

T.C.
MUNZUR ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ



OLASILIK DAĞILIMLARININ RIEMANN GEOMETRİSİ

Merve YAPRAK ONAT

YÜKSEK LİSANS TEZİ
HESAPLAMALI BİLİMLER VE MÜHENDİSLİK ANABİLİM DALI

DANIŞMAN
Doç. Dr. İnan ÜNAL

TUNCELİ-2025

T.C.
MUNZUR ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

OLASILIK DAĞILIMLARININ RIEMANN GEOMETRİSİ

Merve YAPRAK ONAT
(220210029)

YÜKSEK LİSANS TEZİ
HESAPLAMALI BİLİMLER VE MÜHENDİSLİK ANABİLİM DALI

DANIŞMAN
Doç. Dr. İnan ÜNAL

TUNCELİ-2025

T.C.
MUNZUR ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

OLASILIK DAĞILIMLARININ RIEMANN GEOMETRİSİ

Merve YAPRAK ONAT
YÜKSEK LİSANS TEZİ
HESAPLAMALI BİLİMLER VE MÜHENDİSLİK ANABİLİM DALI

Bu tez 30/05/2025 tarihinde aşağıdaki jüri üyeleri tarafından **oybirliği** ile kabul edilmiştir.

İmza	İmza	İmza
Doç. Dr. Ramazan SARI (Amasya Üniversitesi)	Doç. Dr. İnan ÜNAL (Munzur Üniversitesi)	Doç. Dr. Selim Orhun SUSAM (Munzur Üniversitesi)
BAŞKAN	DANIŞMAN	ÜYE

Bu tez, Hesaplamalı Bilimler ve Mühendislik Anabilim Dalı'nda hazırlanmıştır.

Doç. Dr. Nagihan KARAASLAN AYHAN
Enstitü Müdürü

Bu tez çalışması Munzur Üniversitesi Bilimsel Araştırma Projeleri Koordinasyon Birimi tarafından desteklenmiştir.

Proje No: YLMUB024-13

NOT: Bu tezde kullanılan özgün ve başka kaynaktan yapılan bildirişlerin, çizelge, şekil ve fotoğrafların kaynak gösterilmeden kullanımı, 5846 sayılı “Fikir ve Sanat Eserleri Kanunu”ndaki hükümlere tabidir.

30/05/2025

ETİK İLKE VE KURALLARA UYGUNLUK BEYANNAMESİ

Bu tezin bana ait, özgün bir çalışma olduğunu; çalışmamın hazırlık, veri toplama, analiz ve bilgilerin sunumu olmak üzere tüm aşamalarında bilimsel etik ilke ve kurallara uygun davrandığımı; bu çalışma kapsamında elde edilen tüm veri ve bilgiler için kaynak gösterdiğimi ve bu kaynaklara kaynakçada yer verdiğimi ve hiçbir şekilde “intihal içermediğini” beyan ederim. Herhangi bir zamanda, çalışmamla ilgili yaptığım bu beyana aykırı bir durumun saptanması durumunda, ortaya çıkacak tüm ahlaki ve hukuki sonuçları kabul ettiğimi bildiririm.

İmza
Merve YAPRAK ONAT

Danışman
Doç. Dr. İnan ÜNAL

TEŞEKKÜR

Yüksek lisans eğitimime başladığım ilk günden itibaren bilgi, tecrübe ve yol gösterici önerileriyle bu çalışmanın şekillenmesinde büyük katkı sağlayan, her aşamada desteklerini esirgemeyen değerli danışmanım Doç. Dr. İnan ÜNAL'a en içten teşekkürlerimi sunarım.

Ayrıca, tez sürecinde karşılaştığım her zorlukta yanımda olan, sabrı, anlayışı ve desteğiyle bana güç veren sevgili eşim Nail ONAT'a minnettarım. Hayatımın her evresinde yanımda olduğu gibi bu süreçte de desteğini esirgemediği için kendisine sonsuz teşekkür ederim.

Bu tez çalışmasını YLMUB024-13 numaralı tez projesi ile maddi olarak destekleyen Munzur Üniversitesi Bilimsel Araştırma Projeleri Koordinasyon Birimine teşekkürlerimi sunarım.

Merve YAPRAK ONAT
TUNCELİ-2025

İÇİNDEKİLER

ETİK İLKE VE KURALLARA UYGUNLUK BEYANNAMESİ.....	I
TEŞEKKÜR.....	II
İÇİNDEKİLER.....	III
ŞEKİLLER LİSTESİ	VI
TABLolar LİSTESİ	VII
SEMBOLLER LİSTESİ	VIII
ÖZET	IX
ABSTRACT	X
1. GİRİŞ.....	1
2. TEMEL KAVRAMLAR.....	4
2.1. Rasgele Değişken Kavramı	4
2.2. Bir Rasgele Değişkenin Beklenen Değeri	6
2.3. Bir Rasgele Değişkenin Varyansı.....	6
2.4. Olasılık Dağılımları	7
2.4.1. Kesikli Olasılık Dağılımları	8
2.4.2. Sürekli Rasgele Değişkenlerin Dağılımları	12
2.5. Log Olabilirlik Fonksiyonu	13
2.6. Öklid Uzayının Topolojisi ve Geometrisi	14
2.7. Manifoldlar	18
2.8. Manifoldların Riemann Geometrisi.....	23
2.8.1. Koneksiyon Kavramı ve Christoffel Sembolleri	23
2.8.2. Riemann Metriği	24
2.8.3. Riemann Eğrilik Tensörü	25
2.8.4. Geodezik Kavramı.....	26
3. OLASILIK MODELLERİ.....	28
3.1. İstatistiksel Model.....	28
3.2. İstatistiksel Modelin Geometrik Temsili	30
3.3. İstatistiksel Modelin Parametrelendirilmesi	32
3.4. Bazı Olasılık Dağılımlarının Geometrik Modelleri.....	33
3.4.1. Üstel Dağılımın Geometrik Modeli.....	33
3.4.2. Normal Dağılımın Geometrik Modeli	33
3.4.3. Gamma Dağılımının Geometrik Modeli	35

3.4.4. Beta Dağılımının Geometrik Modeli.....	36
3.4.5. Kesikli Dağılımın Geometrik Modeli	37
3.4.6. Poisson Dağılımın Geometrik Modeli.....	37
3.4.7. İki Değişkenli Normal Dağılımın Geometrik Modeli	37
3.4.8. Çok Değişkenli Normal Dağılımın Geometrik Modeli.....	39
3.4.9. İki İstatistiksel Modelin Çarpımı.....	40
3.5. Üstel Aile	41
3.5.1. Üstel Aileye Ait Bazı Özel Dağılımlar	42
3.5.2. Üstel Ailenin Geometrik Modellenmesi	47
3.6. Karışım Aile (Mixture Family).....	49
4. OLASILIK DAĞILIMLARININ RIEMANN GEOMETRİSİ.....	52
4.1. Fisher Bilgi Metriği	52
4.1.1. Fisher Bilgi Metriğinin Elde Edilmesi	52
4.1.2. Fisher Metriği Üzerine Sonuçlar	56
4.1.3. Bazı Dağılımların Fisher Metrikleri	57
4.2. Christoffel Sembolleri	59
4.2.1. Bazı Dağılımların Christoffel Sembolleri	59
4.3. Koneksiyon Bileşenleri.....	61
4.3.1. ∇^0 Levi-Civita Koneksiyonu	61
4.3.2. ∇^1 Koneksiyonu	62
4.3.3. ∇^{-1} Koneksiyonu	63
4.3.4. ∇^α Koneksiyonu	64
4.4. Geodezik Denklemleri	65
4.5. Çarpıklık Tensörü (Skewness Tensor)	67
4.6. Otoparalel Eğriler	70
4.7. Jeffrey Önceliği (Jeffrey's Prior).....	72
4.8. Bazı Özel Dağılımların Riemann Geometrisi	73
4.8.1. Normal Dağılım.....	73
4.8.2. Lognormal Dağılımı	76
4.8.3. Gamma Dağılımı	76
4.8.4. Bernoulli Dağılımı.....	77
4.8.5. Geometrik Olasılık Dağılımı	78
4.8.6. Weibull Dağılımı	78
5. RAYLEIGH DAĞILIMININ RIEMANN GEOMETRİSİ	82

5.1. Rayleigh Dağılımı ve Uygulamaları.....	82
5.2. Rayleigh Dağılımının Geometrik Modellenmesi	83
5.3. Rayleigh Dağılım Manifoldunun Geodezikleri	85
5.4. ∇^α Koneksiyonuna Göre Christoffel Sembolleri.....	86
5.5. Çarpıklık Tensörü	89
5.6. Laplace-Betrami Operatörü ve Sonlu Tipli Altmanifoldlar	89
5.7. Weibull Manifoldunun Bir Alt Manifoldu Olarak Rayleigh Manifoldu	92
5.8. Uygulamalar	92
5.8.1 Hessian Denklemi ve Potansiyel Fonksiyonlar	93
5.8.2 Rayleigh Dağılımının Entropisi	94
5.8.3 Sinyal İşleme Alanında Bazı Uygulamalar	96
6. SONUÇ VE ÖNERİLER.....	102
7. KAYNAKLAR	104
ÖZGEÇMİŞ	106

ŞEKİLLER LİSTESİ

Şekil 2.1.	Standart normal dağılım için örnek bir grafik	12
Şekil 3.1.	İmmersiyon ι her bir ξ parametresini bir p_ξ dağılımına eşler.	30
Şekil 3.2.	S istatistiksel modelinin iki parametrelendirmesi	32
Şekil 3.3.	Manifolddaki iki nokta arasında iki Gauss dağılımını parametrize eden bir $\gamma(t)$ eğrisi	34
Şekil 5.1.	Üstel dağılım ile Rayleigh dağılımının 3 boyutlu uzayda modellenmesi ve kesitler	85
Şekil 5.2.	Rayleigh dağılımı üzerinde geodeziklerin modellenmesi	87
Şekil 5.3.	$\mathbb{R}^2$ içinde gömülmüş . Rayleigh istatistiksel manifoldunun 1-tipli altmanifold olarak tasviri.....	91
Şekil 5.4.	Weibull manifoldu üzerinde Rayleigh manifoldu.....	93
Şekil 5.5.	Rayleigh dağılımının potansiyel fonksiyonları.....	94
Şekil 5.6.	Rayleigh kanal değişkenliğinin geometrik görselleştirmesi.....	97
Şekil 5.7.	Sönümleme bileşenlerinin modellenmesi.....	97
Şekil 5.8.	<i>Solda:</i> Anten ayrılığının bir fonksiyonu olarak tahmin edilen $\hat{\xi}$ ölçek parametreleri. <i>Sağda:</i> Rayleigh manifoldu üzerindeki 0,25 m referans noktasından geodezik uzaklıklar.....	99
Şekil 5.9.	Rayleigh bilgi manifoldunda radar hedef tespiti.....	100

TABLÖLER LİSTESİ

Tablo 5.1.	Geometrik uzaklığa dayanarak tahmin edilen parametreler.....	98
Tablo 5.2.	Fisher bilgi geometrisine dayalı Rayleigh ölçek parametreleri arasındaki geodezik mesafe matrisi	99



SEMBOLLER LİSTESİ

τ	: Topolojik yapı
$\mathbb{R}$	: Reel sayılar
C^∞	: Sürekli diferansiyellenebilir fonksiyonlar sınıfı
$\chi(M)$	: Diferansiyellenebilir vektör alanları kümesi
∇	: Koneksiyon
Γ_{ij}^k	: 2.tür Christoffel sembolleri
$\Gamma_{ij,k}$	: 1.tür Christoffel sembolleri
g_{ij}	: Riemann metrik
$f(x)$	: Olasılık yoğunluk fonksiyonu
$F(x)$	: Olasılık dağılım fonksiyonu
$E(x)$	: Bir rasgele değişkenin beklenen değeri
σ^2	: Bir rasgele değişkenin varyansı
$\wp(X)$	: Kuvvet kümesi
J	: Jakobiyen matrisi
$\psi(\xi)$	: Normalizasyon fonksiyonu
E_ξ	: Fisher bilgi metriğinin beklenen değeri

ÖZET

Bu tezde, olasılık dağılımlarının geometrik modellenmesi ve Fisher bilgi metriği kullanılarak Riemann geometrisi araçlarının bu geometrik modeller üzerinde çalıştırılması araştırılmıştır. Tez çalışması bilgi geometrisine bir giriş niteliğinde olup, olasılık teorisi ve Riemann geometrisinin temel araçlarını içermektedir. İstatistik modeller üzerinde metrik kavramı, koneksiyonlar ve eğrilikler incelenmiştir. Özel bir dağılım olarak Rayleigh dağılımı üzerinde elde edilen neticeler irdelenmiştir.

Anahtar kelimeler: Bilgi geometrisi, olasılık dağılımı, Riemann geometrisi, istatistiksel manifold



ABSTRACT

Riemannian Geometry of Probability Distributions

This thesis investigates the geometric modeling of probability distributions and the application of Riemannian geometry tools on these geometric models using the Fisher information metric. Serving as an introduction to information geometry, the study encompasses fundamental tools from probability theory and Riemannian geometry. Concepts such as metrics, connections, and curvatures on statistical models are examined. Results obtained specifically for the Rayleigh distribution are also discussed.

Anahtar kelimeler: Information geometry, probability distribution, Riemannian geometry, statistical manifold



1. GİRİŞ

Olasılık teorisi tüm bilimlerin başvurduğu önemli çalışma alanlarının başında gelir. Teoriye yapılan katkıların ortaya koyduğu istatistiksel sonuçlar ile birçok alanda önemli başarılar elde edilmiştir. Finanstan sağlığa, sosyal bilimlerden temel bilimlere ve yapay zekadan zirai bilimlere kadar hemen her alanda olasılık teorisi uygulamaları vardır. Olasılık teorisinin üzerinde yoğunlaştığı önemli başlıklardan biri olasılık dağılımlarıdır. Olasılık dağılımları, bir rassal değişkenin olası değerlerinin olasılıklarını temsil eden, istatistik ve olasılık teorisi alanlarında kullanılan matematiksel modellerdir. Sürekli ve kesikli rasgele değişkenlerin birçok özellikleri bu dağılımlar yardımı ile incelenebilir.

Olasılık dağılımları cebirsel ve analitik özelliklerinin yanı sıra geometrik bir yapıya da sahiptir. Tek parametrelili dağılımlar bir eğri ile temsil edilebilirken çok parametrelili dağılımlar yüzeyler şeklinde modellenenmektedir. Dahası yüksek parametrelili dağılımlar bir manifold özelliği taşıyabilmektedir. Dağılımların geometrisi, farklı geometrik araçların dağılımlar üzerinde çalışılması düşüncesine yöneltmiştir. Bu amaçla çeşitli çalışmalar yapılmıştır ve yapılmaktadır.

Riemann geometrisi ise diferansiyel geometrinin bir dalıdır ve çok-boyutlu uzayların eğriliklerini inceleyen bir alan olarak öne çıkar. Riemann geometrisi, uzayın eğriliği ve bu eğrilikle bağlantılı geometrik neticelerin incelenmesiyle ilgilenir. Öklid dışı geometrilerin en önemli örneklerinden biri olan Riemann geometrisi son yıllarda birçok farklı alanda uygulanabilir olmuştur. Bir uzay üzerinde Riemann geometrisi yapılabilmesi için bu uzayın bir diferansiyellenebilir manifold yapısı taşıması ve üzerinde bir metrik tanımlanabilmesi gereklidir. Bazı olasılık dağılımları diferansiyellenebilir manifold biçiminde tanımlanabilmektedir. Ayrıca bu dağılımlar üzerinde Fisher metriği tanımlanmaktadır. Fisher metriği, dejener olmayan bir Riemann metriğidir. Sonuç olarak, olasılık dağılımları bir Riemann manifoldu olarak modellenenilmekte ve bu manifold üzerinde Riemann geometrisi araçları çalıştırılmaktadır.

Bilgi geometrisi, modern geometrinin araçlarını kullanarak bilgi dünyasını incelemeyi amaçlayan bir yaklaşımdır. Bu alan, başlangıçta olasılık dağılımları ailesinin sahip olduğu doğal diferansiyel geometrik yapıyı araştırmak üzere ortaya çıkmıştır. Olasılık dağılımlarının geometrik yapısı son derece zengin ve dikkat çekicidir; yalnızca geleneksel istatistiksel analiz sınırlarının ötesine geçmekle kalmaz, aynı zamanda modellerin altında yatan yapıları da ortaya koyar. Bilgi geometrisi, diferansiyel geometri açısından değerli bir

araştırma alanı olmasının yanı sıra, istatistiksel çıkarıma derin içgörüler kazandırarak makine öğrenmesi ve bilgi kuramı gibi alanlarda da etkili bir rol oynar.

Bilgi geometrisi, istatistiksel modellerin yapısal özelliklerini anlamak amacıyla olasılık dağılımlarının diferansiyel geometrik yapısını inceleyen disiplinler arası bir alandır. Bu yaklaşım, özellikle Shun-ichi Amari'nin 1970'li yıllarda başlattığı çalışmalara dayanır ve bilgi kuramı ile diferansiyel geometrinin birleşiminden doğmuştur (Amari & Nagaoka, 2000). Bilgi geometrisinin temel yapı taşlarından biri, olasılık dağılımları manifolduna uygulanan Fisher bilgi metriği ile oluşturulan Riemannian manifold yapısıdır. Bu metrik, parametrelili istatistiksel modellerin doğal geometriye sahip olduğunu gösterir (Nielsen, 2020). Amari'nin katkılarında biri de yalnızca metrik değil, aynı zamanda çiftli affine bağlantılar (dually coupled affine connections) kavramını getirmesidir (Amari, 2016). Bu yapı, manifold üzerinde hem eğrilik hem de divergans gibi önemli kavramların daha genel biçimde tanımlanmasını mümkün kılar. Böylece bilgi geometrisi, yalnızca geometrik bir teori olmaktan öte, istatistiksel çıkarım, optimizasyon, makine öğrenmesi, sinyal işleme ve nörobilim gibi alanlara yönelik güçlü bir çerçeve sunar (Amari, 1998; Ay ve ark., 2017). Bu alan, klasik analiz yöntemlerinin ötesine geçerek, istatistiksel modellerin içsel yapısını açığa çıkarır. Örneğin, üstel aile ve karışım ailesi gibi model sınıflarının düz manifoldlar (dually flat manifolds) olarak tanımlanabilmesi, çeşitli algoritmaların — örneğin doğal gradyan inişi (natural gradient descent) — geometrik temelde analizine olanak tanır (Amari, 1998). Günümüzde bilgi geometrisi, özellikle doğal gradyan yöntemleri, Bayesçi çıkarımın geometrik yorumları, entropiye dayalı ölçümler, bilgiye dayalı uzaklık metrikleri ve topolojik veri analizi gibi konularda yoğun biçimde kullanılmaktadır. Bu yönüyle bilgi geometrisi, yalnızca kuramsal değil, aynı zamanda uygulamalı bilimlerde de etkili bir araç haline gelmiştir.

Bilgi geometrisi alanında birçok çalışma yapılmıştır. Bu çalışmalardan bazıları şu şekilde özetlenmiştir: Abdel-All ve arkadaşları, 2003'teki çalışmalarında rasgele yürüyüş dağılımının geometrisini incelemişlerdir (Abdel-All ve ark., 2003). Calin ve Udrişte 2014 yılında olasılık dağılımlarının ve istatistiksel modellerin diferansiyel geometrisini konu alan kapsamlı bir kitap yazmışlardır (Calin ve Udrişte, 2014). Branea ve Radulescu 2023 yılında yayınladıkları çalışmalarında Bernoulli ve üstel dağılımla ilgili önemli geometrik sonuçlar elde etmişlerdir (Branea ve Radulescu, 2023). Amari 2016'da yayınladığı kitapta bilgi geometrisini ve uygulamalarını kapsamlı bir biçimde incelemiştir (Amari, 2016). 2022

yılında Umut ako yksek lisans tez alıřmasında ters Gauss daėılımlarının Riemann geometrisini alıřmıřtır (ako, 2022).

Bu tez alıřmasında, olasılık daėılımlarının Riemann geometrisi baėlamında incelenmesi amalanmıřtır. Tezin ikinci blmnde, alıřmanın temelini oluřturan olasılık teorisinin temel kavramları ile Riemann geometrisine dair temel yapılar ayrıntılı biimde sunulmuřtur. nc blmde eřitli olasılık modelleri tanıtılmıř, ardından gelen drdnc blmde bu modellerin geometrik temsilleri ele alınmıřtır. Bu baėlamda olasılık daėılımları manifoldlarında tanımlanan Fisher bilgi metriėi temel alınarak, metrik tensr, affine koneksiyonlar, Christoffel sembolleri ve geodezik eėriler gibi diferansiyel geometrinin temel araları kullanılmıř; zel daėılım ailelerinin geometrik yapıları analiz edilmiřtir. Beřinci ve son blmde ise, istatistiksel modelleme ve sinyal iřleme alanlarında yaygın olarak kullanılan Rayleigh daėılımı bir boyutlu bir Riemann manifold olarak modellenmiř ve bu daėılımın geometrik zellikleri ayrıntılı olarak incelenmiřtir. Elde edilen teorik sonular, sinyal iřleme baėlamında rnek bir uygulama zerinden deėerlendirilerek alıřmanın uygulamalı yn ortaya konmuřtur.

2. TEMEL KAVRAMLAR

Bu bölümde, tez çalışmasının temelini oluşturan olasılık kavramları ile Riemann geometrisinin temel kavramları verilecektir. Konu hakkında detaylı bilgi edinmek için (Akdeniz, 2003) ve (Turgut Vanlı, 2022) kaynaklarına bakılabilir.

2.1. Rasgele Değişken Kavramı

Tanım 2.1.1. Değeri bir deney sonucuyla belirtilen bir değişkene rasgele değişken denir (Akdeniz, 2003).

Tanım 2.1.2. X bir rasgele değişken olsun. X in alabileceği değerlerin sayısı sonlu veya sayılabilir sonsuzlukta ise X 'e kesikli rasgele değişken denir (Akdeniz, 2003).

Örnek 2.1.1. Bir zarın atılması deneyinde örnek uzay $S = \{1,2,3,4,5,6\}$ dir. Bu örnek uzayda sonlu sayıda eleman (örnek nokta) vardır, bu örnek uzay üzerinde tanımlanan X rasgele değişkeni kesikli rasgele değişkendir (Akdeniz, 2003).

Tanım 2.1.3. X bir rasgele değişken olsun. X bir aralıkta ya da birden çok aralıkta her değeri alabiliyorsa X e sürekli rasgele değişken denir (Akdeniz, 2003).

Tanım 2.1.4. X , sonlu sayıdaki $x_1, x_2, \dots, x_N$ değerlerini $f(x_i) = P(X = x_i), i = 1, 2, \dots, N$ olasılıkları alabilen kesikli rasgele değişken olsun. Bu durumda

1. $f(x) \geq 0$, tüm x ler için
2. $\sum_{i=1}^N f(x_i) = 1$

koşullarını sağlayan $f(x)$ fonksiyonuna X in olasılık fonksiyonu denir (Akdeniz, 2003).

Tanım 2.1.5. Bir X rasgele değişkeninin dağılım fonksiyonu $F(x)$ ile gösterilir ve X in x e eşit ya da daha küçük olması olasılığıdır. Yani

$$F(x) = P(X \leq x) = \sum_{x_i \leq x} f(x_i)$$

dir (Akdeniz, 2003).

Tanım 2.1.6. X sayılabilir sonsuzluktaki $x_1, x_2, \dots, x_N, \dots$ değerlerini alan kesikli bir rasgele değişken olsun. Bu değerlere karşılık gelen olasılıklar $f(x_i) = P(X = x_i), i = 1, 2, \dots, N, \dots$ olmak üzere

- (i) $f(x) \geq 0$, tüm x ler için
- (ii) $\sum_{i=1}^{\infty} f(x_i) = 1$

koşullarını sağlayan $f(x)$ fonksiyonuna X rasgele değişkeninin olasılık dağılımı ya da olasılık fonksiyonu denir (Akdeniz, 2003).

Tanım 2.1.7. $X, (-\infty, \infty)$ aralığında tanımlanan sürekli rasgele değişken olsun.

(i) $f: \mathbb{R} \rightarrow [0,1], f(x) \geq 0, (-\infty < x < \infty)$,

(ii) $\int_{-\infty}^{+\infty} f(x)dx = 1$

koşullarını sağlayan $f(x)$ fonksiyonuna X rasgele değişkeninin olasılık yoğunluk fonksiyonu denir.

Ayrıca, $f(x)$ (i) ve (ii) yi sağlarsa sürekli X rasgele değişkeninin, $-\infty < c < d < +\infty$ olmak üzere c ve d arasında bulunması olasılığı $P(c < X < d) = \int_c^d f(x)dx$, $f(x)$ eğrisi, x -ekseni ve $x = c, x = d$ doğruları ile sınırlanan alandır. Ayrıca,

$$P(c < X < d) = P(c \leq x \leq d) = P(c \leq X < d) = P(c < X \leq d)$$

dir. Sürekli X rasgele değişkenin belli bir x değerini alması olasılığı sıfırdır. Yani $P(X = x) = 0$ dir (Akdeniz, 2003).

Tanım 2.1.8. $X, f(x)$ olasılık yoğunluk fonksiyonuna sahip sürekli rasgele değişken olsun. X in birikimli dağılım fonksiyonu

$$F(x) = P(X \leq x) = \int_{-\infty}^x f(s)ds$$

biçimindedir (Akdeniz, 2003).

Teorem 2.1.1.

(i) Sürekli bir X rasgele değişkeninin olasılık yoğunluk fonksiyonu f ve dağılım fonksiyonu F olsun. Bu takdirde bütün x değerlerinde diferansiyellenebilir F fonksiyonu için

$$f(x) = \frac{d}{dx}(F(x))$$

ve $a \leq b$ olmak üzere herhangi a ve b gerçel sayıları için $P(a \leq X \leq b) = F(b) - F(a)$ dir.

(ii) $X, x_1 < x_2 < \dots$ gibi sıralı $x_1, x_2, \dots$ değerlerini alabilen kesikli rasgele değişken olsun. $F(x)$, X ' in dağılım fonksiyonu ise bu takdirde $f(x_1) = F(x_1)$ ve $f(x_i) = P(X = x_i) = F(x_i) - F(x_{i-1})$ dir (Akdeniz, 2003).

2.2. Bir Rasgele Değişkenin Beklenen Değeri

Tanım 2.2.1. X , olasılık fonksiyonu f olan kesikli bir rasgele değişken olsun. X in $E(X)$ ile gösterilen beklenen değeri aşağıdaki gibi tanımlanır:

$$E(X) = x_1 \cdot f(x_1) + x_2 \cdot f(x_2) + \dots + x_N \cdot f(x_N) = \sum_{i=1}^N x_i f(x_i).$$

X rasgele değişkeni sayılabilir sonsuzluktaki $x_1, x_2, \dots, x_n, \dots$ sonuçlarını alıyorsa

$$E(X) = x_1 \cdot f(x_1) + \dots + x_N \cdot f(x_N) + \dots = \sum_{i=1}^{\infty} x_i f(x_i)$$

olur (Akdeniz, 2003).

$E(X)$ sayısına ayrıca X rasgele değişkeninin ortalaması ya da kitle (population) ortalaması denir. Beklenen değeri ya da ortalamayı göstermek için genellikle μ harfi kullanılır. $X, Y, \dots$ gibi birden çok rasgele değişkenle çalışıldığında μ ler için ortalamaları göstermek üzere indis kullanılır. Bu nedenle $\mu_x = E(X)$ ve $\mu_y = E(Y)$ yazılır.

Tanım 2.2.2. X sürekli bir rasgele değişken olsun. $f(x)$, X in olasılık yoğunluk fonksiyonu olmak üzere, X in beklenen değeri (ortalaması), $E(X)$

$$E(x) = \int_{-\infty}^{+\infty} x \cdot f(x) \cdot dx, -\infty < x < +\infty$$

şeklinde tanımlanır (Akdeniz, 2003).

2.3. Bir Rasgele Değişkenin Varyansı

Bir rasgele değişkenin beklenen değeri ya da ortalaması olasılık fonksiyonunun merkezi hakkında bilgi verir. Fakat ortalama değer bir deneyden bir diğerine rasgele değişkenin değerlerinin dağılımı, değişimi ya da yayılması ile ilgili bilgi vermez. Bunun için varyans kullanılır.

Tanım 2.3.1. X kesikli bir rasgele değişken olsun. X in ortalaması $E(X) = \mu$ ise X in varyansı, $\text{Var}(X)$ veya σ_x^2 ile gösterilir ve aşağıdaki gibi tanımlanır:

(i) X kesikli rasgele değişken ise

$$\sigma_x^2 = \text{Var}(X) = \sum_{i=1}^N (x_i - \mu)^2 \cdot f(x_i),$$

(ii) X sürekli rasgele değişken ise,

$$\sigma_x^2 = \text{Var}(X) = \int_{-\infty}^{+\infty} (x_i - \mu)^2 \cdot f(x_i) \cdot dx$$

dir. Yani, X in varyansı, X in kendi ortalamasından sapmasının karesinin ortalamasıdır.

X, μ ortalamalı kesikli ya da sürekli bir rasgele değişken olsun. X in standart sapması, σ_x , varyansın kareköküdür, ve aşağıdaki eşitlikle verilmiştir (Akdeniz, 2003):

$$\sigma_x = \sqrt{\text{Var}(X)}.$$

Tanım 2.3.2. X ve Y rasgele değişkenlerinin ortalamaları μ_x ve μ_y olmak üzere

$$E[(X - \mu_x) \cdot (Y - \mu_y)]$$

beklenen değerine X ve Y arasındaki kovaryans denir ve σ_{xy} ya da $\text{Cov}(X, Y)$ ile gösterilir.

X ve Y arasındaki kovaryansı hesaplamak için

$$\text{Cov}(X, Y) = E(XY) - E(X) \cdot E(Y) = E(XY) - \mu_x \cdot \mu_y$$

formülü kullanılır. Kovaryans X ve Y nin birbiriyle nasıl bir ilişkiye sahip olduğunu gösterir. X ve Y istatistiksel olarak birbirinden bağımsız iseler $E(XY) = E(X) \cdot E(Y)$ olacağından, bu durumda $\text{Cov}(X, Y) = 0$ bulunur. Bunun tersi doğru değildir. Yani, $\text{Cov}(X, Y) = 0$ olması X ve Y nin bağımsızlığını gerektirmez (Akdeniz, 2003).

Tanım 2.3.3. X ve Y , $E(X)$, $E(Y)$ ortalamaları ve $\text{Var}(X)$, $\text{Var}(Y)$ varyanslarına sahip olsunlar.

X ve Y arasındaki korelasyon ρ_{xy} ile gösterilir ve

$$\rho_{xy} = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X) \text{Var}(Y)}}$$

ile tanımlanır. ρ_{xy} 'ye Pearson korelasyon katsayısı denir ve $[-1, 1]$ aralığında değer alır. (Akdeniz, 2003).

2.4. Olasılık Dağılımları

Olasılık dağılımları, bir olayın farklı değerlerinin olasılıklarını tanımlayan matematiksel modellerdir. İstatistik ve olasılık teorisinde rastgele değişkenlerin davranışını açıklamak için kullanılırlar. Bu dağılımlar, bir olayın olasılıklarını temsil eden grafik veya matematiksel fonksiyonlar aracılığıyla ifade edilir. Olasılık dağılımları, istatistiksel analizlerde, modelleme çalışmalarında ve tahminleme süreçlerinde yaygın olarak kullanılır.

2.4.1. Kesikli Olasılık Dağılımları

Kesikli olasılık dağılımı, rastgele bir olayın belirli değerlerinin olasılıklarını gösteren matematiksel bir modeldir. Bu tür dağılımlar, ayrık veya belirli değer kümesine sahip rastgele değişkenlerin olasılık dağılımlarını ifade eder. Örnek olarak, zar atma durumu binom dağılımıyla, bir kart destesinden bir kart seçme durumu ise olasılık kütle fonksiyonu ile ifade edilebilir. Kesikli olasılık dağılımları, istatistiksel analizlerde ve olasılık teorisinde önemli bir rol oynar.

Bernoulli Dağılımı

Bir X rasgele değişkeni için yalnız iki sonuç varsa X e Bernoulli değişkeni denir. 1 değeri belli bir denemenin başarılı olması, 0 ise başarısızlığa karşılıktır.

Örnek 2.4.1. Paranın atılması, kavanozdan top çekilmesi.

Tanım 2.4.1. X rasgele değişkeni 0 ve 1 değerlerini alsın. X 'in olasılık fonksiyonu başarı olasılığı p , ($0 < p < 1$) ve başarısızlık olasılığı $1 - p = q$ olmak üzere

$$f(x) = P(X = x) = p^x$$

biçimindedir. Bu dağılıma Bernoulli dağılıma denir (Akdeniz, 2003).

Bernoulli dağılımının ortalaması (X 'in beklenen değeri) ve varyansı $\mu = E(X) = p$ ve $\sigma^2 = p \cdot q$ dir.

Binom Dağılımı

Birbirinden bağımsız n Bernoulli denemesinden başarılı olanların toplam sayısı X rasgele değişkeni olsun. Aşağıdaki koşulları sağlayan X e binom rasgele değişkeni denir (Akdeniz, 2003):

- (i) Deney n özdeş denemeden oluşmaktadır.
- (ii) Her bir deneme için yalnızca iki sonuç vardır: başarı (T) ve başarısızlık (F).
- (iii) Bir tek deneme için başarı olasılığı olan p her deneme için aynıdır. Başarısızlık olasılığı $q = 1 - p$ dir.
- (iv) Denemeler birbirinden bağımsızdır.

Teorem 2.4.1. Birbirinden bağımsız n Bernoulli denemesi için X , her bir denemede başarı olasılığı p , başarısızlık olasılığı q olan binom rasgele değişkeni ise, X 'in olasılık fonksiyonu:

$$f(x) = \binom{n}{x} \cdot p^x \cdot q^{n-x}, x = 0, 1, 2, \dots, n$$

dir. Bu dağılıma binom dağılımı denir (Akdeniz, 2003).

Geometrik Dağılım

Bir deneyin bağımsız Bernoulli denemelerden oluştuğu kabul edilsin. İlk başarının elde edilmesi için gereken denemelerin sayısı geometrik rasgele değişkendir.

Teorem 2.4.2. X , bir tek denemede başarısızlık olasılığı $q = 1 - p$ ve başarı olasılığı p olan geometrik rasgele değişken ise, X in olasılık fonksiyonu

$$f(x) = P(X = x) = q^{x-1} \cdot p, x = 1, 2, \dots$$

dir. İlk başarının elde edilmesi için gereken denemelerin sayısı X , 1, 2, 3, ...eğerlerinden biri olabilir. $x - 1$ ilk başarıdan önceki başarısızlıkların sayısı olsun. Bu durum aşağıdaki gibi gösterilebilir.

$$\underbrace{FF \dots F}_{x-1} T$$

O halde $x - 1$ başarısızlığı, başarının takip ettiği dizinin olasılığı $q^{x-1} \cdot p$ dir. Bu nedenle X rasgele değişkeninin olasılık fonksiyonu

$$f(x) = P(X = x) = q^{x-1} \cdot p, x = 1, 2, 3, \dots$$

dir. 1, 2, 3, ... denemede ilk başarının elde edilmesi olasılıkları aşağıdaki sonsuz serideki ardışık terimlere karşılık gelir:

$$f(x) = p + q \cdot p + q^2 p + \dots$$

Böylece, olasılıklar toplamı

$$\sum_{x=1}^{\infty} f(x) = \sum_{x=1}^{\infty} P(X = x) = p \cdot (1 + q + q^2 + \dots) = p \cdot \left(\frac{1}{1-q}\right) = 1$$

dir(Akdeniz, 2003).

Teorem 2.4.3. X rasgele değişkeninin olasılık fonksiyonu aşağıdaki gibi olsun.

$$f(x) = q^{x-1} p^x = 1, 2, \dots$$

Bu takdirde geometrik dağılım için ortalama değer ve varyans, sırasıyla $E(x) = \frac{1}{p}$ ve $\sigma^2 = \frac{q}{p^2}$ dir (Akdeniz, 2003).

Poisson Dağılımı

Poisson dağılımı, belirli bir zaman ya da alan içinde rastgele gerçekleşen olayların sayısını modellemek için kullanılır ve bu olaylar sayılabilir (kesikli) değerler alır. Sürekli bir zaman ya da uzayda gerçekleşen, birbirinden bağımsız olayların sayısını incelerken uygulanır.

Tanım 2.4.2. Verilmiş bir zaman aralığında bir alanda ya da hacimde başarıların sayısı, X rasgele değişkeni olsun. Aşağıdaki koşullar sağlanırsa X Poisson dağılımına aittir denir (Akdeniz, 2003) :

- (i) Deney, verilmiş bir birim zaman aralığı, alan ya da hacimde bir olayın elde edilmiş sayılarının sayılmasıyla oluşur.
- (ii) İki ayrık birim zamanda, alanda ya da hacimde elde edilebilecek başarıların sayıları birbirinden bağımsızdır.
- (iii) Birim zaman, alan ve hacimdeki başarı olasılığı, tüm birimler için aynıdır.
- (iv) Çok küçük bir zaman aralığı, alan ya da hacimde iki ya da daha çok başarının olması hemen hemen olanaksızdır. Yani bu durumda birden çok başarının olması olasılığı sıfıra yaklaşır.
- (v) Bir birim zaman, alan ya da hacimde bir sonucun ortalama elde edilmiş sayısı λ dır.

Örnek 2.4.2. Aşağıdaki örnekler verilebilir:

- Büyük bir şehirde ender rastlanan bin hastalıktan her yıl meydana gelen ölümlerin sayısı,
- Büyük bir şehirde trafiğin yoğun olduğu bir kavşakta aylık otomobil kazalarının sayısı,
- Bir üretim malındaki kusurların sayısı.

$X, 0, 1, 2, \dots$ değerlerini alabilen bir Poisson rasgele değişkeni olsun. X in olasılık fonksiyonu

$$f(x) = P(X = x) = \frac{e^{-\lambda} \cdot \lambda^x}{x!}, x = 0, 1, 2, \dots \lambda > 0$$

dır. Bu dağılıma Poisson dağılımı denir. Burada $e = 2,71828$ ve λ dağılımın ortalamasıdır.

Düzgün (Uniform) Dağılım

Düzgün dağılım, rastgele bir değişkenin belirli bir aralıktaki değerlerinin eşit olasılıklarla gerçekleştiği bir olasılık dağılımıdır. Bu dağılımın temel özelliği, olasılık yoğunluğunun belirli bir aralıkta sabit olmasıdır. Yani, her değer aralığı eşit olasılığa sahiptir. Düzgün dağılım genellikle $[a, b]$ aralığındaki rastgele değişkenleri ifade etmek için kullanılır, burada "a" minimum değeri, "b" ise maksimum değeri temsil eder. Bu tür bir dağılım, rastgele olayların homojen ve dengeli bir şekilde gerçekleştiği durumları modelleme amacıyla kullanılır.

Tanım 2.4.3. X rasgele değişkeni tümü eşit olasılıklı N sonuca sahipse, X e kesikli düzgün rasgele değişken denir (Akdeniz, 2003).

Örnek 2.4.3.

1. Bir parayı bir kez atalım. $X = 0$ yazı gelmesi sonucunu $X = 1$ tura gelmesi sonucunu gösterebilir. Burada X kesikli düzgün rasgele değişkendir.
2. Bir zar bir kez atılsın. Bu takdirde, X üste gelen sayı kesikli düzgün rasgele değişkendir.
3. 5'lik bir desteden bir kart çekelim. O halde X , çekilen herhangi bir kart ise bir kesikli düzgün rasgele değişkendir (Akdeniz, 2003).

Tanım 2.4.4. X rasgele değişkeninin alabileceği değerler $x_1, x_2, \dots, x_N$ olsun. X in olasılık fonksiyonu

$$f(x) = P(X = x) = \frac{1}{N}, x = x_1, x_2, \dots, x_N$$

dir. Bu dağılıma kesikli düzgün dağılım denir (Akdeniz, 2003).

Teorem 2.4.4 $f(x) = \frac{1}{N}, x = x_1, x_2, \dots, x_N$ kesikli düzgün dağılımın ortalaması ve varyansı

$$\mu = E(X) = \frac{N + 1}{2}$$

biçimindedir (Akdeniz, 2003).

2.4.2. Sürekli Rasgele Değişkenlerin Dağılımları

Normal Dağılım

Normal dağılım, istatistiksel analizlerde sıkça kullanılan bir olasılık dağılımıdır. Belli bir ortalama değer etrafında simetrik bir şekilde dağılan verileri temsil eder. Normal dağılımın belirgin bir özelliği, dağılımın şeklinin bir çan gibi simetrik olmasıdır; bu nedenle, "çan eğrisi" olarak da adlandırılır. Normal dağılım, merkezi limit teoremi gibi önemli istatistiksel teoremlerin temelini oluşturur ve birçok gerçek dünya olayının istatistiksel analizinde yaygın olarak kullanılır. Normal dağılıma sahip bir veri seti, genellikle belirli bir ortalama ve standart sapma ile tanımlanır. İstatistiksel testlerde ve tahminlemede kullanılan birçok yöntem, normal dağılıma dayanır ve bu dağılım, istatistiksel analizlerin temel bir aracıdır (Akdeniz, 2003).

Tanım 2.4.5. Sürekli bir X rasgele değişkeni için olasılık yoğunluk fonksiyonu

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, -\infty < x < \infty, -\infty < \mu < \infty, \sigma > 0$$

ise X normal dağılıma aittir denir. Burada μ dağılımın ortalaması ve σ dağılımın standart sapmasıdır. X normal dağılıma aitse $X \sim N(\mu, \sigma^2)$ simgesiyle gösterilir. $f(x)$ olasılık yoğunluk fonksiyonu olduğundan,

$$\int_{-\infty}^{+\infty} \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx = 1$$

yazılır (Akdeniz, 2003).

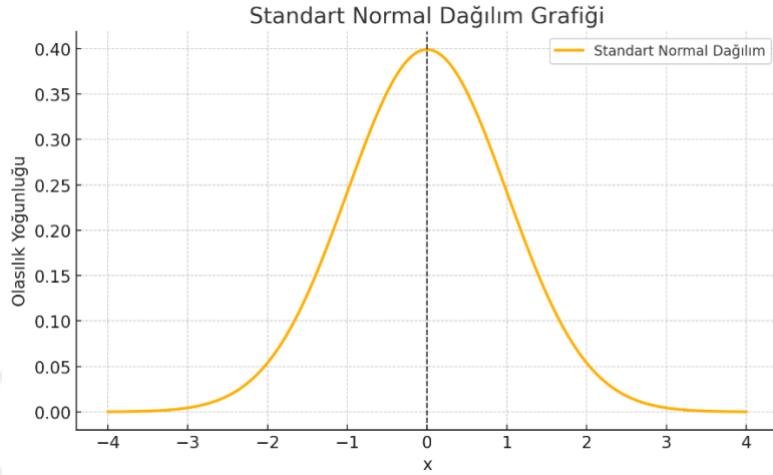
Standart Normal Dağılım

Standart normal dağılım, ortalaması 0 ve standart sapması 1 olan özel bir normal dağılım türüdür. Matematiksel olarak tüm normal dağılımlar bu forma dönüştürülebilir. Bu dönüşüme standartlaştırma denir. Standart normal dağılım ile normal dağılım aynı dağılım türünün farklı halleridir. Aralarındaki fark, temel olarak parametrelerindeki (ortalama ve standart sapma) farklılıktır.

Tanım 2.4.6. Ortalaması $\mu = 0$, varyansı $\sigma^2 = 1$ ve olasılık yoğunluk fonksiyonu

$$f(x) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}} - \infty < x < +\infty$$

olan normal dağılıma standart normal dağılım veya birim normal dağılım denir. Normal dağılımı $X \sim N(0,1)$ olarak gösterilir. $f(x)$ olasılık yoğunluk fonksiyonu olduğundan, $f(x)$ eğrisi altında ve x -ekseninin yukarısında kalan alan 1 e eşittir (Şekil 2.1.).



Şekil 2.1. Standart normal dağılım için örnek bir grafik

Cebirsel olarak

$$\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} \cdot e^{\frac{-x^2}{2}} \cdot dx = 1$$

yazılabilir (Akdeniz, 2003).

Teorem 2.4.5. $X \sim N(\mu, \sigma^2)$ normal dağılıma sahipse $Z = \frac{X-\mu}{\sigma}$ değişkeni, $N(0,1)$ standart normal dağılıma sahiptir (Akdeniz, 2003).

2.5. Log Olabilirlik Fonksiyonu

Olabilirlik yöntemi, bilinmeyen bir parametrenin tahminini yaparken, elde edilen örneklem verilerine en uygun olan değeri seçmeye dayanır. Bu yöntemin temel ilkesi, gözlenen verilerin gerçekleşme olasılığını (veya olasılık yoğunluğunu) en büyük yapan parametre değerinin, söz konusu bilinmeyen parametre için en iyi tahmin olduğudur.

En çok olabilirlik yöntemi, parametre tahmininde sıkça kullanılan güçlü bir tekniktir. Bu yöntemde, olabilirlik fonksiyonu olarak adlandırılan ve örnekleme ait olasılıkların bir fonksiyonu olan ifade maksimize edilir. Bu fonksiyonu maksimum yapan parametre değeri, en çok olabilirlik tahmin edicisi (MLE) olarak adlandırılır. Burada tezde kullanılacak olan log olabilirlik fonksiyonu tanıtılacaktır.

Tanım 2.5.1. $X_1, X_2, \dots, X_n$ olasılık veya olasılık yoğunluk fonksiyonu $f(x; \theta)$ olan kitleden bir örneklem olsun. θ nın olabilirlik fonksiyonu (likelihood function) $\tilde{X} = (X_1, X_2, \dots, X_n)$ olmak üzere,

$$L(\theta \mid \tilde{X} = \tilde{x}) = f(\tilde{x}; \theta) = \prod_{i=1}^n f(x_i; \theta)$$

dir (Işlak, 2024).

Olabilirlik fonksiyonunu maksimum yapan değer (örneklemenin bir fonksiyonu) θ nın en çok olabilirlik tahmin edicisidir. Yani, θ nın en çok olabilirlik tahmin edicisi

$$\hat{\theta}_n(\tilde{X}) = \max_{\theta \in \Theta} L(\theta \mid \tilde{X})$$

dir. Olabilirlik fonksiyonunun maksimizasyonu yerine genellikle fonksiyonun logaritması (log-likelihood, $\ell(\theta) = \log(L(\theta \mid \tilde{X} = \tilde{x}))$) maksimize edilir. θ nın en çok olabilirlik tahmin edicisi olabilirlik fonksiyonunu veya log-olabilirlik fonksiyonunu maksimum yapan değerdir. Yani,

$$\hat{\theta}_n(\tilde{X}) = \max_{\theta \in \Theta} \ell(\theta) = \max_{\theta \in \Theta} L(\theta \mid \tilde{X})$$

dir.

2.6. Öklid Uzayının Topolojisi ve Geometrisi

Bu alt bölümde Öklid uzayı üzerindeki bazı temel kavramlar verilecektir.

Tanım 2.6.1. V bir vektör uzayı olmak üzere

$$\begin{aligned} \langle, \rangle: V \times V &\rightarrow \mathbb{R} \\ (u, v) &\rightarrow \langle u, v \rangle \end{aligned}$$

şeklinde tanımlı $\langle, \rangle$ fonksiyonu

$$(i) \langle u, v \rangle = \langle v, u \rangle$$

$$(ii) \forall \lambda, \mu \in \mathbb{R} \text{ için } \langle \lambda u + \mu v, z \rangle = \lambda \langle u, z \rangle + \mu \langle v, z \rangle$$

$$(iii) v \neq 0 \text{ için } \langle v, v \rangle \geq 0, v = 0 \Leftrightarrow \langle v, v \rangle = 0$$

özelliklerini sağladığında $\langle, \rangle$ fonksiyonuna V üzerinde bir iç çarpım fonksiyonu, $(V, \langle, \rangle)$ ikilisine bir iç çarpım uzayı ve $\langle u, v \rangle$ sayısına da u ile v nin iç çarpımı denir (Turgut Vanlı, 2022).

Tanım 2.6.2. V bir vektör uzayı olmak üzere,

$$\begin{aligned} \|\cdot\|: V &\rightarrow \mathbb{R} \\ v &\rightarrow \|v\| \end{aligned}$$

fonksiyonu

- (i) $\forall v \in V$ için $\|v\| \geq 0$; $\|v\| = 0 \Leftrightarrow v = 0$
- (ii) $\forall a \in \mathbb{R}$ ve $\forall v \in V$ için $\|av\| = |a| \|v\|$
- (iii) $\forall u, v \in V$ için $\|u + v\| \leq \|u\| + \|v\|$ (üçgen eşitsizliği)

özelliklerini sağladığında, $\|\cdot\|$ fonksiyonuna norm, $\|v\|$ sayısına v vektörünün normu (boyu) $(V, \|\cdot\|)$ ikilisine de normlu vektör uzayı denir (Turgut Vanlı, 2022).

Tanım 2.6.3. V bir vektör uzayı olsun. Eğer $d: V \times V \rightarrow \mathbb{R}$ fonksiyonu

- (i) $d(x, y) \geq 0$
- (ii) $d(x, y) = 0 \Leftrightarrow x = y$
- (iii) $d(x, y) = d(y, x)$
- (iv) $d(x, z) \leq d(x, y) + d(y, z)$

özelliklerini sağlıyorsa d fonksiyonuna V kümesi üzerinde bir metrik, (V, d) ikilisine bir metrik uzay ve $d(x, y)$ sayısına da x ile y elemanları arasındaki uzaklık denir (Turgut Vanlı, 2022).

Tanım 2.6.4. X herhangi bir küme olsun. X 'in alt kümelerinin kümesi $\wp(X)$ olmak üzere $\tau \subseteq \wp(X)$ olsun. Yani τ elemanları X in bazı alt kümeleri olan bir küme olsun. Bu durumda

- (i) $\emptyset, X \in \tau$
- (ii) $U, V \in \tau$ için $U \cap V \in \tau$
- (iii) $\forall i \in I$ için $V_i \in \tau$ ise $\bigcup_{i \in I} V_i \in \tau$

özellikleri sağlanıyorsa τ ya X üzerinde bir topoloji (X, τ) ikilisine bir topolojik uzay denir. τ nun elemanlarına açık küme ve açık kümenin tümleyenine de kapalı küme denir (Turgut Vanlı, 2022).

Tanım 2.6.5. X ve Y birer topolojik uzay olmak üzere $f: X \rightarrow Y$ birebir ve örten bir fonksiyon olsun. Eğer f ve f in tersi f^{-1} sürekli ise f fonksiyonuna homeomorfizm denir. Eğer $f: X \rightarrow Y$ bir homeomorfizm ise X uzayı Y uzayına homeomorfikdir denir ve $X \approx Y$ ile gösterilir (Turgut Vanlı, 2022).

Tanım 2.6.6. X ve Y iki topolojik uzay olmak üzere $f: X \rightarrow Y$ bir fonksiyon olsun. X topolojik uzayından Y deki $f(X)$ alt uzayına bir homeomorfizm mevcut ise X uzayı Y uzayı içine gömülür ve f fonksiyonuna da gömme fonksiyonu denir (Turgut Vanlı, 2022).

Tanım 2.6.7. Bir topolojik uzayın farklı iki elemanının ayık komşulukları varsa bu uzaya T_2 - uzayı veya Hausdorff uzayı denir (Turgut Vanlı, 2022).

Teorem 2.6.1. X Hausdorff uzay ve Y bir topolojik uzay olmak üzere $f: X \rightarrow Y$ bir homeomorfizm ise Y Hausdorff uzaydır (Turgut Vanlı, 2022).

Teorem 2.6.2. Hausdorff uzayın kapalı her alt uzayı Hausdorff uzaydır (Turgut Vanlı, 2022).

Tanım 2.6.8. X topolojik uzayının bir alt kümesi A ve X in alt kümeler ailesi Q olsun (Turgut Vanlı, 2022).

- (i) Eğer Q daki kümelerin birleşimi A kümesini içeriyorsa Q koleksiyonuna A alt kümesini örtüyor denir.
- (ii) Q koleksiyonu A alt kümesini örtüyor ve Q koleksiyonundaki her bir küme açık ise Q koleksiyonuna A kümesinin açık örtüsü denir.
- (iii) Q koleksiyonu A kümesinin örtüsü olsun. Q' , A kümesini örtecek şekilde Q nun alt koleksiyonu ise Q' ye Q nun alt örtüsü denir.

Tanım 2.6.9. X Hausdorff uzayının her açık örtüsünün sonlu bir alt örtüsü varsa, X uzayına kompakt uzay denir (Turgut Vanlı, 2022).

Tanım 2.6.10. $f: \mathbb{R}^n \rightarrow \mathbb{R}$ bir fonksiyon ve $\mathbb{R}^n$ nin standart bazı $\{e_1, e_2, \dots, e_n\}$ olsun.

$$\begin{aligned} \nabla f: \mathbb{R}^n &\rightarrow \mathbb{R}^n \\ x &\rightarrow \nabla f(x) = \sum_{i=1}^n \frac{\partial f}{\partial x_i} e_i \end{aligned}$$

şeklinde tanımlı dönüşümüne f fonksiyonun gradiyenti denir (Turgut Vanlı, 2022).

Teorem 2.6.3 $f: \mathbb{R}^n \rightarrow \mathbb{R}$ bir fonksiyon ve $x, v \in \mathbb{R}^n$ olsun. f nin x noktasında v doğrultusundaki yöne göre türevi $D_v f(x) = \langle \nabla f(x), v \rangle$ dir (Turgut Vanlı, 2022).

Teorem 2.6.4 $f: \mathbb{R}^n \rightarrow \mathbb{R}, x \in \mathbb{R}^n$ ve $v = \sum_{i=1}^n v_i e_i$ olsun. f fonksiyonunun x noktasında v doğrultusunda yöne göre türevi

$$D_v f(x) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(x) v_i$$

dir (Turgut Vanlı, 2022).

Teorem 2.6.5. $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ bir fonksiyon ve $x, v \in \mathbb{R}^n$ olsun. f fonksiyonunun x noktasında v doğrultusunda yöne göre türevi

$$D_v f(x) = \sum_{k=1}^m \left(\sum_{i=1}^n \frac{\partial f_k}{\partial x_i}(x) v_i \right) e_k$$

dir (Turgut Vanlı, 2022).

Tanım 2.6.11. $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ bir fonksiyon ve $x_0 \in \mathbb{R}^n$ olsun. Eğer

$$\frac{\lim_{x \rightarrow x_0} \|f(x) - f(x_0) - L(x - x_0)\|}{\|x - x_0\|} = 0$$

olabilecek biçimde bir $L: \mathbb{R}^n \rightarrow \mathbb{R}^m$ lineer dönüşümü varsa f ye x_0 da diferensiyellenebilir, L ye de x_0 da f nin türevi denir ve $Df(x_0)$ ile gösterilir (Turgut Vanlı, 2022).

Tanım 2.6.12. $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ bir fonksiyon ve $f = (f_1, f_2, \dots, f_m)$ olsun. f_k ların $x_0 \in \mathbb{R}^n$ noktasındaki kısmi türevlerinden oluşan

$$\left[\frac{\partial f_k}{\partial x_i}(x_0) \right]_{m \times n} \quad (1 \leq k \leq m, 1 \leq i \leq n)$$

reel matrisine f nin x_0 noktasındaki Jakobiyen matrisi denir ve $J_f(x_0)$ şeklinde gösterilir. Yani

$$J_f(x_0) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(x_0) & \dots & \frac{\partial f_1}{\partial x_n}(x_0) \\ \frac{\partial f_2}{\partial x_1}(x_0) & \dots & \frac{\partial f_2}{\partial x_n}(x_0) \\ \dots & \dots & \dots \\ \frac{\partial f_m}{\partial x_1}(x_0) & \dots & \frac{\partial f_m}{\partial x_n}(x_0) \end{bmatrix}$$

dir (Turgut Vanlı, 2022).

Tanım 2.6.13. $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ ve $x_0 \in \mathbb{R}^n$ olsun. Herhangi bir $k \in \mathbb{Z}^+$ için, eğer varsa f nin x_0 da k . mertebeden türevi

$$D^k f(x_0) = D(D^{k-1}f)(x_0)$$

şeklinde tanımlanır. Eğer $D^k f(x_0)$ varsa f ye de x_0 da k -yinci mertebeden diferensiyellenebilir denir. $k \in \mathbb{Z}^+$ için eğer

$$\begin{array}{ccc} D^k f = D(D^{k-1}f): \mathbb{R}^n & \rightarrow & L^k(\mathbb{R}^n, \mathbb{R}^m) \\ x & \rightarrow & D^k f(x) \end{array}$$

varsa $D^k f$ ye f nin k -yinci mertebeden türevi ya da diferensiyeli denir. Eğer $D^k f$ varsa f ye de k -yinci mertebeden diferensiyellenebilir denir. Ayrıca $D_{j_k \dots j_1} f, (1 \leq j_p \leq n, 1 \leq p \leq k)$ bu tür kısmi türevlere f nin k -yi mertebeden kısmi türevleri denir.

$$D_{j_k \dots j_1} f = D_{j_k}(D_{j_{k-1} \dots j_1} f)$$

dir. $D^0 f(x_0) = f(x_0)$ kabul edilir (Turgut Vanlı, 2022).

Tanım 2.6.14. $f: \mathbb{R}^n \rightarrow \mathbb{R}^m, x_0 \in \mathbb{R}^n$ ve $k \in \mathbb{Z}^+$ olsun. Eğer f , her $k \in \mathbb{Z}^+$ için x_0 da C^k -diferensiyellenebilir ise f ye x_0 da C^∞ -diferensiyellenebilirdir veya düzgündür denir (Turgut Vanlı, 2022).

Tanım 2.6.15. $C^0(\mathbb{R}^n, \mathbb{R}^m) = \{f|f: \mathbb{R}^n \rightarrow \mathbb{R}^m, \mathbb{R}^m \text{ de sürekli} \}$
uzayına $\mathbb{R}^n$ den $\mathbb{R}^m$ e giden C^0 - sınıfından fonksiyonlar uzayı

$$C^1(\mathbb{R}^n, \mathbb{R}^m) = \{f|f: \mathbb{R}^n \rightarrow \mathbb{R}^m, \mathbb{R}^n \text{ de diferensiyellenebilir} \}$$

uzayına $\mathbb{R}^n$ den $\mathbb{R}^m$ e giden C^1 - sınıfından fonksiyonlar uzayı

$$C^k(\mathbb{R}^n, \mathbb{R}^m) = \{f|f: \mathbb{R}^n \rightarrow \mathbb{R}^m, k. \text{ mertebeden kısmi türevler var ve sürekli} \}$$

uzayına $\mathbb{R}^n$ den $\mathbb{R}^m$ e giden C^k -sınıfından fonksiyonlar uzayı

$$C^\infty(\mathbb{R}^n, \mathbb{R}^m) = \{f|f: \mathbb{R}^n \rightarrow \mathbb{R}^m \text{ her mertebeden kısmi türevler var ve sürekli} \},$$

uzayına $\mathbb{R}^n$ den $\mathbb{R}^m$ e giden C^∞ -sınıfından (düzgün) fonksiyonlar uzayı denir (Turgut Vanlı, 2022).

2.7. Manifoldlar

Manifold kavramı Riemann geometrisinin en temel ögesidir. Gauss'un yüzeyler üzerine elde ettiği bazı sonuçları Riemann daha genel kavram olan manifolda taşımıştır. Bu alt bölümde manifoldlara ait temel kavramlar verilecektir. Daha fazlası için (Turgut Vanlı, 2022)'ye bakılabilir.

Tanım 2.7.1. M bir topolojik uzay olsun. M nin her bir p noktasının $\mathbb{R}^n$ in bir V açık alt kümesine homeomorfik olan bir U açık komşuluğu varsa M ye bir topolojik manifold denir. (φ, U) ikilisine M topolojik uzayın bir n - boyutlu haritası, U ya koordinat komşuluğu denir. φ ye de koordinat dönüşümü (veya koordinat sistemi) denir. $\varphi(p) = 0$ ise (φ, U) haritasına $p \in U$ noktasında merkezleyendir denir (Turgut Vanlı, 2022).

Tanım 2.7.2. Bir topolojik uzay M ve M nin n -boyutlu bir haritası (φ, U) olsun. $\mathbb{R}^n$ uzayındaki doğal koordinat fonksiyonları $(x_1, x_2, \dots, x_n)$ ise $u_i = x_i \circ \varphi$ eşitliğiyle tanımlı $(u_1, u_2, \dots, u_n)$ fonksiyonlarına (φ, U) haritasının koordinat fonksiyonları $p \in M$ olmak üzere $(u_1(p), u_2(p), \dots, u_n(p))$ sayısına p noktasının (φ, U) haritasına göre koordinatları denir (Turgut Vanlı, 2022).

Tanım 2.7.3. Bir topolojik manifold M ve $(\varphi_\alpha, U_\alpha), (\varphi_\beta, U_\beta)$ M nin $U_\alpha \cap U_\beta \neq \emptyset$ olacak şekilde iki n -boyutlu haritası olsun. $\varphi_\alpha(U_\alpha \cap U_\beta)$ kümesi $\varphi_\alpha(U_\alpha)$ de $\varphi_\beta(U_\alpha \cap U_\beta)$ kümesi $\varphi_\beta(U_\beta)$ de açık ve

$$\varphi_{\beta\alpha}: \varphi_\beta \circ \varphi_\alpha^{-1}: \varphi_\alpha(U_\alpha \cap U_\beta) \rightarrow \varphi_\beta(U_\alpha \cap U_\beta)$$

dönüşümü bir homeomorfizm ise bu iki haritaya topolojik olarak örtüşürler denir ve $\varphi_{\beta\alpha}$ fonksiyonuna da haritalar arasında dönüşüm fonksiyonu denir (Turgut Vanlı, 2022).

Tanım 2.7.4. M bir topolojik uzay ve M nin haritalarının ailesi $\{(\varphi_\alpha, U_\alpha)\}_{\alpha \in I}$ olsun. Bu ailedeki haritalar topolojik örtüşüyor ise bu aileye M topolojik uzay üzerinde bir topolojik atlas denir (Turgut Vanlı, 2022).

Tanım 2.7.5. $(\varphi_\alpha, U_\alpha)$ ve (φ_β, U_β) , M topolojik manifoldun iki haritası olsun.

$$\varphi_{\alpha\beta} = \varphi_\alpha \circ \varphi_\beta^{-1}: \varphi_\beta(U_\alpha \cap U_\beta) \rightarrow \varphi_\alpha(U_\alpha \cap U_\beta),$$

$$\varphi_{\beta\alpha} = \varphi_\beta \circ \varphi_\alpha^{-1}: \varphi_\alpha(U_\alpha \cap U_\beta) \rightarrow \varphi_\beta(U_\alpha \cap U_\beta),$$

dönüşümleri C^∞ – sınıfından (yani düzgün) ise bu iki haritaya C^∞ – sınıfında veya düzgün örtüşürler denir (Turgut Vanlı, 2022).

Tanım 2.7.6. M topolojik manifoldun haritalarının ailesi $A = \{(\varphi_\alpha, U_\alpha)\}_{\alpha \in I}$ olsun. $M = \bigcup_{\alpha \in I} U_\alpha$ ve A içindeki her iki harita C^∞ - sınıfından (düzgün) örtüşüyor ise A ya M nin bir C^∞ – sınıfından (düzgün) atlası denir (Turgut Vanlı, 2022).

Tanım 2.7.7. Bir M topolojik manifold üzerinde iki C^∞ - atlas A_1 ve A_2 olsun. $A_1 \cup A_2$ de M üzerinde bir C^∞ – atlas ise bu iki atlas eşdeğer (denk) atlaslar (C^∞ – bağdaşıktır) denir ve $A_1 \sim A_2$ ile gösterilir (Turgut Vanlı, 2022).

Tanım 2.7.8. Denklik sınıfı $[A]$ ya M üzerinde A nın doğurduğu C^∞ - diferensiyellenebilir yapı (C^∞ - yapı) veya düzgün yapı denir (Turgut Vanlı, 2022).

Tanım 2.7.9. M topolojik uzay üzerinde tanımlı tüm düzgün atlasları içeren en geniş atlas tam (veya maksimal) atlas denir (Turgut Vanlı, 2022).

Tanım 2.7.10. M Hausdorff uzayı üzerinde tanımlı tam atlas ile birlikte topolojik manifoldda düzgün (veya C^∞) manifold denir. Buradaki tam atlas M üzerinde düzgün (C^∞ – sınıfından) yapı denir. Yani,

(i) M Hausdorff uzay

(ii) $\varphi_\alpha: U_\alpha \subset M \xrightarrow{\text{homeomorfizm}} V \subset \mathbb{R}^n$

(iii) $M = \bigcup_{\alpha \in I} U_\alpha$

(iv) $\varphi_{\beta\alpha}: \varphi_\alpha(U_\alpha \cap U_\beta) \rightarrow \varphi_\beta(U_\alpha \cap U_\beta)$ ve $\varphi_{\alpha\beta} = (\varphi_{\beta\alpha})^{-1}$ dönüşümü düzgün (C^∞ - sınıfından) yani $\varphi_{\alpha\beta}$ dönüşümü C^∞ - sınıfından bir difeomorfizmdir.

(v) $\varphi_{\alpha\beta} \in C^\infty$, örtendir. $\varphi_\beta(U_\alpha \cap U_\beta)$ nin koordinatlar $x = (x_1, x_2, \dots, x_n)$ ve $\varphi_\alpha(U_\alpha \cap U_\beta)$ nin koordinatları $y = (y_1, y_2, \dots, y_n)$ ise y_i fonksiyonların her biri x in C^∞ - sınıfından dönüşümleridir, ve

$$\det \left[\frac{\partial y_i}{\partial x_i} \right] \neq 0 \text{ yani } \det = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} & \dots & \frac{\partial y_1}{\partial x_n} \\ \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2} & \dots & \frac{\partial y_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial y_n}{\partial x_1} & \frac{\partial y_n}{\partial x_2} & \dots & \frac{\partial y_n}{\partial x_n} \end{bmatrix} \neq 0. \text{ (} \varphi_{\beta\alpha} \text{ da i ve ii şıklarını sağlar).}$$

(vi) $A = \{(U_\alpha, \varphi_\alpha)\}_{\alpha \in I}$ maksimal (tam) düzgün (C^∞ - sınıfından) atlasıdır.

Burada M nin koordinat komşulukları $\mathbb{R}^n$ de bir alt kümeye homeomorfik olduğundan M düzgün manifoldu n -boyuta sahiptir (Turgut Vanlı, 2022).

Tanım 2.7.11. n - boyutlu bir düzgün manifold $M, f: M \rightarrow \mathbb{R}$ herhangi bir fonksiyon ve $p \in M$ olsun. $f \circ \varphi^{-1}: \mathbb{R}^n \rightarrow \mathbb{R}$ fonksiyonu $\varphi(p)$ noktasında düzgün olacak şekilde M nin p noktasını içine alan bir (φ, U) haritası varsa, f fonksiyonuna p noktasında düzgündür denir. Burada $f \circ \varphi^{-1}$ fonksiyonuna f fonksiyonunun φ ye bağlı bir koordinat gösterimi denir (Turgut Vanlı, 2022).

Tanım 2.7.12. $f: M \rightarrow \mathbb{R}$ bir dönüşüm olsun. Eğer f, M nin her p noktasında düzgün ise f ye M üzerinde düzgün (C^∞ - sınıfından diferensiyellenebilir) denir. Bu fonksiyonların kümesi $\mathcal{F}(M)$ veya $C^\infty(M, \mathbb{R})$ ile gösterilir (Turgut Vanlı, 2022).

Tanım 2.7.13. M, N sırasıyla m, n boyutlu iki düzgün manifold ve $\Phi: M \rightarrow N$ bir dönüşüm olsun. Eğer $\psi \circ \Phi \circ \varphi^{-1}: \mathbb{R}^m \rightarrow \mathbb{R}^n$ dönüşümü $\varphi(p)$ noktasında düzgün olacak şekilde $p \in M$ noktasının bir (φ, U) haritası ve $\Phi(p) \in N$ noktasının da bir (ψ, V) haritası varsa Φ ye $p \in M$ noktasında düzgündür denir. $\forall p \in M$ noktasında Φ dönüşümü düzgün ise Φ ye M üzerinde düzgündür denir (Turgut Vanlı, 2022).

Tanım 2.7.14. M, N sırasıyla m ve n boyutlu iki düzgün manifold olmak üzere $\Phi: M \rightarrow N$ dönüşümü verilsin. Eğer,

- (i) Φ düzgün (C^∞ - sınıfından)
- (ii) Φ 1-1 ve örten
- (iii) Φ^{-1} düzgün

ise Φ ye bir C^∞ - sınıfından diffeomorfizm denir. M ve N ye de C^∞ - sınıfından diffeomorfiktir denir (Turgut Vanlı, 2022).

Tanım 2.7.15. M n - boyutlu bir manifold olmak üzere, $p \in M$ olsun. $\sigma: C^\infty(p) \rightarrow \mathbb{R}$ dönüşümünü göz önüne alınsın. Eğer $\forall f, g \in C^\infty(p)$ ve $\forall \lambda, \mu \in \mathbb{R}$ için

- (i) $\sigma(\lambda f + \mu g) = \lambda \sigma(f) + \mu \sigma(g)$ (Lineerlik özelliği)
- (ii) $\sigma(f \cdot g) = f(p) \cdot \sigma(g) + g(p) \cdot \sigma(f)$ (Leibnitz kuralı)

ise σ dönüşümüne p de M nin bir tanjant vektörü denir. M nin p noktasındaki tanjant vektörlerin kümesi $T_p(M)$ ile gösterilir.

$$\begin{aligned}\oplus: T_p(M) \times T_p(M) &\longrightarrow T_p(M) \\ (\sigma, \sigma') &\longrightarrow (\sigma \oplus \sigma')(f) = \sigma(f) + \sigma'(f) \\ \odot: \mathbb{R} \times T_p(M) &\longrightarrow T_p(M) \\ (\lambda, \sigma) &\longrightarrow (\lambda \odot \sigma)(f) = \lambda \sigma(f)\end{aligned}$$

olmak üzere $((T_p(M), \oplus), (\mathbb{R}, +, \cdot), \odot)$ altılısı bir vektör uzayıdır. Bu uzaya M nin p noktasındaki tanjant (teğet) uzayı denir (Turgut Vanlı, 2022).

Tanım 2.7.16. $f: M \rightarrow \mathbb{R}$ düzgün bir dönüşüm ve $p \in M$ için M nin bir haritası (φ, U) olsun.

$$\left. \frac{\partial f}{\partial u_i} \right|_p = \left. \frac{\partial (f \circ \varphi^{-1})}{\partial x_i} \right|_{\varphi(p)}$$

ifadesine f fonksiyonun M üzerinde i - yinci değişkene göre p noktasındaki kısmi türevi denir (Turgut Vanlı, 2022).

Teorem 2.7.1. M üzerinde,

$$\begin{aligned}\partial_i(p): C^\infty(p) &\longrightarrow \mathbb{R} \\ f &\longrightarrow \partial_i(f) = \left. \frac{\partial f}{\partial u_i} \right|_p\end{aligned}$$

şeklinde tanımlı dönüşüm M nin p noktasında bir tanjant (teğet) vektörüdür (Turgut Vanlı, 2022).

Teorem 2.7.2. $T_p(M) = sp\{\partial_1(p), \partial_2(p), \dots, \partial_n(p)\}$ ve $\text{boy } T_p(M) = \text{boy}(M)$ dir (Turgut Vanlı, 2022).

Tanım 2.7.17. n – boyutlu düzgün manifold M olsun. M nin her bir noktasına bir teğet vektör bağlayan dönüşüme M üzerinde bir vektör alanı denir. Yani

$$\begin{aligned}X: M &\rightarrow \bigcup_{p \in M} T_p(M) \\ p &\rightarrow X(p) = X_p\end{aligned}$$

şeklinde tanımlı X dönüşümü M üzerinde bir vektör alanıdır ve vektör alanlarının kümesi $\chi(M)$ ile gösterilir (Turgut Vanlı, 2022).

Tanım 2.7.18. $\forall p \in M$ ve $\lambda \in \mathbb{R}$ için

$$\oplus: \chi(M) \times \chi(M) \rightarrow \chi(M) \quad (X \oplus Y)(p) = X(p) + Y(p) = X_p + Y_p$$

$$\odot: \mathbb{R} \times \chi(M) \rightarrow \chi(M), \quad (\lambda X)(p) = \lambda(X(p)) = \lambda X_p$$

işlemleri ile $((\chi(M), \oplus), (\mathbb{R}, +, \cdot), \odot)$ altılısı bir vektör uzayıdır. Bu uzaya M manifoldunun vektör alanları uzayı denir (Turgut Vanlı, 2022).

Teorem 2.7.3. $\chi(M) = sp\{\partial_1, \partial_2, \dots, \partial_n\}$ dir (Turgut Vanlı, 2022).

Tanım 2.7.19. M manifoldu üzerinde herhangi bir vektör alanı $X = \sum_{i=1}^n X_i \frac{\partial}{\partial u_i}$ olmak üzere

$$\begin{aligned} X_i: M &\rightarrow \mathbb{R} \\ p &\rightarrow X_i(p) \end{aligned}$$

dönüşümüne X vektör alanının bileşeni denir. Eğer bu bileşen fonksiyonları düzgün ise X vektör alanı düzgündür denir.

Ayrıca

$$X(p) = \left(\sum_{i=1}^n X_i \partial_i \right) (p) = \sum_{i=1}^n X_i(p) \partial_i(p) = \sum_{i=1}^n X_i(p) \frac{\partial}{\partial u_i} \Big|_p = X_p$$

dir (Turgut Vanlı, 2022).

Tanım 2.7.20. Bir M manifoldu üzerinde,

$$\begin{aligned} X[f]: M &\rightarrow \mathbb{R} \\ p &\rightarrow X[f](p) = X_p(f) \end{aligned}$$

şeklinde tanımlı dönüşüme f fonksiyonun X vektör alanı yönündeki türevi denir (Turgut Vanlı, 2022).

Teorem 2.7.4. $\forall f \in \mathcal{F}(M)$ ve $\forall X \in \chi(M)$ için

$$x[f] = \sum_{i=1}^n X_i \frac{\partial f}{\partial u_i}$$

dir.

Tanım 2.7.21. M üzerinde tanımlı $[,]$ dönüşümü

$$\begin{aligned} [,]: \chi(M) \times \chi(M) &\rightarrow \chi(M) \\ (X, Y) &\rightarrow [X, Y] \end{aligned}$$

ile tanımlanmak üzere, $[,]$ dönüşümü

(i) Bilineer.

(ii) Anti-simetriktir yani $[X, Y] = -[Y, X]$

(iii) $[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0$ (Birinci Bianchi özdeşliği)

şartlarını sağlıyorsa $\chi(M)$ üzerinde Lie (parantez) çarpımı olarak adlandırılır $[X, Y]$ vektör alanına da X ile Y vektör alanlarının Lie çarpımı denir (Turgut Vanlı, 2022).

Teorem 2.7.5. $\forall f \in \mathcal{F}(M)$ ve $\forall X, Y \in \chi(M)$ için

$$[X, Y](f) = X[Y[f]] - Y[X[f]]$$

şeklinde tanımlı dönüşüm bir Lie çarpımıdır (Turgut Vanlı, 2022).

Teorem 2.7.6. $[\partial_i, \partial_j] = 0$ dır (Turgut Vanlı, 2022).

2.8. Manifoldların Riemann Geometrisi

Bir manifold üzerinde metriğin tanımlanması, manifold üzerinde ölçüm yapılmasını olanaklı kılar. Böyle manifoldlar Riemann manifoldları olarak anlandırılır. Bu bölümde Riemann manifoldları üzerinde metrik kullanılarak tanımlanan bazı temel kavramlar verilecektir.

2.8.1. Koneksiyon Kavramı ve Christoffel Sembolleri

Tanım 2.8.1. n - boyutlu düzgün bir manifold M olsun.

$$\begin{aligned} \nabla: \chi(M) \times \chi(M) &\rightarrow \chi(M) \\ (X, Y) &\rightarrow \nabla(X, Y) = \nabla_X Y \end{aligned}$$

dönüşümü $\forall \lambda, \mu \in \mathbb{R}, \forall f \in \mathcal{F}(M)$ ve $\forall X, Y, Z \in \chi(M)$ için,

- (i) $\nabla_{\lambda X + \mu Y} Z = \lambda \nabla_X Z + \mu \nabla_Y Z$
- (ii) $\nabla_X (\lambda Y + \mu Z) = \lambda \nabla_X Y + \mu \nabla_X Z$
- (iii) $\nabla_{fX} Y = f \nabla_X Y$
- (iv) $\nabla_X (fY) = f \nabla_X Y + X[f]Y$

şartlarını sağlanıyorsa bu dönüşüme M üzerinde bir koneksiyon (veya lineer koneksiyon) denir. $\nabla_X Y$ vektör alanına da Y vektör alanının X vektör alanı yönündeki kovaryant türevi denir (Turgut Vanlı, 2022).

(i). ve (iii). şartlarından ∇ nın birinci değişkene göre $\mathcal{F}(M)$ lineer olduğu, (iv) şartından ise ∇ nın ikinci değişken için $\mathcal{F}(M)$ lineerlik özelliğinin geçerli olmadığı görülmektedir.

Özel olarak $X = \partial_k, Y = \partial_l$ baz vektör alanlarını alınsın $\nabla_{\partial_k} \partial_l$ vektör alanının bileşenleri Γ_{kl}^m ile gösterilir. O zaman

$$\nabla_{\partial_k} \partial_l = \sum_{m=1}^n \Gamma_{kl}^m \partial_m$$

dir. Böylece $\nabla_{\partial_k} \partial_l$ nin bileşenleri

$$\Gamma_{kl}^m: U \subset M \rightarrow \mathbb{R}, m = 1, 2, \dots, n$$

olur. Bu fonksiyonlara Christoffel sembolleri denir (Turgut Vanlı, 2022).

Teorem 2.8.1. $\forall X, Y \in \chi(M)$ için $X = \sum_{k=1}^n X_k \partial_k$ ve $Y = \sum_{l=1}^n Y_l \partial_l$ olmak üzere Y vektör alanının X vektör alanı yönündeki kovaryant türevi

$$\nabla_X Y = \sum_{m=1}^n \left(\sum_{k=1}^n X_k \partial_k [Y_m] + \sum_{k,l=1}^n X_k Y_l \Gamma_{kl}^m \right) \partial_m$$

dir (Turgut Vanlı, 2022).

Tanım 2.8.2.

$$\begin{aligned} T: \chi(M) \times \chi(M) &\rightarrow \chi(M) \\ (X, Y) &\rightarrow T(X, Y) = \nabla_X Y - \nabla_Y X - [X, Y] \end{aligned}$$

şeklinde tanımlı dönüşüme ∇ nın torsiyonu (burulması) denir (Turgut Vanlı, 2022).

Tanım 2.8.3. M manifoldu üzerinde bir lineer koneksiyon ∇ olsun. Eğer $T = 0$ ise ∇ ya burulmasız (torsiyonsuz) veya simetriktir denir. Dalosayısıyla, $T = 0$ ise

$$[X, Y] = \nabla_X Y - \nabla_Y X$$

dir (Turgut Vanlı, 2022).

2.8.2. Riemann Metriği

Tanım 2.8.4. n – boyutlu düzgün bir manifold M olsun. $\forall p \in M$ için

$$\begin{aligned} g_p: T_p(M) \times T_p(M) &\rightarrow \mathbb{R} \\ (X_p, Y_p) &\rightarrow g_p(X_p, Y_p) \end{aligned}$$

dönüşümünü göz önüne alalım.

$$\begin{aligned} g: \chi(M) \times \chi(M) &\rightarrow \mathcal{F}(M) \\ (X, Y) &\rightarrow g(X, Y): M \rightarrow \mathbb{R} \\ p &\rightarrow g_p(X_p, Y_p) \end{aligned}$$

şeklinde tanımlı g dönüşümü

(i) Bilineer

(ii) Simetrik

(iii) Pozitif tanımlı yani, $\forall X \in \chi(M)$ için $g(X, X) > 0$ ve $g(X, X) = 0 \Leftrightarrow X = 0$

şartlarını sağlıyorsa g 'ye M üzerinde bir metrik tensör veya Riemann metrik tensörü ve (M, g) ikilisine de bir Riemann manifoldu denir (Turgut Vanlı, 2022).

Teorem 2.8.2. n - boyutlu (M, g) Riemann manifoldu üzerinde ∇ lineer koneksiyonunun bileşenleri

$$\Gamma_{kl}^r = \frac{1}{2} \sum_{h=1}^n g^{hr} \left\{ \frac{\partial g_{lh}}{\partial u_k} + \frac{\partial g_{kh}}{\partial u_l} - \frac{\partial g_{kl}}{\partial u_h} \right\}$$

dir. Burada $[g^{hr}] = [g_{hr}]^{-1}$ dir (Turgut Vanlı, 2022).

Tanım 2.8.5. (M, g) manifoldu üzerinde bir lineer koneksiyon ∇ olsun. $\forall X, Y, Z \in \chi(M)$ için

- i. $\nabla_X Y - \nabla_Y X = [X, Y]$ (Sıfır torsiyon özelliği veya simetri özelliği)
- ii. $X[g(Y, Z)] = g(\nabla_X Y, Z) + g(Y, \nabla_X Z)$ (Metrik ile bağdaşabilme özelliği)

şartları sağlanıyorsa ∇ ya M üzerinde (Riemann koneksiyonu veya) LeviCivita koneksiyonu denir (Turgut Vanlı, 2022).

$(\mathbb{R}^n, g)$ Riemann manifoldu üzerinde herhangi iki vektör alan $X = \sum_{k=1}^n X_k e_k$ ve $Y = \sum_{m=1}^n Y_m e_m$ olmak üzere, tüm Christoffel sembolleri sıfır olduğundan

$$\nabla_X Y = \sum_{m=1}^n \sum_{k=1}^n X_k \partial_k [Y_m] \partial_m$$

dir (Turgut Vanlı, 2022).

2.8.3. Riemann Eğrilik Tensörü

Tanım 2.8.6.

$$\begin{aligned} R: \chi(M) \times \chi(M) \times \chi(M) &\rightarrow \chi(M) \\ (X, Y, Z) &\rightarrow R(X, Y)Z \end{aligned}$$

olmak üzere $R(X, Y)Z = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X, Y]} Z$ dönüşümüne M nin (Riemann) eğrilik tensör alanı denir (Turgut Vanlı, 2022).

Teorem 2.8.3. $\forall f, g \in F(M)$ ve $\forall X, Y, Z, W \in \chi(M)$ için,

- (i) $R(fX + gY, Z)W = fR(X, Z)W + gR(Y, Z)W$
- (ii) $R(X, fY + gZ)W = fR(X, Y)W + gR(X, Z)W$
- (iii) $R(X, Y)(fZ + gW) = fR(X, Y)Z + gR(X, Y)W$

dir. Yani R eğrilik tensörü birinci, ikinci ve üçüncü bileşene göre $\mathcal{F}(M)$ lineerdir (Turgut Vanlı, 2022).

Teorem 2.8.4. Bir n - boyutlu (M, g) Riemann manifoldunun eğriliğinin bileşenleri

$$R_{jkl}^i = \frac{\partial \Gamma_{lj}^i}{\partial u_k} - \frac{\partial \Gamma_{kj}^i}{\partial u_l} + \sum_{m=1}^n (\Gamma_{km}^i \Gamma_{lj}^m - \Gamma_{kj}^m \Gamma_{lm}^i)$$

dir (Turgut Vanlı, 2022).

Teorem 2.8.5. Bir n - boyutlu (M, g) Riemann manifoldunun eğriliği R olmak üzere

$$R_{jkl}^m + R_{klj}^m + R_{ljk}^m = 0$$

dir. Bu eşitliğe 1. Bianchi eşitliği denir (Turgut Vanlı, 2022).

Teorem 2.8.6. Bir n - boyutlu (M, g) Riemann manifoldunun eğriliği R olsun. Bu durumda

$$(i) R_{jmk l} = -R_{jml k}$$

$$(ii) R_{k j m l} = R_{m l k j}$$

$$(iii) R_{j m k l} = -R_{m j k l}$$

dir (Turgut Vanlı, 2022).

Tanım 2.8.7. $X_p, Y_p \in T_p(M)$ ve $\pi = Sp\{X_p, Y_p\} \subset T_p(M)$ olsun.

$$K(X_p, Y_p) = \frac{g_p(R(X_p, Y_p)Y_p, X_p)}{g_p(X_p, X_p)g_p(Y_p, Y_p) - g_p(X_p, Y_p)^2}$$

şeklinde tanımlı K ya π düzleminin kesitsel eğriliği denir (Turgut Vanlı, 2022).

Teorem 2.8.7. (M, g) Riemann manifoldunun eğriliği R ve kesitsel eğriliği K olsun. Bu durumda

$$K = 0 \Leftrightarrow R = 0$$

dir (Turgut Vanlı, 2022).

Tanım 2.8.8. (M, g) Riemann manifoldunun eğrilik tensörü $R = 0$ ise manifoldda düzdür (flat) denir.

2.8.4. Geodezik Kavramı

M bir Riemann manifoldu olsun. Eğer $\alpha: I \rightarrow M$ eğrisinin teğet vektör alanı α boyunca paralel ise eğriye geodeziktir denir. Öklidyen uzayda doğrular birer geodeziktir. Keyfi bir manifoldda geodezik kavramı öklidyen uzayda doğruya karşılık gelen geometrik nesneler olarak düşünülebilir. Örneğin kürede iki noktayı birbirine bağlayan en kısa uzunluğa sahip eğri, yani kürenin geodeziği, büyük çemberlerdir. Bu eğriler kürenin merkezinden geçen düzlemler ile kürenin kesişimi olan eğrilerdir (Şahin, 2012).

Teorem 2.8.8. M bir Riemann manifoldu ve $\alpha: I \rightarrow M$ eğrisinin denklemleri $\xi^i = \xi^i(s)$ olsun. Bu durumda α eğrisinin bir geodezik olması için gerek ve yeter şart

$$\frac{d^2 \xi^i}{ds^2} + \sum_{j,k=1}^n \Gamma_{jk}^i \frac{d\xi^j}{ds} \frac{d\xi^k}{ds} = 0, i = 1, \dots, n$$

olmasıdır. Teoremde verilen diferansiyel denkleme geodezik denklemi denir.

Tanım 2.8.9. α eğrisinin p ve q noktaları arasında bir geodezik boyunca uzaklık ise

$$S(p, q) = \int_p^q \sqrt{\sum_{ij} g_{ij}(\xi) d\xi^i d\xi^j}$$

Şeklindedir (Kobayashi ve Nomizu ,1963)



3. OLASILIK MODELLERİ

Bir olasılık modeli geometrik bir yapı olarak verilebilen, olasılık dağılımlarının ailesini içeren bir yapıdır. Ailenin her bir elemanı aslında bir olasılık dağılımıdır. Olasılık model ise bu aile üzerine inşa edilir. Dağılımın parametreleri özelleştirilerek, ailenin tek bir elemanı elde edilebilmektedir. Olasılık dağılımlarının ailesi düzgün (smooth) parametreler ile tanımlanırsa, model çok boyutlu bir yüzey olarak değerlendirilebilir. Böylece, bir geometrik nesne ortaya çıkmış olur. Bu geometrik nesne bir yüzey ve genel olarak bir manifold biçimindedir. Manifold üzerinde diferansiyel geometri araçları tanımlanarak, modelin geometrisi incelenebilmektedir. Bu bölümde olasılık modellerinin tanımı, özellikleri ve geometrisi ele alınacaktır.

3.1. İstatistiksel Model

Tanım 3.1.1. X bir küme olsun. X in bir $\mathcal{A}$ sınıfı için aşağıdaki özellikler sağlanırsa bu $\mathcal{A}$ sınıfına X üzerinde bir cebirdir denir (Balcı, 2012):

- (i) $X \in \mathcal{A}$
- (ii) Her $E \in \mathcal{A}$ için $E^c = X \setminus E \in \mathcal{A}$
- (iii) $k = 1, 2, \dots, n$ için $E_k \in \mathcal{A} \Rightarrow \bigcup_{k=1}^n E_k \in \mathcal{A}$

Eğer her $n \in \mathbb{N}$ için $E_n \in \mathcal{A}$ iken $\bigcup_{n=1}^{\infty} E_n \in \mathcal{A}$ ise $\mathcal{A}$ cebirine bir σ – cebiri adı verilir.

Eğer $\mathcal{A}, X$ üzerinde bir σ cebiri ise, yukarıdaki özelliklerden,

- (i) $\emptyset \in \mathcal{A}$
- (ii) $k = 1, 2, \dots, n$ için $E_k \in \mathcal{A} \Rightarrow \bigcap_{k=1}^n E_k \in \mathcal{A}$
- (iii) $k \in \mathbb{N}$ için $E_k \in \mathcal{A} \Rightarrow \bigcap_{k=1}^{\infty} E_k \in \mathcal{A}$
- (iv) Her $A, B \in \mathcal{A}$ için $A \setminus B \in \mathcal{A}$

önergeleri geçerlidir (Balcı, 2012).

Tanım 3.1.2. Bir X kümesi, X in alt kümelerinin bir $\mathcal{A}$ σ -cebiri ve $\mathcal{A}$ üzerinde tanımlı bir μ ölçüsünden oluşan $(X, \mathcal{A}, \mu)$ üçlüsüne bir ölçü uzayı adı verilir (Balcı, 2012).

$\mathcal{S}$, sonlu ya da sonsuz bir küme, $\mathcal{F}$, $\mathcal{S}$ 'nin alt kümelerinin bir ailesi ve $(\mathcal{S}, \mathcal{F})$ bir ölçülebilir uzay olsun. $\mathcal{F}$ 'nin her bir elemanına negatif olmayan bir sayı atanabilir. Bu sayı her bir kümenin boyutunu, ölçüsünü (size) verir.

Tanım 3.1.3. $P: \mathcal{F} \rightarrow \mathbb{R} \cup \{\pm\infty\}$ bir ölçü olmak üzere $(\mathcal{S}, \mathcal{F}, P)$ olasılık uzayı verilsin. $\mathcal{F}$ nin elemanlarına olay, $\mathcal{S}$ kümesine de örnek uzay denir (Balcı, 2012).

Reel değerli bir X rassal değişkeni $X: \mathcal{S} \rightarrow \mathbb{R}$ olasılık uzayında ölçülebilir bir fonksiyondur. Yani, herhangi iki $a, b \in \mathbb{R} \cup \{\pm\infty\}$ için $\{s \in \mathcal{S}; a < X(s) < b\} \in \mathcal{F}$ dir.

X kesikli bir rassal değişken ve $X: \mathcal{S} \rightarrow \chi = \{x^1, x^2, \dots\}$ olmak üzere, $p: \chi \rightarrow [0,1]$ olasılık fonksiyonu

$$p_k = p(x^k) = P(X = x^k) = P(\{s \in \mathcal{S}; X(s) = x^k\}) \text{ ve } \sum_k p(x^k) = 1$$

ile tanımlanır. X sürekli bir rassal değişken ve $X: \mathcal{S} \rightarrow \chi \subset \mathbb{R}^n$, $p: \chi \rightarrow \mathbb{R}$ olasılık yoğunluk fonksiyonu ise,

$$p(x) \geq 0, (\forall x \in \chi) \text{ ve } \int_{\chi} p(x) dx = 1$$

şeklinde ifade edilir (Amari ve Nagaoka, 2000)

$\mathcal{S}, X$ kümesi üzerinde, her bir elemanı bir olasılık dağılımı olan olasılık dağılımlarının bir ailesi olsun $\mathcal{S}$. reel değerli $\xi = \{\xi^1, \xi^2, \dots, \xi^n\}$ değişkenleri tarafından parametrelendirilebilir.

$\mathcal{S}$ olasılık dağılımlarının ailesi

$$\mathcal{S} = \{p_{\xi} = p(x; \xi) | \xi = \{\xi^1, \xi^2, \dots, \xi^n\} \in \mathbb{E}\}$$

ile gösterilir. Burada $\mathbb{E}, \mathbb{R}^n$ nin alt kümesi ve $\xi \rightarrow p_{\xi}$ dönüşümü 1-1 dir.

Tanım 3.1.4. $\mathcal{S} = \{p_{\xi} = p(x; \xi) | \xi = \{\xi^1, \xi^2, \dots, \xi^n\} \in \mathbb{E}\}$ olmak üzere $\mathcal{S}$ ye n -boyutlu istatistiksel model veya parametrik model denir ve $\mathcal{S} = \{p_{\xi}\}$ veya $\mathcal{S} = \{p(x; \xi)\}$ şeklinde gösterilir (Amari ve Nagaoka, 2000).

Örnek 3.1.1. Bir parametrelili üstel dağılımın olasılık yoğunluk fonksiyonu $p(x; \xi) = \xi \cdot e^{-\xi x}$ biçimindedir. Burada ξ dağılımın parametresidir. Bu dağılım için $X = \mathbb{R}$ ve $\mathbb{E} = \{\xi | \xi > 0\}$ biçimindedir (Amari ve Nagaoka 2000).

Örnek 3.1.2. İki parametrelili normal dağılımın olasılık yoğunluk fonksiyonu

$$p(x; \xi) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

ile verilir. Bu dağılım için $X = \mathbb{R}, \xi = (\mu, \sigma)$ olup parametre uzayı

$$\mathbb{E} = \{(\mu, \sigma) | 0 < \mu < \infty, 0 < \sigma < \infty\}$$

şeklindedir (Amari ve Nagaoka, 2000). Dağılımın parametrelerinin tanımından hareketle $\mathbb{E}$ uzayının, analitik düzlemin birinci bölgesi olduğu görülmektedir.

3.2. İstatistiksel Modelin Geometrik Temsili

Bir istatistiksel model, parametrelendirilmiş bir yüzey olarak temsil edilebilir. Bu yüzeyin her noktası bir olasılık yoğunluğunu temsil eder. $\mathbb{E} \subset \mathbb{R}^n$ alt kümesine parametre uzayı denir. $\mathcal{S}$ kümesi

$$\mathcal{P}(\mathcal{X}) = \left\{ f \mid f: \mathcal{X} \rightarrow \mathbb{R}, f \geq 0, \int_{\mathcal{X}} f dx = 1 \right\}$$

biçiminde tanımlanan sonsuz boyutlu fonksiyon uzayının bir alt kümesidir.

Geometrik temsil için yüzey tanımına ihtiyaç duyulmaktadır.

$$\iota: \mathbb{E} \rightarrow \mathcal{P}(\mathcal{X}), \quad \iota(\xi) = p_{\xi}$$

dönüşümü bir immersiyon olsun (Şekil 3.1). Yani

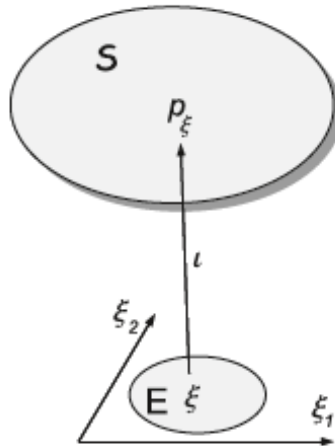
- I. ι bire-birdir
- II. $\text{rank} n = \dim \mathbb{E}$ dir.

Bu durumda $\frac{\partial p_{\xi}}{\partial \xi^1}, \dots, \frac{\partial p_{\xi}}{\partial \xi^n}$ fonksiyonları lineer bağımsızdır. Bu koşul, $\mathcal{X} \subset \mathbb{R}$ özel durumu için, Wronskian'ın tek bir noktada sıfır olmaması anlamına gelir. Yani $\forall \xi$ için

$$\begin{vmatrix} \varphi_1(x_0, \xi) & \dots & \varphi_n(x_0, \xi) \\ \varphi'_1(x_0, \xi) & \dots & \varphi'_n(x_0, \xi) \\ \dots & \dots & \dots \\ \varphi_1^{(n-1)}(x_0, \xi) & \dots & \varphi_n^{(n-1)}(x_0, \xi) \end{vmatrix} \neq 0$$

dir. Burada $x_0 \in \mathcal{X}$ için, $\varphi_j(x, \xi) = \frac{\partial p_{\xi}(x)}{\partial \xi^j}$ ve $\varphi_j^{(k)}(x_0, \xi) = \partial_x^k \varphi_j(x, \xi)|_{x=x_0}$ dir.

Dolayısıyla, $\{\varphi_j\}$ tüm ξ 'ler için x in bağımsız fonksiyonlardır.



Şekil 3.1. İmmersiyon ι her bir ξ parametresini bir p_{ξ} dağılımına eşler.

$\mathcal{S}$ kümesinin yukarıda belirtilen (I)-(II) özelliklerini sağlaması ile istatistiksel model bir geometrik yapıya kavuşmuştur.

$\mathcal{S} = \{p_\xi\}$ model yüzeyi için temel $\varphi_j(x, \xi) = \partial_j p_\xi(x)$ fonksiyonları vektör alanları rolünü oynar. $\mathcal{S}$ 'nin $\mathcal{X}$ üzerindeki gerçek değerli düzgün fonksiyonlar uzayı $\mathcal{F}(\mathcal{X}, \mathbb{R})$ ile gösterilir.

φ_j vektör alanları $f: \mathcal{S} \rightarrow \mathcal{F}(\mathcal{X}, \mathbb{R})$ fonksiyonu için $\varphi_j(f) = \frac{\partial(f(p_\xi))}{\partial \xi^j}$ biçimindedir. Sonuç olarak, $\varphi_j(fg) = f\varphi_j(g) + \varphi_j(f)g$ Leibniz kuralı sağlanır.

$\mathcal{S} = \{p_\xi\}$ model yüzeyi için **log-olabilirlik fonksiyonu** önemli bir rol oynar. Log-olabilirlik fonksiyonu $\ell: \mathcal{S} \rightarrow \mathcal{F}(\mathcal{X}, \mathbb{R})$, $\ell(p_\xi)(x) = \ln p_\xi(x)$ biçiminde tanımlanır. Bazen, kolaylık Ayrıca,

$$\partial_j \ell_x(\xi) = \frac{\partial \ln p_\xi(x)}{\partial \xi^j} = \varphi_j(\ell_x(\xi)), \quad 1 \leq j \leq n$$

biçimindedir. Log-olabilirlik fonksiyonuna literatürde skor fonksiyonları olarak da anılır.

Log-olabilirlik fonksiyonu, sezgisel olarak, p_ξ 'de bulunan bilginin ξ^j yönünde nasıl değiştiğini açıklar.

$\mathcal{X}$ örnek uzayı sınırlıysa türev ve integral yer değiştirebilir. Böylece

$$\int_{\mathcal{X}} \varphi_j(x) dx = \partial_j \int_{\mathcal{X}} p_\xi(x) dx = \partial_j 1 = 0, \quad 1 \leq j \leq n \quad (3.1)$$

olarak elde edilir. Diğer yandan ayırık dağılımlar söz konusu olduğunda

$$\sum_{k \geq 1} \varphi_j(x_k) = \partial_j \sum_{k \geq 1} p_\xi(x_k) = \partial_j 1 = 0, \quad 1 \leq j \leq n \quad (3.2)$$

biçimindedir. Eğer $\mathcal{X}$ sonlu ise türev ve toplam yer değiştirebilir. Aksi takdirde, serinin ξ içinde düzgün yakınsak olduğu varsayılmalıdır.

Teorem 3.2.1. $\mathcal{S} = \{p_\xi\}$ istatistiksel modelinin düzgünlük koşulunu (II) sağlaması için

gerek ve yeter koşul $\mathbb{E}$ içindeki herhangi bir ξ için $\{\partial_1 \ell_x(\xi), \partial_2 \ell_x(\xi), \dots, \partial_n \ell_x(\xi)\}$

kümesinin $n - \text{lineer}$ bağımsız olmasıdır.

Önerme 3.2.1. (I) ve (II) koşullarının sağlandığı kabul edilsin. Eğer $E_\xi[\cdot]$, p_ξ 'ye göre beklentiyi gösteriyorsa, o zaman

$$E_\xi[\partial_j \ell_x(\xi)] = 0, \quad 1 \leq j \leq n$$

dır.

İspat. Doğrudan hesaplama ile

$$\begin{aligned}
E_{\xi}[\partial_j \ell_x(\xi)] &= E_{\xi} \left[\frac{\partial_j p(x; \xi)}{p(x; \xi)} \right] = \int_{\mathcal{X}} \partial_j p(x; \xi) dx \\
&= \partial_j \left(\int_{\mathcal{X}} p(x; \xi) dx \right) = \partial_j(1) = 0
\end{aligned}$$

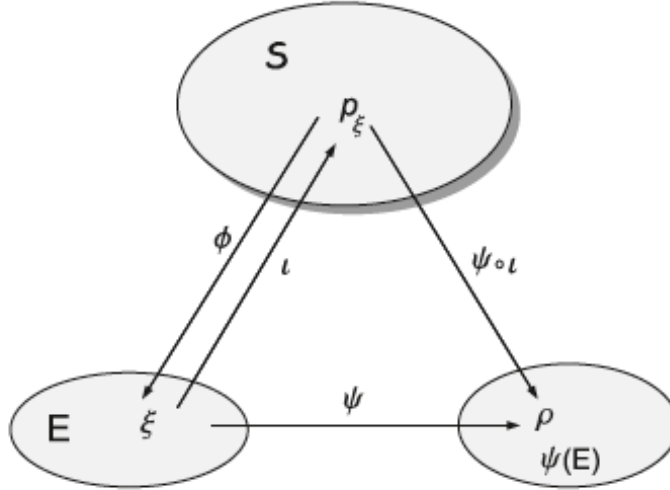
olarak elde edilir. Benzer şekilde, ayrık durumda da

$$\begin{aligned}
E_{\xi}[\partial_j \ell_x(\xi)] &= \sum_{k \geq 1} p_{\xi}(x_k) \partial_j \ln p_{\xi}(x_k) \\
&= \sum_{k \geq 1} \partial_j p_{\xi}(x_k) = \partial_j \sum_{k \geq 1} p_{\xi}(x_k) = 0
\end{aligned}$$

olarak bulunur.

3.3. İstatistiksel Modelin Parametrelendirilmesi

İstatistiksel model $\mathcal{S} = \{p_{\xi}\}$ bire-bir eşlemenin görüntüsü olduğundan $\iota: \mathbb{E} \rightarrow \mathcal{P}(\mathcal{X}), \iota(\xi) = p_{\xi}$, ters fonksiyonu $\phi: \mathcal{S} \rightarrow \mathbb{E} \subset \mathbb{R}^n, \phi(p_{\xi}) = \xi$ biçimindedir. ϕ eşlemesi her p_{ξ} dağılımına bir ξ parametresi atadığından, bu istatistiksel model için bir koordinat sistemi olarak ele alınabilir. Bir istatistiksel model birden çok parametrisasyon ile temsil edilebilir (Şekil 3.2).



Şekil 3.2. $\mathcal{S}$ istatistiksel modelinin iki parametrelendirmesi

$\mathcal{S}$ üzerinde koordinat sistemi şu şekilde değiştirilebilir: $\psi: \mathbb{E} \rightarrow \psi(\mathbb{E}) \subset \mathbb{R}^n$ bir diffeomorfizm olsun. O zaman $\psi \circ \iota: \mathcal{S} \rightarrow \psi(\mathbb{E}) \subset \mathbb{R}^n$ başka bir koordinat sistemidir ve bu yeni koordinat sistemi $\rho = \psi(\xi)$ olarak gösterilirse, istatistiksel model

$\mathcal{S} = \{p_{\psi^{-1}(\rho)}; \rho \in \psi(\mathbb{E})\}$ olarak da yazılabilir. Modelin parametrelendirmesi değişse bile geometrik özellikleri bundan etkilenmez. Bir parametrelendirmede kanıtlanan geometrik sonuçlar tüm parametrelendirmeler için geçerli olduğundan, çalışmak için uygun bir parametrelendirme seçmek yararlıdır.

3.4. Bazı Olasılık Dağılımlarının Geometrik Modelleri

Aşağıda geometrik bir anlam kazanmış istatistiksel model örnekleri verilmiştir.

3.4.1. Üstel Dağılımın Geometrik Modeli

$\mathcal{X} = [0, \infty)$ ve $\mathbb{E} = (0, \infty)$ olmak üzere parametresi ξ olan üstel dağılım $p(x; \xi) = \xi e^{-\xi x}$ ile verilsin ve $\iota: \mathbb{E} \rightarrow \mathcal{P}([0, \infty))$, $\iota(\xi) = p(x; \xi)$ olarak tanımlansın.

Birebirlik: Bu eşlemenin 1-1 olduğu açıktır.

Regülerlik : Her $x \in [0, \infty)$ için $\partial_{\xi} p_{\xi}(x) = (1 - \xi x)e^{-\xi x} \neq 0$ dir. Dönüşüm regülerdir.

Sonuç olarak $\mathcal{S} = \{p(x; \xi)\}$ istatistiksel modeli $\mathcal{P}([0, \infty))$ içinde bir **eğri** olarak düşünülebilir. Ayrıca, $\psi: (0, \infty) \rightarrow (0, \infty)$, $\psi(\xi) = \frac{1}{\xi}$ dönüşümü bir diffeomorfizm olup yeni bir parametrelendirme verir. Yeni parametizasyona göre $p(x; \rho) = \frac{1}{\rho} e^{-\frac{x}{\rho}}$ olarak yazılabilir (Calin ve Udrişte, 2014).

3.4.2. Normal Dağılımın Geometrik Modeli

$\mathcal{X} = \mathbb{R}$ ve $\mathbb{E} = \mathbb{R} \times (0, \infty)$ olmak üzere normal dağılımın olasılık yoğunluk fonksiyonu $p(x; \xi^1, \xi^2) = p(x; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$, $x \in \mathcal{X}$, $(\xi^1, \xi^2) \in \mathbb{E}$ ile verilsin.

Birebirlik: $\iota: \mathbb{E} \rightarrow \mathcal{S}$, $\iota(\xi) = p(x; \mu, \sigma)$ eşlemesinin bire bir olduğunu göstermek için $(\bar{\mu}, \bar{\sigma}) \in \mathbb{E}$ olmak üzere

$$p(x; \mu, \sigma) = p(x, \bar{\mu}, \bar{\sigma})$$

olduğu kabul edilsin. Buradan log-olabilirlik fonksiyonu kullanıldığında

$$\ln p(x; \mu, \sigma) = \ln p(x, \bar{\mu}, \bar{\sigma})$$

olduğu görülür. Polinomların eşitliği ilkesinden

$$\frac{(x - \mu)^2}{2\sigma^2} + \ln \sigma = \frac{(x - \bar{\mu})^2}{2\bar{\sigma}^2} + \ln \bar{\sigma}$$

olarak edilir ki bu da $\mu = \bar{\mu}$ ve $\sigma = \bar{\sigma}$ sonucuna varır. O halde dönüşüm bire birdir.

Regülerlik: Normal dağılım için log-olabilirlik fonksiyonu

$$\ell_x(\mu, \sigma) = \ln p(x; \mu, \sigma) = -\frac{(x - \mu)^2}{2\sigma^2} - \ln \sigma - \ln \sqrt{2\pi},$$

biçimindedir. Buradan kısmi türevler alınır,

$$\begin{aligned}\partial_\mu \ell_x(\mu, \sigma) &= \frac{x - \mu}{\sigma^2}, \\ \partial_\sigma \ell_x(\mu, \sigma) &= \frac{(x - \mu)^2}{\sigma^3} - \frac{1}{\sigma}\end{aligned}$$

ifadeleri elde edilir. $y = x - \mu$ alındığında $\frac{y}{\sigma^2}$ ve $y^2\sigma^3 - \frac{1}{\sigma}$ polinomlarının y 'nin lineer bağımsız fonksiyonları olduğu kolaylıkla görülebilir. O halde dönüşüm regülerdir.

Sonuç olarak, (I) ve (II) koşulları sağlandığından normal dağılımın $\mathcal{S} = \{p(x; \xi)\}$ istatistiksel modeli iki boyutlu bir **yüzeye** karşılık gelir (Şekil 3.3) (Calin ve Udrişte, 2014).

Ayrıca

$$\psi: \mathbb{R} \times (0, \infty) \rightarrow \mathbb{R} \times (0, \infty), \psi(\mu, \sigma) = \left(\mu + \sigma, \frac{1}{\sigma}\right)$$

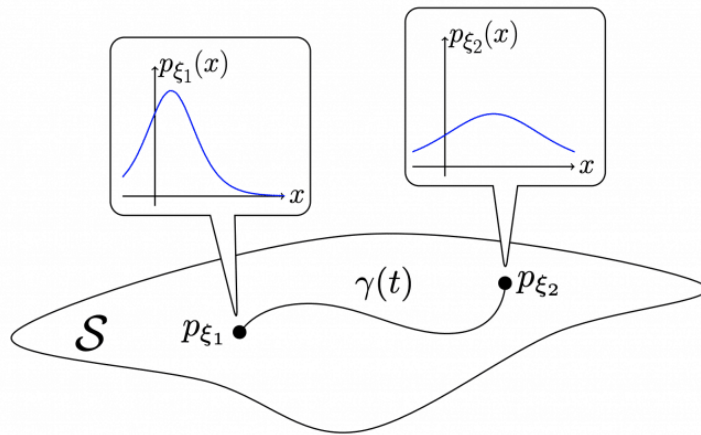
dönüşümü bir diffeomorfizmdir. Bu dönüşümün tersi,

$$\psi^{-1}(\rho, \eta) = \left(\rho - \frac{1}{\eta}, \frac{1}{\eta}\right)$$

olup buradan normal dağılımın

$$p(x; \rho; \eta) = \frac{\eta}{\sqrt{2\pi}} e^{-\frac{1}{2}(x\eta - \rho\eta + 1)^2}, \quad (\rho, \eta) \in \mathbb{E}$$

parametrelendirilmesine ulaşılır. Farklı parametrizasyonlar olsa da $p(x; \mu, \sigma)$ ile $p(x; \rho; \eta)$ geometrik olarak aynıdır.



Şekil 3.3. Manifolddaki iki nokta arasında Gauss dağılımını parametrize eden bir $\gamma(t)$ eğrisi

(Miyamoto, 2024).

3.4.3. Gamma Dağılımının Geometrik Modeli

Örnek uzayı $\mathcal{X} = [0, \infty)$ ve parametre uzayı $\mathbb{E} = (0, \infty) \times (0, \infty)$ ile verilen Gamma dağılımının olasılık yoğunluk fonksiyonu

$$p_{\xi}(x) = p_{\alpha, \beta}(x) = \frac{1}{\beta^{\alpha} \Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\beta}}$$

ile verilir, burada

$$\Gamma(\alpha) = \int_0^{\infty} t^{\alpha-1} e^{-t} dt, \alpha > 0$$

ile tanımlanan Gamma fonksiyonudur.

Gamma dağılım ailesi $(\xi^1, \xi^2) = (\alpha, \beta) \in \mathbb{E}$ parametreleri ile verilir.

Birebirlik: $p(x; \alpha, \beta) = p(x; \bar{\alpha}, \bar{\beta})$ olsun. O zaman $\ln(p(x; \alpha, \beta)) = \ln(p(x; \bar{\alpha}, \bar{\beta}))$ dir.

Böylece

$$-\alpha \ln \beta - \ln \Gamma(\alpha) + (\alpha - 1) \cdot \ln x - x = -\bar{\alpha} \ln \bar{\beta} - \ln \Gamma(\bar{\alpha}) + (\bar{\alpha} - 1) \cdot \ln x - \frac{x}{\bar{\beta}}$$

olarak elde edilir. Buradan terim terime eşitleme ile

$$-\alpha \ln \beta - \ln \Gamma(\alpha) = -\bar{\alpha} \ln \bar{\beta} - \ln \Gamma(\bar{\alpha})$$

ifadesi bulunur ve $\alpha = \bar{\alpha}, \beta = \bar{\beta}$ olur. Yani, fonksiyon birebirdir.

Regülerlik: Olasılık yoğunluk fonksiyonu $p_{(\alpha, \beta)} = p(x; \alpha, \beta) = \frac{1}{\beta^{\alpha} \Gamma(\alpha)} x^{\alpha-1} \cdot e^{-\frac{x}{\beta}}$

biçiminde yazılabilir. ψ fonksiyonu

$$\psi: (0, \infty) \times (0, \infty) \rightarrow (0, \infty) \times (0, \infty)$$

$$(\alpha, \beta) \rightarrow \psi(\alpha, \beta) = \left(\alpha + \beta, \frac{1}{\beta} \right)$$

ile tanımlansın. Bu durumda, $\psi^{-1}(p, r) = \left(p - \frac{1}{r}, \frac{1}{r} \right)$ olur. Böylece, $\alpha = p - \frac{1}{r}, \beta = \frac{1}{r}$

olarak elde edilir. Bu ifadeler dikkate alınır,

$$\begin{aligned} p(x; p \cdot \eta) &= \frac{1}{\left(\frac{1}{\eta}\right)^{\frac{p-1}{\eta}} \Gamma\left(p - \frac{1}{\eta}\right)} \cdot x^{p - \frac{1}{\eta} - 1} \cdot e^{-\frac{x}{\eta}} \\ &= \frac{\eta^{p - \frac{1}{\eta}}}{\Gamma\left(p - \frac{1}{\eta}\right)} \cdot x^{\frac{p\eta - 1 - \eta}{\eta}} \cdot e^{-\eta x} \end{aligned}$$

olur. Buradan,

$$\begin{aligned}
\ln(p(x; \alpha, \beta)) &= \ln\left(\frac{1}{\beta^\alpha \Gamma(\alpha)} \cdot x^{\alpha-1} \cdot e^{-\frac{x}{\beta}}\right) \\
&= \ln\left(\frac{1}{\beta^\alpha \Gamma(\alpha)}\right) + \ln(x^{\alpha-1}) + \ln\left(e^{-\frac{x}{\beta}}\right) \\
&= -\ln\left(\beta^\alpha \Gamma(\alpha) + (\alpha-1) \ln x - \frac{x}{\beta} \ln e\right) \\
&= -\ln\left(\beta^\alpha \Gamma(\alpha) + (\alpha-1) \ln x - \frac{x}{\beta}\right) \\
&= -(\ln(\beta^\alpha) + \ln \Gamma(\alpha)) + (\alpha-1) \ln x - \frac{x}{\beta}
\end{aligned}$$

elde edilir böylece log-olabilirlik fonksiyonu

$$l_x(\alpha, \beta) = -\alpha \ln \beta - \ln \Gamma(\alpha) + (\alpha-1) \ln x - \frac{x}{\beta}$$

dir. Bu fonksiyonun parametrelere göre türevleri alınır,

$$\begin{aligned}
\partial_\alpha l_x &= -\ln \beta - \partial_\alpha \Gamma(\alpha) + \ln x \\
\partial_\beta l_x &= -\alpha \cdot \frac{1}{\beta} + \frac{x}{\beta^2}
\end{aligned}$$

bulunur. Bunlar lineer bağımsızdır. O halde regülerlik koşulu gerçekleşir.

Sonuç olarak, gamma dağılım ailesi $(\xi^1, \xi^2) = (\alpha, \beta) \in \mathbb{E}$ parametreleri ile $\mathcal{P}([0, \infty))$ içinde iki boyutlu bir **yüzey** oluşturur.

Teorem 3.2.2. Üstel dağılım gamma dağılımı üzerinde bir eğridir.

İspat. Gamma dağılımında $\alpha = 1$ alınsın. $\Gamma(1) = 1$ olduğundan

$$p_{1,\beta}(x) = \frac{1}{1\beta^1\Gamma(1)} x^0 \cdot e^{-\frac{x}{\beta}} = \frac{1}{\beta} \cdot e^{-\frac{x}{\beta}}$$

olur. $\frac{1}{\beta} = \xi$ alınırsa $p_{1,\beta}(x) = \xi \cdot e^{-\xi x}$ bulunur. Yani üstel dağılım elde edilir. Sonuç olarak, üstel dağılım 2 boyutlu gamma dağılım yüzeyi üzerinde 1 boyutlu bir eğridir.

3.4.4. Beta Dağılımının Geometrik Modeli

Örnek uzayı $\mathcal{X} = [0,1]$ parametre uzayı $\mathbb{E} = (0, \infty) \times (0, \infty)$ olan beta dağılımının olasılık yoğunluk fonksiyonu

$$p_{a,b}(x) = \frac{1}{B(a,b)} x^{a-1} (1-x)^{b-1}$$

ile verilir. Burada,

$$B(a, b) = \int_0^1 t^{a-1} (1-t)^{b-1} dt, \quad a, b > 0$$

ile tanımlanan beta fonksiyonudur ve $(a, b) \in \mathbb{E} = (0, \infty) \times (0, \infty)$ dir. Beta dağılımı $\mathcal{P}([0,1])$ içinde iki boyutlu bir yüzey oluşturur.

3.4.5. Kesikli Dağılımın Geometrik Modeli

Örnek uzay $\mathcal{X} = \{x^1, x^2, \dots, x^n\}$ sonlu bir küme olsun.

$$\mathbb{E} = \left\{ (\xi^1, \dots, \xi^{n-1}); \xi^i > 0, \sum_{i=1}^{n-1} \xi^i < 1 \right\}$$

parametre uzayı için olasılık yoğunluk fonksiyonu

$$p(x^i; \xi) = \begin{cases} \xi^i, & \text{if } 1 \leq i \leq n-1 \\ 1 - \sum_{i=1}^{n-1} \xi^i, & \text{if } i = n. \end{cases}$$

ile tanımlanırsa kesikli bir olasılık dağılımı elde edilir. $\mathbb{R}^{n-1}$ 'in doğal tabanı $\{e_1, \dots, e_{n-1}\}$ olmak üzere $\partial_j p_\xi = (e_j, -1)^T$ olarak elde edilir. Bu da $\{\partial_j p_\xi\}, j \in \{1, \dots, n-1\}$ kümesinin lineer bağımsız olduğunu gösterir. Böylece regülerlik koşulu sağlanır. ι 'nın bire bir olduğu kolaylıkla görülebilir (Calin ve Udriște, 2014).

3.4.6. Possion Dağılımın Geometrik Modeli

Örnek uzayı $\mathcal{X} = \{0, 1, 2, \dots\}$ ve parametre uzayı $\mathbb{E} = (0, \infty)$ olan Possion dağılımının olasılık yoğunluk fonksiyonu

$$p(x; \lambda) = \frac{\lambda^x}{x!} e^{-\lambda}$$

biçimindedir. Possion dağılımının istatistiksel modeli bir boyutludur. Skor fonksiyonu tek değişkenli olup $\partial_\lambda \ell_x(\lambda) = \frac{x}{\lambda} - 1$ dir. Yani regülerlik sağlanır. Bire bir olma koşulu da kolaylıkla görülebilir.

3.4.7. İki değişkenli Normal Dağılımın Geometrik Modeli

İki değişkenli normal dağılım

$$p(x_1, x_2; \xi) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-\frac{Q(x)}{2}}$$

olasılık yoğunluk fonksiyonu ile verilir, burada

$$Q(x) = \frac{1}{1-\rho^2} \left[\frac{(x_1 - \mu_1)^2}{\sigma_1^2} - 2\rho \frac{(x_1 - \mu_1)(x_2 - \mu_2)}{\sigma_1\sigma_2} + \frac{(x_2 - \mu_2)^2}{\sigma_2^2} \right]$$

biçiminde tanımlanır. Dağılımın örnek uzayı $\mathcal{X} = \mathbb{R}^2$ dir.

Dağılımın parametreleri aşağıdaki koşulları sağlar;

- (i) μ_1 ve μ_2 her reel sayı olabilir
- (ii) σ_1 ve σ_2 pozitifdir.
- (iii) $\rho, [0,1]$ aralığındadır.

Böylece parametre uzayı $\mathbb{E} = \mathbb{R}^2 \times (0, \infty) \times (0, \infty) \times [0,1]$ dir. O halde, $p(x_1, x_2, \xi)$ 5-boyutlu istatistiksel bir modeldir.

İki değişkenli Gauss dağılımının log-olabilirlik fonksiyonu

$$\begin{aligned} \ln[p(x_1, x_2, \xi)] &= \ln \left(\frac{1}{2\pi \cdot \sigma_1 \sigma_2 \sqrt{1-\rho^2}} \cdot e^{-\frac{\theta(x)}{2}} \right) \\ &= \ln \left(\frac{1}{2\pi \sigma_1 \sigma_2 \sqrt{1-\rho^2}} \right) + \ln \left(e^{-\frac{\theta(x)}{2}} \right) \\ &= -\ln(2\pi \cdot \sigma_1 \sigma_2 \sqrt{1-\rho^2}) - \frac{\theta(x)}{2} \cdot \ln e \\ &= -\left[\ln 2\pi + \ln \sigma_1 + \ln \sigma_2 + \ln \sqrt{1-\rho^2} \right] - \frac{\theta(x)}{2} \\ &= -\ln 2\pi - \ln \sigma_1 - \ln \sigma_2 + \frac{1}{2} \ln(1-\rho^2) - \frac{\theta(x)}{2} \end{aligned}$$

biçimindedir. Yani

$$l_x(\xi) = -\ln 2\pi - \ln \sigma_1 - \ln \sigma_2 + \frac{1}{2} \ln(1-\rho^2) - \frac{\theta(x)}{2}$$

dir. Kısmi türevler

μ_1 için

$$\begin{aligned} \frac{\partial Q(x)}{\partial \mu_1} &= \frac{2(x_1 - \mu_1)}{(1-\rho^2)\sigma_1^2} - \frac{2\rho(x_2 - \mu_2)}{(1-\rho^2)\sigma_1\sigma_2} \\ \frac{\partial \theta(x)}{\partial \mu_1} &= -\frac{1}{2} \frac{\partial Q(x)}{\partial \mu_1} \\ \frac{\partial \theta(x)}{\partial \mu_1} &= -\frac{1}{2} \left[\frac{2(x_1 - \mu_1)}{(1-\rho^2)\sigma_1^2} - \frac{2\rho(x_2 - \mu_2)}{(1-\rho^2)\sigma_1\sigma_2} \right]. \end{aligned}$$

μ_2 için

$$\begin{aligned}\frac{\partial Q(x)}{\partial \mu_2} &= -\frac{2\rho(x_1 - \mu_1)}{(1 - \rho^2)\sigma_1\sigma_2} + \frac{2(x_2 - \mu_2)}{(1 - \rho^2)\sigma_2^2} \\ \frac{\partial \theta(x)}{\partial \mu_2} &= -\frac{1}{2} \frac{\partial Q(x)}{\partial \mu_2} \\ \frac{\partial \theta(x)}{\partial \mu_2} &= -\frac{1}{2} \left[-\frac{2\rho(x_1 - \mu_1)}{(1 - \rho^2)\sigma_1\sigma_2} + \frac{2(x_2 - \mu_2)}{(1 - \rho^2)\sigma_2^2} \right].\end{aligned}$$

σ_1 için

$$\begin{aligned}\frac{\partial(-\ln \sigma_1)}{\partial \sigma_1} &= -\frac{1}{\sigma_1} \\ \frac{\partial Q(x)}{\partial \sigma_1} &= -\frac{2(x_1 - \mu_1)^2}{(1 - \rho^2)\sigma_1^3} + \frac{2\rho(x_1 - \mu_1)(x_2 - \mu_2)}{(1 - \rho^2)\sigma_1^2\sigma_2} \\ \frac{\partial \theta(x)}{\partial \sigma_1} &= -\frac{1}{\sigma_1} - \frac{1}{2} \frac{\partial Q(x)}{\partial \sigma_1} \\ \frac{\partial \theta(x)}{\partial \sigma_1} &= -\frac{1}{\sigma_1} - \frac{1}{2} \left[-\frac{2(x_1 - \mu_1)^2}{(1 - \rho^2)\sigma_1^3} + \frac{2\rho(x_1 - \mu_1)(x_2 - \mu_2)}{(1 - \rho^2)\sigma_1^2\sigma_2} \right].\end{aligned}$$

σ_2 için

$$\begin{aligned}\frac{\partial(-\ln \sigma_2)}{\partial \sigma_2} &= -\frac{1}{\sigma_2}, \\ \frac{\partial Q(x)}{\partial \sigma_2} &= -\frac{2(x_2 - \mu_2)^2}{(1 - \rho^2)\sigma_2^3} + \frac{2\rho(x_1 - \mu_1)(x_2 - \mu_2)}{(1 - \rho^2)\sigma_1\sigma_2^2}, \\ \frac{\partial \theta(x)}{\partial \sigma_2} &= -\frac{1}{\sigma_2} - \frac{1}{2} \frac{\partial Q(x)}{\partial \sigma_2}, \\ \frac{\partial \theta(x)}{\partial \sigma_2} &= -\frac{1}{\sigma_2} - \frac{1}{2} \left[-\frac{2(x_2 - \mu_2)^2}{(1 - \rho^2)\sigma_2^3} + \frac{2\rho(x_1 - \mu_1)(x_2 - \mu_2)}{(1 - \rho^2)\sigma_1\sigma_2^2} \right].\end{aligned}$$

ρ için

$$\begin{aligned}\frac{\partial \ln(1 - \rho^2)}{\partial \rho} &= -\frac{2\rho}{1 - \rho^2} \\ \frac{\partial Q(x)}{\partial \rho} &= -\frac{2\rho}{(1 - \rho^2)^2} \left[\frac{(x_1 - \mu_1)^2}{\sigma_1^2} - 2\rho \frac{(x_1 - \mu_1)(x_2 - \mu_2)}{\sigma_1\sigma_2} + \frac{(x_2 - \mu_2)^2}{\sigma_2^2} \right] + \frac{2(x_1 - \mu_1)(x_2 - \mu_2)}{(1 - \rho^2)\sigma_1\sigma_2}, \\ \frac{\partial \theta(x)}{\partial \rho} &= -\frac{\rho}{1 - \rho^2} - \frac{1}{2} \frac{\partial Q(x)}{\partial \rho}.\end{aligned}$$

biçiminde olur. Bu türevlerin lineer bağımsız olduğu görülebilir.

3.4.8. Çok Değişkenli Normal Dağılımın Geometrik Modeli

Çok değişkenli normal dağılımın olasılık yoğunluk fonksiyonu

$$p(x; \xi) = \frac{1}{(2\pi)^{\frac{k}{2}} (\det A)^{\frac{1}{2}}} e^{-\frac{1}{2}(x-\mu)^t A^{-1}(x-\mu)}$$

ile verilir. A simetrik bir matris olduğundan köşegen üzerinde k tane benzersiz eleman vardır. Köşegen üst tarafında $1 + 2 + 3 + \dots + (k - 1)$ tane eleman olur. Bu durumda köşegen hariç eleman sayısı;

$$1 + 2 + 3 + \dots + (k - 1) = \frac{(k - 1) \cdot k}{2} = \frac{k^2}{2} - \frac{k}{2}$$

olup toplam eleman sayısı $k + \frac{k^2}{2} - \frac{k}{2} - \frac{k}{2} + \frac{k^2}{2} = \frac{1}{2}k(k + 1)$ dir. O zaman $\mathbb{E}$ deki eleman sayısı k tane μ den, $\frac{1}{2}k(k + 1)$ tane de A dan gelir ve $n = k + \frac{1}{2} \cdot k(k + 1)$ olur. Yani $\text{boy}(\mathbb{E}) = n = k + \frac{1}{2}k(k + 1)$ dir.

Böylece örnek uzay $\mathcal{X} = \mathbb{R}^k$ olup parametre uzayı

$$\mathbb{E} = \{(\mu, A); \mu \in \mathbb{R}^k, A \in \mathbb{R}^{k \times k} \text{ pozitif tanımlı}\}$$

ile verilir.

Bu modelin log-olabilirlik fonksiyonu

$$\ell_x(\mu, A) = \ln p(x; \xi) = -\frac{1}{2}(x - \mu)^t A^{-1}(x - \mu) - \frac{1}{2} \ln(\det A) - \frac{k}{2} \ln(2\pi)$$

biçimindedir. Burada $\mu = (\mu_1, \dots, \mu_k)$, $A = (A_{ij})$, $A^{-1} = (A^{ij})$ ve $u_j = x_j - \mu_j$ ile gösterilir. Buradan

$$\begin{aligned} \partial_{\mu_r} \ell_x(\mu, A) &= -\sum_{j=1}^k \frac{A^{rj} + A^{jr}}{2} (x_j - \mu_j) \\ &= -\sum_{j=1}^k \frac{A^{rj} + A^{jr}}{2} u_j, \\ \partial_{A^{\alpha\beta}} \ell_x(\mu, A) &= \frac{1}{2} A_{\alpha\beta} - \frac{1}{2} (x_\alpha - \mu_\alpha)(x_\beta - \mu_\beta) \\ &= \frac{1}{2} A_{\alpha\beta} - \frac{1}{2} u_\alpha u_\beta. \end{aligned}$$

eşitlikleri elde edilir. $\partial_{\mu_r} \ell$ ve $\partial_{A^{\alpha\beta}} \ell$, u_j cinsinden doğrusal olarak bağımsız polinomlardır. Yani regülerlik koşulu sağlanır (Calin ve Udrişte, 2014).

3.4.9. İki İstatistiksel Modelin Çarpımı

İki istatistiksel modelin kartezyen çarpımı yeni bir istatistiksel model verir.

$\mathcal{S} = \{p_\xi: \mathcal{X} \rightarrow \mathbb{R}; \xi \in \mathbb{E}\}$ ve $\mathcal{U} = \{q_\theta: \mathcal{Y} \rightarrow \mathbb{R}; \theta \in \mathbb{F}\}$ iki istatistiksel model olsun ve

$$\mathcal{S} \times \mathcal{U} = \{f_{\xi, \theta}: \mathcal{X} \times \mathcal{Y} \rightarrow \mathbb{R}; (\xi, \theta) \in \mathbb{E} \times \mathbb{F}\}$$

biçiminde tanımlansın, burada $f_{\xi,\theta}(x,y) = p_{\xi}(x)q_{\theta}(y)$ dir. Tanımlanan fonksiyon için $f_{\xi,\theta}(x,y) \geq 0$ dir ve

$$\begin{aligned}\iint_{\mathcal{X} \times \mathcal{Y}} f_{\xi,\theta}(x,y) dx dy &= \iint_{\mathcal{X} \times \mathcal{Y}} p_{\xi}(x) q_{\theta}(y) dx dy \\ &= \int_{\mathcal{X}} p_{\xi}(x) dx \int_{\mathcal{Y}} q_{\theta}(y) dy = 1\end{aligned}$$

olup olasılık dağılım fonksiyonu olma koşulunu sağlar. O halde $f_{\xi,\theta}$, $\mathcal{X} \times \mathcal{Y}$ üzerinde bir olasılık yoğunluk fonksiyonudur. Ayrıca modelin log-olabilirlik fonksiyonu

$$\ell_{(x,y)}(\xi, \theta) = \ln f_{\xi,\theta}(x,y) = \ln p_{\xi}(x) + \ln q_{\theta}(y) = \ell_x(\xi) + \ell_y(\theta)$$

biçiminde elde edilir. Diğer yandan

$$\begin{aligned}\partial_{\xi^j} \ell_{(x,y)}(\xi, \theta) &= \partial_{\xi^j} \ell_x(\xi) \\ \partial_{\theta^i} \ell_{(x,y)}(\xi, \theta) &= \partial_{\theta^i} \ell_y(\theta)\end{aligned}$$

dir. $\{\partial_{\xi^j} \ell_x(\xi)\}, \{\partial_{\theta^i} \ell_y(\theta)\}$ fonksiyon sistemleri lineer bağımsız olduğundan, $\{\partial_{\xi^j} \ell_{(x,y)}(\xi, \theta), \partial_{\theta^i} \ell_{(x,y)}(\xi, \theta)\}$, fonksiyon sistemi $\mathcal{X} \times \mathcal{Y}$ üzerinde lineer bağımsızdır. Dolayısıyla $\mathcal{S} \times \mathcal{U}$, $\mathcal{S}$ ve $\mathcal{U}$ çarpım istatistiksel manifoldu olarak adlandırılan istatistiksel bir model haline gelir.

Eğer p_{ξ} ve q_{θ} sırasıyla $\mathcal{X}$ ve $\mathcal{Y}$ örnek uzaylarına sahip iki bağımsız rastgele değişken X ve Y 'nin olasılık yoğunlukları ise, o zaman bunların ortak olasılık yoğunluğu $f_{\xi,\theta}(x,y)$, $\mathcal{S} \times \mathcal{U}$ çarpım istatistiksel manifolduna aittir (Calin ve Udrişte, 2014).

3.5. Üstel Aile

Üstel aile (exponential family), birçok yaygın olasılık dağılımını kapsayan geniş bir sınıftır. İstatistiksel modelleme ve makine öğrenimi gibi alanlarda büyük öneme sahip olan bu sınıf, dağılımların geometrik yapısını anlamak ve analiz etmek için güçlü bir araçtır. Üstel aileler, olasılık yoğunluk fonksiyonlarının belirli bir formda ifade edilmesiyle tanımlanır. Genel olarak, bir üstel ailede olasılık yoğunluğu, parametre vektörü ξ ye bağlı olarak şu şekilde tanımlanır:

$$p(x; \xi) = e^{C(x) + \xi^i F_i(x) - \psi(\xi)}$$

burada, $C(x)$, x 'e bağlı bir fonksiyon, ξ^i doğal parametreler ve $F_i(x)$ yeterli istatistik fonksiyonlarını temsil eder. $\psi(\xi)$ log-partition fonksiyonu olarak adlandırılır ve normalizasyonu sağlar ve $x \in \mathcal{X}$, $\psi(\xi) < \infty$ olmak üzere

$$\psi(\xi) = \ln \left(\int_{\mathcal{X}} e^{C(x) + \xi^T F_i(x)} dx \right)$$

biçiminde tanımlanır.

Üstel aileler, Fisher bilgisi ve doğal metriklerin hesaplanabilirliği açısından Riemann geometrisi ile olan bağlantısı nedeniyle dikkat çeker. Özellikle, olasılık dağılımlarının geometrik yapısını anlamak için üstel aileler güçlü bir matematiksel araç sunar. Bu dağılımlar, log-partition fonksiyonu üzerinden doğal bir geometri sağlar.

Üstel aile (exponential family) olasılık dağılımları, geniş bir dağılım sınıfını kapsar ve oldukça genel bir yapı sunar. Bu genellik, üstel ailenin farklı olasılık dağılımlarını tek bir çerçeve altında ifade edebilme yeteneğinden kaynaklanır.

3.5.1. Üstel Aileye Ait Bazı Özel Dağılımlar

Bu bölümde üstel aileye ait bazı dağılımlar incelenecektir.

1. (Üstel Dağılım) $C(x) = 0, F_1(x) = -x, \xi^1 = \xi, \psi(\xi) = -\ln \xi$ alınırsa $p(x; \xi) = \xi \cdot e^{-\xi x}$ olasılık yoğunluk fonksiyonu ile tanımlanan bir parametreleri üstel dağılım üstel aileye ait olur.

2. (İki Parametrelili Normal Dağılım) Normal dağılımın olasılık yoğunluk fonksiyonu

$$p(x; \xi) = \frac{1}{\sqrt{2\pi}\sigma} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

biçiminde olup

$$p(x; \xi) = e^{\frac{\mu}{\sigma^2}x - \frac{1}{2\sigma^2}x^2 + \frac{\mu^2}{2\sigma^2} + \ln \frac{1}{\sqrt{2\pi}\sigma}}$$

formunda yazılabilir. Buradan, $(\xi^1, \xi^2) \in \mathbb{E} = \mathbb{R} \times (-\infty, 0)$ olmak üzere

$$C(x) = 0, F_1(x) = x, F_2(x) = x^2, \xi^1 = \frac{\mu}{\sigma^2}, \xi^2 = \frac{-1}{2\sigma^2}, \psi(\xi) = \frac{-(\xi^1)^2}{4\xi^2} + \frac{1}{2} \ln \left(\frac{-\pi}{\xi^2} \right)$$

olarak alınırsa, iki parametrelili normal dağılımın üstel dağılım ailesine ait olduğu görülür.

3. (Bernoulli Dağılımı) Bernoulli dağılımının olasılık yoğunluk fonksiyonu,

$$p(x; p) = p^x (1-p)^{1-x}, x \in \{0,1\}, 0 < p < 1.$$

biçimindedir. Buradan

$$p(x; p) = e^{(x \ln(p) + (1-x) \ln(1-p))}$$

şeklinde yazılabilir. İfade açılırsa,

$$p(x; p) = e^{(x \ln(p) + \ln(1-p) - x \ln(1-p))},$$

olur ve

$$p(x; p) = e^{(x \ln(\frac{p}{1-p}) + \ln(1-p))}$$

elde edilir. Bernoulli dağılımı

$$\xi = \ln\left(\frac{p}{1-p}\right), F(x) = x, C(x) = 0 \text{ ve } \psi(\xi) = -\ln(1-p)$$

olacak biçimde üstel aileye ait bir dağılımdır.

4. (Poisson Dağılımı) Poisson dağılımının olasılık yoğunluk fonksiyonu,

$$p(x, \lambda) = \frac{\lambda^x}{x!} \cdot e^{-\lambda} = e^{\ln\left(\frac{\lambda^x}{x!}\right) - \lambda}$$

biçiminde yazılabilir. Diğer yandan

$$\begin{aligned} \ln\left(\frac{\lambda^x}{x!}\right) - \lambda &= \ln \lambda^x - \ln(x!) - \lambda \\ &= x \cdot \ln \lambda - \ln(x!) - \lambda \\ &= -\ln(x!) + \ln \lambda x - \lambda \end{aligned}$$

olur. Sonuç olarak,

$$C(x) = -\ln(x!), \xi^1 = \ln \lambda, F_1(x) = x, \psi(\xi) = \lambda$$

seçilirse Poisson dağılımı üstel aileye ait olur.

5. (Gamma Dağılımı) Gamma dağılımının olasılık yoğunluk fonksiyonu,

$$P(x; \alpha, \beta) = \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-\frac{1}{\beta}x} = e^{\ln\left(\frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1}\right) - \frac{1}{\beta}x}$$

biçiminde yazılabilir. Buradan,

$$\begin{aligned} \ln\left(\frac{1}{\beta^\alpha \Gamma(\alpha)} \cdot x^{\alpha-1}\right) - \frac{1}{\beta}x &= \ln\left(\frac{1}{\beta^\alpha \cdot \Gamma(\alpha)}\right) + \ln(x^{\alpha-1}) - \frac{1}{\beta}x \\ &= -\ln(\beta^\alpha \cdot \Gamma(\alpha)) + (\alpha - 1) \ln x - \frac{1}{\beta}x \\ &= -\left(\ln \beta^\alpha + \ln \Gamma(\alpha) + (\alpha - 1) \cdot \ln x - \frac{1}{\beta}x\right) \\ &= -\alpha \ln \beta - \ln \Gamma(\alpha) + (\alpha - 1) \ln x - \frac{1}{\beta}x \\ &= -\alpha \ln \beta - \ln \Gamma(\alpha) + \alpha \ln x - \ln x - \frac{1}{\beta}x \\ &= -\ln x + \alpha \ln x - \frac{1}{\beta}(x) - \alpha \ln \beta - \ln \Gamma(\alpha) \end{aligned}$$

olacağından

$$C(x) = -\ln x, \xi^1 = \alpha, \xi^2 = -\frac{1}{\beta}, F_1(x) = \ln x, F_2(x) = x \text{ ve}$$

$\psi(\xi) = -\alpha \ln \beta - \ln \Gamma(\alpha)$ eşitlikleri ile gamma dağılımı üstel aileye ait olur.

6. (Beta Dağılımı) Beta dağılımının olasılık yoğunluk fonksiyonu,

$$\begin{aligned} P(x; a, b) &= \frac{1}{\beta(a, b)} \cdot x^{a-1} \cdot (1-x)^{b-1} \\ &= e^{\ln\left(\frac{1}{\beta(a, b)} x^{a-1} \cdot (1-x)^{b-1}\right)} \end{aligned}$$

biçiminde yazılabilir. Buradan,

$$\begin{aligned} P(x; a, b) &= \ln\left(\frac{1}{\beta(a, b)} \cdot x^{a-1} \cdot (1-x)^{b-1}\right) \\ &= \ln\left(\frac{1}{\beta(a, b)}\right) + \ln x^{a-1} + \ln(1-x)^{b-1} \\ &= -\ln(\beta(a, b)) + (a-1) \ln x + (b-1) \cdot \ln(1-x) \\ &= -\ln(B(a, b)) + a \ln x - \ln x + b \cdot \ln(1-x) - \ln(1-x) \\ &= -\ln x - \ln(1-x) + a \ln x + b \ln(1-x) - \ln(B(a, b)) \end{aligned}$$

olarak yazılabilir. Böylece, $C(x) = -(\ln x + \ln(1-x) = -\ln(x \cdot (1-x)))$

$$\xi^1 = a, \quad \xi^2 = b, \quad F_1(x) = \ln x, \quad F_2(x) = \ln(1-x), \quad \psi(\xi) = \ln(B(a, b))$$

olup beta dağılımı üstel aileye aittir.

7. (Erlang Dağılımı) Erlang dağılımının olasılık yoğunluk fonksiyonu, $x \geq 0, \lambda \geq 0$ için

$$p(x; k, \lambda) = \frac{\lambda^k x^{k-1} e^{-\lambda x}}{(k-1)!} = e^{\ln\left(\frac{\lambda^k \cdot x^{k-1} \cdot e^{-\lambda x}}{(k-1)!}\right)}$$

olarak yazılabilir. Buradan,

$$\begin{aligned} p(x; k, \lambda) &= \ln\left(\frac{\lambda^k \cdot x^{k-1} \cdot e^{-\lambda x}}{(k-1)!}\right) = \ln(\lambda^k \cdot x^{k-1} \cdot e^{-\lambda x}) - \ln((k-1)!) \\ &= \ln(\lambda^k) + \ln(x^{k-1}) + \ln(e^{-\lambda x}) - \ln(k-1) \\ &= k \cdot \ln \lambda + (k-1) \ln x - \lambda x - \ln(k-1)! \\ &= -\ln x + k_1 \ln x - \lambda x + k_1 \ln \lambda - \ln((k-1)!) \end{aligned}$$

elde edilir. Böylece

$$C(x) = -\ln x, \quad \xi^1 = k, \quad \xi^2 = -\lambda, \quad F_1(x) = \ln x, \quad F_2(x) = -x$$

bulunur. Erlang dağılımının üstel aileye ait olduğu görülür.

8. (Ters Gauss Dağılımı) Ters Gauss dağılımı (Wald dağılımı olarak da bilinir) $(0, \infty)$ üzerinde iki parametrelili bir sürekli olasılık dağılım ailesidir. Bu dağılım, istatistikte ve mühendislikte geniş bir uygulama alanına sahiptir.

$\mathcal{X} = (0, \infty)$ ve $\mathbb{E} = (0, \infty) \times (0, \infty)$ olmak üzere, x rastgele değişkenli ve (μ, λ) iki parametrelili ters Gauss dağılımının olasılık yoğunluk fonksiyonu, $\forall x \in \mathcal{X}, \forall (\xi^1, \xi^2) \in \mathbb{E}$ olmak üzere

$$p(x; \xi^1, \xi^2) = p(x; \mu, \lambda) = \left(\frac{\lambda}{2\pi x^3} \right)^{\frac{1}{2}} \cdot \exp \left\{ -\frac{\lambda(x - \mu)^2}{2\mu^2 x} \right\}$$

şeklinde ifade edilebilir. Böylece $\xi = (\xi^1, \xi^2) = (\mu, \lambda)$ olacağından

$$\begin{aligned} \ln(x, \xi) = \ln P(x; \mu, \lambda) &= \frac{1}{2} \ln x - \frac{1}{2} \ln(2\pi x^3) - \frac{\lambda(x^2 - 2\mu x + \mu^2)}{2\mu^2 x} \\ &= \frac{1}{2} \ln \lambda - \frac{1}{2} \ln(2\pi x^3) - \frac{\lambda x}{2\mu^2} + \frac{\lambda}{\mu} - \frac{\lambda}{2x} \end{aligned}$$

dir. Bu fonksiyonunun üstel fonksiyonu alınır, yoğunluk fonksiyonu aşağıdaki gibi de ifade edilebilir;

$$\begin{aligned} p(x, \xi) = e^{\ln p(x, \xi)} &= e^{\frac{1}{2} \ln \lambda - \frac{1}{2} \ln(2\pi x^3) - \frac{\lambda}{2\mu^2} + \frac{\lambda}{\mu} - \frac{\lambda}{2x}} \\ &= \frac{1}{\sqrt{2\pi x^3}} \cdot e^{\left\{ -\frac{\lambda}{2\mu^2} x - \frac{\lambda}{2x} + \frac{\lambda}{\mu} + \frac{1}{2} \ln x \right\}} \end{aligned}$$

O halde $p_\xi(x) = p(x, \xi) = h(x) \cdot e^{\xi^1 F_1(x) - \psi(\xi)}$ olmak üzere

$$C(x) = \frac{1}{\sqrt{2\pi x^3}}, \quad \xi^1 = \frac{\lambda}{2\mu^2}, \quad \xi^2 = \frac{-\lambda}{2}, \quad F_1(x) = x, \quad F_2(x) = \frac{1}{x}$$

olur ve böylece ters Gauss dağılımının üstel aileye ait olduğu görülür.

9. (Genelleştirilmiş Ters Gauss Dağılımı) Genelleştirilmiş Ters Gauss Dağılımı (GIG), sürekli bir olasılık dağılımı olup, standart Ters Gauss Dağılımının bir genelleştirmesidir. Bu dağılım, istatistiksel modelleme ve Bayesian analizde yaygın olarak kullanılır. GIG, aşağıdaki yoğunluk fonksiyonu ile tanımlanır:

$$p(x; a, b, \lambda) = \frac{\left(\frac{a}{b} \right)^{\frac{\lambda}{2}}}{2K_\lambda(\sqrt{ab})} x^{\lambda-1} \cdot e^{-\frac{(ax + \frac{b}{x})}{2}}, \quad x > 0.$$

Burada K_λ modifiye Bessel fonksiyonlarıdır. $\mathcal{X} = \mathbb{R}^+$ biçimindedir ve parametreler $a > 0$ $x, b > 0$ ve $\lambda \in \mathbb{R}$ olmak üzere parametre uzayı

$$\mathbb{E} = \mathbb{R}^+ \times \mathbb{R}^+ \times \mathbb{R} = (0, \infty) \times (0, \infty) \times \mathbb{R}$$

dir. Ayrıca,

$$e^{c(x)+\xi f_i(x)-\psi(\xi)} = \frac{\left(\frac{a}{b}\right)^{\frac{\lambda}{2}}}{2k\lambda\sqrt{ab}} \cdot x^{\lambda-1} \cdot e^{\frac{-(ax+\frac{b}{x})}{2}}$$

olarak değerlendirilirse

$$c(x) + \xi_i^i(x) - \psi(\xi) = \ln\left(\left(\frac{a}{b}\right)^{\frac{\lambda}{2}}\right) - \ln(2k\lambda\sqrt{ab}) + (\lambda - 1) \ln x - \left(\frac{ax + \frac{b}{x}}{2}\right)$$

olur. Buradan

$$c(x) + \xi_i^i(x) - \psi(\xi) = \frac{p}{2 \ln\left(\frac{a}{b}\right)} - \ln\left(2k_\lambda(\sqrt{a_0})\right) + (b - 1) \ln x - \left(\frac{ax + \frac{b}{x}}{2}\right)$$

ifadesi gelir. Böylece

$$\begin{aligned} c(x) + \xi^1 F_1(x) + \xi^2 F_2(x) + \xi^3 \cdot F_3(x) - \psi(\xi) \\ = \lambda \cdot \ln x - \ln x - \frac{a}{2} x - \frac{b}{2} \cdot \frac{1}{x} + \frac{\lambda}{2} \ln\left(\frac{a}{b}\right) - \ln(2k\lambda\sqrt{ab}) \\ = -\ln x + a \cdot \left(\frac{-x}{2}\right) + b \cdot \left(-\frac{1}{2x}\right) + \lambda \ln x + \frac{\lambda}{2} \ln\left(\frac{a}{b}\right) - \ln(2k\lambda\sqrt{ab}) \end{aligned}$$

olur. Buradan

$$c(x) = -\ln x, \quad F_1(x) = -\frac{x}{2}, \quad F_2(x) = -\frac{1}{2x}, \quad F_3(x) = \ln x$$

ve

$$\psi(\xi) = \frac{\lambda}{2} \ln\left(\frac{a}{b}\right) - \ln(2k\lambda\sqrt{ab})$$

olduğundan genelleştirilmiş ters Gauss dağılımı üstel aileye aittir.

10. (Ters Gamma Dağılımı) Ters gamma dağılımı, gamma dağılımına göre dağılan bir değişkenin tersinin dağılımı olan pozitif reel doğru üzerinde sürekli olasılık dağılımlarının iki parametrelili bir ailesidir. Ters gamma dağılımının olasılık yoğunluk fonksiyonu

$$p(x, \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{\left(\frac{-\beta}{x}\right)}$$

biçimindedir. $x > 0$, $\alpha > 0$, $\beta > 0$ olduğundan örnek uzay $\mathcal{X} = (0, +\infty)$ ve parametre uzayı $\mathbb{E} = (0, \infty) \times (0, +\infty)$ biçimindedir.

$$e^{c(x)+\xi^i F_i(x)-\psi(\varepsilon)} = \frac{\beta^\alpha}{\Gamma(\alpha)} \cdot x^{-\alpha-1} \cdot e^{\left(\frac{-\beta}{\varepsilon}\right)}$$

olarak ele alınırsa,

$$C(x) + \xi^1 f_i(x) - \psi(\varepsilon) = \ln(\beta^\alpha) - \ln(\Gamma(\alpha)) + \ln(x^{-1}) - \frac{\beta}{x}$$

$$c(x) + \xi^1 \cdot f_1(x) + \xi^2 \cdot f_2(x) - \psi(\xi) = \alpha \cdot \ln \beta - \ln(\Gamma(\alpha)) + (-\alpha - 1) \cdot \ln x - \frac{\beta}{x}$$

$$= -\ln x + \alpha(-\ln x) + \beta \left(-\frac{1}{x}\right) + \alpha \ln \beta - \ln(\Gamma(\alpha))$$

$$c(x) = -\ln x, \quad f_1(x) = -\ln x, \quad f_2(x) = -\frac{1}{x}, \quad \psi(\xi) = \alpha \cdot \ln \beta - \ln(\Gamma(\alpha))$$

olacağından genelleştirilmiş ters gamma dağılımı üstel aileye aittir.

11. (Rayleigh Dağılımı) X rasgele değişkeni negatif değerler almayan mutlak sürekli bir rasgele değişken olmak üzere dağılımın olasılık yoğunluk fonksiyonu

$$p(x; \xi) = \frac{x}{\xi^2} e^{-\frac{x^2}{2\xi^2}}$$

biçimindedir. Böylece, örnek uzay $\mathcal{X} = [0, \infty)$ ve parametre uzayı $\mathbb{E} = (0, \infty)$ dir. Üstel ailenin genel formu göz önüne alınırsa,

$$e^{C(x)} + \xi^1 F_1(x) - \psi(\xi) = \frac{x}{\xi^2} \cdot e^{-\frac{x^2}{2\xi^2}}$$

ve böylece

$$C(x) + \xi^1 F_1(x) - \psi(\xi) = \ln x - \ln \xi^2 + \left(-\frac{x^2}{2\xi^2}\right)$$

$$= \ln x - 2 \ln \xi - \frac{1}{2\xi^2} x^2$$

ifadesinden

$$C(x) = \ln x, \quad \xi' = -\frac{1}{2\xi^2}, \quad F_1(x) = x^2, \quad \psi(\xi) = 2 \ln \xi$$

elde edilir. O halde Rayleigh dağılımı üstel aileye aittir.

Yukarıda üstel aileye ait, çok karşılaşılan bazı dağılım örnekleri verilmiştir. Literatürde bu aileye ait çok sayıda dağılım vardır (bakınız URL-1).

3.5.2. Üstel Ailenin Geometrik Modellenmesi

İstatistik teoride önemli yer tutan, üstel dağılım ailesinin genel bilgileri ve örnekleri yukarıda verildi. Bu bölümde, üstel ailenin bir istatistiksel manifold olarak modellenmesi sunulacaktır.

Üstel aile doğal olarak ξ^j parametreleri ile parametrelendirilmiştir. Parametre uzayı $\mathbb{E}$, $\mathbb{R}^n$ 'in bir açık alt kümesidir.

Teorem 3.5.1. Olasılık yoğunluk fonksiyonu $p(x; \xi) = e^{C(x) + \xi^i F_i(x) - \psi(\xi)}$ olmak üzere $\mathcal{S} = \{p(x; \xi)\}$ istatistiksel modeli bir istatistiksel manifolddur (Calin ve Udrişte, 2014).

İspat . Üstel ailenin bir manifold olduğunu ispatlamak için (I) ve (II) koşullarının sağlandığı gösterilmelidir. $\iota: \mathbb{E} \rightarrow \mathcal{S}$, $\iota(\xi) = \ln p(x; \xi)$ dönüşümü tanımlansın.

Birebirlik: $\iota(\xi) = \iota(\theta)$ olsun. O zaman $\ln p(x; \xi) = \ln p(x; \theta)$ olur. Böylece

$$\xi^i F_i(x) - \psi(\xi) = \theta^i F_i(x) - \psi(\theta)$$

eşitliği yazılabilir. $\{F_i(x)\} \cup \{1\}$ kümesi lineer bağımsız olduğundan $\xi^i = \theta^i$ elde edilir. Sonuç olarak dönüşüm birebirdir.

Regülerlik: Log-olabilirlik fonksiyonu $\ell_x(\xi) = C(x) + \xi^i F_i(x) - \psi(\xi)$ 'nin türevi alınırsa $\partial_j \ell_x(\xi) = F_j(x) - \partial_j \psi(\xi)$ ifadesi elde edilir. $\{F_j(x)\}$ kümesi lineer bağımsız olduğundan $\{\partial_j \ell_x(\xi)\}$ kümesi de lineer bağımsızdır. O halde regülerlik koşulu sağlanır. Sonuç olarak $\mathcal{S} = \{p(x; \xi)\}$ bir istatistiksel manifolddur.

Normalizasyon fonksiyonu $C(x)$ ve $F_i(x)$ 'e bağlıdır. Bu fonksiyon manifoldun geometrisi hakkında önemli bilgiler içerir. Aşağıdaki önerme fonksiyon ve beklenen değer arasındaki kullanışlı bir ilişkiyi vermektedir.

Önerme 3.5.1. (3.1) koşulunun sağlandığı kabul edilsin ve $E_\xi[\cdot]$, p_ξ 'nin beklenen değeri olmak üzere $\{F_i(x)\}$ fonksiyonları için integrallenebilme koşulu sağlasın. O zaman

$$\partial_j \psi(\xi) = E_\xi[F_j], \quad 1 \leq j \leq n$$

dir (Calin ve Udrişte, 2014).

İspat. $\int_{\mathcal{X}} p(x; \xi) dx = 1$ ile verilen normalizasyon koşulunda her iki tarafın diferansiyeli alınırsa $\int_{\mathcal{X}} \partial_{\xi^j} p(x; \xi) dx = 0$ olacağından

$$\int_{\mathcal{X}} p(x; \xi) (F_j(x) - \partial_{\xi^j} \psi(\xi)) dx = 0$$

bulunur. Böylece

$$\int_{\mathcal{X}} p(x; \xi) F_j(x) dx = \partial_{\xi^j} \psi(\xi) \int_{\mathcal{X}} p(x; \xi) dx$$

olacağından $E_\xi[F_j] = \partial_{\xi^j} \psi(\xi) = \partial_j \psi(\xi)$ sonucuna ulaşılır.

Sonuç. Aşağıdaki özellikler mevcuttur:

- (i) $\partial_j \ell_x(\xi) = F_j(x) - E_\xi[F_j]$,
- (ii) $E_\xi[\partial_i \ell_x(\xi) \partial_j \ell_x(\xi)] = \text{Cov}_\xi(F_i, F_j)$,
- (iii) $\partial_i \partial_j \ell_x(\xi) = -\partial_i \partial_j \psi(\xi)$.

Yukarıdaki sonuç ile verilen eşitlikler log-olabilirlik fonksiyonu hakkında kullanışlı neticelerdir. (iii) numaralı eşitlik log-olabilirlik fonksiyonunun konkavlığını sağlar. Buradan şu sonuç verilebilir:

Sonuç. Üstel dağılımın log-olabilirlik fonksiyonunu konkavdır.

Konkav bir fonksiyonun global maksimumu tek olur. Konkavlık, optimizasyon ve maksimum olabilirlik tahmini (MLE) gibi istatistiksel yöntemlerin etkinliğini artırır. Üstel dağılımın log-olabilirlik fonksiyonunun konkav olması, maksimum olabilirlik tahmini, numerik optimizasyon, veri uyumu ve istikrarlı parametre tahmini açısından önemli faydalar sunar. Bu özellik, özellikle makine öğrenimi ve istatistiksel modelleme gibi alanlarda önemli bir rol oynar. Pratt, bu konuda önemli bir çalışma yapmıştır (Pratt, 1981).

3.6. Karışım Aile (Mixture Family)

Tanım 3.6.1. n -sayıda $F_i: \mathcal{X} \rightarrow \mathbb{R}$ düzgün fonksiyonları $\mathcal{X}$ üzerinde lineer bağımsız olsunlar ve

$$\int_{\mathcal{X}} F_j(x) dx = 0, 1 \leq j \leq n \quad (3.3)$$

koşulunu sağlasınlar. $\mathcal{X}$ üzerinde tanımlanan bir başka $C(x)$ fonksiyonu için

$$\int_{\mathcal{X}} C(x) = 1$$

koşulu gerçeklensin.

$$p(x; \xi) = C(x) + \xi^i F_i(x)$$

biçiminde tanımlanan olasılık dağılımı ailesine karışım ailesi (mixture family) denir (Calin ve Udrişte, 2014).

Teorem 3.6.1. $p(x; \xi) = C(x) + \xi^i F_i(x)$ olmak üzere $\mathcal{S} = \{p(x; \xi)\}$ ile verilen karışım ailesi istatistiksel modeli bir istatistiksel manifolddur (Calin ve Udrişte, 2014).

İspat. $\iota: \mathbb{E} \rightarrow \mathcal{P}(\mathcal{X})$ dönüşümü verilsin. $\{F_i(x)\}$ fonksiyonları lineer bağımsız olduğundan birebirlik regülerlik koşulları doğrudan sağlanır.

ξ^j pramaterelerine karışım parametreleri denir. $\mathbb{E}$ parametre uzayının boyutuna karışım ailesinin mertebesi denir.

Karışım ailesinden elde edilen, $\mathcal{S} = \{p(x; \xi)\}$ istatistiksel modeli $\mathcal{P}(\mathcal{X})$ 'in **bir afin alt uzayıdır.**

(3.3) integral koşulu $\partial_{\xi^j} p(x; \xi) = F_j(x)$ ve $\int \partial_{\xi^j} p(x; \xi) dx = 0$ gerçeğiyle uyumludur.

Aşağıdaki önerme karışım ailesi için beklenen değerin özelliklerini vermektedir.

Önerme 3.6.1. $E_{\xi}[\cdot]$, p_{ξ} 'ye göre beklenen değer olmak üzere

$$(i) \partial_j \ell_x(\xi) = \frac{F_j(x)}{p(x)}$$

$$(ii) \partial_i \partial_j \ell_x(\xi) = -\frac{F_i(x)F_j(x)}{p(x;\xi)^2}$$

$$(iii) \partial_i \partial_j \ell_x(\xi) = -\partial_i \ell_x(\xi) \partial_j \ell_x(\xi)$$

dir (Calin ve Udrişte, 2014).

İspat.

(i) $\partial_{\xi^j} p(x; \xi) = F_j(x)$ ifadesi kullanılır ve türev alınır

$$\partial_j \ell_x(\xi) = \partial_{\xi^j} \ln p(x; \xi) = \frac{\partial_{\xi^j} p(x; \xi)}{p(x; \xi)} = \frac{F_j(x)}{p(x; \xi)}$$

elde edilir.

(ii) (i)'de bulunan eşitliğin türevi alındığında,

$$\begin{aligned} \partial_i \partial_j \ell_x(\xi) &= \partial_{\xi^i} \frac{F_j(x)}{p(x; \xi)} = F_j(x) \frac{-\partial_{\xi^i} p(x; \xi)}{p(x; \xi)^2} \\ &= -\frac{F_i(x)F_j(x)}{p(x; \xi)^2} \end{aligned}$$

bulunur.

(iii) (ii) ve (i) den

$$\partial_i \partial_j \ell_x(\xi) = -\frac{F_i(x)}{p(x; \xi)} \frac{F_j(x)}{p(x; \xi)} = -\partial_i \ell_x(\xi) \partial_j \ell_x(\xi)$$

eşitliği elde edilir.

Örnek 3.6.1. (Dağılımlarının Karışımı) $\mathcal{X}$ üzerinde doğrusal olarak bağımsız olan $p_1, p_2, \dots, p_n, p_{n+1}$ biçimindeki $n + 1$ olasılık yoğunlukları göz önüne alınsın ve $\mathbb{E} = \{\xi \in \mathbb{R}^n | \xi^j > 0, \sum_{j=1}^n \xi^j < 1\}$ olsun. Bu durumda aşağıdaki ağırlıklı ortalama da bir olasılık yoğunluğudur:

$$\begin{aligned} p(x; \xi) &= \xi^i p_i(x) + \left(1 - \sum_{i=1}^n \xi^i\right) p_{n+1}(x) \\ &= p_{n+1}(x) + \sum_{i=1}^n \xi^i (p_i(x) - p_{n+1}(x)). \end{aligned}$$

Bu ifade, $C(x) = p_{n+1}(x)$, ve $F_i(x) = p_i(x) - p_{n+1}(x)$ olan bir karışım ailesi haline

gelir.

Örnek 3.6.2. Yukarıdaki genel örneğe uygun olarak aşağıdaki örnekler verilebilir.

$$p_{n+1}(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

için $p_i(x)$, varyansları veya ortalamaları farklı normal dağılımlar olsunlar. Örneğin:

$$p_i(x) = \frac{1}{\sqrt{2\pi\sigma_i^2}} e^{-\frac{(x-\mu_i)^2}{2\sigma_i^2}}$$

burada μ_i ve σ_i farklı parametrelerdir. Bu durumda karışım yoğunluğu şu şekilde olur:

$$p(x; \xi) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} + \sum_{i=1}^n \xi^i \left(\frac{1}{\sqrt{2\pi\sigma_i^2}} e^{-\frac{(x-\mu_i)^2}{2\sigma_i^2}} - \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \right).$$

$n = 2$ için 3 farklı bileşenden oluşan bir karışım ailesinin olasılık yoğunluk fonksiyonu

$$p(x; \xi) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} + \xi^1 \left(\frac{1}{\sqrt{\pi}} e^{-2(x-1)^2} - \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \right) + \xi^2 \left(\frac{1}{\sqrt{\pi}} e^{-2(x+1)^2} - \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \right)$$

biçiminde olur. Burada, $\xi^1 > 0, \xi^2 > 0$ ve $\xi^1 + \xi^2 < 1$ olup karıştırılan dağılımlar

$$p_3(x) \rightarrow \mathcal{N}(0,1),$$

$$p_1(x) \rightarrow \mathcal{N}(1,0.5),$$

$$p_2(x) \rightarrow \mathcal{N}(-1,0.5).$$

biçimindedir. Bu model, bir veri kümesinin normal olmayan davranışını açıklamak için kullanışlıdır. Örneğin, $\mathcal{N}(0,1)$ standart davranışı temsil ederken, diğer bileşenler (p_1, p_2) uç değerleri veya farklı modları modelleyebilir. Bu karışım, özellikle veri analizinde (örneğin, anormal durumları modellemek veya alt popülasyonları belirlemek için) etkili bir modeldir.

4. OLASILIK DAĞILIMLARININ RIEMANN GEOMETRİSİ

Bu bölümde, bir geometrik yapı olarak modellenen olasılık dağılımları üzerinde Riemann metriği ve eğrilik özellikleri ele alınacaktır. Detaylı bilgi için (Calin ve Udrişte, 2014) kaynağına bakılabilir.

4.1. Fisher Bilgi Metriği

Fisher bilgisi, istatistiksel modellerde bir parametrenin ne kadar iyi tahmin edilebileceğini ölçen bir kavramdır. Bu bilgi, parametrelerin etrafındaki veri dağılımının yerel eğriliği yansıtarak, parametre tahminlerinin hassasiyet düzeyini belirler. Ayrıca, istatistiksel tahminlerin verimliliğini değerlendirmede önemli bir rol oynar. Fisher bilgisi parametre uzayının geometrik modellenmesinde bir Riemann metriği vermektedir.

4.1.1. Fisher Bilgi Metriğin Elde Edilmesi

Bir parametre ξ için Fisher bilgisi, log-olabilirlik fonksiyonunun ξ ya göre ikinci türevinin negatif beklentisi olarak tanımlanır.

Log-olabilirlik genellikle konkav bir fonksiyondur. Log-olabilirlik fonksiyonunun ikinci türevinin beklentisi, modelin parametre çevresindeki "eğriliğini" ve bu eğriliğin tahmin ediciye etkisini ifade eder. Bu, parametre çevresindeki duyarlılığı ölçer ve aşağıdaki şekilde tanımlanır:

$$I(\xi) = -E \left[\frac{\partial^2}{\partial \xi^2} \ln f(X; \xi) \right].$$

Bu ifadeler, Fisher bilgisinin parametre tahminindeki kesinlik ve doğruluğu nasıl etkilediğini açıklar..

Önerme 4.1.1. Fisher bilgisi $I(\xi)$, $\partial_\xi \ell_\xi$ 'nin varyansıdır.

İspat. Log-olabilirlik fonksiyonu $\ell_\xi = \ln p(x; \xi)$ olmak üzere

$$E \left[\frac{\partial}{\partial \xi} \ell_\xi(x) \right] = 0$$

olduğu biliniyor. Log-olabilirlik fonksiyonunun türevinin varyansı şu şekilde ifade edilir:

$$\text{Var} \left(\frac{\partial}{\partial \xi} \ell_\xi(x) \right) = E \left[\left(\frac{\partial}{\partial \xi} \ell_\xi(x) \right)^2 \right] - \left(E \left(\frac{\partial}{\partial \xi} \ell_\xi(x) \right) \right)^2$$

dolayısıyla

$$Var\left(\frac{\partial}{\partial \xi} \ell_{\xi}(x)\right) = E\left[\left(\frac{\partial}{\partial \xi} \ell_{\xi}(x)\right)^2\right].$$

olur. Bu, Fisher bilgisinin tanımıyla örtüşür, yani

$$I(\xi) = E\left[\left(\frac{\partial}{\partial \xi} \ell_{\xi}(x)\right)^2\right]$$

dir. Dolayısıyla, Fisher bilgisi $I(\xi)$ log-olabilirlik fonksiyonunun türevinin varyansıdır.

$\mathcal{S} = \{p(x; \xi)\}$ istatistiksel model için, Fisher bilgi matrisi $I(\xi) = E_{\xi}[\partial_i \ell(\xi) \partial_j \ell(\xi)]$ şeklinde yazılabilir. Her $i, j \in \{1, \dots, n\}$ için Fisher bilgi matrisi

$$g_{ij}(\xi) = E_{\xi}[\partial_i \ell(\xi) \partial_j \ell(\xi)]$$

şeklinde gösterilsin. Bir olasılık yoğunluk fonksiyonun beklenen değerinin tanımı göz önüne alınırsa Fisher bilgi matrisi daha açık olarak,

$$g_{ij}(\xi) = \int_{\mathcal{X}} \partial_{\xi^i} \ln p_{\xi}(x) \cdot \partial_{\xi^j} \ln p_{\xi}(x) \cdot p_{\xi}(x) dx \quad (4.1)$$

şeklinde ifade edilebilir (Calin ve Udrişte, 2014).

Teorem 4.1.1. Fisher bilgi matrisi olasılık yoğunluk fonksiyonlarının karekökü cinsinden aşağıdaki gibi ifade edilebilir (Calin ve Udrişte, 2014);

$$g_{ij}(\xi) = 4 \int_{\mathcal{X}} \partial_{\xi^i} \sqrt{p_{\xi}(x)} \cdot \partial_{\xi^j} \sqrt{p_{\xi}(x)} dx.$$

İspat (4.1) ifadesindeki metriğin tanımındaki türevler alınır

$$g_{ij}(\xi) = \int_{\mathcal{X}} \frac{\partial_{\xi^i} p_{\xi}(x)}{p_{\xi}(x)} \cdot \frac{\partial_{\xi^j} p_{\xi}(x)}{p_{\xi}(x)} \cdot p_{\xi}(x) dx$$

olur. Bu ifade uygun çarpanlarla

$$g_{ij}(\xi) = 4 \int_{\mathcal{X}} \frac{\partial_{\xi^i} p_{\xi}(x)}{\sqrt{p_{\xi}(x)}} \cdot \frac{\partial_{\xi^j} p_{\xi}(x)}{2\sqrt{p_{\xi}(x)}} dx$$

biçiminde düzenlenebilir. Böylece

$$g_{ij}(\xi) = 4 \int_{\mathcal{X}} \partial_{\xi^i} \sqrt{p_{\xi}(x)} \cdot \partial_{\xi^j} \sqrt{p_{\xi}(x)} dx$$

eşitliğine ulaşılır

Yukarıdaki teoremin ifadesi kesikli değişkenler için aşağıdaki şekilde ifade edilir;

$$g_{ij}(\xi) = 4 \partial_{\xi^i} \sqrt{p_{\xi}(x_k)} \cdot \partial_{\xi^j} \sqrt{p_{\xi}(x_k)}.$$

Örnek 4.1.1. $\xi^i \in (0, 1)$ olmak üzere

X	x_1	x_2	x_3
$P(X = x_k)$	ξ^1	ξ^2	$1 - \xi^1 - \xi^2$

şeklinde verilen tabloda, ξ^1 , ξ^2 ve $1 - \xi^1 - \xi^2$ olasılıklarıyla ortaya çıkan üç sonucu x_1, x_2, x_3 olan bir rastgele değişken X 'i göstermektedir. Bu, ξ^1 ve ξ^2 iki parametreye bağlı istatistiksel bir modeldir. Bu durumda Fisher bilgi matrisi

$$\begin{aligned}
 g_{ij}(\xi) &= 4 \sum_k \partial_{\xi^i} \sqrt{p_{\xi}(x_k)} \cdot \partial_{\xi^j} \sqrt{p_{\xi}(x_k)} = 4 \left(\partial_{\xi^i} \sqrt{\xi^1} \cdot \partial_{\xi^j} \sqrt{\xi^1} + \partial_{\xi^i} \sqrt{\xi^2} \cdot \partial_{\xi^j} \sqrt{\xi^2} \right) \\
 &\quad + 4 \left(\partial_{\xi^i} \sqrt{1 - \xi^1 - \xi^2} \cdot \partial_{\xi^j} \sqrt{1 - \xi^1 - \xi^2} \right) \\
 &= \frac{\delta_{i1j1}}{\xi^1} + \frac{\delta_{i2j2}}{\xi^2} + \frac{1}{\sqrt{1 - \xi^1 - \xi^2}}.
 \end{aligned}$$

ve buradan

$$(g_{ij}) = \begin{pmatrix} \frac{1}{\xi^1} + \frac{1}{\sqrt{1 - \xi^1 - \xi^2}} & \frac{1}{\sqrt{1 - \xi^1 - \xi^2}} \\ \frac{1}{\sqrt{1 - \xi^1 - \xi^2}} & \frac{1}{\xi^2} + \frac{1}{\sqrt{1 - \xi^1 - \xi^2}} \end{pmatrix}$$

olarak bulunur. Her $\xi^i \in (0, 1)$ için $\det(g_{ij}) \neq 0$ olduğundan matris dejenere değildir (Calin ve Udrişte, 2014).

Önerme 4.1.2. Herhangi bir istatistiksel model üzerindeki Fisher bilgi matrisi simetrik, pozitif tanımlı ve dejenere değildir. Yani bir Riemann metriğidir (Calin ve Udrişte, 2014).

İspat. Metriğin tanımı simetriyi verir. Bu nedenle öncelikle Fisher bilgi matrisinin bir Riemann metriği vereceği kanıtlanacaktır. g nin bir metrik olabilmesi için, pozitif tanımlılık, simetri koşullarının gerçekleşmesi gerekir. Her $\xi, v \in T_{\xi}S, v \neq 0$ için ,

$$g_{ij}(\xi) = 4 \int_{\mathcal{X}} \partial_{\xi^i} \sqrt{p_{\xi}(x)} \cdot \partial_{\xi^j} \sqrt{p_{\xi}(x)} dx$$

ifadesinden

$$g(v, v) = \sum_i g_{ii} v^i v^i = 4 \sum_i \int_{\mathcal{X}} v^i \partial_{\xi^i} \sqrt{p_{\xi}(x)} \cdot v^i \partial_{\xi^i} \sqrt{p_{\xi}(x)} dx$$

yazılabilir Böylece

$$g(v, v) = \int_{\mathcal{X}} \sum_i \left(v^i \partial_{\xi^i} \sqrt{p_{\xi}(x)} \right)^2 dx \geq 0$$

biçiminde olur. Dolayısıyla (g_{ij}) negatif değildir. Diğer yandan $g(v, v) = 0$ alınırsa

$$\sum_i v^i \partial_{\xi^i} \sqrt{p_{\xi}(x)} = 0$$

ve böylece $\{\partial_{\xi^i} p\}$ doğrusal bağımsız olduğundan $\forall i = 1, \dots, n$, için $v^i = 0$ elde edilir. Aynı işlemler zıt yönde takip edilerek tekrarlandığında metriğin non-dejenere olma koşulu sağlanmış olur. (g_{ij}) dejenere olmadığından, pozitif tanımlıdır.

Tanım 4.1.1. Fisher bilgi matrisi, $\mathcal{S}$ yüzeyinde bir Riemann metriğinin katsayılarını sağlar. $g_{ij}(\xi)$ metriğine Fisher bilgi metriği denir.

Fisher bilgi metriği, mesafelerin, açıların ölçülmesine ve istatistiksel modeller üzerindeki bağıntıların tanımlanmasına olanak tanır. Fisher bilgisi tanımında, bu bilginin log-olabilirlik fonksiyonunun ikinci türevinin negatif beklenen değeri olduğu ifade edilmişti. Aşağıdaki teorem Fisher bilgisi ile Fisher metriğinin çakıştığını ifade etmektedir.

Teorem 4.1.2. Fisher metriği

$$g_{ij}(\xi) = -E_{\xi}[\partial_i \partial_j l(\xi)] = -E_{\xi}[\partial_{\xi^i} \partial_{\xi^j} \ln p_{\xi}]$$

biçiminde ifade edilebilir (Amari ve Nagaoka, 2000).

İspat. $\int_{\mathcal{X}} p(x; \xi) dx = 1$ olduğundan,

$$\int_{\mathcal{X}} \partial_{\xi^i} p(x; \xi) dx = 0$$

elde edilir. Buradan,

$$E_{\xi}[\partial_{\xi^i} \ln p_{\xi}] = \int_{\mathcal{X}} \partial_{\xi^i} \ln p(x; \xi) \cdot p_{\xi}(x) dx = \int_{\mathcal{X}} \partial_{\xi^i} p(x; \xi) dx = 0$$

olur. Bu ifadenin ξ^j ye göre diferansiyeli alınırsa

$$\int_{\mathcal{X}} \partial_{\xi^j} \partial_{\xi^i} \ln p(x; \xi) \cdot p_{\xi}(x) dx + \int_{\mathcal{X}} \partial_{\xi^i} \ln p(x; \xi) \cdot \partial_{\xi^j} p_{\xi}(x) dx = 0$$

dir. Buradan

$$E_{\xi}[\partial_{\xi^i} \partial_{\xi^j} \ln p_{\xi}] + \int_{\mathcal{X}} \partial_{\xi^i} \ln p(x; \xi) \cdot \partial_{\xi^j} \ln p(x; \xi) \cdot p_{\xi}(x) dx = 0$$

ve

$$E_{\xi}[\partial_{\xi^i} \partial_{\xi^j} \ln p_{\xi}] + g_{ij}(\xi) = 0.$$

elde edilir.

4.1.2. Fisher Metriği Üzerine Sonuçlar

Teorem 4.1.3. Fisher metriği örnek uzayın yeniden parametrizasyonları altında değişmezdir (Calin ve Udrişte, 2014).

İspat. $X, \mathcal{X} \subseteq \mathbb{R}^n$ örnek uzayına sahip bir rastgele değişken olsun ve bununla ilişkili yoğunluk fonksiyonu p_ξ istatistiksel bir model oluştursun. Örnek uzayların tersinir dönüşümünü $f: \mathcal{X} \rightarrow \mathcal{Y} \subseteq \mathbb{R}^n$, ve $Y = f(X)$ rastgele değişkeni ile ilişkili yoğunluk fonksiyonu $\tilde{p}_\xi(y)$ ile gösterilsin. Bu durumda iki olasılık yoğunluk fonksiyonu arasında

$$p_\xi(x) = \tilde{p}_\xi(y) \left| \frac{df(x)}{dx} \right|.$$

ilişkisi vardır. Burada $\left| \frac{df(x)}{dx} \right|$, f nin Jakobiyanını göstermektedir. Diğer yandan, log-olabilirlik fonksiyonları

$$\begin{aligned} \tilde{\ell}(\xi) &= \ln \tilde{p}_\xi(y) = \ln \tilde{p}_\xi(f(x)) \\ \ell(\xi) &= \ln p_\xi(x) = \tilde{\ell}(\xi) + \ln \left| \frac{df(x)}{dx} \right| \end{aligned}$$

biçimindedir. f nin ξ parametresine bağlı olmadığı varsayılırsa

$$\partial_{\xi^i} \ell(\xi) = \partial_{\xi^i} \tilde{\ell}(\xi)$$

olarak elde edilir. Böylece, Fisher bilgisi

$$\begin{aligned} g_{ij}(\xi) &= \int_{\mathcal{X}} \partial_{\xi^i} \ell(\xi) \partial_{\xi^j} \ell(\xi) p_\xi(x) dx \\ &= \int_{\mathcal{X}} \partial_{\xi^i} \tilde{\ell}(\xi) \partial_{\xi^j} \tilde{\ell}(\xi) \tilde{p}_\xi(f(x)) \left| \frac{df(x)}{dx} \right| dx \\ &= \int_{\mathcal{Y}} \partial_{\xi^i} \tilde{\ell}(\xi) \partial_{\xi^j} \tilde{\ell}(\xi) \tilde{p}_\xi(y) dy \\ &= \tilde{g}_{ij}(\xi), \end{aligned}$$

olur. Bu da istenen değişmezlik özelliğini kanıtlar.

Teorem 4.1.4. Fisher metriği, parametreler uzayının yeniden parametrizasyonları altında kovaryanttır (Calin ve Udrişte, 2014).

İspat. $\xi = (\xi^1, \dots, \xi^n)$ ve $\theta = (\theta^1, \dots, \theta^n)$ koordinat kümeleri $\xi = \xi(\theta)$ tersinir ilişkisi ile ilişkili olsun, yani, $\xi^j = \xi^j(\theta^1, \dots, \theta^n)$. ve $\tilde{p}_\theta(x) = p_{\xi(\theta)}(x)$ olsun. Zincir kuralına göre

$$\partial_{\theta^i} \tilde{p}_\theta = \frac{\partial \xi^k}{\partial \theta^i} \partial_{\xi^k} p_\xi, \quad \partial_{\theta^j} \tilde{p}_\theta = \frac{\partial \xi^r}{\partial \theta^j} \partial_{\xi^r} p_\xi,$$

olur ve böylece

$$\begin{aligned}\tilde{g}_{ij}(\theta) &= \int_{\mathcal{X}} \frac{1}{\tilde{p}_{\theta}(x)} \partial_{\theta^i} p_{\theta}(x) \partial_{\theta^j} p_{\theta}(x) dx \\ &= \left[\int_{\mathcal{X}} \frac{1}{p_{\xi(\theta)}(x)} \partial_{\xi^k} p_{\xi}(x) \partial_{\xi^r} p_{\xi}(x) dx \right] \frac{\partial \xi^k}{\partial \theta^i} \frac{\partial \xi^r}{\partial \theta^j} \\ &= g_{kr}(\xi)|_{\xi=\xi(\theta)} \frac{\partial \xi^k}{\partial \theta^i} \frac{\partial \xi^r}{\partial \theta^j}.\end{aligned}$$

elde edilir.

Son iki teorem, Fisher metriğinin iki ayırt edici özelliğe sahip olduğunu ortaya koymaktadır:

- (i) g_{ij} , $\mathcal{X}$ örnek uzayının yeniden parametризasyonları altında değişmezdir.
- (ii) g_{ij} , $\mathbb{E}$ parametreler uzayının yeniden parametризasyonları altında kovaryanttır (yani,

2-kovaryant tensör olarak dönüşür).

Yukarıda belirtilen özellikleri karşılayan bir metrik tekdir ve dolayısıyla Fisher metriğidir (Corcuera ve Giummole, 1998). Bu sonuçlar, Fisher metriğinin olasılık dağılımlarının geometrik modelleri olarak istatistiksel manifoldlar üzerindeki verimliliğini ortaya koymaktadır.

4.1.3. Bazı Dağılımların Fisher Metrikleri

Bu bölümde bazı özel dağılımların Fisher metriklerinin hesaplamaları verilecektir.

1. (Üstel Dağılım) $p_{\xi}(x) = \xi e^{-\xi x}$ olduğundan $\partial_{\xi}^2 \ln p_{\xi} = -\frac{1}{\xi^2}$ dir. Bu nedenle

$$g = g_{11} = -E[\partial_{\xi}^2 \ln p_{\xi}] = \int_0^{\infty} \frac{1}{\xi^2} p_{\xi}(x) dx = \frac{1}{\xi^2}$$

bulunur.

$\xi_0 = \xi(0)$ ve $\xi_1 = \xi(1)$ noktalarını birleştiren $\xi = \xi(t)$ eğrisinin uzunluğu

$$\int_0^1 \sqrt{g(\xi(t), \xi(t))} dt = \int_0^1 \sqrt{g_{11} \xi^2(t)} dt = \int_0^1 \left| \frac{\dot{\xi}(t)}{\xi(t)} \right| dt = \left| \ln \frac{\xi_1}{\xi_0} \right|$$

dir. Bu, tek boyutlu istatistiksel model $\mathcal{X} = (0, \infty)$ üzerinde $d(x_1, x_2) = |\ln x_2 - \ln x_1|$ hiperbolik uzaklık fonksiyonunu verir.

2. (Üstel Aile) $p(x; \xi) = e^{C(x) + \xi^i F_i(x) - \psi(\xi)}$ üstel ailesi için Fisher bilgi metriğinin katsayıları

$$g_{ij}(\xi) = E_{\xi}[\partial_i \ell_x(\xi) \partial_j \ell_x(\xi)] = \text{Cov}_{\xi}(F_i, F_j)$$

biçimindedir. Katsayılar $F_i(x)$ fonksiyonları cinsinden gösterildiği gibi ψ cinsinden

$$g_{ij}(\xi) = -E_{\xi}[\partial_i \partial_j \ell_x(\xi)] = E_{\xi}[\partial_i \partial_j \psi(\xi)] = \partial_i \partial_j \psi(\xi)$$

biçiminde elde edilir (Calin ve Udrişte, 2014).

Üstel dağılımın, üstel aileye ait olduğu biliniyor. 1-parametrelili üstel dağılım için $\psi(\xi) = -\ln(\xi)$ olduğu göz önüne alınır ve bu ifadenin iki kez türevi alınırsa $g = \frac{1}{\xi^2}$ olarak elde edilir. Bu da önceki örnek ile örtüşür. O halde üstel aileye mensup dağılımların Fisher metrikleri bu bağıntıda da elde edilebilir.

3. (Karışım Ailesi) $p(x; \xi) = C(x) + \xi^i F_i(x)$ olasılık dağılım ailesinin Fisher metriği katsayıları

$$g_{ij}(\xi) = -E_{\xi}[\partial_i \partial_j \ell_x(\xi)] = E_{\xi} \left[\frac{F_i(x) F_j(x)}{p(x; \xi)^2} \right] = \int_x \frac{F_i(x) F_j(x)}{p(x; \xi)} dx$$

olarak elde edilir (Calin ve Udrişte, 2014).

4. (Bernoulli Dağılımı) Bernoulli dağılımı geometrik modeli şu şekilde verilir:

$$\mathcal{S} = \{p(\xi; k) = \xi^k (1 - \xi)^{1-k} : 0 < \xi < 1, k \in \{0, 1\}\}.$$

Bernoulli olasılık yoğunluk fonksiyonu için log-olabilirlik fonksiyonu

$$\ell_k(\xi) = \ln p(\xi; k) = \ln(\xi^k (1 - \xi)^{1-k}) = k \ln \xi + (1 - k) \ln(1 - \xi)$$

olarak elde edilir. Parametreye göre birinci ve ikinci türevleri alınırsa,

$$\begin{aligned} \frac{\partial \ell_k(\xi)}{\partial \xi} &= \frac{k - \xi}{\xi(1 - \xi)}. \\ \frac{\partial^2 \ell_k(\xi)}{\partial \xi^2} &= -\frac{k}{\xi^2} - \frac{1 - k}{(1 - \xi)^2}. \end{aligned}$$

olur. Fisher metrik katsayılarını hesaplamak için

$$g_{ij}(\xi) = -E \left[\frac{\partial^2 \ell_x(\xi)}{\partial \xi^i \partial \xi^j} \right]$$

ifadesinden yararlanılır. Bu durumda

$$g_{11}(\xi) = E \left[\frac{k}{\xi^2} + \frac{1 - k}{(1 - \xi)^2} \right]$$

olur. Beklenen değer

$$\begin{aligned} E[k] &= \xi \\ E[1 - k] &= 1 - \xi \end{aligned}$$

olur. Sonuç olarak,

$$g_{11}(\xi) = \frac{1}{\xi} + \frac{1}{1 - \xi} = \frac{1}{\xi(1 - \xi)}.$$

olarak bulunur (Branea ve Radulescu, 2023).

4.2. Christoffel Sembolleri

İstatistiksel model $\mathcal{S}$ üzerindeki en temel bağlantı Christoffel sembolleri ile tanımlanır. Bir istatistiksel modelin g_{ij} Riemann metriği Fisher bilgi metriği olmak üzere, birinci tür Christoffel sembolleri

$$\Gamma_{ij,k} = \frac{1}{2}(\partial_i g_{jk} + \partial_j g_{ik} - \partial_k g_{ij})$$

ile verilir. Burada $\partial_i = \partial_{\xi^i}$ yi göstermektedir.

Christoffel sembollerini hesaplamadan önce, Fisher bilgi matrisinin türevi için birkaç eşdeğer ifadeye ihtiyaç vardır.

Teorem 4.2.1. $\ell = \ell_x(\xi)$ ve $\partial_i = \partial_{\xi^i}$ olmak üzere Fisher bilgi matrisi aşağıdaki özelliklere sahiptir (Calin ve Udrişte, 2014):

- (i) $\partial_k g_{ij}(\xi) = -E_{\xi}[\partial_i \partial_j \partial_k \ell] - E_{\xi}[(\partial_i \partial_j \ell)(\partial_k \ell)],$
- (ii) $\partial_k g_{ij}(\xi) = E_{\xi}[(\partial_k \partial_i \ell)(\partial_j \ell)] + E_{\xi}[(\partial_k \partial_j \ell)(\partial_i \ell)] + E_{\xi}[(\partial_i \ell)(\partial_j \ell)(\partial_k \ell)],$
- (iii) $\partial_k g_{ij}(\xi) = 4 \int_{\mathcal{X}} \partial_j \partial_k \sqrt{p} \cdot \partial_i \sqrt{p} dx + 4 \int_{\mathcal{X}} \partial_i \partial_k \sqrt{p} \cdot \partial_j \sqrt{p} dx.$

Fisher metriği tanımında kısmi türevler alınarak teoremdaki eşitlikler kolaylıkla elde edilebilir.

Önerme 4.2.1. Birinci tip Christoffel sembollerinin aşağıdaki eşdeğer ifadeleri geçerlidir (Calin ve Udrişte, 2014):

- (i) $2\Gamma_{ij,k}(\xi) = E_{\xi}[(\partial_i \partial_j \ell)(\partial_k \ell)] - E_{\xi}[(\partial_k \partial_j \ell)(\partial_i \ell)] - E_{\xi}[(\partial_k \partial_i \ell)(\partial_j \ell)] - E_{\xi}[\partial_i \partial_j \partial_k \ell]$
- (ii) $\Gamma_{ij,k}(\xi) = E_{\xi}[(\partial_i \partial_j \ell + \frac{1}{2} \partial_i \ell \partial_j \ell) \partial_k \ell];$
- (iii) $\Gamma_{ij,k}(\xi) = 4 \int_{\mathcal{X}} \partial_i \partial_j \sqrt{p(x; \xi)} \cdot \partial_k \sqrt{p(x; \xi)} dx.$

Yukarıda bahsedilen bağıntılar ayırık örnek uzayı durumunda da çalışır.

4.2.1. Bazı Dağılımların Christoffel Sembolleri

Bu bölümde bazı özel dağılımların Christoffel sembollerinin hesaplamaları verilecektir.

1. (Ayrık olasılık Modeli) Daha önce tanımlanan ayırık olasılık modeli göz önüne alınsın.

Eğer $i \neq j$ ise, o zaman

$$\begin{aligned}\Gamma_{ij,k}(\xi) &= 4 \sum_{k=1}^2 \left(\underbrace{\partial_{\xi^i} \partial_{\xi^j} \sqrt{\xi^r}}_{=0} \cdot \partial_{\xi^k} \sqrt{\xi^r} \right) + 4 \partial_{\xi^i} \partial_{\xi^j} \sqrt{1 - \xi_1 - \xi_2} \cdot \partial_{\xi^k} \sqrt{1 - \xi_1 - \xi_2} \\ &= \partial_{\xi^i} (1 - \xi^1 - \xi^2)^{-\frac{1}{2}} \cdot \frac{1}{\sqrt{1 - \xi^1 - \xi^2}} = \frac{1}{2(1 - \xi^1 - \xi^2)^2}\end{aligned}$$

bulunur. Diğer yandan $i = j$ için toplamın ikinci terimi yukarıdaki ile aynı olduğundan

$$\begin{aligned}\Gamma_{ii,k}(\xi) &= 4 \sum_{r=1}^2 \partial_{\xi^i}^2 \sqrt{\xi^r} \cdot \partial_{\xi^k} \sqrt{\xi^r} + \frac{1}{2(1 - \xi^1 - \xi^2)^2} \\ &= -\frac{1}{2} \sum_{r=1}^2 \frac{\delta_i^r \delta_r^k}{\xi^{r^2}} + \frac{1}{2(1 - \xi^1 - \xi^2)^2}.\end{aligned}$$

olur. Eğer $i = j \neq k$ ise yukarıdaki formülün ilk terimi yok olur, yani $\Gamma_{ii,k}(\xi) = \frac{1}{2(1 - \xi^1 - \xi^2)^2}$

olarak bulunur (Calin ve Udrişte, 2014).

2. (Üstel Aile) Üstel aile için Fisher metriği katsayıları $\psi(\xi)$ cinsinden verilmişti. Bu durumda Christoffel sembolleri de $\psi(\xi)$ fonksiyonu cinsinden hesaplanacaktır. Gerekli eşitlikler kullanılırsa

$$\begin{aligned}E_{\xi}[(\partial_i \partial_j \ell)(\partial_k \ell)] &= -E_{\xi}[(\partial_i \partial_j \psi)(F_k(x) - \partial_k \psi)] \\ &= -(\partial_i \partial_j \psi)E_{\xi}[F_k(x)] + (\partial_i \partial_j \psi)(\partial_k \psi) \\ &= (\partial_i \partial_j \psi)(\partial_k \psi - E_{\xi}[F_k(x)]) = 0.\end{aligned}$$

olarak bulunur. Böylece

$$\Gamma_{ij,k} = \frac{1}{2} E_{\xi}[\partial_i \partial_j \partial_k \ell] = \frac{1}{2} E_{\xi}[\partial_i \partial_j \partial_k \psi] = \frac{1}{2} \partial_i \partial_j \partial_k \psi(\xi)$$

olur (Calin ve Udrişte, 2014).

3. (Karışım Ailesi) Karışım ailesinin Christoffel sembolleri

$$\begin{aligned}\Gamma_{ij,k}(\xi) &= E_{\xi} \left[\left(\partial_i \partial_j \ell + \frac{1}{2} \partial_i \ell \partial_j \ell \right) \partial_k \ell \right] = E_{\xi} \left[\left(-\partial_i \ell \partial_j \ell + \frac{1}{2} \partial_i \ell \partial_j \ell \right) \partial_k \ell \right] \\ &= -\frac{1}{2} E_{\xi}[\partial_i \ell \partial_j \ell \partial_k \ell] = -\frac{1}{2} \int \frac{F_i(x) F_j(x) F_k(x)}{p(x; \xi)^2} dx.\end{aligned}$$

olarak elde edilir (Calin ve Udrişte, 2014).

4. (Üstel Dağılım) Üstel dağılımın Christoffel sembolleri

$$\Gamma_{11,1} = \frac{1}{2} \left(\frac{\partial g_{11}}{\partial \xi} + \frac{\partial g_{11}}{\partial \xi} - \frac{\partial g_{11}}{\partial \xi} \right) = -\frac{1}{2} \cdot \frac{2}{\xi^3} = -\frac{1}{\xi^3},$$

ve

$$\Gamma_{11}^1 = \frac{1}{2} g^{11} \left(\frac{\partial g_{11}}{\partial \xi} + \frac{\partial g_{11}}{\partial \xi} - \frac{\partial g_{11}}{\partial \xi} \right) = -\frac{1}{2} \cdot \xi^2 \cdot \frac{2}{\xi^3} = -\frac{1}{\xi}.$$

biçimindedir . Burada g^{11} Fisher matrsinin tersidir (Branea ve Radulescu, 2023).

5. (Bernoulli dağılımı) Dağılımın Fisher metriği katsayıları elde edilmiştir. Bu katsayılar ve Christoffel sembollerinin tanımı göz önüne alınırsa

$$\Gamma_{11,1} = \frac{1}{2} \left(\frac{\partial g_{11}}{\partial \xi} + \frac{\partial g_{11}}{\partial \xi} - \frac{\partial g_{11}}{\partial \xi} \right) = \frac{1}{2} \left(-\frac{1-2\xi}{\xi^2(1-\xi)^2} \right) = \frac{2\xi-1}{2\xi^2(1-\xi)^2}.$$

$$\Gamma_{11}^1 = \frac{1}{2} g^{11} \left(\frac{\partial g_{11}}{\partial \xi} + \frac{\partial g_{11}}{\partial \xi} - \frac{\partial g_{11}}{\partial \xi} \right),$$

olarak elde edilir. Burada g^{11} Fisher matrsinin tersidir ve Bernoulli dağılımı için $g^{11} = \xi(1-\xi)$ formundadır (Branea ve Radulescu, 2023).

4.3. Koneksiyon Bileşenleri

Riemann geometrisinin en önemli araçlarından biri koneksiyon. $\mathcal{S}$ istatistiksel modeli üzerinde farklı koneksiyonlar tanımlanmaktadır. Bu bölümde $\mathcal{S}$ istatistiksel modeli üzerindeki koneksiyonlar incelenecektir.

Bilgi geometrisinde, bir olasılık dağılımı manifoldundaki koneksiyon α -parametresiyle tanımlanan $\nabla^{(\alpha)}$ koneksiyon ailesi olarak ifade edilir. Bu koneksiyonlar, manifoldun geometrik yapısına ve farklı koordinat sistemlerine göre davranışını tanımlar. Buradaki α , manifold üzerindeki dual yapılar arasındaki geçişi kontrol eden bir parametredir. Örneğin, üstel aile ile karışım ailesi arasında bir düallite ilişkisi vardır.

4.3.1. $\nabla^{(0)}$ Levi-Civita Bağlantısı

$\mathcal{S}$ istatistiksel modeli temsil eden Riemann manifoldu olsun. $\mathcal{S}$ üzerinde bir Riemann metriği

$$g(X_\xi, Y_\xi) = \sum_{i,j=1}^n g_{ij}(\xi) a^i(\xi) b^j(\xi), \quad p_\xi \in \mathcal{S}$$

biçiminde tanımlanır. Burada $X_\xi = \sum_{i=1}^n a^i(\xi) \partial_i$ ve $Y_\xi = \sum_{j=1}^n b^j(\xi) \partial_j$, $\mathcal{S}$ üzerinde vektör alanlarıdır. g tensörü Fisher-Riemann metriği olarak adlandırılır. Onunla ilişkili Levi-Civita bağlantısı $\nabla^{(0)}$ ile gösterilir ve şu şekilde tanımlanır:

$$g\left(\nabla_{\partial_i}^{(0)} \partial_j, \partial_k\right) = \Gamma_{ij,k}.$$

$\nabla^{(0)}$ bir metrik koneksiyondur, yani metrik ile bağdaşabilir, bu lokal bileşenlerle

$$\partial_k g_{ij} = \Gamma_{ki,j} + \Gamma_{kj,p}$$

biçiminde yazılabilir. $\nabla^{(\alpha)}$ üst simgesi bağlantının bir parametresini gösterir. α parametresi ile parametrelendirilmiş çeşitli bağlantılar tanımlanacaktır. $\alpha = 0$ durumu Fisher metriği ile ilişkili Levi-Civita bağlantısına karşılık gelir (Calin ve Udrişte, 2014).

4.3.2. $\nabla^{(1)}$ Bağlantısı

Fisher metriği g kullanılarak, $\nabla^{(1)}$ -bağlantısı şu şekilde tanımlanır:

$$g(\nabla_{\partial_i}^{(1)} \partial_j, \partial_k) = E_{\xi}[(\partial_i \partial_j \ell)(\partial_k \ell)].$$

Böylece Christoffel sembolleri

$$\Gamma_{ij,k}^{(1)}(\xi) = E_{\xi}[(\partial_i \partial_j \ell)(\partial_k \ell)] \quad (4.2)$$

biçimindedir (Calin ve Udrişte, 2014).

S manifoldunun verilen koneksiyonlara göre düz olması ξ koordinat sistemine göre ele alınmaktadır.

Önerme 4.3.1. Üstel aile istatistiksel modeli $\nabla^{(1)}$ -düzdür (Calin ve Udrişte, 2014).

İspat. Üstel aile istatistiksel modelinin $\nabla^{(1)}$ -düz olduğunu göstermek için $\Gamma_{ij,k}^{(1)}(\xi) = 0$ olduğu elde edilmelidir. Eşitlik (4.2) kullanılırsa,

$$\begin{aligned} \Gamma_{ij,k}^{(1)}(\xi) &= E_{\xi}[(\partial_i \partial_j \ell)(\partial_k \ell)] = -E_{\xi}[\partial_i \partial_j \psi(\xi)(\partial_k \ell)] \\ &= -\partial_i \partial_j \psi(\xi) E_{\xi}[(\partial_k \ell)] = 0. \end{aligned}$$

bulunur.

Sonuç olarak, $\nabla^{(1)}$ bağlantısına göre üstel bir ailenin burulması ve eğriliği her yerde yok olur.

Önerme 4.3.2. Karışım ailesi $\nabla^{(1)}$ -düz olması için gerek ve yeter koşul $\nabla^{(0)}$ -düz olmasıdır (Calin ve Udrişte, 2014).

İspat. Christoffel sembollerinin tanımlarından,

$$\begin{aligned} \Gamma_{ij,k}^{(0)}(\xi) &= -\frac{1}{2} E_{\xi}[(\partial_i \ell)(\partial_j \ell)(\partial_k \ell)] \\ \Gamma_{ij,k}^{(1)}(\xi) &= E_{\xi}[(\partial_i \partial_j \ell)(\partial_k \ell)] = -E_{\xi}[(\partial_i \ell)(\partial_j \ell)(\partial_k \ell)] \end{aligned}$$

olur. Bu durumda ancak ve ancak $\Gamma_{ij,k}^{(0)}(\xi) = 0$ ise $\Gamma_{ij,k}^{(1)}(\xi) = 0$ olur. Bu da ispatı tamamlar.

4.3.3. $\nabla^{(-1)}$ Bağlantısı

Fisher metriği g kullanılarak, $\mathcal{S}$ istatistiksel modeli üzerindeki $\nabla^{(-1)}$ –bağlantısı şu şekilde tanımlanır:

$$g\left(\nabla_{\partial_i}^{(-1)} \partial_j, \partial_k\right) = \Gamma_{ij,k}^{(-1)} = E_{\xi}[(\partial_i \partial_j \ell + \partial_i \ell \partial_j \ell)(\partial_k \ell)]$$

burada ℓ log-likelihood fonksiyonudur (Calin ve Udrişte, 2014).

Herhangi bir karışım ailesinde $\partial_i \partial_j \ell = -\partial_i \ell \partial_j \ell$ dir. Buradan aşağıdaki önerme elde edilir.

Önerme 4.3.3. Karışım ailesi $\nabla^{(-1)}$ -düzdür (Calin ve Udrişte, 2014).

Önerme 4.3.4. Yukarıdaki üç bağlantı arasındaki ilişki

$$\nabla^{(0)} = \frac{1}{2}(\nabla^{(-1)} + \nabla^{(1)})$$

denklemi ile verilir (Calin ve Udrişte, 2014).

İspat. İspat için

$$\Gamma_{ij,k}^{(0)} = \frac{1}{2}(\Gamma_{ij,k}^{(-1)} + \Gamma_{ij,k}^{(1)})$$

olduğu gösterilmelidir. Yukarıdaki ispatlardan yararlanılırsa

$$\begin{aligned} \Gamma_{ij,k}^{(-1)} + \Gamma_{ij,k}^{(1)} &= E_{\xi}[(\partial_i \partial_j \ell + \partial_i \ell \partial_j \ell)(\partial_k \ell)] + E_{\xi}[(\partial_i \partial_j \ell)(\partial_k \ell)] \\ &= E_{\xi}[(2 \partial_i \partial_j \ell + \partial_i \ell \partial_j \ell)(\partial_k \ell)] \\ &= 2E_{\xi}\left[\left(\partial_i \partial_j \ell + \frac{1}{2} \partial_i \ell \partial_j \ell\right)(\partial_k \ell)\right] \\ &= 2\Gamma_{ij,k}^{(0)}. \end{aligned}$$

olarak elde edilir.

Sonuç. Üstel aile $\nabla^{(-1)}$ düz olması için gerek ve yeter koşul $\nabla^{(0)}$ düz olmasıdır (Calin ve Udrişte, 2014).

Önerme 4.3.5. $\mathcal{S} = \{p_{\xi}\}$ istatistiksel modeli üzerindeki herhangi bir X, Y, Z vektör alanı için

$$Zg(X, Y) = g(\nabla_Z^{(1)} X, Y) + g(X, \nabla_Z^{(-1)} Y)$$

dir (Calin ve Udrişte, 2014).

İspat. Bağlantıyı bir baz üzerinde kanıtlamak yeterlidir. $X = \partial_i$, $Y = \partial_j$ ve $Z = \partial_k$ seçildiğinde, göstermemiz gereken bağlantı

$$\partial_k g_{ij} = \Gamma_{ki,j}^{(1)} + \Gamma_{kj,i}^{(-1)}$$

dir. Böylece

$$\begin{aligned}
\Gamma_{ki,j}^{(1)} + \Gamma_{kj,i}^{(-1)} &= E_{\xi}[(\partial_k \partial_i \ell)(\partial_j \ell)] + E_{\xi}[(\partial_k \partial_j \ell + \partial_k \ell \partial_j \ell)(\partial_i \ell)] \\
&= E_{\xi}[(\partial_k \partial_i \ell)(\partial_j \ell)] + E_{\xi}[(\partial_k \partial_j \ell)(\partial_i \ell)] \\
&\quad + E_{\xi}[(\partial_i \ell)(\partial_j \ell)(\partial_k \ell)] \\
&= \partial_k g_{ij}(\xi)
\end{aligned}$$

olur. İspat tamamlanır.

4.3.4. $\nabla^{(\alpha)}$ Bağlantısı

$\nabla^{(-1)}$ ve $\nabla^{(1)}$, $\mathcal{S}$ üzerinde, sırasıyla karışım ailesinin ve üstel ailenin düz olduğu iki özel bağlantıdır. Bu bağlantıların ortasında Fisher metriğine göre Levi-Civita bağlantısı olan $\nabla^{(0)}$ bağlantısı vardır. İnterpolasyon yaparak, aşağıdaki 1-parametrelili bağlantı ailesi tanımlanır:

$$\nabla^{(\alpha)} = \frac{1+\alpha}{2} \nabla^{(1)} + \frac{1-\alpha}{2} \nabla^{(-1)}.$$

$\alpha = -1, 0, 1$ için sırasıyla $\nabla^{(-1)}, \nabla^{(0)}$ ve $\nabla^{(1)}$ bağlantılarını elde edilir. Fisher metriği g kullanılarak, bağlantı bileşenleri

$$\Gamma_{ij,k}^{(\alpha)} = g(\nabla_{\partial_i}^{(\alpha)} \partial_j, \partial_k)$$

aşağıdaki önerme ile verilmektedir.

Önerme 4.3.6. $\Gamma_{ij,k}^{(\alpha)}$ bileşenleri aşağıdaki gibi yazılabilir

$$\Gamma_{ij,k}^{(\alpha)} = E_{\xi} \left[\left(\partial_i \partial_j \ell + \frac{1-\alpha}{2} \partial_i \ell \partial_j \ell \right) \partial_k \ell \right].$$

İspat. Tanımdan

$$\begin{aligned}
\Gamma_{ij,k}^{(\alpha)} &= \frac{1+\alpha}{2} \Gamma_{ij,k}^{(1)} + \frac{1-\alpha}{2} \Gamma_{ij,k}^{(-1)} \\
&= \frac{1+\alpha}{2} E_{\xi}[(\partial_i \partial_j \ell)(\partial_k \ell)] + \frac{1-\alpha}{2} E_{\xi}[(\partial_i \partial_j \ell + \partial_i \ell \partial_j \ell)(\partial_k \ell)] \\
&= E_{\xi} \left[(\partial_i \partial_j \ell)(\partial_k \ell) + \frac{1-\alpha}{2} \partial_i \ell \partial_j \ell \partial_k \ell \right] \\
&= E_{\xi} \left[\left(\partial_i \partial_j \ell + \frac{1-\alpha}{2} \partial_i \ell \partial_j \ell \right) \partial_k \ell \right].
\end{aligned}$$

elde edilir.

$\nabla^{(\alpha)}$ simetriktir. Bu doğrudan $\Gamma_{ij,k}^{(\alpha)} = \Gamma_{ji,k}^{(\alpha)}$ bağlantısından çıkar. Ancak, $\nabla^{(\alpha)}$, $\alpha \neq 0$ için metrik değildir. Çift bağlantılar $\nabla^{(\alpha)}$ ve $\nabla^{(-\alpha)}$, $\nabla^{(1)}$ ve $\nabla^{(-1)}$ bağlantıları ile aynı dualite ilişkisindedir.

Önerme 4.3.7. $\mathcal{S} = \{p_{\xi}\}$ istatistiksel modeli üzerindeki herhangi bir X, Y, Z vektör alanı için

$$Zg(X, Y) = g(\nabla_Z^{(\alpha)} X, Y) + g(X, \nabla_Z^{(-\alpha)} Y)$$

dir (Calin ve Udrişte, 2014).

İspat. $\partial_k g_{ij} = \Gamma_{ki,j}^{(\alpha)} + \Gamma_{kj,i}^{(-\alpha)}$ olduğunu kanıtlamak yeterlidir. Tanımdan

$$\begin{aligned}\Gamma_{ki,j}^{(\alpha)} &= E_\xi \left[\left(\partial_k \partial_i \ell + \frac{1-\alpha}{2} \partial_k \ell \partial_i \ell \right) \partial_j \ell \right], \\ \Gamma_{kj,i}^{(-\alpha)} &= E_\xi \left[\left(\partial_k \partial_j \ell + \frac{1+\alpha}{2} \partial_k \ell \partial_j \ell \right) \partial_i \ell \right]\end{aligned}$$

olur. Toplamın alınması ve terimlerin azaltılmasıyla aşağıdaki sonuç elde edilir

$$\begin{aligned}\Gamma_{ki,j}^{(\alpha)} + \Gamma_{kj,i}^{(-\alpha)} &= E_\xi [(\partial_k \partial_i \ell)(\partial_j \ell)] + E_\xi [(\partial_k \partial_j \ell)(\partial_i \ell)] \\ &\quad + E_\xi [(\partial_i \ell)(\partial_j \ell)(\partial_k \ell)] \\ &= \partial_k g_{ij}(\xi).\end{aligned}$$

böylece ispat tamamlanır.

Önerme 4.3.8. Aşağıdaki ilişkiler geçerlidir

$$\begin{aligned}\nabla^{(\alpha)} &= (1 - \alpha)\nabla^{(0)} + \alpha\nabla^{(1)} \\ &= (1 + \alpha)\nabla^{(0)} - \alpha\nabla^{(-1)} \\ &= \nabla^{(0)} + \frac{\alpha}{2}(\nabla^{(1)} - \nabla^{(-1)}), \\ \nabla^{(0)} &= \frac{1}{2}(\nabla^{(-\alpha)} + \nabla^{(\alpha)}).\end{aligned}$$

Sonuç 4.3.2. $\mathcal{S}$ istatistiksel bir model olsun. O zaman aşağıdaki ifadeler eşdeğerdir:

- (i) $\mathcal{S}$, $\mathbb{R}$ içindeki tüm α için $\nabla^{(\alpha)}$ -düzdür,
- (ii) $\mathcal{S}$ hem $\nabla^{(1)}$ hem de $\nabla^{(-1)}$ -düzdür,
- (iii) $\mathcal{S}$ hem $\nabla^{(0)}$ hem de $\nabla^{(1)}$ -düzdür,
- (iv) $\mathcal{S}$ hem $\nabla^{(0)}$ hem de $\nabla^{(-1)}$ -düzdür.

Genel olarak, eğer bir $\mathcal{S}$ modeli iki farklı α bağlantısına göre düz ise, o zaman tüm α bağlantılarına göre düzdür.

4.4. Geodezik Denklemleri

Bu bölümde dağılım manifoldları üzerindeki geodeziklerin hesaplamaları verilecektir.

Önerme 4.4.1. $\mathcal{S}$ Üstel dağılımın istatistik modeli $(\mathcal{S}, g)$ 'nin geodezikleri c_1, c_2 sabitler olmak üzere

$$\xi(t) = e^{c_1 t + c_2}$$

biçimindedir (Branea ve Radulescu, 2023).

İspat. Geodezik diferansiyel denkleminin tanımından

$$\frac{d^2 \xi}{dt^2} + \Gamma_{11}^1 \frac{d\xi}{dt} \cdot \frac{d\xi}{dt} = 0 \Leftrightarrow \frac{d^2 \xi}{dt^2} - \frac{1}{\xi} \cdot \left(\frac{d\xi}{dt} \right)^2 = 0.$$

dir. Buradan

$$\xi'' - \frac{(\xi')^2}{\xi} = 0$$

diferansiyel denklemi ede edilir. $\xi (\xi \neq 0)$ ile bölünürse

$$\frac{\xi''}{\xi} - \frac{(\xi')^2}{\xi^2} = 0 \Leftrightarrow \frac{\xi'' \cdot \xi - \xi' \cdot \xi'}{\xi^2} = 0 \Leftrightarrow \left(\frac{\xi'}{\xi} \right)' = 0,$$

olup integral alınırsa

$$\begin{aligned} \frac{\xi'}{\xi} = c_1 &\Leftrightarrow \frac{d\xi}{dt} = c_1 \xi \Leftrightarrow \frac{d\xi}{\xi} = c_1 dt \Leftrightarrow \int \frac{d\xi}{\xi} = \int c_1 dt \\ &\Leftrightarrow \ln(\xi) = c_1 t + c_2 \Leftrightarrow \xi(t) = e^{c_1 t + c_2}, \quad c_1, c_2 \text{ sabitler,} \end{aligned}$$

elde edilir. Böylece ispat tamamlanır.

Önerme 4.4.2. Bernoulli dağılım modeli için jeodezikler şu şekilde verilir (Branea ve Radulescu, 2023) :

$$\xi(t) = \frac{1}{2}(1 + \sin(c_1 t + c_2)), \quad c_1, c_2 \in \mathbb{R}.$$

İspat. Geodezik diferansiyel denklemi tanımından

$$\frac{d^2 \xi}{dt^2} + \Gamma_{11}^1 \frac{d\xi}{dt} \cdot \frac{d\xi}{dt} = 0 \Leftrightarrow \frac{d^2 \xi}{dt^2} + \frac{2\xi - 1}{2\xi(1 - \xi)} \frac{d\xi}{dt} \cdot \frac{d\xi}{dt} = 0.$$

dir.

$$\frac{d\xi}{dt} = u \Rightarrow \frac{d^2 \xi}{dt^2} = \frac{du}{d\xi} \frac{d\xi}{dt} \Leftrightarrow \xi'' = u \frac{du}{d\xi}$$

dönüşümü yapılır ve yerlerine yazılırsa,

$$u \frac{du}{d\xi} + \frac{2\xi - 1}{2\xi(1 - \xi)} u^2 = 0$$

eşitliği elde edilir. İki durum söz konusu olur.

$$(1) \quad u = 0 \Leftrightarrow \frac{d\xi}{dt} = 0 \Leftrightarrow \xi(t) = c, \quad c \in \mathbb{R};$$

$$(2) \quad u \neq 0.$$

u 'nun türevi alınırsa,

$$\frac{du}{d\xi} = -\frac{2\xi - 1}{2\xi(1 - \xi)}u \Leftrightarrow \frac{du}{u} = \frac{1 - 2\xi}{2\xi(1 - \xi)}d\xi$$

bulunur. Her iki tarafın integrali alınırsa

$$\ln u = \ln (c\xi(1 - \xi))^{\frac{1}{2}}, c \in \mathbb{R}$$

olur. Buradan

$$u = (c\xi(1 - \xi))^{\frac{1}{2}} \Leftrightarrow \frac{d\xi}{dt} = (c\xi(1 - \xi))^{\frac{1}{2}} \Leftrightarrow \frac{d\xi}{\sqrt{\xi(1 - \xi)}} = c_1 dt, c_1 \in \mathbb{R}$$

bulunur. Her iki taraf da integre edilerek,

$$\begin{aligned} \arcsin(2\xi - 1) &= c_1 t + c_2, \Leftrightarrow \\ \xi(t) &= \frac{1}{2}(1 + \sin(c_1 t + c_2)), c_1, c_2 \in \mathbb{R} \end{aligned}$$

ispat tamamlanır.

4.5. Çarpıklık Tensörü (Skewness Tensor)

$\nabla^{(\alpha)}$ ve $\nabla^{(\beta)}$, $\mathcal{S}$ istatistiksel modeli üzerinde iki bağlantı olmak üzere (2,1)-tipinde genelleştirilmiş fark tensörü

$$K^{(\alpha, \beta)}(X, Y) = \nabla_X^{(\beta)} Y - \nabla_X^{(\alpha)} Y$$

ile tanımlanır.

Önerme 4.5.1. 3-kovaryant, tamamen simetrik bir T tensörü vardır, öyle ki herhangi bir farklı α, β için $g(K^{(\alpha, \beta)}(X, Y), Z) = \frac{\alpha - \beta}{2} T(X, Y, Z)$ dir, burada g Fisher metriğidir.

İspat. Fark tensör tanımı lokal olarak ifade edilirse,

$$\begin{aligned} \Gamma_{ij,k}^{(\beta)}(\xi) - \Gamma_{ij,k}^{(\alpha)}(\xi) &= E_\xi \left[\left(\partial_i \partial_j \ell + \frac{1 - \beta}{2} \partial_i \ell \partial_j \ell \right) \partial_k \ell \right] \\ &\quad - E_\xi \left[\left(\partial_i \partial_j \ell + \frac{1 - \alpha}{2} \partial_i \ell \partial_j \ell \right) \partial_k \ell \right] \\ &= \frac{\alpha - \beta}{2} E_\xi [\partial_i \ell \partial_j \ell \partial_k \ell], \end{aligned}$$

olur. $T_{ijk}(\xi) = E_\xi [\partial_i \ell \partial_j \ell \partial_k \ell]$ olarak alınırsa

$$\Gamma_{ij,k}^{(\beta)}(\xi) - \Gamma_{ij,k}^{(\alpha)}(\xi) = \frac{\alpha - \beta}{2} T_{ijk}(\xi)$$

yazılır. Bu aynı denklemin yerel koordinatlar versiyonudur. T 'nin i, j, k indislerinin herhangi birinde simetrik olduğu, yani tamamen simetrik olduğu görülmektedir. 3-kovaryant, simetrik tensör T

$$T(\partial_i, \partial_j, \partial_k) = T_{ijk} = E_\xi [\partial_i \ell \partial_j \ell \partial_k \ell]$$

çarpıklık tensörü olarak adlandırılır. Bu, log-likelihood fonksiyonunun varyasyonlarının üçüncü dereceden kümülanlarının beklentisini ölçer. İkinci dereceden kümülanların beklentisini ölçen Fisher metriği ile benzerliğine dikkat çekmek gerekir.

Örnek 4.5.1. Üstel bir aile durumunda

$$\begin{aligned} \Gamma_{ij,k}^{(\alpha)} &= E_\xi \left[\left(\partial_i \partial_j \ell + \frac{1-\alpha}{2} \partial_i \ell \partial_j \ell \right) \partial_k \ell \right] \\ &= E_\xi \left[\left(-\partial_i \partial_j \psi(\xi) + \frac{1-\alpha}{2} \partial_i \ell \partial_j \ell \right) \partial_k \ell \right] \\ &= -\partial_i \partial_j \psi(\xi) E_\xi [\partial_k \ell] + \frac{1-\alpha}{2} E_\xi [\partial_i \ell \partial_j \ell \partial_k \ell] \\ &= \frac{1-\alpha}{2} T_{ijk}. \end{aligned}$$

olur (Calin ve Udrişte, 2014).

Örnek 4.5.2. Bir karışım ailesi durumunda

$$\begin{aligned} \Gamma_{ij,k}^{(\alpha)} &= E_\xi \left[\left(\partial_i \partial_j \ell + \frac{1-\alpha}{2} \partial_i \ell \partial_j \ell \right) \partial_k \ell \right] \\ &= E_\xi \left[\left(-\partial_i \ell \partial_j \ell + \frac{1-\alpha}{2} \partial_i \ell \partial_j \ell \right) \partial_k \ell \right] \\ &= -E_\xi \left[\frac{1+\alpha}{2} \partial_i \ell \partial_j \ell \partial_k \ell \right] \\ &= -\frac{1+\alpha}{2} T_{ijk}. \end{aligned}$$

bulunur (Calin ve Udrişte, 2014).

Örnek 4.5.3. $p(x; \xi) = \xi e^{-\xi x}$, $\xi > 0$ ve $x \geq 0$ olsun. O zaman $\partial_\xi \ell = \frac{1}{\xi} - x$ ve çarpıklık tensörü

$$\begin{aligned} T_{111} &= E_\xi \left[(\partial_\xi \ell)^2 \right] = \int_0^\infty \left(\frac{1}{\xi} - x \right)^2 p(x; \xi) dx \\ &= \frac{1}{\xi^2} \int_0^\infty (1 - \xi x)^3 e^{-\xi x} dx = -\frac{2}{\xi^3} \end{aligned}$$

biçiminde verilir (Calin ve Udrişte, 2014).

Önerme 4.5.2. $\mathcal{S} = \{p_\xi(x); x \in \mathcal{X}, \xi \in \mathbb{E}\}$ istatistiksel modeli üzerindeki çarpıklık tensörü şu şekilde verilir (Calin ve Udrişte, 2014):

$$(i) \quad T_{ijk}(\xi) = 27 \int_{\mathcal{X}} \partial_i \left(p_\xi^{\frac{1}{3}}(x) \right) \partial_j \left(p_\xi^{\frac{1}{3}}(x) \right) \partial_k \left(p_\xi^{\frac{1}{3}}(x) \right) dx.$$

$$(ii) \quad T_{ijk}(\xi) = -E_{\xi}[\partial_i \partial_j \partial_k \ell] - E_{\xi}[(\partial_j \partial_k \ell)(\partial_i \ell)] - E_{\xi}[(\partial_k \partial_i \ell)(\partial_j \ell)] - E_{\xi}[(\partial_i \partial_j \ell)(\partial_k \ell)].$$

$$(iii) \quad T_{ijk}(\xi) = 2E_{\xi}[\partial_i \partial_j \partial_k \ell] + \partial_i g_{jk}(\xi) + \partial_j g_{ki}(\xi) + \partial_k g_{ij}(\xi).$$

İspat. (i) Tanım gereği

$$\begin{aligned} T_{ijk}(\xi) &= E_{\xi}[(\partial_i \ell)(\partial_j \ell)(\partial_k \ell)] \\ &= \int_{\mathcal{X}} (\partial_i \ln p(x)) (\partial_j \ln p(x)) (\partial_k \ln p(x)) p(x) dx \\ &= \int_{\mathcal{X}} \frac{\partial_i p(x)}{p(x)} \frac{\partial_j p(x)}{p(x)} \frac{\partial_k p(x)}{p(x)} p(x) dx \\ &= 3^3 \int_{\mathcal{X}} \partial_i \left(p(x)^{\frac{1}{3}} \right) \partial_j \left(p(x)^{\frac{1}{3}} \right) \partial_k \left(p(x)^{\frac{1}{3}} \right) dx. \end{aligned}$$

olarak elde edilir.

(ii) Fisher bilgisinin iki formülünü $g_{ij}(\xi) = E_{\xi}[\partial_i \ell \partial_j \ell]$ ve $g_{ij}(\xi) = -E_{\xi}[\partial_i \partial_j \ell]$ eşitleyerek

$$\int_{\mathcal{X}} \partial_i \ell \partial_j \ell p(x) dx = - \int_{\mathcal{X}} \partial_i \partial_j \ell p(x) dx$$

elde edilir. ξ^k 'ya göre türev alınır

$$\begin{aligned} &\int_{\mathcal{X}} \partial_i \partial_k \ell \partial_j \ell p(x) dx + \int_{\mathcal{X}} \partial_i \ell \partial_j \partial_k \ell p(x) dx + \int_{\mathcal{X}} \partial_i \ell \partial_j \ell \partial_k p(x) dx \\ &= - \int_{\mathcal{X}} \partial_i \partial_j \partial_k \ell p(x) dx - \int_{\mathcal{X}} \partial_i \partial_j \ell \partial_k p(x) dx \Leftrightarrow \end{aligned}$$

$$\begin{aligned} &E_{\xi}[(\partial_i \partial_k \ell)(\partial_j \ell)] + E_{\xi}[(\partial_i \ell)(\partial_j \partial_k \ell)] + \int_{\mathcal{X}} \partial_i \ell \partial_j \ell \partial_k p(x) dx \\ &= E_{\xi}[\partial_i \partial_j \partial_k \ell] - \int_{\mathcal{X}} \partial_i \partial_j \ell \partial_k p(x) dx \Leftrightarrow \\ &E_{\xi}[(\partial_i \partial_k \ell)(\partial_j \ell)] + E_{\xi}[(\partial_j \partial_k \ell)(\partial_i \ell)] + T_{ijk}(\xi) \\ &= -E_{\xi}[\partial_i \partial_j \partial_k \ell] - E_{\xi}[(\partial_i \partial_j \ell)(\partial_k \ell)] \end{aligned}$$

$T_{ijk}(\xi)$ için çözüldüğünde istenen formül elde edilir.

(iii) Önceki önermeden

$$\begin{aligned} T_{ijk}(\xi) &= E_{\xi}[(\partial_i \ell)(\partial_j \ell)(\partial_k \ell)] \\ &= \partial_k g_{ij}(\xi) - E_{\xi}[(\partial_k \partial_i \ell)(\partial_j \ell)] - E_{\xi}[(\partial_k \partial_j \ell)(\partial_i \ell)] \\ &= \partial_k g_{ij}(\xi) + \partial_j g_{ki}(\xi) + E_{\xi}[\partial_i \partial_j \partial_k \ell] + \partial_i g_{jk}(\xi) + E_{\xi}[\partial_i \partial_j \partial_k \ell] \\ &= 2E_{\xi}[\partial_i \partial_j \partial_k \ell] + \partial_i g_{jk}(\xi) + \partial_j g_{ki}(\xi) + \partial_k g_{ij}(\xi). \end{aligned}$$

bulunur.

Çarpıklık tensörü T 'yi tanımlamanın eşdeğer bir yolu $\beta = 0$ seçmektir:

$$\frac{\alpha}{2}T(X, Y, Z) = g\left(\nabla_X^{(0)}Y - \nabla_X^{(\alpha)}Y, Z\right).$$

Şimdiye kadar, çarpıklık tensörünün tüm tanımları iki bağlantı içeriyordu. Aşağıdaki sonuç bu tensörü sadece bir bağlantıdan ve Fisher metriğinden türetmektedir. Doğrusal bir ∇ bağlantısına göre 2-kovaryant bir h tensörünün aşağıdaki türev formülü

$$(\nabla h)(X, Y, Z) = Xh(Y, Z) - h(\nabla_X Y, Z) - h(Y, \nabla_X Z).$$

biçimindedir. ∇h 3-kovaryant bir tensördür (üç vektör alanına etki eder).

Teorem 4.5.1. Metrik g ve tensör T şu şekilde ilişkilidir $\alpha T = \nabla^{(\alpha)}g$ (Calin ve Udrişte, 2014).

İspat. $\mathcal{S}$ 'ye teğet olan herhangi bir X, Y, Z vektör alanı için

$$\alpha T(X, Y, Z) = Xg(Y, Z) - g\left(\nabla_X^{(\alpha)}Y, Z\right) - g\left(Y, \nabla_X^{(\alpha)}Z\right)$$

olduğunun gösterilmesi gerekir. Bu, yerel koordinatlarda şu şekilde yazılabilir:

$$\alpha T_{ijk} = \partial_i g_{jk} - \Gamma_{ij,k}^{(\alpha)} - \Gamma_{ik,j}^{(\alpha)}.$$

Tanım ve özelliklerden,

$$\begin{aligned} \partial_i g_{jk} - \Gamma_{ij,k}^{(\alpha)} - \Gamma_{ik,j}^{(\alpha)} &= E_\xi[(\partial_i \partial_j \ell)(\partial_k \ell)] \\ &\quad + E_\xi[(\partial_i \partial_k \ell)(\partial_j \ell)] + E_\xi[(\partial_i \ell)(\partial_j \ell)(\partial_k \ell)] \\ &\quad - E_\xi[(\partial_i \partial_j \ell) \partial_k \ell] - \frac{1-\alpha}{2} E_\xi[\partial_i \ell \partial_j \ell \partial_k \ell] \\ &\quad - E_\xi[(\partial_i \partial_k \ell) \partial_j \ell] - \frac{1-\alpha}{2} E_\xi[\partial_i \ell \partial_j \ell \partial_k \ell] \\ &= \alpha E_\xi[\partial_i \ell \partial_j \ell \partial_k \ell] \\ &= \alpha T_{ijk} \end{aligned}$$

olarak elde edilir.

Özellikle, $\alpha = 1$ için

$$T = \nabla^{(1)}g,$$

bulunur. Yani, çarpıklık tensörü Fisher metriği g 'nin $\nabla^{(1)}$ türevidir.

Sonuç. Herhangi bir $\epsilon > 0$ için, T tensörü şu şekilde verilir $T = \frac{\nabla^{(\alpha+\epsilon)}g - \nabla^{(\alpha)}g}{\epsilon}$ (Calin ve Udrişte, 2014).

4.6. Otoparalel Eğriler

$\mathbb{E} \subset \mathbb{R}^n$ Parametre uzayında düzgün bir eğri $\xi(s)$ olsun, yani $\xi(s) = (\xi^1(s), \dots, \xi^n(s))$ 'in her bir bileşeni düzgündür. İstatistiksel model $\mathcal{S} = \{p_\xi\}$ üzerindeki bir eğri şu şekilde tanımlanır:

$$\gamma(s) = \iota(\xi(s)) = p_{\xi(s)}, \quad s \in [0, T].$$

$\gamma(s)$ eğrisinin hızı şu şekilde verilir:

$$\dot{\gamma}(s) = \iota_* \left(\dot{\xi}(s) \right) = \frac{d}{ds} \iota \left(\xi(s) \right) = \frac{d}{ds} p_{\xi(s)}.$$

Bir $\gamma: [0, T] \rightarrow \mathcal{S}$ eğrisi, $\dot{\gamma}(s)$, $\gamma(s)$ boyunca paralel taşınıyorsa, yani $\nabla^{(\alpha)}$ bağlantısına göre ivme yok oluyorsa, $\nabla^{(\alpha)}$ -otoparalel olarak adlandırılır

$$\nabla_{\dot{\gamma}(s)}^{(\alpha)} \dot{\gamma}(s) = 0, \quad \forall s \in [0, T].$$

Yerel koordinatlarda $\gamma(s) = (\gamma^k(s))$, otoparalellik şu anlama gelir

$$\dot{\gamma}^k(s) + \Gamma_{ij}^{(\alpha)} \dot{\gamma}^i(s) \dot{\gamma}^j(s) = 0.$$

Burada $\Gamma_{ij}^{(\alpha)k} = \Gamma_{ij,l}^{(\alpha)} g^{lk}$, $\gamma(s)$ boyunca değerlendirilir ve g Fisher metriğini gösterir. Bu denklem bir Ricatti adi diferansiyel denklemdir. Bu denklem sistemi, $\dot{\gamma}(s)$ hız vektörünün $\nabla^{(\alpha)}$ bağlantısına göre kendisine paralel hareket ettiğini belirtir.

Eğer $\alpha = 0$ ise, otoparalel eğriler Fisher metriği g 'ye göre geodezik olurlar. Genel olarak, Ricatti denklem sistemi doğrusal değildir ve bu nedenle açık bir şekilde çözülmesi zordur. Bununla birlikte, açık çözümlerin mümkün olduğu birkaç özel durum vardır.

Örnek 4.6.1. Herhangi bir üstel aile modeli $\nabla^{(1)}$ -düz olduğundan, $\nabla^{(1)}$ -otoparalel eğriler $\ddot{\gamma}(s) = 0$ dir, yani $\gamma_k(s) = c_k s + \gamma_k(0)$ 'ı sağlar ve dolayısıyla düz çizgilerdir. $\Gamma_{ij,k}^{(\alpha)} = \frac{1-\alpha}{2} T_{ijk}$, olduğu kullanılarak $\nabla^{(\alpha)}$ -otoparalel eğrilerin denklemi, Fisher metriği g indisler düşürüldükten sonra

$$g_{rk} \dot{\gamma}^k(s) + \Gamma_{ij,r}^{(\alpha)} \dot{\gamma}^i(s) \dot{\gamma}^j(s) = 0 \Leftrightarrow g_{rk} \dot{\gamma}^k(s) + \frac{1-\alpha}{2} T_{ijr} \dot{\gamma}^i(s) \dot{\gamma}^j(s) = 0$$

olur.

Eğer X , $\gamma(s)$ boyunca bir vektör alanı ise, o zaman X^r ile çarpıp r üzerinden toplanırsa

$$g(\ddot{\gamma}, X) + \frac{1-\alpha}{2} T(\dot{\gamma}, \dot{\gamma}, X) = 0$$

olur. İki durum özellikle önemlidir.

Eğer $X = \dot{\gamma}(s)$ ise, o zaman

$$g(\ddot{\gamma}(s), \dot{\gamma}(s)) + \frac{1-\alpha}{2} T(\dot{\gamma}(s), \dot{\gamma}(s), \dot{\gamma}(s)) = 0$$

dir. $\alpha = 1$ için, $g(\ddot{\gamma}(s), \dot{\gamma}(s)) = 0$ elde edilir, yani hız ivmeye diktir. Bu durum geodezik eğriler için her zaman geçerlidir.

1. Eğer $X = \dot{\gamma}(s)$ ise, o zaman

$$\|\dot{\gamma}(s)\|_g^2 + \frac{1-\alpha}{2} T(\dot{\gamma}(s), \dot{\gamma}(s), \dot{\gamma}(s)) = 0$$

dir.

2. Eğer s, γ boyunca yay uzunluğunu gösteriyorsa, o zaman $\kappa(s) = \|\dot{\gamma}(s)\|$ terimi eğrilik yorumuna sahiptir. Dolayısıyla, önceki bağıntı eğrilik ve çarpıklık tensörü arasında bir ilişki sağlar.

Önerme 4.6.1. $\alpha \neq \beta$ olsun. Eğer bir $\gamma(s)$ eğrisi hem $\nabla^{(\alpha)}$ hem de $\nabla^{(\beta)}$ -otoparalel ise, o zaman $\gamma(s)$ eğrisi boyunca herhangi bir X vektör alanı için $T(\dot{\gamma}(s), \dot{\gamma}(s), X) = 0$ dır (Calin ve Udrişte, 2014).

İspat. $\nabla_{\dot{\gamma}(s)}^{(\alpha)} \dot{\gamma}(s) = 0$ ve $\nabla_{\dot{\gamma}(s)}^{(\beta)} \dot{\gamma}(s) = 0$ olduğundan, fark tensörü için

$$K^{(\alpha, \beta)}(\dot{\gamma}(s), \dot{\gamma}(s)) = 0$$

ve dolayısıyla $T(\dot{\gamma}(s), \dot{\gamma}(s), X) = 0$ olur.

4.7. Jeffrey Önceliği (Jeffrey's Prior)

Jeffreys önceliği Bayesçi istatistikte kullanılan nesnel (bilgi içermeyen) bir öncelik dağılımıdır. $\mathcal{S} = \{p_\xi; \xi \in \mathbb{E}\}$ istatistiksel bir model olsun ve $G(\xi) = \det g_{ij}(\xi)$ Fisher bilgi matrisinin determinantını gösterebilir. Ayrıca

$$\text{Vol}(\mathcal{S}) = \int_{\mathbb{E}} \sqrt{G(\xi)} d\xi < \infty$$

olduğu kabul edilsin. O zaman

$$Q(\xi) = \frac{1}{\text{Vol}(\mathcal{S})} \sqrt{G(\xi)}$$

ifadesi Jeffrey önceliği olarak adlandırılan $\mathbb{E}$ üzerinde bir olasılık dağılımı tanımlar. Bu dağılım parametreler uzayının yeniden parametreleri altında değişmezdir.

4.8. Bazı Özel Dağılımların Riemann Geometrisi

Bu bölümde bazı özel dağılımlar için hesaplamalar verilecektir.

4.8.1. Normal Dağılım

Normal dağılım Bilgi Geometrisi açısından önemli bir yere sahiptir. Amari, bu dağılımı 2-boyutlu bir manifold şeklinde alarak Bilgi Geometrisi çalışmalarını başlatmıştır. Normal dağılımın olasılık yoğunluk fonksiyonu daha önce verilmişti. Bu fonksiyon kullanılarak normal dağılımının metriği, Christoffel sembolleri ve geodezikleri verilecektir. Normal dağılım için Fisher bilgi metriği

$$g_{ij} = \begin{pmatrix} \frac{1}{\sigma^2} & 0 \\ 0 & \frac{2}{\sigma^2} \end{pmatrix}$$

biçimindedir. Levi-Civita koneksiyonu için birinci ve ikinci türden sıfır olmayan Christoffel sembolleri

$$\Gamma_{11,2} = \frac{1}{\sigma^3}, \quad \Gamma_{12,1} = -\frac{1}{\sigma^3}, \quad \Gamma_{22,2} = -\frac{2}{\sigma^3}$$
$$\Gamma_{ij}^1 = \begin{pmatrix} 0 & -\frac{1}{\sigma} \\ -\frac{1}{\sigma} & 0 \end{pmatrix}, \quad \Gamma_{1j}^2 = \begin{pmatrix} \frac{1}{2\sigma} & 0 \\ 0 & -\frac{1}{\sigma} \end{pmatrix}$$

şeklinde hesaplanır (Calin ve Udrişte, 2014). c bir reel sayı olmak üzere geodezikler $\sigma > 0$ için

$$\sigma(s)^2 + (\mu(s) - K)^2 \frac{E}{c^4} = c^2$$

biçimindeki yarım elipsler olacaktır. Burada, K integral sabitidir.

Diğer yandan, $c = 0$ durumunda, σ

$$\frac{\dot{\sigma}}{\sigma} = \sqrt{E} \Leftrightarrow \frac{d}{ds} \ln \sigma = \sqrt{E}$$

koşullarını sağlar. Buradan çözüm $\mu = K$ sabitken

$$\sigma(s) = \sigma(0)e^{\sqrt{E}s}$$

biçimindedir. Bu durumda geodezikler dikey yarım çizgilerdir (Detaylar için bkz. (Calin ve Udriște, 2014).)

Önerme 4.8.1. Ortalamaları eşit, $\mu_0 = \mu_1$ ve standart sapmaları farklı σ_0 ve σ_1 olan iki normal dağılım alınsın. O zaman ilk dağılımı ikincisine dönüştüren en küçük bilgi dönüşümü,

$$\sigma(s) = \sigma_1^{\frac{s}{\tau}} \sigma_0^{1-\frac{s}{\tau}}, \quad s \in [0, \tau].$$

sabit ortalama ve standart sapmaya sahip bir normal dağılımdır (Calin ve Udriște, 2014).

Diğer yandan α koneksiyonu için birinci türden Christoffel katsayıları

$$\begin{aligned} \Gamma_{11,1}^{(\alpha)} &= \Gamma_{21,2}^{(\alpha)} = \Gamma_{12,2}^{(\alpha)} = \Gamma_{22,1}^{(\alpha)} = 0 \\ \Gamma_{11,2}^{(\alpha)} &= \frac{1-\alpha}{\sigma^3}, \quad \Gamma_{12,1}^{(\alpha)} = \Gamma_{21,1}^{(\alpha)} = -\frac{1+\alpha}{\sigma^3}, \quad \Gamma_{22,2}^{(\alpha)} = -\frac{2(1+2\alpha)}{\sigma^3}. \end{aligned}$$

dir. Ayrıca ikinci türden Christoffel katsayıları,

$$\begin{aligned} \Gamma_{ij}^{1(\alpha)} &= g^{11}\Gamma_{ij,1}^{(\alpha)} + g^{12}\Gamma_{ij,2}^{(\alpha)} = \sigma^2\Gamma_{ij,1}^{(\alpha)} \\ &= \sigma^2 \begin{pmatrix} 0 & -\frac{1+\alpha}{\sigma^3} \\ -\frac{1+\alpha}{\sigma^3} & 0 \end{pmatrix} = \begin{pmatrix} 0 & -\frac{1+\alpha}{\sigma} \\ -\frac{1+\alpha}{\sigma} & 0 \end{pmatrix} \\ \Gamma_{ij}^{2(\alpha)} &= g^{21}\Gamma_{ij,1}^{(\alpha)} + g^{22}\Gamma_{ij,2}^{(\alpha)} \\ &= \frac{\sigma^2}{2} \begin{pmatrix} \frac{1-\alpha}{\sigma^3} & 0 \\ 0 & -2\frac{1+2\alpha}{\sigma^3} \end{pmatrix} = \begin{pmatrix} \frac{1-\alpha}{2\sigma} & 0 \\ 0 & -\frac{1+2\alpha}{\sigma} \end{pmatrix}. \end{aligned}$$

şeklindedir. Böylece α -otoparalel eğriler için Riccati denklemleri

$$\begin{aligned} \ddot{\mu} - \frac{2(1+\alpha)}{\sigma} \dot{\sigma} \dot{\mu} &= 0 \\ \ddot{\sigma} + \frac{1-\alpha}{2\sigma} \dot{\mu}^2 - \frac{1+2\alpha}{\sigma} \dot{\sigma}^2 &= 0 \end{aligned}$$

biçimindedir. Denklemin çözümü yapılırsa, μ -bileşeni şu şekilde verilir:

$$\mu = c \int \sigma^{2(1+\alpha)}(s) ds.$$

Bu denklemin açık bir şekilde çözülebildiği α 'nın birkaç özel değeri vardır.

1. Durum: $\alpha = -1$ için denklem şu hale gelir:

$$-v - K = (\pm s + s_0) \sqrt{\frac{1-\alpha}{2}} c.$$

Bu denklem çözülürse K sabiti için

$$\sigma^2(s) = C^2 - \left((\pm s + s_0) \sqrt{\frac{1-\alpha}{2}} c + K \right)^2$$

bulunur. Böylece $\mu(s) = cs + \mu(0)$ olur.

2. Durum: $\alpha = \frac{1}{2}$ için

$$\int \frac{dv}{(C^2 - v^2)^{\frac{3}{2}}} = \frac{v}{C^2 \sqrt{C^2 - v^2}}$$

denklemini çözülürse

$$-\frac{v}{C^2 \sqrt{C^2 - v^2}} = (\pm s + s_0) \frac{c}{2} + K$$

olup

$$\sigma(s) = \frac{C}{\sqrt{1 + C^4 \left((\pm s + s_0) \frac{c}{2} + K \right)^2}}.$$

elde edilir. Böylece μ -bileşeni

$$\mu(s) = c \int \sigma^3(s) ds$$

integrali ile verilir.

Normal dağılımın Jeffrey önceliği için

$$\mathcal{S}_\mu = \{p_\xi; E[p_\xi = \mu], \text{Var}[p_\xi] > 1\} = \{p_{(\mu, \sigma)}; \sigma > 1\}$$

istatistiksel modeli kullanılacaktır. Bu da üst-yarım düzlemde dikey bir yarım çizgiyi temsil eder (Calin ve Udrişte). Metriğin determinanı

$$G(\xi) = \det g_{ij}(\xi) = \det \begin{pmatrix} \frac{1}{\sigma^2} & 0 \\ 0 & \frac{2}{\sigma^2} \end{pmatrix} = \frac{2}{\sigma^4}$$

biçimindedir. Böylece hacim:

$$\text{Vol}(\mathcal{S}_\mu) = \int_1^\infty \sqrt{G(\xi)} d\sigma = \int_1^\infty \frac{\sqrt{2}}{\sigma^2} d\sigma = \sqrt{2} < \infty$$

biçimindedir. Bu nedenle $\mathcal{S}_\mu$ üzerindeki öncelik

$$Q(\sigma) = \frac{\sqrt{G(\sigma)}}{\text{Vol}(\mathcal{S}_\mu)} = \frac{1}{\sigma^2}$$

biçiminde olur.

4.8.2. Lognormal Dağılımı

Lognormal dağılımın olasılık yoğunluk fonksiyonu

$$p_{\mu,\sigma}(x) = \frac{1}{\sqrt{2\pi\sigma x}} e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}}, \quad x > 0,$$

biçiminde ifade edilir. Fisher bilgi matrisi (Fisher-Riemann metriği)

$$g = \begin{pmatrix} g_{\mu\mu} & g_{\mu\sigma} \\ g_{\sigma\mu} & g_{\sigma\sigma} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sigma^2} & 0 \\ 0 & \frac{2}{\sigma^2} \end{pmatrix}$$

şeklindedir. Bu bir normal dağılım modelinin Fisher metriği ile çakışır. Dolayısıyla, yukarıda elde edilen sonuçlar lognormal dağılım için de geçerlidir (Calin ve Udrişte, 2014).

4.8.3. Gamma Dağılımı

Gamma dağılımının $(\alpha, \beta) \in (0, \infty) \times (0, \infty), x \in (0, \infty)$ için istatistiksel modeli aşağıdaki yoğunluk ailesi ile tanımlanır:

$$p_\varepsilon(x) = p_{\alpha,\beta}(x) = \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\beta}}.$$

Bu modelin incelenmesinde bazı özel fonksiyonlara ihtiyaç vardır. Bu fonksiyonlar

$$\psi(\alpha) = \frac{\Gamma'(\alpha)}{\Gamma(\alpha)}, \quad \psi_1(\alpha) = \psi'(\alpha)$$

biçiminde olup sırasıyla digamma ve trigamma fonksiyonları olarak adlandırılır (Calin ve Udrişte, 2014). Dirichlet'in integral gösterimi ile

$$\psi(\alpha) = \int_0^\infty [e^{-t} - (1+t)^{-\alpha}] t^{-1} dt, \quad \alpha > 0$$

biçiminde olup trigamma fonksiyonu için

$$\psi_1(\alpha) = \psi'(\alpha) = \int_0^\infty \frac{\ln(1+t)}{t(1+t)^\alpha} dt$$

dir. Ayrıca trigamma fonksiyonu Hurwitz zeta fonksiyonu olarak

$$\psi_1(\alpha) = \zeta(2, \alpha) = \sum_{n \geq 0} \frac{1}{(\alpha + n)^2}$$

biçiminde ifade edilir. Bu bağıntı $\alpha \notin \{0, -1, -2, -3, \dots\}$ için geçerlidir ve $\alpha > 0$ olduğundan Bilgi Geometrisi açısından da geçerlidir. Diğer yandan

$$\int_0^{\infty} p_{\varepsilon}(x) dx = 1, \quad \int_0^{\infty} x p_{\varepsilon}(x) dx = \alpha\beta$$

bağıntıları geçerlidir. Fisher-Riemann metriğinin bileşenleri log-olabilirlik fonksiyonunun türevleri yardımıyla

$$\begin{aligned} g_{\alpha\alpha} &= -E[\partial_{\alpha}^2 \ell_x(\xi)] = \int_0^{\infty} \psi'(\alpha) p_{\xi}(x) dx = \psi'(\alpha) = \psi_1(\alpha), \\ g_{\beta\beta} &= -E[\partial_{\beta}^2 \ell_x(\xi)] = -\int_0^{\infty} \left(\frac{\alpha}{\beta^2} - \frac{2x}{\beta^3} \right) p_{\xi}(x) dx \\ &= -\frac{\alpha}{\beta^2} + \frac{2}{\beta^3} \int_0^{\infty} x p_{\xi}(x) dx = \frac{\alpha}{\beta^2}, \\ g_{\alpha\beta} &= -E[\partial_{\alpha\beta} \ell_x(\xi)] = \int_0^{\infty} \frac{1}{\beta} p_{\xi}(x) dx = \frac{1}{\beta} \end{aligned}$$

biçiminde olur. Böylece aşağıdaki sonuç verilir.

Önerme 4.8.2. Gamma dağılımı için Fisher bilgi matrisi (Fisher-Riemann metriği)

$$g = \begin{pmatrix} \psi_1(\alpha) & \frac{1}{\beta} \\ \frac{1}{\beta} & \frac{\alpha}{\beta^2} \end{pmatrix} = \begin{pmatrix} \sum_{n \geq 0} \frac{1}{(\alpha + n)^2} & \frac{1}{\beta} \\ \frac{1}{\beta} & \frac{\alpha}{\beta^2} \end{pmatrix}$$

dir (Calin ve Udriște, 2014). (Burada α gama dağılımının parametresidir ve α bağlantıları ile ilgisi yoktur.)

4.8.4. Bernoulli Dağılımı

Örnek uzayı $\mathcal{X} = \{0, 1, \dots, n\}$ ve parametre uzayı $\mathbb{E} = [0, 1]$ olmak üzere Bernoulli ya da binom dağılımı

$$p(k; \xi) = \binom{n}{k} \xi^k (1 - \xi)^{n-k}$$

ile verilir. Burada ξ parametresi başarı olasılığını ifade eder. Fisher bilgisi

$$\begin{aligned} g_{11}(\xi) &= -E_{\xi}[\partial_{\xi}^2 \ell(\xi)] = \sum_{k=0}^n p(k; \xi) \partial_{\xi}^2 \ell_k(\xi) \\ &= \sum_{k=0}^n \frac{k}{\xi^2} p(k; \xi) + \sum_{k=0}^n \frac{(n-k)}{(1-\xi)^2} p(k; \xi) \\ &= \frac{n}{\xi} + \frac{n(1-\xi)}{(1-\xi)^2} = \frac{n}{\xi(1-\xi)}, \end{aligned}$$

biçiminde hesaplanır. Varyansın $n\xi(1-\xi)$ olduğu kullanılarak,

$$g_{11}(\xi) = \frac{n^2}{\text{Var}(p_\xi)}$$

sonucu elde edilir. Bu da etkin bir tahmin ediciye karşılık gelen Cramér-Rao tipi bir özdeşliktir (Calin ve Udrişte, 2014).

4.8.5. Geometrik Olasılık Dağılımı

$\mathcal{X} = \{1, 2, 3, \dots\}$, $\mathbb{E} = [0, 1]$ ve $p(k; \xi) = (1 - \xi)^{k-1} \xi$, $k \in \mathcal{X}, \xi \in \mathbb{E}$ olsun. Log-olabilirlik fonksiyonunun kısmi türevleri kullanılarak Fisher bilgisi

$$\begin{aligned} g_{11}(\xi) &= -E_\xi[\partial_\xi^2 \ell(\xi)] \\ &= \sum_{k \geq 1} \frac{(k-1)p(k; \xi)}{(\xi-1)^2} + \sum_{k \geq 1} \frac{1}{\xi^2} p(k; \xi) \\ &= \frac{1}{\xi^2(1-\xi)}, \end{aligned}$$

şeklinde olur. Burada ortalama için $\sum_{k \geq 1} k p(k; \xi) = \frac{1}{\xi}$ ifadesi kullanılmıştır (Calin ve Udrişte, 2014).

4.8.6. Weibull Dağılımı

İsveçli fizikçi Wallodi Weibull, 1939 yılında malzemelerin çekme gerilmesi davranışlarını açıklamak amacıyla Weibull dağılımını geliştirmiştir (Oral, 2010). Bu dağılım, günümüzde özellikle rüzgar hızı verilerinin modellenmesinde yaygın olarak kullanılmaktadır. İki parametrelili Weibull dağılımı, rüzgar hızının istatistiksel özelliklerini temsil etmede en sık başvurulan yöntemlerden biridir. Bunun temel nedenleri arasında dağılımın rüzgar hızına oldukça iyi uyum sağlaması, esnek bir matematiksel yapıya sahip olması, az sayıda parametre ile tanımlanabilmesi, parametrelerin kolaylıkla belirlenebilmesi ve bir yükseklik için elde edilen parametrelerle farklı yüksekliklerdeki dağılımların da tahmin edilebilmesi gibi avantajlar yer almaktadır (Dokur ve Kurban, 2015). İki parametrelili Weibull dağılımına ait olasılık yoğunluk fonksiyonu şu şekilde tanımlanır:

$$p_\xi(x) = \frac{\xi_2}{\xi_1} \left(\frac{x}{\xi_1}\right)^{\xi_2-1} e^{-\left(\frac{x}{\xi_1}\right)^{\xi_2}}, \xi = (\xi_1, \xi_2)$$

Bu ifadede, x rüzgar hızını, ξ_2 şekil parametresini, ξ_1 ise ölçek parametresini temsil eder. $f(x)$ ise rüzgar hızına karşılık gelen olasılık yoğunluğunu ifade etmektedir. Bu dağılımın istatistik modeli

$$\mathcal{S} = \left\{ p_{\xi}(x) \left| p_{\xi}(x) = \left[\frac{\xi_2}{\xi_1} \left(\frac{x}{\xi_1} \right)^{\xi_2-1} \right] \exp \left\{ - \left(\frac{x}{\xi_1} \right)^{\xi_2} \right\}, x \in [0, +\infty), (\xi_1, \xi_2) \in \mathbb{R}^+ \times \mathbb{R}^+ \right\}$$

biçiminde verilebilir. Bu modelde (a,b) koordinat sisteminin rolünü oynar ve 2 boyutlu bir manifold yapısı olarak değerlendirilir. Dağılımın log olabilirlik fonksiyonu

$$\ln p_{\xi}(x) = -\xi_1^{-\xi_2} x^{\xi_2} + \ln \xi_2 + (\xi_2 - 1) \ln x - \xi_1$$

$l(x) = \ln p(x)$ olarak alınırsa için $\partial_i l = \frac{\partial}{\partial \xi_i} l(x), i = 1, 2$ olmak üzere

$$\begin{aligned} \partial_1 \partial_1 l &= -(\xi_2^2 + \xi_2) \xi_1^{-\xi_2-2} x^{\xi_2} + \frac{\xi_2}{\xi_1^2} \\ \partial_1 \partial_2 l &= \partial_2 \partial_1 l = x^{\xi_2} \xi_1^{-\xi_2-1} + \xi_2 \xi_1^{-\xi_2-1} x^{\xi_2} \ln x - \xi_2 \xi_1 \log \xi_1^{\xi_2} x \\ \partial_2 \partial_2 l &= -\xi_1^{-\xi_2} x^{\xi_2} \left(\ln \frac{x}{\xi_1} \right)^2 - \frac{1}{\xi_2^2} \end{aligned}$$

biçiminde elde edilir. Böylece doğrudan hesaplama ile aşağıdaki beklenen değerleri bulunur:

$$\begin{aligned} E[x^{\xi_2}] &= \xi_1^{\xi_2} \\ E[x^{\xi_2} \ln x] &= \frac{\xi_1^{\xi_2}}{\xi_2} (1 - \lambda) + \xi_1^{\xi_2} \ln \xi_1 \\ E \left[x^{\xi_2} \left(\ln \frac{x}{\xi_1} \right)^2 \right] &= \frac{\xi_1^{\xi_2}}{\xi_2^2} \left(\lambda^2 - 2\lambda + \frac{\pi^2}{6} \right). \end{aligned}$$

Burada λ Euler sabitidir ve $E(f)$ f 'nin beklentisi anlamına gelir. Ayrıca

$$\begin{aligned} E[\partial_1 \partial_1 l] &= -\frac{\xi_2^2}{\xi_1^2} \\ E[\partial_1 \partial_2 l] &= E[\partial_2 \partial_1 l] = \frac{1 - \lambda}{\xi_1} \\ E[\partial_2 \partial_2 l] &= -\frac{\left(\lambda^2 - 2\lambda + \frac{\pi^2}{6} + 1 \right)}{b^2} \end{aligned}$$

ifadelerine ulaşılır. Böylece Fisher bilgi matrisi ve tersi

$$(g_{ij}) = \begin{pmatrix} \frac{\xi_2^2}{\xi_1^2} & \frac{\lambda - 1}{\xi_1} \\ \frac{\lambda - 1}{\xi_1} & \frac{\lambda^2 - 2\lambda + \frac{\pi^2}{6} + 1}{\xi_2^2} \end{pmatrix}$$

ve

$$(g^{ij}) = \begin{pmatrix} \frac{6\xi_1^2 \left(\lambda^2 - 2\lambda + \frac{\pi^2}{6} + 1 \right)}{\pi^2 b^2} & \frac{6\xi_1(1-\lambda)}{\pi^2} \\ \frac{6\xi_1(1-\lambda)}{\pi^2} & \frac{6\xi_2^2}{\pi^2} \end{pmatrix}$$

olarak elde edilir.

Buradan aşağıdaki sonuca ulaşılır (Cao ve ark., 2014).

Önerme 4.8.3. Weibull dağılım manifoldunun yay uzunluğu elemanının karesi ve alan elemanı

$$ds^2 = \frac{\xi_2^2}{\xi_1^2} d\xi_1^2 + 2 \frac{\xi - 1}{a} d\xi_1 d\xi_2 + \frac{\lambda^2 - 2\lambda + \frac{\pi^2}{6} + 1}{\xi_2^2} d\xi_2^2$$

$$dA = \sqrt{\det(g_{ij})} d\xi_1 d\xi_2 = \frac{\pi}{\sqrt{6}\xi_1} d\xi_1 d\xi_2$$

dir (Cao ve ark., 2014).

Teorem 4.8.1. Weibull dağılım manifoldunun Ricci eğrilikleri ve Gauss eğriliği

$$R_{11} = -\frac{6\xi_2^2}{\pi^2 \xi_1^2}, R_{12} = R_{21} = \frac{6(1-\lambda)}{\pi^2 \xi_1}, R_{22} = -\frac{6 \left(\lambda^2 - 2\lambda + \frac{\pi^2}{6} + 1 \right)}{\pi^2 \xi_2^2}$$

$$K = -\frac{6}{\pi^2}$$

dir (Cao ve ark., 2014).

İspat. Christoffel sembolleri

$$\Gamma_{ij}^k = g^{kl} \left(\frac{\partial g_{li}}{\partial \theta_j} + \frac{\partial g_{lj}}{\partial \theta_i} - \frac{\partial g_{ij}}{\partial \theta_l} \right)$$

denklemleri kullanılarak

$$\Gamma_{11}^1 = \frac{6 \left(\lambda \xi_2 - \xi_2 - \frac{\pi^2}{6} \right)}{\pi^2 \xi_1}, \Gamma_{11}^2 = -\frac{\xi_2^3}{\pi^2 \xi_1^2}$$

$$\Gamma_{21}^1 = \Gamma_{12}^1 = \frac{6 \left(\lambda^2 - 2\lambda + \frac{\pi^2}{6} + 1 \right)}{\pi^2 \xi_2}, \Gamma_{21}^2 = \Gamma_{12}^2 = \frac{6b(1-\lambda)}{\pi^2 \xi_1}$$

$$\Gamma_{22}^1 = -\frac{6\xi_1(1-\lambda) \left(\lambda^2 - 2\lambda + \frac{\pi^2}{6} + 1 \right)}{\pi^2 \xi_2^3}, \Gamma_{22}^2 = -\frac{6 \left(\lambda^2 - 2\lambda + \frac{\pi^2}{6} + 1 \right)}{\pi^2 \xi_2}$$

olarak elde edilir. Diğer yandan manifoldun eğriliği

$$R_{ijkl} = (\partial_j \Gamma_{ik}^s - \partial_i \Gamma_{jk}^s) g_{sl} + (\Gamma_{jtl} \Gamma_{ik}^t - \Gamma_{itl} \Gamma_{jk}^t)$$

eşitliğinden

$$R_{1212} = -R_{2112} = -R_{1221} = R_{2121} = -\frac{1}{\xi_1^2}$$

olarak elde edilir. Böylece Ricci eğrilikleri

$$R_{11} = -\frac{6\xi_2^2}{\pi^2 \xi_1^2}, R_{12} = R_{21} = \frac{6(1-\lambda)}{\pi^2 \xi_1}, R_{22} = -\frac{6\left(\lambda^2 - 2\lambda + \frac{\pi^2}{6} + 1\right)}{\pi^2 \xi_2^2}$$

dir. Gauss eğriliği

$$K = \frac{R_{1212}}{\det(g_{ij})}$$

eşitliği kullanılarak

$$K = -\frac{6}{\pi^2}$$

biçiminde elde edilir.

5. RAYLEIGH DAĞILIMININ RIEMANN GEOMETRİSİ

Bu bölümde, tek parametrelili bir dağılım olan Rayleigh dağılımının Riemann geometrisi incelenecektir.

5.1. Rayleigh Dağılımı ve Uygulamaları

Rayleigh dağılımı, genellikle düzlemdeki bir vektörün büyüklüğü (genel uzunluğu) onun yatay ve dikey bileşenleriyle ilişkili olduğunda ortaya çıkar. Örneğin, rüzgar hızının iki boyutta modellenmesi ele alınsın. Rüzgarın yatay ve dikey bileşenleri sırasıyla X ve Y ile gösterilsin. Eğer bu bileşenler birbirinden bağımsız, sıfır ortalamalı ve eşit varyansa sahip normal dağılımlı rastgele değişkenlerse, yani

$$X \sim \mathcal{N}(0, \sigma^2), Y \sim \mathcal{N}(0, \sigma^2), X \perp Y$$

ise, bu durumda vektörün büyüklüğü olan

$$R = \sqrt{X^2 + Y^2}$$

rasgele değişkeni Rayleigh dağılımına uyar. R 'nin olasılık yoğunluk fonksiyonu

$$f_R(r) = \frac{r}{\sigma^2} e^{-\frac{r^2}{2\sigma^2}}, r \geq 0$$

şeklinde ifade edilir. Burada σ 'ya ölçekleme parametresi adı verilir. Dolayısıyla, iki bağımsız normal dağılımlı bileşene sahip bir vektörün normu Rayleigh dağılımına sahiptir. Bu dağılımın doğal bir örneği, rüzgar hızının iki boyutta analiz edilmesidir. Eğer rüzgar hızının yatay ve dikey bileşenlerinin sıfır ortalamalı, eşit varyanslı ve birbirinden bağımsız normal dağılımlı olduğu varsayılırsa, rüzgar hızının büyüklüğü (vektörün uzunluğu) Rayleigh dağılımını takip eder.

Rayleigh dağılımı üzerine birçok istatistiksel çalışma yapılmıştır. Bu çalışmalar, dağılımın parametrik kestirimi, genelleştirilmiş biçimleri ve çeşitli uygulamalardaki kullanımı gibi konuları kapsamaktadır. Örneğin, Kundu ve Raqab, Rayleigh dağılımının genelleştirilmiş halini tanımlanmış ve bu yeni modelin ömür verilerine uygulanabilirliğini incelenmiştir (Kundu ve Raqab, 2005). Özellikle mühendislik ve güvenilirlik analizi alanlarında, Rayleigh dağılımının parametrelerine yönelik maksimum olabilirlik (MLE), moment ve Bayesçi tahmin yöntemleri detaylı olarak araştırılmıştır (Basu, 1988; Ali ve Woo, 2002). Bunun yanı sıra, sinyal işleme, radar sistemleri ve kablosuz haberleşme gibi uygulamalı alanlarda Rayleigh dağılımı yaygın olarak kullanılmaktadır. Bu bağlamda, fading modelleri çerçevesinde dağılımın istatistiksel özellikleri ve performans ölçütleri analiz edilmektedir (Simon, ve Alouini, 2005).

5.2. Rayleigh Dağılımının Geometrik Modellenmesi

Rayleigh dağılımı için örnek uzay $\chi = [0, +\infty)$ ve parametrik uzay $\mathbb{E} = (0, \infty)$ dır. Olasılık yoğunluk fonksiyonu şu şekilde yazılabilir:

$$p(x; \xi) = \frac{x}{\xi^2} e^{-\frac{x^2}{2\xi^2}}$$

İlişkili istatistiksel model $\mathcal{S} = \{p(x; \xi)\}$ 1 –boyutludur ve bu nedenle dağılım geometrik olarak $P([0, \infty))$ 'de bir eğri gibi düşünülebilir.

$$P(\chi) = P([0, +\infty)) = \left\{ f \mid f: x \rightarrow \mathbb{R}, f \geq 0, \int_{\chi} f dx = 1 \right\}$$

olarak verilsin. O zaman

$$\begin{aligned} i: \mathbb{E} &\rightarrow P(\chi) \\ \xi &\rightarrow i(\xi) = p_{\xi} \end{aligned}$$

dönüşümü tanımlanabilir. $\left\{ \frac{\partial p_{\xi}}{\partial \xi} \right\}$ kümesinin lineer bağımsız olduğu açıktır. Dolayısıyla rank $i = 1$ olup i birebirdir. Böylece (I)-(II) koşulları sağlanır. Diğer yandan $\varphi_j(x, \xi) = \partial_j p_{\xi}(x)$ temel vektör alanları

$$\varphi_1(x, \xi) = \partial_{\xi} p_{\xi}(x) = \frac{x}{\xi^3} e^{-\frac{x^2}{2\xi^2}} \left(-2 + \frac{x^2}{\xi^2} \right)$$

biçimindedir. φ_j vektör alanı düzgün dönüşümdür, yani $t: \mathcal{S} \rightarrow P(\chi, \mathbb{R})$ için

$$\varphi_j(t) = \frac{\partial (t(p_{\xi}))}{\partial \xi^j}$$

olup $t, s \in C^{\infty}(\chi, \mathbb{R})$ için

$$\varphi_i(t \cdot s) = t \cdot \varphi_j(s) + \varphi_j(t) \cdot s$$

Leibniz kuralı sağlanır.

Model için log-olabilirlik fonksiyonu

$$\ell: \mathbb{E} \rightarrow P(\chi), \ell(p_{\xi})(x) = \ell_x(\xi) = \ln p_{\xi}$$

ile verilsin. O zaman

$$\ell_x(\xi) = \ln x - 2 \ln \xi - \frac{x^2}{2\xi^2}$$

olur ve bu ifadenin türevi

$$\partial_{\xi} \ell_x(\xi) = -\frac{2}{\xi} + \frac{x^2}{\xi^3}$$

dir.

Fisher bilgisi

$$I(\xi) = E_{\xi} \left[\partial_{\xi}^i \ell_{\xi}(x) \partial_{\xi}^j \ell_{\xi}(x) \right]$$

olacağından $g_{ij} = -E_{\xi} [\partial_{\xi}^3 \ln p_{\xi}]$ dir. Diğer yandan

$$\partial_{\xi}^2 \ln p_{\xi} = \partial_{\xi} \left(\frac{-2}{\xi} + \frac{1}{\xi^3} x^2 \right) = \frac{2}{\xi^2} - \frac{3}{\xi^4} x^2$$

olur. Böylece doğrudan hesaplama ile Riemann metriği

$$g_{ij} = \frac{4}{\xi^2}$$

biçiminde elde edilir. $\xi(t)$, parametre uzayında düzgün bir yol olsun. $\xi_0 = \xi(0)$ ve $\xi_1 = \xi(1)$ noktaları arasındaki geodezik uzaklık

$$S = \int_{\xi_0}^{\xi_1} \sqrt{g(\xi)} d\xi$$

dir. Böylece Rayleigh dağılımı için bu uzaklık

$$S = 2 \ln \left| \frac{\xi_1}{\xi_0} \right|$$

dir. Bu ifade iki paramtere arasındaki uzaklığı verir. Böylece,

$$\mathcal{S}_R = \left\{ p(x; \xi) = \frac{x}{\xi^2} e^{-\frac{x^2}{2\xi^2}} \mid \xi \in (0, \infty) \right\}$$

olmak üzere Rayleigh dağılım manifoldu $(\mathcal{S}_R, g)$ biçiminde tanımlanır.

Rayleigh dağılımı için elde edilen metrik bileşeni üstel dağılımda benzemektedir. Dağılımlar geometrik olarak da benzerdir (Şekil 5.1). Aşağıdaki teorem bu benzerliği vurgulamaktadır.

Teorem 5.2.1. $\mathcal{S}_E$ üstel dağılım manifoldunu göstermek üzere $\Phi: \mathcal{S}_R \rightarrow \mathcal{S}_E$

$$\Phi(\xi) = \lambda = \frac{1}{2\xi^2}$$

dönüşümü tanımlansın.

- i. Φ bir diffeomorfizmdir.
- ii. Fisher metriği $g_R(\xi) = 8\lambda$ eşitliğini sağlar. Burada λ yeni parametredir.
- iii. $(\mathcal{S}_R, g_R = \frac{4}{\xi^2})$ ve $(\mathcal{S}_E, g_E = \frac{1}{\lambda^2})$ manifoldları izometrik değildir ancak bir λ konformal ölçeği ile birbirine bağlıdır.

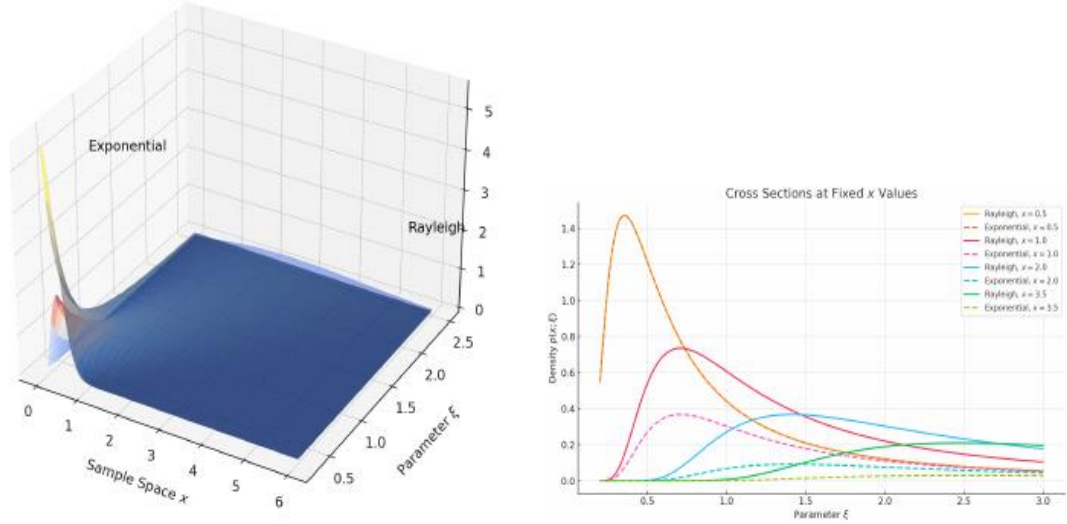
İspat. $Y = X^2$. olsun. O zaman rasgele değişkenlerin standart dönüşümü kullanılarak

$$f_Y(y) = f_X(\sqrt{y}) \frac{1}{2\sqrt{y}}, \quad y \geq 0$$

yazılabilir. Olasılık dağılım fonksiyonunun tanımını göz önüne alınırsa,

$$f_Y(y) = \frac{\sqrt{y}}{\xi^2} e^{-\frac{y}{2\xi^2}} \cdot \frac{1}{2\sqrt{y}} = \frac{1}{2\xi^2} e^{-\frac{y}{2\xi^2}}$$

bulunur. Böylece $\lambda = \frac{1}{2\xi^2}$ ölçek parametresi ile Y üstel dağılıma ait olur. Sonuç olarak $\xi = \frac{1}{\sqrt{2\lambda}}$ ilişkisi vardır. Bu ilişkiye dayanarak tanımlanan Φ dönüşümü bir diffeomorfizm olur. Ayrıca $g_R(\xi) = 8\lambda$ olduğu kolayca görülebilir.



Şekil 5.1. Üstel dağılım ile Rayleigh dağılımının 3 boyutlu uzayda modellenmesi ve kesitler.

5.3. Rayleigh Dağılım Manifoldunun Geodezikleri

Önerme 5.3.1. $\mathcal{S}_R$ Rayleigh dağılım manifoldunun birinci ve ikinci çeşit Christoffel sembolleri

$$\Gamma_{111} = -\frac{4}{\xi^3} \text{ ve } \Gamma_{11}^1 = -\frac{1}{\xi}$$

biçimindedir.

İspat. Birinci çeşit Christoffel sembolü için

$$\Gamma_{ijk} = \frac{1}{2} (\partial_j g_{ik} + \partial_k g_{ij} - \partial_i g_{jk})$$

denkleminde

$$\Gamma_{111} = \frac{1}{2} \cdot \frac{dg_{11}}{d\xi}$$

olur. Diğer yandan

$$\frac{dg_{11}}{d\xi} = \frac{d}{d\xi} \left(\frac{4}{\xi^2} \right) = -\frac{8}{\xi^3}$$

olacağından

$$\Gamma_{111} = \frac{1}{2} \cdot \left(-\frac{8}{\xi^3} \right) = -\frac{4}{\xi^3}$$

olarak bulunur. Ayrıca

$$\Gamma_{jk}^i = g^{il} \Gamma_{ljk}.$$

olup

$$\Gamma_{11}^1 = g^{11} \Gamma_{111},$$

biçimindedir. Metriğin tersi

$$g^{11} = (g_{11})^{-1} = \frac{\xi^2}{4},$$

olacağından $\Gamma_{11}^1 = -\frac{1}{\xi}$ olarak elde edilir.

Geodezik denklemlerinin

$$\frac{d^2 \xi}{dt^2} + \Gamma_{11}^1(\xi) \left(\frac{d\xi}{dt} \right)^2 = 0$$

biçiminde olduğu bilinmektedir. Rayleigh dağılımının ikinci çeşit Christoffel sembolü üstel dağılım ile aynı değer olarak elde edildi. Dolayısıyla üstel dağılımın geodeziklerinin bulunması ile aynı adımlar takip edilirse Rayleigh dağılımı için geodezikler (Şekil 5.2) aşağıdaki teorem ile verilir.

Teorem 5.3.1. Rayleigh dağılım manifoldunun $\xi(t)$ geodezik eğrisi $\xi(t) = \xi_0 e^{kt}$ biçiminde verilir. Burada $\xi_0 > 0$ başlangıç noktası ve $k \in \mathbb{R}$ başlangıçtaki hızı veren integrasyon sabitidir.

5.4. $\nabla^{(\alpha)}$ Koneksiyonuna Göre Christoffel Sembolleri

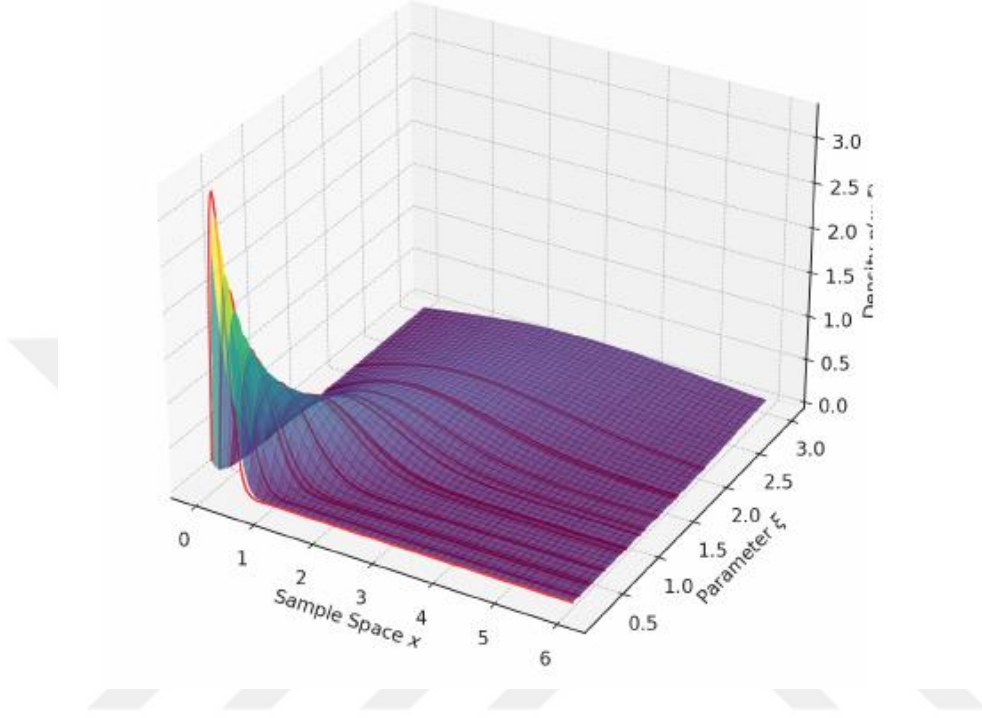
Log-olabilirlik fonksiyonu

$$\ell(x; \xi) = \ln x - 2 \ln \xi - \frac{x^2}{2\xi^2}$$

biçiminde olup ξ parametresine göre türevi alınır

$$\partial_{\xi} \ell = -\frac{2}{\xi} + \frac{x^2}{\xi^3}, \quad \partial_{\xi}^2 \ell = \frac{2}{\xi^2} - \frac{3x^2}{\xi^4}$$

ifadeleri elde edilir. Diğer yandan Rayleigh dağılımının $\mathbb{E}[X^2] = 2\xi^2$ ve $\mathbb{E}[X^4] = 8\xi^4$ momentleri dikkate alınırsa $\nabla^{(-1)}$ koneksiyonuna göre Christoffel sembolü



Şekil 5.2. Rayleigh dağılımı üzerinde geodeziklerin modellenmesi

$$\Gamma_{111}^{(1)}(\xi) = -\frac{12}{\xi^3}$$

olarak bulunur. Diğer yandan Rayleigh dağılımının üstel aile üyesi olarak

$$\theta = -\frac{1}{2\xi^2}$$

ile yeniden parametrelendirilmesiyle log-olabilirlik fonksiyonu

$$\ell(x; \theta) = \theta x^2 - \psi(\theta) + C(x)$$

biçiminde yeniden yazılabilir. Burada $\psi(\theta) = -\ln(-2\theta)$ ve $C(x) = \ln x$ dir. Böylece θ parametresine göre türevler

$$\partial_{\theta} \ell = x^2 - \psi'(\theta), \quad \partial_{\theta}^2 \ell = -\psi''(\theta)$$

biçiminde olur. Böylece Christoffel sembolleri

$$\Gamma_{111}^{(1)}(\theta) = E[(\partial_{\theta}^2 \ell)(\partial_{\theta} \ell)] = -\psi''(\theta)E[x^2 - \psi'(\theta)]$$

biçiminde elde edilir. $E[x^2] = \psi'(\theta)$ olacağından

$$\Gamma_{111}^{(1)}(\theta) = 0$$

olarak bulunur. Bu sonuç üstel aile hakkındaki genel durum ile uyumludur.

$\nabla^{(-1)}$ koneksiyonuna göre Christoffel sembolü

$$\Gamma_{111}^{(-1)} = E_{\xi} \left[\left(\partial_{\xi}^2 \ell + (\partial_{\xi} \ell)^2 \right) (\partial_{\xi} \ell) \right]$$

biçimindedir. Diğer yandan

$$\partial_{\xi}^2 \ell + (\partial_{\xi} \ell)^2 = \frac{6}{\xi^2} - \frac{7x^2}{\xi^4} + \frac{x^4}{\xi^6}$$

türevinden

$$\left(\partial_{\xi}^2 \ell + (\partial_{\xi} \ell)^2 \right) (\partial_{\xi} \ell) = \left(\frac{6}{\xi^2} - \frac{7x^2}{\xi^4} + \frac{x^4}{\xi^6} \right) \left(-\frac{2}{\xi} + \frac{x^2}{\xi^3} \right)$$

olarak bulunur. Beklenen değerin özelliklerinden $\mathbb{E}[X^2] = 2\xi^2$, $\mathbb{E}[X^4] = 8\xi^4$, ve $\mathbb{E}[X^6] = 48\xi^6$ olacağından

$$\Gamma_{111}^{(-1)} = \frac{4}{\xi^3}$$

olarak bulunur.

Diğer yandan $\nabla^{(-1)}$ için Christoffel sembolü

$$\Gamma_{111}^{(-1)}(\theta) = \mathbb{E}_{\theta} \left[\left(\partial_{\theta}^2 \ell + (\partial_{\theta} \ell)^2 \right) \partial_{\theta} \ell \right]$$

biçimindedir. $\partial_{\theta}^2 \ell + (\partial_{\theta} \ell)^2 = -\psi''(\theta) + (x^2 - \psi'(\theta))^2$ türev ifadesi dikkate alınır

$$\Gamma_{111}^{(-1)}(\theta) = \mathbb{E} \left[\left(-\psi''(\theta) + (x^2 - \psi'(\theta))^2 \right) (x^2 - \psi'(\theta)) \right]$$

bulunur. Ayrıca $\mathbb{E}[x^2] = \psi'(\theta)$, olduğundan

$$\mathbb{E}[x^2 - \psi'(\theta)] = 0$$

bulunur. Böylece Christoffel sembolü

$$\Gamma_{111}^{(-1)}(\theta) = \mathbb{E} \left[(x^2 - \psi'(\theta))^3 \right]$$

halini alır. Bu x^2 'nin üçüncü momentidir. x^2 , $\psi'(\theta) = -\frac{1}{\theta}$ ortalaması ile üstel dağılımı

aittir. Üçüncü momenti $\mu_3 = -\frac{2}{\theta^3}$ olduğundan

$$\Gamma_{111}^{(-1)}(\theta) = -\frac{2}{\theta^3}$$

biçiminde elde edilir.

Christoffel sembolleri arasındaki

$$\Gamma_{ijk}^{(\alpha)} = \frac{1+\alpha}{2} \Gamma_{ijk}^{(1)} + \frac{1-\alpha}{2} \Gamma_{ijk}^{(-1)}$$

bağıntısından aşağıdaki sonuç elde edilir.

Önerme 5.4.1. ξ parametresi ile parametrelendirilmiş Rayleigh dağılım manifoldu için $\nabla^{(\alpha)}$ koneksiyonunun $\Gamma_{111}^{(\alpha)}$ Christoffel sembolü

$$\Gamma_{111}^{(\alpha)} = \frac{-4(1 + 2\alpha)}{\xi^3}$$

biçimindedir.

Böylece, α sayısının özel değerleri

- $\alpha = -1$ için $\Gamma_{111}^{(-1)} = \frac{4}{\xi^3}$,
- $\alpha = 0$, için $\Gamma_{111}^{(0)} = -\frac{4}{\xi^3}$,
- $\alpha = 1$, için $\Gamma_{111}^{(1)} = -\frac{12}{\xi^3}$,

olur.

5.5. Çarpıklık Tensörü

Manifold 1-boyutlu olduğundan çarpıklık tensörünün bir tane bileşeni olacaktır. Bu bileşen

$$T_{111} = E_{\xi}[\partial_1 \ell \partial_1 \ell \partial_1 \ell]$$

ifadesinden,

$$T_{111} = -\frac{8}{\xi^3} \mathbb{E}[1] + \frac{12}{\xi^5} \mathbb{E}[X^2] - \frac{6}{\xi^7} \mathbb{E}[X^4] + \frac{1}{\xi^9} \mathbb{E}[X^6]$$

biçiminde gelir. Buradaki beklenen değerler

$$\mathbb{E}[X^2] = 2\xi^2, \mathbb{E}[X^4] = 8\xi^4, \mathbb{E}[X^6] = 48\xi^6$$

biçiminde olup yerlerine yazılırsa

$$T_{111} = -\frac{8}{\xi^3} + \frac{24}{\xi^3} - \frac{48}{\xi^3} + \frac{48}{\xi^3} = \frac{16}{\xi^3}$$

olarak bulunur. ξ pozitif değerler aldığından çarpıklık tensörü de pozitif değerler alır, ayrıca ξ küçüldükçe çarpıklığın arttığı tersine ξ büyüdükçe çarpıklığın azaldığı görülmektedir.

5.6. Laplace-Betrami Operatörü ve Sonlu Tipli Altmanifoldlar

Bir Riemann manifoldu (M, g) ile ilişkili Laplace-Beltrami operatörü şu şekilde tanımlanır:

$$\Delta f := -\frac{1}{\sqrt{\det g}} \sum_{i,j=1}^n \frac{\partial}{\partial x_i} \left(g^{ij} \sqrt{\det g} \frac{\partial f}{\partial x_j} \right).$$

Burada g^{ij} , g_{ij} metriğinin tersidir. Düzgün bir f fonksiyonu için $\Delta f = 0$ ise f 'ye harmonik fonksiyon denir.

Tanım 5.6.1. Kompakt bir $M \subset \mathbb{R}^m$ Riemann altmanifoldunun yer vektörü x Laplace-Beltrami operatörünün öz fonksiyonların sonlu toplamı şeklinde yazılabiliyorsa bu altmanifoldta sonlu tipten altmanifold denir (Chen, 1984).

Daha açık ifadeyle eğer

$$P(\Delta)H = \Delta^k H + c_1 \Delta^{k-1} H + \dots + c_k H = 0$$

biçiminde aşık olmaya bir P polinomu varsa M sonlu tipten bir altmanifolddur, burada H ortalama eğrilik vektörü olup $c_i \in \mathbb{R}$ dir. En küçük k değeri manifoldun sonlu tipini ifade eder.

Bu çalışmada, Rayleigh istatistiksel manifoldunun Öklidyen uzayının altmanifoldu olarak sonlu tipte olup olmadığı araştırılmıştır.

Önerme 5.6.1. Rayleigh istatistiksel manifoldu üzerinde tanımlanan diferansiyellenebilir bir $f: \mathcal{S}_R \rightarrow \mathbb{R}$ fonksiyonun Laplace-Beltrami operatörü

$$\Delta f = -\frac{\xi}{4} \left(\frac{\partial f}{\partial \xi} + \xi \frac{\partial^2 f}{\partial \xi^2} \right)$$

biçimindedir.

İspat. Rayleigh istatistiksel manifoldunun metrik bileşenlerinin

$$g(\xi) = \frac{4}{\xi^2}, \Rightarrow \det g = \frac{4}{\xi^2}, g^{\xi\xi} = \frac{\xi^2}{4}, \sqrt{\det g} = \frac{2}{\xi}$$

biçiminde olduğu bilinmektedir. Bu ifadeler Laplace-Beltrami operatörünün tanımında yerlerine yazılırsa

$$\Delta f = -\frac{\xi}{2} \frac{\partial}{\partial \xi} \left(\frac{2}{\xi} \cdot \frac{\xi^2}{4} \cdot \frac{\partial f}{\partial \xi} \right) = -\frac{\xi}{2} \frac{\partial}{\partial \xi} \left(\frac{\xi}{2} \cdot \frac{\partial f}{\partial \xi} \right) = -\frac{\xi}{2} \left(\frac{1}{2} \frac{\partial f}{\partial \xi} + \frac{\xi}{2} \frac{\partial^2 f}{\partial \xi^2} \right)$$

bulunur. Buradan

$$\Delta f = -\frac{\xi}{4} \left(\frac{\partial f}{\partial \xi} + \xi \frac{\partial^2 f}{\partial \xi^2} \right)$$

olarak elde edilir.

Bu önerme yardımıyla aşağıdaki teoreme ulaşılır.

Teorem 5.6.1. Rayleigh istatistiksel manifoldu $\mathbb{R}^2$ de 1-tipli bir altmanifolddur.

İspat. $x: \mathcal{S}_R \rightarrow \mathbb{R}^2$, $x(\xi) = (\cos(2 \ln \xi), \sin(2 \ln \xi))$ immersiyonu verilsin. İmmersiyonun birinci ve ikinci türevleri

$$\frac{\partial x}{\partial \xi} = \left(-\frac{2}{\xi} \sin(2 \ln \xi), \frac{2}{\xi} \cos(2 \ln \xi) \right)$$

$$\frac{\partial^2 x}{\partial \xi^2} = \left(\frac{2 \cos(2 \ln \xi) - 2 \sin(2 \ln \xi)}{\xi^2}, \frac{-2 \sin(2 \ln \xi) - 2 \cos(2 \ln \xi)}{\xi^2} \right)$$

biçiminde elde edilir. Bu ifadeler ve metrik değerleri yerlerine yazılırsa,

$$\Delta x = -\frac{1}{\sqrt{\det g}} \frac{d}{d\xi} \left(\sqrt{\det g} \cdot g^{\xi\xi} \cdot \frac{dx}{d\xi} \right) = -\frac{\xi}{2} \frac{d}{d\xi} \left(\frac{\xi}{2} \cdot \frac{dx}{d\xi} \right) = x(\xi)$$

eşitliği elde edilir. Ayrıca, $\Delta x = -nH$ eşitliğinden dolayı $n = 1 \Rightarrow H = -\Delta x = -x(\xi)$ olacaktır. Böylece

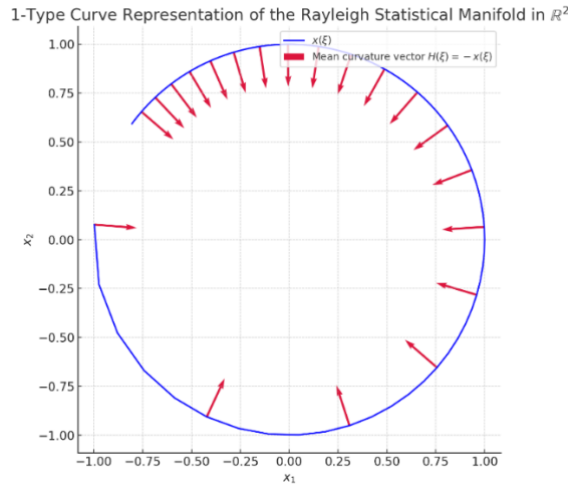
$$\Delta H = (-2 \cos(2 \ln \xi), -2 \sin(2 \ln \xi))$$

olarak bulunur. Sonuç olarak ortalama eğrilik vektörü H

$$-4\Delta H - H = 0$$

eşitliğini sağlar. Böylece (Chen,1984)’de Teorem 2.2 den $x(\xi)$ eğrisi 1-tipli bir eğri olur (Şekil 5.3).

Sonuç. Rayleigh istatistiksel manifoldunun Laplace-Beltrami operatörünün öz değeri 1 dir.



Şekil 5.3. $\mathbb{R}^2$ içinde gömülmüş . Rayleigh istatistiksel manifoldunun 1-tipli altmanifold olarak tasviri

İmmersiye (gömme) fonksiyonundaki logaritmik argüman nedeniyle, eğri $\mathbb{R}^2$ içinde spiral benzeri bir yapı sergiler ve ξ arttıkça birbirini izleyen dönüşler giderek birbirine yaklaşır. Şekilde mavi spiral eğri, gömülmenin logaritmik salınımını yansıtan $x(\xi)$

daldırmasının görüntüsünü göstermektedir. Kırmızı oklar ortalama eğrilik vektör alanını H göstermektedir.

5.7. Weibull Manifoldunun Bir Alt Manifoldu Olarak Rayleigh Manifoldu

$k = 2$, $\lambda = \sqrt{2}\xi$, seçimi ile Rayleigh manifoldu Weibull manifoldunun 1-boyutlu bir altmanifoldu olarak değerlendirilebilir. Bu durumda Rayleigh manifoldu

$$\mathcal{M}_R := \{(\lambda, k) \in \mathcal{M}_W \mid k = 2\}$$

biçiminde bir gömülü manifolddur (Şekil 5.4). Böylece bu manifoldun eğriliği sıfır olur, yani $\text{Riem}(\mathcal{S}_R) = 0$ dır. Diğer yandan manifoldun Weibull manifoldda gömülmesi dış eğriliğin araştırılmasını olanaklı kılar. Özel olarak ikinci temel form, şekil operatörü ve normal eğrilik gibi kavramlar ile manifold analiz edilebilir. Bu büyüklükler, Rayleigh modelinin içsel geometrisinin ötesinde, ortam Weibull modelinde geometrik olarak nasıl konumlandığına dair fikir vermektedir.

$\varphi: \mathbb{R}^+ \rightarrow \mathcal{M}_W$, düzgün immersiyonu

$$\varphi(\xi) = (\xi_1 = \sqrt{2}\xi, \xi_2 = 2)$$

ile tanımlansın. Bu durumda birinci ve ikinci türevleri

$$\frac{d\varphi}{d\xi} = (\sqrt{2}, 0), \quad \frac{d^2\varphi}{d\xi^2} = (0, 0)$$

şeklindedir. Böylece ikinci temel form için N eğri boyunca normal vektör olmak üzere

$$\Pi(X, X) = \left\langle \frac{d^2\varphi}{d\xi^2}, N \right\rangle = 0$$

eşitliği elde edilir. Buradan şu sonuca ulaşılır.

Teorem 5.7.1. Rayleigh istatistiksel manifoldu, Weibull manifoldunun total geodezik altmanifoldudur.

5.8. Uygulamalar

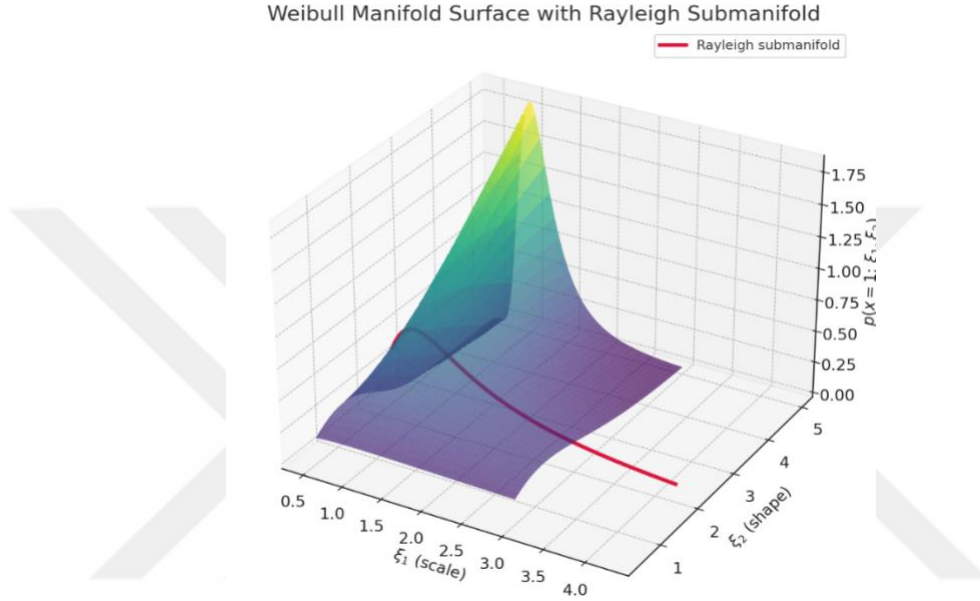
Bu bölümde Rayleigh manifoldu kullanılarak bilgi geometrisinin bazı uygulamaları ele alınacaktır.

5.8.1. Hessian Denklemi ve Potansiyel Fonksiyonlar

Rayleigh manifoldu üzerinde tanımlı düzgün bir $f: \mathcal{S}_R \rightarrow \mathbb{R}$ fonksiyonu verilsin. $Hess(f) = 0$, Hessian denklemini sağlayan f fonksiyonlarına potansiyel fonksiyonlar denir.

Rayleigh manifoldunun potansiyel fonksiyonları için aşağıdaki sonuç elde edilmiştir.

Teorem 5.8.1. C, D reel sabitler olmak üzere Rayleigh manifoldunun potansiyel fonksiyonları $f(\xi) = C \ln \xi + D$ biçimindedir.



Şekil 5.4. Weibull manifoldu üzerinde Rayleigh manifoldu

İspat. Fisher metriğinin bileşenleri ile Christoffel sembolü kullanılarak Hessian denklemi

$$\nabla_{\partial_\xi} \nabla_{\partial_\xi} f = \partial_\xi^2 f - \Gamma_{11}^1 \partial_\xi f = 0$$

biçiminde yazılır. Bu diferansiyel denklem

$$f''(\xi) + \frac{1}{\xi} f'(\xi) = 0$$

biçiminde olup $u(\xi) = f'(\xi)$ dönüşümü uygulanırsa

$$u'(\xi) + \frac{1}{\xi} u(\xi) = 0$$

adi diferansiyel denklemi elde edilir. $\mu(\xi) = \xi$ integral çarpanı uygulanırsa

$$\partial_\xi (\xi u(\xi)) = 0$$

olup, $C \in \mathbb{R}$ için

$$\xi u(\xi) = C$$

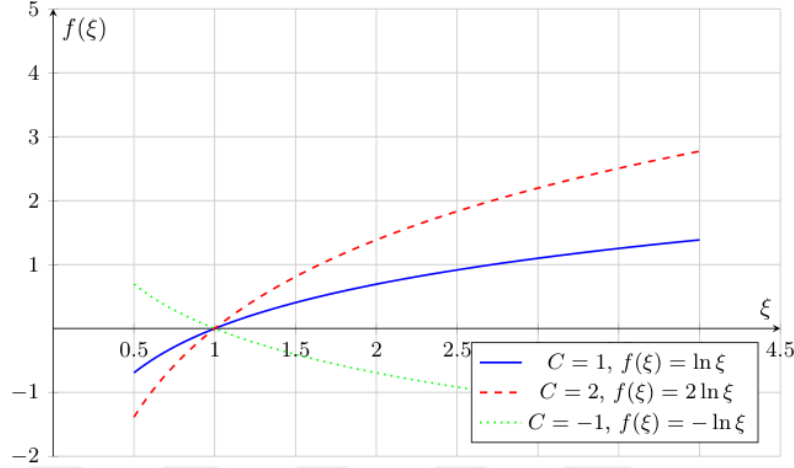
dir. Böylece

$$u(\xi) = \frac{C}{\xi}$$

olacağından

$$f(\xi) = C \ln \xi + D$$

fonksiyon ailesi elde edilir (Şekil 5.5).



Şekil 5.5. Rayleigh dağılımının potansiyel fonksiyonları

Rayleigh manifoldunun üzerinden $\nabla^{(\alpha)}$ koneksiyonunun Christoffel sembolleri

$$\Gamma_{111}^{(\alpha)} = \frac{-4(1+2\alpha)}{\xi^3}.$$

biçiminde olduğundan $f(\xi) = \ln \xi$ fonksiyonu için Hessian denklemi

$$f''(\xi) - \Gamma_{111}^{(\alpha)} f'(\xi) = -\frac{1}{\xi^2} + \frac{4(1+2\alpha)}{\xi^4}$$

şeklini alır. Bu ifade genel olarak sıfır değildir. Bu sonuç, $f(\xi) = \ln \xi$ fonksiyonun yalnızca $\nabla^{(0)}$ Levi-Civita koneksiyonu altında afin kaldığını göstermektedir. Bu fonksiyon Rayleigh istatistiksel manifoldunda önemli bir rol oynamaktadır. Fisher metriği tarafından indüklenen afin düz yapıyı yansıtan Levi-Civita bağlantısı $\nabla^{(0)}$ altında Hessian denklemini çözer. Keyfi α için genel $\nabla^{(\alpha)}$ bağlantıları altında afin olmamasına rağmen, $\nabla^{(1)}$ ve $\nabla^{(-1)}$ geometrileri arasında köprü kurarak karışım ve üstel aileler arasındaki dualistik yapı ile ilgili potansiyel bir fonksiyon olarak hareket eder.

5.8.2. Rayleigh Dağılımının Entropisi

Entropi, bir sistemin düzensizliğini, belirsizliğini veya bilgi eksikliğini nicel olarak ifade eden temel bir kavramdır. Bilgi geometrisinde, bir parametrik olasılık dağılım ailesi $\mathcal{S} = \{p(x; \theta)\}$ ele alındığında, entropi fonksiyonu şu şekilde tanımlanır:

$$H(\theta) = - \int p(x; \theta) \log p(x; \theta) dx$$

Bu fonksiyon, dual geometrik yapıların kurulmasında temel rol oynar. üstel aile için, entropi konveks bir potansiyel fonksiyondur ve affine bağlantılarla uyumlu bir geometri oluşturur. Entropi minimum olduğunda, dağılım en "keskin" bilgiyi içerir. Fisher bilgi maksimum olduğunda, parametre hakkında gözlemler en fazla bilgiyi taşır. Bu nedenle entropi ve Fisher bilgi metriği genellikle birbirini tamamlayan karakteristikler olarak görülür. Bilgi geometrisinde geodezikler, olasılık dağılımları arasındaki en kısa yolları temsil eder. Entropi, bu yollar boyunca bilgi kaybını veya artışını izlemek için bir araç olarak kullanılabilir. Örneğin, minimum entropi yolları, bilgi açısından en ekonomik geçişleri gösterir. Bu bölümde Rayleigh dağılımının entropisi hesaplanacaktır.

Teorem 5.8.2. Rayleigh dağılımının entropisi

$$H(\xi) = 1 + \ln \frac{\xi}{\sqrt{2}} + \frac{\gamma}{2}$$

dir. Burada, γ Euler sabitidir.

İspat. Entropinin tanımı gereği

$$H(\xi) = - \int_0^\infty f(x) \left(\ln x - 2 \ln \xi - \frac{x^2}{2\xi^2} \right) dx$$

yazılabilir. Buradan,

$$H(\xi) = - \left(\int_0^\infty f(x) \ln x dx - 2 \ln \xi \int_0^\infty f(x) dx - \frac{1}{2\xi^2} \int_0^\infty x^2 f(x) dx \right)$$

olur. İlk parça $\ln x$ in beklenen değeri olup $\ln \xi + \frac{\gamma}{2}$ dir, ikinci parça olasılık yoğunluk fonksiyonu olmasından dolayı 1 dir ve son parça $2\xi^2$ olarak hesaplanır. Bu ifadeler yerlerine yazılırsa

$$H(\xi) = 1 + \ln \frac{\xi}{\sqrt{2}} + \frac{\gamma}{2}$$

olarak elde edilir.

Üstel dağılıma benzer şekilde, Rayleigh dağılımının entropisi de parametrelendirme seçimine bağlı olarak değişir. Özellikle, $\lambda = \frac{1}{2\xi^2}$ dönüşümü altında, entropi fonksiyonu λ cinsinden ifade edildiğinde farklı bir biçim alacaktır. Bu durum, olasılık dağılımına özgü olmakla birlikte entropinin istatistiksel modelin yeniden parametризasyonları altında değişmez olmadığı genel ilkesini göstermektedir.

$H(\xi)$ 'nin kritik noktalarını bulmak için, ξ 'ye göre türevini alınırsa $\frac{dH}{d\xi} = \frac{1}{\xi}$ olur. Her $\xi > 0$ için $\frac{1}{\xi} > 0$ olduğundan, $H(\xi)$ 'nin türevi tamamen pozitifdir ve ξ ye kesinlikle artandır. Sonuç olarak, $H(\xi)$, $(0, \infty)$ içinde herhangi bir yerel kritik noktaya sahip değildir. Entropi ξ ile monoton olarak artar ve ξ büyüdükçe daha büyük değerlere ulaşır.

5.8.3. Sinyal İşleme Alanında Bazı Uygulamalar

Rayleigh dağılımı bir çok alanda farklı uygulamalara sahiptir. Bu bölümde sinyal işleme alanındaki bazı uygulamalar Bilgi Geometrisi bağlamında ele alınacaktır.

Geodezik Yollarla Kanal İzleme

İki Rayleigh dağılımı arasındaki geodezik uzaklığı veren metrik hiperbolik tipte bir metriğe indirgenir:

$$d(\sigma_1, \sigma_2) = 2 \left| \ln \left(\frac{\sigma_2}{\sigma_1} \right) \right|$$

Bu geometrik formülasyon, farklı sinyal ortamlarını karşılaştırmak için doğal ve anlaşılır bir yol sağlar. Pratik açıdan, bir kablosuz iletişim senaryosundaki farklı bölgeler (örneğin, açık alan, kentsel ortam, iç mekan) farklı σ değerlerine karşılık gelebilir. Rayleigh manifoldu üzerindeki bu noktalar arasındaki geodezik yolun izlenmesi, kanalın nasıl geliştiğini anlamak için nicel bir araç sağlar.

Üç farklı ortamda hareket eden bir mobil cihaz ele alınsın. Bu ortamlar için

- A noktası (açık alan): $\sigma_1 = 1.0$
- B Noktası (yarı kentsel): $\sigma_2 = 1.5$
- C Noktası (kapalı alan): $\sigma_3 = 2.2$

bilgisi verilsin.

Fisher bilgi metriği aracılığıyla hesaplanan jeodezik mesafeler şunlardır:

$$\begin{aligned} d_{AB} &\approx 0.81 \\ d_{BC} &\approx 0.63 \\ d_{AC} &\approx 1.57. \end{aligned}$$

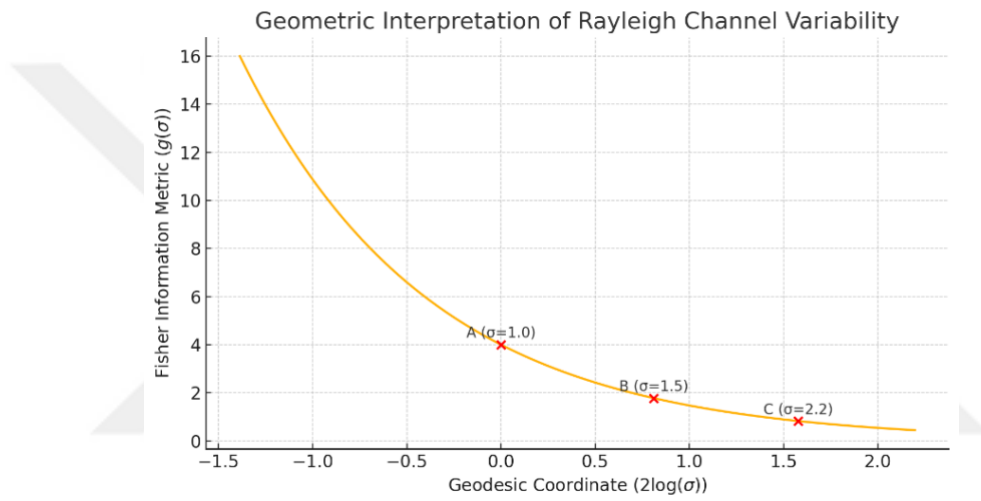
Bu değerler, farklı kanal durumları arasında geçiş yapmak için gereken istatistiksel "eforu" veya değişikliği yansıtır ve uyarlanabilir modülasyon, kanal tahmini veya sinyal bozulması için erken uyarı sistemlerini bilgilendirebilir.

Şekil 5.6'da gösterildiği gibi, her ortam Rayleigh bilgi manifoldunda bir noktaya karşılık gelen farklı bir σ değerine neden olur. Bu noktalar arasındaki jeodezik mesafeler

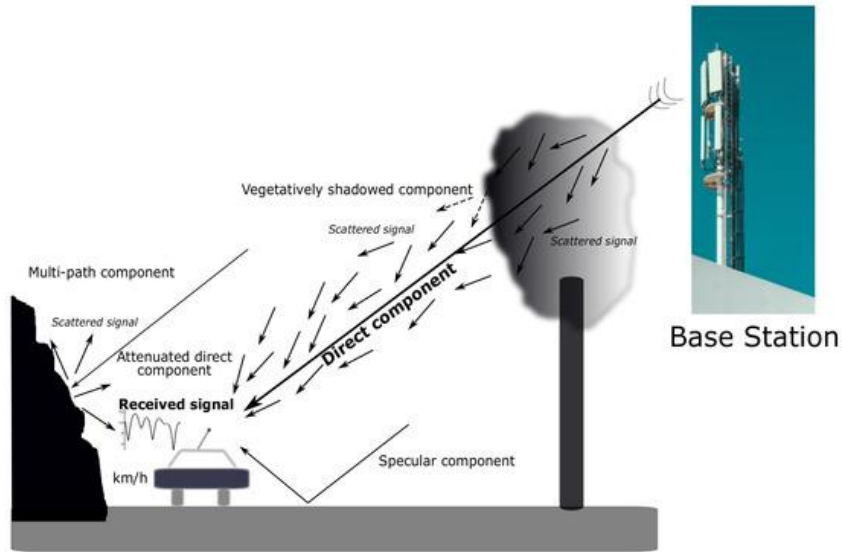
kablosuz kanalın istatistiksel evrimini ölçer. Eğri, parametre uzayı üzerindeki Fisher bilgi metriğini temsil eder ve kırmızı noktalar, farklı ölçek parametreleri σ ile karakterize edilen farklı kanal durumlarına (A, B, C) karşılık gelir.

Sinyal Sönümlenmesi (Signal Fading)

Rayleigh dağılımı kablosuz iletişim sistemlerinde, özellikle de çok yollu sönümlenme koşulları altında alınan sinyallerin büyüklüğünü modellerken sıklıkla karşımıza çıkar. Bu tür senaryolarda alınan sinyal her biri bağımsız ve aynı dağılımlı Gauss dağılımını izleyen iki ortogonal bileşenin sonuç vektörü olarak düşünülebilir. Dolayısıyla bu vektörünün büyüklüğü Rayleigh dağılımına aittir (Şekil 5.7) .



Şekil 5.6. Rayleigh kanal değişkenliğinin geometrik görselleştirmesi.



Şekil 5.7. Sönümlenme bileşenlerinin modellenmesi (Gomez, 2020)

Şekil 5.7'de sabit bir kaynak (yer baz istasyonundaki verici) ve birkaç sinyal bileşeninin dahil olduğu mobil bir alıcı (araç) için basitleştirilmiş bir sönümlenme modeli gösterilmektedir. Bu bileşen çok az yönlülük gösterir ve büyüklüğünün genellikle Rayleigh dağılımlı olduğu varsayılırken fazının düzgün dağıldığı varsayılır. Bu uygulamada çeşitli fiziksel konfigürasyonlar altında bir Rayleigh sönümlenme kanalından gerçek dünya ölçümlerinin analizi yapılacaktır. Veri seti (Palczynska ve Noga 2021) den alınmıştır. Veri kümesi, iletişim kanalındaki çok yollu sönümlenmenin modellenmesine ilişkin araştırmasının bir parçası olan simülasyonların sonuçlarını içermektedir. Rayleigh sönümlenme zarfı, LabVIEW programlama ortamında Monte-Carlo simülasyonu (MCS) kullanılarak oluşturulmuştur. Alınan veriler Python programlama dili kullanılarak analiz edilmiştir. Veri kümesinde 16 dosya bulunmaktadır. İlk olarak her bir dosya için tahmini ξ parametresi hesaplanmıştır. Parametre tahmini için

$$\hat{\xi} = \sqrt{\frac{1}{2} E[X^2]}$$

eşitliğinden yararlanılmıştır. Alınan veriler Python programlama dili kullanılarak analiz edilmiştir. Böylece geometrik uzaklığa dayanarak tahmin edilen parametreler her bir dosya için tablodaki şekilde elde edilmiştir.

Uzaklık (m)	ξ	Uzaklık (m)	ξ
0.25	0.501	1.33	1.156
0.33	0.579	1.50	1.227
0.50	0.706	1.67	1.292
0.67	0.817	1.75	1.319
0.75	0.864	2.00	1.414
1.00	1.002	2.25	1.500
1.25	1.119	2.50	1.577
		2.75	1.657
		3.00	1.731

Tablo 5.1. Geometrik uzaklığa dayanarak tahmin edilen parametreler

Diğer yandan bir başlangıç noktası ξ_0 alınarak ve geodezik mesafe kullanılarak Tablo 5.2 elde edilir:

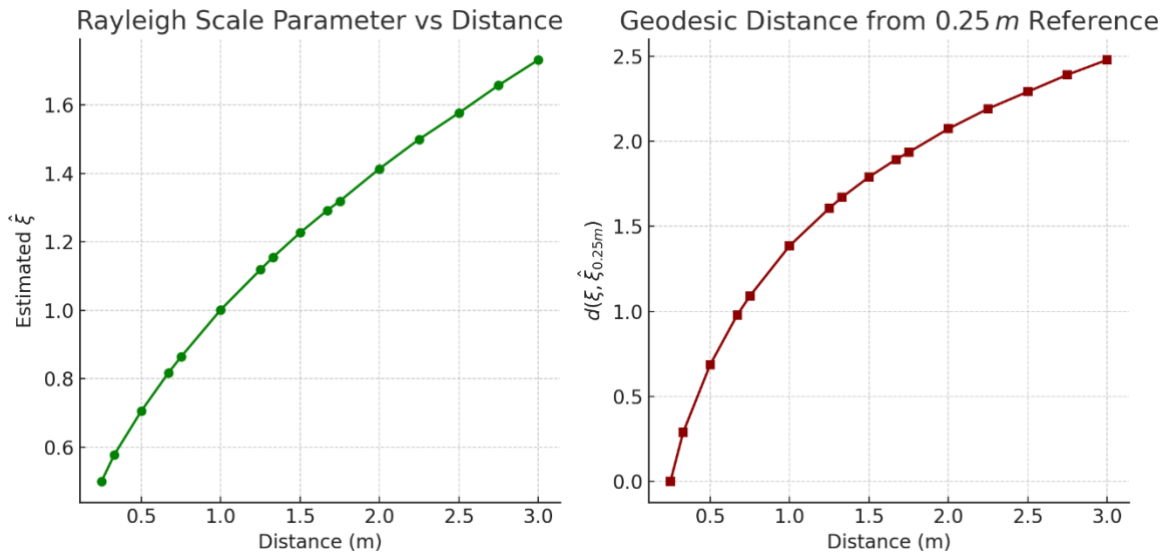
Elde edilen değerler Şekil 5.8’de verilen grafiğe aktarılmıştır. Bu grafiklere bakılarak şu yorumlar yapılabilir;

- Tahmini $\hat{\xi}$ fiziksel mesafe ile artar,
- Geodezik mesafe içsel istatistiksel varyasyonu yakalar,
- Rayleigh manifoldu karşılaştırma için doğal bir alan sağlar.

	0.25	0.33	0.5	0.67	0.75	1.0	1.25	1.33	1.5	1.67	1.75	2.0	2.25	2.5	2.75	3.0
0.25	0.000	0.285	0.660	0.972	1.105	1.476	1.773	1.854	1.978	2.082	2.122	2.292	2.422	2.530	2.642	2.744
0.33	0.285	0.000	0.372	0.686	0.820	1.192	1.490	1.570	1.694	1.798	1.837	2.007	2.137	2.245	2.357	2.459
0.5	0.660	0.372	0.000	0.315	0.450	0.825	1.123	1.204	1.330	1.434	1.473	1.645	1.774	1.882	1.994	2.096
0.67	0.972	0.686	0.315	0.000	0.134	0.510	0.809	0.889	1.016	1.120	1.159	1.331	1.460	1.568	1.680	1.782
0.75	1.105	0.820	0.450	0.134	0.000	0.375	0.675	0.754	0.882	0.985	1.025	1.197	1.326	1.434	1.546	1.648
1.0	1.476	1.192	0.825	0.510	0.375	0.000	0.300	0.380	0.508	0.611	0.651	0.823	0.952	1.060	1.172	1.274
1.25	1.773	1.490	1.123	0.809	0.675	0.300	0.000	0.080	0.208	0.311	0.351	0.523	0.652	0.760	0.872	0.974
1.33	1.854	1.570	1.204	0.889	0.754	0.380	0.080	0.000	0.128	0.231	0.271	0.443	0.572	0.680	0.792	0.894
1.5	1.978	1.694	1.330	1.016	0.882	0.508	0.208	0.128	0.000	0.104	0.144	0.316	0.445	0.553	0.665	0.767
1.67	2.082	1.798	1.434	1.120	0.985	0.611	0.311	0.231	0.104	0.000	0.040	0.212	0.341	0.449	0.561	0.663
1.75	2.122	1.837	1.473	1.159	1.025	0.651	0.351	0.271	0.144	0.040	0.000	0.172	0.301	0.409	0.521	0.623
2.0	2.292	2.007	1.645	1.331	1.197	0.823	0.523	0.443	0.316	0.212	0.172	0.000	0.130	0.238	0.350	0.452
2.25	2.422	2.137	1.774	1.460	1.326	0.952	0.652	0.572	0.445	0.341	0.301	0.130	0.000	0.108	0.220	0.322
2.5	2.530	2.245	1.882	1.568	1.434	1.060	0.760	0.680	0.553	0.449	0.409	0.238	0.108	0.000	0.112	0.214
2.75	2.642	2.357	1.994	1.680	1.546	1.172	0.872	0.792	0.665	0.561	0.521	0.350	0.220	0.112	0.000	0.102
3.0	2.744	2.459	2.096	1.782	1.648	1.274	0.974	0.894	0.767	0.663	0.623	0.452	0.322	0.214	0.102	0.000

Tablo 5.2. Fisher bilgi geometrisine dayalı Rayleigh ölçek parametreleri arasındaki geodezik mesafe matrisi

Sonuçlar, fiziksel mesafe arttıkça, tahmini ölçek parametresinin büyüdüğünü ve çok yollu etkilerden kaynaklanan daha güçlü dağılmayı yansıttığını doğrulamaktadır. Geodezik analiz, değişen koşullar altında kanal davranışının koordinatsız, bilgi teorik bir karşılaştırmasını sağlar.

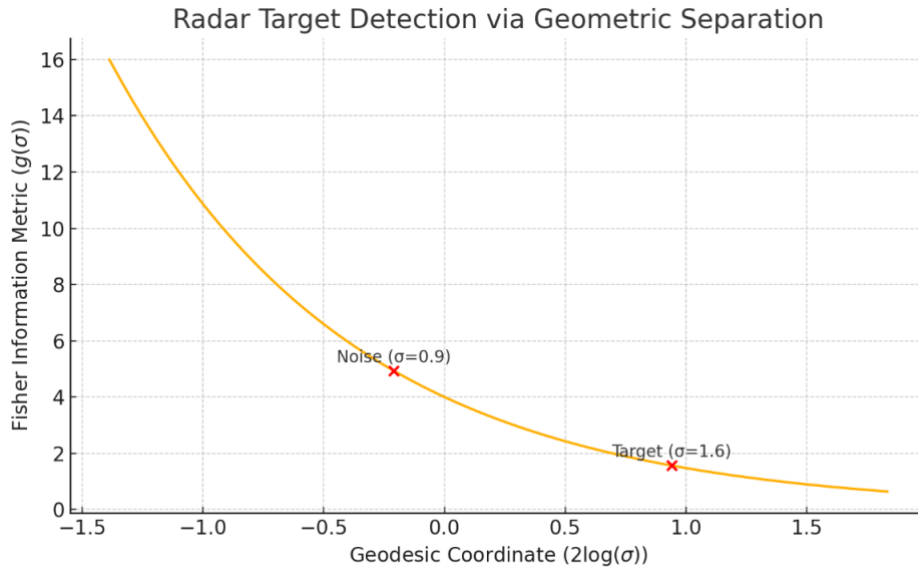


Şekil 5.8. *Solda:* Anten ayrılığının bir fonksiyonu olarak tahmin edilen $\hat{\xi}$ ölçek parametreleri. *Sağda:* Rayleigh manifoldu üzerindeki 0,25 m referans noktasından geodezik uzaklıklar.

Radar Sistemlerinde Hedef Tespitinin Geometrik Yorumu

Kablosuz sönümlenme modellerine ek olarak, Rayleigh dağılımı, radar sistemlerinde, özellikle görüş hattı olmayan veya karmaşık koşullar altında, alınan yankı işaretlerinin zarfını modellemek için yaygın olarak kullanılır. Klasik hedef tespit stratejileri tipik olarak alınan sinyal gücünü önceden tanımlanmış bir eşik değeriyle karşılaştırır. Bununla birlikte, Fisher bilgi metriğinin neden olduğu diferansiyel-geometrik yapı ile donatıldığında, Rayleigh dağılımları arka plan gürültüsü ile hedef yansımalarını ayırt etmek için daha ilkeli bir yaklaşıma izin verir.

Radar hedef tespiti için geodezik mesafenin kullanımı Zhao ve Wang tarafından gösterilmiştir, burada tespit görevini istatistiksel bir manifold üzerinde bir ayırma problemi olarak yeniden yorumlamışlardır (Zhao ve Wang, 2015). Yöntemleri, Fisher bilgi tabanlı mesafeleri kullanarak gözlenen dağılımı hedef var ve hedef yok hipotezleriyle karşılaştırmaktadır. Bu çerçeve, sinyal ortamları arasında ayırım yapmak için Rayleigh manifoldunun geodezik yapısını kullanma yaklaşımımızı desteklemektedir.



Şekil 5.9. Rayleigh bilgi manifoldunda radar hedef tespiti.

Sadece gürültülü geri dönüşlerin $\sigma_1 = 0.9$ parametresine sahip bir Rayleigh dağılımını takip ettiği ve hedefin mevcut olduğu geri dönüşlerin $\sigma_2 = 1.6$ parametresini

takip ettiđi varsayıldığında, Rayleigh manifoldu üzerindeki bu iki istatistiksel durum arasındaki geodezik mesafe řu řekilde verilir:

$$d(\sigma_1, \sigma_2) = 2 \left| \ln \left(\frac{1.6}{0.9} \right) \right| \approx 1.23.$$

Bu mesafe, sadece varyans karşılaştırmasının ötesinde iki ortam arasındaki istatistiksel farklılıđı ölçer. Şekil 5.9 bu ayrımı Rayleigh bilgi manifoldu bağlamında göstermektedir. İki rejim -arka plan ve hedef- parametre manifolduna gömölüdür ve aralarındaki ayrım geometrik olarak yorumlanabilir bir kavram haline gelir. Bu tür mesafeler eşik tabanlı dedektörlere dahil edilebilir veya değışiklik tespit algoritmalarında kullanılabilir.

Her nokta radar dönüşlerinin istatistiksel bir rejimini temsil eder (gürültü ve hedef) ve jeodezik ayrılıkları algılama zorluđunu ölçer.



6. SONUÇ VE ÖNERİLER

Öklid dışı geometriler ilk ortaya atıldığında bilim dünyası tarafından şüpheyile karşılanmıştı. Uzun yıllar beşinci postulatın ispatını arayan bilim insanları, bu yeni yaklaşımı kabullenmekte zorlandılar, çünkü beşinci postulatın olmadığı bir geometri fikri, bazıları için verilen tanımları ve ispatlanan yeni geometrik sonuçları reddetmek anlamına geliyordu. Öklid dışı geometriye dair en güçlü itirazlardan biri ise uygulanabilir olmayacağı yönündeydi. Öklid geometrisinin doğa ile mükemmel uyumu ve hem teoride hem de pratikteki kusursuz işleyişi, bu itirazı güçlendiriyordu. Ancak Riemann'ın yaptığı çalışma ile ortaya koyduğu yeni yaklaşım, Öklid dışı geometrileri ayrı bir boyuta taşıdı. Aksiyom tartışmasından öte, daha güçlü bir zemine oturan manifold kavramı ile Öklid dışı geometrilerin uygulanabilirliğine dair itirazlar zayıflamaya başladı. Evrenin doğal olarak düz olmayışı, bu eğriliğin formalize edilmesi ile Öklid geometrisinin çözmekte yetersiz kaldığı bazı problemlerin çözümüne dair yeni bir yol ortaya koydu. Öklid dışı geometriyi kapsamlı bir şekilde uygulayan ilk bilim insanlarından biri Albert Einstein'di. Einstein, rölativite teorisini Riemann geometrisine dayandırıyor. Bugün geldiğimiz noktada, Riemann geometrisi teorik fizikten matematiğe, veri biliminden yapay zekaya, sağlıktan mühendisliğe kadar birçok alanda uygulanmaktadır.

Bu tez çalışmasında, Öklid dışı geometrilerin istatistiksel modellemeye uygulanabilirliği, özellikle Riemann geometrisinin olasılık dağılımları üzerindeki etkisi bağlamında ele alınmıştır. Olasılık kuramının diferansiyel geometrik bir yapıya sahip olabileceği fikri, klasik istatistik anlayışının ötesine geçerek modellerin altında yatan yapısal ilişkileri anlamaya yönelik yeni bir bakış açısı sunmuştur. Fisher bilgi metriği ile tanımlanan Riemannian manifoldlar, olasılık dağılımlarının sadece sayısal değil, geometrik bir düzlemde de incelenmesine olanak tanımaktadır. Bu kapsamda, Rayleigh dağılımı gibi önemli bir dağılım, bir boyutlu manifold olarak modellenmiş ve geometri temelli yorumlamalarla teorik olduğu kadar pratik anlamda da uygulanabilir sonuçlar elde edilmiştir. Elde edilen bulgular, sinyal işleme gibi uygulamalı alanlarda bilgi geometrisinin ne denli güçlü bir araç olabileceğini göstermektedir. Böylece bu tez, Riemann geometrisinin sadece soyut matematiksel bir çerçeve değil, aynı zamanda istatistiksel çıkarım ve veri analizi gibi alanlarda da etkin bir yöntem olduğunu ortaya koymuştur.

Bu çalışma, bilgi geometrisinin özellikle Riemann geometrisi temelli yaklaşımlarının olasılık dağılımları üzerindeki etkilerini göstermesi bakımından önemli bir başlangıç noktası

oluşturmaktadır. Gelecek çalışmalar, bu geometrik yaklaşımın çok boyutlu dağılımlar üzerindeki etkilerini inceleyerek daha geniş kapsamlı modellere ulaşabilir. Özellikle Fisher metriğinin farklı dağılım aileleri üzerindeki davranışı ve bu yapıların uygulamalı algoritmalara entegrasyonu, makine öğrenmesi, sinyal işleme ve istatistiksel çıkarım gibi alanlarda yeni yöntemlerin geliştirilmesine olanak sağlayabilir. Ayrıca, bu tür geometrik yaklaşımların bilgi teorisi, biyoinformatik ve veri sıkıştırma gibi alanlarda kullanımı da araştırmaya değerlidir. Rayleigh dağılımı üzerinde yapılan geometrik incelemeler, benzer şekilde diğer dağılımlar için de uygulanarak dağılım temelli modellere yeni yorumlar kazandırabilir. Bu bağlamda, bilgi geometrisinin disiplinler arası çalışmalarda daha aktif şekilde kullanılması hem kuramsal gelişimi hem de pratik katkıları açısından teşvik edilmelidir.



7. KAYNAKLAR

- Abdel-All, N., Abd-Ellah, H., and Moustafa, H.** 2003. Information Geometry of Random Walk Distribution. *Publicationes Mathematicae* (63).
- Ali, M. M. and Woo, J. K.** (2002). Bayesian estimation for the Rayleigh distribution. *Journal of Statistical Computation and Simulation*, 72(10), 765–776.
- Amari, S.,** 1985. Differential Geometrical Methods in Statistics. In: Springer Lecture Notes in Statistics.
- Amari S. and Nagaoka H.,** 2000. Methods of Information Geometry. American Mathematical Society, Oxford University Press (2000).
- Amari, S. I.** 2016. Information geometry and its applications . Springer.
- Akdeniz, F.** 2022. Olasılık ve istatistik. Akademisyen Kitabevi.
- Basu, A. P.** (1988). Estimation of the scale parameter of the Rayleigh distribution based on censored samples. *IEEE Transactions on Reliability*, 37(2), 211–214.
- Branca, I. A. and Radulescu, I,** 2023. A Geometric Approach of Probability Distributions. *Romanian Journal of Mathematics and Computer Science* 1:13, 26-36.
- Calin, Ovidiu and Constantin Udriște.** 2014. Geometric modeling in probability and statistics. Berlin, Germany:: Springer,
- Cao, L., Sun, H. and Wang, X.** 2008. The geometric structures of the Weibull distribution manifold and the generalized exponential distribution manifold. *Tamkang Journal of Mathematics*, 39:1, 45-51.
- Chen, B.Y.** 1984. Total mean curvature and submanifolds of finite type, World Scientific Publishing Company,
- Corcuera J.M., Giummol'e F.,**1998. A Characterization of Monotone and Regular Divergences, *Annals of the Institute of Statistical Mathematics* 50:3, 433–450.
- Çako U.** 2022. Ters Gauss Dağılımının Geometrisi, *Yüksek Lisans Tezi*, Kahramanmaraş Sütçü İmam Üniversitesi, Kahramanmaraş.
- Dodson C.T.J.,** 1987. Systems of connections for parametric models. In Proc. Workshop on Geometrization of Statistical Theory. Lancaster, October 28-31 (C. T. J. Dodson, eds.), ULDS Publications, Univ. Lancaster.
- Dokur, E. and Kurban, M.** 2015. Wind speed potential analysis based on Weibull distribution. *Balkan Journal of Electrical and Computer Engineering*, 3, 231--235.
- Gomez-Deniz, E. and Gomez, L.** 2020 The Rayleigh-Birnbaum Saunders distribution: A general fading model. *Symmetry*, 12(3), 389.

- Işık, Ü.** 2024 Temel Olasılık ve İstatistik-1, Nesin Yayınevi, İzmir.
- Kobayashi, S. and Nomizu, K.** 1996. Foundations of differential geometry, volume 1 (Vol. 1). John Wiley & Sons.
- Kundu, D. and Raqab, M. Z.** 2005. Generalized Rayleigh distribution: properties and applications. Communications in Statistics - Theory and Methods, 34(3), 497–512.
- Palczynska, B. and Noga, K. M.** 2021. The Rayleigh model of the multipath fading channel (1–) [Dataset]. Gdansk University of Technology. <https://doi.org/10.34808/2m4t-zd14>
- Simon, M. K. and Alouini, M. S.** (2005). Digital Communication over Fading Channels. John Wiley & Sons.
- URL-1** 2025, <https://www.randomservices.org/random/special/GeneralExponential.html>, Exponential Family, 5 Ocak 2025
- Vanlı Turgut, A.** 2023. Diferensiyel Geometri, Kişisel Yayın, Ankara
- Zhao, X. and Wang, S.** 2015. Information Geometry Method to Radar Target Detection. Journal of Signal Processing, 31(6).

ÖZGEÇMİŞ

