söyleyebileceğimiz tek şey, tesadüfi örneğin herhangi bir gerçekleşmesi olan $x_1, x_2, \dots, x_n$ de farklı dağılımdan gelen bir gözlemin mevcut olduğudur. Sezgisel olarak bu gözlemin en büyük $R_i(x_i, \bar{x}, \Sigma)$ değerini veren gözlem olduğunu söyleyebiliriz. Eğer uç gözlem $x_{(n)}$ 'in (n.sıra istatistiği) uyuşmazlığını (aykırılığını) test etmek istiyorsak,

$$R_{(n)}(x_{(n)}, \bar{x}, \Sigma) = (x_{(n)} - \bar{x})' \Sigma^{-1} (x_{(n)} - \bar{x})$$

istatistiğinin dağılımı üzerinden değerlendirme yapılması gerekmektedir.

4. AYKIRI GÖZLEMLERİN TEK TEK VARLIĞININ İRDELENMESİNE ÇOK DEĞİŞKENLİ ÇOKLU REGRESYON ANALİZİNDE BAYESGİL YAKLAŞIM

Chaloner ve Brant(10) tarafından çoklu doğrusal regresyon analizinde aykırı gözlemleri belirlemek için önerilmiş yöntem bölüm(2.2.2.1.2)'de kısaca değinilmişti. Bu yöntem göre , gözlemlerin hepsi öngörülen model altında gerçekleşmiştir. Ancak bazıları büyük tesadüfi hataya sahip olabilir. Buna göre bir gözlemi aykırı olarak niteleyebilmek için o gözleme ilişkin tesadüfi hatanın büyüklüğünün bir kritik değerle tanımlanması gerekmektedir.

Öngörülen modelde tanımlanmış olan, tesadüfi hatanın olasılık yoğunluk fonksiyonu üzerinden n çaplı tesadüfi örneğin herhangi bir gerçekleşişinde gözlemlerden hiçbirinin aykırı olmama olasılığı çok büyük olarak seçilirse , örneğin $0.95 = [pr(\epsilon_i < k), \text{ tüm } i \text{ 'ler için}]$, herhangi bir gözlemin aykırı olmama olasılığı $pr(\epsilon_i < k) = (0.95)^{1/n}$, aykırı olması olasılığı da $pr(\epsilon_i > k) = 1 - (0.95)^{1/n}$ olarak elde edilir. Bu olasılığı sağlayan $k = q\sigma$, kritik değer olarak tanımlanmaktadır. Chaloner ve Brant gerçekleşmiş hata (Realized Error (2)) 'nın sonsal fonksiyonu üzerinden elde edilen $pr((\epsilon_i / \Sigma, Y) > k)$ sonsal olasılığının , ϵ_i 'nin öngörülen modelde tanımlanan dağılımı üzerinden elde edilen $pr(\epsilon_i > k)$ olasılığından büyük olmasını, i nci gözlemin aykırı olduğunun göstergesi olarak tanımlamaktadır.

Varbanov(39) , Chaloner ve Brant(10)'in yaptığı çalışmayı çok değişkenli-çoklu doğrusal regresyon analizine genişletmiştir.

4.1. Çok Değişkenli-Çoklu Regresyon Analizinde Aykırı Gözlemlerin Uyumsuzluklarının Tek Tek İrdelendiği Bayesgil Bir Yaklaşım

Y_i , p değişkenli i nci bağımlı değişken ($i=1,2,...,n$) olmak üzere,

$$Y = \begin{bmatrix} Y_1^T \\ Y_2^T \\ \vdots \\ Y_n^T \end{bmatrix} = [Y^{(1)}, Y^{(2)}, \dots, Y^{(p)}] = \begin{bmatrix} Y_{11}, Y_{12}, \dots, Y_{1p} \\ Y_{21}, Y_{22}, \dots, Y_{2p} \\ \vdots \\ Y_{n1}, Y_{n2}, \dots, Y_{np} \end{bmatrix}$$

(nxp) boyutlu veri matrisini gösterebilir. Burada $Y^{(j)}$, ($j=1,2,\dots,p$) j nci değişkendeki n tane bağımsız gözlemden oluşan vektördür.

Çok değişkenli-çoklu doğrusal regresyon modeli, X , (nxq) boyutlu tasarım(dizayn) matrisi olmak üzere,

$$Y^{(j)} = X\theta_j + \varepsilon^{(j)} \quad ; j=1,2,\dots,p$$

şeklinde yazılabilir. Matris biçiminde bu model,

$$Y = X\Theta + E \quad , \quad \Theta = (\theta_1, \theta_2, \dots, \theta_p) \quad [4.1]$$

olarak ifade edilmektedir. (nxp) boyutlu E matrisinin i nci satırı, Y_i nin tesadüfi hatalarını içermektedir ve

$$E = \begin{bmatrix} \varepsilon_1 \Sigma^{1/2} \\ \varepsilon_2 \Sigma^{1/2} \\ \vdots \\ \varepsilon_n \Sigma^{1/2} \end{bmatrix}$$

formunda ifade edilir. Burada, ε_i 'lerin bağımsız , parametreleri 0 ve I olan p boyutlu, aynı normal dağılıma sahip tesadüfi değişkenler olduğu varsayılmaktadır. Bu durumda $Y=X\Theta+E$ modeli ,

$$Y_i = X_i \Theta + \varepsilon_i \Sigma^{1/2} \quad ; i=1,2,\dots,n$$

şeklinde de yazılabilir. $[\varepsilon_i = (\varepsilon_{1i}, \varepsilon_{2i}, \dots, \varepsilon_{pi}) = (Y_i - X_i \Theta) \Sigma^{-1/2} \sim N(0, I)]$ dır.

Bölüm 3'te ifade edilen sebeplerden dolayı, Varbanov(39) tarafından , i nci gözlemin aykırı olmadığının test edilmesi için ,

$$R_i = \varepsilon_i \varepsilon_i' = (Y_i - X_i \Theta) \Sigma^{-1} (Y_i - X_i \Theta)'$$

miktarının(ölçüsünün) sonsal dağılımı kullanılmaktadır. ε_i 'lerin bağımsız aynı dağılımlı , ortalaması 0 ve varyans kovaryans matrisi I olan p değişkenli normal dağılıma sahip olduğu varsayımı altında , R_i istatistiği , p serbestlik dereceli merkezi ki-kare dağılımına sahip olur.

R_i istatistiğinin öngörülen modeldeki varsayımlara uygun olarak elde edilen dağılımı (p serbestlik dereceli merkezi ki-kare dağılımı) üzerinden n çaplı tesadüfi örneğin herhangi bir gerçekleşişinde hiçbir gözlemin aykırı olmaması olasılığı büyük bir değer olarak seçilirse (örneğin 0.95) ,

$$\begin{aligned} 0.95 &= \text{pr}\{R_i \leq k \text{ , tüm } i \text{ 'ler için } \} \\ &= \text{pr}\{R_1 \leq k\} . \text{pr}\{R_2 \leq k\} . \dots . \text{pr}\{R_n \leq k\} = \{F_p(k)\} \end{aligned}$$

eşitliğinden, herhangi bir gözlemin aykırı olmamasına ilişkin $\text{pr}\{R_i \leq k\} = \{F_p(k)\}^{1/n}$ olasılığından k kritik değeri elde edilir.

Gerçekleşmiş ancak gözlenemeyen hatanın karesel formu olan R_i miktarının(ölçüsünün), k kritik değerinden büyük olmasının sonsal olasılığı $\text{pr}((R_i / \Sigma, Y) > k)$ 'nin , öngörülen modeldeki varsayımlara uygun olarak R_i (tesadüfi hata ε_i 'nin karesel formu) istatistiğinin dağılımı üzerinden elde edilen $\text{pr}(R_i > k)$ olasılığından daha büyük olması, i nci gözlemin aykırı olabileceğini ifade etmektedir.

Bir diğer yaklaşım,

$$H_{0i} : R_i \geq k , Y_i \text{ aykırı gözlem ise } (i=1,2,\dots,n)$$

$$H_{1i} : R_i < k$$

hipotezini test etmek için Bayes Faktörünün kullanılmasıdır. H_{0i} 'nin H_{1i} 'e karşı test edilmesi için kullanılan Bayes Faktörü,

$$B_i = \frac{\frac{p_i}{1-p_i}}{\frac{1-\{F_p(k)\}^{1/n}}{\{F_p(k)\}^{1/n}}} = \frac{p_i \{F_p(k)\}^{1/n}}{(1-p_i)(1-\{F_p(k)\}^{1/n})}$$

sonsal oran(posterior odd)'ların önsel oran(Prior odd)'lara oranıdır. Buradaki p_i sonsal olasılığı $\text{pr}((R_i / \Sigma, Y) > k)$ 'yı ifade etmektedir ve elde edilişi Bölüm 4.6'da verilmektedir. Kass ve Raftery(28), B_i 'nin 10 dan büyük olmasının H_{0i} 'nin doğruluğuna dair büyük bir kanıt , 100 den büyük olmasının ise çok güçlü bir kanıt olduğunu ifade etmektedirler.

4.2. Gerçekleşmiş Hatanın Karesel Formu R_i 'nin Sonsal Dağılımı

Gerçekleşmiş Hatanın karesel formu olan $R_i = \varepsilon_i \varepsilon_i'$ miktarının (ölçüsünün) sonsal dağılımının elde edilmesi için, $(\varepsilon_i / \Sigma, Y)$ 'nin sonsal dağılımının bulunması gereklidir. $\varepsilon_i = (Y_i - X_i \Theta) \Sigma^{-1/2}$, gerçekleşmiş hata Θ ve Σ 'nin doğrusal bir fonksiyonu olduğu için, Θ ve Σ 'nin bileşik sonsal dağılımı

$$p(\Theta, \Sigma / Y) = \ell(\Theta, \Sigma / Y) p(\Theta, \Sigma)$$

elde edilmelidir. Burada $\ell(\Theta, \Sigma / Y)$, Y verilmişken Θ ve Σ 'nin olabilirlik fonksiyonunu, $p(\Theta, \Sigma)$ ise bileşik önsel dağılımı ifade etmektedir.

Σ istenmeyen (nuisance) parametre olarak düşünülüyor ise, yani sadece Θ için çıkarsama yapmak önem taşıyorsa, Θ ve Σ 'ya ilişkin sonsal dağılım,

$$p(\Theta, \Sigma / Y) = p(\Theta / \Sigma, Y) p(\Sigma / Y)$$

Σ ve Y biliniyorken Θ 'nın sonsal dağılımı ile Y biliniyorken Σ 'nin marjinal sonsal dağılımının çarpımı olarak yazılır. Bölüm (4.3) 'de önsel dağılım ve Bölüm (4.5)'de olabilirlik fonksiyonu ile ilgili bilgiler verildikten sonra Bölüm(4.6)'te $p(\Theta / \Sigma, Y)$ üzerinden Gerçekleşmiş Hata $(\varepsilon_i / \Sigma, Y)$ 'nin ve Bölüm(4.7)'de de Gerçekleşmiş Hatanın karesel miktarının sonsal dağılımı ile ilgili bilgiler verilmektedir.

4.3. Önsel Dağılım

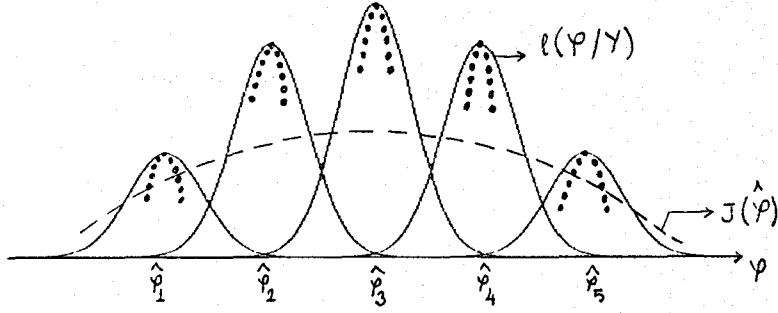
Önsel bilginin kişiden kişiye değişiklik gösterdiği veya mevcut olmadığı durumlarda, Sonsal dağılım üzerinde veriden elde edilen bilginin ağırlığının daha yüksek olması tercih edilir(35,38). Yani olabilirlik fonksiyonunun sonsal dağılım üzerinde daha etkili(baskın) olması tercih edilir. Önsel'in tek düze yada tek düzeye yakın bir aileden olması, olabilirlik fonksiyonunun sonsal üzerindeki etkisinin fazla olmasını sağlayacaktır. Bu düşünceyle önerilmiş olan Başvuru (Reference) önsel, deneyden

elde edilen gözlemlerin ve önsel'in bir fonksiyonu olan Fayda (utility) fonksiyonunu maksimum yapacak önsel fonksiyon olarak tanımlanmaktadır. Fayda fonksiyonunu maksimum yapan önsel fonksiyon da yaklaşık tek düze olarak karşımıza çıkmakta ve bu Başvuru önsel olarak isimlendirilmektedir(24). Jeffrey kuralı da parametreye ilişkin yaklaşık tek düze önsel fonksiyonun elde edilmesini sağlayacak metriği bulmamızı sağlar.

4.4. Jeffrey Kuralı

Aynı $p(y/\varphi)$ (φ dağılımın konum (location) veya ölçüm (scale) parametrelerinden herhangi birinin tesadüfi değişken, diğerinin de sabit olduğu durumu ifade etmektedir) dağılımına sahip ve bağımsız tesadüfi değişkenlerden oluşan $Y_1, Y_2, \dots, Y_n$ tesadüfi örneğinin konum veya ölçüm parametrelerine ilişkin olabilirlik fonksiyonunun verisi dönüştürülmüş (data translated) olmasını sağlayacak metrik, bu parametrelere ilişkin bölgesel tek düze (locally uniform) önsel fonksiyon olmaktadır. Bu sebepile olabilirlik fonksiyonunun verisi dönüştürülmüş olduğu metriği belirlemekle, bölgesel tek düze olan önsel fonksiyonu da elde etmiş oluruz.

$Y_1, Y_2, \dots, Y_n$ bağımsız ve aynı $p(y/\varphi)$ dağılımına sahip tesadüfi bir örnek olsun. Dağılım bazı düzenlilik koşullarına (regularity conditions) uyduğu takdirde, yeterince büyük n (tesadüfi örneğin herhangi bir gerçekleşişinin çapı) için φ 'nin ve bire-bir dönüşümlerinin olabilirlik fonksiyonları normale yakınsar (22,23).



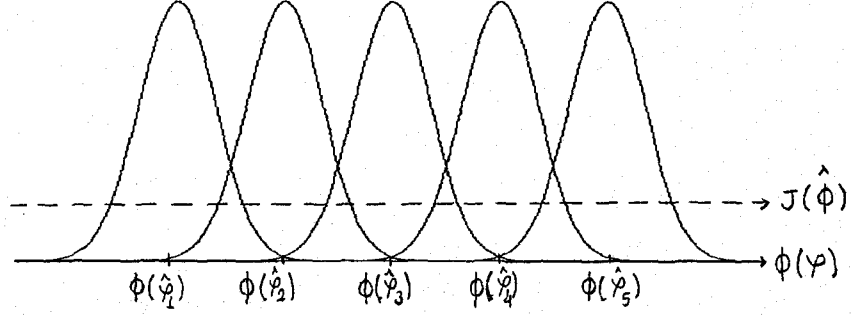
Şekil 4.1. φ metriğindeki olabilirlik fonksiyonları

Şekil 4.1’de noktalarla çizilmiş paraboller, olabilirlik fonksiyonunun φ ’nın en çok olabilirlik tahmin edicisi olan $\hat{\varphi}$ etrafındaki ikinci derece Taylor açılımını ifade etmektedir. Olabilirlik fonksiyonunun logaritmasının $\hat{\varphi}$ etrafındaki ikinci derece Taylor açılımı,

$$\begin{aligned} L(\varphi / Y) &= \log \ell(\varphi / Y) = \log \prod_{i=1}^n p(Y_i / \varphi) \\ &\cong L(\hat{\varphi} / Y) - \frac{n}{2} (\varphi - \hat{\varphi})^2 \left(-\frac{1}{n} \frac{\partial^2 L}{\partial \varphi^2} \right)_{\hat{\varphi}} \end{aligned}$$

dır. $J(\hat{\varphi}) = \left(-\frac{1}{n} \frac{\partial^2 L}{\partial \varphi^2} \right)_{\hat{\varphi}}$ ifadesi sadece $\hat{\varphi}$ ’nın bir fonksiyonudur. Şekil 4.1’de

kesikli çizgilerle çizilmiş eğri $J(\hat{\varphi})$ ’yı ifade etmektedir. Şekil 4.2 ‘de görüldüğü üzere, $\phi(\varphi)$ gibi φ ’nin bire-bir bir dönüşümünün sonucunda ortaya çıkan metrik üzerinde, olabilirlik fonksiyonunun verisi dönüştürülmüş olduğunu düşünelim .



Şekil 4.2. ϕ metriğindeki olabilirlik fonksiyonları

Yani $J(\phi(\varphi))$ 'nın bir sabit olduğunu düşünelim. Bu durumda,

$$J(\hat{\phi}) = \left(-\frac{1}{n} \frac{\partial^2 L}{\partial \phi^2} \right)_{\hat{\phi}} = \left(-\frac{1}{n} \frac{\partial^2 L}{\partial \phi^2} \right)_{\hat{\phi}} \left(\frac{d\phi}{d\hat{\phi}} \right)_{\hat{\phi}}^2 = J(\hat{\phi}) \left(\frac{d\phi}{d\hat{\phi}} \right)_{\hat{\phi}}^2$$

ifadesinde $\left| \frac{d\phi}{d\hat{\phi}} \right|_{\hat{\phi}} \propto J^{-1/2}(\hat{\phi})$ olursa, $J(\hat{\phi})$, $\hat{\phi}$ 'dan bağımsız bir sabit olarak

elde edilir. Bu ise olabilirlik fonksiyonunun ϕ metriğinde yaklaşık verisi

dönüştürülmüş olduğu anlamına gelir. $p(\varphi) = \left| \frac{d\phi}{d\varphi} \right| \propto J^{1/2}(\hat{\phi})$ 'de φ metriğinde

bilgi-vermeyen (non-informative) önsel dağılımı ifade etmektedir. Olabilirlik fonksiyonunun verisi dönüştürülmüş olduğu ϕ metriği,

$$\frac{d\phi}{d\varphi} \propto J^{1/2}(\varphi) \quad , \quad \phi \propto \int J^{1/2}(t) dt$$

olarak elde edilir.

Hata terimlerinin (E_i) bağımsız, aynı dağılımlı parametreleri sıfır ve Σ olan çok değişkenli normal dağılıma sahip olduğu varsayımı altındaki çok değişkenli-çoklu

doğrusal regresyon analizinde, konum ve ölçüm parametrelerine ilişkin önsel fonksiyonların bağımsızlığı göz önüne alındığında, Jeffrey kuralına göre merkezi parametresi için bilgi vermeyen önsel fonksiyon ,

$$p(\Theta) = c = \text{sabit} , \quad \Theta = (\theta_1, \theta_2, \dots, \theta_p)$$

ve ölçüm parametresi Σ 'nın $\frac{1}{2} p(p+1)$ elemanı içinde

$$p(\Sigma) \propto |F(\Sigma)|^{1/2}$$

dir. Burada $F(\Sigma)$ Fisher'in Bilgi (Fisher Information) ölçüsüdür. $J(\Sigma)$, örnek çapı büyüdükçe $F(\Sigma)$ 'e yakınsamaktadır ((8),sayfa41).

Olabilirlik fonksiyonunda Σ^{-1} 'e göre türevler daha rahat alındığından Σ için bilgi matrisinin Σ^{-1} üzerinden elde edilmesi ,

$$|F(\Sigma)| = |F(\Sigma^{-1})| \left| \frac{\partial \Sigma}{\partial \Sigma^{-1}} \right|^{-2}$$

dönüşümü ile mümkün olur. $\left| \frac{\partial \Sigma}{\partial \Sigma^{-1}} \right| = \left| \frac{\partial(\sigma_{11}, \sigma_{12}, \dots, \sigma_{pp})}{\partial(\sigma^{11}, \sigma^{12}, \dots, \sigma^{pp})} \right|$, Σ 'nın σ_{ij}

elemanlarından Σ^{-1} 'in σ^{ij} elemanlarına yapılan dönüşümün Jacobianıdır. Σ^{-1} 'e ilişkin Fisher'in bilgi ölçüsünün

$$|F(\Sigma^{-1})| \propto \left| \frac{\partial \Sigma}{\partial \Sigma^{-1}} \right| = |\Sigma|^{p+1}$$

olduğu bilinmektedir ((8),Appendix A.8.2). Böylece Σ parametresine ilişkin bilgi vermeyen önsel,

$$p(\Sigma) = |\Sigma|^{-\frac{1}{2}(p+1)}$$

şeklinde elde edilir.

4.5. Olabilirlik Fonksiyonu

Çok değişkenli-çoklu doğrusal regresyon analizinde parametreleri sıfır ve Σ olan aynı çok değişkenli normal dağılıma sahip, bağımsız n tane tesadüfi hata vektörünün bileşik olasılık yoğunluk fonksiyonu,

$$\begin{aligned} p(E / \Sigma, \Theta) &= \prod_{i=1}^n p(E_i / \Sigma, \Theta) \\ &= (2\pi)^{-\frac{np}{2}} |\Sigma|^{-\frac{n}{2}} \exp\left(-\frac{1}{2} \sum_{i=1}^n E_i \Sigma^{-1} E_i'\right), \quad -\infty < E_{im} < \infty \\ &\quad i = 1, 2, \dots, n \\ &\quad m = 1, 2, \dots, p \end{aligned}$$

olan çok değişkenli normal dağılımdır. Bu dağılım fonksiyonunun üstel kısmındaki karesel form,

$$\sum_{i=1}^n E_i \Sigma^{-1} E_i' = iZ \Sigma^{-1} S(\Theta) = \sum_{m=1}^p \sum_{j=1}^p \sigma^{mj} S_{mj}(\theta_m, \theta_j)$$

biçiminde yazılabilir. Burada,

$$S(\Theta) = \{S_{mj}(\theta_m, \theta_j)\}, \quad m, j = 1, 2, \dots, p$$

($p \times p$) boyutlu simetrik bir matristir. Bu matrisin (m, j) nci elemanı

$$S_{mj}(\theta_m, \theta_j) = (Y_m - X\theta_m)'(Y_j - X\theta_j)$$

$$\begin{aligned}
&= (Y_m - X\hat{\theta}_m + X\hat{\theta}_m - X\theta_m)'(Y_j - X\hat{\theta}_j + X\hat{\theta}_j - X\theta_j) \\
&= [(Y_m - X\hat{\theta}_m)' - (X\theta_m - X\hat{\theta}_m)'][(Y_j - X\hat{\theta}_j) - (X\theta_j - X\hat{\theta}_j)] \\
&= (Y_m - X\hat{\theta}_m)'(Y_j - X\hat{\theta}_j) - (Y_m - X\hat{\theta}_m)'(X\theta_j - X\hat{\theta}_j) - \\
&\quad - (X\theta_m - X\hat{\theta}_m)'(Y_j - X\hat{\theta}_j) + (X\theta_m - X\hat{\theta}_m)'(X\theta_j - X\hat{\theta}_j)
\end{aligned}$$

dır. Bu eşitlikte yer alan

$$\begin{aligned}
(Y_m - X\hat{\theta}_m)'(X\theta_j - X\hat{\theta}_j) &= Y_m'X\theta_j - Y_m'X\hat{\theta}_j - \hat{\theta}_m'X'X\theta_j + \hat{\theta}_m'X'X\hat{\theta}_j \\
&= Y_m'X(\theta_j - \hat{\theta}_j) - \hat{\theta}_m'X'X(\theta_j - \hat{\theta}_j)
\end{aligned}$$

ifadesi, $\hat{\theta}_m = (X'X)^{-1}X'Y_m$, θ_m 'nin en çok olabilirlik tahmin edicisi olmak üzere

$$\begin{aligned}
&= (Y_m'X - \hat{\theta}_m'X'X)(\theta_j - \hat{\theta}_j) \\
&= (Y_m'X - Y_m'X(X'X)^{-1}X'X)(\theta_j - \hat{\theta}_j) \\
&= (Y_m'X - Y_m'X)(\theta_j - \hat{\theta}_j) = 0
\end{aligned}$$

olduğundan,

$$S_{mj}(\theta_m, \theta_j) = (Y_m - X\hat{\theta}_m)'(Y_j - X\hat{\theta}_j) + (\theta_m - \hat{\theta}_m)'X'X(\theta_j - \hat{\theta}_j)$$

formunda yazılabilir. Böylece

$$S(\Theta) = A + (\Theta - \hat{\Theta})'X'X(\Theta - \hat{\Theta})$$

iki matrisin toplamı şeklinde ifade edilir. $[A = \{a_{mj}\}]$,

$a_{mj} = (Y_m - X\hat{\theta}_m)'(Y_j - X\hat{\theta}_j)$ örneğe ilişkin kovaryans matrisi ile orantılı (p x p) boyutlu simetrik matris]

Bu durumda, gözlemler verilmişken Θ ve Σ 'ya ilişkin olabilirlik fonksiyonu,

$$\begin{aligned}\ell(\Theta, \Sigma / Y) &\propto p(E / \Theta, \Sigma) \\ &\propto |\Sigma|^{-\frac{n}{2}} \exp\left[-\frac{1}{2} i z \Sigma^{-1} S(\Theta)\right]\end{aligned}$$

formunda ifade edilir.

4.6. $p(\Theta / \Sigma, Y)$ Üzerinden i nci Gözleme İlişkin Gerçekleşmiş Hatanın Sonsal Dağılımı

Çok değişkenli-çoklu doğrusal regresyon analizinde bileşik bilgi-vermeyen önsel $p(\Theta, \Sigma) = p(\Theta)p(\Sigma) = |\Sigma|^{-(p+1/2)}$ 'un ve olabilirlik fonksiyonu $\ell(\Theta, \Sigma / Y)$ 'nin çarpımından, Θ ve Σ parametrelerinin bileşik sonsal dağılımı,

$$\begin{aligned}p(\Theta, \Sigma / Y) &\propto |\Sigma|^{-\frac{1}{2}(n+p+1)} \exp\left\{-\frac{1}{2} i z \Sigma^{-1} \left[A + (\Theta - \hat{\Theta})' X' X (\Theta - \hat{\Theta})\right]\right\} \\ &, \quad -\infty < \Theta < \infty, \quad \Sigma > 0\end{aligned}$$

olarak elde edilir.

Σ ve Y biliniyorken Θ 'nın Sonsal dağılımı

$\hat{\Theta}$, Θ için yeterli bir istatistik olduğunda $\ell(\Theta / \Sigma, \hat{\Theta}) \propto p(\hat{\Theta} / \Sigma, \Theta)$ olarak ifade edilmektedir. Buradan, Σ ve Y biliniyor olması koşulu altında Θ 'nın sonsal dağılımı,

$$p(\Theta / \Sigma, Y) \propto \ell(\Theta / \Sigma, \hat{\Theta}) p(\Theta) \quad , \quad p(\Theta) = \text{sabit}$$

$$\begin{aligned}
&\propto \ell(\Theta / \Sigma, \hat{\Theta}) \\
&\propto p(\hat{\Theta} / \Sigma, \Theta) \\
&\propto \exp \left[-\frac{1}{2} i z \Sigma^{-1} (\Theta - \hat{\Theta})' X' X (\Theta - \hat{\Theta}) \right]
\end{aligned} \tag{4.6}$$

pq boyutlu Normal dağılımdır $((\Theta / \Sigma, Y) \sim N(\hat{\Theta}, \Sigma \otimes (X^T X)^{-1}))$.

Y Biliniyorken Σ 'nın Marjinal Sonsal Dağılımı

Σ 'nın marjinal Sonsal dağılımı,

$$\begin{aligned}
p(\Sigma / Y) &= \int_{\Theta / \Sigma, Y} p(\Theta, \Sigma / Y) d(\Theta / \Sigma, Y) \\
&= \int_{\Theta / \Sigma, Y} |\Sigma|^{-\frac{1}{2}(n+p+1)} \exp \left\{ -\frac{1}{2} i z \Sigma^{-1} A \right\} \exp \left\{ -\frac{1}{2} i z \Sigma^{-1} (\Theta - \hat{\Theta})' X' X (\Theta - \hat{\Theta}) \right\} d(\Theta / \Sigma, Y) \\
&= |\Sigma|^{-\frac{1}{2}(n+p+1)} \exp \left\{ -\frac{1}{2} i z \Sigma^{-1} A \right\} |\Sigma|^{\frac{q}{2}} \\
&= |\Sigma|^{-\frac{1}{2}(n+p-q+1)} \exp \left\{ -\frac{1}{2} i z \Sigma^{-1} A \right\}
\end{aligned}$$

$(n+p-q+1)$ serbestlik dereceli Ters-Wishard (Inverted-Wishard) tır
 $((\Sigma / Y) \sim W^{-1}(A, n - q - p + 1))$.

(Σ^{-1} / Y) 'nin Marjinal Sonsal dağılımı da, Σ 'dan Σ^{-1} 'e dönüşümün Jacobian'ı
 $|\Sigma|^{p+1}$ olmak üzere,

$$\begin{aligned}
p(\Theta, \Sigma^{-1} / Y) &= p(\Theta, \Sigma / Y) |J| = p(\Theta, \Sigma / Y) |\Sigma|^{p+1} \\
&= |\Sigma|^{-\frac{1}{2}(n+p+1)} \exp\{iz\Sigma^{-1}A\} \exp\left\{-\frac{1}{2}iz\Sigma^{-1}(\Theta - \hat{\Theta})'X'X(\Theta - \hat{\Theta})\right\} |\Sigma|^{p+1} \\
&= |\Sigma|^{-\frac{1}{2}(n-p-1)} \exp\left\{-\frac{1}{2}iz\Sigma^{-1}A\right\} \exp\left\{-\frac{1}{2}iz\Sigma^{-1}(\Theta - \hat{\Theta})'X'X(\Theta - \hat{\Theta})\right\}
\end{aligned}$$

Θ ve Σ^{-1} 'in birleşik sonsal dağılımından

$$\begin{aligned}
p(\Sigma^{-1} / Y) &= |\Sigma|^{-\frac{1}{2}(n-p-1)} \exp\left\{-\frac{1}{2}iz\Sigma^{-1}A\right\} \int_{\Theta} \exp\left\{-\frac{1}{2}iz\Sigma^{-1}(\Theta - \hat{\Theta})'X'X(\Theta - \hat{\Theta})\right\} d(\Theta / \Sigma, Y) \\
&= |\Sigma|^{-\frac{1}{2}(n-p-1)} \exp\left\{-\frac{1}{2}iz\Sigma^{-1}A\right\} |\Sigma|^{\frac{q}{2}} \\
&= |\Sigma^{-1}|^{\frac{1}{2}(n-p-q-1)} \exp\left\{-\frac{1}{2}iz\Sigma^{-1}A\right\} \\
&= |\Sigma^{-1}|^{\frac{1}{2}(n-p-q+1)-1} \exp\left\{-\frac{1}{2}iz\Sigma^{-1}A\right\}
\end{aligned}$$

$(n-p-q+1)$ serbestlik dereceli Wishart Dağılımı $((\Sigma^{-1} / Y) \sim W(A^{-1}, n-p-q+1))$ olarak elde edilir.

i nci gözleme ilişkin Gerçekleşmiş Hata $\varepsilon_i = (Y_i - X_i \Theta) \Sigma^{-1/2}$ 'nın Σ biliniyor olması koşulu altındaki sonsal dağılımı, Σ biliniyor olması koşulu altında Θ 'nın sonsal dağılımı üzerinden,

$$p(\varepsilon_i / \Sigma, Y) \propto \exp\left\{-\frac{1}{2} \left\{X_i (X'X)^{-1} X_i'\right\} [\varepsilon_i - \hat{\varepsilon}_i]' I [\varepsilon_i - \hat{\varepsilon}_i]\right\}$$

olarak elde edilmiştir (39). Burada $\hat{\varepsilon}_i = (Y_i - X_i \hat{\Theta}) \Sigma^{-1/2}$ i nci gözleme ilişkin gerçekleşmiş hatanın tahminini ifade etmektedir.

4.7. i nci Gözleme İlişkin Gerçekleşmiş Hatanın Karesel Miktarının Sonsal Dağılımı

$(\epsilon_i / \Sigma, Y) \sim N_p \left((Y_i - X_i \hat{\Theta}) \Sigma^{-1/2} = \hat{\epsilon}_i, (X_i (X'X)^{-1} X'_i)^{-1} I \right)$ olduğundan ve

$R_i = \epsilon_i \epsilon'_i$ olmak üzere $T_i = \frac{R_i}{X_i (X'X)^{-1} X'_i}$ ifadesi, serbestlik derecesi p, merkezi olmama parametresi,

$$\lambda_i = \frac{1}{X_i (X'X)^{-1} X'_i} (\hat{\epsilon}_i \hat{\epsilon}'_i),$$

olan, merkezi olmayan ki-kare (χ^2) dağılımına sahiptir (1).

Böylece, Σ üzerinden, $h_i = X_i (X'X)^{-1} X'_i$ olmak üzere,

$$p_i = E_{\Sigma/Y} \left\{ \text{pr}(T_i > \frac{k}{h_i} / \Sigma, Y) \right\}$$

olasılığı veya $(\Sigma^{-1} / Y) \sim W(S^{-1}, p, n - q - p + 1)$ olmak üzere, Σ^{-1} üzerinden aynı olasılık,

$$p_i = E_{\Sigma^{-1}/Y} \left\{ \text{pr}(T_i > \frac{k}{h_i} / \Sigma^{-1}, Y) \right\}$$

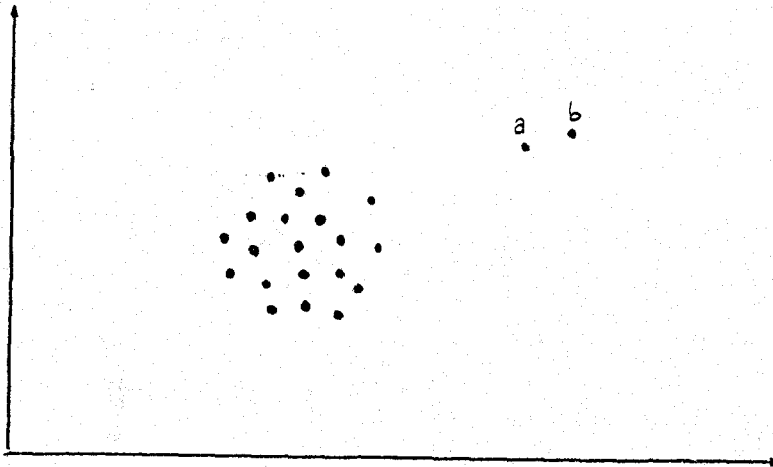
şeklinde ifade edilir.

Uygulamada p_i sonsal olasılıkları Monte-Carlo simülasyon tekniğine uygun olarak MAT-LAB'ta yazılmış programın kullanılması ile elde edilecektir. Programın yazılımı EK.1'de yer almaktadır.



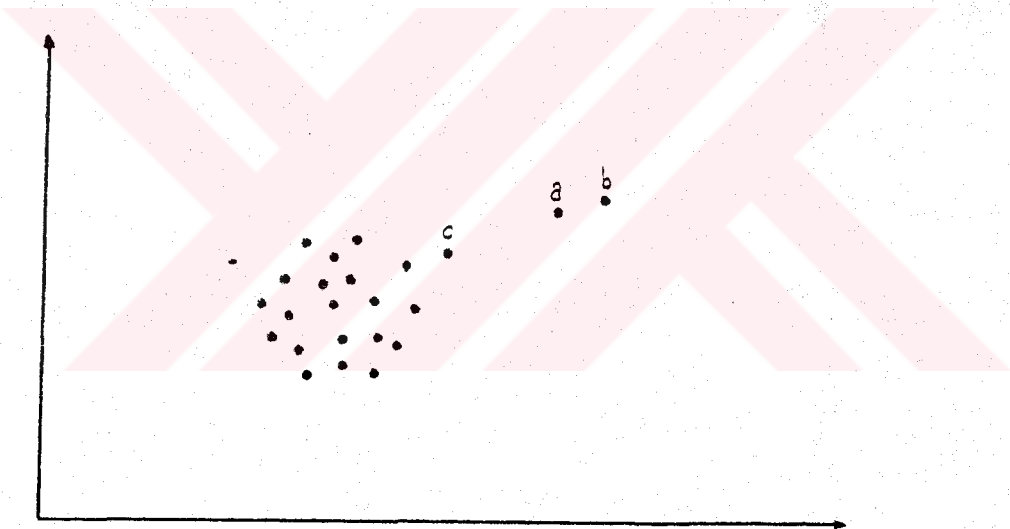
5. GİZLEME (masking) VE YANILGIYA-DÜŞÜRME (swamping) PROBLEMLERİ

Klasik ve Bayesgil yaklaşımlarla, çoklu doğrusal regresyon analizinde aykırı gözlemlerin tek tek incelenmesine ilişkin pek çok çalışma literatürde yer almaktadır. Son yıllarda ise üzerinde durulan konu gizleme ve yanılgiya- düşürme problemleridir. Bu problemlerin çözümü için gözlemlerin gruplar halinde aykırı olup olmadıklarının incelenmesi gerekmektedir. Gözlemlerin gruplar halinde aykırı olup olmadıklarının incelenmesi iki problemi de beraberinde getirmektedir. Bu problemlerden ilki, tesadüfi örneğin herhangi bir gerçekleşişinde yer alan gözlemlerin tüm mümkün alt gruplarının aykırı olup olmadıklarına ilişkin testlerin yapılmasının getirdiği işlem zorluğudur. İkincisi de, gözlemlerin gruplar halinde aykırı olup olmadıklarını test ederek gizleme problemi önlenmek istenirken yanılgiya-düşürme problemiyle karşılaşılmasıdır. Yani, “d” tesadüfi örneğin herhangi bir gerçekleşişindeki gerçek(bilinmeyen) aykırı gözlem sayısı olmak üzere, d 'den daha az sayıda gözlem içeren bir grubun aykırı olup olmadığı inceleniyor olsun. Gerçekte aykırı olup bu grubun dışında yer alan aykırı gözlem ya da gözlemler grubun içinde yer alan aykırı gözlemleri, aykırı olarak nitelememizi engelleyebilir. Bu durum gizleme problemi olarak isimlendirilmektedir (Şekil 5.1).



Şekil 5.1 (Gizleme)Gerçekte iki gözlem aykırı iken, gözlemlerin tek tek aykırı olup olmadıkları incelenir ise a'dan dolayı b, b'den dolayı da a gözlemi gizlenmiş olur

Oysa biz tesadüfî örneğin herhangi bir gerçekleşişinin içerdđđi gerek aykırı gözlem sayısını bilmemekteyiz. Dolayısı ile ilk bakışta tesadüfî örneğin herhangi bir gerçekleşişinde yer alan gözlemler üzerinden oluşturulacak tüm mümkün alt gruplarının aykırı olup olmadığının incelenmesi ve en fazla sayıda aykırı gözlem içerdđđini düşündüğümüz alt grubun, tesadüfî örneğin mevcut gerçekleşişindeki tüm aykırı gözlemleri içerdđđi düşüncesiyle gizleme problemini giderdiğimizi düşünebiliriz. Fakat “d” den daha fazla sayıda gözlem içeren bir alt grup aykırı olarak nitelenmiş olabilir. Bu grupta yer alan “d” tane gerek aykırı gözlem, gerekte aykırı olmayıp ta bu grupta yer alan gözlemleri de aykırı olarak nitelememize sebep olmuş olur. Bu durum, yanılgıya-düşürme problemi olarak isimlendirilmektedir (Şekil 5.2).



Şekil 5.2 (Yanılığa-Düşürme) Gerçekte iki gözlem aykırı iken, gözlemlerin üçer üçer aykırı olup olmadıkları incelenir ise a ve b gözlemleri c'nin de aykırı olarak nitelenmesine sebep olur

Yukarıda ifade edilen iki problemten ilkinin çözümüne ilişkin çeşitli yöntemler mevcuttur. Yöntemlerden biri Justel and Pena(27) tarafından önerilmiştir. Bu yöntemde işlem kalabalığının önüne geçmek için Gibbs örnekleme(sampling)(27) kullanılmaktadır.

İkinci problemin çözümüne ilişkin olarak klasik yaklaşımda önerilmiş yöntemlerden biri Barnett and Lewis (3) tarafından verilmiştir. Bu yöntem test istatistiğinin güç

fonksiyonundan hareketle örneğin içerdiği gerçek aykırı gözlem sayısına ilişkin bir tahminin elde edilmesini sağlamaktadır. Sadece elde edilen bu tahmin miktarında gözlem içeren alt grupların aykırı olup olmadığının incelenmesi, gizleme problemini engellerken yanlışya düşürme problemiyle karşılaşılmasını sağlayacaktır.

Çok değişkenli doğrusal regresyon analizinde, Klasik veya Bayesgil yaklaşımlar kullanılarak aykırı gözlemlerin belirlenmesine ilişkin çalışmalar literatürde çok az sayıda yer almakta ve bunlarda gözlemlerin tek tek aykırı olup olmadığını irdelemektedir. Varbanov(39) tarafından önerilmiş olan Bayesgil yöntem bu çalışmalardan biridir.

Bu Bayesgil yöntemde kullanılan mantıktan hareketle, gizleme ve yanlışya düşürme problemlerinin önlenmesi için gerekli ilk adım olan, gruplar halinde gözlemlerin aykırı olup olmadıklarının irdelenmesine , geçilebilir.

6. AYKIRI GÖZLEMLERİN GRUPLAR HALİNDE VARLIĞININ İRDELENMESİNE ÇOK DEĞİŞKENLİ ÇOKLU REGRESYON ANALİZİNDE BAYESGİL YAKLAŞIM

Bölüm (4.1)'de çok değişkenli-çoklu doğrusal regresyon modeli,

$$Y_{n \times p} = X_{n \times q} \theta_{q \times p} + E_{n \times p}, \quad E \sim N(0; \Sigma \otimes I_{n \times n})$$

şeklinde veya

$$Y_{n \times p} = X_{n \times q} \theta_{q \times p} + \varepsilon_{n \times p} \Sigma_{p \times p}^{1/2}, \quad \varepsilon_i = (\varepsilon_{i1}, \varepsilon_{i2}, \dots, \varepsilon_{ip}) \sim N(0; I_{p \times p})$$

olarak tanımlanmıştır. Matrislerin açık yazılması ile bu model,

$$\begin{bmatrix} Y_{11}, Y_{12}, \dots, Y_{1p} \\ Y_{21}, Y_{22}, \dots, Y_{2p} \\ \vdots \\ Y_{n1}, Y_{n2}, \dots, Y_{np} \end{bmatrix} = \begin{bmatrix} X_{11}, X_{12}, \dots, X_{1q} \\ X_{21}, X_{22}, \dots, X_{2q} \\ \vdots \\ X_{n1}, X_{n2}, \dots, X_{nq} \end{bmatrix} \begin{bmatrix} \theta_{11}, \theta_{12}, \dots, \theta_{1p} \\ \theta_{21}, \theta_{22}, \dots, \theta_{2p} \\ \vdots \\ \theta_{q1}, \theta_{q2}, \dots, \theta_{qp} \end{bmatrix} + \begin{bmatrix} \varepsilon_{11}, \varepsilon_{12}, \dots, \varepsilon_{1p} \\ \varepsilon_{21}, \varepsilon_{22}, \dots, \varepsilon_{2p} \\ \vdots \\ \varepsilon_{n1}, \varepsilon_{n2}, \dots, \varepsilon_{np} \end{bmatrix} \Sigma^{1/2}$$

biçiminde ifade edilir. X matrisinin ilk sütunu birler sütunu olarak değerlendirilmektedir. Bölüm (4.6)'daki [4.6] nolu eşitlikte verilen Σ koşul altındaki Θ 'nın sonsal dağılımı,

$$\Theta^* = [\theta_1; \theta_2; \dots; \theta_q] = [\theta_{11}, \theta_{12}, \dots, \theta_{1p}; \theta_{21}, \theta_{22}, \dots, \theta_{2p}; \dots; \theta_{q1}, \theta_{q2}, \dots, \theta_{qp}]$$

olmak üzere,

$$p(\Theta^* / \Sigma, Y) \propto \exp \left\{ -\frac{1}{2} (\Theta^* - \hat{\Theta}^*)' (X'X) \otimes \Sigma^{-1} (\Theta^* - \hat{\Theta}^*) \right\}$$

şeklinde de ifade edilebilmektedir. Aynı sonsal fonksiyon

$$\Theta^{**} = [\theta_1; \theta_2; \dots; \theta_p] = [\theta_{11}, \theta_{21}, \dots, \theta_{q1}; \theta_{12}, \theta_{22}, \dots, \theta_{q2}; \dots; \theta_{1p}, \theta_{2p}, \dots, \theta_{qp}]$$

olmak üzere,

$$p(\Theta^{**} / \Sigma, Y) \propto \exp \left\{ -\frac{1}{2} (\Theta^{**} - \hat{\Theta}^{**})' \Sigma^{-1} \otimes (X'X) (\Theta^{**} - \hat{\Theta}^{**}) \right\}$$

formunda da yazılmaktadır.

[4.6] nolu eşitlikten gerçekleşmiş ancak gözlenemeyen hataların sonsal dağılımını elde edebilmek için, $Y = X\Theta + E$ modelinden hareketle $\Theta = (X'X)^{-1} X'(Y - E)$ dönüşümü uygulandığı takdirde,

$$\begin{aligned} p(\epsilon / \Sigma, Y) &\propto \exp \left\{ -\frac{1}{2} \text{iz} \Sigma^{-1} \left[(X'X)^{-1} X'(Y - E) - \hat{\Theta} \right] (X'X) \left[(X'X)^{-1} X'(Y - E) - \hat{\Theta} \right] \right\} \\ &\propto \exp \left\{ -\frac{1}{2} \text{iz} \left\{ X \left[(X'X)^{-1} X'(Y - E) - \hat{\Theta} \right] \Sigma^{-1} \left[(X'X)^{-1} X'(Y - E) - \hat{\Theta} \right] X' \right\} \right\} \end{aligned}$$

olarak elde edilir. $H = X(X'X)^{-1} X'$ olmak üzere,

$$\begin{aligned} &\propto \exp \left\{ -\frac{1}{2} \text{iz} \left\{ [HY - HE - X\hat{\Theta}] \Sigma^{-1} [HY - HE - X\hat{\Theta}]' \right\} \right\} \\ &\propto \exp \left\{ -\frac{1}{2} \text{iz} \left\{ H[(Y - X\hat{\Theta}) - E] \Sigma^{-1} [(Y - X\hat{\Theta}) - E]' H' \right\} \right\} \end{aligned}$$

biçiminde yazılabilir.

$E = \varepsilon \Sigma^{1/2}$ olduğu için,

$$\propto \exp \left\{ -\frac{1}{2} \text{iz} \left\{ H \left[(Y - X\hat{\Theta}) - \varepsilon \Sigma^{1/2} \right] \Sigma^{-1} \left[(Y - X\hat{\Theta}) - \varepsilon \Sigma^{1/2} \right]' H' \right\} \right\}$$

$$\propto \exp \left\{ -\frac{1}{2} \text{iz} \left\{ H \left[(Y - X\hat{\Theta}) \Sigma^{-1/2} - \varepsilon \Sigma^{1/2} \Sigma^{-1/2} \right] I_{p \times p} \left[(Y - X\hat{\Theta}) \Sigma^{-1/2} - \varepsilon \Sigma^{1/2} \Sigma^{-1/2} \right]' H' \right\} \right\}$$

formunda ifade edilebilir. $(Y - X\hat{\Theta}) \Sigma^{-1/2} = \hat{\varepsilon}$ ile ifade edilmek üzere, Σ biliniyor olması koşulu altında gerçekleşmiş ancak gözlenemeyen hataların bileşik sonsal dağılımı

$$\propto \exp \left\{ -\frac{1}{2} \text{iz} \left\{ H [\hat{\varepsilon} - \varepsilon] I_{p \times p} [\hat{\varepsilon} - \varepsilon]' H' \right\} \right\}$$

biçiminde veya

$$\propto \exp \left\{ -\frac{1}{2} \text{iz} \left\{ I_{p \times p} [\hat{\varepsilon} - \varepsilon]' H' H [\hat{\varepsilon} - \varepsilon] \right\} \right\}$$

şeklinde yazılıp, H 'ın eş-güçlü (idempotent) bir matris olmasından dolayı da

$$\propto \exp \left\{ -\frac{1}{2} \text{iz} \left\{ I_{p \times p} [\hat{\varepsilon} - \varepsilon]' H [\hat{\varepsilon} - \varepsilon] \right\} \right\}$$

formunda ifade edilebilir.

Tesadüfi örneğin herhangi bir gerçekleşinde yer alan gözlemlerin tüm mümkün alt gruplarından birinin aykırı bir alt grup olduğunu, bu grupta yer alan gözlemlerin gerçekleşmiş ancak gözlenemeyen hatalarının bileşik marjinal fonksiyonları üzerinden, gerçekleşmiş ancak gözlenemeyen bu hataların karesel formlarının dağılımını elde edebilmek, ε 'nın

$$\varepsilon_{1 \times np}^* = [\varepsilon_1; \varepsilon_2; \dots; \varepsilon_i; \dots; \varepsilon_n] = [\varepsilon_{11}, \varepsilon_{12}, \dots, \varepsilon_{1p}; \varepsilon_{21}, \varepsilon_{22}, \dots, \varepsilon_{2p}; \dots; \varepsilon_{n1}, \varepsilon_{n2}, \dots, \varepsilon_{np}]$$

formunda ifade edilmesini gerektirmektedir. Buna göre, ε^* 'ın sonsal dağılımı,

$$p(\varepsilon^* / \Sigma, Y) \propto \exp \left\{ -\frac{1}{2} (\varepsilon^* - \hat{\varepsilon}^*) (H \otimes I_{p \times p}) (\varepsilon^* - \hat{\varepsilon}^*)' \right\} \quad [6]$$

olarak elde edilir ($\varepsilon^* \sim N(\hat{\varepsilon}^*, (H \otimes I_{p \times p})^{-1})$).

Gerçekleşmiş ancak gözlenemeyen hataların karesel formu olan $R = \varepsilon^* (\varepsilon^*)'$ miktarının (ölçüsünün) sonsal dağılımını elde edebilmek için varyans kovaryans matrisi $(H \otimes I_{p \times p})$ 'nın, $I_{np \times np}$ formuna gelmesini sağlayacak bir dönüşümün uygulanması gerekmektedir.

U ; $(np \times np)$ 'lik $(H \otimes I_{p \times p})$ matrisinin öz vektörler matrisi

Λ ; $(H \otimes I_{p \times p})$ matrisinin öz değerlerinden oluşan $(np \times np)$ 'lik köşegen matris olmak üzere, U ve Λ

$$U' (H \otimes I_{p \times p}) U = \Lambda I_{np \times np}$$

$$\Lambda^{-1/2} U' (H \otimes I_{p \times p}) U \Lambda^{-1/2} = I_{np \times np}$$

denklemlerinden elde edilirler. [6] nolu eşitlikte

$$\varepsilon^* = Z\Lambda^{-1/2}U'$$

dönüşümü uygulandığı taktirde, dönüşümün jacobian'ı sabit olduğu için ε^* yerine $Z\Lambda^{-1/2}U'$ yazıldığında, Σ biliniyor olması koşulu altında Z 'nin sonsal fonksiyonu,

$$p(Z/\Sigma, Y) = c. \exp\left\{-\frac{1}{2}\left\{\left[Z\Lambda^{-1/2}U' - \varepsilon^*\right](H \otimes I)\left[Z\Lambda^{-1/2}U' - \varepsilon^*\right]'\right\}\right\}$$

olarak elde edilir. Bu sonsal fonksiyon üzerinde işlemler devam ettirilirse,

$$p(Z/\Sigma, Y) = c. \exp\left\{-\frac{1}{2}\left\{\left[Z - \varepsilon^* \Lambda^{1/2}(U')^{-1}\right]\Lambda^{-1/2}U'(H \otimes I)U\Lambda^{-1/2}\left[Z - \varepsilon^* \Lambda^{1/2}(U')^{-1}\right]'\right\}\right\}$$

formu ile karşılaşılır. $\Lambda^{-1/2}U'(H \otimes I)U\Lambda^{-1/2} = I$ olduğu için, sonsal fonksiyon

$$p(Z/\Sigma, Y) = c. \exp\left\{-\frac{1}{2}\left\{\left[Z - \varepsilon^* \Lambda^{1/2}U'\right]\left[Z - \varepsilon^* \Lambda^{1/2}U'\right]'\right\}\right\}$$

biçiminde elde edilir. Bu ise np boyutlu çok değişkenli normal dağılımdır $((Z/\Sigma, Y) \sim N_{np}(\varepsilon^* \Lambda^{1/2}U'; I_{np \times np}))$.

$R = ZZ'$ miktarının (ölçüsünün) Σ biliniyor olması koşulu altındaki sonsal dağılımı, serbestlik derecesi np ve merkezi olmama parametresi

$$\lambda = \left[\varepsilon^* \Lambda^{-1/2}U'\right]\left[\varepsilon^* \Lambda^{-1/2}U'\right]'$$

olan ki-kare dağılımıdır.

Tesadüfi örneğin herhangi bir gerçekleşişinde yer alan gözlemlerin tüm mümkün alt gruplarından birinin aykırı olduğunu irdelleyebilmek için, bu alt grupta yer alan gözlemlerin gerçekleşmiş ancak gözlenemeyen hatalarının karesel miktarı (ölçüsü) olan $R_A = Z_A Z'_A$ 'nın Σ biliniyor olması koşulu altındaki sonsal dağılımının belirlenmesi gerekmektedir. Burada A , 2^n tane farklı durumdan α sayıda aykırı gözlemi barındıran $\binom{n}{\alpha}$ kombinasyonu kadar mümkün durumdan herhangi birini ifade etmektedir ($\alpha=1,2,\dots,n$). α 'nın tüm mümkün değerleri için uygulanabilir olan bu yöntem, gösterimi daha kolay olduğu için $\alpha=2$ durumu üzerinden aşağıda ifade edilecektir.

$\alpha=2$ durumu için, $\binom{n}{2}$ kombinasyonu kadar mümkün alt gruptan herhangi birinin aykırı olmadığı irdeleniyor olsun. i.nci ve j.nci gözlemler için,

$$Z_A = Z_{ij} = [Z_i; Z_j] = [Z_{i1}, Z_{i2}, \dots, Z_{ip}; Z_{j1}, Z_{j2}, \dots, Z_{jp}] \quad , \quad i, j = 1, 2, \dots, n$$

olmak üzere

$$R_A = R_{ij} = Z_{ij} Z'_{ij}$$

şeklinde ifade edilir. R_{ij} 'nin Σ biliniyor olması koşulu altındaki sonsal dağılımı, serbestlik derecesi $\alpha p = 2p$, merkezi olmama parametresi

$$\lambda_A = \lambda_{ij} = \left[\hat{\varepsilon}_{ij}^* \Lambda_{H_{ij} \otimes I}^{-1/2} U'_{H_{ij} \otimes I} \right] \left[\hat{\varepsilon}_{ij}^* \Lambda_{H_{ij} \otimes I}^{-1/2} U'_{H_{ij} \otimes I} \right]$$

olan ki-kare dağılımıdır. Burada $\Lambda_{H_{ij} \otimes I_{p \times p}}$ ve $U_{H_{ij} \otimes I_{p \times p}}$ sırasıyla i.nci ve j.nci gözlemlere ilişkin $(2p \times 2p)$ boyutlu $H_{ij} \otimes I_{p \times p}$ matrisinin öz değer (köşegen) ve öz vektör matrisleridir. Merkezi olmama parametresi,

$$\lambda_A = \lambda_{ij} = \hat{\varepsilon}_{ij}^* (H_{ij} \otimes I)^{-1} (\hat{\varepsilon}_{ij}^*)'$$

şeklinde de ifade edilebilir.

Dönüşümün sonunda kritik değer, $K = \left(\sqrt{\frac{k}{\alpha p}}; \sqrt{\frac{k}{\alpha p}}; \dots; \sqrt{\frac{k}{\alpha p}} \right)_{1 \times \alpha p = 1 \times 6}$ olmak üzere

$$k_Z = K(H_{ij} \otimes I)^{-1} K'$$

formunda elde edilir.

Merkezi olmayan bu sonsal dağılım üzerinden,

$$p_A = p_{ij}^* = E_{\Sigma/Y} \{ \Pr(R_{ij} > k_Z / \Sigma, Y) \}$$

ya da

$$p_A = p_{ij} = E_{\Sigma^{-1}/Y} \{ \Pr(R_{ij} > k_Z / \Sigma^{-1}, Y) \}$$

sonsals olasılıkları elde edilir.

Öngörülen model altında, $R_A = R_{ij} = \varepsilon_{ij}^* \varepsilon_{ij}^*$ istatistiğinin dağılımı $\alpha p = 2p$ serbestlik dereceli ki-kare dağılımıdır. n tane gözlemden hiçbirinin aykırı olmaması olasılığı

$$0.95 = \prod_{i=1}^n \Pr(R_i \leq k) \quad , \quad i=1,2,\dots,n$$

şeklinde seçilirse, $\alpha = 2$ durumu için, n gözlemden herhangi iki tanesinin aykırı olması olasılığı

$$\Pr(R_{ij} \geq k) = \Pr(R_i, R_j \geq k) = \Pr(R_i \geq k) \Pr(R_j \geq k) = (1 - (0.95)^{1/n})^2$$

olarak belirlenir. Elde edilen p_{ij} sonsal olasılıklarının $(1 - (0.95)^{1/n})^2$ 'den büyük olması, n tane gözlemin i .nci ve j .nci gözlemlerinden oluşan alt grubunun aykırı olduğu anlamını taşımaktadır.

Uygulamada, α 'nın tüm mümkün değerleri için p_A sonsal olasılıkları Monte-Carlo simülasyon tekniğine uygun olarak MAT-LAB'ta yazılmış program yoluyla elde edilmektedir. Programın yazılımı EK.2'de yer almaktadır.

7.UYGULAMA

Çok değişkenli-çoklu doğrusal regresyon analizinde tesadüfi örneğin herhangi bir gerçekleşişinde yer alan gözlemlerin tek tek ve gruplar halinde aykırı olup olmadıklarının (uyuşmazlıkların) irdelenmesinde, üzerinde durulan Bayesgil yöntemi uygulayabilmek için, Bruce E.,Barrett and Robert F.Ling(9) tarafından farklı bir yöntemle değerlendirilmiş olan “Rohwer’s data” isimli veri kullanılmıştır.

Bu veri üç bağımlı ve beş bağımsız değişken içermektedir. Bağımlı değişkenler,

Y_1 ; Peabody Picture Vocabulary Test (PRVT)

Y_2 ; Prograssive Matrices Test (RPMT)

Y_3 ; Student Achievement Test (SAT)

ve bağımsız değişkenlerde,

X_1 ; Named

X_2 ; Still

X_3 ; Named Still

X_4 ; Named Action

X_5 ; Sentence Still

şeklinde ifade edilmektedir. Tesadüfi olarak seçilmiş 32 öğrenciye ilişkin bağımlı ve bağımsız değişken değerleri EK.3’te yer almaktadır.

Çapı 32 olan tesadüfi örneğin, bu mevcut gerçekleşişinde gözlemlerin tek tek aykırı olup olmadığının (uyuşmazlığının) irdelenmesi için yazılımı EK.1’de verilmiş olan simülasyon programının kullanılması sonucu elde edilen p_i sonsal olasılıkları EK.3’te yer almaktadır. Gözlemlerin tek tek aykırı olup olmadıkları irdelenirken, i .nci gözleme ilişkin p_i sonsal olasılığının, tek bir gözlemin aykırı olmasına ilişkin öngörölmüş 0.0016 önsel olasılığından daha büyük olup olmadığına bakılmaktadır. Eğer p_i sonsal olasılığı 0.0016 ’den daha büyük ise i .nci gözlem aykırı olarak değerlendirilmektedir. EK.3’te yer alan sonuçlara bakıldığında 5.nci, 14.ncü ve

25.nci gözlemlere ilişkin $p_3=0.0018$, $p_{14}=0.0024$ $p_{25}=0.0143$ sonsal olasılıkları 0.0016 önsel olasılığından daha büyük oldukları görülmektedir. Buna göre, 5.nci, 14.ncü ve 25.nci gözlemleri tek tek irdelemenin sonucunda aykırı olarak değerlendirebiliriz. Benzer sonuç Bruce E.,Barrett and Robert F.Ling(9) tarafından elde edilmiştir. EK.5'te yer alan izdüşüm matrisi bu gözlemlerin merkeze ve birbirlerine olan uzaklıklarına ilişkin fikir edinmekte yardımcı olacaktır.

Eğer tesadüfi örneğin mevcut gerçekleşişi gerçekte 2 yada daha fazla sayıda aykırı gözlemi içeriyorsa, tek tek irdelemenin sonucunda 5,14 ve 25.nci gözlemlere ilişkin elde edilmiş sonsal olasılıklar, gerçekte mevcut olan aykırı gözlem sayısı birden fazla olduğu için diğer aykırı gözlemlerin gizlemesinden dolayı büyük çıkmamış olabilirler. Bu sebepten dolayı gözlemlerin ikiyeşerli gruplar halinde aykırı olduklarının irdelenmesi yerinde olacaktır.

Gözlemlerin ikiyeşerli gruplar halinde aykırı olduklarının irdelenmesinde ise EK.2'de yer alan simülasyon programının kullanılması sonucu elde edilen ve EK.4'te verilmiş olan sonuçlar elde edilmiştir. (5,14), (5,25) ve (14,25) gözlemlerinden oluşan ikiyeşerli gruplara ilişkin sonsal olasılıkların 0.000002566 önsel olasılığından büyük olması bu grupların aykırı olduğunu göstermektedir. Ancak tek tek irdelemenin yapılması durumunda aykırı görünmeyen gözlemlerin ikiyeşerli irdelemede pek çok ikiyeşerli aykırı grubun içerisinde yer aldığı görülmektedir. Bu durum tek tek irdelemede aykırı görülen gözlemler tarafından bu gözlemlerin yanlışlığı düşürüldüğü yada tek tek irdeleme yaparken birbirlerini maskeleydikleri anlamına gelebilir. Örneğin 10.ncu gözlem tek tek irdelemeler sonucunda aykırı bulunmamışken 14.ncü gözlemle oluşturdukları (10,14) ikili grup aykırı bulunmuştur. 10.ncu gözlemin 14.ncü gözlem tarafından yanlışlığı düşürüldüğünü söyleyebiliriz. Yine tek tek irdelemede 7 ve 9.ncu gözlemlerin aykırı olmadıkları görülmüşken (7,9) ikili grubunun aykırı bulunması bu iki gözlemin birbirini gizleyen gözlemler oldukları söylenebilir. 32 gözlemin 2'li tüm mümkün alt gruplarının sayısının çok büyük olması sonuçlar üzerinde yorum yapılmasını zorlaştırmaktadır.

Tek tek yapılmış irdellemelerde aykırı bulunmuş 5,14 ve 25.nci gözlemlerin dahil olduğu 25 ikili grup aykırı olarak belirlenmiştir. Bu 25 grupta yer alan gözlemlerden bazılarının 5,14 ve 25.nci gözlemler tarafından yanlışya-düşürüldükleri söylenebilir. Yanlışya-düşürülmediğini düşündüğümüz diğer gözlemler ise bu 25 lik grupta hiç yer almayan başka gözlemler tarafından gizlenmiş gözlemler olarak düşünebiliriz. Örneğin aykırı bulunmuş (2,25) grubunda 25.nci gözlemin 2.nci gözlemi yanlışya düşürdüğünü düşünebiliriz. Çünkü 2.nci gözlem 25.nci gözlemin haricinde diğer gözlemlerle ikişerli olarak irdelendiğinde 14 ,21 ve 31.nci gözlemlerle birlikte aykırı görülmüştür. Eğer 2.nci gözlem 14 ve 25.nci gözlemler tarafından yanlışya-düşürülmüş bir gözlem değilse tek tek irdelenmede 21.nci ve 31.nci gözlemler tarafından gizlenmiş bir gözlem demektir. Tüm mümkün ikili gruplar incelendiğinde 8,9,14,21,23,25 ve 31.nci gözlemlerin aykırı bulunmuş ikili grupların pek çoğunun içerisinde yer aldığı görülmektedir. Tesadüfi örneğin bu mevcut gerçekleşişinin gerçekte kaç tane aykırı gözlem içerdiğini(yada içermediğini) bilmediğimiz için ikiden fazla sayıda gözlem içeren grupların aykırı olduklarını irdelemeye devam etmek gerekmektedir. Aykırı olduğu irdelenmek istenen gruptaki gözlem sayısı arttıkça, mümkün alt grupların sayısı artmaktadır. Bu durumda kullanılacak simülasyon programından sonuçları almak, düşük işlem gücüne sahip bilgisayarlarda çok uzun süre almaktadır. Dolayısı ile uygulamada gözlemlerin tek tek ve ikişerli gruplar halinde irdelenmesi gerçekleştirilmiştir.

8. SONUÇ VE ÖNERİLER

Çok değişkenli-çoklu doğrusal regresyon analizinde Varbanov(39) tarafından gözlemlerin tek tek aykırı olup olmadıklarını irdelemeye yönelik Bayesgil bir yöntem ileri sürülmüştür. Belli bir hipotez testi prosedürüne bağlı kalmaksızın, amaca uygun olduğu düşünülen bir istatistiğin kullanılması ile gözlemlerin tek tek aykırı olup olmadıklarını irdelemeye yönelik bu yöntem, sezgisel uyumsuzluk testi görüşü başlığı altında yer verilmiştir. Bu yöntemde çok değişkenli gözlemleri merkeze olan uzaklıkları bakımından sıralamaya olanak sağladığı düşüncesinden dolayı, bir alt-sıralama ölçüsü olan gerçekleşmiş ancak gözlenemeyen hataların karesel formları kullanılmaktadır. Öngörülen model üzerinden, i .nci gözleme ilişkin karesel formun “ k ” kritik değerinden daha büyük olmasının sonsal olasılığı, önsel olasılıktan daha büyük ise bu gözlem aykırı olarak değerlendirilmektedir. Sadece öngörülen modelin varlığına bağlı bu yöntem, aykırı gözlemi de tanımladığı düşünülen alternatif bir modelin varlığını gerektirmemektedir. Çünkü kullanılan aykırı gözlem tanımı “aynı model tarafından üretilmiş ancak diğerlerinden aşırı derecede uzak olan gözlem” olarak yapılmaktadır. Aykırı gözlemi de tanımladığı düşünülen alternatif modellerin çokluğu, en uygun alternatif model olarak hangisinin kullanılacağı konusunda bir dezavantaj oluşturabilir. Dolayısı ile bu Bayesgil yöntem, tanım gereği alternatif bir model tanımına ihtiyaç duymaması bakımından söz konusu dezavantajı taşımamaktadır. Ayrıca uygulanabilirliği Bölüm 2.2.2’de özetlenen diğer Bayesgil yöntemlere göre daha kolaydır.

Bilgi-vermeyen önsel fonksiyonun haricinde başka önsel fonksiyonlarda kullanılabilir. Yalnız kullanılan herhangi bir önsel fonksiyon sonsal fonksiyonun tanıdık bir formda çıkmamasına yada elde edilememesine neden olabilir. Merkezi parametre için normal, ölçüm parametreleri içinde Ters-Wishart dağılımının kullanılması sonucu elde edilen sonsal dağılım, bilgi-vermeyen önsel dağılımın kullanılmasıyla elde edilen sonsal dağılım ile aynı aileden olmaktadır. Ancak güçlü bir önsel bilginin mevcut olmayışı olsa bile istatistik olarak “gözlemlerin, tesadüfi örneğin mevcut gerçekleşişinin ortalamasına olan uzaklıkları” aykırı gözlemleri belirlemekte kullanıldığı için güçlü bir önsel fonksiyonun kullanılmasına ihtiyaç

duyulmamaktadır. Yani aykırı gözlemi belirlemekte kullanılan istatistikten dolayı bilgi-vermeyen önsel fonksiyon kullanılmaktadır.

Aykırı gözlemlerle ilgili olarak gizleme ve yanılgıya-düşürme son yıllarda üzerinde durulan önemli problemlerdir. 5.Bölümde ayrıntılı olarak tanımlanmış bu problemlerle ilgilenabilmek için gözlemlerin gruplar halinde aykırı olup olmadıklarının irdelenmesi gerekmektedir. Bu amaca uygun olarak, Varbanov tarafından önerilmiş yöntemin, gözlemlerin gruplar halinde karşılaştırılmasına izin verecek şekilde genişletilmesi tezin amacını oluşturmaktadır.

Uygulamada tek tek ve ikişerli gruplar halinde aykırı gözlemlerin varlığı irdelenmiştir. Aykırı olduğu irdelenecek gruptaki gözlem sayısının artması, önsel olasılığın çok küçük olarak belirlenmesine sebep olmaktadır. Dolayısı ile önsel olasılık üzerinden belirlenen “k” kritik değeri çok büyük çıkmaktadır. Gerçekleşmiş ancak gözlenemeyen hatanın karesel formunun bu kritik değerden daha büyük olması sonsal olasılığının mevcut bilgisayar işlemcileri ile elde edilmesi çok uzun zaman almaktadır. Bu sebepten dolayı üç ve daha fazla sayıda gözlem içeren grupların aykırı olduklarının irdelenmesine uygulamada yer verilememiştir. Ayrıca gruptaki gözlem sayısının artmasına bağlı olarak önsel olasılığın düşmesi çok gözlem içeren grupların aykırı bulunmasını zorlaştırmaktadır.

İki ve daha fazla gözlem içeren grupların aykırı olduklarını irdelemeye yönelik geliştirilen bu yöntemden sonra, grupta yer alan gözlem sayısının artması ile orantılı olarak artacak tüm mümkün grupların irdelenmelerinin getirdiği işlem çokluğunun önüne geçmek gibi problemlerle ilgilenilebilir. Literatürde tesadüfi örneğin herhangi bir gerçekleşinde yer alabilecek aykırı gözlem sayısına ilişkin bir tahmin elde edebilmeye yönelik çalışmalar mevcuttur. Bu tür çalışmalar elde edilen tahmin değerinden farklı sayıda gözlem içeren grupların aykırılıklarının irdelenmesinin gerekliliğini ortadan kaldırmaktadır. Bu da işlem sayısının azalması anlamına gelir. Gibbs örneklemesi (Gibbs sampling)(27) gibi yöntemlerin kullanılması ile de işlem zorluğunun önüne geçmeye yönelik çalışmalar mevcuttur.

KAYNAKLAR

1. Anderson T.W., An Introduction to Multivariate Statistical Analysis, *John Wiley&Sons.Inc*, New York1958.
2. Arnold Zellner, Bayesian Analysis of Regression Error Terms, *J.A.S.A.*, Volume 70, Number 349, Theory and Methods Section, 1975.
3. Barnett,V., and Lewis,T., Outlier in Statistical Data, *John Wiley & Sons*, England, 1994.
4. Barnett,V., The ordering of multivariate data (with discussion),*J.Roy.Statist.Soc.A*, 139, 318-354, 1976b.
5. Barnett,V., Some outlier tests for multivariate samples, *Souht African Statist. J.*,13,29-52, 1979.
6. Box,G.E.P., Sampling and Bayesian inference in scientific modelling and robustness (with discussion), *J.R.Statist.Soc.B*, 26, 161-175,1980.
7. Box G.E.P.,Tiao,G.C., A Bayesian approach to some outlier problems, *Biometrika*, 55, 1, p.119,1968.
8. Box,G.E.P.,Tiao,G.C., Bayesian inference in statistical analysis, *Reading,MA, Addison-Wesley*,1973.
9. Bruce,E.,Barrett and Robert,F.Ling, General classes of influence measures for multivariate regression, *J.A.S.A.*, March 1992, Vol.87,No.417,Theory and Methods, 1992.
10. Chaloner,K., Brant,R., A Bayesian approach to outlier detection and residual analysis, *Biometrika*, 75, 651-9, 1988.
11. Dawid,A.P., Posterior expectations for large observations, *Biometrika*, 60, 664-667,1973.
12. Geisser,S., Bayesian Estimation in Multivariate Analysis, *Ann.Math.Statist*, 36, 150-9,1965.
13. Geisser,S., Discussion of a paper by G.E.P.Box, *J.R.Statist.Soc.A*. 143, 416-7, 1980.
14. Geisser,S., On the predicting of observables: a selective update, *In Bayesian Statistics 2*, Ed.J.M.Bernardo, M.H.DeGroot, D.V.Lindley and A.F.M.Smith, pp.203-30, Amsterdam, North Holland, 1985.

15. Geisser, S., Influential observations, diagnostics and discordancy tests, *J. Appl. Statist.*, 14, 133-42, 1987.
16. Geisser, S., Predictive approaches to discordancy testing, *In Bayesian and Likelihood Influence*, Essays in Honor of George Barnard, Ed. S. Geisser, J. S. Hodges, S. J. Press and A. Zellner, pp. 321-35, Amsterdam, North Holland, 1990.
17. Geisser, S., Diagnostics, divergences and perturbations analysis, *The IMA Volumes in Mathematics and its Applications*, 33(1), Ed. W. Stahel and S. Weisberg, pp. 89-100, New York, Springer Verlag, 1991.
18. Guttman, I., Care and Handling of univariate or Multivariate outliers in Detecting Spuriousity- a Bayesian Approach, *Technometrics*, 15, 723-738, 1973b.
19. Guttman, I., Dutter, R. & Freeman, P. R., Care and handling of univariate outliers in the general linear model to detect spuriousity- a Bayesian approach, *Technometrics*, 20, 187-93, 1978.
20. Guttman, I. & Pena, D., Outlier and influence : Evaluation of posteriors of parameters in linear model, *In Bayesian Statistics 3*, Ed. J. M. Bernardo, M. H. DeGroot, D. V., Lindley and A. F. M. Smith, pp. 631-40, Oxford University Press, 1988.
21. Guttman, I. & Pena D., A Bayesian look at diagnostics in the univariate linear model, *Taipai, Statistica Sinica*, To appear, 1993.
22. Johnson, R. A., An Asymptotic Expansions for Posterior Distributions, *Ann. Math. Statist.*, 38, 1899, 1967.
23. Johnson, R. A., Asymptotic Expansions Associated with Posterior Distributions, *Ann. Math. Statist.*, 41, 851, 1970.
24. J. M. Bernardo, Adrian F. M. Smith, Bayesian Theory, *John Wiley & Sons*, 1994.
25. Johnston, W. & Geisser, S., A predictive view of the detection and characterization of influential observations in regression analysis, *J. Am. Statist. Assoc.*, 78, 137-44, 1983.
26. Johnston, W. & Geisser, S., Estimative influence measures of the multivariate general linear model, *J. Statist. Plan. Inf.*, 11, 33-56, 1985.
27. Justel, A. & Pena, D., Bayesian unmasking in linear models, *Computational Statistics & Data Analysis*, 36, 69-84, 2001.

28. Kass,R.,Raftery,A., Bayes factors, *JASA*, 90, 773-95, 1995.
29. Lindley,D.V., The choice of variables in multiple regression (with discussion), *J.Roy.Statist.Soc.B*, 30, 31-66, 1968.
30. O'Hagan,A., On outlier rejection phenomena in Bayes Inference, *J.Roy.Statist, Soc.B*, 41, 358-367, 1979.
31. O'Hagan,A., Kendall's Advanced Theory of Statistics, Volume 2B. Bayesian Inference, *John Wiley&Sons Inc*, New York, 1994.
32. Pena D.,Guttman I., Comparing probabilistic methods for outlier detection in linear models, *Biometrika*, 80, 3, pp.603-10, 1993.
33. Pena, D. & Tiao, G.C., Bayesian robustness functions for linear models, *In Bayesian Statistics 4*, Ed.J.M.Bernardo, J.O.Berger, A.P. Dawid and A.F.M.Smith, pp.365-88, Oxford University Press, 1992.
34. Pettit, L. I. & Smith , A.F.M., Outliers and Influential observations in linear models, *In Bayesian Statistics 2*, Ed. J.M.Bernardo, M.H. DeGroot, D.V. , Lindley and A.F.M. Smith, pp.473-94, Amsterdam, Elsevier, 1985.
35. Peter M.Lee , Bayesian Statistics, *Oxford University press*.New York, 1989.
36. Shapiro,S.S., and Wilk,M.B., An analysis of variance test for normality (complete samples), *Biometrika*, 52, 591-611, 1965.
37. Shapiro,S.S., and Wilk,M.B., An analysis of variance test for the exponential distribution(complete samples), *Technometrics*, 14, 355-370, 1972.
38. T.Leonard,J.S.J.Hsu, Bayesian Methods, *Cambridge University press*, England.1999.
39. Varbanov A., Bayesian approach to outlier detection in multivariate normal samples and linear models, *Commun.Statist.-Theory Meth.*, 27(3), 547-557, 1998.



EKLER

EK.1

```
y=[68,15,24;82,11,8;82,13,88;91,18,82;82,13,90;100,15,77;100,13,58;96,12,14;63,1
0,1;91,18,98;87,10,8;105,21,88;87,14,4;76,16,14;66,14,38;74,15,4;68,13,64;98,16,8
8;63,15,14;94,16,99;82,18,50;89,15,36;80,19,88;61,11,14;102,20,24;71,12,24;102,1
6,24;96,13,50;55,16,8;96,18,98;74,15,98;78,19,50]
```

```
X=[1,0,10,8,21,22;1,7,3,21,28,21;1,7,9,17,31,30;1,6,11,16,27,25;1,20,7,21,28,16;1,4
,11,18,32,29;1,6,7,17,26,23;1,5,2,11,22,23;1,3,5,14,24,20;1,16,12,16,27,30;1,5,3,17,
25,24;1,2,11,10,26,22;1,1,4,14,25,19;1,11,5,18,27,22;1,0,0,3,16,11;1,5,8,11,12,15;1,
1,6,10,28,23;1,1,9,12,30,18;1,0,13,13,19,16;1,4,6,14,27,19;1,4,5,16,21,24;1,1,6,15,2
3,28;1,5,8,14,25,24;1,4,5,11,16,22;1,5,7,17,26,15;1,0,4,8,16,14;1,4,17,21,27,31;1,5,8
,20,28,26;1,4,7,19,20,13;1,4,7,10,23,19;1,2,6,14,25,17;1,5,10,18,27,26]
```

```
Q=(inv(X'*X))*X'*y
```

```
u=((y-X*Q)')*(y-X*Q)
```

```
S=inv(u)
```

```
k=chi2inv(0.9984,3) .
```

```
for i=1:32
```

```
    p=0;
```

```
    top=0;
```

```
    xi=X(i,:);
```

```
    yi=y(i,:);
```

```
    x=inv(X'*X);
```

```
    H=(xi*x*xi');
```

```
    H1=inv(H);
```

```
    k1=k*H1;
```

```
    for t=1:20000
```

```
        [W]=wishrnd(S,24);
```

```
        L=H1*((yi-xi*Q)*W*(yi-xi*Q)');
```

```
        pr=ncx2cdf(k1,3,L);
```

```
        a=1-pr;
```

```
        top=top+a;
```

```
    p=top/t;
```

```
end  
B=(p*(0.9984))/((0.0016)*(1-p));  
[i,p,B]  
end
```



EK.2

```

y=[68,15,24;82,11,8;82,13,88;91,18,82;82,13,90;100,15,77;100,13,58;96,12,14;63,1
0,1;91,18,98;87,10,8;105,21,88;87,14,4;76,16,14;66,14,38;74,15,4;68,13,64;98,16,8
8;63,15,14;94,16,99;82,18,50;89,15,36;80,19,88;61,11,14;102,20,24;71,12,24;102,1
6,24;96,13,50;55,16,8;96,18,98;74,15,98;78,19,50]
X=[1,0,10,8,21,22;1,7,3,21,28,21;1,7,9,17,31,30;1,6,11,16,27,25;1,20,7,21,28,16;1,4
,11,18,32,29;1,6,7,17,26,23;1,5,2,11,22,23;1,3,5,14,24,20;1,16,12,16,27,30;1,5,3,17,
25,24;1,2,11,10,26,22;1,1,4,14,25,19;1,11,5,18,27,22;1,0,0,3,16,11;1,5,8,11,12,15;1,
1,6,10,28,23;1,1,9,12,30,18;1,0,13,13,19,16;1,4,6,14,27,19;1,4,5,16,21,24;1,1,6,15,2
3,28;1,5,8,14,25,24;1,4,5,11,16,22;1,5,7,17,26,15;1,0,4,8,16,14;1,4,17,21,27,31;1,5,8
,20,28,26;1,4,7,19,20,13;1,4,7,10,23,19;1,2,6,14,25,17;1,5,10,18,27,26]
I=[1,0,0;0,1,0;0,0,1];
HH=X*(inv(X'*X))*X'
Q=(inv(X'*X))*X'*y
u=((y-X*Q))*(y-X*Q);
S=inv(u);
k=chi2inv(0.999997434,6);
k1=(k/6)^(1/2);
for i=1:32
    p=0;
    top=0;
    xi=X(i,:);
    for j=i+1:32
        p=0;
        top=0;
        xj=X(j,:);
        xij=[xi;xj];
        Hij=xij*(inv(X'*X))*xij';
        k2=[k1;k1;k1;k1;k1;k1];
        H=kron(Hij,I);
        Hters=inv(H);

```

```

k3=(k2)'*(Hters)*(k2);
for t=1:20000
    [W]=wishrnd(S,24);
    e=(y-X*Q)*(W^(1/2));
    ei=e(i,:);
    ej=e(j,:);
    eij=[ei,ej];
    L=(eij)*(Hters)*((eij)');
    pr=ncx2cdf(k3,6,L);
    a=1-pr;
    top=top+a;
    p=top/t;
end
B=(p*(0.999997434))/((0.000002566)*(1-p));
[i,j,p,B]
end
end

```

EK.3

y =

68	15	24
82	11	8
82	13	88
91	18	82
82	13	90
100	15	77
100	13	58
96	12	14
63	10	1
91	18	98
87	10	8
105	21	88
87	14	4
76	16	14
66	14	38
74	15	4
68	13	64
98	16	88
63	15	14
94	16	99
82	18	50
89	15	36
80	19	88
61	11	14
102	20	24
71	12	24
102	16	24
96	13	50
55	16	8
96	18	98
74	15	98
78	19	50

X =

1	0	10	8	21	22
1	7	3	21	28	21
1	7	9	17	31	30
1	6	11	16	27	25
1	20	7	21	28	16
1	4	11	18	32	29
1	6	7	17	26	23
1	5	2	11	22	23
1	3	5	14	24	20
1	16	12	16	27	30
1	5	3	17	25	24
1	2	11	10	26	22
1	1	4	14	25	19
1	11	5	18	27	22
1	0	0	3	16	11
1	5	8	11	12	15
1	1	6	10	28	23
1	1	9	12	30	18
1	0	13	13	19	16
1	4	6	14	27	19
1	4	5	16	21	24
1	1	6	15	23	28
1	5	8	14	25	24
1	4	5	11	16	22
1	5	7	17	26	15
1	0	4	8	16	14
1	4	17	21	27	31
1	5	8	20	28	26
1	4	7	19	20	13
1	4	7	10	23	19
1	2	6	14	25	17
1	5	10	18	27	26

Q =

39.6971	13.2438	-28.4675
0.0673	0.0593	3.2571
0.3700	0.4924	2.9966
-0.3744	-0.1640	-3.8591
1.5230	0.1190	5.6662
0.4102	-0.1212	-0.6227

u =

1.0e+004 *		
0.3899	0.0281	0.1280
0.0281	0.0177	0.0624
0.1280	0.0624	1.7134

S =

0.0003	-0.0004	-0.0000
-0.0004	0.0072	-0.0002
-0.0000	-0.0002	0.0001

k = 13.7931

<u>i</u>	<u>p_i</u>	<u>B</u>
1.0000	0.0000	0.0014
2.0000	0.0000	0.0000
3.0000	0.0000	0.0000
4.0000	-0.0000	-0.0000
5.0000	0.0018	1.1252
6.0000	0.0000	0.0000
7.0000	0.0000	0.0000
8.0000	0.0000	0.0144
9.0000	0.0000	0.0013
10.0000	0.0000	0.0019
11.0000	0.0000	0.0000
12.0000	0.0000	0.0018
13.0000	0.0000	0.0000
14.0000	0.0024	1.5012
15.0000	0.0000	0.0000
16.0000	0.0000	0.0012
17.0000	0.0000	0.0143
18.0000	0.0000	0.0000
19.0000	0.0000	0.0063
20.0000	0.0000	0.0000
21.0000	0.0004	0.1338
22.0000	0.0000	0.0000
23.0000	0.0000	0.0000
24.0000	0.0000	0.0000
25.0000	0.0143	9.0526
26.0000	0.0000	0.0000
27.0000	0.0003	0.1025
28.0000	0.0000	0.0000
29.0000	0.0004	0.1241
30.0000	0.0000	0.0000
31.0000	0.0004	0.1117
32.0000	0.0000	0.0004

EK.4

<u>i</u>	<u>i</u>	<u>p_{ij}</u>	<u>B</u>				
1.0000	2.0000	0.0000	0.0000				
1.0000	3.0000	0.0000	0.0000				
1.0000	4.0000	0.0000	0.0000				
1.0000	5.0000	-0.0000	-0.0000				
1.0000	6.0000	0.0000	0.0000				
1.0000	7.0000	-0.0000	-0.0000				
1.0000	8.0000	0.0000	0.0000				
1.0000	9.0000	0.0000	0.0000				
1.0000	10.0000	0.0000	0.0012				
1.0000	11.0000	0.0000	0.0000				
1.0e+005 *	0.0000	0.0001	0.0000	6.3593			
1.0000	13.0000	0.0000	0.0000				
1.0000	14.0000	0.0000	0.0000				
1.0000	15.0000	0.0000	0.0083				
1.0000	16.0000	0.0000	0.0662				
1.0000	17.0000	0.0000	0.0056				
1.0000	18.0000	0.0000	0.0025				
1.0000	19.0000	0.0000	0.0043				
1.0000	20.0000	0.0000	0.0000				
1.0000	21.0000	0.0000	0.0000				
1.0000	22.0000	0.0000	0.0000				
1.0000	23.0000	0.0002	59.0634				
1.0000	24.0000	0.0000	0.0005				
1.0000	25.0000	0.0000	0.0000				
1.0000	26.0000	0.0000	0.1306				
1.0000	27.0000	0.0000	0.0078				
1.0000	28.0000	-0.0000	-0.0000				
1.0000	29.0000	0.0000	0.0000				
1.0e+004 *	0.0001	0.0030	0.0000	5.9592			
1.0000	31.0000	0.0000	0.0064				
1.0000	32.0000	0.0000	0.0000				
2.0000	3.0000	0.0000	0.0000				
2.0000	4.0000	-0.0000	-0.0000				
2.0000	5.0000	0.0000	0.0000				
2.0000	6.0000	-0.0000	-0.0000				
2.0000	7.0000	0.0000	2.3502				
2.0000	8.0000	0.0000	0.0000				
2.0000	9.0000	0.0000	0.0316				
2.0000	10.0000	0.0000	0.0000				
2.0000	11.0000	0.0000	0.0002				
2.0000	12.0000	-0.0000	-0.0000				
2.0000	13.0000	0.0000	0.0001				
1.0e+003 *	0.0020	0.0140	0.0000	3.6727			
2.0000	15.0000	0.0000	0.0000				
2.0000	16.0000	-0.0000	-0.0000				
2.0000	17.0000	0.0000	0.0000				
2.0000	18.0000	-0.0000	-0.0000				
2.0000	19.0000	-0.0000	-0.0000				
2.0000	20.0000	0.0000	0.0003				
2.0000	21.0000	0.0008	320.8649				
2.0000	22.0000	0.0000	0.0000				
2.0000	23.0000	-0.0000	-0.0000				
2.0000	24.0000	-0.0000	-0.0000				
1.0e+003 *	0.0020	0.0250	0.0000	2.6772			
2.0000	26.0000	-0.0000	-0.0000				
2.0000	27.0000	-0.0000	-0.0000				
2.0000	28.0000	0.0000	0.0115				
2.0000	29.0000	0.0000	0.1697				
2.0000	30.0000	0.0000	0.0000				
2.0000	31.0000	0.0011	448.3336				
2.0000	32.0000	0.0000	0.0148				
3.0000	4.0000	0.0000	0.0000				
3.0000	5.0000	0.0000	0.0000				
3.0000	6.0000	0.0002	96.0008				
3.0000	7.0000	0.0000	0.1036				
3.0000	8.0000	0.0000	5.1339				
3.0000	9.0000	0.0000	0.0000				
3.0000	10.0000	0.0000	0.0468				
3.0000	11.0000	0.0000	0.0000				
3.0000	12.0000	0.0000	2.2079				
3.0000	13.0000	0.0000	0.0000				
3.0000	14.0000	0.0014	534.7502				
3.0000	15.0000	-0.0000	-0.0000				
3.0000	16.0000	-0.0000	-0.0000				
3.0000	17.0000	0.0000	0.0019				
3.0000	18.0000	0.0000	0.0000				
3.0000	19.0000	0.0000	0.0000				
3.0000	20.0000	0.0000	0.0000				
3.0000	21.0000	0.0000	0.0041				
3.0000	22.0000	0.0000	0.0003				
3.0000	23.0000	0.0012	482.5461				
3.0000	24.0000	0.0000	0.0000				
3.0000	25.0000	0.0000	0.0000				
3.0000	26.0000	0.0000	0.0000				
3.0000	27.0000	0.0000	0.0058				
3.0000	28.0000	0.0000	0.0000				
3.0000	29.0000	0.0000	0.0000				
3.0000	30.0000	0.0000	0.0000				
3.0000	31.0000	0.0000	0.0000				
3.0000	32.0000	0.0007	278.6991				
4.0000	5.0000	-0.0000	-0.0000				
4.0000	6.0000	-0.0000	-0.0000				
4.0000	7.0000	0.0000	0.0000				
4.0000	8.0000	-0.0000	-0.0000				
4.0000	9.0000	0.0000	0.0000				
4.0000	10.0000	-0.0000	-0.0000				
4.0000	11.0000	-0.0000	-0.0000				
4.0000	12.0000	0.0000	0.0000				
4.0000	13.0000	-0.0000	-0.0000				
4.0000	14.0000	0.0000	0.0000				
4.0000	15.0000	-0.0000	-0.0000				
4.0000	16.0000	-0.0000	-0.0000				
4.0000	17.0000	0.0000	0.0000				
4.0000	18.0000	-0.0000	-0.0000				
4.0000	19.0000	-0.0000	-0.0000				
4.0000	20.0000	0.0000	0.0000				
4.0000	21.0000	0.0000	0.0000				
4.0000	22.0000	-0.0000	-0.0000				
4.0000	23.0000	0.0000	0.7019				
4.0000	24.0000	-0.0000	-0.0000				
4.0000	25.0000	0.0000	0.0000				
4.0000	26.0000	-0.0000	-0.0000				
4.0000	27.0000	0.0000	0.0000				
4.0000	28.0000	-0.0000	-0.0000				
4.0000	29.0000	-0.0000	-0.0000				
4.0000	30.0000	0.0000	0.0000				
4.0000	31.0000	-0.0000	-0.0000				
4.0000	32.0000	0.0002	64.1929				
5.0000	6.0000	-0.0000	-0.0000				
5.0000	7.0000	-0.0000	-0.0000				
5.0000	8.0000	0.0000	0.0006				
5.0000	9.0000	0.0000	0.0000				
5.0000	10.0000	0.0001	39.1247				
5.0000	11.0000	-0.0000	-0.0000				
5.0000	12.0000	0.0000	0.0000				
5.0000	13.0000	-0.0000	-0.0000				

1.0e+005 * 0.0001 0.0001 0.0000 1.4148

5.0000 5.0000 0.0000 0.0000
 5.0000 16.0000 0.0000 0.0033
 5.0000 17.0000 0.0000 0.0000
 5.0000 18.0000 0.0000 0.0000
 5.0000 19.0000 0.0000 0.0001
 5.0000 20.0000 0.0000 0.0000
 5.0000 21.0000 0.0000 0.0001
 5.0000 22.0000 -0.0000 -0.0000
 5.0000 23.0000 0.0000 0.0000
 5.0000 24.0000 0.0000 0.0000

1.0e+004 * 0.0005 0.0025 0.0000 2.3010

5.0000 26.0000 -0.0000 -0.0000
 5.0000 27.0000 0.0000 0.0001
 5.0000 28.0000 -0.0000 -0.0000
 5.0000 29.0000 0.0000 4.5000
 5.0000 30.0000 0.0000 0.0000
 5.0000 31.0000 0.0000 0.3079
 5.0000 32.0000 0.0000 0.0000

6.0000 7.0000 0.0000 0.0000
 6.0000 8.0000 -0.0000 -0.0000
 6.0000 9.0000 0.0000 0.0000
 6.0000 10.0000 0.0000 0.0000
 6.0000 11.0000 -0.0000 -0.0000
 6.0000 12.0000 0.0000 0.0016
 6.0000 13.0000 -0.0000 -0.0000
 6.0000 14.0000 0.0000 0.0002
 6.0000 15.0000 0.0000 0.0000
 6.0000 16.0000 -0.0000 -0.0000
 6.0000 17.0000 0.0000 0.0953
 6.0000 18.0000 0.0000 0.0000
 6.0000 19.0000 -0.0000 -0.0000
 6.0000 20.0000 0.0000 0.0000
 6.0000 21.0000 0.0000 0.0000
 6.0000 22.0000 -0.0000 -0.0000
 6.0000 23.0000 0.0000 0.7276
 6.0000 24.0000 -0.0000 -0.0000
 6.0000 25.0000 0.0000 0.0152
 6.0000 26.0000 -0.0000 -0.0000
 6.0000 27.0000 0.0000 0.0000
 6.0000 28.0000 -0.0000 -0.0000
 6.0000 29.0000 -0.0000 -0.0000
 6.0000 30.0000 0.0000 0.0000
 6.0000 31.0000 0.0000 0.0155

1.0e+004 * 0.0006 0.0032 0.0000 4.6032

7.0000 8.0000 0.0000 0.0000
 1.0e+004 * 0.0007 0.0009 0.0000 8.9093
 7.0000 10.0000 0.0000 0.0000
 7.0000 11.0000 0.0000 0.0000
 7.0000 12.0000 0.0000 0.0000
 7.0000 13.0000 0.0000 0.0164

1.0e+005 * 0.0001 0.0001 0.0000 1.0879

7.0000 15.0000 -0.0000 -0.0000
 7.0000 16.0000 -0.0000 -0.0000
 7.0000 17.0000 0.0000 0.0000
 7.0000 18.0000 -0.0000 -0.0000
 7.0000 19.0000 -0.0000 -0.0000
 7.0000 20.0000 0.0000 0.0000
 7.0000 21.0000 0.0002 58.9471
 7.0000 22.0000 0.0000 0.0000

1.0e+004 * 0.0007 0.0023 0.0000 2.3206

7.0000 24.0000 -0.0000 -0.0000
 7.0000 25.0000 0.0000 0.0276
 7.0000 26.0000 -0.0000 -0.0000
 7.0000 27.0000 0.0000 0.0000
 7.0000 28.0000 0.0000 0.0000
 7.0000 29.0000 0.0000 0.0002
 7.0000 30.0000 0.0000 0.0000
 7.0000 31.0000 0.0000 0.0026

1.0e+005 * 0.0001 0.0003 0.0000 5.5760

8.0000 9.0000 0.0001 27.1429
 8.0000 10.0000 0.0000 0.0999
 8.0000 11.0000 0.0000 0.0126
 8.0000 12.0000 0.0000 0.0000
 8.0000 13.0000 0.0000 0.0000
 8.0000 14.0000 0.0024 945.9460
 8.0000 15.0000 0.0001 49.5032
 8.0000 16.0000 0.0000 0.0000
 8.0000 17.0000 0.0008 296.9898
 8.0000 18.0000 0.0000 0.0000
 8.0000 19.0000 0.0000 0.0000
 8.0000 20.0000 0.0000 0.0000

1.0e+004 * 0.0008 0.0021 0.0000 2.7972

8.0000 22.0000 0.0000 5.4234
 8.0000 23.0000 0.0007 266.8087

1.0e+003 * 0.0080 0.0240 0.0000 5.2649

8.0000 25.0000 0.0000 0.0004
 8.0000 26.0000 0.0000 0.0001
 8.0000 27.0000 0.0000 0.0000
 8.0000 28.0000 0.0000 0.0000
 8.0000 29.0000 -0.0000 -0.0000
 8.0000 30.0000 0.0000 0.0001
 8.0000 31.0000 0.0000 0.0001
 8.0000 32.0000 -0.0000 -0.0000

9.0000 10.0000 0.0000 0.0000
 9.0000 11.0000 0.0003 105.3146
 9.0000 12.0000 0.0000 0.0000

1.0e+005 * 0.0001 0.0001 0.0000 6.7568

9.0000 14.0000 0.0000 0.0196
 9.0000 15.0000 0.0000 0.3534
 9.0000 16.0000 0.0000 0.0000
 9.0000 17.0000 0.0000 0.0000
 9.0000 18.0000 0.0000 0.0001
 9.0000 19.0000 0.0000 0.0000

1.0e+005 * 0.0001 0.0002 0.0000 1.5501

1.0e+003 * 0.0090 0.0210 0.0000 9.6673

9.0000 22.0000 0.0000 9.5403
 9.0000 23.0000 0.0008 317.5466
 9.0000 24.0000 0.0000 0.0000

1.0e+003 * 0.0090 0.0250 0.0000 1.5177

9.0000 26.0000 0.0000 0.2136
 9.0000 27.0000 0.0000 0.0000
 9.0000 28.0000 0.0001 21.4785
 9.0000 29.0000 0.0000 0.0015
 9.0000 30.0000 0.0001 38.3810

1.0e+005 * 0.0001 0.0003 0.0000 5.2199

9.0000 32.0000 0.0000 0.0025

10.0000 11.0000 -0.0000 -0.0000
 10.0000 12.0000 0.0000 0.0024
 10.0000 13.0000 0.0000 0.0000
 10.0000 14.0000 0.0007 283.1825
 10.0000 15.0000 0.0000 0.0000
 10.0000 16.0000 0.0000 0.0002
 10.0000 17.0000 0.0000 0.0001
 10.0000 18.0000 0.0000 0.0000
 10.0000 19.0000 0.0000 0.0000
 10.0000 20.0000 0.0000 0.0000
 10.0000 21.0000 0.0000 0.0055
 10.0000 22.0000 -0.0000 -0.0000
 10.0000 23.0000 0.0000 0.0063
 10.0000 24.0000 0.0000 0.0000
 10.0000 25.0000 0.0000 0.0038
 10.0000 26.0000 -0.0000 -0.0000
 10.0000 27.0000 0.0000 0.0244
 10.0000 28.0000 -0.0000 -0.0000
 10.0000 29.0000 -0.0000 -0.0000
 10.0000 30.0000 0.0000 0.0000

10.0000	31.0000	0.0000	0.0000		
10.0000	32.0000	0.0000	0.0000		
11.0000	12.0000	0.0000	0.0000		
11.0000	13.0000	0.0000	0.0000		
1.0e+003 *	0.0110	0.0140	0.0000	2.2749	
11.0000	15.0000	-0.0000	-0.0000		
11.0000	16.0000	-0.0000	-0.0000		
11.0000	17.0000	0.0000	0.0000		
11.0000	18.0000	-0.0000	-0.0000		
11.0000	19.0000	-0.0000	-0.0000		
11.0000	20.0000	0.0000	0.0000		
1.0e+006 *	0.0000	0.0000	0.0000	1.5961	
11.0000	22.0000	0.0001	34.3451		
11.0000	23.0000	0.0000	0.0128		
11.0000	24.0000	0.0000	0.0000		
11.0000	25.0000	0.0000	0.0001		
11.0000	26.0000	-0.0000	-0.0000		
11.0000	27.0000	-0.0000	-0.0000		
11.0000	28.0000	0.0000	0.0000		
11.0000	29.0000	0.0000	0.0000		
11.0000	30.0000	0.0000	0.0000		
11.0000	31.0000	0.0000	0.0549		
11.0000	32.0000	0.0000	0.1137		
12.0000	13.0000	0.0000	0.0000		
12.0000	14.0000	0.0000	0.0000		
12.0000	15.0000	0.0000	0.0002		
12.0000	16.0000	0.0000	0.0000		
1.0e+004 *	0.0012	0.0017	0.0000	9.8213	
12.0000	18.0000	0.0013	509.1295		
12.0000	19.0000	0.0001	24.2173		
12.0000	20.0000	0.0000	0.0003		
12.0000	21.0000	-0.0000	-0.0000		
12.0000	22.0000	-0.0000	-0.0000		
12.0000	23.0000	0.0000	0.0204		
12.0000	24.0000	-0.0000	-0.0000		
12.0000	25.0000	0.0000	0.4027		
12.0000	26.0000	0.0000	0.0000		
12.0000	27.0000	0.0000	0.0018		
12.0000	28.0000	0.0000	0.0000		
12.0000	29.0000	-0.0000	-0.0000		
12.0000	30.0000	0.0000	0.4299		
12.0000	31.0000	0.0001	21.0087		
12.0000	32.0000	0.0000	0.0000		
13.0000	14.0000	0.0000	0.0000		
13.0000	15.0000	0.0000	0.0000		
13.0000	16.0000	-0.0000	-0.0000		
13.0000	17.0000	0.0000	0.3750		
13.0000	18.0000	0.0000	0.0000		
13.0000	19.0000	-0.0000	-0.0000		
1.0e+004 *	0.0013	0.0020	0.0000	2.3393	
13.0000	21.0000	0.0000	0.1340		
13.0000	22.0000	0.0000	0.0000		
13.0000	23.0000	0.0000	0.0000		
13.0000	24.0000	-0.0000	-0.0000		
13.0000	25.0000	0.0000	0.0542		
13.0000	26.0000	0.0000	0.0000		
13.0000	27.0000	0.0000	0.0000		
13.0000	28.0000	0.0000	0.0001		
13.0000	29.0000	0.0000	0.0003		
13.0000	30.0000	0.0000	0.0000		
1.0e+007 *	0.0000	0.0000	0.0000	1.2253	
13.0000	32.0000	0.0000	0.0000		
14.0000	15.0000	0.0000	0.0452		
14.0000	16.0000	0.0000	0.0117		
14.0000	17.0000	0.0000	0.0314		
14.0000	18.0000	0.0000	0.0005		
14.0000	19.0000	0.0000	0.0000		
14.0000	20.0000	0.0008	300.5392		
14.0000	21.0000	0.0001	22.4841		
14.0000	22.0000	0.0000	0.0001		
14.0000	23.0000	0.0000	0.0018		
14.0000	24.0000	0.0000	0.0539		
14.0000	25.0000	0.0001	34.4975		
14.0000	26.0000	0.0000	0.0000		
14.0000	27.0000	0.0000	0.0001		
14.0000	28.0000	0.0002	67.0873		
14.0000	29.0000	0.0000	0.9844		
14.0000	30.0000	0.0000	0.0059		
14.0000	31.0000	0.0000	13.6588		
14.0000	32.0000	0.0000	0.0009		
15.0000	16.0000	0.0000	0.0000		
15.0000	17.0000	0.0000	0.8445		
15.0000	18.0000	0.0000	0.0000		
15.0000	19.0000	0.0000	0.0000		
15.0000	20.0000	0.0000	0.0000		
15.0000	21.0000	0.0000	0.0011		
15.0000	22.0000	-0.0000	-0.0000		
15.0000	23.0000	0.0000	0.0000		
15.0000	24.0000	0.0000	0.0000		
15.0000	25.0000	0.0000	1.6537		
15.0000	26.0000	0.0000	7.4563		
15.0000	27.0000	-0.0000	-0.0000		
15.0000	28.0000	-0.0000	-0.0000		
15.0000	29.0000	0.0000	0.0000		
15.0000	30.0000	0.0000	0.4338		
15.0000	31.0000	0.0000	6.9472		
15.0000	32.0000	-0.0000	-0.0000		
16.0000	17.0000	-0.0000	-0.0000		
16.0000	18.0000	-0.0000	-0.0000		
16.0000	19.0000	0.0001	27.8949		
16.0000	20.0000	0.0000	0.0000		
16.0000	21.0000	0.0000	17.9648		
16.0000	22.0000	-0.0000	-0.0000		
16.0000	23.0000	0.0000	0.0000		
16.0000	24.0000	0.0001	51.5886		
16.0000	25.0000	0.0000	0.9073		
16.0000	26.0000	0.0000	0.0000		
16.0000	27.0000	0.0000	0.0009		
16.0000	28.0000	-0.0000	-0.0000		
16.0000	29.0000	0.0004	154.8905		
16.0000	30.0000	0.0000	0.0000		
16.0000	31.0000	0.0000	0.0005		
16.0000	32.0000	0.0000	0.0000		
1.0e+003 *	0.0170	0.0180	0.0000	4.7751	
17.0000	19.0000	0.0000	0.0000		
1.0e+003 *	0.0170	0.0200	0.0000	4.0055	
17.0000	21.0000	0.0000	0.0000		
17.0000	22.0000	0.0000	0.0001		
17.0000	23.0000	0.0000	0.0888		
17.0000	24.0000	0.0000	0.0000		
17.0000	25.0000	0.0000	0.1554		
17.0000	26.0000	0.0000	0.0000		
17.0000	27.0000	0.0000	0.0000		
17.0000	28.0000	0.0000	0.0000		
17.0000	29.0000	0.0000	0.0000		
1.0e+004 *	0.0017	0.0030	0.0000	1.0240	
17.0000	31.0000	0.0001	48.4856		
17.0000	32.0000	0.0000	0.0000		
18.0000	19.0000	0.0000	0.0000		
18.0000	20.0000	0.0000	0.1811		
18.0000	21.0000	0.0000	0.0000		
18.0000	22.0000	-0.0000	-0.0000		
18.0000	23.0000	0.0000	0.0000		
18.0000	24.0000	-0.0000	-0.0000		

EK.5

H (izdüşüm matrisi)

Kolon (1- 8)

0.1670	-0.1191	0.0297	0.0577	-0.1050	0.0370	-0.0066	0.0297
-0.1191	0.2185	0.0343	-0.0128	0.0988	0.0372	0.0746	0.0469
0.0297	0.0343	0.1417	0.0670	0.0135	0.1216	0.0470	0.0624
0.0577	-0.0128	0.0670	0.0731	0.0471	0.0771	0.0301	-0.0096
-0.1050	0.0988	0.0135	0.0471	0.5682	-0.0468	0.0529	-0.0078
0.0370	0.0372	0.1216	0.0771	-0.0468	0.1543	0.0449	-0.0098
-0.0066	0.0746	0.0470	0.0301	0.0529	0.0449	0.0453	0.0332
0.0297	0.0469	0.0624	-0.0096	-0.0078	-0.0098	0.0332	0.1766
0.0108	0.0704	0.0207	0.0066	-0.0033	0.0237	0.0348	0.0514
0.0726	-0.0881	0.1340	0.1086	0.2674	0.0378	0.0302	0.0783
-0.0461	0.1426	0.0564	-0.0096	-0.0133	0.0298	0.0595	0.1109
0.1405	-0.0969	0.0570	0.0757	-0.0186	0.0782	-0.0005	-0.0095
-0.0026	0.1017	0.0165	-0.0058	-0.0451	0.0383	0.0375	0.0445
-0.0539	0.1072	0.0579	0.0210	0.2086	0.0127	0.0542	0.0717
0.0986	-0.0284	-0.0486	-0.0405	-0.0004	-0.0938	-0.0108	0.1427
0.0860	-0.0647	-0.0952	0.0179	0.0809	-0.1146	0.0040	0.0194
0.0857	-0.0040	0.0905	0.0328	-0.0722	0.0885	0.0171	0.0760
0.0726	-0.0036	0.0420	0.0540	0.0222	0.1016	0.0113	-0.0518
0.1132	-0.0738	-0.0732	0.0622	-0.0235	0.0086	0.0002	-0.1210
0.0137	0.0587	0.0335	0.0229	0.0616	0.0469	0.0305	0.0192
-0.0026	0.0850	0.0255	0.0013	-0.0545	0.0041	0.0483	0.0920
0.0363	0.0615	0.0703	0.0143	-0.1843	0.0607	0.0458	0.1036
0.0533	0.0042	0.0603	0.0440	0.0131	0.0476	0.0293	0.0493
0.0697	-0.0130	0.0002	0.0029	-0.0429	-0.0545	0.0222	0.1279
-0.0303	0.0961	-0.0257	0.0204	0.1433	0.0210	0.0364	-0.0571
0.0813	-0.0136	-0.0608	-0.0145	-0.0382	-0.0650	0.0032	0.0641
0.0746	-0.0286	0.0707	0.1159	-0.0854	0.1373	0.0416	-0.1013
-0.0305	0.1148	0.0656	0.0359	-0.0115	0.0871	0.0600	0.0115
-0.0558	0.1249	-0.1002	-0.0007	0.1069	-0.0380	0.0433	-0.0811
0.0878	-0.0401	0.0246	0.0347	0.0511	0.0099	0.0082	0.0476
0.0126	0.0671	-0.0036	0.0127	0.0222	0.0322	0.0285	-0.0062
0.0220	0.0425	0.0645	0.0575	-0.0049	0.0841	0.0441	-0.0034

Kolon (9-16)

0.0108	0.0726	-0.0461	0.1405	-0.0026	-0.0539	0.0986	0.0860
0.0704	-0.0881	0.1426	-0.0969	0.1017	0.1072	-0.0284	-0.0647
0.0207	0.1340	0.0564	0.0570	0.0165	0.0579	-0.0486	-0.0952
0.0066	0.1086	-0.0096	0.0757	-0.0058	0.0210	-0.0405	0.0179
-0.0033	0.2674	-0.0133	-0.0186	-0.0451	0.2086	-0.0004	0.0809
0.0237	0.0378	0.0298	0.0782	0.0383	0.0127	-0.0938	-0.1146
0.0348	0.0302	0.0595	-0.0005	0.0375	0.0542	-0.0108	0.0040
0.0514	0.0783	0.1109	-0.0095	0.0445	0.0717	0.1427	0.0194
0.0513	-0.0417	0.0628	0.0081	0.0675	0.0327	0.0634	0.0033
-0.0417	0.4516	-0.0261	0.0834	-0.1223	0.1165	-0.0332	0.0942
0.0628	-0.0261	0.1454	-0.0689	0.0773	0.0757	0.0214	-0.0246
0.0081	0.0834	-0.0689	0.1705	0.0051	-0.0289	0.0663	-0.0005
0.0675	-0.1223	0.0773	0.0051	0.1037	0.0226	0.0718	-0.0441
0.0327	0.1165	0.0757	-0.0289	0.0226	0.1265	0.0160	-0.0004
0.0634	-0.0332	0.0214	0.0663	0.0718	0.0160	0.3325	0.0974
0.0033	0.0942	-0.0246	-0.0005	-0.0441	-0.0004	0.0974	0.3318
0.0467	0.0059	0.0254	0.1169	0.0701	0.0075	0.1136	-0.1106
0.0376	-0.0596	-0.0540	0.1580	0.0749	-0.0116	0.0685	-0.1120
0.0041	-0.0462	-0.0929	0.0963	-0.0014	-0.0783	-0.0042	0.1890
0.0456	-0.0241	0.0242	0.0541	0.0670	0.0410	0.0654	-0.0477
0.0458	0.0002	0.1168	-0.0637	0.0416	0.0387	0.0093	0.0811
0.0490	-0.0202	0.1242	-0.0270	0.0551	0.0020	-0.0072	0.0141
0.0231	0.0838	0.0285	0.0500	0.0147	0.0299	0.0251	0.0209
0.0281	0.0928	0.0724	-0.0234	-0.0037	0.0230	0.1025	0.1874
0.0434	-0.0806	0.0032	0.0202	0.0680	0.0440	0.0178	0.0086
0.0476	-0.0555	0.0176	0.0309	0.0499	-0.0130	0.1859	0.1507
-0.0161	0.0822	-0.0196	0.0585	-0.0360	-0.0460	-0.2277	0.0738
0.0404	-0.0207	0.0885	-0.0260	0.0544	0.0425	-0.0898	-0.0354
0.0450	-0.1463	0.0292	-0.0656	0.0611	0.0230	-0.0242	0.1513

0.0240	0.0770	-0.0173	0.0959	0.0169	0.0198	0.1231	0.0458
0.0518	-0.0939	0.0228	0.0389	0.0805	0.0160	0.0629	-0.0122
0.0212	0.0421	0.0376	0.0253	0.0203	0.0212	-0.0751	0.0041

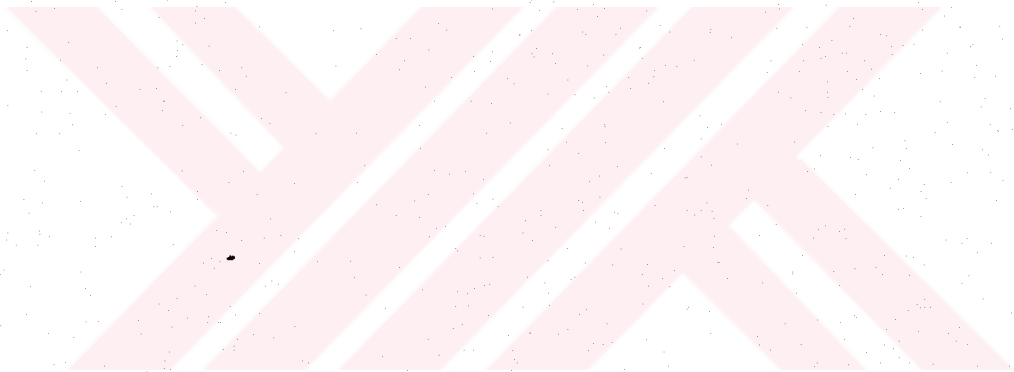
Kolon (17-24)

0.0857	0.0726	0.1132	0.0137	-0.0026	0.0363	0.0533	0.0697
-0.0040	-0.0036	-0.0738	0.0587	0.0850	0.0615	0.0042	-0.0130
0.0905	0.0420	-0.0732	0.0335	0.0255	0.0703	0.0603	0.0002
0.0328	0.0540	0.0622	0.0229	0.0013	0.0143	0.0440	0.0029
-0.0722	0.0222	-0.0235	0.0616	-0.0545	-0.1843	0.0131	-0.0429
0.0885	0.1016	0.0086	0.0469	0.0041	0.0607	0.0476	-0.0545
0.0171	0.0113	0.0002	0.0305	0.0483	0.0458	0.0293	0.0222
0.0760	-0.0518	-0.1210	0.0192	0.0920	0.1036	0.0493	0.1279
0.0467	0.0376	0.0041	0.0456	0.0458	0.0490	0.0231	0.0281
0.0059	-0.0596	-0.0462	-0.0241	0.0002	-0.0202	0.0838	0.0928
0.0254	-0.0540	-0.0929	0.0242	0.1168	0.1242	0.0285	0.0724
0.1169	0.1580	0.0963	0.0541	-0.0637	-0.0270	0.0500	-0.0234
0.0701	0.0749	-0.0014	0.0670	0.0416	0.0551	0.0147	-0.0037
0.0075	-0.0116	-0.0783	0.0410	0.0387	0.0020	0.0299	0.0230
0.1136	0.0685	-0.0042	0.0654	0.0093	-0.0072	0.0251	0.1025
-0.1106	-0.1120	0.1890	-0.0477	0.0811	0.0141	0.0209	0.1874
0.1732	0.1491	-0.0441	0.0808	-0.0231	0.0336	0.0489	-0.0231
0.1491	0.2635	0.0808	0.1193	-0.1067	-0.0779	0.0232	-0.1369
-0.0441	0.0808	0.2984	0.0086	-0.0119	-0.0291	0.0104	0.0168
0.0808	0.1193	0.0086	0.0788	-0.0167	-0.0154	0.0222	-0.0413
-0.0231	-0.1067	-0.0119	-0.0167	0.1402	0.1426	0.0316	0.1318
0.0336	-0.0779	-0.0291	-0.0154	0.1426	0.1938	0.0463	0.1229
0.0489	0.0232	0.0104	0.0222	0.0316	0.0463	0.0446	0.0446
-0.0231	-0.1369	0.0168	-0.0413	0.1318	0.1229	0.0446	0.2064
0.0054	0.1208	0.0951	0.0805	-0.0241	-0.0635	-0.0036	-0.0725
0.0250	0.0128	0.0921	0.0240	0.0492	0.0313	0.0179	0.1084
-0.0528	-0.0109	0.1903	-0.0382	0.0563	0.0949	0.0458	0.0229
0.0046	-0.0010	0.0033	0.0233	0.0766	0.0928	0.0270	0.0065
-0.1127	0.0003	0.1831	0.0268	0.0595	-0.0109	-0.0226	0.0112
0.0793	0.0836	0.0386	0.0473	-0.0188	-0.0184	0.0394	0.0305
0.0544	0.1132	0.0613	0.0733	-0.0024	-0.0065	0.0102	-0.0345
0.0117	0.0166	0.0472	0.0146	0.0467	0.0652	0.0367	0.0147

Kolon (25-32)

-0.0303	0.0813	0.0746	-0.0305	-0.0558	0.0878	0.0126	0.0220
0.0961	-0.0136	-0.0286	0.1148	0.1249	-0.0401	0.0671	0.0425
-0.0257	-0.0608	0.0707	0.0656	-0.1002	0.0246	-0.0036	0.0645
0.0204	-0.0145	0.1159	0.0359	-0.0007	0.0347	0.0127	0.0575
0.1433	-0.0382	-0.0854	-0.0115	0.1069	0.0511	0.0222	-0.0049
0.0210	-0.0650	0.1373	0.0871	-0.0380	0.0099	0.0322	0.0841
0.0364	0.0032	0.0416	0.0600	0.0433	0.0082	0.0285	0.0441
-0.0571	0.0641	-0.1013	0.0115	-0.0811	0.0476	-0.0062	-0.0034
0.0434	0.0476	-0.0161	0.0404	0.0450	0.0240	0.0518	0.0212
-0.0806	-0.0555	0.0822	-0.0207	-0.1463	0.0770	-0.0939	0.0421
0.0032	0.0176	-0.0196	0.0885	0.0292	-0.0173	0.0228	0.0376
0.0202	0.0309	0.0585	-0.0260	-0.0656	0.0959	0.0389	0.0253
0.0680	0.0499	-0.0360	0.0544	0.0611	0.0169	0.0805	0.0203
0.0440	-0.0130	-0.0460	0.0425	0.0230	0.0198	0.0160	0.0212
0.0178	0.1859	-0.2277	-0.0898	-0.0242	0.1231	0.0629	-0.0751
0.0086	0.1507	0.0738	-0.0354	0.1513	0.0458	-0.0122	0.0041
0.0054	0.0250	-0.0528	0.0046	-0.1127	0.0793	0.0544	0.0117
0.1208	0.0128	-0.0109	-0.0010	0.0003	0.0836	0.1132	0.0166
0.0951	0.0921	0.1903	0.0033	0.1831	0.0386	0.0613	0.0472
0.0805	0.0240	-0.0382	0.0233	0.0268	0.0473	0.0733	0.0146
-0.0241	0.0492	0.0563	0.0766	0.0595	-0.0188	-0.0024	0.0467
-0.0635	0.0313	0.0949	0.0928	-0.0109	-0.0184	-0.0065	0.0652
-0.0036	0.0179	0.0458	0.0270	-0.0226	0.0394	0.0102	0.0367
-0.0725	0.1084	0.0229	0.0065	0.0112	0.0305	-0.0345	0.0147
0.1571	0.0292	0.0011	0.0384	0.1621	0.0218	0.1015	0.0218
0.0292	0.1533	-0.0547	-0.0297	0.0757	0.0647	0.0496	-0.0195
0.0011	-0.0547	0.3673	0.1122	0.0826	-0.0337	-0.0141	0.1369
0.0384	-0.0297	0.1122	0.1119	0.0694	-0.0313	0.0317	0.0774
0.1621	0.0757	0.0826	0.0694	0.3043	-0.0296	0.0896	0.0384
0.0218	0.0647	-0.0337	-0.0313	-0.0296	0.0866	0.0336	-0.0023

0.1015 0.0496 -0.0141 0.0317 0.0896 0.0336 0.0892 0.0175
0.0218 -0.0195 0.1369 0.0774 0.0384 -0.0023 0.0175 0.0732



ÖZGEÇMİŞ

17-02-1970 yılında Ankara’da doğmuş olan Ufuk Ekiz Gazi Üniversitesi Fen-Edebiyat.Fakültesi İstatistik bölümünden 1993 yılında mezun olmuştur. 1994 yılında mezun olduğu bölümde araştırma görevlisi olarak çalışmaya başlamış ve 1997 yılında Gazi Üniversitesi Fen Bilimleri Enstitüsü İstatistik Bölümü İstatistik Teorisi Ana Bilim Dalında 1993’te hak kazanmış olduğu Yüksek Lisansını tamamlamıştır. Aynı yıl Gazi Üniversitesi Fen Bilimleri Enstitüsü İstatistik Bölümü İstatistik Teorisi Ana Bilim Dalında doktora yapmaya hak kazanmıştır.

