

**ANKARA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

DOKTORA TEZİ

**GENELLEŞTİRİLMİŞ HİPERBOLİK DAĞILIMLARA DAYALI
DİSKRİMİNANT ANALİZİ**

Abdullah YILMAZ

İSTATİSTİK ANABİLİM DALI

**ANKARA
2013**

Her Hakkı Saklıdır

TEZ ONAYI

Abdullah YILMAZ tarafından hazırlanan “**GENELLEŐTİRİLMİŐ HİPERBOLİK DAĞILIMLARA DAYALI DİSKRİMİNANT ANALİZİ**” adlı tez çalışması 10.05.2013 tarihinde aőağıdaki jüri tarafından **oy birliğı** ile Ankara Üniversitesi Fen Bilimleri Enstitüsü İstatistik Anabilim Dalı’nda **DOKTORA TEZİ** olarak kabul edilmiştir.

Danışman: Prof. Dr. Olcay ARSLAN

Juri Üyeleri :

Başkan: Prof. Dr. Olcay ARSLAN

Ankara Üniversitesi İstatistik Anabilim Dalı

Üye: Doç. Dr. Cemal ATAKAN

Ankara Üniversitesi İstatistik Anabilim Dalı

Üye: Doç. Dr. Meral ÇETİN

Hacettepe Üniversitesi İstatistik Anabilim Dalı

Üye: Prof. Dr. Birdal ŞENOĞLU

Ankara Üniversitesi İstatistik Anabilim Dalı

Üye: Doç. Dr. Serpil Aktaş ALTUNAY

Hacettepe Üniversitesi İstatistik Anabilim Dalı

Yukarıdaki sonucu onaylarım.

Prof. Dr. İbrahim DEMİR

Enstitü Müdürü

ÖZET

Doktora Tezi

GENELLEŞTİRİLMİŞ HİPERBOLİK DAĞILIMLARA DAYALI DİSKRİMİNANT ANALİZİ

Abdullah YILMAZ

Ankara Üniversitesi
Fen Bilimleri Enstitüsü
İstatistik Anabilim Dalı

Danışman : Prof.Dr. Olcay ARSLAN

Bu tez çalışmasında Çok Değişkenli Genelleştirilmiş Hiperbolik Dağılım (GHD) ailesinin bazı özel durumlarının diskriminant analizinde kullanılması ele alınmıştır. GHD ailesinden bazı dağılımlar için diskriminant fonksiyonları elde edilmiş ve bu fonksiyonların sınıflandırmadaki başarısı örneklerle gösterilmiştir.

Çalışma sekiz bölümden oluşmaktadır. Birinci bölüm giriş kısmına ayrılmıştır. Bu bölümde diskriminant analizinde kullanılan klasik yöntemler kısaca özetlenmiştir. İkinci bölümde GHD ile ilgili genel tanımlamalar ile bu dağılımın parametrelerinin tahminine ilişkin bilgiler verilmiştir. Üçüncü, dördüncü, beşinci ve altıncı bölümlerde sırasıyla Çok Değişkenli Genelleştirilmiş Hiperbolik Çarpık-t Dağılımı, Varyans Gamma Dağılımı, Çarpık Laplace Dağılımı ve Normal Ters Gauss Dağılımı ele alınmıştır. Bu dağılımlar için ilgili bölümlerde diskriminant fonksiyonları elde edilmiştir. Yedinci bölümde, önceki bölümlerde elde edilen diskriminant fonksiyonlarının karşılaştırılması değişik senaryolar altında ele alınmış, son bölümde ise elde edilen bulgular özetlenmiştir.

Mayıs 2013, 96 sayfa

Anahtar Kelimeler: Genelleştirilmiş Hiperbolik Dağılımlar, Diskriminant Analizi, Normal Ortalama Varyans Karma Dağılımlar, Çarpık-t Dağılımı, Varyans Gamma Dağılımı, Çarpık Laplace Dağılımı, Ters Gauss Dağılımı, EM algoritması

ABSTRACT

Ph. D. Thesis

DISCRIMINANT ANALYSIS BASED ON GENERALIZED HYPERBOLIC DISTRIBUTIONS

Abdullah YILMAZ

Ankara University
Graduate School of Natural and Applied Sciences
Department of Statistics

Supervisor : Prof.Dr. Olcay ARSLAN

In this thesis the family of Generalized Hyperbolic Distributions (GHD) are used in discriminant analysis. For some special cases of this family, the discriminant functions are derived.

Thesis includes eight chapters. First chapter reserved for introductions. In this chapter, conventional methods of discriminant analysis are briefly summarized. In the second section, general definitions of GHD and its parameter estimation are studied. Generalized hyperbolic multivariate Skew-t Distribution, Variance Gamma Distribution, multivariate Skew Laplace Distribution and multivariate Inverse Gaussian Distribution are studied in chapter three, four, five and six, respectively. Discriminant functions based on these distributions are obtained from the relevant sections. In chapter seven, comparison of discriminant functions obtained in the previous chapters are discussed under different scenarios. The results obtained are summarized in the last section.

May 2013, 96 pages

Key Words: Generalized Hyperbolic Distributions, Discriminant Analysis, Normal Mean-Variance Mixture, Skew-t Distribution, Variance Gamma Distribution, Skew Laplace Distribution, Inverse Gaussian Distribution, EM Algorithm

TEŞEKKÜR

Bu tez çalışması konusunda çalışma imkanımı sağlayan ve çalışmalarım süresince yakın ilgi ve desteğini esirgemeyen Danışman Hocam Sayın Prof. Dr. Olcay ARSLAN (Ankara Üniversitesi İstatistik Anabilim Dalı)'a en derin saygılarımı ve teşekkürlerimi sunmayı bir borç bilirim.

Ayrıca maddi ve manevi olarak beni hiçbir zaman yalnız bırakmayan aileme ve çalışma arkadaşlarıma da saygı ve teşekkürlerimi sunarım.

Abdullah YILMAZ

Ankara, Mayıs 2013

İÇİNDEKİLER

ÖZET.....	i
ABSTRACT	ii
TEŞEKKÜR	iii
ŞEKİLLER DİZİNİ	vi
ÇİZELGELER DİZİNİ	viii
1. GİRİŞ	1
1.1 Sınıflandırma Problemi	2
1.2 Fisher Yaklaşımı	3
1.3 Normal Dağılımlı Kitleler İçin Sınıflandırma	5
2. GENELLEŞTİRİLMİŞ ÇOK DEĞİŞKENLİ HİPERBOLİK DAĞILIMLAR ...	9
2.1 Tanımlamalar	9
2.1.1 Digamma Fonksiyonu	9
2.1.2 Üçüncü Türden Düzeltilmiş Bessel Fonksiyonu	10
2.1.3 Çok Değişkenli Normal Dağılım	11
2.1.4 Ters Gamma Dağılımı	12
2.1.5 Genelleştirilmiş Ters Gauss Dağılımı	12
2.1.6 Ters Gauss Dağılımı.....	14
2.1.7 Çok Değişkenli Normal Ortalama-Varyans Karma Ailesi	14
2.1.8 Çok Değişkenli Genelleştirilmiş Hiperbolik Dağılım Ailesi	15
2.2 Çok Değişkenli Genelleştirilmiş Hiperbolik Dağılım Ailesinin Bazı Özel Durumları	20
2.3 Parametrelerin Tahmini ve EM-Algoritması	20
3. ÇOK DEĞİŞKENLİ GENELLEŞTİRİLMİŞ HİPERBOLİK ÇARPIK-t DAĞILIMI.....	29
3.1 Çarpık-t Dağılımından Rasgele Sayı Üretilmesi	33
3.2 Çarpık-t Dağılımının Parametrelerinin Tahmini İçin EM-Algoritması.....	33
3.3 Çok Değişkenli Genelleştirilmiş Hiperbolik Çarpık-t Dağılımına Sahip Kitlelerde Diskriminant Fonksiyonu.....	38
3.4 Çarpık-t Dağılımı İçin Simülasyon.....	40
4. VARYANS GAMMA DAĞILIMI.....	42

4.1 Varyans Gamma Dağılımından Rasgele Sayı Üretilmesi.....	45
4.2 Varyans Gamma Dağ. Parametrelerinin Tahmini İçin EM-Algoritması.....	46
4.3 Varyans Gamma Dağılımına Sahip Kitlelerde Diskriminant Fonksiyonu.....	50
4.4 Varyans Gamma Dağılımı İçin Simülasyon	52
5. ÇARPIK LAPLACE DAĞILIMI.....	55
5.1 Çarpık Laplace Dağılımından Rasgele Sayı Üretilmesi	58
5.2 Çarpık Laplace Dağılımının Parametrelerinin Tahmini İçin EM-Algoritması.....	58
5.3 Çarpık Laplace Dağılımına Sahip Kitlelerde Diskriminant Fonksiyonu	62
5.4 Çarpık Laplace Dağılımı İçin Simülasyon Sonuçları	63
6. ÇOK DEĞİŞKENLİ NORMAL TERS GAUSS (MNIG) DAĞILIMI	65
6.1 MNIG Dağılımından Rasgele Sayıların Üretilmesi.....	68
6.2 MNIG Dağılımının Parametrelerinin Tahmini İçin EM-Algoritması	69
6.3 MNIG Dağılımına Sahip Kitlelerde Diskriminant Fonksiyonu.....	74
6.4 MNIG Dağılımı İçin Simülasyon Sonuçları.....	75
7. UYGULAMA.....	78
7.1 Simülasyon Yoluyla Üretilen Verilerin Sınıflandırılması	78
7.2 Iris Çiçeği Verileri Üzerine Uygulama	83
7.3 Sporcu Verileri Üzerine Uygulama	87
8. SONUÇ.....	90
KAYNAKLAR	93
ÖZGEÇMİŞ.....	96

ŞEKİLLER DİZİNİ

Şekil 3.1 İki değişkenli genelleştirilmiş çarpık-t dağılımının μ parametresine göre değişimi.....	31
Şekil 3.2 İki değişkenli genelleştirilmiş çarpık-t dağılımının Σ parametresine göre değişimi.....	32
Şekil 3.3 İki değişkenli genelleştirilmiş çarpık-t dağılımının γ parametresine göre değişimi.....	32
Şekil 3.4 Tek değişkenli genelleştirilmiş çarpık-t dağılımının ν parametresine göre değişimi.....	32
Şekil 3.5 n=10 örnek ile çarpık-t dağılımının parametrelerinin EM tahmini.....	36
Şekil 3.6 n=20 örnek ile çarpık-t dağılımının parametrelerinin EM tahmini.....	36
Şekil 3.7 n=30 örnek ile çarpık-t dağılımının parametrelerinin EM tahmini.....	36
Şekil 3.8 n=50 örnek ile çarpık-t dağılımının parametrelerinin EM tahmini.....	37
Şekil 3.9 n=100 örnek ile çarpık-t dağılımının parametrelerinin EM tahmini.....	37
Şekil 4.1 Varyans gamma dağılımının μ parametresine göre değişimi.....	44
Şekil 4.2 Varyans gamma dağılımının Σ parametresine göre değişimi.....	44
Şekil 4.3 Varyans gamma dağılımının γ parametresine göre değişimi.....	44
Şekil 4.4 Varyans gamma Dağılımının λ ve ψ parametresine göre değişimi.....	45
Şekil 4.5 n=10 örnek ile varyans gamma dağılımının parametrelerinin EM tahmini.....	48
Şekil 4.6 n=20 örnek ile varyans gamma dağılımının parametrelerinin EM tahmini.....	49
Şekil 4.7 n=30 örnek ile varyans gamma dağılımının parametrelerinin EM tahmini.....	49
Şekil 4.8 n=50 örnek ile varyans gamma dağılımının parametrelerinin EM tahmini.....	49
Şekil 4.9 n=100 örnek ile varyans gamma dağılımının parametrelerinin EM tahmini.....	50
Şekil 5.1 Çarpık Laplace dağılımının μ parametresine göre değişimi.....	57
Şekil 5.2 Çarpık Laplace dağılımının Σ parametresine göre değişimi.....	57
Şekil 5.3 Çarpık Laplace dağılımının γ parametresine göre değişimi.....	57
Şekil 5.4 n=10 örnek ile çarpık Laplace dağılımının parametrelerinin EM tahmini.....	60
Şekil 5.5 n=20 örnek ile çarpık Laplace dağılımının parametrelerinin EM tahmini.....	60
Şekil 5.6 n=30 örnek ile çarpık Laplace dağılımının parametrelerinin EM tahmini.....	60

Şekil 5.7 n=50 örnek ile çarpık Laplace dağılımının parametrelerinin EM tahmini.	61
Şekil 5.8 n=100 örnek ile çarpık Laplace dağılımının parametrelerinin EM tahmini. ...	61
Şekil 6.1 MNIG dağılımının μ parametersine göre değişimi.	67
Şekil 6.2 MNIG dağılımının Σ parametersine göre değişimi.	67
Şekil 6.3 MNIG dağılımının γ parametresine göre değişimi.	67
Şekil 6.4 MNIG dağılımının χ ve ψ parametresine göre değişimi.	68
Şekil 6.5 n=10 örnek ile MNIG dağılımının parametrelerinin EM tahmini.	71
Şekil 6.6 n=20 örnek ile MNIG dağılımının parametrelerinin EM tahmini.	72
Şekil 6.7 n=30 örnek ile MNIG dağılımının parametrelerinin EM tahmini.	72
Şekil 6.8 n=50 örnek ile MNIG dağılımının parametrelerinin EM tahmini.	72
Şekil 6.9 n=100 örnek ile MNIG dağılımının parametrelerinin EM tahmini.	73
Şekil 7.1 Örnek 1 için diskriminant fonksiyonları.	79
Şekil 7.2 Örnek 2 için diskriminant fonksiyonları.	80
Şekil 7.3 Örnek 3 için diskriminant fonksiyonları.	81
Şekil 7.4 Örnek 5 için diskriminant fonksiyonları.	82
Şekil 7.5 Örnek 6 için diskriminant fonksiyonları.	82
Şekil 7.6 Iris çiçeği için diskriminant fonksiyonları.	85
Şekil 7.7 Iris çiçeği için tahmini parametrelerle çizilmiş Varyans Gamma dağılımlarının olasılık yoğunluk fonksiyonları.	86
Şekil 7.8 Sporcu verileri için diskriminant fonksiyonları.	88
Şekil 7.9 Sporcu verileri için tahmini parametrelerle çizilmiş Varyans Gamma dağılımlarının olasılık yoğunluk fonksiyonları.	89
Şekil 8.1 EM algoritmasında süreler (100 yineleme).	90
Şekil 8.2 Em algoritmasında süreler (yakınsama koşulu altında).	91

ÇİZELGELER DİZİNİ

Çizelge 3.1 Çarpık-t dağılımı için hatalı sınıflandırma oranları.	40
Çizelge 3.2 Çarpık-t dağılımı için $\underline{\mu}_1$ ve $\underline{\mu}_2$ hariç aynı parametreler ile hatalı sınıflandırma oranları.	41
Çizelge 4.1 Varyans gamma dağılımı için hatalı sınıflandırma oranları.	53
Çizelge 4.2 Varyans gamma dağılımı için $\underline{\mu}_1$ ve $\underline{\mu}_2$ hariç aynı parametreler ile hatalı sınıflandırma oranları.	54
Çizelge 5.1 Çarpık Laplace dağılımı için hatalı sınıflandırma oranları.	64
Çizelge 5.2 Çarpık Laplace dağılımı için $\underline{\mu}_1$ ve $\underline{\mu}_2$ hariç aynı parametreler ile hatalı sınıflandırma oranları.	64
Çizelge 6.1 MNIG dağılımı için diskriminant analizi hatalı sınıflandırma oranları.	76
Çizelge 6.2 MNIG dağılımı için $\underline{\mu}_1$ ve $\underline{\mu}_2$ hariç aynı parametreler ile diskriminant analizi hatalı sınıflandırma oranları.	77
Çizelge 7.1 Örnek 1'e ait hatalı sınıflandırma oranları.	78
Çizelge 7.2 Örnek 2'ye ait hatalı sınıflandırma oranları.	79
Çizelge 7.3 Örnek 3'e ait hatalı sınıflandırma oranları.	80
Çizelge 7.4 Örnek 4'e ait hatalı sınıflandırma oranları.	81
Çizelge 7.5 Örnek 5'e ait hatalı sınıflandırma oranları.	81
Çizelge 7.6 Iris çiçeği için hatalı sınıflandırma oranları (4 parametre ve her iki gruptan 50 birimlik örneklem ile).	84
Çizelge 7.7 Iris çiçeği için hatalı sınıflandırma oranları (2 parametre ve her iki gruptan 10 birimlik örneklem ile).	84
Çizelge 7.8 Iris çiçeği için tahmini parametre değerleri (2 parametre ve her iki gruptan 10 birimlik örneklem ile).	85
Çizelge 7.9 Iris çiçeği için hatalı sınıflandırma oranları (2 parametre ve 80 birimlik örneklem ile)	86
Çizelge 7.10 Sporcu verileri için hatalı sınıflandırma oranları.	87
Çizelge 7.11 Sporcu verileri için tahmini parametre değerleri (2 parametre ve her iki gruptan 15 birimlik örneklem ile).	88
Çizelge 8.1 Örnek 1 için sınıflandırma süreleri.	92

1. GİRİŞ

Diskriminant analizi, birey üzerinden ölçümler alındığında bu ölçümlere bağlı olarak bireyi, daha önceden bilinen sonlu sayıdaki ayırık gruplardan (sınıflardan, kitlelerden) birisine atayan istatistiksel bir yöntemdir. Temel olarak, incelenen bireyin sınıfının belirlenmesinde kullanılacak diskriminant fonksiyonlarının bulunması ve bu fonksiyonları kullanarak oluşturulacak atama kuralına uygun biçimde söz konusu bireyin atanacağı grubun belirlenmesi amaçlanır.

Diskriminant analizi ile ilgili ilk çalışma, lineer diskriminant fonksiyonunun (LDF) literatüre katılmasını sağlayan Fisher (1936)'e aittir. Bu çalışmada Fisher'in dağılımdan bağımsız olarak geliştirdiği yöntem, gözlemlerin alındığı sınıfların eşit varyanslı normal dağılım gösteren kitleler olması durumu ile örtüşmektedir. Fisher'in çalışmasından sonra sınıflandırma problemi birçok araştırmacı tarafından çeşitli değişikliklerle tekrar ele alınmıştır. Geisser (1964) , Logan ve Gupta (1993)'nın Bayesçi metotları, Anderson (1984), Gupta (1986) ile Lee ve Chen(2009)'in olabilirlik oran testi kriterini, Freed ve Glover (1981)'in lineer programlama yaklaşımını kullanmaları bu çalışmalara örnek olarak verilebilir.

Verilerin Normal dağıldığı varsayımı altında sınıflandırma problemini ele alan bir çok çalışma yapılmıştır. John (1961) ve O'Neill (1978)'in çalışmaları bunlara örnek olarak verilebilir. Buna karşın değişik dağılım varsayımları ile sınıflandırma konusunu ele alan çalışma sayısı nispeten daha azdır. Sutradhar (1990) lineer diskriminant fonksiyonunun çok değişkenli t-dağılımı varsayımı altındaki davranışlarını incelemiş, Thabane ve Drekić (2004) genelleştirilmiş çok değişkenli Bessel dağılımı gösteren iki kitle için sınıflandırma problemini ele almışlardır. Benzer çalışmalar Gompertz dağılımı (Moustafa ve Ramadan 2005), budanmış (truncated) t-dağılımı (Batsidis 2010), ters Gauss dağılımı (Amoha ve Kocherlakota 1986) ve gamma dağılımı (Mahmoud ve Moustafa 1995) için yapılmıştır. Bu çalışmada ise '*çok değişkenli genelleştirilmiş hiperbolik dağılım*' varsayımı altında sınıflandırma problemi ele alınmıştır.

Çalışmanın birinci bölümünün devamında diskriminant analizinde kullanılan klasik yöntemler kısaca özetlenmiştir. İkinci bölümde çok değişkenli genelleştirilmiş hiperbolik dağılımlara ilişkin genel tanımlamalar ile bu dağılımın parametrelerinin tahmini ile ilgili bilgiler verilmiştir. Üçüncü, dördüncü, beşinci ve altıncı bölümlerde sırasıyla çok değişkenli genelleştirilmiş hiperbolik çarpık-t dağılımı, varyans gamma dağılımı, çarpık Laplace dağılımı ve normal ters Gauss Dağılımı tanıtılmıştır. İlgili bölümlerde parametrelerin tahmini için EM algoritmaları ve diskriminant fonksiyonları elde edilmiştir. Yedinci bölümde, önceki bölümlerde elde edilen diskriminant fonksiyonlarının karşılaştırılması değişik senaryolar altında ele alınmıştır. Sekizinci bölümde ise çalışma kapsamındaki bulgular özetlenmiştir.

1.1 Sınıflandırma Problemi

Diskriminant analizinde her bir kitle ölçümlere karşılık gelen rasgele vektörün olasılık dağılımı ile karakterize edilir. $\Pi_1, \Pi_2, \dots, \Pi_l$ l tane farklı kitleyi ve $\underline{X} \in \mathbb{R}^d$ bireyler üzerinden alınan ölçümleri gösteren rasgele vektör olmak üzere, $\underline{X}$ 'in olasılık yoğunluk fonksiyonunu $f_{\underline{X}_i}(\underline{x}; \underline{\theta}_i)$ ile göstereceğiz. Bu olasılık yoğunluk fonksiyonları biçimsel olarak bilindiğinde diskriminant analizi, $\underline{\theta}_i$ parametresinin tahmin edilmesi ve atama kuralının oluşturulması ile gerçekleştirilir ($i = 1, 2, \dots, l$).

l tane farklı kitlenin birleştirilmesiyle oluşturulmuş gruptan rasgele alınan bir bireyin Π_i kitlesine ait olması olasılığını q_i ile gösterelim. Buna göre $q_i = P(\underline{X} \in \Pi_i)$ olasılığı ‘önsel(prior) olasılık’ olarak adlandırılır. Dolayısıyla tanım gereği

$$\sum_{i=1}^l q_i = 1$$

olacağı açıktır. q_i önsel olasılıklarının seçiminde genelde iki yaklaşım kullanılmaktadır. Bunlardan ilki l tane kitlenin tümü için q_i 'leri eşit kabul ederek $\hat{q}_i = 1/l$ biçiminde seçilmesi, diğeri ise kitlelerden alınan örneklem hacimleri ile her bir kitle için $\hat{q}_i = n_i/n$

biçiminde seçilmesidir. Burada n_i , i . kitleden alınan örneklem sayısını, n ise toplam örnek sayısını göstermektedir ($i = 1, 2, \dots, l$) (Hubert ve Drisen 2004).

Atama işleminde, gözlemin ait olduğu asıl kitleden farklı bir kitleye atanması durumu ‘*hatalı sınıflandırma*’ olarak adlandırılır. Sınıflandırma probleminin, $\underline{X} \in \mathbb{R}^d$ vektörünün aldığı değerlere göre, örneklem uzayını l ayrık bölgeye parçaladığı göz önüne alınırsa, gerçekte j .kitleye ait olan bireyin yanlışlıkla i .kitleye atanması olasılığı

$$p_{i/j} = \int_{B_i} f_j(\underline{x}; \underline{\theta}_j) d\underline{x} \quad (1.1)$$

dır. Burada $\bigcup_{i=1}^l B_i = \mathbb{R}^d$, $B_i \cap B_j = \emptyset$ ve $i \neq j$, $i, j = 1, 2, \dots, l$ dir.

Her hatalı sınıflandırma beraberinde bir maliyet getirecektir. Bu maliyete ‘*hatalı sınıflandırma maliyeti*’ denir.

1.2 Fisher Yaklaşımı

Fisher’in dağılımdan bağımsız olarak geliştirdiği yönteminde $\underline{X} = [X_1, X_2, \dots, X_d]'$ birimler üzerinden gözlenebilen d boyutlu rastgele vektör olmak üzere, $\underline{x}$ gözlemleri

$$Y = \underline{a}' \underline{X} = a_1 X_1 + a_2 X_2 + \dots + a_d X_d \quad (1.2)$$

dönüşümü ile tek değişkenli y gözlemlerine dönüştürülür. Burada $\underline{a}$ vektörü mümkün olan en iyi sınıflandırmayı gerçekleştirecek şekilde seçilmelidir. Fisher’in bu yaklaşımında çok değişkenli gözlem verileri, sınıflandırmada kullanılacak tek değişkenli bir skor değerine çevrilir.

Π_1 ve Π_2 sırasıyla $\underline{\mu}_1$ ve $\underline{\mu}_2$ ortalamalı, ortak Σ varyans-kovaryans matrisine sahip kitleler olsun. Bu durumda $\underline{\mu}_{1Y}$ ve $\underline{\mu}_{2Y}$ sırasıyla Π_1 ve Π_2 kitlelerinde olan $\underline{X}$ rasgele değişkeni ile elde edilen Y rasgele değişkeninin koşullu beklenen değerleri olmak üzere

$$\begin{aligned}\underline{\mu}_{1Y} &= E(Y|\Pi_1) = E(\underline{a}'\underline{X}|\Pi_1) = \underline{a}'\underline{\mu}_1 \\ \underline{\mu}_{2Y} &= E(Y|\Pi_2) = E(\underline{a}'\underline{X}|\Pi_2) = \underline{a}'\underline{\mu}_2\end{aligned}$$

olacaktır. Benzer şekilde Y 'nin varyansı her iki grup için de aynı olup

$$\sigma_Y^2 = Var(\underline{a}'\underline{X}) = \underline{a}'\underline{\Sigma}\underline{a}$$

olarak elde edilecektir. Fisher'in yaklaşımında $\underline{a}$ vektörü, $\underline{\mu}_{1Y}$ ve $\underline{\mu}_{2Y}$ beklenen değerleri arasındaki karesel farkın Y 'nin varyansına oranının en büyük olmasını sağlayacak biçimde seçilir. Bu oran

$$j(\underline{a}) = \frac{(\underline{a}'\underline{\mu}_1 - \underline{a}'\underline{\mu}_2)^2}{\underline{a}'\underline{\Sigma}\underline{a}} \quad (1.3)$$

olarak tanımlanır. Buradan

$$\max_{\underline{a}} j(\underline{a}) \quad , \underline{a} \neq \underline{0} \quad (1.4)$$

maksimum probleminin çözülmesiyle $\underline{a} = \underline{\Sigma}^{-1}(\underline{\mu}_1 - \underline{\mu}_2)$ olarak elde edilir (Anderson 1984). $\underline{x} \in \mathbb{R}^d$ gözlemi için diskriminant skoru ise $y = \underline{a}'\underline{x} = \underline{x}'\underline{a} = \underline{x}'\underline{\Sigma}^{-1}(\underline{\mu}_1 - \underline{\mu}_2)$ olarak elde edilir. Buradan '*Fisher Lineer Diskriminant Fonksiyonu*'

$$d^{LD}(\underline{x}) = \underline{x}'\underline{\Sigma}^{-1}(\underline{\mu}_1 - \underline{\mu}_2) \quad (1.5)$$

biçiminde elde oluşturulur. Atama kuralı ise dönüşüm uygulandıktan sonra grup ortalamaları arasındaki uzaklığın orta noktasının $1/2(\underline{a}'\underline{\mu}_1 + \underline{a}'\underline{\mu}_2) = 1/2\Delta^2$ olduğu göz önüne alınarak

$$\begin{cases} \text{Eğer } d^{LD}(\underline{x}) \geq \frac{1}{2} \Delta^2 \text{ ise } \underline{x} \in \Pi_1 \\ \text{Eğer } d^{LD}(\underline{x}) < \frac{1}{2} \Delta^2 \text{ ise } \underline{x} \in \Pi_2 \end{cases} \quad (1.6)$$

biçiminde oluşturulur. Dikkat edilecek olursa $\Delta^2 = (\underline{a}'\underline{\mu}_1 + \underline{a}'\underline{\mu}_2) = \underline{a}'(\underline{\mu}_1 + \underline{\mu}_2)$ değerinde $\underline{a} = \Sigma^{-1}(\underline{\mu}_1 - \underline{\mu}_2)$ yerine yazıldığında $\Delta^2 = (\underline{\mu}_1 - \underline{\mu}_2)' \Sigma^{-1}(\underline{\mu}_1 + \underline{\mu}_2)$ olarak elde edilir. Böylece (1.6) ile verilen atama kuralının, eşit varyans-kovaryans matrisine sahip Normal dağılımlı iki kitle için (1.10) ile verilen atama kuralı ile aynı olduğu görülür.

1.3 Normal Dağılımlı Kitleler İçin Sınıflandırma

Bu bölümde iki çok değişkenli normal dağılım için sınıflandırma problemi hakkında kısa bilgiler verilecektir.

Π_1 ve Π_2 sırasıyla $N_d(\underline{\mu}_1, \Sigma_1)$ ve $N_d(\underline{\mu}_2, \Sigma_2)$ dağılımına sahip kitleler olsun. $\underline{X} \in \mathbb{R}^d$ rasgele vektörü için

$$f_{\underline{X}_i}(\underline{x}; \underline{\mu}_i, \Sigma_i) = \frac{1}{(2\pi)^{d/2} |\Sigma_i|^{1/2}} \exp \left\{ -\frac{1}{2} (\underline{x} - \underline{\mu}_i)' \Sigma_i^{-1} (\underline{x} - \underline{\mu}_i) \right\}, i=1,2 \quad (1.7)$$

dir. Burada $\Sigma_i \in \mathbb{R}^{d \times d}$ pozitif tanımlı bir matris ve $\underline{\mu}_i \in \mathbb{R}^d$ dir ($i=1,2$). Bu iki kitlenin birisinden geldiği bilinen bir $\underline{x}_0 \in \mathbb{R}^d$ gözlemi için atama kuralı, $C_{i/j}$ ile gerçekte j . kitleden olan bir bireyin i . kitleye hatalı sınıflandırma maliyetini, $q_i = P(\underline{X} \in \Pi_i)$ i . kitleye ilişkin önsel olasılığı göstermek üzere ($i, j=1,2$ ve $i \neq j$)

$$\begin{cases} \text{Eğer } \frac{f_{\underline{X}_1}(\underline{x}_0)}{f_{\underline{X}_2}(\underline{x}_0)} \geq \frac{q_2}{q_1} \frac{C_{1/2}}{C_{2/1}} \text{ ise, } \underline{x}_0 \in \Pi_1 \\ \text{Eğer } \frac{f_{\underline{X}_1}(\underline{x}_0)}{f_{\underline{X}_2}(\underline{x}_0)} < \frac{q_2}{q_1} \frac{C_{1/2}}{C_{2/1}} \text{ ise, } \underline{x}_0 \in \Pi_2 \end{cases}$$

biçimindedir. Çalışmanın devamında hatalı sınıflandırma maliyetleri her iki kitle için eşit, önsel olasılıklarında iki kitle için aynı olduğu kabul edilecektir. Böylece $C_{1/2} = C_{2/1}$ ve $q_1 = q_2$ alınarak atama kuralı

$$\begin{cases} \text{Eğer } f_{\underline{x}_1}(\underline{x}_0)/f_{\underline{x}_2}(\underline{x}_0) \geq 1 \text{ ise, } \underline{x}_0 \in \Pi_1 \\ \text{Eğer } f_{\underline{x}_1}(\underline{x}_0)/f_{\underline{x}_2}(\underline{x}_0) < 1 \text{ ise, } \underline{x}_0 \in \Pi_2 \end{cases} \quad (1.8)$$

biçiminde elde edilecektir (Welch 1939).

$\Sigma_1 = \Sigma_2 = \Sigma$ olması durumunda, logaritmik fonksiyonun monoton artan olma özelliği kullanılarak

$$\log \{f_{\underline{x}_1}(\underline{x})/f_{\underline{x}_2}(\underline{x})\} = \underline{x}'\Sigma^{-1}(\underline{\mu}_1 - \underline{\mu}_2) - \frac{1}{2}(\underline{\mu}_1 + \underline{\mu}_2)'\Sigma^{-1}(\underline{\mu}_1 - \underline{\mu}_2)$$

biçiminde yazılabilir. Burada $\Delta^2 = (\underline{\mu}_1 - \underline{\mu}_2)'\Sigma^{-1}(\underline{\mu}_1 - \underline{\mu}_2)$ olmak üzere, lineer diskriminant fonksiyonu

$$d^{LD}(\underline{x}) = \underline{x}'\Sigma^{-1}(\underline{\mu}_1 - \underline{\mu}_2) \quad (1.9)$$

olarak, atama kuralı ise

$$\begin{cases} \text{Eğer } d^{LD}(\underline{x}) \geq \frac{1}{2}\Delta^2 \text{ ise } \underline{x} \in \Pi_1 \\ \text{Eğer } d^{LD}(\underline{x}) < \frac{1}{2}\Delta^2 \text{ ise } \underline{x} \in \Pi_2 \end{cases} \quad (1.10)$$

biçiminde elde edilir. $\Sigma_1 \neq \Sigma_2$ olması durumunda

$$\log \left\{ f_{X_1}(\underline{x}) / f_{X_2}(\underline{x}) \right\} = -\frac{1}{2} \log \frac{|\Sigma_1|}{|\Sigma_2|} + \frac{1}{2} \left[(\underline{x} - \underline{\mu}_2)' \Sigma_2^{-1} (\underline{x} - \underline{\mu}_2) - (\underline{x} - \underline{\mu}_1)' \Sigma_1^{-1} (\underline{x} - \underline{\mu}_1) \right]$$

olup, $D_{x,i} = (\underline{x} - \underline{\mu}_i)' \Sigma_i^{-1} (\underline{x} - \underline{\mu}_i)$, $i=1,2$ gösterimi kullanılarak karesel diskriminant fonksiyonu

$$d^{QD}(\underline{x}) = D_{x,2} - D_{x,1} \quad (1.11)$$

alındığında, atama kuralı

$$\begin{cases} \text{Eğer } d^{QD}(\underline{x}) < \log \frac{|\Sigma_1|}{|\Sigma_2|} \text{ ise} & \underline{x} \in \Pi_1 \\ \text{Eğer } d^{QD}(\underline{x}) \geq \log \frac{|\Sigma_1|}{|\Sigma_2|} \text{ ise} & \underline{x} \in \Pi_2 \end{cases} \quad (1.12)$$

biçiminde elde edilir.

$\Sigma_1 = \Sigma_2 = \Sigma$ olması durumunda eğer parametreler bilinmiyorsa (1.9) ve (1.10)'de $\underline{\mu}_1, \underline{\mu}_2$ ve Σ parametreleri yerlerine sırasıyla

$$\begin{aligned} \overline{X}_1 &= \frac{1}{n_1} \sum_{i=1}^{n_1} X_{1i} \\ \overline{X}_2 &= \frac{1}{n_2} \sum_{i=1}^{n_2} X_{2i} \end{aligned} \quad (1.13)$$

$$S = \frac{1}{n_1 + n_2 - 2} \left[\sum_{i=1}^{n_1} (\underline{X}_{1i} - \overline{X}_1)(\underline{X}_{1i} - \overline{X}_1)' + \sum_{i=1}^{n_2} (\underline{X}_{2i} - \overline{X}_2)(\underline{X}_{2i} - \overline{X}_2)' \right]$$

tahmin edicileri kullanılır. $\Sigma_1 \neq \Sigma_2$ olması durumunda parametreler bilinmiyorsa (1.11) ve (1.12)'de $\underline{\mu}_1, \underline{\mu}_2$ ve Σ_1, Σ_2 parametreleri yerlerine sırasıyla

$$\begin{aligned}
\bar{X}_1 &= \frac{1}{n_1} \sum_{i=1}^{n_1} X_{1i} \\
\bar{X}_2 &= \frac{1}{n_2} \sum_{i=1}^{n_2} X_{2i} \\
S_1 &= \frac{1}{n_1 - 1} \left[\sum_{i=1}^{n_1} (X_{1i} - \bar{X}_1)(X_{1i} - \bar{X}_1)' \right] \\
S_2 &= \frac{1}{n_2 - 1} \left[\sum_{i=1}^{n_2} (X_{2i} - \bar{X}_2)(X_{2i} - \bar{X}_2)' \right]
\end{aligned} \tag{1.14}$$

tahmin edicileri kullanılır.

2. GENELLEŞTİRİLMİŞ ÇOK DEĞİŞKENLİ HİPERBOLİK DAĞILIMLAR

Birçok istatistiksel teknik, gözlemlerin alındığı kitlenin ya da kitlelerin Normal dağıldığı varsayımına dayanmaktadır. Fakat gerçek hayatta sıklıkla verilerin normal dağılım göstermemesi problemiyle karşılaşilmektedir. Bu varsayımdaki bozulmanın göz ardı edilmesi, gerçek dünya ile uyuşmayan sonuçlar elde edilmesine yol açmaktadır. Bu çalışmada söz konusu bozulmanın önceki bölümde kısaca anlatılan diskriminant analizi üzerine etkisini gidermeyi amaçlanmıştır. Bu sebeple kitlelerin '*çok değişkenli genelleştirilmiş hiperbolik dağılım*' göstermesi durumunda atama kurallarının oluşturulmasına çalışılacaktır.

Çok değişkenli genelleştirilmiş hiperbolik dağılım ailesi oldukça esnek özelliklere sahiptir. '*Normal ortalama-varyans karma dağılım*' ailesinin özel bir hali olup, Student- t , çarpık- t , varyans-gamma, ters Gauss gibi bilindik bir çok dağılımı kapsamaktadır. Parametrelerin uygun seçimleriyle bu dağılımların basıklık ve çarpıklıklarının farklı seçimleri söz konusu olabilmektedir.

Genelleştirilmiş hiperbolik dağılım ailesine dahil bazı dağılımlar risk analizi ve finansal verilerin modellenmesinde sıklıkla kullanılmaktadır. Bu çalışmada çok değişkenli genelleştirilmiş hiperbolik dağılım ve bu dağılımın parametrelerinin farklı seçimleriyle elde edilecek dağılımları kullanarak diskriminant fonksiyonlarının oluşturulması hedeflenmektedir.

2.1 Tanımlamalar

Bu bölümde ilerleyen kısımlarda kullanılacak bazı temel kavram ve tanımlar verilecektir.

2.1.1 Digamma fonksiyonu

Digamma fonksiyonu gamma fonksiyonunun logaritmik türevi olarak tanımlanır (Abramowitz ve Stegun 1964). Digamma fonksiyonunu

$$\psi(x) = \frac{d}{dx} \log(\Gamma(x)) \quad (2.1)$$

biçiminde gösterilecektir.

2.1.2 Üçüncü türden düzeltilmiş Bessel fonksiyonu (modified Bessel function of third kind)

Üçüncü türden düzeltilmiş Bessel fonksiyonunu (kısaca Bessel fonksiyonu) için integral gösterimi Barndorff-Nielsen(1978) tarafından

$$K_{\lambda}(t) = \frac{1}{2} \int_0^{\infty} y^{\lambda-1} \exp\left\{-\frac{t}{2}(y + y^{-1})\right\} dy, \quad t > 0 \quad (2.2)$$

biçiminde verilir. Bu fonksiyon için

$$K_{\lambda}(x) \sim \Gamma(\lambda) 2^{\lambda-1} x^{-\lambda}, \quad x \downarrow 0 \quad (2.3)$$

$$K_{\lambda}(x) \sim \Gamma(-\lambda) 2^{-\lambda-1} x^{\lambda}, \quad x \downarrow 0 \quad (2.4)$$

asimptotik formülleri (Abramowitz ve Stegun 1964), çalışma içerisinde bazı limit durumlarının incelenmesinde faydalı olacaktır. Ayrıca Bessel fonksiyonu λ parametresine göre simetri özelliği de taşımaktadır. Yani

$$K_{\lambda}(x) = K_{-\lambda}(x) \quad (2.5)$$

dir. Ayrıca $\lambda = 1/2$ özel durumu için Bessel fonksiyonunu değeri

$$K_{1/2}(x) = x^{-1/2} e^{-x} \sqrt{\pi/2} \quad (2.6)$$

eşitliği ile verilir. Bessel fonksiyonu için kullanılacak bir diğer özellik ise, Bessel fonksiyonunun logaritmik türevi ile ilgili

$$\frac{\partial}{\partial x} \log K_{\lambda}(x) = \frac{\lambda}{x} - \frac{K_{\lambda+1}(x)}{K_{\lambda}(x)} \quad (2.7)$$

$$\frac{\partial}{\partial x} \log K_{\lambda}(x) = -\frac{\lambda}{x} - \frac{K_{\lambda-1}(x)}{K_{\lambda}(x)} \quad (2.8)$$

eşitlikleridir (Abramowitz ve Stegun 1964, Blaesild 1981).

2.1.3 Çok değişkenli normal dağılım

$\underline{X}$ rasgele vektörü

$$f_{\underline{X}}(\underline{x}; \underline{\mu}, \underline{\Sigma}) = \frac{1}{(2\pi)^{d/2} |\underline{\Sigma}|^{1/2}} \exp \left\{ -\frac{1}{2} (\underline{x} - \underline{\mu})^t \underline{\Sigma}^{-1} (\underline{x} - \underline{\mu}) \right\}, \underline{x} \in \mathbb{R}^d \quad (2.9)$$

olasılık yoğunluk fonksiyonuna sahipse, d boyutlu normal dağılıma sahiptir denir ve $\underline{X} \sim N_d(\underline{\mu}, \underline{\Sigma})$ ile gösterilir. Burada $\underline{\mu} \in \mathbb{R}^d$ ve $\underline{\Sigma} \in \mathbb{R}^{d \times d}$ pozitif tanımlı singüler olmayan bir matristir. Bu durumda

$$E(\underline{X}) = \underline{\mu}$$

$$Cov(\underline{X}) = \underline{\Sigma}$$

beklenen değer ve varyans-kovaryans matrisidir.

2.1.4 Ters gamma (inverse gamma) dağılımı

X rasgele değişkeni

$$f_X(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{-\alpha-1} \exp\left\{-\frac{\beta}{x}\right\}, x > 0, \alpha > 0, \beta > 0 \quad (2.10)$$

olasılık yoğunluk fonksiyonuna sahipse ‘ters gamma’ dağılımına sahiptir denir ve $X \sim IGamma(\alpha, \beta)$ ile gösterilir. X rasgele değişkeninin beklenen değer ve varyansı ise

$$E(X) = \beta/(\alpha - 1) \quad (2.11)$$

$$Var(X) = \beta^2 / ((\alpha - 1)^2 (\alpha - 2)) \quad , \alpha > 2 \quad (2.12)$$

dir. Ayrıca $Y \sim Gamma(\alpha, \beta)$ dağılımına sahip ise $1/Y \sim IGamma(\alpha, \beta)$ dağılımı gösterir (McNeil vd. 2005).

2.1.5 Genelleştirilmiş ters Gauss (generalized inverse Gaussian) dağılımı

X rasgele değişkeni

$$f_X(x; \lambda, \chi, \psi) = \frac{\chi^{-\lambda} (\sqrt{\chi\psi})^\lambda}{2K_\lambda(\sqrt{\chi\psi})} x^{\lambda-1} \exp\left(-\frac{1}{2}\left(\chi\frac{1}{x} + \psi x\right)\right), x > 0 \quad (2.13)$$

olasılık yoğunluk fonksiyonuna sahipse ‘genelleştirilmiş ters Gauss’ dağılımına sahiptir denir ve $X \sim N^-(\lambda, \chi, \psi)$ ile gösterilir. Burada $K_\lambda(\cdot)$ üçüncü türden λ parametrelili değiştirilmiş Bessel fonksiyonunu göstermektedir. Ayrıca λ, χ, ψ parametreleri

$$\begin{cases} \lambda < 0 \text{ ise } \chi > 0, \psi \geq 0 \\ \lambda = 0 \text{ ise } \chi > 0, \psi > 0 \\ \lambda \geq 0 \text{ ise } \chi \geq 0, \psi > 0 \end{cases}$$

özelliklerini sağlamaktadır.

$X \sim N^-(\lambda, \chi, \psi)$ olmak üzere $E(X^\alpha)$ beklenen değeri

$$E(X^\alpha) = \int_0^\infty x^\alpha f_X(x; \lambda, \chi, \psi) dx = \frac{\chi^{-\lambda} (\sqrt{\chi\psi})^\lambda}{2K_\lambda(\sqrt{\chi\psi})} \int_0^\infty x^{\alpha+\lambda-1} \exp\left(-\frac{1}{2}\left(\chi \frac{1}{x} + \psi x\right)\right) dx$$

eşitliğinde (2.13) olasılık yoğunluk fonksiyonu dikkate alınırsa

$$E(X^\alpha) = \frac{\chi^{-\lambda} (\sqrt{\chi\psi})^\lambda}{2K_\lambda(\sqrt{\chi\psi})} \frac{2K_{\lambda+\alpha}(\sqrt{\chi\psi})}{\chi^{-\lambda-\alpha} (\sqrt{\chi\psi})^{\lambda+\alpha}} = \left(\frac{\chi}{\psi}\right)^{\alpha/2} \frac{K_{\lambda+\alpha}(\sqrt{\chi\psi})}{K_\lambda(\sqrt{\chi\psi})}$$

elde edilir. Buna göre

$$E(X^\alpha) = \left(\frac{\chi}{\psi}\right)^{\alpha/2} \frac{K_{\lambda+\alpha}(\sqrt{\chi\psi})}{K_\lambda(\sqrt{\chi\psi})} \quad (2.14)$$

dır. Genelleştirilmiş ters Gauss dağılımına ilişkin bir diğer kullanışlı özellik ise

$$E(\log X) = \frac{dE(X^\alpha)}{d\alpha} \Big|_{\alpha=0} \quad (2.15)$$

olmasıdır. Yazılan son iki eşitlik $\chi > 0$ ve $\psi > 0$ durumu için geçerlidir. Dikkat edilmesi gereken başka bir nokta ise (2.15) değerinin nümerik yollarla bulunması gerekliliğidir (McNeil vd. 2005).

2.1.6 Ters Gauss (inverse Gaussian) dağılımı

X rasgele değişkeni

$$f(x) = \sqrt{\frac{\lambda}{2\pi}} x^{-\frac{3}{2}} \exp\left\{-\frac{\lambda}{2\mu^2 x} (x - \mu)^2\right\}, x > 0 \quad (2.16)$$

olasılık yoğunluk fonksiyonuna sahip olduğunda ‘ters Gauss’ dağılımına sahiptir denir ve $X \sim IG(\mu, \lambda)$ ile gösterilir. X rasgele değişkeninin beklenen değer ve varyansı

$$E(X) = \mu \quad (2.17)$$

$$Var(X) = \mu^3 / \lambda \quad (2.18)$$

eşitlikleri ile verilir (Chhikara ve Folks 1989).

2.1.7 Çok değişkenli normal ortalama-varyans karma ailesi

$\underline{X}$ rasgele vektörü

$$\underline{X} = \underline{m}(W) + \sqrt{W} \mathbf{A} \underline{Z} \quad (2.19)$$

dönüşümü ile tanımlansın. Eğer

- i. $\underline{Z} \sim N_k(\underline{0}, \mathbf{I}_k)$
- ii. $W \geq 0$ negatif değerler almayan ve $\underline{Z}$ ’den bağımsız bir rasgele değişken

- iii. $\mathbf{A} \in \mathbb{R}^{d \times k}$ bir matris
- iv. $\underline{m}: \mathbb{R}^+ \rightarrow \mathbb{R}^d$ ölçülebilir bir fonksiyon

özellikleri sağlanıyorsa, $\underline{X}$ rasgele vektörüne ‘*çok değişkenli normal ortalama-varyans karma dağılımına*’ sahiptir denir. Ayrıca W rasgele değişkeni *karıştırıcı değişken* (mixing variable) olarak adlandırılır.

Kolayca görüleceği üzere

$$(\underline{X}|W = w) \sim N_d(\underline{m}(w), w\Sigma) \quad (2.20)$$

koşullu dağılımına sahiptir. Burada $\Sigma = \mathbf{A}\mathbf{A}'$ dır.

2.1.8 Çok değişkenli genelleştirilmiş hiperbolik dağılım ailesi

Genelleştirilmiş hiperbolik dağılım ailesi, ilk olarak Barndorff-Nielsen (1977) tarafından önerilmiştir. Barndorff-Nielsen bu çalışmasında esen rüzgardaki kum tanelerinin büyüklüklerinin dağılımını ele almıştır. Devam eden çalışmalarda bu dağılım ailesi fizik, biyoloji, istatistik gibi bir çok bilim alanında kullanılmıştır. Dağılımın kalın kuyruklu ve çarpık yapıda oluşu, finansal verilerin modellenmesinde de kullanışlı sonuçlar vermektedir. Bu alandaki ilk çalışma Eberlein ve Keller tarafından (1995) verilmiş, sonraki tarihlerde genelleştirilmiş hiperbolik dağılım ailesinin finansal veriler üzerine uygulandığı bir çok çalışma yapılmıştır. Bu çalışmalara örnek olarak Prause (1999) McNeil (2005) ile bu kaynakların referansları verilebilir.

Bu dağılım ailesi normal ortalama-varyans karma dağılım ailesinin bir alt sınıfı olup, karıştırıcı değişken dağılımının genelleştirilmiş ters Gauss (GIG) alınmasıyla elde edilir. GIG dağılımının özel parametre seçimleriyle normal, gamma, ters gamma ve ters Gauss dağılımı gibi özel dağılımların elde edilebiliyor olması, genelleştirilmiş hiperbolik

dağılım ailesinde de etkilerini göstermekte; GIG dağılımının özel halleri kullanılarak genelleştirilmiş hiperbolik dağılım ailesinin alt sınıfları elde edilebilmektedir.

Normal ortalama-varyans karma dağılımında W karıştırıcı değişkeninin dağılımının genelleştirilmiş ters gauss, $\underline{m}$ ölçülebilir fonksiyonunun ise

$$\underline{m}(W) = \underline{\mu} + W \underline{\gamma} \quad (2.21)$$

biçiminde seçilmesi durumunda, (2.19) dönüşümü ile verilen $\underline{X}$ rasgele vektörü ‘çok değişkenli genelleştirilmiş hiperbolik dağılıma’ sahiptir.

$\underline{m}$ fonksiyonunun bu şekilde seçimiyle (2.19) dönüşümü, $\underline{\mu} \in \mathbb{R}^d$ ve $\underline{\gamma} \in \mathbb{R}^d$ parametre vektörleri olmak üzere

$$\underline{X} = \underline{\mu} + W \underline{\gamma} + \sqrt{W} \mathbf{A} \underline{Z} \quad (2.22)$$

olarak yazılır. (2.20)’deki koşullu dağılım ise

$$(\underline{X} | W = w) \sim N_d(\underline{\mu} + W \underline{\gamma}, w \mathbf{\Sigma}) \quad (2.23)$$

halini alır. Buradan $\underline{X}$ rasgele vektörünün olasılık yoğunluk fonksiyonu,

$$f_{\underline{X}}(\underline{x}) = \int_0^\infty f_{\underline{X}|W}(\underline{x}|w) h_W(w) dw$$

integrali ile bulunabilir. Buna göre $\underline{X}|W$ koşullu olasılık yoğunluk fonksiyonu $f_{\underline{X}|W}(\underline{x}|w)$ için (2.23) ve (2.9)’dan

$$\begin{aligned}
f_{\underline{x}|w}(\underline{x}|w) &= \frac{1}{(2\pi)^{d/2} |w\mathbf{\Sigma}|^{1/2}} \exp \left\{ -\frac{1}{2} (\underline{x} - \underline{\mu} - w\underline{\gamma})' (w\mathbf{\Sigma})^{-1} (\underline{x} - \underline{\mu} - w\underline{\gamma}) \right\} \\
&= \frac{1}{(2\pi)^{d/2} |\mathbf{\Sigma}|^{1/2} w^{d/2}} \exp \left\{ -\frac{1}{2} (\underline{x} - \underline{\mu} - w\underline{\gamma})' (w\mathbf{\Sigma})^{-1} (\underline{x} - \underline{\mu} - w\underline{\gamma}) \right\}, \underline{x} \in \mathbb{R}^d
\end{aligned}$$

elde edilir. Üstel ifadenin düzenlenmesiyle

$$f_{\underline{x}|w}(\underline{x}|w) = \frac{w^{-d/2}}{(2\pi)^{d/2} |\mathbf{\Sigma}|^{1/2}} \exp \left\{ (\underline{x} - \underline{\mu})' \mathbf{\Sigma}^{-1} \underline{\gamma} - \frac{(\underline{x} - \underline{\mu})' \mathbf{\Sigma}^{-1} (\underline{x} - \underline{\mu})}{2w} - \frac{\underline{\gamma}' \mathbf{\Sigma}^{-1} \underline{\gamma}}{2/w} \right\}$$

yazılabilir. Son eşitlik ile verilen $f_{\underline{x}|w}(\underline{x}|w)$ koşullu olasılık yoğunluk fonksiyonu ve $W \sim N^-(\lambda, \chi, \psi)$ rasgele değişkeninin (2.13)'de verilen olasılık yoğunluk fonksiyonunu kullanarak

$$\begin{aligned}
f_{\underline{x}}(\underline{x}) &= \int_0^\infty f_{\underline{x}|w}(\underline{x}|w) h_w(w) dw \\
&= \int_0^\infty \frac{w^{-d/2}}{(2\pi)^{d/2} |\mathbf{\Sigma}|^{1/2}} \exp \left\{ (\underline{x} - \underline{\mu})' \mathbf{\Sigma}^{-1} \underline{\gamma} - \frac{(\underline{x} - \underline{\mu})' \mathbf{\Sigma}^{-1} (\underline{x} - \underline{\mu})}{2w} - \frac{\underline{\gamma}' \mathbf{\Sigma}^{-1} \underline{\gamma}}{2/w} \right\} \times \\
&\quad \frac{\chi^{-\lambda} (\sqrt{\chi\psi})^\lambda}{2K_\lambda(\sqrt{\chi\psi})} w^{\lambda-1} \exp \left(-\frac{1}{2} \left(\chi \frac{1}{w} + \psi w \right) \right) dw
\end{aligned}$$

biçiminde elde edilir. Düzenleme yapıldıktan sonra $D_{\underline{x}} = (\underline{x} - \underline{\mu})' \mathbf{\Sigma}^{-1} (\underline{x} - \underline{\mu})$ olmak üzere

$$\begin{aligned}
f_{\underline{x}}(\underline{x}) &= \frac{\chi^{-\lambda} (\sqrt{\chi\psi})^\lambda}{(2\pi)^{d/2} |\mathbf{\Sigma}|^{1/2} K_\lambda(\sqrt{\chi\psi})} \exp \left\{ (\underline{x} - \underline{\mu})' \mathbf{\Sigma}^{-1} \underline{\gamma} \right\} \times \\
&\quad \frac{1}{2} \int_0^\infty w^{\lambda-d/2-1} \exp \left\{ -\frac{(\underline{x} - \underline{\mu})' \mathbf{\Sigma}^{-1} (\underline{x} - \underline{\mu})}{2w} - \frac{\underline{\gamma}' \mathbf{\Sigma}^{-1} \underline{\gamma}}{2/w} \right\} \exp \left(-\frac{1}{2} \left(\chi \frac{1}{w} + \psi w \right) \right) dw
\end{aligned}$$

$$= \frac{\chi^{-\lambda} (\sqrt{\chi\psi})^\lambda}{(2\pi)^{d/2} |\Sigma|^{1/2} K_\lambda(\sqrt{\chi\psi})} \exp\{(\underline{x} - \underline{\mu})' \Sigma^{-1} \underline{\gamma}\} \times \\ \frac{1}{2} \int_0^\infty w^{\lambda-d/2-1} \exp\left\{-\frac{D_{\underline{x}} + \chi}{2w} - \frac{\underline{\gamma}' \Sigma^{-1} \underline{\gamma} + \psi}{2/w}\right\} dw$$

olarak elde edilir. Son eşitlikte $w = \frac{\sqrt{D_{\underline{x}} + \chi}}{\sqrt{\underline{\gamma}' \Sigma^{-1} \underline{\gamma} + \psi}} w^*$ dönüşümü yapılırsa

$$dw = \frac{\sqrt{D_{\underline{x}} + \chi}}{\sqrt{\underline{\gamma}' \Sigma^{-1} \underline{\gamma} + \psi}} dw^* \text{ olmak üzere}$$

$$f_{\underline{x}}(\underline{x}) = \frac{\chi^{-\lambda} (\sqrt{\chi\psi})^\lambda}{(2\pi)^{d/2} |\Sigma|^{1/2} K_\lambda(\sqrt{\chi\psi})} \exp\{(\underline{x} - \underline{\mu})' \Sigma^{-1} \underline{\gamma}\} \left(\frac{\sqrt{(D_{\underline{x}} + \chi)}}{\sqrt{(\underline{\gamma}' \Sigma^{-1} \underline{\gamma} + \psi)}} \right)^{\lambda-d/2} \times \\ \frac{1}{2} \int_0^\infty (w^*)^{\lambda-d/2-1} \exp\left\{-\frac{\sqrt{(D_{\underline{x}} + \chi)(\underline{\gamma}' \Sigma^{-1} \underline{\gamma} + \psi)}}{2} \left(w^* + \frac{1}{w^*}\right)\right\} dw^*$$

elde edilir. Burada (2.2)'de verilen Bessel Fonksiyonu için integral gösterimi dikkate alınırsa eşitliğin sağ tarafındaki integral için

$$\frac{1}{2} \int_0^\infty (w^*)^{\lambda-d/2-1} \exp\left\{-\frac{\sqrt{(D_{\underline{x}} + \chi)(\underline{\gamma}' \Sigma^{-1} \underline{\gamma} + \psi)}}{2} \left(w^* + \frac{1}{w^*}\right)\right\} dw^* \\ = K_{\lambda-d/2} \left(\sqrt{(D_{\underline{x}} + \chi)(\underline{\gamma}' \Sigma^{-1} \underline{\gamma} + \psi)} \right)$$

olduğu görülür. Böylece

$$f_{\underline{X}}(\underline{x}) = \frac{\chi^{-\lambda} (\sqrt{\chi\psi})^{\lambda}}{(2\pi)^{d/2} |\underline{\Sigma}|^{1/2} K_{\lambda}(\sqrt{\chi\psi})} \exp\{(\underline{x} - \underline{\mu})' \underline{\Sigma}^{-1} \underline{\gamma}\} \left(\frac{\sqrt{(D_{\underline{x}} + \chi)}}{\sqrt{(\underline{\gamma}' \underline{\Sigma}^{-1} \underline{\gamma} + \psi)}} \right)^{\lambda-d/2} \times \\ K_{\lambda-d/2} \left(\sqrt{(D_{\underline{x}} + \chi)(\underline{\gamma}' \underline{\Sigma}^{-1} \underline{\gamma} + \psi)} \right)$$

halini alır. Sabitlerin düzenlenmesiyle çok değişkenli genelleştirilmiş hiperbolik dağılıma sahip $\underline{X}$ rasgele değişkeninin olasılık yoğunluk fonksiyonu

$$c = \frac{\psi^{\lambda} (\sqrt{\chi\psi})^{-\lambda} (\underline{\gamma}' \underline{\Sigma}^{-1} \underline{\gamma} + \psi)^{\frac{d}{2}-\lambda}}{(2\pi)^{d/2} |\underline{\Sigma}|^{1/2} K_{\lambda}(\sqrt{\chi\psi})}$$

olmak üzere

$$f_{\underline{X}}(\underline{x}) = c \frac{\exp\{(\underline{x} - \underline{\mu})' \underline{\Sigma}^{-1} \underline{\gamma}\}}{\left(\sqrt{(D_{\underline{x}} + \chi)(\underline{\gamma}' \underline{\Sigma}^{-1} \underline{\gamma} + \psi)} \right)^{\frac{d}{2}-\lambda}} K_{\lambda-\frac{d}{2}} \left(\sqrt{(D_{\underline{x}} + \chi)(\underline{\gamma}' \underline{\Sigma}^{-1} \underline{\gamma} + \psi)} \right) \quad (2.24)$$

biçiminde elde edilmiş olur. $\underline{X}$ rasgele değişkeninin olasılık yoğunluk fonksiyonu yukarıdaki gibi olduğunda $\underline{X} \sim GH_d(\lambda, \chi, \psi, \underline{\mu}, \underline{\Sigma}, \underline{\gamma})$ gösterimi kullanılır.

(2.20)'den kolayca görülebileceği üzere $E(\underline{X}|\underline{W}) = \underline{\mu} + \underline{W} \underline{\gamma}$ ve $Cov(\underline{X}|\underline{W}) = \underline{W} \underline{\Sigma}$ olup bu değerler yardımıyla çok değişkenli genelleştirilmiş hiperbolik dağılıma sahip $\underline{X}$ rasgele vektörü için

$$E(\underline{X}) = E(E(\underline{X}|\underline{W})) = \underline{\mu} + E(\underline{W}) \underline{\gamma} \quad (2.25)$$

$$Cov(\underline{X}) = E(Cov(\underline{X}|\underline{W})) + Cov(E(\underline{X}|\underline{W})) \\ = E(\underline{W}) \underline{\Sigma} + Var(\underline{W}) \underline{\gamma} \underline{\gamma}' \quad (2.26)$$

beklenen değer ve kovaryans ifadeleri elde edilir. Dikkat edileceği üzere $\underline{\gamma} \neq \underline{0}$ olması durumunda $\underline{\mu}$ ve $\underline{\Sigma}$ parametreleri, $\underline{X}$ 'in beklenen değer vektöründen ve varyans-kovaryans matrisinden farklı birer parametredir. Son iki ifade (2.14) kullanılarak açılırsa

$$E(\underline{X}) = \underline{\mu} + \left(\frac{\chi}{\psi} \right)^{1/2} \frac{K_{\lambda+1}(\sqrt{\chi\psi})}{K_{\lambda}(\sqrt{\chi\psi})} \underline{\gamma} \quad (2.27)$$

$$Cov(\underline{X}) = \left(\frac{\chi}{\psi} \right)^{1/2} \frac{K_{\lambda+1}(\sqrt{\chi\psi})}{K_{\lambda}(\sqrt{\chi\psi})} \underline{\Sigma} + \left(\frac{\chi}{\psi} \frac{K_{\lambda+2}(\sqrt{\chi\psi})}{K_{\lambda}(\sqrt{\chi\psi})} - \frac{\chi}{\psi} \frac{K_{\lambda+1}^2(\sqrt{\chi\psi})}{K_{\lambda}^2(\sqrt{\chi\psi})} \right) \underline{\gamma\gamma'} \quad (2.28)$$

elde edilir.

2.2 Çok Değişkenli Genelleştirilmiş Hiperbolik Dağılım Ailesinin Bazı Özel Durumları

Çok değişkenli genelleştirilmiş hiperbolik dağılım ailesi oldukça esnek özelliklere sahiptir. Aşağıda parametrelerin özel seçimleriyle elde edilecek özel durumlar özetlenmiştir (McNeil vd. 2005).

- Eğer $\lambda = 1$ seçilirse, marjinal dağılımları tek değişkenli hiperbolik dağılımlar olan çok değişkenli dağılım elde edilir.
- Eğer $\lambda = -1/2$ seçilirse, normal ters Gauss dağılımı (NIG) elde edilir.
- Eğer $\lambda > 0$ ve $\chi = 0$ seçilirse, varyans-gamma dağılımı elde edilir.
- Eğer $\lambda = -1$, $\chi = \nu$ ve $\psi = 0$ seçilirse çarpık-t dağılımı elde edilir.

2.3 Parametrelerin Tahmini ve EM Algoritması

Çok değişkenli genelleştirilmiş hiperbolik dağılım ailesi, sahip olduğu esnek özelliklere karşın, uygulamada parametrelerin tahmin edilmesi gerektiği durumlarda üzerinde çalışılması zor bir dağılım ailesi haline gelmektedir. Özellikle boyut sayısı arttıkça olabilirlik fonksiyonu maksimizasyonu karmaşık hale gelmektedir. Parametre tahmininde karşılaşılan diğer bir zorluk ise tahmin edilmeye çalışılan parametrelerden

bazılarının, gözlenemeyen karıştırıcı dağılımın parametreleri olmasından kaynaklanmaktadır. EM algoritması direk olarak gözlenemeyen parametrelerle biçimlendirilmiş dağılımlarda, en çok olabilirlik tahmin edicilerini bulmayı amaçlayan iteratif bir yöntemdir. Yöntem E-Adımı ve M-Adımı olarak isimlendirilen iki adımdan oluşur. İlk adım olan E-Adımında dağılımın log-olabilirlik fonksiyonu, gözlenemeyen değişkenlere göre tahmin edilir. Bu adımda $\underline{X}_1, \underline{X}_2, \dots, \underline{X}_n$ 'lerin gözlem değerleri kullanılır. M-Adımında ilgilendiğimiz parametre değerleri, ilk adımda tahmin edilen log-olabilirlik fonksiyonunu maksimum yapacak şekilde hesaplanır. Bu hesap değerleri E-Adımının yeni girdileri olarak kabul edilip iterasyon yenilenir. Parametre değerlerinde yakınsama sağlanıncaya kadar yukarıdaki işlemler tekrar edilir.

Bu bölümde McNeil (2005) tarafından verilen çok değişkenli genelleştirilmiş hiperbolik dağılım için EM (Expectation-Estimation) algoritmasını ele alınacaktır. Sonraki bölümlerde çarpık-t, varyans gamma, çarpık laplace, normal ters Gauss dağılımları için söz konusu algoritma özelleştirilecek ve bu dağılımlar için EM algoritmaları oluşturulacaktır. EM algoritmaları için Hu (2005) ve Arslan (2010) kaynaklarına bakılabilir.

$\underline{X}_1, \underline{X}_2, \dots, \underline{X}_n \sim GH_d(\lambda, \chi, \psi, \underline{\mu}, \underline{\Sigma}, \underline{\gamma})$ dağılımından rasgele bir örneklem olsun. GH dağılımının tahmin etmeye çalışılan parametrelerinin kümesi $\zeta = (\lambda, \chi, \psi, \underline{\mu}, \underline{\Sigma}, \underline{\gamma})$ ile gösterilsin. Olabilirlik fonksiyonunun logaritması

$$\log L(\zeta; \underline{x}_1, \underline{x}_2, \dots, \underline{x}_n) = \sum_{i=1}^n \log f_{\underline{x}_i}(\underline{x}_i; \zeta) \quad (2.29)$$

olarak yazılabilir. $\zeta = (\lambda, \chi, \psi, \underline{\mu}, \underline{\Sigma}, \underline{\gamma})$ parametrelerinin tahmin edicileri, olabilirlik fonksiyonunun maksimum yapılmasıyla bulunabilir. Ancak parametre sayısının fazla olması özellikle boyut sayısı arttıkça bu maksimizasyon probleminin çözümünü oldukça zorlaştırmaktadır. Bu yüzden daha önceki bölümde bahsedilen, koşullu normal dağılım gösterimini kullanarak

$$f_{\underline{x}_i}(\underline{x}_i; \lambda, \chi, \psi, \underline{\mu}, \underline{\Sigma}, \underline{\gamma}) = f_{\underline{x}_i|w_i}(\underline{x}_i | w_i; \underline{\mu}, \underline{\Sigma}, \underline{\gamma}) h_{w_i}(w_i; \lambda, \chi, \psi) \quad (2.30)$$

biçiminde yazılsın. Burada $\underline{X}|W \sim N_d(\underline{\mu} + w\underline{\gamma}, w\underline{\Sigma})$ ve $h_w(w; \lambda, \chi, \psi)$ (2.13)'de verilen genelleştirilmiş ters Gauss dağılımının olasılık yoğunluk fonksiyonudur. Böylece (2.29) ile verilen log-olabilirlik fonksiyonu

$$\begin{aligned} \log L(\zeta; \underline{x}_1, \underline{x}_2, \dots, \underline{x}_n, w_1, w_2, \dots, w_n) = \\ \sum_{i=1}^n \log f_{\underline{x}_i|w_i}(\underline{x}_i | w_i; \underline{\mu}, \underline{\Sigma}, \underline{\gamma}) + \sum_{i=1}^n \log h_{w_i}(w_i; \lambda, \chi, \psi) \end{aligned} \quad (2.31)$$

gibi yazılabilecektir. Dolayısıyla $\underline{\mu}, \underline{\Sigma}, \underline{\gamma}$ parametreleri

$$L_1 = \sum_{i=1}^n \log f_{\underline{x}_i|w_i}(\underline{x}_i | w_i; \underline{\mu}, \underline{\Sigma}, \underline{\gamma})$$

ifadesinin, λ, χ, ψ parametreleri ise

$$L_2 = \sum_{i=1}^n \log h_{w_i}(w_i; \lambda, \chi, \psi)$$

ifadesinin maksimum yapılmasıyla ayrı ayrı elde edilebilir. Fakat L_1 ifadesinin maksimum yapılması için w_i gözlenemeyen karıştırıcı değişkenlerinin kullanılması gerekmektedir. Bu noktada w_i 'lerin gözlenebilir olduğu varsayılarak bu değerler optimize edilir. (2.31) eşitliğinde ilgili dağılımların olabilirlik fonksiyonu yardımıyla

$$\begin{aligned} L_1 = \sum_{i=1}^n \log f_{\underline{x}_i|w_i}(\underline{x}_i | w_i; \underline{\mu}, \underline{\Sigma}, \underline{\gamma}) = -\frac{n}{2} \log |\underline{\Sigma}| - \frac{d}{2} \sum_{i=1}^n \log w_i + \sum_{i=1}^n (\underline{x}_i - \underline{\mu})' \underline{\Sigma}^{-1} \underline{\gamma} \\ - \frac{1}{2} \sum_{i=1}^n \frac{1}{w_i} D_{\underline{x}_i} - \frac{1}{2} \underline{\gamma}' \underline{\Sigma}^{-1} \underline{\gamma} \sum_{i=1}^n w_i \end{aligned} \quad (2.32)$$

$$L_2 = \sum_{i=1}^n \log h_{w_i}(w_i; \lambda, \chi, \psi) = (\lambda - 1) \sum_{i=1}^n \log w_i - \frac{\chi}{2} \sum_{i=1}^n w_i^{-1} - \frac{\psi}{2} \sum_{i=1}^n w_i - \frac{n\lambda}{2} \log \chi + \frac{n\lambda}{2} \log \psi - n \log \left(2K_\lambda(\sqrt{\chi\psi}) \right) \quad (2.33)$$

olacaktır. L_1 'in $\underline{\mu}, \underline{\Sigma}, \underline{\gamma}$ parametrelerine göre maksimum yapılması için

$$\frac{\partial L_1}{\partial \underline{\mu}} = \underline{0}, \quad \frac{\partial L_1}{\partial \underline{\gamma}} = \underline{0}, \quad \frac{\partial L_1}{\partial \underline{\Sigma}} = \underline{0}$$

kısmi türevleri alınarak denklem sistemi oluşturulur.

$$\frac{\partial L_1}{\partial \underline{\mu}} = \left(-n \underline{\Sigma}^{-1} \underline{\gamma} - \sum_{i=1}^n \frac{1}{w_i} (-\underline{x}_i' \underline{\Sigma}^{-1} + \underline{\mu}' \underline{\Sigma}^{-1}) \right) = \underline{0} \text{ eşitliğinden } \underline{\mu} \text{ çekilirse}$$

$$\hat{\underline{\mu}} = \frac{n^{-1} \sum_{i=1}^n w_i^{-1} \underline{x}_i - \hat{\underline{\gamma}}}{n^{-1} \sum_{i=1}^n w_i^{-1}}, \quad (2.34)$$

$$\frac{\partial L_1}{\partial \underline{\gamma}} = \left(\sum_{i=1}^n (\underline{x}_i - \underline{\mu})' \underline{\Sigma}^{-1} - \underline{\gamma}' \underline{\Sigma}^{-1} \sum_{i=1}^n w_i \right) = \underline{0} \text{ eşitliğinden } \underline{\gamma} \text{ çekilip, } \underline{\mu} \text{ parametresinin (2.34)}$$

'de verilen tahmini yerine yazılırsa

$$\hat{\underline{\gamma}} = \frac{n^{-1} \sum_{i=1}^n w_i^{-1} (\bar{x} - \underline{x}_i)}{n^{-2} \left(\sum_{i=1}^n w_i \right) \left(\sum_{i=1}^n w_i^{-1} \right) - 1}, \quad (2.35)$$

$$\begin{aligned} \frac{\partial L_1}{\partial \underline{\Sigma}} = & -\frac{n}{2} \underline{\Sigma}^{-1} - \sum_{i=1}^n \underline{\Sigma}^{-1} (\underline{x}_i - \underline{\mu}) \underline{\gamma}' + \\ & \sum_{i=1}^n w_i^{-1} \underline{\Sigma}^{-1} (\underline{x}_i - \underline{\mu}) (\underline{x}_i - \underline{\mu})' \underline{\Sigma}^{-1} + \frac{1}{2} \underline{\Sigma}^{-1} \underline{\gamma} \underline{\gamma}' \underline{\Sigma}^{-1} \sum_{i=1}^n w_i = \underline{0} \end{aligned}$$

eşitliğinden ise Σ 'nın çekilmesiyle*

$$\hat{\Sigma} = \frac{1}{n} \sum_{i=1}^n w_i^{-1} (\underline{x}_i - \hat{\underline{\mu}})(\underline{x}_i - \hat{\underline{\mu}})' - \frac{1}{n} \sum_{i=1}^n w_i \hat{\underline{\gamma}} \hat{\underline{\gamma}}' \quad (2.36)$$

tahmin edicileri elde edilmiş olur. Benzer şekilde L_2 'in λ, χ, ψ parametrelerine göre maksimum yapılması için bu parametrelere göre kısmi türevleri alınır. Ancak bu noktadan sonra λ 'nın sabit olduğunu düşüneceğiz. Bu yüzden diğer iki parametreye göre

$$\frac{\partial L_2}{\partial \chi} = 0, \quad \frac{\partial L_2}{\partial \psi} = 0$$

kısmi türevleri alınarak denklem sistemi oluşturulur. Buna göre

$$\frac{\partial L_2}{\partial \chi} = -\frac{1}{2} \sum_{i=1}^n w_i^{-1} - \frac{n\lambda}{\chi} + n \sqrt{\frac{\psi}{\chi}} \frac{K_{\lambda+1}(\theta)}{K_{\lambda}(\theta)} = 0, \quad \theta = \sqrt{\chi\psi} \quad (2.37)$$

$$\frac{\partial L_2}{\partial \psi} = -\frac{1}{2} \sum_{i=1}^n w_i + \frac{n\chi K_{\lambda+1}(\theta)}{2\theta K_{\lambda}(\theta)} = 0 \quad (2.38)$$

denklemleri elde edilir. Eşitliklerin her iki tarafı n^{-1} ile çarpıldıktan sonra (2.38)'den elde edilen

$$\chi = n^{-1} \sum_{i=1}^n w_i \frac{\theta K_{\lambda}(\theta)}{K_{\lambda+1}(\theta)}$$

değeri (2.37)'da yerine yazılırsa

* $\frac{\partial}{\partial \mathbf{X}} \log(\det(\mathbf{X})) = (\mathbf{X}^{-1})'$, $\frac{\partial}{\partial \mathbf{X}} \underline{a}' \mathbf{X} \underline{b} = -(\mathbf{X}^{-1})' \underline{a} \underline{b}' (\mathbf{X}^{-1})'$

$$K_{\lambda}^2(\theta)\theta\left(1-n^{-2}\sum_{i=1}^nw_i\sum_{j=1}^nw_j^{-1}\right)-2\lambda K_{\lambda+1}(\theta)K_{\lambda}(\theta)=0 \quad (2.39)$$

eşitliği elde edilir. Son eşitlikten θ 'nın çekilmesiyle

$$\hat{\chi} = \frac{n^{-1}\theta\sum_{i=1}^nw_iK_{\lambda}(\theta)}{K_{\lambda+1}(\theta)} \quad (2.40)$$

$$\hat{\psi} = \theta^2/\chi \quad (2.41)$$

değerleri elde edilir (Hu 2005).

Buraya kadar olan kısımda w_i değerlerinin gözlenebilir olduğu varsayılmıştı. Bu noktada (2.31) ile verilen log-olabilirlik fonksiyonunun $\underline{X}_1, \underline{X}_2, \dots, \underline{X}_n$ 'lerin gözlem değerlerini kullanarak, bu değerler bilinirken koşullu beklenen değeri alınacaktır. E-adımının k .iterasyonu için

$$Q = E\left(\log L(\zeta; \underline{x}_1, \underline{x}_2, \dots, \underline{x}_n, W_1, W_2, \dots, W_n) | \underline{x}_1, \underline{x}_2, \dots, \underline{x}_n; \zeta^{[k]}\right)$$

M-Adımında bu fonksiyonun maksimum yapılmasıyla yeni $\zeta^{[k+1]}$ parametre değerleri elde edilir. Bu işlem için $E(W_i | \underline{x}_i; \zeta^{[k]})$, $E(W_i^{-1} | \underline{x}_i; \zeta^{[k]})$ ve $E(\log(W_i) | \underline{x}_i; \zeta^{[k]})$ koşullu beklenen değerlerini, (2.32) ve (2.33) da verilen L_1 ve L_2 log-olabilirlik fonksiyonlarında w_i , w_i^{-1} ve $\log w_i$ değerlerinin yerine yazılması gerekir. Beklenen değerleri

$$\delta_i^{[l]} = E(W_i^{-1} | \underline{x}_i; \zeta^{[l]}) , \eta_i^{[l]} = E(W_i | \underline{x}_i; \zeta^{[l]}) , \xi_i^{[l]} = E(\log(W_i) | \underline{x}_i; \zeta^{[l]}) \quad (2.42)$$

notasyonları ile gösterelim. Ayrıca bu değerlerin ortalamaları için

$$\bar{\delta}^{[.]} = \frac{1}{n} \sum_{i=1}^n \delta_i^{[.]} \quad , \quad \bar{\eta}^{[.]} = \frac{1}{n} \sum_{i=1}^n \eta_i^{[.]} \quad , \quad \bar{\xi}^{[.]} = \frac{1}{n} \sum_{i=1}^n \xi_i^{[.]} \quad (2.43)$$

sembolleri kullanılsın (McNeil vd. 2005). (2.42)'de verilen beklenen değer ifadeleri (2.14) ve (2.15) eşitliklerini kullanarak

$$\delta_i^{[k]} = \left(\frac{D_{\underline{x}_i}^{[k]} + \chi^{[k]}}{\psi^{[k]} + \underline{\gamma}^{[k]}, \Sigma^{[k]-1} \underline{\gamma}^{[k]}} \right)^{-\frac{1}{2}} \frac{K_{\lambda - \frac{d}{2} - 1} \left(\sqrt{(D_{\underline{x}_i}^{[k]} + \chi^{[k]}) (\psi^{[k]} + \underline{\gamma}^{[k]}, \Sigma^{[k]-1} \underline{\gamma}^{[k]})} \right)}{K_{\lambda - \frac{d}{2}} \left(\sqrt{(D_{\underline{x}_i}^{[k]} + \chi^{[k]}) (\psi^{[k]} + \underline{\gamma}^{[k]}, \Sigma^{[k]-1} \underline{\gamma}^{[k]})} \right)} \quad (2.44)$$

$$\eta_i^{[k]} = \left(\frac{D_{\underline{x}_i}^{[k]} + \chi^{[k]}}{\psi^{[k]} + \underline{\gamma}^{[k]}, \Sigma^{[k]-1} \underline{\gamma}^{[k]}} \right)^{\frac{1}{2}} \frac{K_{\lambda - \frac{d}{2} + 1} \left(\sqrt{(D_{\underline{x}_i}^{[k]} + \chi^{[k]}) (\psi^{[k]} + \underline{\gamma}^{[k]}, \Sigma^{[k]-1} \underline{\gamma}^{[k]})} \right)}{K_{\lambda - \frac{d}{2}} \left(\sqrt{(D_{\underline{x}_i}^{[k]} + \chi^{[k]}) (\psi^{[k]} + \underline{\gamma}^{[k]}, \Sigma^{[k]-1} \underline{\gamma}^{[k]})} \right)} \quad (2.45)$$

$$\begin{aligned} \xi_i^{[k]} = & \frac{1}{2} \log \left(\frac{D_{\underline{x}_i}^{[k]} + \chi^{[k]}}{\psi^{[k]} + \underline{\gamma}^{[k]}, \Sigma^{[k]-1} \underline{\gamma}^{[k]}} \right) \\ & + \frac{\frac{\partial}{\partial \alpha} K_{\lambda - \frac{d}{2} + \alpha} \left(\sqrt{(D_{\underline{x}_i}^{[k]} + \chi^{[k]}) (\psi^{[k]} + \underline{\gamma}^{[k]}, \Sigma^{[k]-1} \underline{\gamma}^{[k]})} \right)}{K_{\lambda - \frac{d}{2}} \left(\sqrt{(D_{\underline{x}_i}^{[k]} + \chi^{[k]}) (\psi^{[k]} + \underline{\gamma}^{[k]}, \Sigma^{[k]-1} \underline{\gamma}^{[k]})} \right)} \Big|_{\alpha=0} \end{aligned} \quad (2.46)$$

olarak hesaplanır. Böylece L_1 'in maksimum yapılması için (2.34), (2.35) ve (2.36) denklemlerinden parametrelerin $k+1$. adım için yeni değerleri

$$\hat{\underline{\gamma}}^{[k+1]} = \frac{n^{-1} \sum_{i=1}^n \delta_i^{[k]} (\bar{\underline{x}} - \underline{x}_i)}{\frac{\bar{\delta}^{[k]} - \bar{\eta}^{[k]}}{\eta} - 1}, \quad (2.47)$$

$$\hat{\underline{\mu}}^{[k+1]} = \frac{n^{-1} \sum_{i=1}^n \delta_i^{[k]} \underline{x}_i - \hat{\underline{\gamma}}^{[k+1]}}{\bar{\delta}^{[k]}}, \quad (2.48)$$

$$\hat{\Sigma}^{[k+1]} = \frac{1}{n} \sum_{i=1}^n \delta_i^{[k]} \left(\underline{x}_i - \hat{\underline{\mu}}^{[k+1]} \right) \left(\underline{x}_i - \hat{\underline{\mu}}^{[k+1]} \right)^t - \bar{\eta}^{[k]} \hat{\underline{\gamma}}^{[k+1]} \hat{\underline{\gamma}}^{[k+1]}, \quad (2.49)$$

olarak hesaplanır. Benzer olarak L_2 'nin de koşullu beklenen değeri maksimum yapılmak istenirse (2.40) ve (2.41) denklemlerinden parametrelerin $k+1$. adım için yeni değerleri

$$\hat{\chi}^{[k+1]} = \frac{\theta^{[k+1]} \bar{\eta}^{[k]} K_{\lambda}(\theta^{[k+1]})}{K_{\lambda+1}(\theta^{[k+1]})} \quad (2.50)$$

$$\hat{\psi}^{[k+1]} = \frac{\theta^{[k+1]^2}}{\chi^{[k+1]}} \quad (2.51)$$

eşitlikleri ile hesaplanır. Aşağıda çok değişkenli genelleştirilmiş hiperbolik dağılımın parametrelerinin tahmini için EM algoritması yer almaktadır.

- E Adımı

- o k.iterasyon'da tüm gözlemler için (2.44),(2.45) ve (2.46)'u kullanarak

$$\delta_i^{[k]}, \eta_i^{[k]}, \xi_i^{[k]}$$

değerlerini hesapla.

- o k.iterasyon'da

$$\bar{\delta}^{[k]} = \frac{1}{n} \sum_{i=1}^n \delta_i^{[k]}, \quad \bar{\eta}^{[k]} = \frac{1}{n} \sum_{i=1}^n \eta_i^{[k]}, \quad \bar{\xi}^{[k]} = \frac{1}{n} \sum_{i=1}^n \xi_i^{[k]}$$

değerlerini bul.

- M Adımı

- o Parametrelerin yeni değerlerini

$$\hat{\underline{\gamma}}^{[k+1]}, \hat{\underline{\mu}}^{[k+1]}, \hat{\Sigma}^{[k+1]}, \hat{\chi}^{[k+1]}, \hat{\psi}^{[k+1]}$$

(2.47),(2.48),(2.49),(2.50) ve (2.51)'i kullanarak hesapla.

- o Yakınsama sağlanmadıysa $k=k+1$ alarak E-Adımına dön.

Bu algoritma genelleştirilmiş hiperbolik dağılımın özel halleri için değiştirilebilir. İterasyonun başlangıç değerleri olarak $\underline{\mu}$ için örneklem ortalaması, $\underline{\Sigma}$ için örneklem Varyans-kovaryans matrisi veya bu parametrelerin MCD robust tahminleri kullanılabilir. $\underline{\gamma}, \underline{\chi}, \underline{\psi}$ parametreleri için ise sıfıra çok yakın değerler seçilebilir.

3. ÇOK DEĞİŞKENLİ GENELLEŞTİRİLMİŞ HİPERBOLİK ÇARPIK-T DAĞILIMI

Tek değişkenli Student-t dağılımı finansal verilerin modellenmesinde, normal dağılıma kıyasla değişik serbestlik dereceleri ile kuyruk kalınlığının ayarlanabilmesinden dolayı sıklıkla kullanılmaktadır. Student-t dağılımının çarpık durumları için birçok tanımlama mevcuttur. Birbirlerine alternatif olan bu tanımlamalara örnek olarak Fernandez ve Steel (1998), Jones ve Faddy (2003) ile Azzalini ve Capitanio (2003)'nun çalışmaları verilebilir. Aas ve Haff (2006)'ın genelleştirilmiş hiperbolik dağılım ailesinin bir özel hali olarak ele aldığı ve çalışmanın bu bölümünde kaynak olarak alınan çalışma haricinde, dağılım ile ilgili diğer çalışmalar için Hu (2005), Hu ve Kercheval (2010) ile McNeil vd. (2005) kaynaklarına bakılabilir.

$\underline{X} \sim GH_d(\lambda, \chi, \psi, \underline{\mu}, \underline{\Sigma}, \gamma)$ dağılımına sahip olsun. Bu durumda, $W \sim N^-(\lambda, \chi, \psi)$ olmak üzere, $\underline{X}$ rasgele vektörünün

$$\underline{X} = \underline{\mu} + W\underline{\gamma} + \sqrt{W}\mathbf{A}\underline{Z}$$

dönüşümü ile gösterilebileceği verilmişti. W 'nin dağılımında $\lambda = -v/2$, $\chi = v$ ve $\psi = 0$ parametre seçimi yapılırsa elde edilecek dağılım '*çok değişkenli genelleştirilmiş hiperbolik çarpık-t dağılım*' dır. Buna göre $\underline{X}$ rasgele vektörü

$$c = \frac{2^{1-(v+d)/2}}{\Gamma\left(\frac{v}{2}\right)(\pi v)^{d/2} |\underline{\Sigma}|^{1/2}}$$

olmak üzere

$$f_{\underline{X}}(\underline{x}) = c \frac{K_{\frac{v+d}{2}} \left[\sqrt{(v+D_{\underline{x}})(\underline{\gamma}'\underline{\Sigma}^{-1}\underline{\gamma})} \right] \exp\left\{(\underline{x}-\underline{\mu})'\underline{\Sigma}^{-1}\underline{\gamma}\right\}}{\left[\sqrt{(v+D_{\underline{x}})(\underline{\gamma}'\underline{\Sigma}^{-1}\underline{\gamma})} \right]^{-(v+d)/2} \left(1+\frac{1}{v}D_{\underline{x}}\right)^{(v+d)/2}} \quad (3.1)$$

olasılık yoğunluk fonksiyonuna sahiptir. $\underline{X}$ rasgele vektörü (3.1)'de verilen olasılık yoğunluk fonksiyonuna sahip olduğunda $\underline{X} \sim Skew-t_d(v, \underline{\mu}, \underline{\Sigma}, \underline{\gamma})$ olarak gösterilecektir. Çalışmanın ilerleyen kısımlarında çok değişkenli genelleştirilmiş hiperbolik çarpık-t dağılımını kısaca '*çarpık-t dağılımı*' olarak ifade edilecektir.

$\lambda = -v/2$, $\chi = v$ ve $\psi = 0$ parametre değerleri W 'nin dağılımında yerlerine yazılırsa $W \sim N\left(\lambda = -\frac{v}{2}, \chi = v, \psi = 0\right)$ dağılımının olasılık yoğunluk fonksiyonu

$$\begin{aligned} f_w\left(w; -\frac{v}{2}, v, \psi \rightarrow 0\right) &= \lim_{\psi \rightarrow 0} \frac{v^{-\left(\frac{v}{2}\right)} \left(\sqrt{v\psi}\right)^{\left(\frac{v}{2}\right)}}{2K_{\left(\frac{v}{2}\right)} \left(\sqrt{v\psi}\right)} w^{\left(\frac{v}{2}\right)-1} \exp\left(-\frac{1}{2}\left(v\frac{1}{w} + \psi w\right)\right) \\ &= \frac{v^{\frac{v}{2}} \left(\sqrt{v\psi}\right)^{\left(\frac{v}{2}\right)}}{2\Gamma\left(\frac{v}{2}\right) 2^{\frac{v}{2}-1} \left(\sqrt{v\psi}\right)^{\frac{v}{2}}} w^{\left(\frac{v}{2}\right)-1} \exp\left(-\frac{v}{2w}\right) \\ &= \frac{\left(\frac{v}{2}\right)^{\frac{v}{2}}}{\Gamma\left(\frac{v}{2}\right)} w^{\left(\frac{v}{2}\right)-1} \exp\left(-\frac{v}{2w}\right), w > 0 \end{aligned}$$

biçiminde elde edilir. Buradan

$$W \sim IGamma\left(\alpha = \frac{v}{2}, \beta = \frac{v}{2}\right) \quad (3.2)$$

olduğu görülür. Ters gamma dağılımı için Bölüm 2.1.4’de verilen beklenen değer ve varyans ifadelerinden

$$E(W) = \frac{\beta}{\alpha - 1} = \frac{v}{v - 2} \quad \text{ve} \quad Var(W) = \frac{\beta^2}{(\alpha - 1)^2 (\alpha - 2)} = \frac{2v^2}{(v - 2)^2 (v - 4)}$$

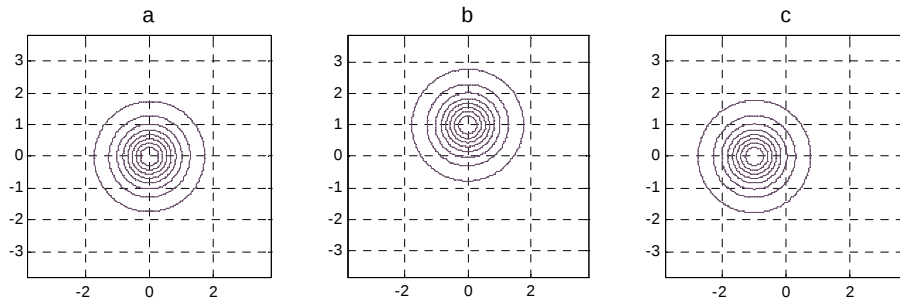
elde edilir. Böylece (2.25) ve (2.26) kullanılarak çarpık-t dağılımı için

$$\begin{aligned} E(\underline{X}) &= \underline{\mu} + E(W)\underline{\gamma} \\ &= \underline{\mu} + \frac{v}{v - 2}\underline{\gamma} \quad , v > 2 \end{aligned} \quad (3.3)$$

$$\begin{aligned} Cov(\underline{X}) &= E(W)\underline{\Sigma} + Var(W)\underline{\gamma}\underline{\gamma}' \\ &= \frac{v}{v - 2}\underline{\Sigma} + \frac{2v^2}{(v - 2)^2 (v - 4)}\underline{\gamma}\underline{\gamma}' \quad , v > 4 \end{aligned} \quad (3.4)$$

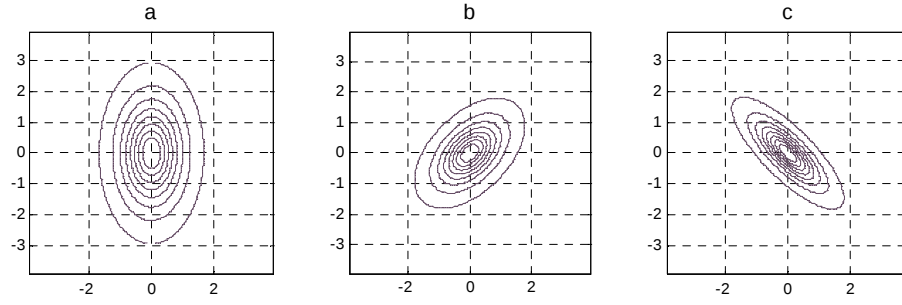
beklenen değeri ve kovaryans değerleri elde edilir.

$Skew-t_2(v, \underline{\mu}, \underline{\Sigma}, \underline{\gamma})$ dağılımı için olasılık yoğunluk fonksiyonunun iz düşüm grafiği $v, \underline{\mu}, \underline{\Sigma}, \underline{\gamma}$ parametrelerinin değişik seçimleri ile şekil 3.1-3.4’de verilmiştir.



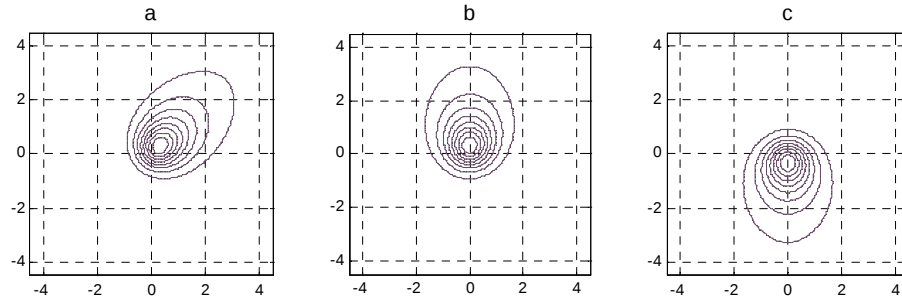
Şekil 3.1 İki değişkenli genelleştirilmiş çarpık-t dağılımının $\underline{\mu}$ parametresine göre değişimi

a. $\underline{\mu} = [0 \ 0]'$, b. $\underline{\mu} = [0 \ 1]'$, c. $\underline{\mu} = [-1 \ 0]'$



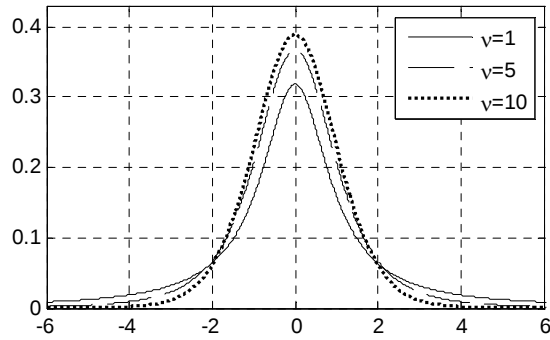
Şekil 3.2 İki değişkenli genelleştirilmiş çarpık-t dağılımının Σ parametresine göre değişimi

$$\text{a. } \Sigma = \begin{bmatrix} 1 & 0.01 \\ 0.01 & 3 \end{bmatrix} \quad \text{b. } \Sigma = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix} \quad \text{c. } \Sigma = \begin{bmatrix} 1 & -0.9 \\ -0.9 & 1 \end{bmatrix}$$



Şekil 3.3 İki değişkenli genelleştirilmiş çarpık-t dağılımının $\underline{\gamma}$ parametresine göre değişimi

$$\text{a. } \underline{\gamma} = [1 \ 1]' \quad , \quad \text{b. } \underline{\gamma} = [0 \ 1]' \quad , \quad \text{c. } \underline{\gamma} = [0 \ -1]'$$



Şekil 3.4 Tek değişkenli genelleştirilmiş çarpık-t dağılımının ν parametresine göre değişimi

Grafiklerden görüleceği üzere $\underline{\mu}$ parametresi dağılımın konumunu, Σ parametresi saçılımını, $\underline{\gamma}$ parametresi ise dağılımın çarpıklığını değiştirmektedir. ν parametresinin

yüksek değerleri, dağılımı sivri ve ince kuyruklu hale getirirken, değer azaldıkça basıklık artmakta ve kalın kuyruklu bir dağılım elde edilmektedir.

3.1 Çarpık-t Dağılımından Rasgele Sayı Üretilmesi

Çalışmanın ilerleyen kısımlarında kullanılacak simülasyonlarda, çarpık-t dağılımından üretilmiş sayılara ihtiyaç duyulacaktır. Bunun için aşağıdaki adımlar izlenir.

Girdiler: $v, \underline{\mu}_{d \times 1}, \underline{\gamma}_{d \times 1}, \underline{\Sigma}_{d \times d}$

$\underline{\Sigma}$ matrisinin Cholesky ayrışımından $\underline{\Sigma} = \mathbf{A}\mathbf{A}^t$ olacak şekilde $\mathbf{A}_{d \times d}$ bul.

1. $\underline{Z} \sim N_k(\underline{0}, \mathbf{I}_k)$ dağılımından sayı üret.
2. $W \sim IGamma(v/2, v/2)$ dağılımından sayı üret.
3. $\underline{X} = \underline{\mu} + W\underline{\gamma} + \sqrt{W}\mathbf{A}\underline{Z}$ değerini hesapla.

Çıktı: $\underline{X}$

$W \sim IGamma(v/2, v/2)$ dağılımından sayı üretmek için Bölüm 2.1.4 'de verilen bilgiler kullanılabilir.

3.2 Çarpık-t Dağılımının Parametrelerinin Tahmini İçin EM-Algoritması

Çok değişkenli genelleştirilmiş hiperbolik dağılım için EM-Algoritması Bölüm 2.3'de verilmişti. $\lambda = -v/2$, $\chi = v$ ve $\psi = 0$ parametre değerleri algoritmada yerlerine yazılarak çarpık-t dağılımı için EM-Algoritması oluşturulur.

v parametresinin tahmini için (2.33)' da verilen log-olabilirlik fonksiyonu

$$L_2 = (\lambda - 1) \sum_{i=1}^n \log w_i - \frac{\chi}{2} \sum_{i=1}^n w_i^{-1} - \frac{\psi}{2} \sum_{i=1}^n w_i - n \log \left(\frac{2K_\lambda(\sqrt{\chi\psi})}{\chi^{-\lambda/2} \psi^{\lambda/2}} \right)$$

biçiminde yazıldıktan sonra $\psi = 0, \lambda = -v/2, \chi = v$ durumu için (2.4) özelliği kullanılarak

$$L_2 = -\left(\frac{\nu}{2} + 1\right) \sum_{i=1}^n \log w_i - \frac{\nu}{2} \sum_{i=1}^n w_i^{-1} - n \log \left(2^{\frac{\nu}{2}} \Gamma\left(\frac{\nu}{2}\right) \nu^{-\frac{\nu}{2}} \right)$$

elde edilir. İfade ν 'ye göre türevlenip sıfıra eşitlendiğinde

$$\frac{\partial}{\partial \nu} L_2 = -n^{-1} \sum_{i=1}^n \log(w_i) - n^{-1} \sum_{i=1}^n w_i^{-1} - \psi\left(\frac{\nu}{2}\right) + 1 = 0 \quad (3.5)$$

elde edilir. Eşitliği sağlayan ν değeri ν 'nün yeni değeri olarak alınır. (3.5) eşitliğinde (2.43)'de verilen beklenen değer ifadeleri kullanılarak EM-Algoritmasının bu parametreyle ilgili adımı tamamlanmış olur. (Burada $\psi(\cdot)$ digamma fonksiyonunu göstermektedir.)

- E-Adımı

- o k.iterasyon'da tüm gözlemler için aşağıdakileri hesapla

$$\delta_i^{[k]} = \left(\frac{D_{\underline{x}_i}^{[k]} + \nu^{[k]}}{\underline{\gamma}^{[k]}, \underline{\Sigma}^{[k-1]} \underline{\gamma}^{[k]}} \right)^{-\frac{1}{2}} \frac{K_{\frac{\nu+d+2}{2}} \left(\sqrt{(D_{\underline{x}_i}^{[k]} + \nu^{[k]}) (\underline{\gamma}^{[k]}, \underline{\Sigma}^{[k-1]} \underline{\gamma}^{[k]})} \right)}{K_{\frac{\nu+d}{2}} \left(\sqrt{(D_{\underline{x}_i}^{[k]} + \nu^{[k]}) (\underline{\gamma}^{[k]}, \underline{\Sigma}^{[k-1]} \underline{\gamma}^{[k]})} \right)}$$

$$\eta_i^{[k]} = \left(\frac{D_{\underline{x}_i}^{[k]} + \nu^{[k]}}{\underline{\gamma}^{[k]}, \underline{\Sigma}^{[k-1]} \underline{\gamma}^{[k]}} \right)^{\frac{1}{2}} \frac{K_{\frac{\nu+d-2}{2}} \left(\sqrt{(D_{\underline{x}_i}^{[k]} + \nu^{[k]}) (\underline{\gamma}^{[k]}, \underline{\Sigma}^{[k-1]} \underline{\gamma}^{[k]})} \right)}{K_{\frac{\nu+d}{2}} \left(\sqrt{(D_{\underline{x}_i}^{[k]} + \nu^{[k]}) (\underline{\gamma}^{[k]}, \underline{\Sigma}^{[k-1]} \underline{\gamma}^{[k]})} \right)}$$

$$\xi_i^{[k]} = \frac{1}{2} \log \left(\frac{D_{\underline{x}_i}^{[k]} + \nu^{[k]}}{\underline{\gamma}^{[k]}, \underline{\Sigma}^{[k-1]} \underline{\gamma}^{[k]}} \right) + \frac{\frac{\partial}{\partial \alpha} K_{-\frac{\nu+d}{2} + \alpha} \left(\sqrt{(D_{\underline{x}_i}^{[k]} + \nu^{[k]}) (\underline{\gamma}^{[k]}, \underline{\Sigma}^{[k-1]} \underline{\gamma}^{[k]})} \right)}{K_{\frac{\nu+d}{2}} \left(\sqrt{(D_{\underline{x}_i}^{[k]} + \nu^{[k]}) (\underline{\gamma}^{[k]}, \underline{\Sigma}^{[k-1]} \underline{\gamma}^{[k]})} \right)} \Big|_{\alpha=0}$$

- o k.iterasyon'da aşağıdaki değerleri bul

$$\bar{\delta}^{[k]} = \frac{1}{n} \sum_{i=1}^n \delta_i^{[k]} \quad , \quad \bar{\eta}^{[k]} = \frac{1}{n} \sum_{i=1}^n \eta_i^{[k]} \quad , \quad \bar{\xi}^{[k]} = \frac{1}{n} \sum_{i=1}^n \xi_i^{[k]}$$

- M-Adımı

- o Parametrelerin yeni değerlerini bul

$$\underline{\gamma}^{[k+1]} = \frac{n^{-1} \sum_{i=1}^n \delta_i^{[k]} (\bar{x} - \underline{x}_i)}{\bar{\delta}^{[k]} \bar{\eta}^{[k]} - 1} \quad , \quad \underline{\mu}^{[k+1]} = \frac{n^{-1} \sum_{i=1}^n \delta_i^{[k]} \underline{x}_i - \underline{\gamma}^{[k+1]}}{\bar{\delta}^{[k]}}$$

$$\underline{\Sigma}^{[k+1]} = \frac{1}{n} \sum_{i=1}^n \delta_i^{[k]} (\underline{x}_i - \underline{\mu}^{[k+1]}) (\underline{x}_i - \underline{\mu}^{[k+1]})' - \bar{\eta}^{[k]} \underline{\gamma}^{[k+1]} \underline{\gamma}^{[k+1]}$$

o v 'nin yeni değerini aşağıdaki denklemi çözerek bul

$$-\psi\left(\frac{v}{2}\right) + \log\left(\frac{v}{2}\right) + 1 - \bar{\xi}^{[k]} - \bar{\delta}^{[k]} = 0$$

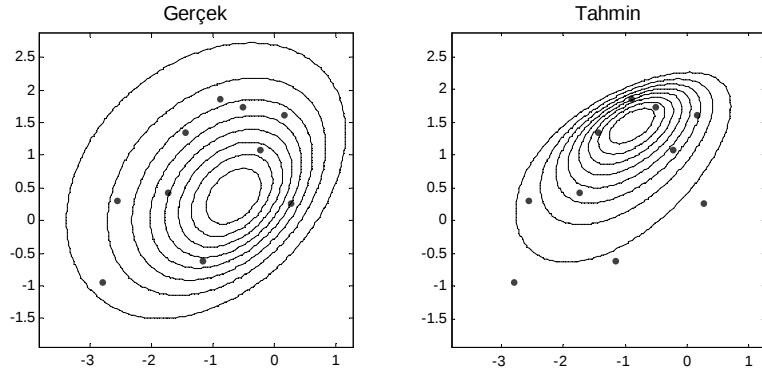
o Yakınsama sağlanmadıysa $k=k+1$ olarak E-Adımına dön.

İterasyonun başlangıç değerleri olarak $\underline{\mu}$ için örneklem ortalaması, $\underline{\Sigma}$ için örneklem Varyans-kovaryans matrisi ve $\underline{\gamma}$ için ise sıfıra çok yakın elemanları olan bir vektör seçilir. Serbestlik derecesi için ise değişik değerler kullanılabilir. Bu çalışmadaki uygulamalarda serbestlik derecesi 1 olarak belirlenmiştir. Bazı çalışmalarda v parametresi bilinen bir sabit olarak alınmakta, bu parametrenin küçük değerlerde seçimi dağılımın diğer parametreleri için daha robust tahminler elde etmeye olanak tanımaktadır.

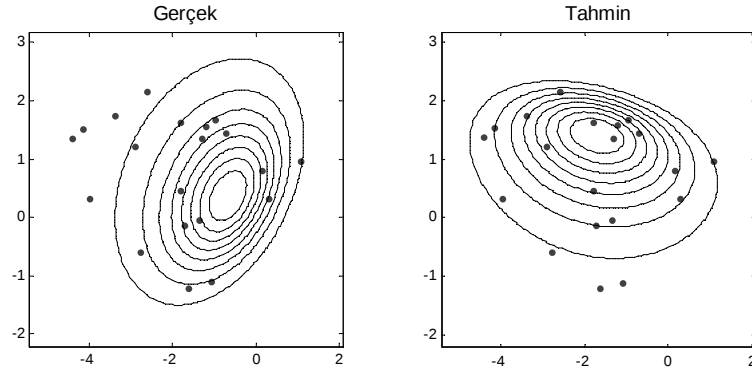
Şekil 3.5-3.9'da değişik hacimlerdeki örneklemlemlere ilişkin parametre tahminleri ile bu tahmin değerleri ve gerçek parametre değerleri kullanılarak çizilmiş olasılık yoğunluk fonksiyonlarının izdüşüm grafikleri görülmektedir. Gerçek parametreler için

$$v = 6 \quad \underline{\mu} = [0; 0] \quad \underline{\gamma} = [-0.9; 0.5] \quad \underline{\Sigma} = [1, 0.5; 0.5, 1]$$

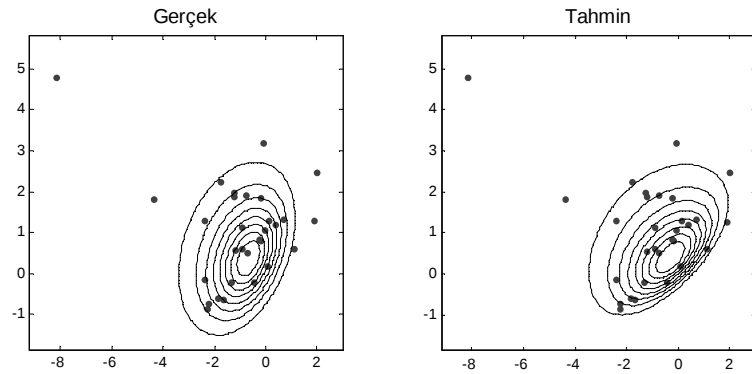
değerleri seçilmiştir.



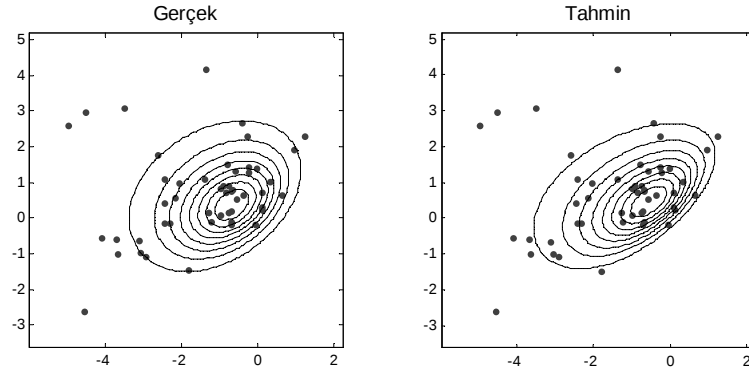
Şekil 3.5 $n=10$ örnek ile çarpık-t dağılımının parametrelerinin EM tahmini
 $\hat{\nu} = 9.29$, $\hat{\underline{\mu}} = [-0.5; 2.7]$, $\hat{\underline{\gamma}} = [-0.4; -1.6]$, $\hat{\underline{\Sigma}} = [0.7, 0.2; 0.2, 0.1]$



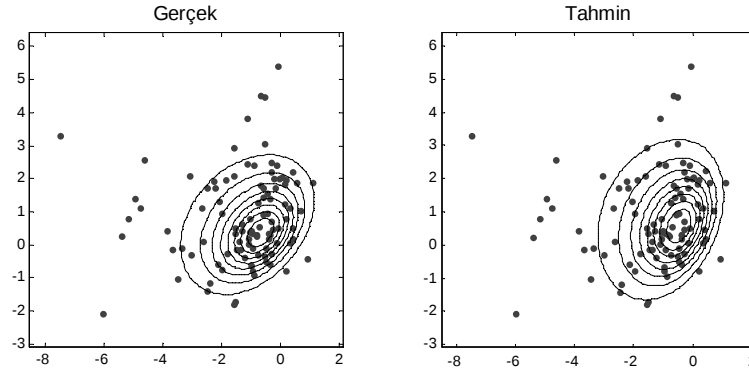
Şekil 3.6 $n=20$ örnek ile çarpık-t dağılımının parametrelerinin EM tahmini
 $\hat{\nu} = 13.68$, $\hat{\underline{\mu}} = [-1.4; 3.0]$, $\hat{\underline{\gamma}} = [-0.4; -1.9]$, $\hat{\underline{\Sigma}} = [2.1, -0.3; -0.3, 0.1]$



Şekil 3.7 $n=30$ örnek ile çarpık-t dağılımının parametrelerinin EM tahmini
 $\hat{\nu} = 8.89$, $\hat{\underline{\mu}} = [0.7; -0.5]$, $\hat{\underline{\gamma}} = [-1.3; 1.2]$, $\hat{\underline{\Sigma}} = [1.5, 0.9; 0.9, 0.6]$



Şekil 3.8 n=50 örnek ile çarpık-t dağılımının parametrelerinin EM tahmini
 $\hat{v} = 8.37$, $\hat{\underline{\mu}} = [0.7; -0.1]$, $\hat{\underline{\gamma}} = [-1.6; 0.5]$, $\hat{\underline{\Sigma}} = [0.9, 0.7; 0.7, 0.9]$



Şekil 3.9 n=100 örnek ile çarpık-t dağılımının parametrelerinin EM tahmini
 $\hat{v} = 7.87$, $\hat{\underline{\mu}} = [0.7; 0.2]$, $\hat{\underline{\gamma}} = [-1.5; 0.5]$, $\hat{\underline{\Sigma}} = [0.6, 0.5; 0.5, 1.4]$

Grafiklerden görülebileceği üzere parametrelerin tahmin değerleri ile çizilen konturların, özellikle 30 birimden daha düşük örneklem hacimlerinde gerçek parametrelerden daha iyi sonuç verdiği anlaşılmaktadır. Söz konusu durum parametre tahminleriyle ve gerçek parametrelerle hesaplanan $E(\underline{X})$ ve $Cov(\underline{X})$ değerlerinde de etkisini göstermektedir. Örneğin 20 birimlik örneklem ile örneklem ortalaması $[-1.8; 0.8]$, örneklem varyans-kovaryans matrisi $[2.2, -0.2; -0.2, 1.0]$ olarak gözlemlendiğinde, parametrelerin tahminleri ile beklenen değer $[-1.8; 0.8]$, varyans-kovaryans matrisi $[2.5, -0.2; -0.2, 1.1]$ hesaplanırken, gerçek parametre değerleri ile beklenen değer $[-1.4; 0.8]$, varyans-kovaryans matrisi $[3.3, -0.3; -0.26, 2.1]$ olarak hesaplanmıştır.

3.3 Çok Değişkenli Genelleştirilmiş Hiperbolik Çarpık-t Dağılımına Sahip Kitlelerde Diskriminant Fonksiyonu

Bu bölümde çarpık-t dağılımına sahip iki kitle için atama kuralının oluşturulması üzerinde durulacaktır. Π_1 ve Π_2 sırasıyla $Skew-t_d(v_1, \underline{\mu}_1, \underline{\Sigma}_1, \underline{\gamma}_1)$ ve $Skew-t_d(v_2, \underline{\mu}_2, \underline{\Sigma}_2, \underline{\gamma}_2)$ dağılımlı kitleler olsunlar. Bu kitleler için ($i=1,2$ olmak üzere)

$$f_{\underline{X}_i}(\underline{x}) = \frac{2^{1-(v_i+d)/2} \left(1 + \frac{1}{v_i} D_{\underline{x},i}\right)^{-\frac{v_i+d}{2}} K_{\frac{v_i+d}{2}} \left[\sqrt{(v_i + D_{\underline{x},i}) (\underline{\gamma}_i' \underline{\Sigma}_i^{-1} \underline{\gamma}_i)} \right]}{\Gamma\left(\frac{v_i}{2}\right) (\pi v_i)^{d/2} |\underline{\Sigma}_i|^{1/2} \left[\sqrt{(v_i + D_{\underline{x},i}) (\underline{\gamma}_i' \underline{\Sigma}_i^{-1} \underline{\gamma}_i)} \right]^{\frac{v_i+d}{2}}} \times \exp\left\{(\underline{x} - \underline{\mu}_i)' \underline{\Sigma}_i^{-1} \underline{\gamma}_i\right\} \quad (3.6)$$

yazılabilir. Bu iki kitlenin birisinden geldiği bilinen bir $\underline{x} \in \mathbb{R}^d$ gözlemi için atama kuralı

$$\begin{cases} E\ddot{g}er & f_{\underline{X}_1}(\underline{x})/f_{\underline{X}_2}(\underline{x}) \geq 1 & ise, & \underline{x} \in \Pi_1 \\ E\ddot{g}er & f_{\underline{X}_1}(\underline{x})/f_{\underline{X}_2}(\underline{x}) < 1 & ise, & \underline{x} \in \Pi_2 \end{cases} \quad (3.7)$$

biçiminde olup $f_{\underline{X}_1}(\underline{x})/f_{\underline{X}_2}(\underline{x})$ oranı

$$\begin{aligned} \frac{f_{\underline{X}_1}(\underline{x})}{f_{\underline{X}_2}(\underline{x})} &= 2^{\frac{v_2-v_1}{2}} \left(\frac{v_2}{v_1}\right)^{d/2} \frac{\sqrt{|\underline{\Sigma}_2|} \Gamma(v_2/2)}{\sqrt{|\underline{\Sigma}_1|} \Gamma(v_1/2)} \frac{K_{\frac{v_1+d}{2}} \left[\sqrt{(v_1 + D_{\underline{x},1}) (\underline{\gamma}_1' \underline{\Sigma}_1^{-1} \underline{\gamma}_1)} \right]}{K_{\frac{v_2+d}{2}} \left[\sqrt{(v_2 + D_{\underline{x},2}) (\underline{\gamma}_2' \underline{\Sigma}_2^{-1} \underline{\gamma}_2)} \right]} \\ &\quad \times \frac{\left[\sqrt{(v_2 + D_{\underline{x},2}) (\underline{\gamma}_2' \underline{\Sigma}_2^{-1} \underline{\gamma}_2)} \right]^{-\frac{v_2+d}{2}} \left(1 + \frac{1}{v_2} D_{\underline{x},2}\right)^{\frac{v_2+d}{2}}}{\left[\sqrt{(v_1 + D_{\underline{x},1}) (\underline{\gamma}_1' \underline{\Sigma}_1^{-1} \underline{\gamma}_1)} \right]^{-\frac{v_1+d}{2}} \left(1 + \frac{1}{v_1} D_{\underline{x},2}\right)^{\frac{v_1+d}{2}}} \\ &\quad \times \exp\left\{\underline{x}' (\underline{\Sigma}_1^{-1} \underline{\gamma}_1 - \underline{\Sigma}_2^{-1} \underline{\gamma}_2) + \underline{\mu}_2' \underline{\Sigma}_2^{-1} \underline{\gamma}_2 - \underline{\mu}_1' \underline{\Sigma}_1^{-1} \underline{\gamma}_1\right\} \end{aligned} \quad (3.8)$$

olarak elde edilir. Buradan diskriminant fonksiyonu

$$d_d^{Skew-t}(\underline{x}) = \frac{K_{\frac{v_1+d}{2}} \left[\sqrt{(v_1 + D_{\underline{x},1})} (\underline{\gamma}_1' \underline{\Sigma}_1^{-1} \underline{\gamma}_1) \right]}{K_{\frac{v_2+d}{2}} \left[\sqrt{(v_2 + D_{\underline{x},2})} (\underline{\gamma}_2' \underline{\Sigma}_2^{-1} \underline{\gamma}_2) \right]} \frac{\left[\sqrt{(v_2 + D_{\underline{x},2})} (\underline{\gamma}_2' \underline{\Sigma}_2^{-1} \underline{\gamma}_2) \right]^{\frac{v_2+d}{2}}}{\left[\sqrt{(v_1 + D_{\underline{x},1})} (\underline{\gamma}_1' \underline{\Sigma}_1^{-1} \underline{\gamma}_1) \right]^{\frac{v_1+d}{2}}} \quad (3.9)$$

$$\times \frac{\left(1 + 1/v_2 D_{\underline{x},2}\right)^{\frac{v_2+d}{2}}}{\left(1 + 1/v_1 D_{\underline{x},2}\right)^{\frac{v_1+d}{2}}} \exp\left\{\underline{x} \left(\underline{\Sigma}_1^{-1} \underline{\gamma}_1 - \underline{\Sigma}_2^{-1} \underline{\gamma}_2\right)\right\}$$

olarak yazılabilir. Böylece atama kuralı

$$\left\{ \begin{array}{l} \text{Eğer } d_d^{Skew-t}(\underline{x}) \geq 2^{\frac{v_1-v_2}{2}} \left(\frac{v_1}{v_2}\right)^{d/2} \frac{\sqrt{|\underline{\Sigma}_1|}}{\sqrt{|\underline{\Sigma}_2|}} \frac{\Gamma(v_1/2)}{\Gamma(v_2/2)} e^{\underline{\mu}_1' \underline{\Sigma}_1^{-1} \underline{\gamma}_1 - \underline{\mu}_2' \underline{\Sigma}_2^{-1} \underline{\gamma}_2} \text{ ise } \underline{x} \in \Pi_1 \\ \text{Eğer } d_d^{Skew-t}(\underline{x}) < 2^{\frac{v_1-v_2}{2}} \left(\frac{v_1}{v_2}\right)^{d/2} \frac{\sqrt{|\underline{\Sigma}_1|}}{\sqrt{|\underline{\Sigma}_2|}} \frac{\Gamma(v_1/2)}{\Gamma(v_2/2)} e^{\underline{\mu}_1' \underline{\Sigma}_1^{-1} \underline{\gamma}_1 - \underline{\mu}_2' \underline{\Sigma}_2^{-1} \underline{\gamma}_2} \text{ ise } \underline{x} \in \Pi_2 \end{array} \right. \quad (3.10)$$

olarak oluşturulabilir. $v_1 = v_2$, $\underline{\gamma}_1 = \underline{\gamma}_2$ ve $\underline{\Sigma}_1 = \underline{\Sigma}_2$ olması durumunda ise

$$\frac{f_{\underline{\gamma}_1}(\underline{x})}{f_{\underline{\gamma}_2}(\underline{x})} = \frac{K_{\frac{v+d}{2}} \left[\sqrt{(v + D_{\underline{x},1})} (\underline{\gamma}' \underline{\Sigma}^{-1} \underline{\gamma}) \right]}{K_{\frac{v+d}{2}} \left[\sqrt{(v + D_{\underline{x},2})} (\underline{\gamma}' \underline{\Sigma}^{-1} \underline{\gamma}) \right]} \frac{\left[\sqrt{(v_2 + D_{\underline{x},2})} \right]^{\frac{v+d}{2}}}{\left[\sqrt{(v_1 + D_{\underline{x},1})} \right]^{\frac{v+d}{2}}} \times \exp\left\{(\underline{\mu}_2 - \underline{\mu}_1)' \underline{\Sigma}^{-1} \underline{\gamma}\right\} \quad (3.11)$$

biçiminde olacaktır. Bu durumda diskriminant fonksiyonu

$$d_s^{Skew-t}(\underline{x}) = \log \left\{ \frac{K_{\frac{v+d}{2}} \left[\sqrt{(v + D_{\underline{x},1})} (\underline{\gamma}' \underline{\Sigma}^{-1} \underline{\gamma}) \right]}{K_{\frac{v+d}{2}} \left[\sqrt{(v + D_{\underline{x},2})} (\underline{\gamma}' \underline{\Sigma}^{-1} \underline{\gamma}) \right]} \left(\frac{1 + \frac{1}{v} D_{\underline{x},2}}{1 + \frac{1}{v} D_{\underline{x},1}} \right)^{-\frac{3}{4}(v+d)} \right\} \quad (3.12)$$

biçimindedir. Böylece bir $\underline{x} \in \mathbb{R}^d$ gözlemi için atama kuralı

$$\begin{cases} \text{Eğer } d_s^{Skew-t}(\underline{x}) \geq (\underline{\mu}_1 - \underline{\mu}_2)' \underline{\Sigma}^{-1} \underline{\gamma} & \underline{x} \in \Pi_1 \\ \text{Eğer } d_s^{Skew-t}(\underline{x}) < (\underline{\mu}_1 - \underline{\mu}_2)' \underline{\Sigma}^{-1} \underline{\gamma} & \underline{x} \in \Pi_2 \end{cases} \quad (3.13)$$

halini alacaktır.

3.4 Çarpık-t Dağılımı İçin Simülasyon

Simülasyon çalışmasında farklı örneklem hacimlerinde $Skew-t_2(v, \underline{\mu}, \underline{\Sigma}, \underline{\gamma})$ dağılımından sayı üretip, bu rasgele örnekler ile bir önceki bölümde ele alınan atama kuralları ile sınıflandırma yapılmıştır. Bu işlem her bir örnek hacminde 100 kez tekrarlanıp, gözlenen hatalı sınıflandırma oranlarının ortalaması alınmıştır.

Çizelge 3.1’de parametrelerin farklı olduğu durum için (3.10)’ile verilen atama kuralı kullanılarak elde edilen hatalı sınıflandırma oranları görülmektedir. Çizelgenin son iki sütununda (3.10)’ile verilen atama kuralının, sırasıyla gerçek parametre değerleri kullanılarak ve parametrelerin tahminleri kullanılarak elde edilmiş hatalı sınıflandırma oranları görülmektedir. Çizelgenin ilk iki sütununda ise verilerin normal dağıldığı varsayımı ile (1.10) ve (1.12) ile verilen atama kurallarının hatalı sınıflandırma oranları verilmiştir. Bu simülasyon çalışmasında iki boyutlu rasgele vektörün dağılımı birinci kitle için $Skew-t_2(v_1=5, \underline{\mu}_1=[2;4], \underline{\Sigma}_1=[1,0.9;0.9,1], \underline{\gamma}_1=[0.9;0.9])$, ikinci kitle için $Skew-t_2(v_2=9, \underline{\mu}_2=[5;5], \underline{\Sigma}_2=[1,-0.9;-0.9,1], \underline{\gamma}_2=[-0.9;0.1])$ alınmıştır.

Çizelge 3.1 Çarpık-t dağılımı için hatalı sınıflandırma oranları

n	Lineer	Karesel	Skew-t (Gerçek par. ile)	Skew-t (Tahmini par. ile)
10	0.31	0.13	0.13	0.11
20	0.35	0.13	0.12	0.10
30	0.35	0.14	0.13	0.11
50	0.36	0.13	0.11	0.10
100	0.32	0.14	0.13	0.13

Çizelge 3.1’den görülebileceği, tahmini parametreler kullanılarak (3.10)’ile verilen atama kuralına göre sınıflandırma yapıldığında lineer ve karesel diskriminant fonksiyonlarından daha iyi hatalı sınıflandırma oranı elde edildiği görülmektedir.

Çizelge 3.2’de $v=5, \Sigma = [1, 0.1; 0.1, 1], \gamma = [0.9; 0.1]$ olmak üzere birinci kitlenin dağılımı $Skew-t_2(v, \underline{\mu}_1 = [2; 2], \Sigma, \gamma)$, ikinci kitlenin dağılımı ise $Skew-t_2(v, \underline{\mu}_2 = [0; 0], \Sigma, \gamma)$ olarak seçilmiştir. Çizelgenin son iki sütununda (3.13)’ile verilen atama kuralının, sırasıyla gerçek parametre değerleri kullanılarak ve parametrelerin tahminleri kullanılarak elde edilmiş hatalı sınıflandırma oranları görülmektedir.

Çizelge 3.2 Çarpık-t dağılımı için $\underline{\mu}_1$ ve $\underline{\mu}_2$ hariç aynı parametreler ile hatalı sınıflandırma oranları

n	Lineer	Karesel	Skew-t (Gerçek par. ile)	Skew-t (Tahmini par. ile)
10	0.13	0.12	0.14	0.10
20	0.15	0.16	0.14	0.13
30	0.13	0.13	0.13	0.11
50	0.15	0.15	0.14	0.13
100	0.16	0.16	0.14	0.13

Parametrelerin eşit alındığı durum (konum parametresi hariç) ile tüm parametrelerin farklı olduğu durum için yapılan simülasyon çalışmaları karşılaştırılacak olursa d_d^{Skew-t} diskriminant fonksiyonunun, lineer ve karesel diskriminant fonksiyonlara göre başarısının, d_s^{Skew-t} diskriminant fonksiyonundan daha iyi olduğu görülmektedir. Bu durum uygulamada d_d^{Skew-t} ’nin d_s^{Skew-t} ’den daha başarılı olacağını göstermektedir. Bu noktada her iki örnekte de parametre değerlerinin, gözlemleri olabildiğince ayrımı zor hale getirecek şekilde seçildiği belirtilmelidir. Ayrımı kolay olacak şekilde seçilen parametre değerleri ile yapılacak sınıflandırmada d_d^{Skew-t} ve d_s^{Skew-t} ’nin, lineer ve karesel diskriminant fonksiyonları ile hemen hemen aynı hatalı sınıflandırma oranlarına sahip oldukları gözlenmiştir. Bu durum için çalışmanın son kısmında verilmiş örneklerle bakılabilir.

4. VARYANS GAMMA DAĞILIMI

Varyans gamma dağılımı (genelleştirilmiş Laplace dağılımı ya da Bessel dağılımı), genelleştirilmiş hiperbolik dağılım ailesinin bir üyesidir. Ailenin diğer üyeleri gibi kuyruk davranışı bakımından normal dağılımdan daha kalın kuyruklu bir dağılımdır. Dağılım finansal literatürde ilk olarak Madan ve Seneta (1990) tarafından kullanılmıştır.

Daha önce bahsedildiği gibi $\underline{X} = \underline{\mu} + W\underline{\gamma} + \sqrt{W}\mathbf{A}\underline{Z}$ dönüşümü ile verilen $\underline{X}$ rasgele vektörü $\underline{\mu}, \underline{\gamma} \in \mathbb{R}^d$, $\underline{Z} \sim N_d(\underline{0}, \mathbf{I})$ ve $\mathbf{A}\mathbf{A}' = \underline{\Sigma}$ olmak üzere eğer $W \sim N^-(\lambda, \chi, \psi)$ dağılırsa $\underline{X} \sim GH_d(\underline{\mu}, \underline{\Sigma}, \underline{\gamma}, \lambda, \chi, \psi)$ olarak dağılır. Burada W ve $\underline{X}$ 'in dağılımları sırasıyla (2.13) ve (2.24)'deki gibidir. W 'nun dağılımında $\lambda > 0$ ve $\chi = 0$ parametre değerleri alınırsa $\underline{X}$ rasgele vektörünün dağılımı 'varyans gamma' olur ve

$$c = \frac{\psi^\lambda (\psi + \underline{\gamma}'\underline{\Sigma}^{-1}\underline{\gamma})}{(2\pi)^{d/2} |\underline{\Sigma}|^{1/2} \Gamma(\lambda) 2^{\lambda-1}}$$

olmak üzere

$$f_{\underline{X}}(\underline{x}) = c \frac{K_{\lambda-d/2} \left(\sqrt{D_{\underline{x}}(\psi + \underline{\gamma}'\underline{\Sigma}^{-1}\underline{\gamma})} \right)}{\sqrt{D_{\underline{x}}(\psi + \underline{\gamma}'\underline{\Sigma}^{-1}\underline{\gamma})}^{d/2-\lambda}} \exp \left\{ (\underline{x} - \underline{\mu})' \underline{\Sigma}^{-1} \underline{\gamma} \right\} \quad (4.1)$$

olasılık yoğunluk fonksiyonuna sahiptir. Buna göre $\underline{X}$ rasgele vektörü (4.1) ile verilen olasılık yoğunluk fonksiyonuna sahip olduğunda $\underline{X} \sim VG_d(\underline{\mu}, \underline{\Sigma}, \underline{\gamma}, \lambda, \psi)$ ile gösterilecektir.

$\lambda > 0$ ve $\chi = 0$ parametre değerleri W 'nun dağılımında yerlerine yazılırsa $W \sim N^-(\lambda, \chi = 0, \psi)$ dağılımının olasılık yoğunluk fonksiyonu

$$\begin{aligned}
f_w(w; \lambda, \chi, \psi) &= \frac{\chi^{-\lambda} (\sqrt{\chi\psi})^\lambda}{2K_\lambda(\sqrt{\chi\psi})} w^{\lambda-1} \exp\left(-\frac{1}{2}\left(\chi \frac{1}{w} + \psi w\right)\right), w > 0, \chi \rightarrow 0 \\
&= \lim_{\chi \rightarrow 0} \frac{\chi^{-\lambda} (\sqrt{\chi\psi})^\lambda}{2K_\lambda(\sqrt{\chi\psi})} w^{\lambda-1} \exp\left(-\frac{1}{2}\left(\chi \frac{1}{w} + \psi w\right)\right) \\
&= \frac{\chi^{\frac{\lambda}{2}} \psi^{\frac{\lambda}{2}}}{\Gamma(\lambda) 2^\lambda \sqrt{\chi\psi}^{-\lambda}} w^{\lambda-1} \exp\left(-\frac{\psi}{2} w\right) \\
&= \frac{(\psi/2)^\lambda}{\Gamma(\lambda)} w^{\lambda-1} \exp\left(-\frac{\psi}{2} w\right)
\end{aligned}$$

biçiminde elde edilir. Böylece $W \sim \text{Gamma}(\lambda, \psi/2)$ olduğu görülür.

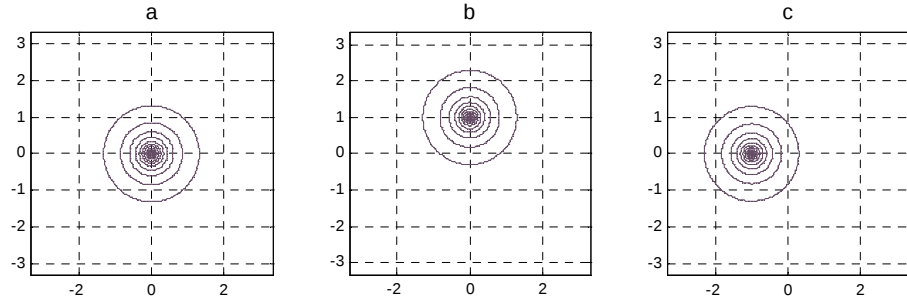
$W \sim \text{Gamma}(\lambda, \psi/2)$ olmak üzere $E(W) = 2\lambda/\psi$ ve $Var(W) = 4\lambda/\psi^2$ olduğundan (2.25) ve (2.26) den

$$E(\underline{X}) = \underline{\mu} + \frac{2\lambda}{\psi} \underline{\gamma} \quad (4.2)$$

$$Cov(\underline{X}) = \frac{2\lambda}{\psi} \underline{\Sigma} + \frac{4\lambda}{\psi^2} \underline{\gamma} \underline{\gamma}', \quad (4.3)$$

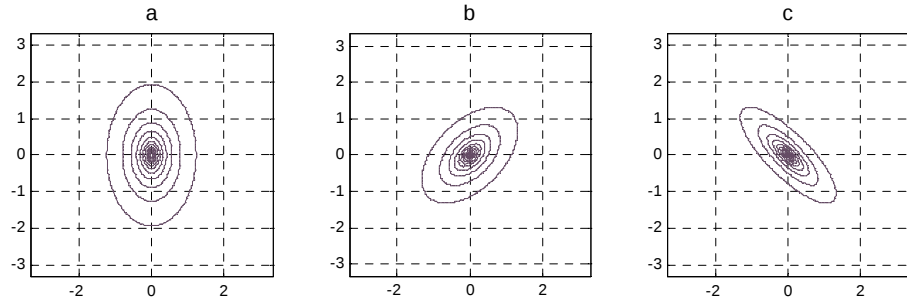
değerleri elde edilir.

$\underline{X} \sim VG_d(\underline{\mu}, \underline{\Sigma}, \underline{\gamma}, \lambda, \psi)$ dağılımı için olasılık yoğunluk fonksiyonlarının ve iz düşümlerinin grafiği $\underline{\mu}, \underline{\Sigma}, \underline{\gamma}, \lambda, \psi$ parametrelerinin değişik seçimleri ile şekil 4.1-4.4'de verilmiştir.



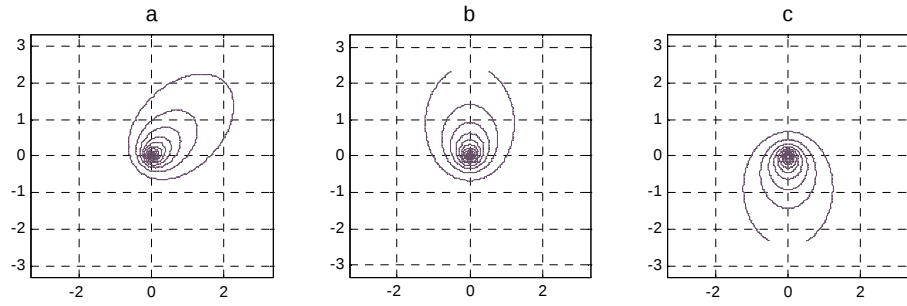
Şekil 4.1 Varyans gamma dağılımının $\underline{\mu}$ parametresine göre değişimi

a. $\underline{\mu} = [0 \ 0]'$, b. $\underline{\mu} = [0 \ 1]'$, c. $\underline{\mu} = [-1 \ 0]'$



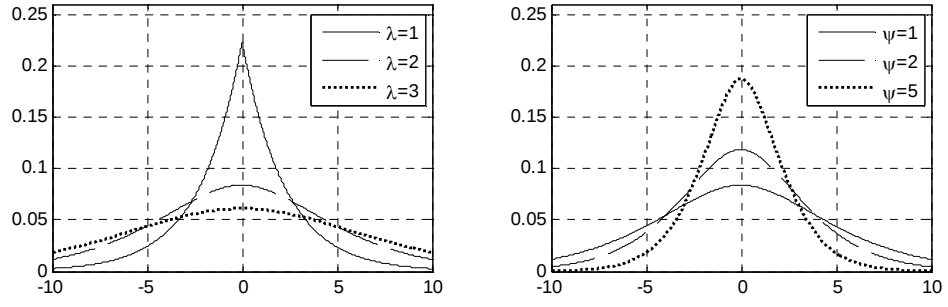
Şekil 4.2 Varyans gamma dağılımının Σ parametresine göre değişimi

a) $\Sigma = \begin{bmatrix} 1 & 0.01 \\ 0.01 & 4 \end{bmatrix}$ b) $\Sigma = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}$ c) $\Sigma = \begin{bmatrix} 1 & -0.9 \\ -0.9 & 1 \end{bmatrix}$



Şekil 4.3 Varyans gamma dağılımının $\underline{\gamma}$ parametresine göre değişimi

a. $\underline{\gamma} = [1 \ 1]'$, b. $\underline{\gamma} = [0 \ 1]'$, c. $\underline{\gamma} = [0 \ -1]'$



Şekil 4.4 Varyans gamma Dağılımının λ ve ψ parametresine göre değişimi

Grafiklerden görüleceği üzere $\underline{\mu}$ parametresi dağılımın konumu, Σ parametresi saçılımını, $\underline{\gamma}$ parametresi dağılımın çarpıklığını değiştirmektedir. λ ve ψ parametreleri ise kuyruk yüksekliğini ve dağılımın basıklığını değiştirmektedir. ψ parametresinin artması dağılımı daha sivri ve ince kuyruklu hale getirirken, λ parametresinin büyük değerleri dağılımı daha basık ve kalın kuyruklu yapmaktadır.

4.1 Varyans Gamma Dağılımından Rasgele Sayı Üretilmesi

Varyans gamma dağılımından sayı üretilmesi için aşağıdaki adımlar izlenir.

Girdiler: $\underline{\mu}_{d \times 1}, \Sigma_{d \times d}, \underline{\gamma}_{d \times 1}, \lambda, \psi$

1. Σ matrisinin Cholesky ayrışımından $\Sigma = \mathbf{A}\mathbf{A}^t$ olacak şekilde $\mathbf{A}_{d \times d}$ matrisini bul.
2. $\underline{Z} \sim N_k(\underline{0}, \mathbf{I}_d)$ dağılımından sayı üret.
3. $W \sim \text{Gamma}(\lambda, \psi/2)$ dağılımından sayı üret.
4. $\underline{X} = \underline{\mu} + W\underline{\gamma} + \sqrt{W}\mathbf{A}\underline{Z}$ değerini hesapla.

Çıktı: $\underline{X}$

4.2 Varyans Gamma Dağılımının Parametrelerinin Tahmini İçin EM-Algoritması

Genelleştirilmiş hiperbolik dağılım ailesinin parametreleri için daha önce EM-Algoritması verilmişti. Varyans gamma dağılımı için algorithmada $\chi = 0, \lambda > 0$ alınarak aşağıdaki güncellemeler yapılır. Bu noktada yine Bessel fonksiyonunun özelliklerinden yararlanılır.

λ parametresinin tahmini için (2.33)' de verilen log-olabilirlik fonksiyonu

$$L_2 = (\lambda - 1) \sum_{i=1}^n \log w_i - \frac{\chi}{2} \sum_{i=1}^n w_i^{-1} - \frac{\psi}{2} \sum_{i=1}^n w_i - n \log \left(\frac{2K_\lambda(\sqrt{\chi\psi})}{\chi^{-\lambda/2}\psi^{\lambda/2}} \right)$$

biçiminde yazılır ve $\chi = 0, \lambda > 0$ durumu için (2.3) özelliği kullanılırsa

$$L_2 = (\lambda - 1) \sum_{i=1}^n \log w_i - \frac{\psi}{2} \sum_{i=1}^n w_i - n \log(2^\lambda \Gamma(\lambda) \psi^{-\lambda})$$

elde edilir. ψ parametresi için bu ifade maksimum yapılırsa

$$\frac{\partial}{\partial \psi} L_2 = -\frac{1}{2} \sum_{i=1}^n w_i + \frac{n\lambda}{\psi} = 0$$

eşitliğinden ψ 'nin çekilmesiyle

$$\psi = \frac{2\lambda}{n^{-1} \sum_{i=1}^n w_i} \quad (4.4)$$

elde edilir. Aynı ifade λ 'ya göre türevlenip sıfıra eşitlenirse

$$n^{-1} \sum_{i=1}^n \log(w_i) + \log\left(n^{-1} \sum_{i=1}^n w_i\right) + \log(\lambda) - \psi(\lambda) = 0 \quad (4.5)$$

elde edilir. Son eşitlikte, eşitliği sağlayan λ değeri, λ 'nın yeni tahmini olarak alınır. (4.4) ve (4.5) eşitliklerinde (2.43)'da verilen beklenen değer ifadeleri kullanılarak EM-Algoritmasının bu parametrelerle ilgili adımı tamamlanmış olur. (Burada $\psi(\cdot)$ digamma fonksiyonunu göstermektedir.)

- E-Adımı

- o k.iterasyon'da tüm gözlemler için aşağıdakileri hesapla

$$\delta_i^{[k]} = \left(\frac{D_{\underline{x}_i}^{[k]}}{\psi^{[k]} + \underline{\gamma}^{[k]}, \Sigma^{[k]-1} \underline{\gamma}^{[k]}} \right)^{-\frac{1}{2}} \frac{K_{\lambda - \frac{d}{2} - 1} \left(\sqrt{D_{\underline{x}_i}^{[k]} (\psi^{[k]} + \underline{\gamma}^{[k]}, \Sigma^{[k]-1} \underline{\gamma}^{[k]})} \right)}{K_{\lambda - \frac{d}{2}} \left(\sqrt{D_{\underline{x}_i}^{[k]} (\psi^{[k]} + \underline{\gamma}^{[k]}, \Sigma^{[k]-1} \underline{\gamma}^{[k]})} \right)}$$

$$\eta_i^{[k]} = \left(\frac{D_{\underline{x}_i}^{[k]}}{\psi^{[k]} + \underline{\gamma}^{[k]}, \Sigma^{[k]-1} \underline{\gamma}^{[k]}} \right)^{\frac{1}{2}} \frac{K_{\lambda - \frac{d}{2} + 1} \left(\sqrt{D_{\underline{x}_i}^{[k]} (\psi^{[k]} + \underline{\gamma}^{[k]}, \Sigma^{[k]-1} \underline{\gamma}^{[k]})} \right)}{K_{\lambda - \frac{d}{2}} \left(\sqrt{D_{\underline{x}_i}^{[k]} (\psi^{[k]} + \underline{\gamma}^{[k]}, \Sigma^{[k]-1} \underline{\gamma}^{[k]})} \right)}$$

$$\xi_i^{[k]} = \frac{1}{2} \log \left(\frac{D_{\underline{x}_i}^{[k]}}{\psi^{[k]} + \underline{\gamma}^{[k]}, \Sigma^{[k]-1} \underline{\gamma}^{[k]}} \right) + \frac{\frac{\partial}{\partial \alpha} K_{\lambda - \frac{d}{2} + \alpha} \left(\sqrt{D_{\underline{x}_i}^{[k]} (\psi^{[k]} + \underline{\gamma}^{[k]}, \Sigma^{[k]-1} \underline{\gamma}^{[k]})} \right) \Big|_{\alpha=0}}{K_{\lambda - \frac{d}{2}} \left(\sqrt{D_{\underline{x}_i}^{[k]} (\psi^{[k]} + \underline{\gamma}^{[k]}, \Sigma^{[k]-1} \underline{\gamma}^{[k]})} \right)}$$

- o k.iterasyon'da aşağıdaki değerleri bul

$$\bar{\delta}^{[k]} = \frac{1}{n} \sum_{i=1}^n \delta_i^{[k]} \quad , \quad \bar{\eta}^{[k]} = \frac{1}{n} \sum_{i=1}^n \eta_i^{[k]} \quad , \quad \bar{\xi}^{[k]} = \frac{1}{n} \sum_{i=1}^n \xi_i^{[k]}$$

- M-Adımı

- o Parametrelerin yeni değerlerini bul

$$\underline{\gamma}^{[k+1]} = \frac{n^{-1} \sum_{i=1}^n \delta_i^{[k]} (\bar{\underline{x}} - \underline{x}_i)}{\bar{\delta}^{[k]} \bar{\eta}^{[k]} - 1} \quad , \quad \underline{\mu}^{[k+1]} = \frac{n^{-1} \sum_{i=1}^n \delta_i^{[k]} \underline{x}_i - \underline{\gamma}^{[k+1]}}{\bar{\delta}^{[k]}}$$

$$\Sigma^{[k+1]} = \frac{1}{n} \sum_{i=1}^n \delta_i^{[k]} (\underline{x}_i - \underline{\mu}^{[k+1]}) (\underline{x}_i - \underline{\mu}^{[k+1]})' - \bar{\eta}^{[k]} \underline{\gamma}^{[k+1]} \underline{\gamma}^{[k+1]}$$

- o λ 'nın yeni değeri $\lambda^{[k+1]}$ aşağıdaki eşitliği sağlayacak şekilde bul.

$$\log(\lambda) - \log(\eta^{[k]}) + \xi_i^{[k]} - \psi(\lambda) = 0$$

- o ψ 'nin yeni değerini bul.

$$\psi^{[k]} = 2\lambda^{[k+1]} / \eta^{[k]}$$

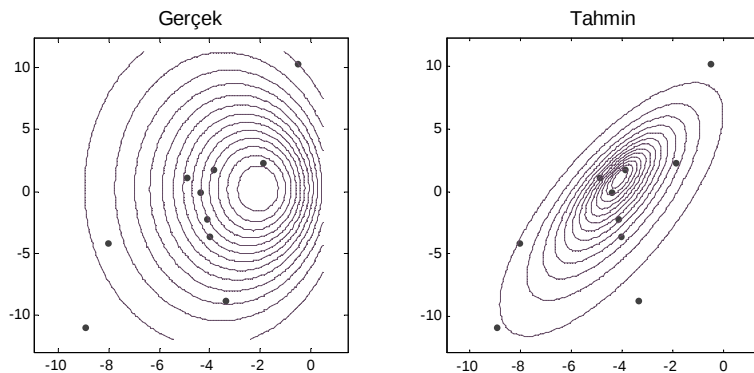
- o Yakınsama sağlanmadıysa $k=k+1$ olarak E-Adımına dön.

İterasyonun başlangıç değerleri olarak $\underline{\mu}$ için örneklem ortalaması, $\underline{\Sigma}$ için örneklem Varyans-kovaryans matrisi, $\underline{\gamma}$ için elemanları 1 olan vektör, λ ve ψ değerleri için ise 10 değeri alınmıştır.

Şekil 4.5-4.9'da $\underline{X} \sim VG_2(\underline{\mu}, \underline{\Sigma}, \underline{\gamma}, \lambda, \psi)$ dağılımından üretilmiş değişik hacimlerdeki örneklerden elde edilmiş $\underline{\mu}, \underline{\Sigma}, \underline{\gamma}, \lambda, \psi$ parametrelerinin tahminleri, gerçek parametreler ve tahminlerine göre çizilmiş kontur grafikleri görülmektedir. Parametre değerleri

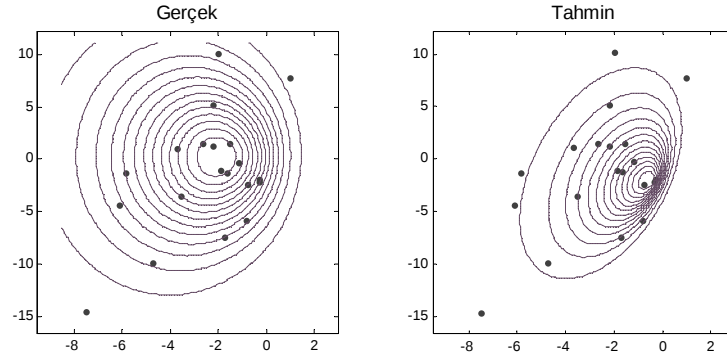
$$\underline{\mu} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \underline{\Sigma} = \begin{bmatrix} 1 & 0.1 \\ 0.1 & 1 \end{bmatrix}, \underline{\gamma} = \begin{bmatrix} -0.9 \\ 0.1 \end{bmatrix}, \lambda = 4, \psi = 2$$

olarak seçilmiştir.



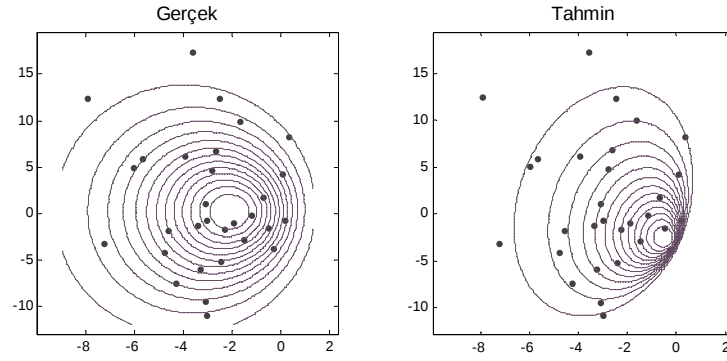
Şekil 4.5 n=10 örnek ile varyans gamma dağılımının parametrelerinin EM tahmini

$$\hat{\underline{\mu}} = \begin{bmatrix} -4.02 \\ 1.47 \end{bmatrix}, \hat{\underline{\Sigma}} = \begin{bmatrix} 3.16 & 5.32 \\ 5.32 & 15.6 \end{bmatrix}, \hat{\underline{\gamma}} = \begin{bmatrix} -0.20 \\ -1.49 \end{bmatrix}, \hat{\lambda} = 1.9, \hat{\psi} = 1.97$$



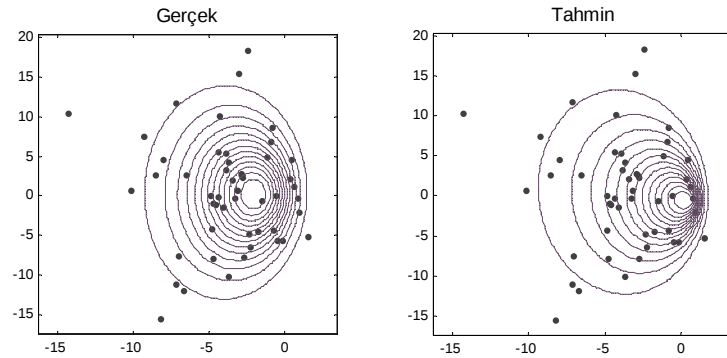
Şekil 4.6 n=20 örnek ile varyans gamma dağılımının parametrelerinin EM tahmini

$$\hat{\underline{\mu}} = \begin{bmatrix} -2.29 \\ -2.16 \end{bmatrix}, \hat{\underline{\Sigma}} = \begin{bmatrix} 1.03 & 2.57 \\ 2.57 & 16.07 \end{bmatrix}, \hat{\underline{\gamma}} = \begin{bmatrix} -1.16 \\ 0.42 \end{bmatrix}, \hat{\lambda} = 1.9, \hat{\psi} = 2.0$$



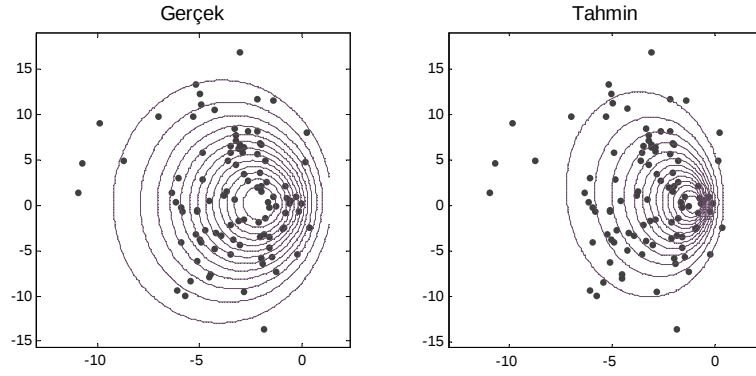
Şekil 4.7 n=30 örnek ile varyans gamma dağılımının parametrelerinin EM tahmini

$$\hat{\underline{\mu}} = \begin{bmatrix} -0.05 \\ -3.25 \end{bmatrix}, \hat{\underline{\Sigma}} = \begin{bmatrix} 0.89 & 3.93 \\ 3.93 & 22.62 \end{bmatrix}, \hat{\underline{\gamma}} = \begin{bmatrix} -1.63 \\ 2.52 \end{bmatrix}, \hat{\lambda} = 1.8, \hat{\psi} = 2.1$$



Şekil 4.8 n=50 örnek ile varyans gamma dağılımının parametrelerinin EM tahmini

$$\hat{\underline{\mu}} = \begin{bmatrix} 2.49 \\ 5.68 \end{bmatrix}, \hat{\underline{\Sigma}} = \begin{bmatrix} 1.68 & -0.56 \\ -0.56 & 0.77 \end{bmatrix}, \hat{\underline{\gamma}} = \begin{bmatrix} -3,14 \\ -3,17 \end{bmatrix}, \hat{\lambda} = 12.1, \hat{\psi} = 13.6$$



Şekil 4.9 n=100 örnek ile varyans gamma dağılımının parametrelerinin EM tahmini

$$\hat{\underline{\mu}} = \begin{bmatrix} -0.60 \\ -0.09 \end{bmatrix}, \hat{\underline{\Sigma}} = \begin{bmatrix} 1.23 & 0.12 \\ 0.12 & 20.88 \end{bmatrix}, \hat{\underline{\gamma}} = \begin{bmatrix} -1.49 \\ 0.71 \end{bmatrix}, \hat{\lambda} = 1.9, \hat{\psi} = 2.1$$

Grafiklerden, 30 birimden daha büyük örneklem hacimlerinde gerçek parametrelerle çizilen konturlar ile parametrelerin tahmini değerleriyle çizilen konturların benzemeye başladıkları görülmektedir. Ancak küçük örneklem hacimlerinde parametre tahminlerinin, üretilen veriye gerçek parametrelerden daha uygun olduğu görülmektedir. Söz konusu durum parametre tahminleriyle ve gerçek parametrelerle hesaplanan $E(\underline{X})$ ve $Cov(\underline{X})$ değerlerinde de etkisini göstermektedir. Örneğin 20 birimlik örneklem ile örneklem ortalaması $[-2.5; -1.4]$, örneklem varyans-kovaryans matrisi $[4.7, 6.4; 6.4, 32.1]$ olarak bulunduğu, parametrelerin tahminleri ile beklenen değer $[-2.5; -1.4]$, varyans-kovaryans matrisi $[4.6, 3.9; 3.9, 30.9]$ hesaplanırken, gerçek parametre değerleri ile beklenen değer $[-3.6; 0.4]$, varyans-kovaryans matrisi $[7.2, 0.04; 0.04, 40.0]$ olarak hesaplanmıştır.

4.3 Varyans Gamma Dağılımına Sahip Kitlelerde Diskriminant Fonksiyonu

Bu bölümde Varyans gamma dağılımına sahip iki kitle için atama kuralının oluşturulması üzerinde duracağız. Π_1 ve Π_2 sırasıyla $VG_d(\underline{\mu}_1, \underline{\Sigma}_1, \underline{\gamma}_1, \lambda_1, \psi_1)$ ve $VG_d(\underline{\mu}_2, \underline{\Sigma}_2, \underline{\gamma}_2, \lambda_2, \psi_2)$ dağılımına sahip kitleler olsun. Bu iki kitlenin birisinden geldiği bilinen bir $\underline{x} \in \mathbb{R}^d$ gözlemi için atama kuralı

$$\begin{cases} \text{Eğer } f_{\underline{x}_1}(\underline{x})/f_{\underline{x}_2}(\underline{x}) \geq 1 \text{ ise, } \underline{x} \in \Pi_1 \\ \text{Eğer } f_{\underline{x}_1}(\underline{x})/f_{\underline{x}_2}(\underline{x}) < 1 \text{ ise, } \underline{x} \in \Pi_2 \end{cases}$$

biçiminde yazılır. $f_{\underline{x}_1}(\underline{x})/f_{\underline{x}_2}(\underline{x})$ oranı

$$\begin{aligned} c &= 2^{\lambda_2 + \lambda_1} \sqrt{\frac{|\Sigma_2|}{|\Sigma_1|}} \frac{\Gamma(\lambda_2)}{\Gamma(\lambda_1)} \frac{\psi_1^{\lambda_1}(\psi_1 + \underline{\gamma}_1' \Sigma_1^{-1} \underline{\gamma}_1)}{\psi_2^{\lambda_2}(\psi_2 + \underline{\gamma}_2' \Sigma_2^{-1} \underline{\gamma}_2)} \\ \frac{f_{\underline{x}_1}(\underline{x})}{f_{\underline{x}_2}(\underline{x})} &= c \frac{K_{\lambda_1 - d/2} \left(\sqrt{D_{\underline{x},1}(\psi_1 + \underline{\gamma}_1' \Sigma_1^{-1} \underline{\gamma}_1)} \right) \sqrt{D_{\underline{x},2}(\psi_2 + \underline{\gamma}_2' \Sigma_2^{-1} \underline{\gamma}_2)}^{d/2 - \lambda_2}}{K_{\lambda_2 - d/2} \left(\sqrt{D_{\underline{x},2}(\psi_2 + \underline{\gamma}_2' \Sigma_2^{-1} \underline{\gamma}_2)} \right) \sqrt{D_{\underline{x},1}(\psi_1 + \underline{\gamma}_1' \Sigma_1^{-1} \underline{\gamma}_1)}^{d/2 - \lambda_1}} \\ &\quad \times \exp \left\{ (\underline{x} - \underline{\mu}_1)' \Sigma_1^{-1} \underline{\gamma}_1 - (\underline{x} - \underline{\mu}_2)' \Sigma_2^{-1} \underline{\gamma}_2 \right\} \end{aligned} \quad (4.6)$$

biçiminde yazılır. Böylece diskriminant fonksiyonu

$$\begin{aligned} d_d^{VG}(\underline{x}) &= \frac{K_{\lambda_1 - d/2} \left(\sqrt{D_{\underline{x},1}(\psi_1 + \underline{\gamma}_1' \Sigma_1^{-1} \underline{\gamma}_1)} \right) \sqrt{D_{\underline{x},2}(\psi_2 + \underline{\gamma}_2' \Sigma_2^{-1} \underline{\gamma}_2)}^{d/2 - \lambda_2}}{K_{\lambda_2 - d/2} \left(\sqrt{D_{\underline{x},2}(\psi_2 + \underline{\gamma}_2' \Sigma_2^{-1} \underline{\gamma}_2)} \right) \sqrt{D_{\underline{x},1}(\psi_1 + \underline{\gamma}_1' \Sigma_1^{-1} \underline{\gamma}_1)}^{d/2 - \lambda_1}} \\ &\quad \times \exp \left\{ (\underline{x} - \underline{\mu}_1)' \Sigma_1^{-1} \underline{\gamma}_1 - (\underline{x} - \underline{\mu}_2)' \Sigma_2^{-1} \underline{\gamma}_2 \right\} \end{aligned} \quad (4.7)$$

ve atama kuralı

$$\begin{cases} \text{Eğer } d_d^{VG}(\underline{x}) \geq 2^{\lambda_1 - \lambda_2} \sqrt{\frac{|\Sigma_1|}{|\Sigma_2|}} \frac{\Gamma(\lambda_1)}{\Gamma(\lambda_2)} \frac{\psi_2^{\lambda_2}(\psi_2 + \underline{\gamma}_2' \Sigma_2^{-1} \underline{\gamma}_2)}{\psi_1^{\lambda_1}(\psi_1 + \underline{\gamma}_1' \Sigma_1^{-1} \underline{\gamma}_1)} & , \underline{x} \in \Pi_1 \\ \text{Eğer } d_d^{VG}(\underline{x}) < 2^{\lambda_1 - \lambda_2} \sqrt{\frac{|\Sigma_1|}{|\Sigma_2|}} \frac{\Gamma(\lambda_1)}{\Gamma(\lambda_2)} \frac{\psi_2^{\lambda_2}(\psi_2 + \underline{\gamma}_2' \Sigma_2^{-1} \underline{\gamma}_2)}{\psi_1^{\lambda_1}(\psi_1 + \underline{\gamma}_1' \Sigma_1^{-1} \underline{\gamma}_1)} & , \underline{x} \in \Pi_2 \end{cases} \quad (4.8)$$

biçiminde elde edilir. $\Sigma_1 = \Sigma_2 = \Sigma$, $\underline{\gamma}_1 = \underline{\gamma}_2 = \underline{\gamma}$, $\lambda_1 = \lambda_2 = \lambda$ ve $\psi_1 = \psi_2 = \psi$ olması durumunda ise

$$\frac{f_{\underline{x}_1}(\underline{x})}{f_{\underline{x}_2}(\underline{x})} = \sqrt{\frac{D_{\underline{x},2}^{-d/2-\lambda}}{D_{\underline{x},1}}} \frac{K_{\lambda-d/2}\left(\sqrt{D_{\underline{x},1}}(\underline{\psi} + \underline{\gamma}'\underline{\Sigma}^{-1}\underline{\gamma})\right)}{K_{\lambda-d/2}\left(\sqrt{D_{\underline{x},2}}(\underline{\psi} + \underline{\gamma}'\underline{\Sigma}^{-1}\underline{\gamma})\right)} \exp\left\{\left(\underline{\mu}_2 - \underline{\mu}_1\right)'\underline{\Sigma}^{-1}\underline{\gamma}\right\}$$

olup diskriminant fonksiyonu

$$d_s^{VG}(\underline{x}) = \frac{d-2\lambda}{4}(\log D_{\underline{x},2} - \log D_{\underline{x},1}) + \log K_{\lambda-d/2}\left(\sqrt{D_{\underline{x},1}}(\underline{\psi} + \underline{\gamma}'\underline{\Sigma}^{-1}\underline{\gamma})\right) - \log K_{\lambda-d/2}\left(\sqrt{D_{\underline{x},2}}(\underline{\psi} + \underline{\gamma}'\underline{\Sigma}^{-1}\underline{\gamma})\right) \quad (4.9)$$

olarak elde edilir. Buradan logaritmik fonksiyonun monoton artan olması özelliğinden yararlanarak atama kuralı

$$\begin{cases} \text{Eğer} & d_s^{VG}(\underline{x}) \geq (\underline{\mu}_1 - \underline{\mu}_2)'\underline{\Sigma}^{-1}\underline{\gamma} & \underline{x} \in \Pi_1 \\ \text{Eğer} & d_s^{VG}(\underline{x}) < (\underline{\mu}_1 - \underline{\mu}_2)'\underline{\Sigma}^{-1}\underline{\gamma} & \underline{x} \in \Pi_2 \end{cases} \quad (4.10)$$

olarak oluşturulur.

4.4 Varyans Gamma Dağılımı İçin Simülasyon

Simülasyon çalışmasında farklı örneklem hacimlerinde $VG_2(\underline{\mu}, \underline{\Sigma}, \underline{\gamma}, \lambda, \underline{\psi})$ dağılımından sayı üretip, bu örnekler ile bir önceki bölümde ele alınan atama kuralları ile sınıflandırma yapılmıştır. Bu işlem her bir örnek hacminde 100 kez tekrarlanıp, gözlenen hatalı sınıflandırma oranlarının ortalaması alınmıştır.

Çizelge 4.1’de iki boyutlu rasgele vektörün dağılımı birinci kitle için $VG_2(\underline{\mu}_1 = [0; 0], \underline{\Sigma}_1 = [10, 1; 1, 10], \underline{\gamma}_1 = [1, -1], \lambda_1 = 3, \underline{\psi}_1 = 1)$, ikinci kitle için $VG_2(\underline{\mu}_2 = [6; -6], \underline{\Sigma}_2 = [20, 1; 1, 20], \underline{\gamma}_2 = [-1, 1], \lambda_2 = 1, \underline{\psi}_2 = 2)$ alınmıştır. Bu simülasyon çalışmasında birinci ve ikinci kitleden değişik hacimlerde veri üretilmiş, lineer diskriminant fonksiyonu ve karesel diskriminant fonksiyonu ile parametrelerin farklı

olduğu durum için (4.8) ile verilen atama kuralı, gerçek parametreler ve parametre tahminleri kullanılarak söz konusu verilerin sınıflandırması yapılmıştır. Tahmini parametreler kullanılarak (4.8) ile verilen atama kuralına göre yapılan sınıflandırmada lineer ve karesel diskriminant daha az hatalı sınıflandırma oranının elde edildiği görülmektedir.

Çizelge 4.1 Varyans gamma dağılımı için hatalı sınıflandırma oranları

n	Lineer	Karesel	VGamma (Gerçek par. ile)	VGamma (Tahmini par. ile)
10	0.40	0.25	0.29	0.21
20	0.44	0.30	0.30	0.26
30	0.45	0.30	0.28	0.26
50	0.45	0.31	0.29	0.27
100	0.48	0.32	0.28	0.27

$\underline{\mu}_1$ ve $\underline{\mu}_2$ hariç diğer parametrelerin aynı olduğu durum için yapılan simülasyon çalışmasında $\Sigma = [2, -0.001; -0.001, 2]$, $\underline{\gamma} = [0.9, -0.9]$, $\lambda = 10$, $\psi = 6$ olmak üzere birinci kitlenin dağılımı $VG_2(\underline{\mu}_1 = [0; 0], \Sigma, \underline{\gamma}, \lambda, \psi)$ ikinci kitlenin dağılımı ise $VG_2(\underline{\mu}_1 = [2; -2], \Sigma, \underline{\gamma}, \lambda, \psi)$ olarak alınmıştır. Varyans gamma dağılımında aynı parametreler için (4.10) ile verilen atama kuralına göre gerçek parametreler ve parametrelerin tahminleri kullanılarak sınıflandırma yapılmış, elde edilen hatalı sınıflandırma oranları ile lineer ve karesel diskriminant fonksiyonları kullanılarak yapılan sınıflandırmada elde edilen hatalı sınıflandırma oranları çizelge 4.2’de verilmiştir. Buna göre 50 ve daha az örneklem hacimlerinde en az hatalı sınıflandırma oranının tahmini parametreler kullanılarak (4.10) ile verilen atama kuralına göre sınıflandırma yapıldığında elde edildiği görülmektedir. 100 birimlik örneklem ile yapılan simülasyon çalışmasında ise karesel diskriminant fonksiyonu ile yapılan sınıflandırma ile tahmini parametreler kullanarak atama kuralı (4.10) ile sınıflandırma yapıldığında elde edilen hatalı sınıflandırma oranlarının aynı olduğu gözlenmiştir.

Çizelge 4.2 Varyans gamma dağılımı için $\underline{\mu}_1$ ve $\underline{\mu}_2$ hariç aynı parametreler ile hatalı sınıflandırma oranları

n	Lineer	Karesel	VGamma (Gerçek par. ile)	VGamma (Tahmini par. ile)
10	0.29	0.29	0.31	0.24
20	0.29	0.29	0.30	0.27
30	0.32	0.32	0.32	0.30
50	0.31	0.31	0.31	0.29
100	0.31	0.30	0.31	0.30

Çarpık-t dağılımı için yapılan simülasyon çalışmasındaki duruma paralel olarak (konum parametresi hariç) parametrelerin eşit alındığı durumda d_s^{VG} 'nin lineer ve karesel diskriminant fonksiyonlarına karşı başarısının, tüm parametrelerin farklı olduğu durumda d_d^{VG} 'ye göre daha düşük olduğu görülmektedir. Her iki örnekte de parametre değerleri, gözlemleri olabildiğince ayrımı zor hale getirecek şekilde seçilmiştir. Ayrımı kolay olacak şekilde seçilen parametre değerleri ile yapılacak sınıflandırmada ele alınan tüm diskriminant fonksiyonlarının hemen hemen aynı hatalı sınıflandırma oranlarına sahip oldukları gözlenmiştir. Bu durum için çalışmanın son kısmında verilmiş örneklere bakılabilir.

5. ÇARPIK LAPLACE DAĞILIMI

Klasik anlamda Laplace dağılımı ilk olarak Laplace(1774) tarafından verilmiştir. Bu dağılım çoğunlukla normal dağılımdan daha kalın kuyruklu verilerin modellenmesinde kullanılmıştır. Bu çalışmalara örnek olarak Esterling (1978), Hsu (1979), Damsleth ve El-Shaarawi (1989) verilebilir. Dağılımın çarpık durumu ise biyoloji, ekonomi, sigortacılık ve finans matematiği gibi bir çok alanda kullanılmakta olup daha geniş bilgi için Kalashnikov (1997), Kotz vd. (2001) ve Arslan (2010) çalışmalarına bakılabilir.

Çarpık Laplace dağılımı, önceki bölümde ele alınan varyans gamma dağılımının özel bir hali olup, genelleştirilmiş hiperbolik dağılım ailesinin bir üyesidir. Varyans gamma dağılımı gibi normal ortalama-varyans karma dağılımında karıştırıcı değişken W 'nin dağılımının gamma dağılması durumunda elde edilir. Çok değişkenli genelleştirilmiş hiperbolik dağılıma sahip bir $\underline{X}$ rasgele vektörü için daha önce (2.22) ile verilen dönüşümde, $W \sim N^-(\lambda, \chi, \psi)$ karıştırıcı rasgele değişkeninin parametreleri $\lambda = (d+1)/2 > 0$, $\chi = 0$ ve $\psi = 1$ olarak alınsın. W 'nin olasılık yoğunluk fonksiyonu önceki bölümdesine benzer olarak

$$f_W(w) = \frac{w^{\frac{d+1}{2}-1}}{2^{\frac{d+1}{2}} \Gamma\left(\frac{d+1}{2}\right)} \exp\left(-\frac{w}{2}\right), w > 0$$

elde edilir. Buradan görülebileceği üzere $W \sim \text{Gamma}(\alpha = (d+1)/2, \beta = 1/2)$ dağılımına sahip olacaktır. Bu durumda $\underline{X}$ rasgele vektörünün dağılımı ise alınan parametre değerleri için $\underline{X} \sim GH_d(\lambda = (d+1)/2, \chi = 0, \psi = 1, \underline{\mu}, \underline{\Sigma}, \underline{\gamma})$ olacaktır. $\underline{X}$ rasgele vektörünün olasılık yoğunluk fonksiyonunu elde etmek için parametre değerlerini yerlerine yazarsak (2.24)'de verilen sabit terim, Bessel fonksiyonu için (2.3) ile verilen özellik kullanılarak

$$c = \frac{\chi^{\frac{\lambda}{2}} (\underline{\gamma}' \underline{\Sigma}^{-1} \underline{\gamma} + 1)^{\frac{d}{2} \frac{d+1}{2}}}{(2\pi)^{d/2} |\underline{\Sigma}|^{1/2} \Gamma\left(\frac{d+1}{2}\right) 2^{\frac{d+1}{2}-1} \chi^{\frac{\lambda}{2}}} = \frac{(\underline{\gamma}' \underline{\Sigma}^{-1} \underline{\gamma} + 1)^{-\frac{1}{2}}}{2^{d-\frac{1}{2}} \pi^{\frac{d}{2}} |\underline{\Sigma}|^{\frac{1}{2}} \Gamma\left(\frac{d+1}{2}\right)}$$

olarak elde edilir. Benzer şekilde

$$f_{\underline{x}}(\underline{x}) = \frac{(\underline{\gamma}' \underline{\Sigma}^{-1} \underline{\gamma} + 1)^{-\frac{1}{2}}}{2^{d-\frac{1}{2}} \pi^{\frac{d}{2}} |\underline{\Sigma}|^{\frac{1}{2}} \Gamma\left(\frac{d+1}{2}\right)} \frac{K_{\frac{1}{2}}\left(\sqrt{D_{\underline{x}}} (\underline{\gamma}' \underline{\Sigma}^{-1} \underline{\gamma} + 1)\right)}{(D_{\underline{x}} (\underline{\gamma}' \underline{\Sigma}^{-1} \underline{\gamma} + 1))^{-\frac{1}{4}}} \exp\left\{(\underline{x} - \underline{\mu})' \underline{\Sigma}^{-1} \underline{\gamma}\right\}$$

elde edilir. $\alpha = (1 + \underline{\gamma}' \underline{\Sigma}^{-1} \underline{\gamma})^{1/2}$ ile gösterilir ve Bessel fonksiyonu için (2.6) ile verilen özel durum göz önüne alınırsa

$$f_{\underline{x}}(\underline{x}) = \frac{|\underline{\Sigma}|^{-\frac{1}{2}}}{2^d \pi^{\frac{d-1}{2}} \alpha \Gamma\left(\frac{d+1}{2}\right)} \exp\left\{-\alpha \sqrt{D_{\underline{x}}} + (\underline{x} - \underline{\mu})' \underline{\Sigma}^{-1} \underline{\gamma}\right\} \quad (5.1)$$

elde edilir. $\underline{X}$ rasgele değişkeninin olasılık yoğunluk fonksiyonu (5.1)'deki gibi olduğunda $\underline{X}$ 'e çok değişkenli çarpık Laplace dağılımına sahiptir denir ve $\underline{X} \sim MSL_d(\underline{\mu}, \underline{\Sigma}, \underline{\gamma})$ ile gösterilir.

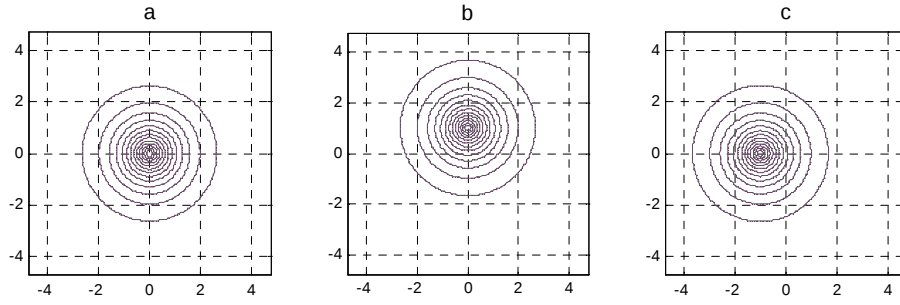
$W \sim Gamma((d+1)/2, 1/2)$ olmak üzere $E(W) = d+1$ ve $Var(W) = 2(d+1)$ olduğundan (2.25) ve (2.26) den $\underline{X}$ 'in beklenen değeri ve varyans-kovaryans matrisi için

$$E(\underline{X}) = \underline{\mu} + (d+1)\underline{\gamma} \quad (5.2)$$

$$Cov(\underline{X}) = (d+1)\underline{\Sigma} + 2(d+1)\underline{\gamma}\underline{\gamma}' \quad (5.3)$$

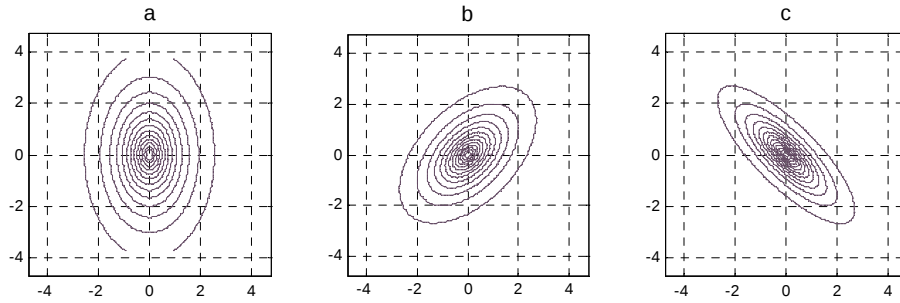
değerleri elde edilir.

$\underline{X} \sim MSL_2(\underline{\mu}, \underline{\Sigma}, \underline{\gamma})$ dağılımı için olasılık yoğunluk fonksiyonlarının ve izdüşümlerinin grafikleri $\underline{\mu}, \underline{\Sigma}, \underline{\gamma}$ parametrelerinin değişik seçimleri şekil 5.1-5.3'de görülmektedir.



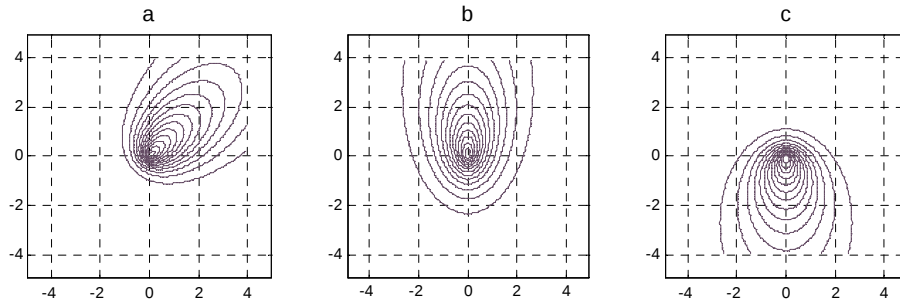
Şekil 5.1 Çarpık Laplace dağılımının $\underline{\mu}$ parametresine göre değişimi

a. $\underline{\mu} = [0 \ 0]'$, b. $\underline{\mu} = [0 \ 1]'$, c. $\underline{\mu} = [-1 \ 0]'$



Şekil 5.2 Çarpık Laplace dağılımının $\underline{\Sigma}$ parametresine göre değişimi

a. $\underline{\Sigma} = \begin{bmatrix} 1 & 0.01 \\ 0.01 & 4 \end{bmatrix}$ b. $\underline{\Sigma} = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}$ c. $\underline{\Sigma} = \begin{bmatrix} 1 & -0.9 \\ -0.9 & 1 \end{bmatrix}$



Şekil 5.3 Çarpık Laplace dağılımının $\underline{\gamma}$ parametresine göre değişimi

a) $\underline{\gamma} = [1 \ 1]'$, b) $\underline{\gamma} = [0 \ 1]'$, c) $\underline{\gamma} = [0 \ -1]'$

Grafiklerden görüleceği üzere $\underline{\mu}$ parametresi dağılımın konumunu, $\underline{\Sigma}$ parametresi saçılımını, $\underline{\gamma}$ parametresi dağılımın çarpıklığını değiştirmektedir.

5.1 Çarpık Laplace Dağılımından Rasgele Sayı Üretilmesi

Çarpık Laplace dağılımından sayı üretilmesi için aşağıdaki adımlar izlenir.

Girdiler: $\underline{\mu}_{d \times 1}, \underline{\Sigma}_{d \times d}, \underline{\gamma}_{d \times 1}, \lambda, \psi$

1. $\underline{\Sigma}$ matrisinin Cholesky ayrışımından $\underline{\Sigma} = \mathbf{A}\mathbf{A}^t$ olacak şekilde $\mathbf{A}_{d \times d}$ matrisini bul.
2. $\underline{Z} \sim N_k(\underline{0}, \mathbf{I}_d)$ dağılımından sayı üret.
3. $W \sim Gamma((d+1)/2, 1/2)$ dağılımından sayı üret.
4. $\underline{X} = \underline{\mu} + W\underline{\gamma} + \sqrt{W}\mathbf{A}\underline{Z}$ değerini hesapla.

Çıktı: $\underline{X}$

5.2 Çarpık Laplace Dağılımının Parametrelerinin Tahmini İçin EM-Algoritması

Varyans gamma dağılımının parametrelerinin tahmini için EM-Algoritması önceki bölümde verilmişti. Çarpık Laplace dağılımı için algorithmada $\lambda = (d+1)/2$, $\psi = 1$ alınarak aşağıdaki güncelleştirmeler yapılır. Bu noktada yine Bessel fonksiyonunun özelliklerinden yararlanılır.

- E-Adımı

o k.iterasyon'da tüm gözlemler için aşağıdakileri hesapla

$$\delta_i^{[k]} = \left(\frac{D_{x_i}^{[k]}}{1 + \underline{\gamma}^{[k]t} \underline{\Sigma}^{[k-1]} \underline{\gamma}^{[k]}} \right)^{-\frac{1}{2}}$$

$$\eta_i^{[k]} = \left(\frac{D_{\underline{x}_i}^{[k]}}{1 + \underline{\gamma}^{[k]'} \underline{\Sigma}^{[k]-1} \underline{\gamma}^{[k]}} \right)^{\frac{1}{2}} \frac{K_{\frac{3}{2}} \left(\sqrt{D_{\underline{x}_i}^{[k]} \left(1 + \underline{\gamma}^{[k]'} \underline{\Sigma}^{[k]-1} \underline{\gamma}^{[k]} \right)} \right)}{\left[D_{\underline{x}_i}^{[k]} \left(1 + \underline{\gamma}^{[k]'} \underline{\Sigma}^{[k]-1} \underline{\gamma}^{[k]} \right) \right]^{\frac{1}{4}} \sqrt{\pi/2}} \times \exp \left\{ \left[D_{\underline{x}_i}^{[k]} \left(1 + \underline{\gamma}^{[k]'} \underline{\Sigma}^{[k]-1} \underline{\gamma}^{[k]} \right) \right]^{\frac{1}{2}} \right\}$$

o k.iterasyon’da aşağıdaki değerleri bul

$$\bar{\delta}^{[k]} = \frac{1}{n} \sum_{i=1}^n \delta_i^{[k]}, \quad \bar{\eta}^{[k]} = \frac{1}{n} \sum_{i=1}^n \eta_i^{[k]}$$

• M-Adımı

o Parametrelerin yeni değerlerini bul

$$\underline{\gamma}^{[k+1]} = \frac{n^{-1} \sum_{i=1}^n \delta_i^{[k]} (\bar{\underline{x}} - \underline{x}_i)}{\bar{\delta}^{[k]} \bar{\eta}^{[k]} - 1}, \quad \underline{\mu}^{[k+1]} = \frac{n^{-1} \sum_{i=1}^n \delta_i^{[k]} \underline{x}_i - \underline{\gamma}^{[k+1]}}{\bar{\delta}^{[k]}}$$

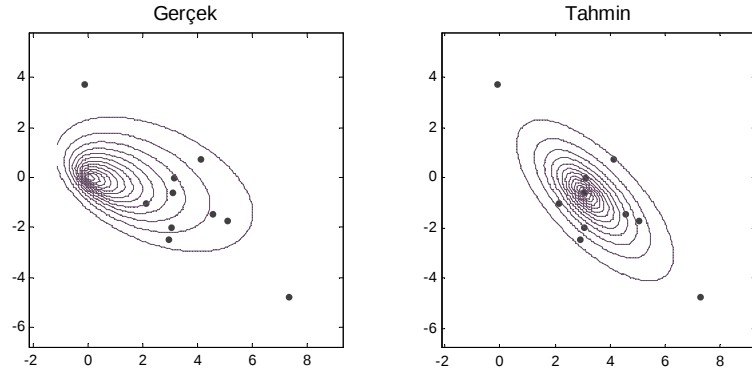
$$\underline{\Sigma}^{[k+1]} = \frac{1}{n} \sum_{i=1}^n \delta_i^{[k]} (\underline{x}_i - \underline{\mu}^{[k+1]})(\underline{x}_i - \underline{\mu}^{[k+1]})' - \bar{\eta}^{[k]} \underline{\gamma}^{[k+1]} \underline{\gamma}^{[k+1]}'$$

İterasyonun başlangıç değerleri olarak $\underline{\mu}$ için örneklem ortalaması, $\underline{\Sigma}$ için örneklem Varyans-kovaryans matrisi, $\underline{\gamma}$ için elemanları 1 olan vektör, λ ve ψ değerleri için ise 10 değeri alınmıştır.

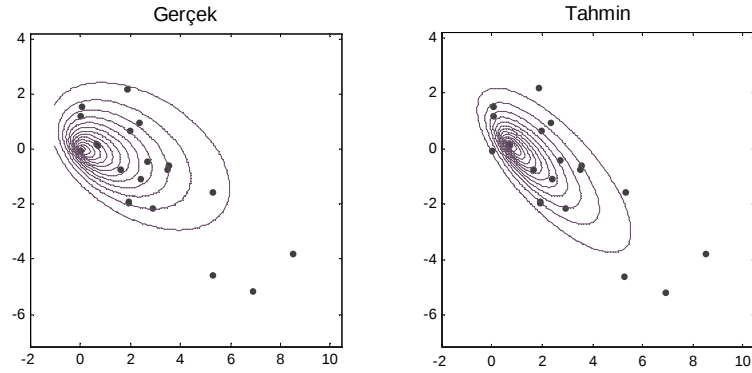
Şekil 5.4-5.8’de $\underline{X} \sim MSL_2(\underline{\mu}, \underline{\Sigma}, \underline{\gamma})$ dağılımından üretilmiş değişik hacimlerdeki örneklerden elde edilmiş $\underline{\mu}, \underline{\Sigma}, \underline{\gamma}$ parametrelerinin tahminleri, gerçek parametreler ve tahminlerine göre çizilmiş izdüşüm grafikleri görülmektedir. Parametre değerleri

$$\underline{\mu} = [0; 0], \quad \underline{\Sigma} = [1, -0.5; -0.5, 1], \quad \underline{\gamma} = [0.9; -0.1]$$

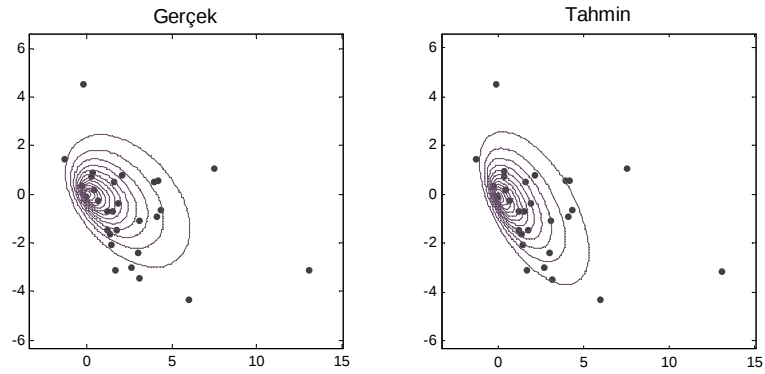
olarak seçilmiştir.



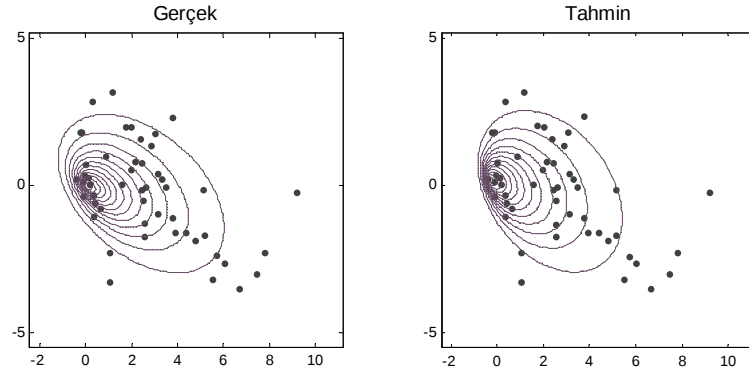
Şekil 5.4 n=10 örnek ile çarpık Laplace dağılımının parametrelerinin EM tahmini
 $\hat{\underline{\mu}} = [3.1; -0.6]$, $\hat{\underline{\Sigma}} = [1.01, -0.86; -0.86, 1.41]$, $\hat{\underline{\gamma}} = [0.15; -0.11]$



Şekil 5.5 n=20 örnek ile çarpık Laplace dağılımının parametrelerinin EM tahmini
 $\hat{\underline{\mu}} = [0.7; -0.2]$, $\hat{\underline{\Sigma}} = [0.84, -0.71; -0.71, 1.11]$, $\hat{\underline{\gamma}} = [0.68; -0.35]$

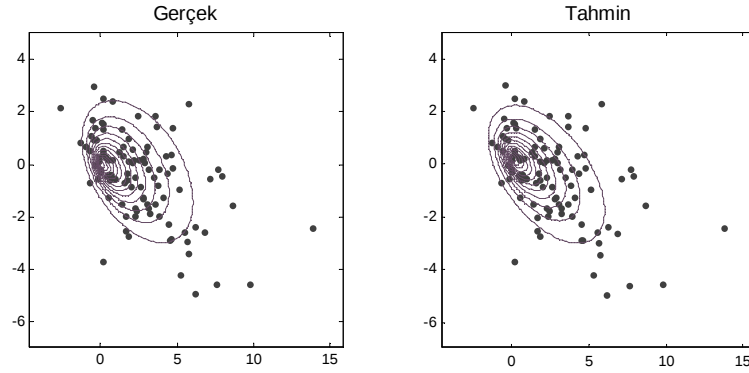


Şekil 5.6 n=30 örnek ile çarpık Laplace dağılımının parametrelerinin EM tahmini
 $\hat{\underline{\mu}} = [-0.04; -0.09]$, $\hat{\underline{\Sigma}} = [0.75, -0.66; -0.66, 1.32]$, $\hat{\underline{\gamma}} = [0.80; -0.18]$



Şekil 5.7 n=50 örnek ile çarpık Laplace dağılımının parametrelerinin EM tahmini.

$$\hat{\underline{\mu}} = [-0.4; -0.2], \hat{\underline{\Sigma}} = [0.33, -0.25; -0.25, 1.02], \hat{\underline{\gamma}} = [1.01; -0.15]$$



Şekil 5.8 n=100 örnek ile çarpık Laplace dağılımının parametrelerinin EM tahmini.

$$\hat{\underline{\mu}} = [-0.1; -0.3], \hat{\underline{\Sigma}} = [1.28, -0.54; -0.54, 0.91], \hat{\underline{\gamma}} = [0.87; -0.25]$$

Grafiklerden özellikle küçük örneklem hacimlerinde gerçek parametreler ile tahminleri arasındaki farklılığın yüksek olduğu görülmektedir. Buna karşın parametre tahminleri kullanılarak çizilen olasılık yoğunluk fonksiyonunun iz düşüm grafikleri her örneklem hacminde veriye daha iyi uymaktadır. Söz konusu durum parametre tahminleriyle ve gerçek parametrelerle hesaplanan $E(\underline{X})$ ve $Cov(\underline{X})$ değerlerinde de etkisini göstermektedir. Örneğin 10 birimlik örneklem ile örneklem ortalaması $[3.51; -0.95]$, örneklem varyans-kovaryans matrisi $[3.8, -3.4; -3.4, 4.9]$ olarak gözlemlendiğinde, parametrelerin tahminleri ile beklenen değer $[3.51; -0.95]$, varyans-kovaryans matrisi

$[3.4, -2.7; -2.7, 4.3]$ hesaplanırken, gerçek parametre değerleri ile beklenen değer $[2.7; -0.3]$, varyans-kovaryans matrisi $[7.9, -2.04; -2.04, 3.06]$ olarak hesaplanmıştır.

5.3 Çarpık Laplace Dağılımına Sahip Kitlelerde Diskriminant Fonksiyonu

Bu bölümde çarpık Laplace dağılımına sahip iki kitle için atama kuralının oluşturulması üzerinde duracağız. Π_1 ve Π_2 sırasıyla $MSL_d(\underline{\mu}_1, \underline{\Sigma}_1, \underline{\gamma}_1)$ ve $MSL_d(\underline{\mu}_2, \underline{\Sigma}_2, \underline{\gamma}_2)$ dağılımına sahip kitleler olsun. Bu iki kitlenin birisinden geldiği bilinen bir $\underline{x} \in \mathbb{R}^d$ gözlemi için atama kuralı

$$\begin{cases} \text{Eğer } f_{\underline{x}_1}(\underline{x})/f_{\underline{x}_2}(\underline{x}) \geq 1 \text{ ise, } \underline{x} \in \Pi_1 \\ \text{Eğer } f_{\underline{x}_1}(\underline{x})/f_{\underline{x}_2}(\underline{x}) < 1 \text{ ise, } \underline{x} \in \Pi_2 \end{cases}$$

biçiminde yazılır. $f_{\underline{x}_1}(\underline{x})/f_{\underline{x}_2}(\underline{x})$ oranı

$$\frac{f_{\underline{x}_1}(\underline{x})}{f_{\underline{x}_2}(\underline{x})} = \frac{\alpha_2 |\underline{\Sigma}_1|^{\frac{1}{2}}}{\alpha_1 |\underline{\Sigma}_2|^{\frac{1}{2}}} \exp \left\{ \alpha_2 \sqrt{D_{\underline{x},2}} - \alpha_1 \sqrt{D_{\underline{x},1}} + (\underline{x} - \underline{\mu}_1)' \underline{\Sigma}_1^{-1} \underline{\gamma}_1 - (\underline{x} - \underline{\mu}_2)' \underline{\Sigma}_2^{-1} \underline{\gamma}_2 \right\}$$

olup diskriminant fonksiyonu

$$d_d^{MSL}(\underline{x}) = \alpha_2 \sqrt{D_{\underline{x},2}} - \alpha_1 \sqrt{D_{\underline{x},1}} + (\underline{x} - \underline{\mu}_2)' \underline{\Sigma}_2^{-1} \underline{\gamma}_2 - (\underline{x} - \underline{\mu}_1)' \underline{\Sigma}_1^{-1} \underline{\gamma}_1 \quad (5.4)$$

biçiminde, atama kuralı ise

$$\begin{cases} \text{Eğer } d_d^{MSL}(\underline{x}) \geq \log \left(\alpha_1 |\underline{\Sigma}_2|^{\frac{1}{2}} \right) - \log \left(\alpha_2 |\underline{\Sigma}_1|^{\frac{1}{2}} \right) , \underline{x} \in \Pi_1 \\ \text{Eğer } d_d^{MSL}(\underline{x}) < \log \left(\alpha_1 |\underline{\Sigma}_2|^{\frac{1}{2}} \right) - \log \left(\alpha_2 |\underline{\Sigma}_1|^{\frac{1}{2}} \right) , \underline{x} \in \Pi_2 \end{cases} \quad (5.5)$$

olarak bulunur. $\Sigma_1 = \Sigma_2 = \Sigma$ ve $\underline{\gamma}_1 = \underline{\gamma}_2 = \underline{\gamma}$ olması durumunda ise $f_{\underline{x}_1}(\underline{x})/f_{\underline{x}_2}(\underline{x})$ oranı

$$\frac{f_{\underline{x}_1}(\underline{x})}{f_{\underline{x}_2}(\underline{x})} = \exp \left\{ \alpha \left(\sqrt{D_{\underline{x},2}} - \sqrt{D_{\underline{x},1}} \right) + (\underline{\mu}_2 - \underline{\mu}_1)' \Sigma^{-1} \underline{\gamma} \right\}$$

olup, diskriminant fonksiyonu

$$d_s^{MSL}(\underline{x}) = \sqrt{D_{\underline{x},1}} - \sqrt{D_{\underline{x},2}} \quad (5.6)$$

biçiminde, atama kuralı ise

$$\begin{cases} \text{Eğer } d_s^{MSL}(\underline{x}) \geq \frac{1}{\alpha} (\underline{\mu}_2 - \underline{\mu}_1)' \Sigma^{-1} \underline{\gamma} & , \underline{x} \in \Pi_1 \\ \text{Eğer } d_s^{MSL}(\underline{x}) < \frac{1}{\alpha} (\underline{\mu}_2 - \underline{\mu}_1)' \Sigma^{-1} \underline{\gamma} & , \underline{x} \in \Pi_2 \end{cases} \quad (5.7)$$

olarak bulunur.

5.4 Çarpık Laplace Dağılımı İçin Simülasyon Sonuçları

Simülasyon çalışmasında farklı örneklem hacimlerinde $MSL_2(\underline{\mu}, \Sigma, \underline{\gamma})$ dağılımından sayı üretip, bu rasgele örnekler ile bir önceki bölümde ele alınan atama kuralları ile sınıflandırma yapılmıştır. Bu işlem her bir örnek hacminde 100 kez tekrarlanıp, gözlenen hatalı sınıflandırma oranlarının ortalaması alınmıştır.

Çizelge 5.1’de birinci kitlenin dağılımı $MSL_2(\underline{\mu}_1 = [0;0], \Sigma_1 = [2, 0.1; 0.1, 2], \underline{\gamma}_1 = [1, -1])$, ikinci kitlenin dağılımı ise $MSL_2(\underline{\mu}_2 = [3; -3], \Sigma_2 = [2, -1; -1, 2], \underline{\gamma}_2 = [-0.1, 1])$ alınmıştır.

Tüm örneklem hacimlerinde parametrelerin tahmini değerlerini kullanılarak atama kuralı (5.5) ile yapılan sınıflandırma ile gözlenen hatalı sınıflandırma oranlarının, lineer ve karesel diskriminant fonksiyonlarından daha az olduğu görülmektedir.

Çizelge 5.1 Çarpık Laplace dağılımı için hatalı sınıflandırma oranları

n	Lineer	Karesel	MSL (Gerçek par. ile)	MSL (Tahmini par. ile)
10	0.32	0.29	0.26	0.20
20	0.34	0.32	0.27	0.25
30	0.35	0.32	0.26	0.24
50	0.35	0.33	0.26	0.25
100	0.36	0.33	0.25	0.25

Çizelge 5.2’de $\underline{\mu}_1$ ve $\underline{\mu}_2$ hariç diğer parametrelerin aynı olduğu durum için yapılan simülasyon çalışmasından elde edilen ortalama hatalı sınıflandırma oranları görülmektedir. Burada $\Sigma = [5, 1; 1, 5]$ ve $\underline{\gamma} = [1, -1]$ olmak üzere birinci kitlenin dağılımı $MSL_2(\underline{\mu}_1 = [0; 0], \Sigma, \underline{\gamma})$ ikinci kitlenin dağılımı ise $MSL_2(\underline{\mu}_2 = [3; -3], \Sigma, \underline{\gamma})$ olarak alınmıştır.

Çizelge 5.2 Çarpık Laplace dağılımı için $\underline{\mu}_1$ ve $\underline{\mu}_2$ hariç aynı parametreler ile hatalı sınıflandırma oranları

n	Lineer	Karesel	MSL (Gerçek par. ile)	MSL (Tahmini par. ile)
10	0.31	0.30	0.28	0.22
20	0.31	0.29	0.26	0.25
30	0.32	0.31	0.27	0.25
50	0.32	0.31	0.26	0.25
100	0.32	0.32	0.27	0.26

Çizelge 5.2’den görülebileceği üzere tüm örnek hacimlerinde, parametrelerin tahmini değerlerini kullanılarak atama kuralı (5.7) ile yapılan sınıflandırma ile gözlenen hatalı sınıflandırma oranlarının, lineer ve karesel diskriminant fonksiyonlarından daha az olduğu gözlenmiştir.

6. ÇOK DEĞİŞKENLİ NORMAL TERS GAUSS (MNIG) DAĞILIMI

Normal ortalama-varyans karma dağılımında karıştırıcı değişken olan W 'nin dağılımının ters Gauss (IG) alınması durumunda normal ters Gauss (NIG) dağılımı elde edilir. İlk olarak Barndorff-Nielsen (1997) tarafından finansal modellemede kullanılan dağılım aynı zamanda sinyal işleme uygulamalarında da sıklıkla kullanılmaktadır. Genelleştirilmiş hiperbolik dağılım ailesinin bir üyesi olmasından dolayı bu ailenin ağır kuyruklu ve çarpık olma özelliklerini taşır.

$\underline{X} = \underline{\mu} + W\underline{\gamma} + \sqrt{W}\mathbf{A}\underline{Z}$ dönüşümü ile verilmiş normal-ortalama varyans karma dağılımında karıştırıcı değişken $W \sim N^-(\lambda = -1/2, \chi, \psi)$ dağılımlı olarak alındığında $\underline{X}$ rasgele vektörünün dağılımı '*çok değişkenli normal ters gauss* (multivariate normal inverse Gaussian, MNIG)' olur ve

$$f_{\underline{X}}(\underline{x}) = \left[\frac{\chi}{2^{d-1}|\mathbf{\Sigma}|} \right]^{\frac{1}{2}} \left[\frac{\underline{\gamma}'\mathbf{\Sigma}^{-1}\underline{\gamma} + \psi}{\pi^2(D_{\underline{x}} + \chi)} \right]^{\frac{d+1}{4}} K_{\frac{d+1}{2}} \left(\sqrt{(D_{\underline{x}} + \chi)(\underline{\gamma}'\mathbf{\Sigma}^{-1}\underline{\gamma} + \psi)} \right) \times \exp \left\{ (\underline{x} - \underline{\mu})'\mathbf{\Sigma}^{-1}\underline{\gamma} + \sqrt{\chi\psi} \right\} \quad (6.1)$$

olasılık yoğunluk fonksiyonuna sahiptir. Buna göre $\underline{X}$ rasgele vektörü (6.1) ile verilen olasılık yoğunluk fonksiyonuna sahip olduğunda $\underline{X} \sim MNIG(\underline{\mu}, \mathbf{\Sigma}, \underline{\gamma}, \chi, \psi)$ ile gösterilecektir.

$W \sim N^-(\lambda = -1/2, \chi, \psi)$ olmak üzere karıştırıcı değişken W 'nin olasılık yoğunluk fonksiyonu

$$\begin{aligned}
f_W(w; \lambda = -1/2, \chi, \psi) &= \frac{\chi^{\frac{1}{2}} (\sqrt{\chi\psi})^{-\frac{1}{2}}}{2K_{\frac{1}{2}}(\sqrt{\chi\psi})} w^{-\frac{1}{2}-1} \exp\left(-\frac{1}{2}\left(\chi \frac{1}{w} + \psi w\right)\right) \\
&= \frac{\chi^{\frac{1}{4}} \psi^{-\frac{1}{4}}}{2\chi^{\frac{1}{4}} \psi^{-\frac{1}{4}} \sqrt{\pi/2}} w^{-\frac{3}{2}} \exp\left(-\frac{1}{2}\left(\chi \frac{1}{w} + \psi w\right) + \sqrt{\chi\psi}\right) \\
&= \left(\frac{\chi}{2\pi w^3}\right)^{\frac{1}{2}} \exp\left(-\frac{1}{2}\left(\chi \frac{1}{w} + \psi w\right) + \sqrt{\chi\psi}\right), w > 0
\end{aligned}$$

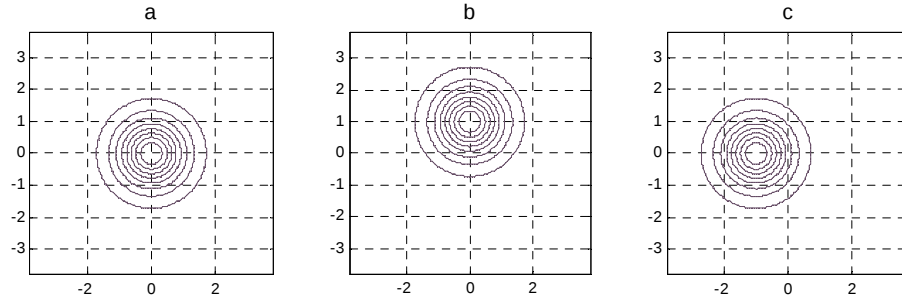
olarak elde edilir. Buradan görülebileceği üzere $W \sim IG(\sqrt{\chi/\psi}, \psi)$ dağılımlıdır. Bölüm 2.1.7’de verilen beklenen değer ve varyans ifadelerinden $E(W) = \sqrt{\chi/\psi}$ ve $Var(W) = \sqrt{\chi\psi}$ olduğu görülür. Böylece (2.25) ve (2.26) kullanılarak $\underline{X} \sim MNIG(\underline{\mu}, \underline{\Sigma}, \underline{\gamma}, \chi, \psi)$ rasgele vektörü için

$$\begin{aligned}
E(\underline{X}) &= \underline{\mu} + E(W)\underline{\gamma} \\
&= \underline{\mu} + \underline{\gamma} \sqrt{\chi/\psi} \quad ,
\end{aligned} \tag{6.2}$$

$$\begin{aligned}
Cov(\underline{X}) &= E(W)\underline{\Sigma} + Var(W)\underline{\gamma}\underline{\gamma}' \\
&= \sqrt{\chi/\psi}\underline{\Sigma} + \underline{\gamma}\underline{\gamma}'\sqrt{\chi\psi}
\end{aligned} \tag{6.3}$$

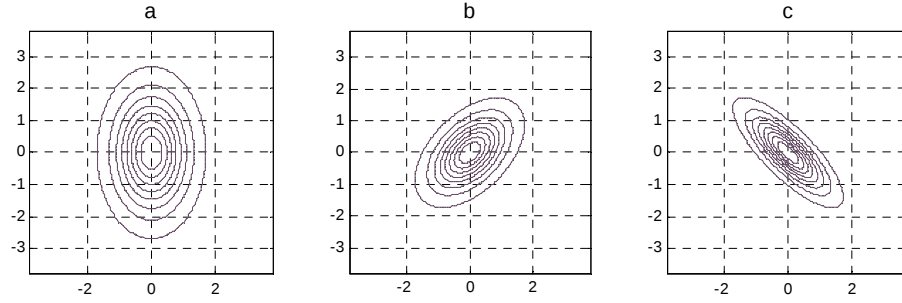
beklenen değeri ve varyans-kovaryans matrisi elde edilir.

Şekil 6.1-6.4’de $MNIG(\underline{\mu}, \underline{\Sigma}, \underline{\gamma}, \chi, \psi)$ dağılımı için parametrelerin olasılık yoğunluk fonksiyonunu nasıl değiştirdiği aşağıdaki grafiklerde görülmektedir.



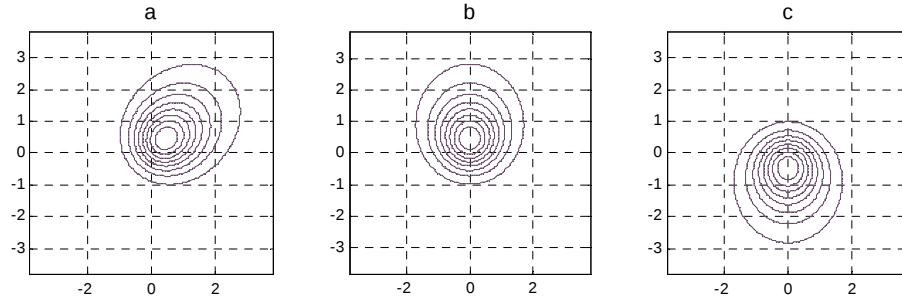
Şekil 6.1 MNIG dağılımının $\underline{\mu}$ parametersine göre değişimi

a. $\underline{\mu} = [0 \ 0]'$, b. $\underline{\mu} = [0 \ 1]'$, c. $\underline{\mu} = [-1 \ 0]'$



Şekil 6.2 MNIG dağılımının $\underline{\Sigma}$ parametersine göre değişimi

a. $\underline{\Sigma} = \begin{bmatrix} 1 & 0.01 \\ 0.01 & 4 \end{bmatrix}$ b. $\underline{\Sigma} = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}$ c. $\underline{\Sigma} = \begin{bmatrix} 1 & -0.9 \\ -0.9 & 1 \end{bmatrix}$

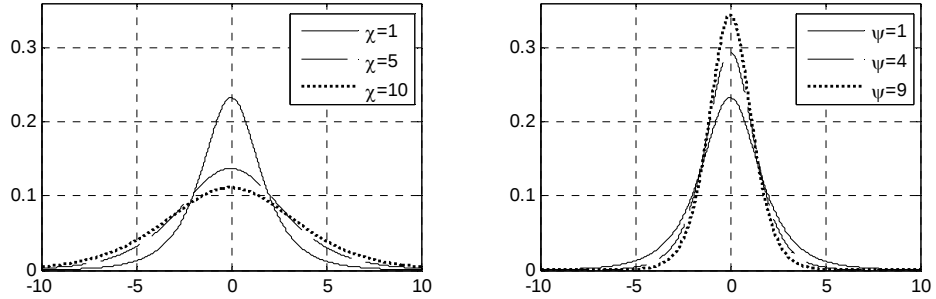


Şekil 6.3 MNIG dağılımının $\underline{\gamma}$ parametresine göre değişimi

a. $\underline{\gamma} = [1 \ 1]'$, b. $\underline{\gamma} = [0 \ 1]'$, c. $\underline{\gamma} = [0 \ -1]'$

Genelleştirilmiş hiperbolik dağılım ailesinin diğer üyelerinde olduğu gibi MNIG dağılımında da $\underline{\mu}$ parametresi dağılımın konumunu, $\underline{\Sigma}$ parametresi saçılımını, $\underline{\gamma}$

parametresi dağılımın çarpıklığını değiştirmektedir. Aşağıdaki grafiklerde ise χ ve ψ parametrelerinin dağılımı nasıl değiştirdiği görülmektedir.



Şekil 6.4 MNIG dağılımının χ ve ψ parametresine göre değişimi

Grafiklerden görülebileceği üzere χ ve ψ parametreleri dağılımın basıklığını değiştirmektedir. Basıklık χ artarken artmakta, ψ artarken azalmaktadır. Bu iki parametrenin farklı seçimleriyle kuyruk kalınlığı farklı dağılımlar elde etmek mümkün olmaktadır.

6.1 MNIG Dağılımından Rasgele Sayıların Üretilmesi

Çok değişkenli normal ters Gauss dağılımından sayı üretilmesi için aşağıdaki adımlar izlenir.

Girdiler: $\underline{\mu}_{d \times 1}, \underline{\Sigma}_{d \times d}, \underline{\gamma}_{d \times 1}, \chi, \psi$

1. $\underline{\Sigma}$ matrisinin Cholesky ayrışımından $\underline{\Sigma} = \underline{\mathbf{A}}\underline{\mathbf{A}}^t$ olacak şekilde $\underline{\mathbf{A}}_{d \times d}$ matrisini bul.
2. $\underline{Z} \sim N_k(\underline{0}, \underline{\mathbf{I}}_d)$ dağılımından sayı üret.
3. $W \sim IG(\sqrt{\chi/\psi}, \psi)$ dağılımından sayı üret.
4. $\underline{X} = \underline{\mu} + W\underline{\gamma} + \sqrt{W}\underline{\mathbf{A}}\underline{Z}$ değerini hesapla.

Çıktı: $\underline{X}$

6.2 MNIG Dağılımının Parametrelerinin Tahmini İçin EM-Algoritması

Çok değişkenli genelleştirilmiş hiperbolik dağılım için EM-Algoritması Bölüm 2.3’de verilmişti. $\lambda = -1/2$ parametre değeri algoritmada yerlerine yazılarak MNIG Dağılımı için EM-Algoritması oluşturulur. Bu noktada Bessel fonksiyonunun Bölüm 2.1.2’de verilmiş olan özelliklerinden yararlanılır.

χ ve ψ parametrelerinin tahmini için (2.33)’ da verilen log-olabilirlik fonksiyonunda $\lambda = -1/2$ alınır. Buradan χ parametresinin tahmini için

$$\frac{\partial}{\partial \chi} L_2 = \frac{n - \chi \sum_{i=1}^n w_i^{-1} + n\sqrt{\chi\psi}}{2\chi}$$

değeri sıfıra eşitlenirse

$$1 - \frac{\chi}{n} \sum_{i=1}^n w_i^{-1} + \sqrt{\chi\psi} = 0 \quad (6.4)$$

denklemini elde edilir. Buradan eşitliği sağlayan χ değeri χ ’nin yeni değeri olarak alınır. Benzer şekilde ψ parametresinin tahmini için

$$\frac{\partial}{\partial \psi} L_2 = \frac{n\chi}{2\sqrt{\chi\psi}} - \frac{1}{2} \sum_{i=1}^n w_i$$

değeri sıfıra eşitlenirse

$$\hat{\psi} = \frac{\left(\frac{1}{n} \sum_{i=1}^n w_i \right)^2}{\chi} \quad (6.5)$$

eşitliğinden ψ 'nin yeni değeri elde edilir. MNIG Dağılımı için EM-Algoritması aşağıdaki gibi oluşturulur.

- E-Adımı

- o k.iterasyon'da tüm gözlemler için aşağıdakileri hesapla

$$\delta_i^{[k]} = \left(\frac{D_{x_i}^{[k]} + \chi^{[k]}}{\psi^{[k]} + \underline{\gamma}^{[k]}, \Sigma^{[k]-1} \underline{\gamma}^{[k]}} \right)^{\frac{1}{2}} \frac{K_{\frac{d+3}{2}} \left(\sqrt{(D_{x_i}^{[k]} + \chi^{[k]}) (\psi^{[k]} + \underline{\gamma}^{[k]}, \Sigma^{[k]-1} \underline{\gamma}^{[k]})} \right)}{K_{\frac{d+1}{2}} \left(\sqrt{(D_{x_i}^{[k]} + \chi^{[k]}) (\psi^{[k]} + \underline{\gamma}^{[k]}, \Sigma^{[k]-1} \underline{\gamma}^{[k]})} \right)}$$

$$\eta_i^{[k]} = \left(\frac{D_{x_i}^{[k]} + \chi^{[k]}}{\psi^{[k]} + \underline{\gamma}^{[k]t} \Sigma^{[k]-1} \underline{\gamma}^{[k]}} \right)^{\frac{1}{2}} \frac{K_{\frac{d-1}{2}} \left(\sqrt{(D_{x_i}^{[k]} + \chi^{[k]}) (\psi^{[k]} + \underline{\gamma}^{[k]}, \Sigma^{[k]-1} \underline{\gamma}^{[k]})} \right)}{K_{\frac{d+1}{2}} \left(\sqrt{(D_{x_i}^{[k]} + \chi^{[k]}) (\psi^{[k]} + \underline{\gamma}^{[k]}, \Sigma^{[k]-1} \underline{\gamma}^{[k]})} \right)}$$

- o k.iterasyon'da aşağıdaki değerleri bul

$$\bar{\delta}^{[k]} = \frac{1}{n} \sum_{i=1}^n \delta_i^{[k]} \quad , \quad \bar{\eta}^{[k]} = \frac{1}{n} \sum_{i=1}^n \eta_i^{[k]}$$

- M-Adımı

- o Parametrelerin yeni değerlerini bul

$$\underline{\gamma}^{[k+1]} = \frac{n^{-1} \sum_{i=1}^n \delta_i^{[k]} (\bar{x} - \underline{x}_i)}{\bar{\delta}^{[k]} \bar{\eta}^{[k]} - 1} \quad , \quad \underline{\mu}^{[k+1]} = \frac{n^{-1} \sum_{i=1}^n \delta_i^{[k]} \underline{x}_i - \underline{\gamma}^{[k+1]}}{\bar{\delta}^{[k]}}$$

$$\Sigma^{[k+1]} = \frac{1}{n} \sum_{i=1}^n \delta_i^{[k]} (\underline{x}_i - \underline{\mu}^{[k+1]}) (\underline{x}_i - \underline{\mu}^{[k+1]})' - \bar{\eta}^{[k]} \underline{\gamma}^{[k+1]} \underline{\gamma}^{[k+1]}$$

- o χ 'nin yeni değerini aşağıdaki denklemi çözerek bul

$$1 - \chi \bar{\delta}^{[k]} + \sqrt{\chi \bar{\psi}} = 0$$

- o ψ 'nin yeni değerini aşağıdaki denklemden bul

$$\hat{\bar{\psi}} = \bar{\eta}^{[k]2} / \chi$$

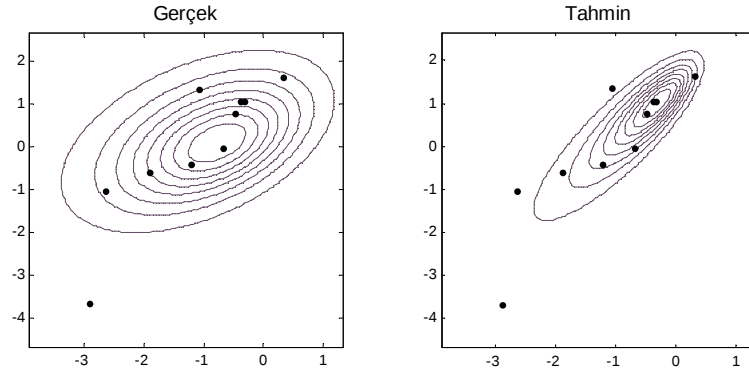
- o Yakınsama sağlanmadıysa $k=k+1$ olarak E-Adımına dön.

İterasyonun başlangıç değerleri olarak $\underline{\mu}$ için örneklem ortalaması, $\underline{\Sigma}$ için örneklem Varyans-kovaryans matrisi ve $\underline{\gamma}$ için ise sıfıra çok yakın elemanları olan bir vektör seçilir. χ ve ψ 'nin başlangıç değerleri ise 10 olarak seçilir.

Şekil 6.5-6.9'da $\underline{X} \sim MNIG(\underline{\mu}, \underline{\Sigma}, \underline{\gamma}, \chi, \psi)$ dağılımından üretilmiş değişik hacimlerdeki örneklerden elde edilmiş $\underline{\mu}, \underline{\Sigma}, \underline{\gamma}, \chi, \psi$ parametrelerinin tahminleri, gerçek parametreler ve tahminlerine göre çizilmiş olasılık yoğunluk fonksiyonlarının izdüşüm grafikleri görülmektedir. Parametrelerin değerleri

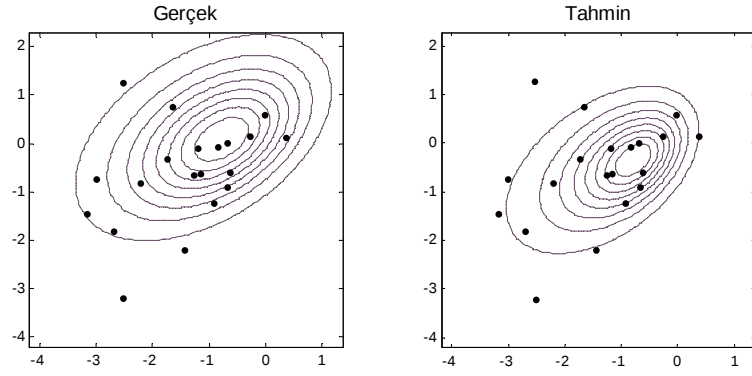
$$\underline{\mu} = [0; 0], \underline{\Sigma} = [1, 0.5; 0.5, 1], \underline{\gamma} = [-0.9; 1], \chi = 6, \psi = 4$$

olarak seçilmiştir.

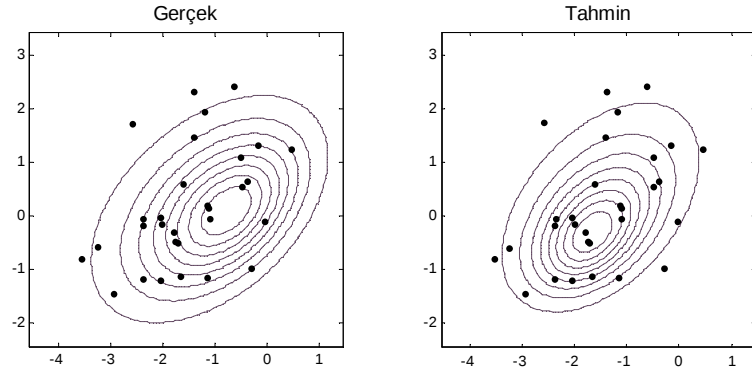


Şekil 6.5 n=10 örnek ile MNIG dağılımının parametrelerinin EM tahmini

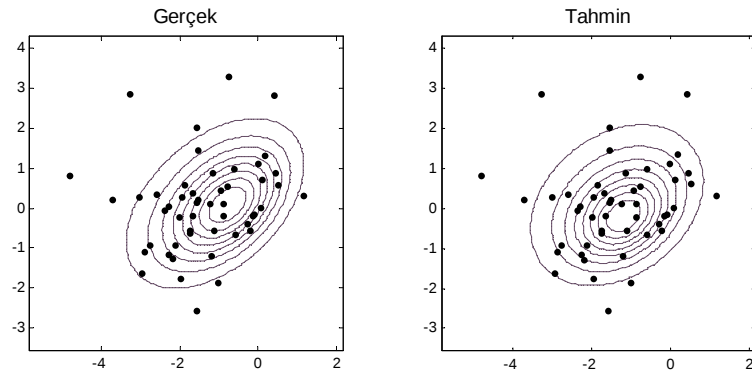
$$\hat{\underline{\mu}} = [0.3; 1.7], \hat{\underline{\Sigma}} = [0.3, 0.2; 0.2, 0.5], \hat{\underline{\gamma}} = [-0.8; -1.0], \hat{\chi} = 3.0, \hat{\psi} = 1.0$$



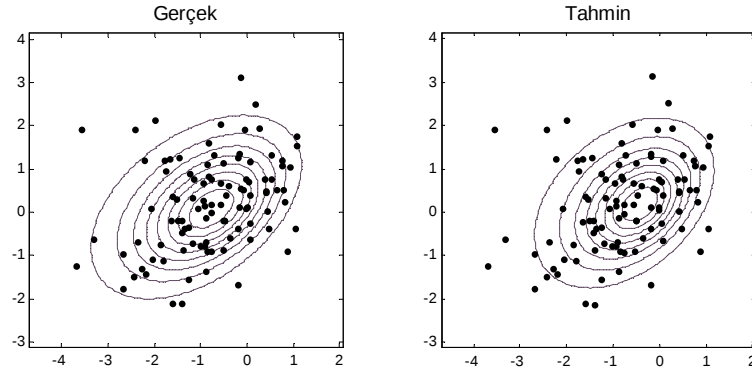
Şekil 6.6 n=20 örnek ile MNIG dağılımının parametrelerinin EM tahmini
 $\hat{\underline{\mu}} = [-0.2; -0.1]$, $\hat{\underline{\Sigma}} = [0.3, 0.2; 0.2, 0.5]$, $\hat{\underline{\gamma}} = [-0.6; -0.2]$, $\hat{\chi} = 4.4$, $\hat{\psi} = 1.0$



Şekil 6.7 n=30 örnek ile MNIG dağılımının parametrelerinin EM tahmini
 $\hat{\underline{\mu}} = [-1.9; -1.0]$, $\hat{\underline{\Sigma}} = [0.4, 0.1; 0.1, 0.3]$, $\hat{\underline{\gamma}} = [0.1; 0.4]$, $\hat{\chi} = 8.0$, $\hat{\psi} = 1.0$



Şekil 6.8 n=50 örnek ile MNIG dağılımının parametrelerinin EM tahmini
 $\hat{\underline{\mu}} = [-1.0; -0.5]$, $\hat{\underline{\Sigma}} = [0.7, 0.2; 0.2, 0.6]$, $\hat{\underline{\gamma}} = [-0.2; 0.3]$, $\hat{\chi} = 5.0$, $\hat{\psi} = 1.0$



Şekil 6.9 n=100 örnek ile MNIG dağılımının parametrelerinin EM tahmini
 $\hat{\underline{\mu}} = [0.0; 0.2]$, $\hat{\underline{\Sigma}} = [0.4, 0.2; 0.2, 0.5]$, $\hat{\underline{\gamma}} = [-0.3; 0.0]$, $\hat{\chi} = 6.0$, $\hat{\psi} = 1.0$

Grafiklerden görülebileceği üzere parametrelerin tahmini değerleri ile çizilen konturların özellikle küçük örneklem hacimlerinde gerçek parametrelerden daha iyi sonuç verdiği anlaşılmaktadır. Aynı durum parametre tahminleriyle ve gerçek parametrelerle hesaplanan $E(\underline{X})$ ve $Cov(\underline{X})$ değerlerinde de etkisini göstermektedir. Örneğin 20 birimlik örneklem ile örneklem ortalaması $[-1.4; -0.6]$, örneklem varyans-kovaryans matrisi $[1.0, 0.4; 0.4, 1.1]$ olarak gözlemlendiğinde, parametrelerin tahminleri ile beklenen değer $[-1.4; -0.6]$, varyans-kovaryans matrisi $[1.4, 0.6; 0.6, 1.1]$ hesaplanırken, gerçek parametre değerleri ile beklenen değer $[-1.1; 0.1]$, varyans-kovaryans matrisi $[5.2, 0.2; 0.2, 1.3]$ olarak hesaplanmıştır.

Değerlerden parametre tahmini ile elde edilen değerlerin örneklemde elde edilen değerlere daha yakın olduğu anlaşılmaktadır. Bu durum EM tahminlerinin gerçek parametrelere uzak olmasına rağmen neden örneklere daha iyi uyduğunu açıklamaktadır.

6.3 MNIG Dağılımına Sahip Kitlelerde Diskriminant Fonksiyonu

Π_1 ve Π_2 sırasıyla $MNIG(\underline{\mu}_1, \underline{\Sigma}_1, \underline{\gamma}_1, \chi_1, \psi_1)$ ve $MNIG(\underline{\mu}_2, \underline{\Sigma}_2, \underline{\gamma}_2, \chi_2, \psi_2)$ dağılımına sahip kitleler olsun. Bu iki kitlenin birisinden geldiği bilinen bir $\underline{x} \in \mathbb{R}^d$ gözlemi için atama kuralı

$$\begin{cases} \text{Eğer } f_{\underline{x}_1}(\underline{x})/f_{\underline{x}_2}(\underline{x}) \geq 1 \text{ ise, } \underline{x} \in \Pi_1 \\ \text{Eğer } f_{\underline{x}_1}(\underline{x})/f_{\underline{x}_2}(\underline{x}) < 1 \text{ ise, } \underline{x} \in \Pi_2 \end{cases}$$

biçiminde yazılır. $f_{\underline{x}_1}(\underline{x})/f_{\underline{x}_2}(\underline{x})$ oranı

$$\begin{aligned} \frac{f_{\underline{x}_1}(\underline{x})}{f_{\underline{x}_2}(\underline{x})} &= \left[\frac{\chi_1}{\chi_2} \frac{|\underline{\Sigma}_2|}{|\underline{\Sigma}_1|} \right]^{\frac{1}{2}} \left[\frac{\underline{\gamma}_1' \underline{\Sigma}_1^{-1} \underline{\gamma}_1 + \psi_1}{\underline{\gamma}_2' \underline{\Sigma}_2^{-1} \underline{\gamma}_2 + \psi_2} \right]^{\frac{d+1}{4}} \left[\frac{D_{\underline{x},2} + \chi_2}{D_{\underline{x},1} + \chi_1} \right]^{\frac{d+1}{4}} \frac{K_{\frac{d+1}{2}} \left(\sqrt{(D_{\underline{x},1} + \chi_1)(\underline{\gamma}_1' \underline{\Sigma}_1^{-1} \underline{\gamma}_1 + \psi_1)} \right)}{K_{\frac{d+1}{2}} \left(\sqrt{(D_{\underline{x},2} + \chi_2)(\underline{\gamma}_2' \underline{\Sigma}_2^{-1} \underline{\gamma}_2 + \psi_2)} \right)} \\ &\quad \times \exp \left\{ (\underline{x} - \underline{\mu}_1)' \underline{\Sigma}_1^{-1} \underline{\gamma}_1 - (\underline{x} - \underline{\mu}_2)' \underline{\Sigma}_2^{-1} \underline{\gamma}_2 \right\} \exp \left\{ \sqrt{\chi_1 \psi_1} - \sqrt{\chi_2 \psi_2} \right\} \end{aligned}$$

biçiminde yazılır. Buradan

$$\begin{aligned} d_d^{MNIG}(\underline{x}) &= \left[\frac{D_{\underline{x},2} + \chi_2}{D_{\underline{x},1} + \chi_1} \right]^{\frac{d+1}{4}} \frac{K_{\frac{d+1}{2}} \left(\sqrt{(D_{\underline{x},1} + \chi_1)(\underline{\gamma}_1' \underline{\Sigma}_1^{-1} \underline{\gamma}_1 + \psi_1)} \right)}{K_{\frac{d+1}{2}} \left(\sqrt{(D_{\underline{x},2} + \chi_2)(\underline{\gamma}_2' \underline{\Sigma}_2^{-1} \underline{\gamma}_2 + \psi_2)} \right)} \\ &\quad \times \exp \left\{ (\underline{x} - \underline{\mu}_1)' \underline{\Sigma}_1^{-1} \underline{\gamma}_1 - (\underline{x} - \underline{\mu}_2)' \underline{\Sigma}_2^{-1} \underline{\gamma}_2 \right\} \end{aligned} \quad (6.6)$$

diskriminant fonksiyonu oluşturulur. Atama kuralı ise

$$\left\{ \begin{array}{l} \text{Eğer } d_d^{MNIG}(\underline{x}) \geq \left[\frac{\chi_2 |\underline{\Sigma}_1|}{\chi_1 |\underline{\Sigma}_2|} \right]^{\frac{1}{2}} \left[\frac{\underline{\gamma}_2' \underline{\Sigma}_2^{-1} \underline{\gamma}_2 + \psi_2}{\underline{\gamma}_1' \underline{\Sigma}_1^{-1} \underline{\gamma}_1 + \psi_1} \right]^{\frac{d+1}{4}} \exp\{\sqrt{\chi_2 \psi_1} - \sqrt{\chi_1 \psi_1}\} \quad , \underline{x} \in \Pi_1 \\ \text{Eğer } d_d^{MNIG}(\underline{x}) < \left[\frac{\chi_2 |\underline{\Sigma}_1|}{\chi_1 |\underline{\Sigma}_2|} \right]^{\frac{1}{2}} \left[\frac{\underline{\gamma}_2' \underline{\Sigma}_2^{-1} \underline{\gamma}_2 + \psi_2}{\underline{\gamma}_1' \underline{\Sigma}_1^{-1} \underline{\gamma}_1 + \psi_1} \right]^{\frac{d+1}{4}} \exp\{\sqrt{\chi_2 \psi_1} - \sqrt{\chi_1 \psi_1}\} \quad , \underline{x} \in \Pi_2 \end{array} \right. \quad (6.7)$$

biçiminde elde edilir. $\underline{\mu}_1$ ve $\underline{\mu}_2$ hariç diğer parametrelerin eşit olması durumunda ise diskriminant fonksiyonu

$$\begin{aligned} d_s^{MNIG}(\underline{x}) = & \frac{d+1}{4} \left[\log(D_{\underline{x},2} + \chi_2) - \log(D_{\underline{x},1} + \chi_1) \right] \\ & + \log K_{\frac{d+1}{2}} \left(\sqrt{(D_{\underline{x},1} + \chi_1)(\underline{\gamma}_1' \underline{\Sigma}_1^{-1} \underline{\gamma}_1 + \psi_1)} \right) \\ & - \log K_{\frac{d+1}{2}} \left(\sqrt{(D_{\underline{x},2} + \chi_2)(\underline{\gamma}_2' \underline{\Sigma}_2^{-1} \underline{\gamma}_2 + \psi_2)} \right) \end{aligned} \quad (6.8)$$

ve atama kuralı

$$\left\{ \begin{array}{l} \text{Eğer } d_s^{MNIG}(\underline{x}) \geq (\underline{\mu}_1 - \underline{\mu}_2)' \underline{\Sigma}^{-1} \underline{\gamma} \quad , \underline{x} \in \Pi_1 \\ \text{Eğer } d_s^{MNIG}(\underline{x}) < (\underline{\mu}_1 - \underline{\mu}_2)' \underline{\Sigma}^{-1} \underline{\gamma} \quad , \underline{x} \in \Pi_2 \end{array} \right. \quad (6.9)$$

biçiminde elde edilir.

6.4 MNIG Dağılımı İçin Simülasyon Sonuçları

Simülasyon çalışmasında farklı örneklem hacimlerinde $MNIG(\underline{\mu}, \underline{\Sigma}, \underline{\gamma}, \chi, \psi)$ dağılımından sayı üretip, bu örnekler ile bir önceki bölümde ele alınan atama kuralları ile sınıflandırma yapılmıştır. Bu işlem her bir örnek hacminde 100 kez tekrarlanıp, gözlenen hatalı sınıflandırma oranlarının ortalaması alınmıştır.

Çizelge 6.1’de iki boyutlu rasgele vektörün dağılımı birinci kitle için $MNIG(\underline{\mu}_1 = [0;0], \underline{\Sigma}_1 = [1,0.5;0.5,1], \underline{\gamma}_1 = [-0.9;0.1], \chi_1 = 6, \psi_1 = 4)$, ikinci kitle için $MNIG(\underline{\mu}_2 = [-2;1], \underline{\Sigma}_2 = [1,-0.1;-0.1,1], \underline{\gamma}_2 = [0.9;-0.9], \chi_2 = 3, \psi_2 = 3)$ alınmıştır.

Çizelgeden tahmini parametreler kullanarak (6.7) ile verilen atama kuralı kullanıldığında yapılan sınıflandırmada gözlenen hatalı sınıflandırma oranlarının, tüm örneklem hacimlerinde lineer ve karesel diskriminant fonksiyonlarından daha az olduğu görülmektedir. Özellikle düşük örneklem hacimlerinde hatalı sınıflandırma oranı lineer ve karesel diskriminant fonksiyonuna göre oldukça düşüktür.

Çizelge 6.1 MNIG dağılımı için diskriminant analizi hatalı sınıflandırma oranları

n	Lineer	Karesel	MNIG (Gerçek par. ile)	MNIG (Tahmini par. ile)
10	0.42	0.32	0.35	0.25
20	0.42	0.34	0.35	0.29
30	0.44	0.35	0.35	0.31
50	0.46	0.37	0.35	0.32
100	0.47	0.38	0.35	0.34

$\underline{\mu}_1$ ve $\underline{\mu}_2$ hariç diğer parametrelerin aynı olduğu durum için yapılan simülasyon çalışmasının sonuçları çizelge 6.2’de görülmektedir. Bu simülasyon çalışmasında birinci ve ikinci kitlenin için sırasıyla $MNIG(\underline{\mu}_1 = [0;0], \underline{\Sigma}, \underline{\gamma}, \chi, \psi)$ ve $MNIG(\underline{\mu}_2 = [-2;1], \underline{\Sigma}, \underline{\gamma}, \chi, \psi)$ alınmıştır. Ortak parametrelerin değerleri ise $\underline{\Sigma} = [1,0.5;0.5,1], \underline{\gamma} = [-0.9;0.1], \chi = 6, \psi = 4$ dir. Tüm örneklem hacimlerinde elde edilen hatalı sınıflandırma oranları benzer olmasına karşın en az hatalı sınıflandırma (6.6) ile verilen atama kuralı, tahmini parametre değerleri ile kullanıldığında elde edilmiştir.

Çizelge 6.2 MNIG dağılımı için $\underline{\mu}_1$ ve $\underline{\mu}_2$ hariç aynı parametreler ile diskriminant analizi hatalı sınıflandırma oranları

n	Lineer	Karesel	MNIG (Gerçek par. ile)	MNIG (Tahmini par. ile)
10	0.09	0.08	0.10	0.07
20	0.10	0.10	0.11	0.09
30	0.10	0.10	0.10	0.09
50	0.10	0.09	0.09	0.09
100	0.10	0.10	0.10	0.09

Önceki bölümlerde yapılan simülasyon çalışmasındaki durumlara benzer olarak (konum parametresi hariç) parametrelerin eşit alındığı durumda d_s^{MNIG} 'nin lineer ve karesel diskriminant fonksiyonlarına karşı başarısının, tüm parametrelerin farklı olduğu durumda d_d^{MNIG} 'ye göre daha düşük olduğu görülmektedir. Her iki örnekte de parametre değerleri, gözlemleri olabildiğince ayrımı zor hale getirecek şekilde seçilmiştir. Ayrımı kolay olacak şekilde seçilen parametre değerleri ile yapılacak sınıflandırmada ele alınan tüm diskriminant fonksiyonlarının hemen hemen aynı hatalı sınıflandırma oranlarına sahip oldukları gözlenmiştir. Bu durum için çalışmanın son kısmında verilmiş örneklere bakılabilir.

7. UYGULAMA

Bu bölümde çalışma kapsamında ele alınan diskriminant fonksiyonlarının hatalı sınıflandırma oranları, değişik senaryolar altında karşılaştırılacaktır.

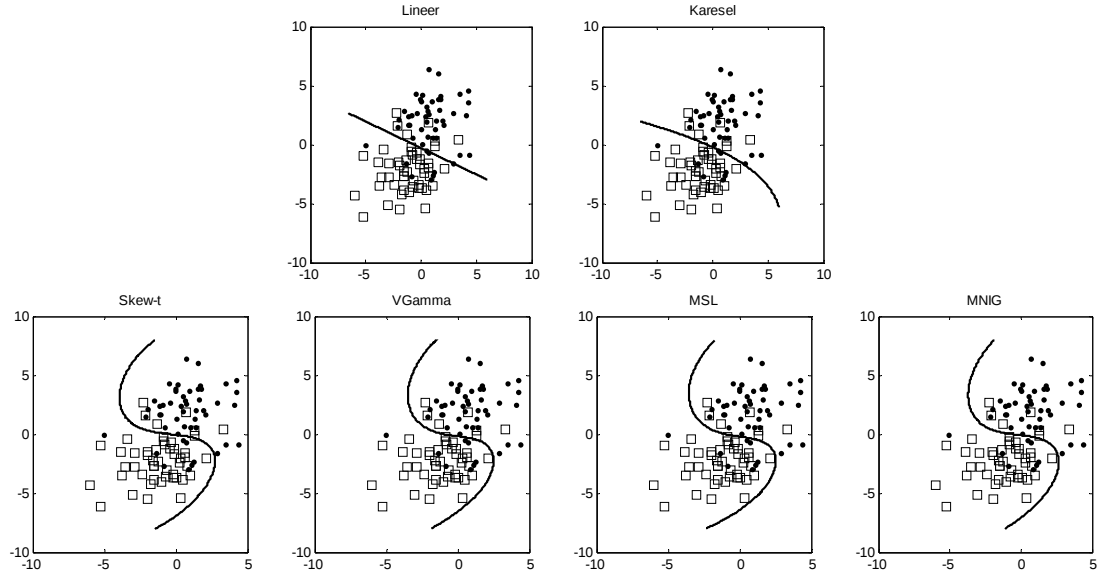
7.1 Simülasyon Yoluyla Üretilen Verilerin Sınıflandırılması

Örnek 1 (Normal dağılım, eş varyans) : Bu simülasyon çalışmasında sırasıyla $\underline{\mu}_1 = [1; 2]$, $\underline{\mu}_2 = [-1; -2]$ ve $\Sigma_1 = \Sigma_2 = [4, 0.5; 0.5, 4]$ parametrelili normal dağılıma sahip iki kitleden 50'şer birimlik örneklem üretilip hatalı sınıflandırma oranları elde edilmiştir. Çizelge 7.1'de 10 deney sonucunda elde edilen ortalama hatalı sınıflandırma oranları görülmektedir.

Çizelge 7.1 Örnek 1'e ait hatalı sınıflandırma oranları

Lineer	Karesel	Skew-t	VGamma	MSL	MNIG
0.12	0.11	0.12	0.11	0.12	0.12

Şekil 7.1'de ele alınan diskriminant fonksiyonlarının belirledikleri bölgeler görülmektedir.



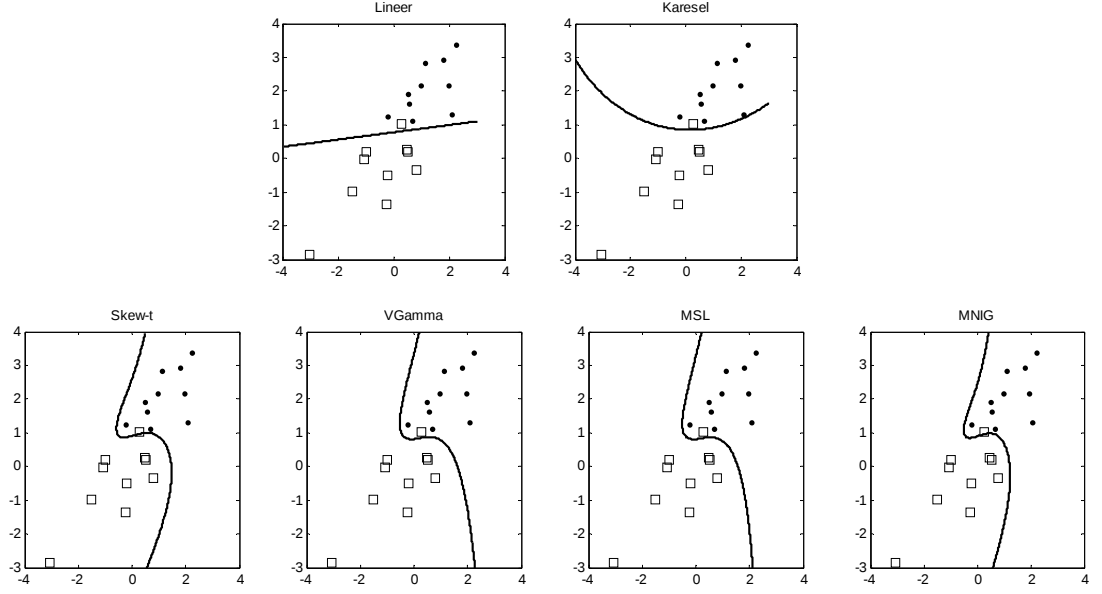
Şekil 7.1 Örnek 1 için diskriminant fonksiyonları

Örnek 2 (Normal dağılım, eş varyans, düşük örneklem hacmi) : Bu simülasyon çalışmasında sırasıyla $\underline{\mu}_1 = [1; 2]$, $\underline{\mu}_2 = [0; 0]$ ve $\Sigma_1 = \Sigma_2 = [1, 0.5; 0.5, 1]$ parametrelili normal dağılıma sahip iki kitleden 10'ar birimlik örneklem üretilip hatalı sınıflandırma oranları elde edilmiştir. Çizelge 7.2'de 20 tekrar sonucunda elde edilen ortalama hatalı sınıflandırma oranları görülmektedir.

Çizelge 7.2 Örnek 2'ye ait hatalı sınıflandırma oranları

Lineer	Karesel	Skew-t	VGamma	MSL	MNIG
0.140	0.125	0.118	0.105	0.108	0.100

Şekil 7.2'de ele alınan diskriminant fonksiyonlarının belirledikleri bölgeler görülmektedir.



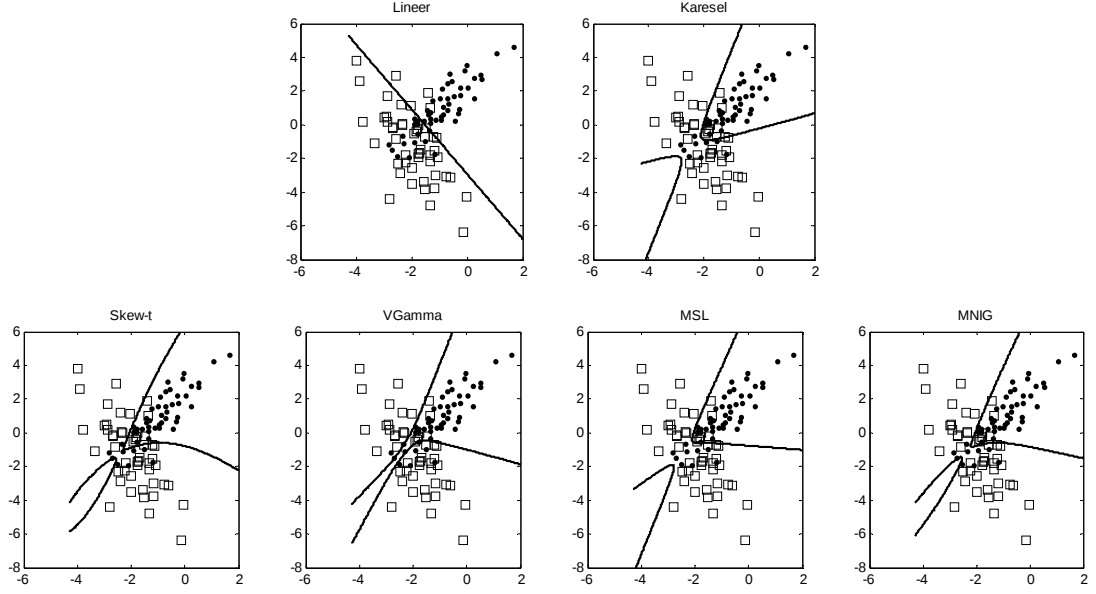
Şekil 7.2 Örnek 2 için diskriminant fonksiyonları

Örnek 3 (Normal dağılım, farklı varyans) : Bu örnekte sırasıyla $\underline{\mu}_1 = [-1; 1]$, $\underline{\Sigma}_1 = [1, 1.5; 1.5, 3]$ ve $\underline{\mu}_2 = [-2; -1]$, $\underline{\Sigma}_2 = [1, -1.5; -1.5, 5]$ parametrelili normal dağılıma sahip iki kitleden 50'şer birimlik örneklem üretilip hatalı sınıflandırma oranları elde edilmiştir. Çizelge 7.3'de 10 tekrar sonucunda elde edilen ortalama hatalı sınıflandırma oranları görülmektedir.

Çizelge 7.3 Örnek 3'e ait hatalı sınıflandırma oranları

Lineer	Karesel	Skew-t	VGamma	MSL	MNIG
0.20	0.15	0.14	0.14	0.15	0.15

Aşağıdaki grafikte ele alınan diskriminant fonksiyonlarının belirledikleri bölgeler görülmektedir.



Şekil 7.3 Örnek 3 için diskriminant fonksiyonları.

Örnek 4 (Normal dağılım, farklı varyans, düşük örneklem) : Bir önceki örnekte ele alınan kitlelerden 10'ar birimlik örneklem simülasyon yoluyla elde edilmiş, 20 tekrar sonucunda elde edilen ortalama hatalı sınıflandırma oranları çizelge 7.4'de verilmiştir.

Çizelge 7.4 Örnek 4'e ait hatalı sınıflandırma oranları

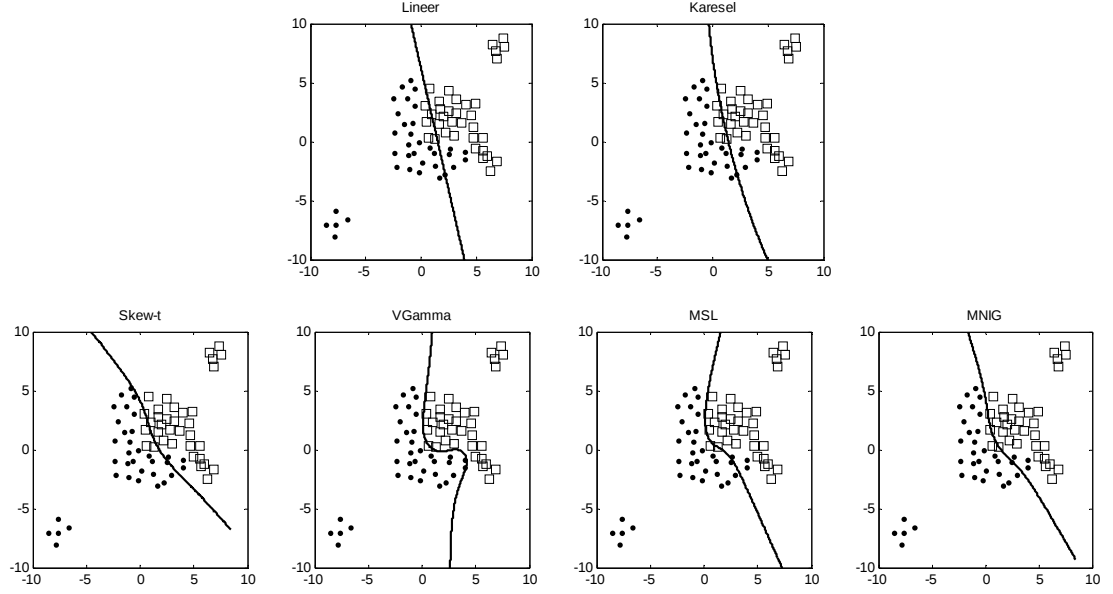
Lineer	Karesel	Skew-t	VGamma	MSL	MNIG
0.130	0.110	0.095	0.080	0.010	0.080

Örnek 5 (Sunı üretilmiş veri) : Bu örnekte suni olarak üretilmiş bir veri setinin sınıflandırılması ele alınmıştır. Çizelge 7.5'de hatalı sınıflandırma oranları görülmektedir.

Çizelge 7.5 Örnek 5'e ait hatalı sınıflandırma oranları

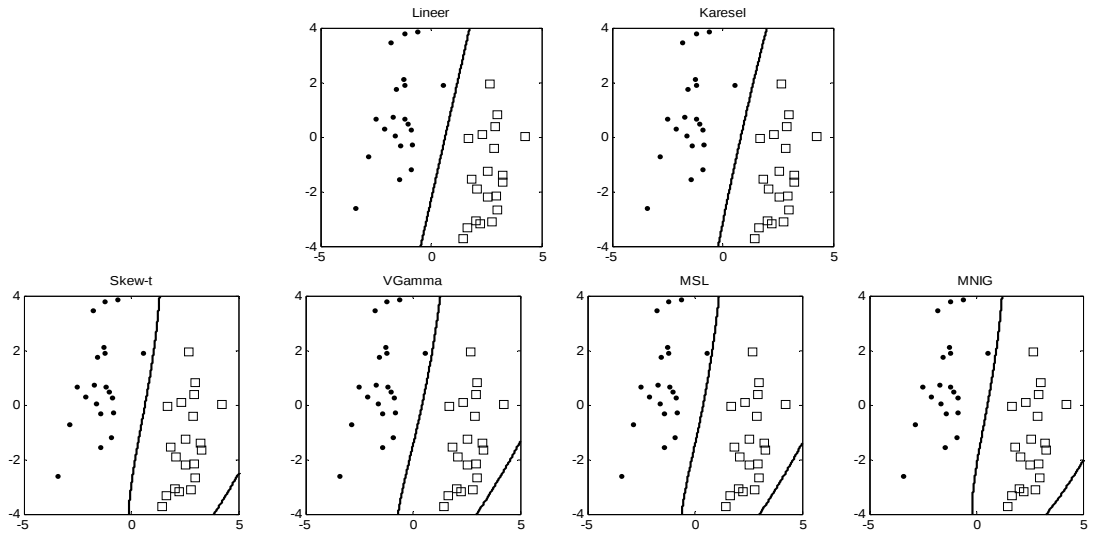
Lineer	Karesel	Skew-t	VGamma	MSL	MNIG
0.129	0.129	0.100	0.014	0.071	0.086

Şekil 7.3'de ele alınan diskriminant fonksiyonlarının belirledikleri bölgeler görülmektedir.



Şekil 7.4 Örnek 5 için diskriminant fonksiyonları

Örnek 6 (Kolay ayrılabilir veri) : Bu simülasyonda sırasıyla $\underline{\mu}_1 = [-1; 1]$, $\underline{\mu}_2 = [2; -1]$ ve $\Sigma_1 = \Sigma_2 = [1, 0.5; 0.5, 3]$ parametrelili normal dağılıma sahip iki kitleden 20'şer birimlik örneklem üretilmiştir. Tüm diskriminant fonksiyonlarının hatasız biçimde sınıflandırma yaptığı bu örneğe ilişkin diskriminant fonksiyonlarının belirledikleri bölgeler aşağıda görülmektedir.



Şekil 7.5 Örnek 6 için diskriminant fonksiyonları

Örneklerden, her iki grup için normal dağılımdan üretilen 50 birimlik örneklerle yapılan simülasyon çalışmalarında, ele alınan tüm diskriminant fonksiyonlarının hemen hemen aynı hatalı sınıflandırma oranına sahip oldukları görülmektedir. Bu durumda skew-t, vgamma, msl ve mnig diskriminant fonksiyonlarının, verilerin normal dağılması durumunda kullanılmasında herhangi bir sakınca olmadığını göstermektedir. Düşük örneklem hacimlerinde ise skew-t, vgamma, msl ve mnig diskriminant fonksiyonlarının, lineer ve karesel diskriminant fonksiyonlarından az bir farkla daha düşük hatalı sınıflandırma oranlarına sahip oldukları görülmektedir.

Aşırı üç değerlerin bulunduğu 4.örnekte özellikle varyans gamma diskriminant fonksiyonunun lineer ve karesel diskriminant fonksiyonlarına karşı başarısı dikkat çekicidir. Bu diskriminant fonksiyonlarından yaklaşık 9 kat daha az hatalı sınıflandırma yapan vgamma diskriminant fonksiyonunun belirlediği bölge, lineer ve karesel diskriminant fonksiyonlarının belirledikleri bölgelerle kıyaslandığında söz konusu başarının nedeni açıkça görülmektedir.

7.2 İris Çiçeği Verileri Üzerine Uygulama

Bu uygulamada, diskriminant analizi ile ilgili ilk çalışmada Fisher (1936) tarafından verilen ve sonrasında birçok çalışmada uygulama verisi olarak ele alınan iris çiçeği verisi kullanılmıştır. Veri setinin orijinal halinde iris çiçeğinin üç cinsinden (setosa, versicolor, virginica) 50'şer gözlem bulunmaktadır. Her örnek üzerinden çanak yaprağı uzunluğu ve genişliği ile taçyaprağı uzunluğu ve genişliği olmak üzere toplam 4 adet özellik ölçülmüştür.

Bu veri setindeki ilk uygulamada iris çiçeğinin setosa ve versicolor cinslerinden yukarıda bahsedilen 4 özellik üzerinden elde edilen gözlemler kullanılarak sınıflandırma yapılmış ve hatalı sınıflandırma oranları aşağıdaki gibi elde edilmiştir.

Çizelge 7.6 İris çiçeği için hatalı sınıflandırma oranları (4 parametre ve her iki gruptan 50 birimlik örneklem ile)

Lineer	Karesel	Skew-t	VGamma	MSL	MNIG
0.030	0.030	0.020	0.010	0.010	0.010

En az hatalı sınıflandırma oranının VGamma, MSL ve MNIG diskriminant fonksiyonları ile elde edildiği görülmektedir.

Bu veri setindeki ikinci uygulamada ise iris çiçeğinin setosa ve versicolor cinslerinden çanak yaprağı uzunluğu ile genişliği ele alınmış ve bu türlerin her birinden 10 tanesi rasgele seçilmiştir (training set). Gözlenen hatalı sınıflandırma oranlarına ilişkin sonuçlar çizelge 7.7’de verilmiştir. Lineer ve karesel diskriminant fonksiyonları örnek hacminin düşük olduğu bu örnekte 0.40 hatalı sınıflandırma yapmışlardır. Skew-t, VGamma ve MSL diskriminant fonksiyonları 0.25 hatalı sınıflandırma oranı ile en az hata oranına sahipken MNIG 0.30 hatalı sınıflandırma oranına sahiptir. Aşağıda her bir dağılım için parametrelerin tahminleri ve diskriminant fonksiyonlarının bu tahminler ile belirledikleri bölgeler görülmektedir.

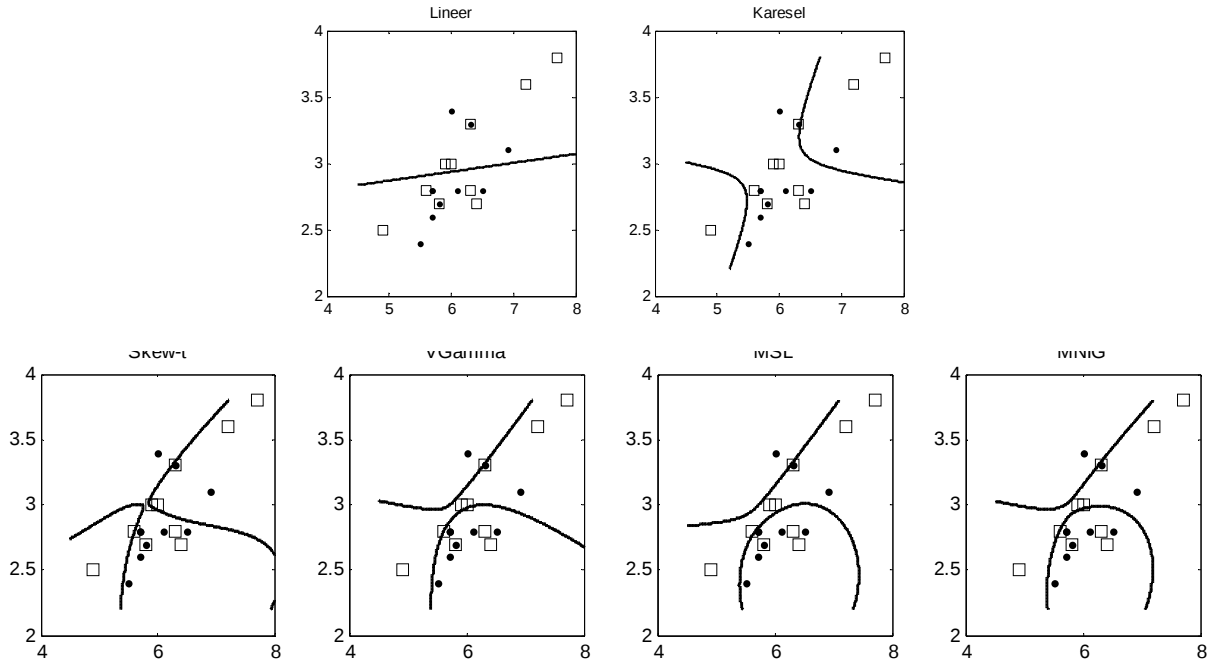
Çizelge 7.7 İris çiçeği için hatalı sınıflandırma oranları (2 parametre ve her iki gruptan 10 birimlik örneklem ile)

Lineer	Karesel	Skew-t	VGamma	MSL	MNIG
0.40	0.40	0.25	0.25	0.25	0.30

Çizelge 7.8’de parametrelerin tahmin değerleri, şekil 7.3’de ise ele alınan diskriminant fonksiyonlarının belirledikleri bölgeler görülmektedir.

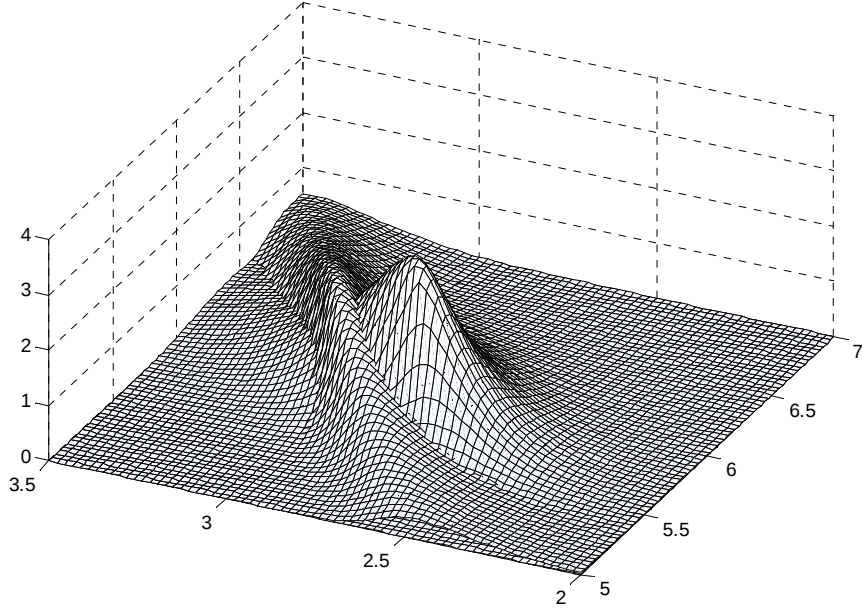
Çizelge 7.8 İris çiçeği için tahmini parametre değerleri (2 parametre ve her iki gruptan 10 birimlik örneklem ile)

	Lineer	Karesel	Skew-t	VGamma	MSL	MNIG
$\hat{\underline{\mu}}_1$	[6.03;2.87]		[5.58;2.66]	[5.60;2.67]	[5.68;2.62]	[5.49;2.58]
$\hat{\underline{\mu}}_2$	[6.21;3.02]		[5.74;3.07]	[5.79;2.99]	[5.90;3.00]	[5.73;3.01]
$\hat{\underline{\Sigma}}_1$	$\begin{bmatrix} 0.40 & 0.18 \\ 0.18 & 0.14 \end{bmatrix}$	$\begin{bmatrix} 0.19 & 0.08 \\ 0.08 & 0.09 \end{bmatrix}$	$\begin{bmatrix} 0.02 & 0.02 \\ 0.02 & 0.05 \end{bmatrix}$	$\begin{bmatrix} 0.03 & 0.02 \\ 0.02 & 0.05 \end{bmatrix}$	$\begin{bmatrix} 0.03 & 0.01 \\ 0.01 & 0.02 \end{bmatrix}$	$\begin{bmatrix} 0.02 & 0.02 \\ 0.02 & 0.05 \end{bmatrix}$
$\hat{\underline{\Sigma}}_2$		$\begin{bmatrix} 0.63 & 0.29 \\ 0.29 & 0.18 \end{bmatrix}$	$\begin{bmatrix} 0.49 & 0.26 \\ 0.26 & 0.14 \end{bmatrix}$	$\begin{bmatrix} 0.33 & 0.17 \\ 0.17 & 0.10 \end{bmatrix}$	$\begin{bmatrix} 0.22 & 0.11 \\ 0.11 & 0.07 \end{bmatrix}$	$\begin{bmatrix} 0.40 & 0.21 \\ 0.21 & 0.12 \end{bmatrix}$
$\hat{\underline{\gamma}}_1$	-	-	[0.19;0.09]	[0.21;0.10]	[0.12;0.09]	[0.32;0.17]
$\hat{\underline{\gamma}}_2$	-	-	[0.33;-0.04]	[0.22;0.02]	[0.11;0.01]	[0.28;0.01]
Diğer	-	-	$\hat{v}_1 = 2.89$ $\hat{v}_2 = 6.51$	$\hat{\lambda}_1 = 1.96$ $\hat{\lambda}_2 = 1.96$ $\hat{\psi}_1 = 1.95$ $\hat{\psi}_2 = 2.00$	-	$\hat{\chi}_1 = 2.70$ $\hat{\chi}_2 = 3.03$ $\hat{\psi}_1 = 1.02$ $\hat{\psi}_2 = 0.98$



Şekil 7.6 İris çiçeği için diskriminant fonksiyonları

Şekil 7.7’de yer alan grafik en az hatalı sınıflandırma oranının elde edildiği dağılımlardan biri olan varyans gamma dağılımının, her iki cins için tahmini parametrelerle çizilen olasılık yoğunluk fonksiyonlarını göstermektedir.



Şekil 7.7 İris çiçeği için tahmini parametrelerle çizilmiş varyans gamma dağılımlarının olasılık yoğunluk fonksiyonları

Uygulamanın devamı olarak, setosa ve versicolor cinslerinden rasgele seçilen ve diskriminant fonksiyonlarının oluşturulmasında kullanılan toplam 20 örnek dışında kalan 80 örnek, yukarıda oluşturulmuş olan diskriminant fonksiyonları ile sınıflandırılmış ve hatalı sınıflandırma oranları çizelge 7.9’daki gibi elde edilmiştir.

Çizelge 7.9 İris çiçeği için hatalı sınıflandırma oranları (2 parametre ve 80 birimlik örneklem ile)

Lineer	Karesel	Skew-t	VGamma	MSL	MNIG
0.4125	0.4000	0.4250	0.3875	0.3875	0.3750

Buna göre kalan 80 örneğin her bir gruptan alınan 10'ar birimlik örnek ile elde edilen diskriminant fonksiyonları ile sınıflandırılmasında en az hatalı sınıflandırma oranının 0.375 ile MNIG diskriminant fonksiyonuna ait olduğu görülmektedir.

7.3 Sporcu Verileri Üzerine Uygulama

Bu uygulamada 15 amatör ve 15 profesyonel sporcu, bir tür çeviklik testi olan 505 testine ve 30 metre koşu testine tabi tutulmuştur. Sporcular üzerinden elde edilen skorlar iki boyutlu bir rasgele vektör olarak ele alınmış ve gözlem değerleri ile sınıflandırma yapılmıştır. Elde edilen hatalı sınıflandırma oranları çizelge 7.10'da verilmiştir.

Çizelge 7.10 Sporcu verileri için hatalı sınıflandırma oranları

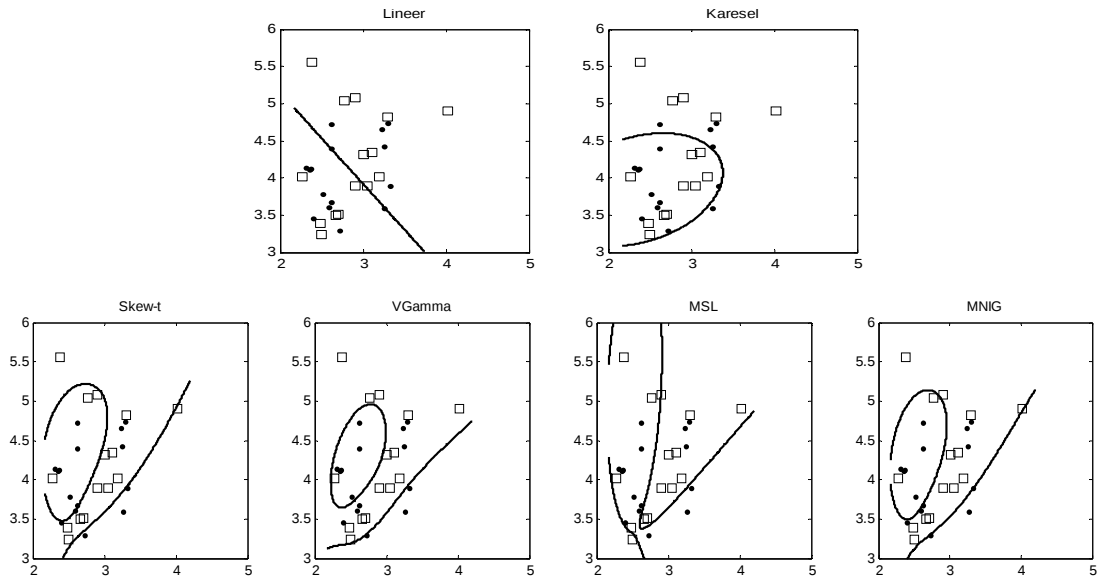
Lineer	Karesel	Skew-t	VGamma	MSL	MNIG
0.4000	0.5000	0.2667	0.2333	0.2333	0.3000

En düşük hatalı sınıflandırma oranının 0.23 ile verilerin varyans gamma yada çarpık Laplace dağıldığı varsayımı altında elde edildiği gözlenmiştir.

Aşağıda her bir dağılım için parametrelerin tahminleri ve diskriminant fonksiyonlarının bu tahminler ile belirledikleri bölgeler görülmektedir.

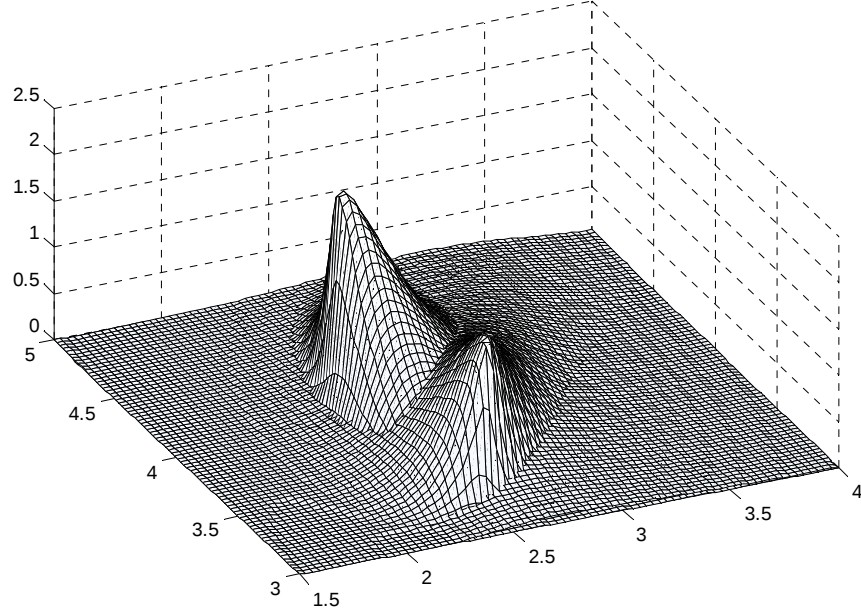
Çizelge 7.11 Sporcu verileri için tahmini parametre değerleri (2 parametre ve her iki gruptan 15 birimlik örneklem ile)

	Lineer	Karesel	Skew-t	VGamma	MSL	MNIG
$\hat{\underline{\mu}}_1$	[2.76; 4.04]		[2.13; 3.72]	[2.31; 4.13]	[2.39; 3.46]	[2.21; 3.84]
$\hat{\underline{\mu}}_2$	[2.88; 4.23]		[2.80; 3.42]	[2.51; 3.24]	[2.69; 3.47]	[2.72; 3.27]
$\hat{\underline{\Sigma}}_1$	$\begin{bmatrix} 0.17 & 0.08 \\ 0.08 & 0.36 \end{bmatrix}$	$\begin{bmatrix} 0.15 & 0.05 \\ 0.05 & 0.22 \end{bmatrix}$	$\begin{bmatrix} 0.02 & -0.02 \\ -0.02 & 0.15 \end{bmatrix}$	$\begin{bmatrix} 0.02 & 0.02 \\ 0.02 & 0.13 \end{bmatrix}$	$\begin{bmatrix} 0.02 & -0.02 \\ -0.02 & 0.05 \end{bmatrix}$	$\begin{bmatrix} 0.01 & -0.00 \\ -0.00 & 0.12 \end{bmatrix}$
$\hat{\underline{\Sigma}}_2$		$\begin{bmatrix} 0.19 & 0.09 \\ 0.09 & 0.50 \end{bmatrix}$	$\begin{bmatrix} 0.14 & 0.11 \\ 0.11 & 0.12 \end{bmatrix}$	$\begin{bmatrix} 0.09 & 0.03 \\ 0.03 & 0.03 \end{bmatrix}$	$\begin{bmatrix} 0.07 & 0.05 \\ 0.08 & 0.07 \end{bmatrix}$	$\begin{bmatrix} 0.13 & 0.09 \\ 0.09 & 0.09 \end{bmatrix}$
$\hat{\underline{\gamma}}_1$	-	-	[0.48; 0.24]	[0.26; -0.05]	[0.13; 0.20]	[0.29; 0.11]
$\hat{\underline{\gamma}}_2$	-	-	[0.04; 0.43]	[0.22; 0.59]	[0.07; 0.26]	[0.09; 0.58]
Diğer	-	-	$\hat{\nu}_1 = 7.03$ $\hat{\nu}_2 = 3.51$	$\hat{\lambda}_1 = 1.68$ $\hat{\lambda}_2 = 1.61$ $\hat{\psi}_1 = 1.88$ $\hat{\psi}_2 = 1.89$	-	$\hat{\chi}_1 = 3.54$ $\hat{\chi}_2 = 2.63$ $\hat{\psi}_1 = 1.00$ $\hat{\psi}_2 = 1.05$



Şekil 7.8 Sporcu verileri için diskriminant fonksiyonları

Şekil 7.9’da yer alan grafik en az hatalı sınıflandırma oranının elde edildiği dağılımlardan biri olan varyans gamma dağılımının, amatör ve profesyonel sporcular için tahmini parametrelerle çizilen olasılık yoğunluk fonksiyonlarını göstermektedir.



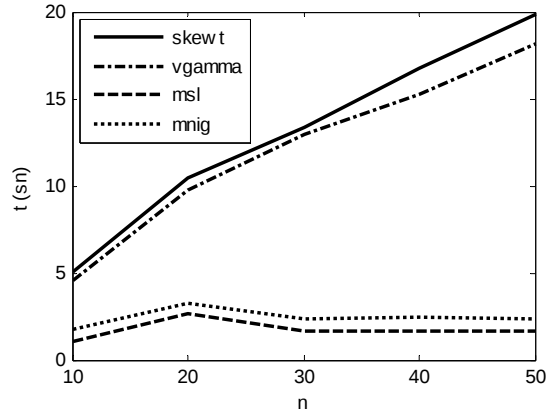
Şekil 7.9 Sporcu verileri için tahmini parametrelerle çizilmiş varyans gamma dağılımlarının olasılık yoğunluk fonksiyonları

Grafikte yukarıda aynı dağılım için çizilen bölgeler açıkça görülebilmektedir.

8. SONUÇ

Bu çalışma kapsamında normal dağılım varsayımının sağlanmadığı, özellikle çarpık dağılımlardan gözlenmiş verilerin diskriminant analizi ile sınıflandırılması ele alınmıştır. Bu amaçla çarpıklık ve basıklık özellikleri bakımından son derece esnek bir aile olan çok değişkenli genelleştirilmiş hiperbolik dağılım ailesinin bazı özel durumları ele alınmış ve gözlemlerin geldiği kitlelerin bu dağılımlar olduğu varsayımı altında diskriminant fonksiyonları elde edilmiştir.

Çok değişkenli genelleştirilmiş hiperbolik dağılım ailesi, sahip olduğu esnek özelliklere karşın, uygulamada parametrelerin tahmin edilmesi gerektiği durumlarda üzerinde çalışılması zor bir dağılım ailesi haline gelmektedir. Parametrelerin tahmininde kullanılan EM algoritmasının yakınsama süresi, tahmin edilen parametre sayısı, gözlem sayısı ve boyut sayısı ile birlikte artmaktadır. Aşağıdaki grafikte ele alınan her bir dağılım için EM algoritmasını 100 yinelemesinin, farklı örneklem hacimlerinde ne kadar sürdüğü görülmektedir.

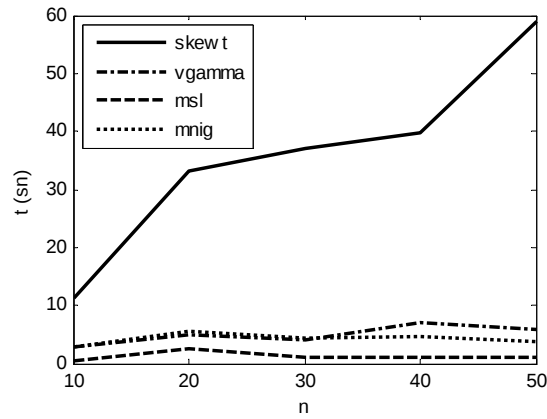


Şekil 8.1 EM algoritmasında süreler (100 yineleme)

Grafikten görüleceği üzere özellikle en çok parametre sayısına sahip varyans gamma (vgamma, 5 parametre) ve çarpık-t (skewt, 4 parametre) dağılımında tahmin süresi, örnek hacmine bağlı olarak artış göstermektedir. Her iki dağılım için kullanılan EM

algoritmalarında özellikle ξ_i değerlerinin hesaplanması sırasında gerek duyulan Bessel fonksiyonunun türevi, algoritmanın en yavaş noktası olarak gözlenmiştir. Buna karşın söz konusu değerin hesaplanmasının ihtiyaç duyulmadığı çarpık Laplace (msl, 3 parametre) ve normal ters Gauss (mnig, 4 parametre) dağılımlarında iterasyon süresinin örnek hacminden fazla etkilenmediği gözlenmiştir.

Aşağıdaki grafik EM algoritmasının yakınsama sürelerinin, çalışma kapsamında ele alınan dağılımlarda, değişik örneklem hacimlerine göre nasıl değiştiğini göstermektedir. Karşılaştırmanın yapılabilmesi için yakınsama şartı, tüm dağılımlarda ortak olan $\underline{\mu}, \underline{\Sigma}$ ve $\underline{\gamma}$ parametrelerindeki toplam değişimin 10^{-2} 'den küçük olması olarak alınmıştır. Ayrıca bu parametrelerin başlangıç değerleri tüm dağılımlar için aynı alınmıştır.



Şekil 8.2 Em algoritmasında süreler (yakınsama koşulu altında)

Özellikle çarpık-t dağılımı için yakınsama süresi örneklem hacmi ile hızla artmaktadır. Aynı durumun çarpık-t gibi düşük iterasyon hızına sahip varyans gamma dağılımında da görülmesinin beklenmesine karşın, bu dağılımda örnek hacminin artması ile yakınsama süresi fazla değişmemektedir. Bu durum varyans gamma dağılımının Çarpık-t dağılımından daha hızlı yakınsadığı anlamına gelmektedir.

EM algoritmalarının yakınsama hızları diskriminant fonksiyonlarının oluşturulması sürelerine direkt olarak etki etmektedir. Örneğin 7.Bölümde ele alınan ilk örnekte sınıflandırma süreleri ve hatalı sınıflandırma oranları çizelge 8.1'deki gibi gözlenmiştir.

Çizelge 8.1 Örnek 1 için sınıflandırma süreleri

	Lineer	Karesel	Skew-t	VGamma	MSL	MNIG
Hatalı Sınıflandırma Oranı	0.12	0.11	0.12	0.11	0.12	0.16
Sınıflandırma Süresi (sn)	0.11	0.04	76.54	10.21	1.99	8.28

Uygulama bölümünde ele alınan örneklerden görülebileceği üzere en az hatalı sınıflandırma oranına sahip olan diskriminant fonksiyonunun, kitlelerin varyans gamma dağılımına sahip olduğu varsayımı altında oluşturulan diskriminant fonksiyonu olduğu gözlenmiştir. Bu durumun parametre sayısı ile ilişkili olduğu düşünülmektedir. Parametre sayısının fazla olması dağılımın şeklinin belirlenmesinde esneklik sağlamakta, böylece gözlemlere uygun bir dağılımın uydurulması daha sağlıklı olmaktadır.

EM algoritmalarının yakınsama süreleri (ya da sınıflandırma süreleri) ile hatalı sınıflandırma oranları birlikte ele alındığında çarpık Laplace (MSL) dağılımı dikkat çekmektedir. Hatalı sınıflandırma oranlarında sınıflandırma süresi göz önünde bulundurulursa, varyans gamma dağılımı ile çarpık Laplace dağılımı arasındaki fark, makul orandadır.

Çalışmada elde edilen ve yukarıda özetlenen bulgular ışığında gerek normal dağılım gösteren verilerde, gerekse çarpık veri yapılarında, 5.Bölümde ele alınan çok değişkenli çarpık Laplace dağılımı ile elde edilmiş diskriminant fonksiyonunun kullanılması önerilmektedir.

Bu tez çalışmasında ele alınmayan ikiden fazla kitle için sınıflandırma kurallarının oluşturulması, hatalı sınıflandırma maliyetlerinin eşit olmaması, önsel olasılıkların eşit olmaması gibi durumlar ile hatalı sınıflandırma olasılıklarının hesaplanması gibi konular sonraki çalışmalarda ele alınabilir.

KAYNAKLAR

- Aas, K. and Haff, I. H. 2006. The generalised hyperbolic skew student's-t distribution. *Journal of Financial Econometrics*, Vol. 2(4); pp. 275-309.
- Abramowitz, M. and Stegun, I. 1964. *Handbook of Mathematical Functions with Formulas, Graps and Mathematical Tables*. Dover Publications, 1046 p., New York.
- Amoha, R. and Kocherlakota., K. 1986. Errors of misclassification associated with the inverse Gaussian distribution. *Communications in Statistics - Theory and Methods*, Vol. 15(2); pp. 589-612.
- Anderson, T. 1984. *An Introduction to Multivariate Statistical Analysis*. Wiley Series, 675 s., Hoboken.
- Arslan, O. 2010. An alternative multivariate skew Laplace distribution: properties and estimation. *Statistical Papers*, Vol. 51(4); pp. 865-887.
- Azzalini, A. and Capitanio, A. 2003. Distributions generated by pertubation of symmetry with emphasis on a multivariate skew t distribution. *Journal of the Royal Statistical Society* , Vol. B65(3); pp. 579-602.
- Barndorff-Nielsen, O.E. 1977. Exponentially decreasing distributions for the logarithm of particle size. *Proceedings of the Royal Society of London*, Vol. 353(1674); pp. 401-419.
- Barndorff-Nielsen, O.E. 1978. Hyperbolic distributions and distributions on hyperbolae. *Scandinavian Journal of Statistics*, Vol. 5(3); pp. 151-157.
- Barndorff-Nielsen, O.E. 1997. Normal inverse Gaussian distributions and stochastic volatility modeling. *Scandinavian Journal of Statistics*, Vol. 24(1); pp. 1-13.
- Batsidis, A. 2010. Errors of misclassification in discrimination with data from truncated t populations. *Statistical Papers*, Vol. 53(2); pp. 281-298.
- Blaesild, P. 1981. The two-dimensional hyperbolic distribution and related distributions,with an application to Johannsen's bean data. *Biometrika*, Vol. 68(1); pp. 251-263.
- Chhikara, R.S. and Folks, J.L. 1989. *Inverse Gaussian Distribution : Theory, Methodology, and Applications*. Marcel Dekker Incorporated, 213 p., New York.
- Damsleth, E. and El-Shaarawi, A. H. 1989. ARMA models with double exponentially distributed noise. *Journal of the Royal Statistical Society*, Vol. B51(1); pp. 61-69.

- Eberlein, E. and Keller, U. 1995. Hyperbolic distributions in finance. *Bernoulli*, Vol. 1(3); pp. 281-299.
- Esterling, R.G. 1978. Exponential responses with double exponential measurement error - A model for steam generator inspection. *Proceedings of the DOE Statistical Symposium*, pp. 90-110.
- Fernandez, C. and Steel, M. 1998. On bayesian modeling of fat tails and skewness. *Journal of the American Statistical Association*, Vol. 93(441); pp. 359-372.
- Fisher, R.A. 1936. The use of multiple measurements in taxonomic problems. *Annals of Eugenics*, Vol. 7(2); pp. 179-188.
- Freed, N. and Glover, F. 1986. Evaluating alternative linear programming models to solve the two-group discriminant problem. *Decision Sciences*, Vol. 17(2); pp. 151-162.
- Geisser, S. 1964. Posterior odds for multivariate normal classification. *Journal of the Royal Statistical Society*, Vol. 26(1); pp.69-76.
- Gupta, A. 1986. On a classification rule for multiple measurements. *Computers & Mathematics with Applications*. Vol. 12(2); pp. 301-308.
- Hsu, D.A. 1979. Long tailed distributions for position errors in navigation. *Journal of the Royal Statistical Society*, Vol. C28(1); pp. 62-72.
- Hu, W. 2005. Calibration Of Multivariate Generalized Hyperbolic Distributions Using The EM Algorithm, With Applications In Risk Management, Portfolio Optimization And Portfolio Credit Risk. PhD Thesis. Florida State Univesity, 97s., Florida.
- Hu, W. and Kercheval, A.N. 2010. Portfolio optimization for Student t and skewed t returns. *Quantitative Finance* , Vol. 10(1); pp. 91-105.
- Hubert, M. and Driessen, K.V. 2004. Fast and robust discriminant analysis. *Computational Statistics & Data Analysis*, Vol. 45(2); pp. 301-320.
- John, S. 1961. Errors in discrimination. *The Annals of Mathematical Statistics*, Vol. 32(4); pp. 1125-1144.
- Jones, M.C. and Faddy, M.J. 2003. A skew extension of the t distribution, with applications. *Journal of the Royal Statistical Society*, Vol. B65(2); pp. 159-174.
- Kalashnikov, V. 1997. Geometric sums: Bounds for rare events with applications. *Kluwer Acadeic Publishers*, 265 p., Dordrecht.
- Kotz, S., Kozubowski, T.J. and Podgorski, K. 2001. An asymmetric multivariate Laplace distribution. *Birkhäuser-The Laplace Distribution and Generalizations*, pp. 239-272, Boston.

- Laplace, P.S. 1774. Memoire sur la probabilité des causes par les événements. Memoirea de Mathematique et de Physique, Vol. 6; pp. 621-656.
- Lee, H.S. and Chen, B. 2009. Generalized likelihood ratio discriminant analysis. Automatic Speech Recognition & Understanding, ASRU 2009; pp. 158-163.
- Logan, T. and Gupta, A. 1993. Bayesian discrimination using multiple observations. Communications in Statistics - Theory and Methods Vol. 22(6); pp. 1735-1754.
- Madan, D.B. and Senata, E. 1990. The variance gamma model for share market returns. The Journal of Business, Vol. 63(4); pp. 511-524.
- Mahmoud, M.A.W. and Moustafa, M. 1995. Errors of misclassification associated with gamma distribution. Mathematical and Computer Modelling, Vol. 22(3); pp. 105-119.
- McNeil, A.J., Frey R. and Embrechts P. 2005. Quantitative Risk Management: Concepts, Techniques and Tools. Princeton University Press, 608 p., Princeton.
- Moustafa, H. and Ramadan, S. 2005. Errors of misclassification and their probability distributions when parent populations are Gompertz. Applied Mathematics and Computation, Vol. 163; pp. 423-442.
- O'Neill, T.J. 1978. Normal discrimination with unclassified observations. Journal of the American Statistical Association, Vol. 73(364); pp. 821-826.
- Prause, K. 1999. The generalized hyperbolic models: Estimation, financial derivatives and risk measurement. PhD Thesis. 167 p., University of Freiburg. Freiburg.
- Sutradhar, B.C. 1990. Discrimination of observations into one of two t populations. Biometrics, Vol. 46(3); pp. 827-835.
- Thabane, L. and Drekcic, S. 2004. Discrimination between two generalized multivariate modified bessel populations. International Journal of Statistical Sciences, Vol. 3; pp. 209-219.
- Welch, B. 1939. Note on discriminant functions. Biometrika, Vol. 31(2); pp. 218-220.

ÖZGEÇMİŞ

Adı Soyadı : Abdullah YILMAZ

Doğum Yeri : Ankara

Doğum Tarihi : 16.03.1980

Medeni Hali : Evli, 1 çocuk babası

Yabancı Dili : İngilizce

Eğitim Durumu (Kurum ve Yıl):

Lise : Sincan Lisesi (1997)

Üniversite : Kırıkkale Üniversitesi, Fen Edebiyat Fakültesi,
İstatistik Bölümü (2002)

Yüksek Lisans : Kırıkkale Üniversitesi, Fen Bilimleri Enstitüsü,
Matematik Anabilim Dalı (2008)

Çalıştığı Kurum / Kurumlar ve Yıl :

- Kırıkkale Üniversitesi, Fen Edebiyat Fakültesi, İstatistik Bölümü,
Araştırma Görevlisi (2003-2010)
- Kırıkkale Üniversitesi, Fen Edebiyat Fakültesi, İstatistik Bölümü,
Öğretim Görevlisi (2010-...)

Yayınları :

Yılmaz A., Arslan O. 2012. Çarpık-t Dağılımı ile Diskriminasyon. Kırıkkale Üniversitesi Bilimde Gelişmeler Dergisi, 1(2);21-32.