

**DAİRESEL VERİLER VE
ARDIŞIK TESTLERDE KULLANIMI**

Kadir Özgür PEKER
Doktora Tezi

Fen Bilimleri Enstitüsü
İstatistik Anabilim Dalı
Kasım – 2002

JÜRİ VE ENSTİTÜ ONAYI

Kadir Özgür PEKER'in **Dairesel Veriler ve Ardışık Testlerde Kullanımı** başlıklı **İstatistik** Anabilim Dalındaki, Doktora Tezi **..01/11/2002....** tarihinde, aşağıdaki jüri tarafından Anadolu Üniversitesi Lisansüstü Eğitim-Öğretim ve Sınav Yönetmeliğinin ilgili maddeleri uyarınca değerlendirilerek kabul edilmiştir.

Adı Soyadı

İmza

Üye(Tez 1. Danışmanı) : Yard. Doç. Dr. Sevil BACANLI

Üye(Tez 2. Danışmanı) : Prof. Dr. Emel ŞIKLAR

Üye : Prof. Dr. Ersoy CANKÜYER

Üye : Prof. Dr. Ahmet ÖZMEN

Üye : Yard. Doç. Dr. Fikret ER

Anadolu Üniversitesi Fen Bilimleri Enstitüsü Yönetim Kurulunun **..06.11.2002...** tarih ve **..37/2.....**sayılı kararıyla onaylanmıştır.

Enstitü Müdürü
Prof. Dr. Orhan ÖZER
Fen Bilimleri Enstitüsü
MÜDÜRÜ

ÖZET

Doktora Tezi

DAİRESEL VERİLER VE ARDIŞIK TESTLERDE KULLANIMI

KADİR ÖZGÜR PEKER

Anadolu Üniversitesi
Fen Bilimleri Enstitüsü
İstatistik Anabilim Dalı

1. Danışman: Yard. Doç. Dr. Sevil BACANLI

2. Danışman: Prof. Dr. Emel ŞIKLAR

2002, 109 Sayfa

Bu çalışmanın amacı, dairesel verilerin özelliklerini tanıtmak, dairesel verilere uygulanan istatistiksel yöntemlerin, doğrusal verilere uygulanan yöntemlerden farklarını ortaya koymak ve dairesel verilerde ardışık testlerin kullanımını incelemektir. Çalışmada, dairesel verilerin özellikleri incelenmiş, dairesel ölçümlerde farklı birimler arasındaki dönüşümler açıklanmış, dairesel verilerde veri gösterimi yöntemleri ve dairesel veriler için temel parametre değerlerinin hesaplanması örneklerle açıklanmıştır. Dairesel veri analizindeki temel olasılık dağılımları verilmiş, von Mises dağılımı ayrıntılı olarak incelenmiştir. Daha sonra, dairesel verilerde ortalama yön testleri incelenmiş ve von Mises dağılımı için ortalama yön testi verilmiştir. Olasılık Oranlarının Ardışık Testi'in kuramsal tanımı, karakteristik işlem fonksiyonu ve ortalama örneklem sayısı fonksiyonu hakkında teorik bilgiler verilmiştir. Ayrıca, von Mises dağılımına sahip anakütleler için ardışık test süreçleri incelenmiştir. Veri türetimi yoluyla, açıklanan yöntemlerin uygulaması yapılmıştır. Gerçek veriler üzerinde uygulamayı göstermek amacıyla, Anadolu Üniversitesi Sivil Havacılık Yüksekokulu Meydan Meteoroloji İstasyon Müdürlüğü'nden, Anadolu Üniversitesi havaalanı için ölçülmüş olan rüzgar yönlerinden oluşan veri değerleri elde edilmiştir. Dairesel veriler için ardışık testlerin rüzgar yönü verilerine uygulanmasıyla elde edilen sonuçlar değerlendirilmiş ve bu testlerin kullanımına ilişkin yorumlar yapılmıştır.

Anahtar Kelimeler: Dairesel Veriler, Von Mises Dağılımı, Ardışık Testler.

ABSTRACT**PhD Thesis****CIRCULAR DATA AND IT'S APPLICATION TO SEQUENTIAL TESTS****KADİR ÖZGÜR PEKER**

**Anadolu University
Graduate School of Natural and Applied Sciences
Statistics Program**

Supervisor 1: Assist. Prof. Sevil BACANLI**Supervisor 2: Prof. Emel ŞIKLAR****2002, 109 Pages**

The aim of this study is to introduce the properties of circular data, to describe the differences between statistical techniques in circular data and linear data, and to investigate the usage of sequential tests in circular statistics. In the study, the properties of circular data is investigated, interrelations between different units of circular measurements is described, circular data display techniques and calculation of descriptive statistics for circular data are explained via examples. The probability distributions in circular data analysis is given and von Mises distribution is examined in detail. Furthermore, mean direction tests for circular data is described and the test of mean direction for von Mises distribution is introduced. The theoretical details, operating characteristic function and average sample number function of Sequential Probability Ratio Test is investigated in detail. Also, sequential test processes for von Mises distributed populations are given. The applications of the techniques by simulation are given. Then, a real life example is shown using the wind directions on Anadolu University airport, obtained from Anadolu University School of Civil Aviation Airfield Meteorology Station Administration. The results obtained by applying wind directions data to the sequential tests for circular data is explained and interpretation of how to use these tests is given.

Keywords: Circular Data, Von Mises Distribution, Sequential Tests.

TEŞEKKÜR

Bu çalışmanın her aşamasında değerli katkı ve eleştirileriyle bana yön veren danışman hocam Sayın Yrd. Doç. Dr. Sevil BACANLI'ya, çalışmaya katkılarından dolayı ikinci danışman hocam Sayın Prof. Dr. Emel ŞIKLAR'a ve değerli hocam Sayın Prof. Dr. Ersoy CANKÜYER'e, çalışmanın gerçekleşmesi için gerekli ortamı hazırlayan Anadolu Üniversitesi Fen Fakültesi İstatistik Bölüm Başkanı, değerli hocam Sayın Prof. Dr. Ali Fuat YÜZER'e, konu ile ilgili kaynaklara ulaşılmasında, bilgisayar programının yazılması ve işletilmesindeki katkılarından dolayı Sayın Yrd. Doç. Dr. Fikret ER'e, tezin yazımı sırasındaki yardımlarından dolayı Sayın Araş. Gör. Halil ERYILMAZ'a, bu süre boyunca, manevi desteklerini gördüğüm değerli hocalarıma, arkadaşlarıma ve aileme sonsuz teşekkürlerimi sunarım. Ayrıca, zorlu çalışma dönemi süresince desteklerini esirgemeyen, moral kaynağım sevgili Fatma ÖZFINDIK'a teşekkürlerimi sunarım.

İÇİNDEKİLER

	<u>Sayfa</u>
ÖZET	i
ABSTRACT	ii
TEŞEKKÜR	iii
İÇİNDEKİLER	iv
ŞEKİLLER DİZİNİ	vii
ÇİZELGELER DİZİNİ	ix
SİMGELER VE KISALTMALAR DİZİNİ	x
1. GİRİŞ.....	1
1.1. Literatür Bilgisi.....	4
2. DAİRESEL VERİLER VE ŞEMATİK GÖSTERİMLERİ.....	8
2.1. Dairesel Ölçümlerin Farklı Birimleri Arasındaki Dönüşümler.....	8
2.2. Gruplanmamış Veriler İçin Şematik Gösterim.....	9
2.3. Gruplanmış Veriler İçin Şematik Gösterim.....	11
2.3.1. Doğrusal histogram.....	11
2.3.2. Dairesel histogram.....	13
2.3.3. Gül şeması.....	14
3. DAİRESEL VERİLER İÇİN TANIMLAYICI İSTATİSTİKLER.....	17
3.1. Ortalama Yön, Bileşke Uzunluğu ve Ortalama Bileşke Uzunluğu	18
3.2. Yoğunlaşma Parametresi.....	26
3.3. Dairesel Varyans.....	27
3.4. Dairesel Standart Sapma.....	30
3.5. Dairesel Saçılım.....	31
3.6. Dairesel Standart Hata.....	32
3.7. Medyan Yönü.....	32
3.8. Mod Yönü.....	35

4.	DAİRESEL VERİLERDE OLASILIK DAĞILIMLARI.....	37
4.1.	Çember Üzerindeki Olasılık Dağılımları.....	39
4.1.1.	Düzgün dağılım.....	40
4.1.2.	Cardioid dağılım.....	41
4.1.3.	Genel sarmal dağılımlar.....	41
4.1.4.	Sarmal Cauchy dağılımı.....	42
4.1.5.	Sarmal normal dağılım.....	43
4.2.	Von Mises (Dairesel Normal) Dağılımı.....	44
5.	DAİRESEL VERİLERDE TEK-ÖRNEKLEM ORTALAMA YÖN TESTLERİ.....	49
5.1.	Tek Örneklem Ortalama Yön Testi.....	49
5.1.1.	Güven aralığının belirlenmesi yoluyla test.....	49
5.1.2.	Örneklem hacmine dayalı test.....	50
5.2.	Von Mises Dağılımına Sahip Anakütelerde Ortalama Yön Testi.....	50
5.2.1.	Güven aralığının belirlenmesi yoluyla test.....	51
5.2.2.	Örneklem hacmine ve yoğunlaşma parametresine dayalı test.....	51
5.2.2.1.	Yoğunlaşma parametresinin bilinmediği durum.....	51
5.2.2.2.	Yoğunlaşma parametresinin bilindiği durum.....	52
6.	OLASILIK ORANLARININ ARDIŞIK TESTİ.....	54
6.1.	Kuramsal Tanım.....	55
6.1.1.	A ve B değişmezlerinin saptanması.....	57
6.1.2.	Karakteristik işlem fonksiyonu.....	58
6.1.3.	Ortalama örneklem sayısı fonksiyonu.....	61
6.2.	Varyansı Bilinen Normal Dağılımın Ortalamasının Ardışık Testi.	63
6.2.1.	Testin çıkarsaması.....	63
6.2.2.	Karakteristik işlem fonksiyonu.....	65
6.2.3.	Ortalama örneklem sayısı fonksiyonu.....	66

7.	DAİRESEL VERİLERDE ORTALAMA YÖN İÇİN ARDIŞIK TESTLER.....	68
7.1.	Yoğunlaşma Parametresi Bilindiğinde Ortalama Yön İçin Ardışık Test (U-Testi)	70
7.1.1.	Kuramsal tanım.....	70
7.1.2.	Karakteristik işlem fonksiyonu.....	72
7.1.3.	U-testi için ortalama örneklem sayısı fonksiyonu.....	74
7.2.	Yoğunlaşma Parametresi Bilinmediğinde Ortalama Yön İçin Ardışık Test (V-Testi)	77
8.	UYGULAMA.....	82
8.1.	Veri Türetimi Uygulaması.....	82
8.1.1.	Veri türetimi işlemi.....	82
8.1.2.	Şematik gösterimler.....	83
8.1.3.	Tanımlayıcı istatistiklerin hesaplanması.....	85
8.1.4.	Tek örneklem ortalama yön testleri.....	86
8.1.5.	Ortalama yön için ardışık testler.....	87
8.2.	Meteorolojik Verilere Uygulama.....	90
8.2.1.	Rüzgar yönü verilerinin şematik gösterimleri.....	91
8.2.2.	Tanımlayıcı istatistiklerin hesaplanması.....	93
8.2.3.	Ortalama yön için ardışık test.....	94
8.3.	Sonuçlar ve Tartışma.....	96
	KAYNAKLAR	98
	EKLER	102
EK-1	Bessel fonksiyonlarına ilişkin formüller	102
EK-2	Von Mises dağılımında verilen \bar{R} değerleri için $\hat{\kappa}$ en çok olabilirlik tahminleri	104
EK-3	Von Mises dağılımında verilen κ değerleri için $\rho = A(\kappa)$ bileşke uzunlukları	105
EK-4	S-Plus Paket Programında dairesel veriler için kullanılan fonksiyonlar	106

ŞEKİLLER DİZİNİ

1.1.	Gözlem değerinin başlangıç noktası ve dönüş yönü seçimine bağlı olması.....	2
2.1.	Galatasaray futbol takımının 1999-2000 sezonunda UEFA kupasında attığı gollerin atıldığı dakikalara ilişkin birinci tür ham veri grafiği.....	10
2.2.	Galatasaray futbol takımının 1999-2000 sezonunda UEFA kupasında attığı gollerin atıldığı dakikalara ilişkin ikinci tür ham veri grafiği.....	11
2.3.	Galatasaray futbol takımının 2001-2002 futbol sezonunda Türkiye Birinci Süper Ligi'nde attığı gollerin atıldığı dakikalara ilişkin doğrusal histogram.....	13
2.4.	Galatasaray futbol takımının 2001-2002 futbol sezonunda Türkiye Birinci Süper Ligi'nde attığı gollerin atıldığı dakikalara ilişkin dairesel histogram.....	14
2.5.	Galatasaray futbol takımının 2001-2002 futbol sezonunda Türkiye Birinci Süper Ligi'nde attığı gollerin atıldığı dakikalara ilişkin gül şeması.....	15
3.1.	Dik ve kutupsal koordinatlar arasındaki ilişki.....	17
3.2.	Aritmetik ortalamanın yanlış yönde oluşu.....	19
3.3.	İki dairesel gözlem için bileşke vektör ve ortalama yön.....	21
4.1.	Sürekli bir dairesel dağılımın dairesel gösterimi.....	38
4.2.	Sürekli bir dairesel dağılımın doğrusal gösterimi.....	38
4.3.	$\mu = 0^\circ$ ve $\kappa = 0.5, 1, 2$ ve 4 değerleri için von Mises yoğunluk fonksiyonları.....	47
4.4.	$\mu = 0^\circ$, $\kappa = 1$ iken, tek-modlu ve simetrik dağılımların karşılaştırılması.....	48
7.1.	$\mu_0 = 0$, $\mu_1 = 90^\circ$ ve $\alpha = \beta = 0.05$ iken U-testi için karakteristik işlem fonksiyonu.....	74
7.2.	Zıt yönlerde, fakat ortalama yönü açısal olarak eşit hipotezler için V-testi.....	80
8.1.	Von Mises dağılımından türetilen verilerden seçilen örneklemin ham veri grafiği.....	83
8.2.	Von Mises dağılımından türetilen verilerden seçilen örneklemin doğrusal histogramı.....	84
8.3.	Von Mises dağılımından türetilen verilerden seçilen örneklemin dairesel histogramı.....	84

8.4.	Von Mises dağılımından türetilen verilerden seçilen örneklemin gül şeması.....	85
8.5.	Rüzgar yönü verilerinin ham veri grafiği.....	91
8.6.	Rüzgar yönü verilerinin doğrusal histogramı.....	92
8.7.	Rüzgar yönü verilerinin dairesel histogramı.....	92
8.8.	Rüzgar yönü verilerinin gül şeması.....	93

ÇİZELGELER DİZİNİ

2.1.	Galatasaray futbol takımının 2001-2002 futbol sezonunda Türkiye Birinci Süper Ligi'nde attığı gollerin atıldığı dakikalara ilişkin gruplanmış veriler.....	12
3.1.	Çizelge 2.1.'de verilen veri seti için ortalama bileşke uzunluğu ve ortalama yönün hesaplanması.....	26
3.2.	Çizelge 2.1.'de verilen veri seti için medyan yönünün hesaplanması.....	34
3.3.	Çizelge 2.1.'de verilen veri seti için mod yönünün hesaplanması...	35
5.1.	n ve κ değerlerine bağlı olarak testin uygulanabildiği durumlar.....	52
6.1.	Olasılık oranlarının ardışık testi için çeşitli h değerlerinde Θ ve $P(\Theta)$ değerleri.....	61
6.2.	Çeşitli h değerlerine karşılık gelen μ ve $P(\mu)$ değerleri.....	66
8.1.	Von Mises dağılımından türetilen verilerden seçilen örneklem.....	83

SİMGELER VE KISALTMALAR DİZİNİ

- θ : Dairesel gözlem değeri
 $\bar{\theta}$: Örneklem ortalama yönü
 μ : Anakütle ortalama yönü
 R : Örneklem bileşke uzunluğu
 ρ : Anakütle bileşke uzunluğu
 \bar{R} : Örneklem ortalama bileşke uzunluğu
 $\bar{\rho}$: Anakütle ortalama bileşke uzunluğu
 κ : Yoğunlaşma parametresi
 $A_1(x)$: Dönüştürülmüş iki Bessel fonksiyonunun oranı
 $I_p(x)$: p -inci sıra dönüştürülmüş Bessel fonksiyonu
 V : Dairesel varyans
 ν : Dairesel standart sapma
 δ : Dairesel saçılım
 σ : Dairesel standart hata
 $\tilde{\theta}$: Medyan yönü
 $\check{\theta}$: Mod yönü
o.y.f. : Olasılık yoğunluk fonksiyonu
VM : Von Mises yoğunluk fonksiyonu
 α : Birinci tip hata olasılığı
 β : İkinci tip hata olasılığı
 A : Ardışık testler için üst sınır
 B : Ardışık testler için alt sınır
 $E(n; \mu)$: Ortalama örneklem sayısı fonksiyonu
 $P(\mu)$: Karakteristik işlem fonksiyonu
 L_n : Olabilirlik oranı
 λ_n : Von Mises dağılımının n -inci gözlem için olabilirlik fonksiyonu

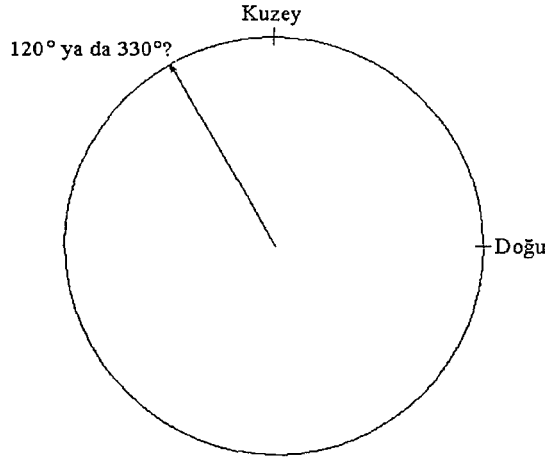
1. GİRİŞ

Günümüzde bir çok bilim dalında ölçümler açısal olarak elde edilmekte ve iki-boyutlu düzlemde yönsel olarak veri toplanması işlemi sıklıkla kullanılmaktadır. Bu tür verilerin özellikle biyoloji, coğrafya, jeoloji, jeofizik, tıp, meteoroloji ve okyanus-bilim alanlarında kullanıldığı görülmektedir. Açısal gözlemler, rassal deneylerde farklı biçimlerde ortaya çıkmaktadırlar. Bu gözlemler, rüzgar yönleri, kuşların serbest bırakıldıktan sonraki gidiş yönleri, hayvanların herhangi bir uyarılmaya karşı verdikleri yönsel tepkiler, okyanustaki akıntı yönleri ve biyoritimler gibi doğrudan ölçümler şeklinde olabilmektedir. Örneğin bir biyolog bir kuşun uçuş yönünü ya da bir hayvanın hareket yönünü ölçerken, bir jeolog da dünyanın manyetik kutup yönüne ilişkin bir araştırma yapabilir. İlk örnekteki yönsel araştırmalar iki boyutlu olarak incelenirken, ikinci örnekteki araştırma üç boyutta incelenmektedir. Yönlere ilişkin bu şekildeki herhangi bir gözlem kümesi, *yönsel (directional)* veri olarak adlandırılmaktadır. Yönsel veriler bazı periyodik ve açığa dönüştürülmüş zaman ölçümlerinde de dolaylı olarak elde edilebilmektedirler. Örnek olarak, belirlenen yıllarda aylık olarak belirli bir hastalığın meydana gelme oranı verilebilir.

İki-boyutlu yönler, uygun biçimde seçilmesi gereken bir “sıfır yönü”ne ve “dönüş doğrultusu”na göre ölçülen açılar olarak gösterilebilmektedirler. Burada sıfır yönü başlangıç noktasını ve dönüş doğrultusu da pozitif yön olarak saat yönünün mü yoksa saat yönünün tersinin mi alınacağını belirtmektedir. Yön kavramında herhangi bir büyüklük söz konusu olmadığından, gözlem değerleri merkezi orijin olan bir birim çemberin çevresi üzerinde noktalarla ya da orijini bu noktalarla birleştiren birim vektörlerle gösterilebilir. Buna göre derece cinsinden ölçülmüş tek bir gözlem θ ($0^\circ < \theta \leq 360^\circ$), birim vektör olacaktır. Bu dairesel gösterimden dolayı, bu şekildeki iki-boyutlu yönlerle ilişkin gözlemler *dairesel (circular)* veri olarak da adlandırılmaktadır. Burada θ ; vektör ile pozitif x- ekseninin, saat yönünün tersi yönünde yaptığı açığı göstermektedir. Vektörün kartezyen koordinatları $(\cos\theta, \sin\theta)$ ve kutupsal koordinatları da $(1, \theta)$ 'dir. Eğer vektörün yönü belli değil ise, yani θ ($0^\circ < \theta \leq 180^\circ$) ve $180^\circ + \theta$ açıları birbirinden ayırt edilemiyorsa, veri *açısal (angular)* veri olarak tanımlanabilir. Benzer olarak,

üç boyutlu yönsel gözlemler ise, üç boyutta birim vektör biçimindeki iki açı yardımıyla (dünya yüzeyi üzerindeki herhangi bir noktanın enlemi ve boylamı yardımıyla gösterilmesinde olduğu gibi), ya da bir birim kürenin yüzeyi üzerindeki noktalar biçiminde gösterilebilmektedirler. Bundan dolayı, üç-boyutlu gözlemler, *küresel (spherical)* veri olarak da adlandırılmaktadırlar.

Yönsel verilerin hem modelleme hem de istatistiksel deneylere ilişkin birçok farklı ve yeni özelliği bulunmaktadır. Örneğin, iki-boyutlu bir yönün, açı ya da birim vektör biçimindeki sayısal gösteriminin tek bir tane olması beklenmemelidir. Çünkü dairesel gözlemin değeri, sıfır yönü ve dönüş doğrultusu durumunun saat yönünde olup olmaması seçimine bağlıdır. Şekil 1.1.'e bakıldığında; bir matematikçi için 120° , Doğu yönü sıfır-yönü kabul edilerek saat yönünün tersine doğru bir dönüş işlemi uygulandıktan sonra belirlenirken, aynı nokta bir jeolog için Kuzey yönü sıfır-yönü kabul edilerek saat yönü doğrultusunda bir dönüş işlemi uygulandıktan sonra belirlenmektedir ve bu nokta 330° 'ye karşılık gelmektedir.



Şekil 1.1. Gözlem değerinin başlangıç noktası ve dönüş yönü seçimine bağlı olması

Bu durum küresel veriler için de geçerlidir. Burada üzerinde durulması gereken önemli nokta, elde edilen sonuçların (özet veriler, çıkarımlar vb.), verilen gözlem değerlerinin bir fonksiyonu olduğu ve bunlara verilen keyfi değerlere bağlı olmamalarıdır. Yani, keyfi seçilen orijine ve dönüş doğrultusuna bağlı olmayan sonuçlar hedeflenmelidir. Yine, bu keyfilikten dolayı doğal bir

düzenleme ya da sıralama bulunmamaktadır. Çünkü, bir yönün diğerinden “daha büyük” olup olmaması, saat yönünden ya da saat yönünün tersinden hangisinin pozitif doğrultu olarak ele alındığına ve de sıfır yönünün nerede olduğuna bağlıdır. Başlangıç noktası ile bitiş noktası çakıştığı, yani $0=2\pi$ olmasından dolayı ve genelde ölçümün, herhangi bir p tamsayısı için θ değeri $\theta+2\pi$ değerine eşit olacak şekilde periyodik olmasından dolayı, yönsel verilerde ele alınan yöntemlerde herhangi iki nokta arasındaki uzaklığın nasıl ölçüleceğine dikkat edilmesi gerekmektedir.

Bu türdeki özelliklerinden dolayı yönsel analiz, tek değişkenli ya da çok değişkenli verilerin standart istatistiksel analizinden oldukça farklıdır. İstatistiksel yöntemlerdeki değişmezliğe ve keyfi sıfır-yönü ve dönüş doğrultusu seçimine ilişkin ölçülere duyulan gereksinim, birçok bilinen tekniği ve ölçüleri tamamen anlamsız olmasa da çoğu zaman hatalı kılmaktadır. Gerçek doğru üzerinde sıkça kullanılan örneklem ortalaması ve varyans gibi özet ölçülerin ve tüm momentlerle kümülanların uygun olmadıkları anlaşılmaktadır. Buna eşdeğer olarak moment üreten fonksiyon ve diğer üreten fonksiyonlar gibi analitik araçlar da yararsız olmaktadır. Korelasyon, regresyon ve de bunların istatistiksel ölçüleri gibi birçok kavram yönsel veri setleri için yeniden düzenlenmek zorundadır. Benzer olarak, yansızlık, kayıp fonksiyonlar, varyans sınırları, testlerin güç fonksiyonlarının monotonluğu, vb. istatistiksel çıkarımlar dikkatle yeniden tanımlanmalıdır.

Yönsel veriler için parametrik istatistiksel çıkarımların büyük bir kısmı yalnızca bir ya da iki modele dayanarak türetilmektedir. Uygulamada çoğu kez karşılaşılan asimetric veri setlerinin modellenmesinde dahi, uygun model sayısının az olması nedeniyle tartışmalar meydana gelmektedir. Son yirmi yılda veri gösterimi, korelasyon, regresyon ve zamana ya da konuma bağlı yapıdaki verilerin analizi problemleri üzerinde durularak, bu verilerin işlenmesinde kullanılan istatistiksel yöntemler hızlı bir biçimde geliştirilmiştir. Buna ek olarak, genel istatistik yöntem-biliminde ve özellikle parametrik olmayan düzleştirme yöntemleri ve bootstrap yöntemlerindeki modern gelişmeler, veri analizcilerinin denetlemesi nispeten güç olan problemlerde ilerleme kaydetmelerine önemli ölçüde katkıda bulunmaktadır.

Konuyla ilgili ilk ciddi çalışma, Mardia (1972) tarafından yapılmıştır. Mardia çalışmalarında dairesel veri setlerinin istatistik teorisinde nasıl kullanılabileceği ya da istatistik teorisinin dairesel veri analizine nasıl uygulanabileceği üzerinde durmuştur. Bu alanda yazılan diğer kitaplar arasında Fisher (1983) de bulunmaktadır.

Dairesel veri analizinde herhangi bir regresyon analizi problemi çeşitli varsayımlar da göz önüne alınarak, dairesel analiz teknikleriyle çözümlenebilecek hale getirilirler. Örneğin, herhangi bir mağazanın bir yıl boyunca yaptığı satışlar 360° olduğu bilinen bir daireye her gün için bir açı gelecek şekilde eşit olarak dağıtılırsa ve bu geniş bir zaman süreci için incelenirse satışlarda yıllara (ya da mevsimlere) göre dairesellik meydana gelmiş olur.

Basit istatistiksel tekniklerin dairesel (yönsel) veri setlerine uyarlanmasıyla, daha karmaşık tekniklerin de kullanılabilmesi sağlanmaktadır. İstatistikçilerin konuya yoğunlaşmaları son yıllarda giderek artmakta ve özellikle parametrik olmayan tekniklere ilişkin çalışmalar yapılmaktadır.

Bütün bunların sonucunda, yönsel verinin bu alanı araştırmacılara çok geniş bir alanda ilerleme olanağı vermektedir ve yeni istatistiksel yöntemler ve araçlar geliştirmede çok verimli bir alan olduğu görülmektedir. Ayrıca doğal, fiziksel, tıbbi ve de sosyal bilimlerde ortaya çıkan problemler için yeni ve farklı uygulamalar geliştirilebilmektedir.

1.1. Literatür Bilgisi

Gumbel ve ark. (1953), dairesel normal dağılımın teorik altyapısını incelemişler ve bu dağılıma ilişkin grafik ve tabloları oluşturmuşlardır.

Gumbel (1954), önceki çalışmasında teorik yapısını incelemiş olduğu dairesel normal dağılımın uygulamadaki kullanımını üzerinde çalışmıştır. Jeofizik, ekonomi ve yaşam istatistikleri verileri üzerinde uygulama gerçekleştirmiştir.

Ajne (1968), dairesel dağılımlar için bir düzgünlük testi geliştirmiştir. Bu çalışmada, bir yarım çember üzerinde maksimum sayıda nokta içerebilen bir test istatistiği üzerinde çalışılmıştır. Sıfır hipotezi H_0 ve alternatif hipotez H_1 altındaki tam ve asimptotik dağılımlar, tablolarıyla birlikte verilmektedir.

Stephens (1969), veri seti bir birim çemberin çevresi üzerindeki noktalardan oluştuğu durumdaki rassallık testlerini tanımlamıştır. Ayrıca, test istatistikleri için anlamlılık noktaları tabloları verilmiştir.

Mardia (1975) çalışmasında, çok-boyutlu durum için temel dağılımlar olarak bilinen von Mises-Fisher dağılımı ve Bingham dağılımı incelenmektedir. Ayrıca, bu konuya ilişkin çeşitli sayısal örnekler verilmektedir.

Collett (1980), tek-değişkenli dairesel ya da yönsel verilerden oluşan örneklemdeki şaşırtıcı değerlerin görünüşüne ilişkin problemleri incelemiştir. Bu değerlerin varlığının nasıl tespit edilebileceği ve olası uyum testlerinin tanımlanması üzerine çalışmıştır. Test süreçlerinin görece performansı Monte Carlo yöntemleri kullanılarak araştırılmıştır. Ayrıca, uyum testlerine ilişkin bir uygulama gerçekleştirilmiştir.

Jupp ve Mardia (1980) çalışmasında, iki-değişkenli dairesel dağılımlar için bir korelasyon katsayısı sunulmaktadır. Bu korelasyon katsayısının bazı özellikleri incelenmiş ve diğer iki-yönlü korelasyon katsayılarıyla karşılaştırılmıştır. Korelasyon katsayısıyla ölçülmüş olan tam bağımlılıktan ortaya çıkan regresyon modelleri üzerinde çalışılmıştır.

Gadsden ve Kanji (1981), von Mises dağılımına sahip anakütleler için bazı ardışık test süreçleri geliştirme çalışması yapmışlardır. Ayrıca testin kullanımının gösterilmesi amacıyla sayısal bir örnek vermişlerdir.

Kotz ve Johnson (1982), hazırladıkları *İstatistiksel Bilimler Ansiklopedisi*'nde, genel olarak yönsel veri analizini, yönsel dağılımlardan dairesel dağılımları ve küresel dağılımları incelemişlerdir.

Gadsden ve Kanji (1982-83) çalışmalarında, farklı bilim dallarında kullanılan dairesel veriler için ardışık süreçlere ilişkin bazı uygulamalar sunmakta ve dairesel veri analizinde ortaya çıkan problemlerde kullanılan yöntemleri kısaca tartışmaktadırlar. Son olarak, inceledikleri testlere ilişkin bazı uygulamalar vermişlerdir.

Fisher ve Lee (1983), dairesel değişkenler arasındaki ilişki miktarını ölçen bir katsayı üzerinde tartışmışlar, bu katsayının asimptotik dağılımını bulmuşlar ve özelliklerini geliştirmişlerdir. Kullanılan diğer istatistiklerle karşılaştırmaları yapılmış ve bazı örnekler verilmiştir.

Cheaney (1983), jeolojide kullanılan istatistiksel yöntemler üzerinde çalışmıştır. Bu çalışmada, iki-boyutlu ve üç-boyutlu yönsel veriler için çeşitli hipotez testleri verilmiş ve bu verilere ilişkin önemli dağılımlar incelenmiştir.

Jeoloji biliminde istatistiksel veri analizi konusundaki çalışmalardan bir diğeri de Davis (1986) tarafından yapılmıştır. Bu çalışmada, yönsel veri analizi genel hatlarıyla incelenmiş, dairesel yönsel verilere ilişkin hipotez testleri, küresel dağılımlar, küresel yönsel verilere ilişkin hipotez testleri konularına değinilmiştir.

Fisher ve ark. (1987), küresel verilerin istatistiksel analizi konusundaki temel çalışmalardan birisini gerçekleştirmişlerdir.

Coblentz ve Richardson (1995), stres rejimindeki ve temel istatistik tekniklerinin kullanımıyla elde edilen stres göstergelerinden bulunan yönsel bilgidaki trendlerin değerlendirilmesi üzerine araştırma gerçekleştirmişlerdir.

Presnell ve ark. (1998), yönsel veriler için küresel olarak tasarlanmış çok değişkenli doğrusal model sunmuşlardır. Bu çalışmada, dairesel veri durumu dikkate alınarak, model için en çok olabilirlik tahminleri karşıt yaklaşımların tam tersine iteratif yöntemler kullanılarak hesaplanmıştır. Sonuçta elde edilen yöntemlerin gerçek uygulamalarla açıklandığı örnekler verilmiştir.

Lund (1999) çalışmasında, dairesel bir yanıt değişkeninin başka bir dairesel değişken ve bir grup doğrusal yardımcı değişken yardımıyla öngörülmesi amacıyla, en küçük kareler regresyonunun dairesel yorumunu sunmaktadır.

Anderson-Cook (2000), faktöriyel bir deneyden elde edilen faktörlerin etkisinin belirlenmesi için C-doğrusal eğrilerini dairesel-doğrusal veriye uydurma yöntemini kullanarak, otomotiv endüstrisine ilişkin endüstriyel bir örnek incelemiştir.

Artes ve ark. (2000) çalışmasında, iki parametrelili dairesel dağılımlar ailesi için, boylamsal yapıdaki dairesel verinin modellenmesinde kullanılan tahmin denklemlerini elde etmektedir. Burada, dairesel ortalama ve bileşke uzunluğunun modellenmesinde kullanılan tahmin denklemleri ayrı ayrı verilmektedir. Bunun ardından, bir karışım modeli için tahmin denklemleri bulunmuştur. Çalışma sonucunda, bu denklemlerden elde edilen tahminlerin sürekli ve asimptotik olarak normal oldukları gösterilmektedir. Sonuçlar, göçmen kuşlardan elde edilen yönsel veriye ilişkin örneklerle açıklanmıştır.

Le Pennec (2000), yönsel verilerle kül akımı kaynaklarının belirlenmesi üzerine çalışma yapmıştır. Çalışmasında, Türkiye'nin Orta Anadolu bölgesinde bulunan Kızılkaya ignimbiritleri (bir tür sert kayaç) üzerine bir uygulama gerçekleştirmiştir.

Sengupta ve Pal (2001), çalışmalarında izotropi (bütün yönlerde aynı özelliklere sahip bir ortamın durumu) hipotezi ya da yönlerin rassallığı için konumsal en güçlü değişmez (*LMPI- locally most powerful invariant*) bir test geliştirmişlerdir. Testin güç fonksiyonunun küresel monotonluğu saptanmıştır ve bu test sürekli dir. Seçilen bazı parametre kombinasyonları için testin güç değerleri, örneklem hacmi düşük olduğunda bile oldukça iyi bir performans gösteren simülasyon sonucu elde edilmiştir.

Lund (2002), dairesel ve doğrusal tahmin edicilerin birleşiminden oluşan bir dairesel yanıt tahmininin elde edilmesinde kullanılan regresyon ağaçları için temel bir yapı oluşturmuştur. Bu çalışmada sunulan yöntemler dairesel yanıt değişkenli regresyon ağacı çözümlemesinde kullanılmaktadır.

2. DAİRESEL VERİLER VE ŞEMATİK GÖSTERİMLERİ

Bu bölümde, dairesel ölçümlere ilişkin dönüşümler ve dairesel ölçümlerde tek örneklem analizinin en temel durumları olan veri gösterimi yöntemleri incelenecektir. Dairesel verilerin şematik gösterimi ilk olarak gruplanmamış verilerde, ikinci olarak da gruplanmış verilerde incelenecektir.

2.1. Dairesel Ölçümlerin Farklı Birimleri Arasındaki Dönüşümler

Teoride, derece olarak belirtilen θ° açısının radyan cinsinden θ açısına dönüştürülmesi tercih edilmektedir. Buna göre;

$$\theta^\circ = \frac{180\theta}{\pi} \quad , \quad 0^\circ < \theta^\circ \leq 360^\circ \quad (2.1)$$

$$\theta = \frac{\pi\theta^\circ}{180} \quad , \quad 0 < \theta \leq 2\pi \quad (2.2)$$

olur. Birim yarıçaplı çember üzerindeki herhangi bir yay, çemberin merkezinde bulunan θ radyan açısının iki ucunun birleşmesinden oluşuyorsa, yayın uzunluğu θ olacaktır.

Herhangi bir veri kümesindeki gözlem değerleri ($0^\circ, 90^\circ$) ya da ($0^\circ, 180^\circ$) aralıklarında yoğunluk gösterebilirler. Örneğin, eksensel veriler ($0^\circ, 180^\circ$) aralığında yer almaktadırlar. Bu durumda her bir açı değeri iki ile çarpılarak ($0^\circ, 360^\circ$) aralığına dönüşüm yapılabilir. Ayrıca, θ° 'nin ($0^\circ, 360^\circ/k$) aralığında yer alan bir değeri,

$$\theta = \frac{k\pi\theta^\circ}{180} \quad (2.3)$$

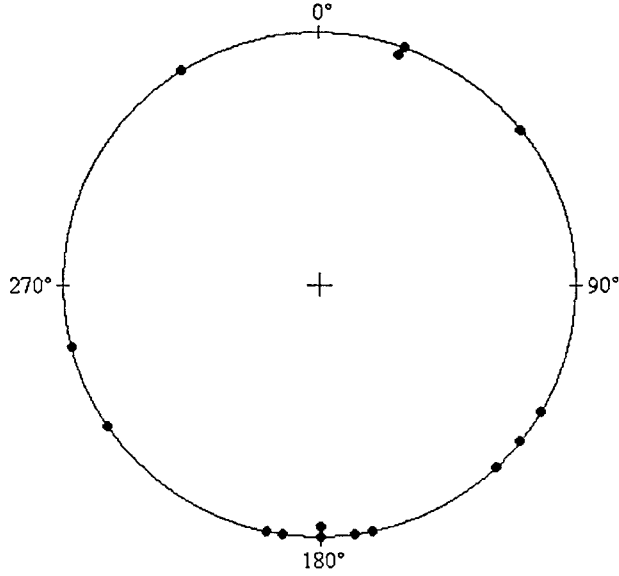
Eşitlik (2.3) yardımı ile ($0, 2\pi$) aralığına dönüştürülebilmektedir. Buradaki k değeri, herhangi bir tamsayıdır.

2.2. Gruplanmamış Veriler İçin Şematik Gösterim

Dairesel gözlemlerin gösterimi gruplanmamış verilerde genellikle iki biçimde olmaktadır ve bu gösterim türlerine genel olarak *ham veri grafikleri* adı verilmektedir. Ham veri grafiklerinin en önemli özelliği küçük bir mod grubunun varlığını ya da özel olarak incelenmesi gereken aykırı değerleri net bir biçimde ortaya koyabilmesidir. Birinci gösterim türünde, her bir gözlem değeri birim çemberin çevresi üzerine nokta olarak yerleştirilir. Eğer veri seti içerisinde herhangi bir gözlemden birden fazla sayıda var ise, bu değerler üst üste noktalar halinde gösterilebilir.

Örnek 2.1.

1999-2000 futbol sezonunda Galatasaray futbol takımının şampiyonu olduğu UEFA kupası karşılaşmalarında attığı gollerin atılış dakikaları şu şekildedir: 5, 5, 13, 30, 32, 34, 42, 43, 45, 45, 47, 48, 59, 64, 82. Bu gol dakikaları, $360 \times t / 90$ derecelik ölçüme karşılık gelen t dakikalık daireSEL ölçümlerdir. Böylece 1 dakika, 4° 'ye karşılık gelmektedir. Buna göre veriler yeniden düzenlendiğinde: $20^\circ, 20^\circ, 52^\circ, 120^\circ, 128^\circ, 136^\circ, 168^\circ, 172^\circ, 180^\circ, 180^\circ, 188^\circ, 192^\circ, 236^\circ, 256^\circ, 328^\circ$ olur. Şekil 2.1.'de verilen grafiğe bakıldığında, ilk yarımın sonları ile ikinci yarımın başları arasında bir kümelenme olduğu gözlenmektedir. Buna göre atılan 15 golden 6 tanesinin 45-inci dakika civarında atıldığı söylenebilir.

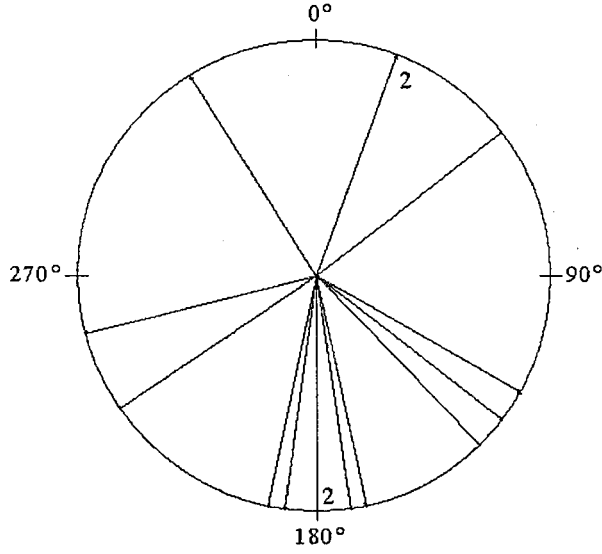


Şekil 2.1. Galatasaray futbol takımının 1999-2000 sezonunda UEFA kupasında attığı gollerin atıldığı dakikalara ilişkin birinci tür ham veri grafiği

Bu gösterime alternatif olarak veri gösterimi, çemberin çevresi üzerindeki gözlenen noktalarla orijinin birleştirilmesi sonucunda elde edilen birim çember yarıçapları çizilerek yapılabilmektedir. Örnek 2.2.'de bu tür bir grafik gösterilmektedir.

Örnek 2.2.

Örnek 2.1.'deki veriler dikkate alındığında, yukarıda açıklaması yapılan ham veri grafiği türü Şekil 2.2.'de görüldüğü gibi olmaktadır. Bir gözlem değerinden birden fazla sayıda olması durumu, o gözlemleri belirten vektörün üzerine tekrar sayısı yazılarak belirtilmektedir.



Şekil 2.2. Galatasaray futbol takımının 1999-2000 sezonunda UEFA kupasında attığı gollerin atıldığı dakikalara ilişkin ikinci tür ham veri grafiği

Burada vektörler birim uzunlukta. Birinci ham veri grafiği türü, doğru üzerindeki veriler için kullanılan genel kartezyen grafik türüne benzer olmasından dolayı çoğunlukla tercih edilmektedir.

2.3. Gruplanmış Veriler İçin Şematik Gösterim

Dairesel verilerde, $(0^\circ, 360^\circ)$ aralığı belirlenen sayıda sınıf-aralığına bölünebilmekte ve her bir sınıfa karşılık gelen sıklık sayılabilmektedir. Sınıf sınırlarının seçimi ve sınıf aralıklarının uzunluğu, doğrusal verilerde olduğu gibidir.

Gruplanmış türdeki daireli verilerin şematik gösterimi genel olarak üç tür histogram biçimi kullanılarak yapılmaktadır. Bunlar; *doğrusal histogram*, *daireli histogram*, ve *gül şeması*'dir.

2.3.1. Doğrusal histogram

Doğrusal histogramda veri daireli olarak değil, doğrusal olacak şekilde oluşturulur. Bir başlangıç noktası (örneğin, eğer veri $(0^\circ, 360^\circ)$ aralığında kaydedilmişse 0° , ya da $(-180^\circ, 180^\circ)$ aralığında ise -180°) ve bir grup aralığı veya genişliği seçilir (örneğin 5° , 10° veya 20°) ve histogram bilinen yolla elde edilir.

Sonuçta histogram 360° genişliğine sahip bir kesite oturtulmuş olmaktadır. Dairesel veriler için doğrusal histogram, bir çemberin çevresinde sarmal olacak şekilde tasarlandıktan sonra değerlendirilmelidir. Bundan dolayı, birinci blok grup genişliğinin sonundan itibaren tekrar edilebilir. Doğrusal histogram, yorumlanmasındaki etkinliğinden dolayı tercih edilen bir histogram türüdür.

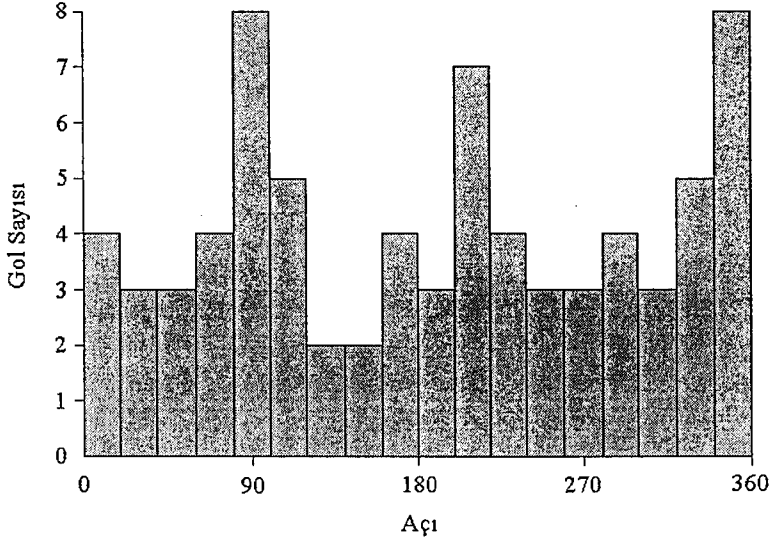
Örnek 2.3.

2001-2002 futbol sezonunda Galatasaray futbol takımının şampiyonu olduğu Türkiye Birinci Süper Ligi karşılaşmalarında attığı gollerin atılış dakikaları veri seti için 20° (5 dakika)'lik grup genişliği seçilerek elde edilen sıklık dağılımı Çizelge 2.1.'de verilmektedir:

Çizelge 2.1. Galatasaray futbol takımının 2001-2002 futbol sezonunda Türkiye Birinci Süper Ligi'nde attığı gollerin atıldığı dakikalara ilişkin gruplanmış veriler

<i>Açı</i>	<i>Gol Sayısı</i>	<i>Açı</i>	<i>Gol Sayısı</i>
$1^\circ - 20^\circ$	4	$181^\circ - 200^\circ$	3
$21^\circ - 40^\circ$	3	$201^\circ - 220^\circ$	7
$41^\circ - 60^\circ$	3	$221^\circ - 240^\circ$	4
$61^\circ - 80^\circ$	4	$241^\circ - 260^\circ$	3
$81^\circ - 100^\circ$	8	$261^\circ - 280^\circ$	3
$101^\circ - 120^\circ$	5	$281^\circ - 300^\circ$	4
$121^\circ - 140^\circ$	2	$301^\circ - 320^\circ$	3
$141^\circ - 160^\circ$	2	$321^\circ - 340^\circ$	5
$161^\circ - 180^\circ$	4	$341^\circ - 360^\circ$	8
		<i>Toplam</i>	75

Çizelge 2.1.'deki verilere ilişkin doğrusal histogram Şekil 2.3.'de gösterilmektedir.



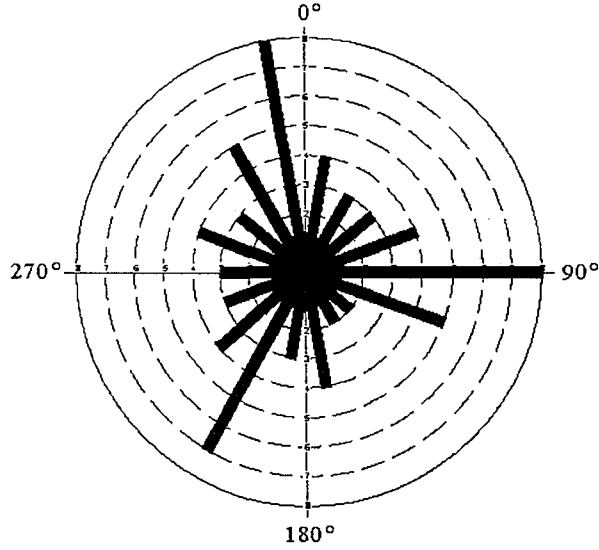
Şekil 2.3. Galatasaray futbol takımının 2001-2002 futbol sezonunda Türkiye Birinci Süper Ligi'nde attığı gollerin atıldığı dakikalara ilişkin doğrusal histogram

2.3.2. Dairesel histogram

Basit bir dairesel histogram bir çemberin çevresine doğrusal histogramın sarılmasıyla elde edilmektedir. İlk önce birim çember çizilir ve bu çemberin çevresine kendi sınıf aralığındaki sıklık miktarıyla orantılı alana sahip olan bir blok yerleştirilir. Histogramdaki her bir blok, kendi grup aralığının orta noktasında bulunmaktadır. Ayrıca, blok tepelerinin orta noktalarının birleştirilmesi ile *sıklık poligonu* da oluşturulabilmektedir.

Örnek 2.4.

Çizelge 2.1.'de kullanılan veri seti için dairesel histogram gösterimi Şekil 2.4.'te verilmektedir.



Şekil 2.4. Galatasaray futbol takımının 2001-2002 futbol sezonunda Türkiye Birinci Süper Ligi'nde attığı gollerin atıldığı dakikalara ilişkin dairesel histogram

2.3.3. Gül şeması

Diğer bir gösterim türü de gül şemasıdır. Tepe noktası orijin olmak üzere bir daire dilimi oluşturulduktan sonra, sınıf aralığının iki ucu bir yay ile birleştirilmektedir. Dolayısıyla, her grup bir daire dilimi biçiminde gösterilmiş olmaktadır. Her dilimin yarıçapı, grubun göreceli sıklığının kareköküyle orantılı olacak şekilde alınmaktadır. Böylece dilimin alanı grup frekansıyla orantılı olmaktadır. Çünkü, her bir daire diliminin alanı sıklığın karesine göre değişim göstermektedir. Sıklıkların yerine sıklıkların karekökleriyle orantılı alanlar elde edebilmek için, sıklıkların karekökleri yarıçap olarak alınmalıdır.

Örnek 2.5.

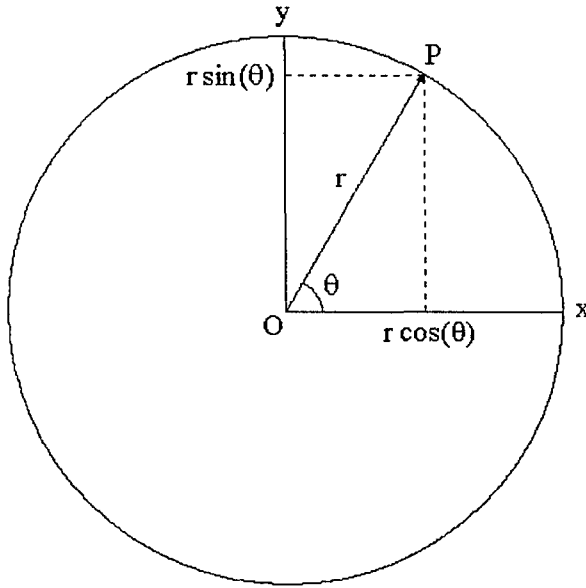
Çizelge 2.1.'de kullanılan veri seti için gül şeması gösterimi Şekil 2.5.'te verilmektedir. Burada her bir dilimin yarıçapı grupların göreceli sıklıklarının kareköküyle orantılıdır ve dairesel genişlik olarak 20° (5 dakika) kullanılmaktadır.

- Histogram kullanımında, göze ya da grup sınırlarının keyfi olarak seçilmesi gerekmektedir. Uygun olmayan bir sınırın seçilmesi ise, mod gruplarının sayısı, büyüklüğü ve konumu bakımından örneklem bilgisinde önemli farklılıklara neden olmaktadır.
- Yapılacak diğer bir seçim her bir gözenin genişliği seçimidir. Gül şemasının, temel dağılımın şeklini açıklamadaki yetersizliğine ilişkin fikirler göz önüne alınarak, bu tip gösterimin gerekli olduğu durumlarda, 10° ya da 20° 'lik (veya çok küçük veri setleri için mümkün olduğu kadar fazla) keyfi bir genişlik kullanılması önerilmektedir (Fisher 1993).

3. DAİRESEL VERİLER İÇİN TANIMLAYICI İSTATİSTİKLER

Bu bölümde, anakütleden elde edilen örneklem yardımıyla bu anakütle hakkında bilgi edinilmesi amacıyla dairesel veriler için bazı temel parametre değerleri tanımlanacaktır.

Dairesel veriler, bir birim çemberin çevresi üzerinde açılar ya da noktalar biçiminde gösterilebilmektedir. Dairesel durum yalnızca iki koordinat değeriyle belirlenebilmektedir. Bu amaçla, O orijinli ve O 'ya göre birbirine dik X,Y eksenli dik koordinat sistemi kullanılabilir. Düzlemdeki herhangi bir P noktası, (X,Y) dik koordinatları ile ya da (r,θ) kutupsal koordinatları ile gösterilebilmektedir. Burada r ; P noktasının orijine olan uzaklığını, θ ise; P noktasının yönünü göstermektedir. Özel bir durum olarak, O noktası için $r = 0$ 'dır ve bu noktanın yönü belirtilemez, yani θ tanımsızdır.



Şekil 3.1. Dik ve kutupsal koordinatlar arasındaki ilişki

Kutupsal koordinatlar ve dik koordinatların birbirlerine dönüşümü işlemi *sinüs* ve *kosinüs* trigonometrik fonksiyonları yardımıyla yapılmaktadır. (r,θ) kutupsal koordinatlarına sahip P noktasını ele alalım. Şekil 3.1.'e göre, P noktasının dik koordinatları;

$$x = r \cos\theta \quad , \quad y = r \sin\theta \quad (3.1)$$

ile verilir. Daha önce belirtildiği gibi, dairesel veri analizinde vektörün uzunluğu ile değil, yönüyle ilgilenilecek ve kullanım kolaylığı açısından bu vektörler birim uzunlukta ($r = 1$) ele alınacaktır. Böylece her bir yöne, birim çemberin çevresi üzerinde bir P noktası karşılık gelecektir. Bu nokta, alternatif olarak yalnızca açıyla da belirtilebilmektedir. Buna göre, eğer P noktası birim çemberin çevresi üzerinde bir nokta ise, kutupsal ve dik koordinatlar arasındaki dönüşüm;

$$(1, \theta) \Leftrightarrow (x = \cos\theta, y = \sin\theta) \quad (3.2)$$

biçiminde yapılabilmektedir.

Burada belirtilmesi gereken önemli bir nokta: $\{u_i\} = \{(x_i, y_i)\}$ olarak ifade edilmesine rağmen, bu şekildeki veriler düzlem üzerinde iki-değişkenli veri olarak düşünülmemelidirler. Çünkü, x_i ve y_i değerleri bütünüyle $x_i^2 + y_i^2 = 1, \forall i$ eşitliğine bağlıdır (Jammalamadaka ve SenGupta 2001 ve Bolton 1995).

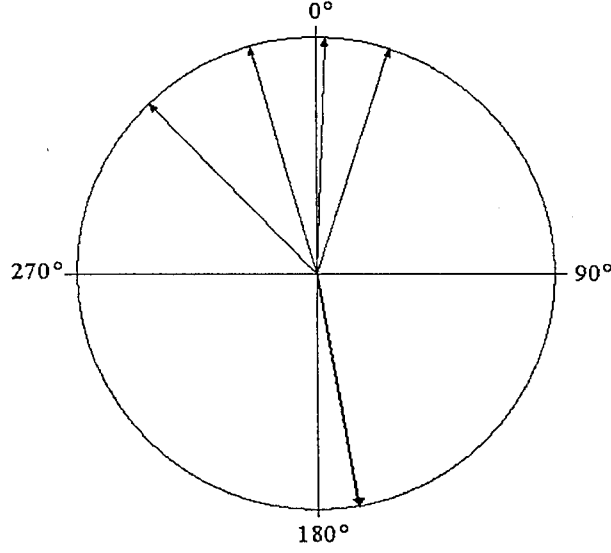
3.1. Ortalama Yön, Bileşke Uzunluğu ve Ortalama Bileşke Uzunluğu

Verilen tek modlu bir dairesel veri setinde ortalama yönü (tercih edilen yön) belirleyebilmek için, açıların aritmetik ortalaması hesaplanmak istenebilir. Fakat örneklem ortalaması dairesel veriler için sıfır yönü ve dönüş doğrultusu seçimine güçlü derecede bağımlıdır. Aşağıdaki örnekte dairesel veriler için tanımlayıcı merkezi ölçü olarak bilinen aritmetik ortalamanın neden uygun olmadığı açıklanmaktadır.

Örnek 3.1

Bir araştırmacının elinde 4 tane kaplumbağa bulunmaktadır. Bu araştırmacı, Kuzey yönünü sıfır yönü olarak kabul etmiş ve saat yönünde bir pozitif dönüş doğrultusu izleyerek, elindeki kaplumbağaları serbest bıraktıktan sonraki gidiş açılarını (yönlerini) kaydetmiştir.

Kaplumbağa :	1	2	3	4
Yön :	18°	2°	344°	316°



Şekil 3.2. Aritmetik ortalamamanın yanlış yönde oluşu

Klasik istatistik teknikleri kullanılarak kaplumbağaların araştırmacıdan ortalama olarak hangi yönde uzaklaştıkları bulunmak istendiğinde, derecelerin toplamını alıp terim sayısına bölmek belirli bir anlam ifade etmeyecektir. Şekil 3.2.'ye bakıldığında, bu örnek için klasik ortalama mantığı ile 170° bulunmaktadır. Bu sonuç veri setiyle karşılaştırıldığında, bulunan değer kaplumbağaların gidiş yönünün tersinde bir yöne karşılık gelmektedir (Şekil 3.2.'de koyu olarak belirtilen vektör). Ayrıca belirtilmesi gereken bir nokta da; aynı veri seti için örneklem genel aritmetik ortalaması, sıfır noktası ve dönüş doğrultusunun seçimine bağlı olarak da farklılık gösterebilmektedir. Örnekte de görüldüğü gibi, verilen bir dairesel veri seti için, doğrusal verilerde genel olarak kullanılan aritmetik ortalama 'merkezi' bir ölçü değildir, sıfır yönü ve dönüş doğrultusu seçiminin bir fonksiyonudur. Bu nedenle dairesel veriler için bir merkez ölçüsü olarak kullanılmasından kaçınılmalıdır. Örneklem ortalamasına bağlı örneklem varyansı s^2 'de de aynı problemle karşılaşmaktadır. Bu da dairesel veriler için daha uygun tanımlayıcı istatistiklerin tanımlanması gerektiğini göstermektedir.

Tek modu bulunan ya da tek bir yöne doğru kümeleşme gösteren bir dairesel veri seti için uygun ve anlamlı bir ortalama yön ölçüsü, gözlem değerleri birim vektörler olarak düşünülerek ve bu vektörlerin bileşke vektörünün yönü

kullanılarak belirlenmektedir. Verilen bir dairesel veri seti için bileşke vektörün yönünün yani *ortalama yön*ün hesaplanması aşağıda verilmektedir:

P_i ; birim çember üzerinde θ_i ($i = 1, \dots, n$) açısına bağlı olarak belirlenen herhangi bir nokta olmak üzere, $\theta_1, \dots, \theta_n$ açılarının ortalama yönü $\bar{\theta}$; $\overline{OP_1}, \dots, \overline{OP_n}$ birim vektörlerinin bileşkesinin yönü olarak tanımlanmaktadır. P_i noktalarının ağırlık merkezi (C, S) ;

$$C = \sum_{i=1}^n \cos \theta_i, \quad S = \sum_{i=1}^n \sin \theta_i$$

olmak üzere, *bileşke uzunluğu* R ;

$$R = \sqrt{C^2 + S^2} \quad (3.3)$$

eşitliğinden, veya ortalama yön biliniyorsa

$$R = \frac{C}{\cos \bar{\theta}} = \frac{S}{\sin \bar{\theta}} \quad (3.4)$$

eşitliği yardımıyla bulunmaktadır ve $\theta_1, \dots, \theta_n$ açılarının vektör bileşkesinin yönü $\bar{\theta}$;

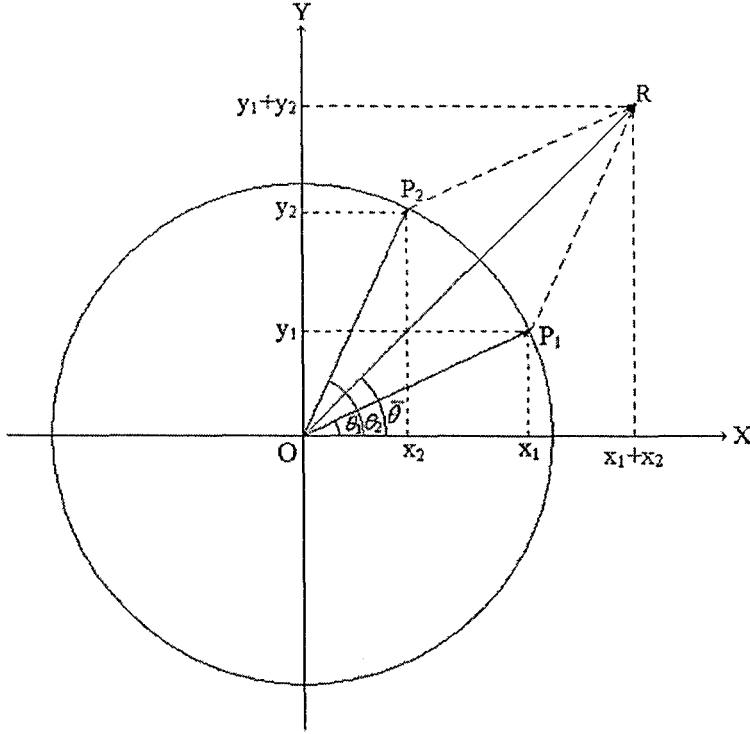
$$\bar{\theta} = \begin{cases} \tan^{-1}(S/C) & , \quad S \geq 0, C > 0 \\ \tan^{-1}(S/C) + \pi & , \quad C < 0 \\ \tan^{-1}(S/C) + 2\pi & , \quad S < 0, C \geq 0 \\ \pi/2 & , \quad S > 0, C = 0 \\ \text{tanımsız} & , \quad S = 0, C = 0 \end{cases} \quad (3.5)$$

biçiminde tanımlanmakta ve *ortalama yön* olarak adlandırılmaktadır. Burada R miktarı $(0, n)$ aralığında değerler almaktadır. *Ortalama bileşke uzunluğu* \bar{R} ise ortalama yön $\bar{\theta}$ ile ilişkilidir ve

$$\bar{R} = R/n \quad (3.6)$$

biçiminde tanımlanır. Ortalama bileşke uzunluğu, (0,1) aralığında değerler almaktadır.

Bu eşitliklerin elde edilebilmesi amacıyla, iki tane dairesel gözlem için bileşke vektörün uzunluğu ve yönü bulunduktan sonra, sonuç genellenerek n tane dairesel gözlem değeri için genel formüller elde edilebilir.



Şekil 3.3. İki dairesel gözlem için bileşke vektör ve ortalama yön

Birim çember üzerinde bulunan herhangi iki P_1 ve P_2 noktasının koordinatları sırasıyla (x_1, y_1) ve (x_2, y_2) ; X eksenine saat yönünün tersi doğrultuda yaptıkları açılar ise θ_1 ve θ_2 olsun. Buna göre r ; vektörlerin boyu yani birim çemberin yarıçapı ($r=1$) olmak üzere;

$$x_1 = r \cos \theta_1 = \cos \theta_1, x_2 = r \cos \theta_2 = \cos \theta_2,$$

$$y_1 = r \sin \theta_1 = \sin \theta_1, y_2 = r \sin \theta_2 = \sin \theta_2$$

yazılabilir. İki vektörün birbirine eklenmesiyle oluşturulan bileşke vektör R 'nin uzunluğu;

$$R = \sqrt{(x_1 + x_2)^2 + (y_1 + y_2)^2}$$

$$= \sqrt{\left(\sum_{i=1}^2 \cos \theta_i\right)^2 + \left(\sum_{i=1}^2 \sin \theta_i\right)^2}$$

olarak elde edilir. Bileşke vektörün yönünü gösteren $\bar{\theta}$ ise;

$$\cos \bar{\theta} = \frac{x_1 + x_2}{\sqrt{(x_1 + x_2)^2 + (y_1 + y_2)^2}}, \quad \sin \bar{\theta} = \frac{y_1 + y_2}{\sqrt{(x_1 + x_2)^2 + (y_1 + y_2)^2}}$$

eşitlikleri yardımıyla elde edilir (Bolton 1995).

Buna göre değerler yerine konulduğunda;

$$\cos \bar{\theta} = \frac{\sum_{i=1}^2 \cos \theta_i}{\sqrt{\left(\sum_{i=1}^2 \cos \theta_i\right)^2 + \left(\sum_{i=1}^2 \sin \theta_i\right)^2}},$$

$$\sin \bar{\theta} = \frac{\sum_{i=1}^2 \sin \theta_i}{\sqrt{\left(\sum_{i=1}^2 \cos \theta_i\right)^2 + \left(\sum_{i=1}^2 \sin \theta_i\right)^2}}$$

Buradan;

$$\tan \bar{\theta} = \frac{\sin \bar{\theta}}{\cos \bar{\theta}} = \frac{\sum_{i=1}^2 \sin \theta_i}{\sum_{i=1}^2 \cos \theta_i}$$

bulunur. Sonuç olarak;

$$\bar{\theta} = \tan^{-1} \left(\frac{\sum_{i=1}^n \sin \theta_i}{\sum_{i=1}^n \cos \theta_i} \right)$$

elde edilir. Bundan sonra, birim çember üzerinde n sayıda gözlem değeri olduğu durum incelenirse;

$$x_i = \cos \theta_i, y_i = \sin \theta_i, i = 1, \dots, n \text{ olmak üzere}$$

bileşke vektör R 'nin uzunluğu;

$$R = \sqrt{\left(\sum_{i=1}^n \cos \theta_i \right)^2 + \left(\sum_{i=1}^n \sin \theta_i \right)^2} = \sqrt{C^2 + S^2}$$

elde edilir. Böylece Eşitlik (3.3) elde edilir. Ortalama yön $\bar{\theta}$ ise;

$$\cos \bar{\theta} = \frac{\sum_{i=1}^n \cos \theta_i}{\sqrt{\left(\sum_{i=1}^n \cos \theta_i \right)^2 + \left(\sum_{i=1}^n \sin \theta_i \right)^2}} = \frac{C}{\sqrt{C^2 + S^2}} \quad (3.7)$$

$$\sin \bar{\theta} = \frac{\sum_{i=1}^n \sin \theta_i}{\sqrt{\left(\sum_{i=1}^n \cos \theta_i \right)^2 + \left(\sum_{i=1}^n \sin \theta_i \right)^2}} = \frac{S}{\sqrt{C^2 + S^2}} \quad (3.8)$$

eşitlikleri yardımıyla bulunur. Sonuç olarak Eşitlik (3.5)

$$\tan \bar{\theta} = \frac{S}{C} \Rightarrow \bar{\theta} = \tan^{-1} \left(\frac{S}{C} \right)$$

elde edilir. Ayrıca Eşitlik (3.3), (3.7) ve (3.8)'den, Eşitlik (3.4) sonucuna da ulaşılabilir.

Dolayısıyla, verilen bir açı kümesi için, ilk olarak bu değerler dik koordinatlara dönüştürülür, yani her bir açının kosinüs ve sinüs değeri hesaplanır. Bileşke vektörün bulunabilmesi için bu değerler toplanır. Böylece bu bileşke vektörün yönü olan $\bar{\theta}$, Eşitlik (3.5) kullanılarak elde edilmektedir. Bu değer verilen bir açı kümesi için dairesel ortalama yönü göstermektedir.

Ayrıca, $R = 0$ durumunda, yani $C = 0$ ve $S = 0$ iken, dairesel ortalama tanımsızdır. Daha geniş bir ifadeyle, eğer bileşke vektör pozitif uzunlukta ise, bu vektörün yönü olan $\bar{\theta}$, dairesel ortalama yön olarak alınmaktadır. Eğer bileşke vektör 0 uzunluğunda ise, ortalama yön bulunmamaktadır. $R = n$ olması ise bütün veri noktalarının çakıştığı anlamına gelmektedir.

Bileşke vektörün sıfır uzunluğa sahip olması durumu, verinin çember üzerinde herhangi bir yöne doğru yoğunlaşma göstermediğini ve *düzgün* dağıldığını belirtmektedir. Bu durumda, verinin herhangi bir tercih edilen yönü ya da ortalama yönü bulunmamaktadır.

Örnek 3.2

Bileşke uzunluğunun, ortalama bileşke uzunluğunun ve ortalama yönün Örnek 3.1'deki veriler için hesaplanması:

θ_i	$\cos \theta_i$	$\sin \theta_i$
18	0.95106	0.30902
2	0.99939	0.03490
344	0.96126	-0.27564
316	0.71934	-0.69466
Toplam	$C = 3.63105$	$S = -0.62638$

Buna göre bileşke uzunluğu;

$$R = \sqrt{C^2 + S^2} = \sqrt{(3.63105)^2 + (-0.62638)^2} = 3.68468$$

ortalama bileşke uzunluğu;

$$\bar{R} = R/n = 3.68468/4 = 0.92117$$

ortalama yön ise $S < 0$ ve $C > 0$ olduğundan Eşitlik (3.5)'e göre;

$$\begin{aligned}\bar{\theta} &= \tan^{-1}(S / C) + 2\pi = \tan^{-1}(-0.62638 / 3.63105) + 2\pi \\ &= \tan^{-1}(-0.17251) + 2\pi = -9.79 + 360 = 350.21^\circ\end{aligned}$$

biçiminde elde edilir.

Gruplanmış verilerde ise genel bir yaklaşım olarak, bir aralıktaki tüm gözlem değerlerinin o aralığın orta noktası olduğu varsayımı kullanılmaktadır. Örneğin, n tane orjinal gözlem değeri k tane sınıfa göre gruplanmış olsun. i -inci sınıfın orta noktası θ_i ve sıklığı f_i , $i = 1, \dots, k$; $\left(\sum_{i=1}^k f_i \right) = n$ olmak üzere;

$$\bar{C}_n = \frac{1}{n} \sum_{i=1}^k f_i \cos \theta_i, \quad \bar{S}_n = \frac{1}{n} \sum_{i=1}^k f_i \sin \theta_i, \quad \bar{R} = \sqrt{\bar{C}^2 + \bar{S}^2} \quad (3.9)$$

Eşitlik (3.9)'da verilen değerler hesaplanır.

Örnek 3.3

Ortalama bileşke uzunluğunun ve ortalama yönün Örnek 2.3'de verilen gruplanmış veri seti için hesaplanması:

Ortalama bileşke uzunluğu ve ortalama yön için gerekli hesaplamalar Çizelge 3.1.'de görülmektedir. Çizelge 3.1. kullanılarak;

$$\bar{C} = 4.0580 / 75 = 0.0541,$$

$$\bar{S} = 1.3266 / 75 = 0.0177,$$

$$\bar{R} = \sqrt{\bar{C}^2 + \bar{S}^2} = \sqrt{(0.0541)^2 + (0.0177)^2} = 0.0569$$

$\bar{S} > 0$ ve $\bar{C} > 0$ olduğundan Eşitlik (3.5)'e göre;

$$\bar{\theta} = \tan^{-1}(\bar{S} / \bar{C}) = \tan^{-1}(0.0177 / 0.0541)$$

$$= \tan^{-1}(0.3272) = 18.12^\circ$$

sonuçlarına ulaşılır.

Çizelge 3.1. Çizelge 2.1.'de verilen veri seti için ortalama bileşke uzunluğu ve ortalama yönün hesaplanması

θ_i (Orta Nokta)	f_i	$\cos \theta_i$	$f_i \cos \theta_i$	$\sin \theta_i$	$f_i \sin \theta_i$
10°	4	0.9848	3.9392	0.1736	0.6944
30°	3	0.8660	2.5980	0.5000	1.5000
50°	3	0.6428	1.9284	0.7660	2.2980
70°	4	0.3420	1.3680	0.9397	3.7588
90°	8	0.0000	0.0000	1.0000	8.0000
110°	5	-0.3420	-1.7100	0.9397	4.6985
130°	2	-0.6428	-1.2856	0.7660	1.5320
150°	2	-0.8660	-1.7320	0.5000	1.0000
170°	4	-0.9848	-3.9392	0.1736	0.6944
190°	3	-0.9848	-2.9544	-0.1736	-0.5208
210°	7	-0.8660	-6.0620	-0.5000	-3.5000
230°	4	-0.6428	-2.5712	-0.7660	-3.0640
250°	3	-0.3420	-1.0260	-0.9397	-2.8191
270°	3	0.0000	0.0000	-1.0000	-3.0000
290°	4	0.3420	1.3680	-0.9397	-3.7588
310°	3	0.6428	1.9284	-0.7660	-2.2980
330°	5	0.8660	4.3300	-0.5000	-2.5000
350°	8	0.9848	7.8784	-0.1736	-1.3888
Toplam	75		4.0580		1.3266

3.2. Yoğunlaşma Parametresi

Yoğunlaşma parametresi κ^2 'nin en çok olabilirlik tahmini $\hat{\kappa}$;

$$A_1(\hat{\kappa}) = R/n = \bar{R} \quad (3.10)$$

Eşitlik (3.10)'un çözümüdür. Burada \bar{R} ; ortalama bileşke uzunluğunu göstermektedir ve

$$A_1(x) = I_1(x) / I_0(x) \quad (3.11)$$

dönüştürülmüş iki Bessel fonksiyonunun oranıdır. Eşitlik (3.10)'un çözümü için uygun bir yaklaşım;

$$\hat{\kappa} = \begin{cases} 2\bar{R} + \bar{R}^3 + 5\bar{R}^5 / 6 & \bar{R} < 0.53 \\ -0.4 + 1.39\bar{R} + 0.43/(1 - \bar{R}) & 0.53 \leq \bar{R} < 0.85 \\ 1/(\bar{R}^3 - 4\bar{R}^2 + 3\bar{R}) & \bar{R} \geq 0.85 \end{cases} \quad (3.12)$$

ile verilmektedir (Fisher 1993). EK-1'de von Mises dağılımında verilen \bar{R} değerleri için $\hat{\kappa}$ en çok olabilirlik tahminleri ve EK-2'de von Mises dağılımında verilen κ değerleri için $\rho = A(\kappa)$ bileşke uzunlukları verilmektedir. Von Mises dağılımı ise Kesim 4.2.'de ayrıntılı olarak incelenecektir.

Örnek 3.4

Yoğunlaşma parametresinin, Örnek 3.1'de verilen veri seti için tahmin edilmesi:

Örnek 3.1'de ortalama bileşke uzunluğu, $\bar{R} = 0.92117$ olarak bulunmuştu. Eşitlik (3.12)'ye göre; $\bar{R} \geq 0.85$ olduğu için,

$$\hat{\kappa} = \frac{1}{(0.92117)^3 - 4(0.92117)^2 + 3(0.92117)} = 6.62445$$

olarak elde edilir.

3.3. Dairesel Varyans

Vektör bileşkesi R 'nin yönü olan $\bar{\theta}$ değerinin ortalama yönü vermesinin yanısıra, tek-modlu veriler için diğer bir yararlı ölçü de verinin bu merkez etrafındaki yoğunlaşma miktarını gösteren bileşke vektör uzunluğu R 'dir. Bütün gözlem noktaları (birim vektörler) aynı yönde büyük bir yoğunlaşma gösteriyor ise, R değerinin büyüklüğü n kadar olacaktır. Bu durumun tersine, veri noktaları herhangi bir yoğunlaşma göstermeden çember üzerinde düzgün bir dağılım gösteriyor ise, R değeri sıfıra çok yakın bir değer alacaktır.

Dairesel varyansın tanımlanabilmesi için öncelikle dairesel uzaklık kavramının açıklanması gerekmektedir.

P_i , birim çember üzerinde θ_i açılı bir nokta ve α sabit bir yön olsun. İlk olarak $\alpha = 0$ olduğu varsayalım ve P de çember üzerinde buna karşılık gelen nokta olsun. P ve P_i arasındaki dairesel saçılımın bir ölçüsü OP_i 'nin OP ile yaptığı iki açıdan küçük olanıdır. Bu ölçüye *daireysel uzaklık* (ξ) adı verilir ve aşağıdaki gibi elde edilir.

$$\xi_i = \min(\theta_i, 2\pi - \theta_i) = \pi - |\pi - \theta_i| \quad (3.13)$$

Buna göre, çember çevresi üzerinde birbirinden π değerinden daha uzak iki nokta bulunamayacağı ya da Eşitlik (3.13)'da verilen dairesel uzaklığın her zaman $[0, \pi]$ aralığında olacağı açıktır.

$1 - \cos \xi_i$, ξ_i 'nin monoton artan bir fonksiyonu olduğu için,

$$D = \frac{1}{n} \sum_{i=1}^n (1 - \cos \xi_i) \quad (3.14)$$

Eşitlik (3.14) ifadesi, P_i noktalarının bir saçılım ölçüsü olarak alınabilmesidir.

Eşitlik (3.13) ve (3.14) kullanılarak, sıfır yönü α olarak değiştirildikten sonra,

$$D = \frac{1}{n} \sum_{i=1}^n \{1 - \cos(\theta_i - \alpha)\} \quad (3.15)$$

elde edilir. Eşitlik (3.15)'deki D saçılımı, $\alpha = \bar{\theta}$ için minimumdur. Bunun gösterilebilmesi için, Eşitlik (3.15)'in α 'ya göre türevi alınıp sifıra eşitlenmelidir.

Buna göre;

$$\sum_{i=1}^n \sin(\theta_i - \alpha) = 0$$

olur. Böylece $\bar{\theta}$ etrafındaki saçılım minimum olmaktadır. Bu minimizasyon özelliği bilinen örneklem varyansıninkine benzerdir. $\bar{\theta}$ etrafındaki D değeri V ile gösterildiğinde;

$$V = 1 - \frac{1}{n} \sum_{i=1}^n \cos(\theta_i - \bar{\theta})$$

olur (Mardia 1972). Eşitlik (3.4) kullanıldığında;

$$V = 1 - \frac{1}{n} \sum_{i=1}^n (\cos \theta_i \cos \bar{\theta} + \sin \theta_i \sin \bar{\theta})$$

$$V = 1 - \frac{1}{n} \left(\cos \bar{\theta} \sum_{i=1}^n \cos \theta_i + \sin \bar{\theta} \sum_{i=1}^n \sin \theta_i \right)$$

$$V = 1 - \frac{1}{n} (C \cos \bar{\theta} + S \sin \bar{\theta})$$

$$V = 1 - \frac{1}{n} (R \cos \bar{\theta} \cos \bar{\theta} + R \sin \bar{\theta} \sin \bar{\theta})$$

$$V = 1 - \frac{1}{n} R$$

$$V = 1 - \bar{R} \tag{3.16}$$

sonucuna ulaşılır. Eşitlik (3.16), *örneklem dairesel varyansı* olarak tanımlanır. Doğrusal veri varyansında olduğu gibi, dairesel varyans değeri küçüldükçe, dağılım homojenleşir. Fakat, bilinen varyanstan farklı olarak, dairesel varyansta $0 \leq V \leq 1$ 'dir. Ayrıca, $\bar{R} = 0$ ve $V = 1$ olması, maksimal bir dağılımın varlığı şeklinde yorumlanmamalıdır.

n birimlik dairesel gözlem $\bar{\theta}$ ortalama yönü etrafında kümelenmiş ise, \bar{R} ortalama bileşke uzunluğu 1'e yakın bir değer alacaktır. Dolayısıyla, dairesel varyans da sıfıra yakın olacaktır. Yönler geniş bir dağılım göstermiş ise, \bar{R} değeri küçük ve dairesel varyans 1'e yakın olacaktır. Ayrıca dairesel varyansın bir özelliği de, sıfır yönü değiştirilse bile dairesel varyans değerinin sabit kalmasıdır.

Örnek 3.5

Örnek 3.1'deki veriler için dairesel varyans;

$$V = 1 - \bar{R} = 1 - 0.92117 = 0.07883$$

olarak elde edilir. Bu değer, Eşitlik (2.1)'den, 4.52° 'ye karşılık gelmektedir. Bu sifıra yakın çıkan sonuç, kaplumbağaların seçtikleri gidiş yönlerinin homojen olduğunu, fazla değişim göstermediğini belirtmektedir.

3.4. Dairesel Standart Sapma

Doğru üzerindeki varyansın değişim aralığı $(0, \infty)$ iken, dairesel varyans $(0, 1)$ aralığında değerler almaktadır. V değerinin $(0, \infty)$ aralığına uygun bir dönüşümü;

$$\nu = \{-2 \log_e(1 - V)\}^{1/2} \quad (3.17)$$

biçiminde tanımlanmaktadır.

ν ölçüsü, doğru üzerindeki bilinen örneklem standart sapmasına benzemektedir. Bununla birlikte, V ve R değerleri kuramsal araştırmalarda ν değerinden daha kullanışlı olmaktadır.

V 'nin küçük değerleri için, Eşitlik (3.17)

$$\nu = (2V)^{1/2} \quad (3.18)$$

biçimindedir. Eşitlik (3.17)'te, θ 'nin değişim aralığının $(0, 2\pi)$ olduğu varsayılmaktadır. θ 'nin değişim aralığı $(0, 2\pi/l)$ ise, dairesel standart sapma;

$$\nu = \{-2 \log_e(1 - V)\}^{1/2} / l \quad (3.19)$$

biçiminde verilir (Mardia 1972).

Örnek 3.6

Örnek 3.1'deki veriler için dairesel standart sapma;

$$\begin{aligned} \nu &= \{-2 \log_e(1 - V)\}^{1/2} = \{-2 \log_e(1 - 0.07883)\}^{1/2} \\ &= 0.40524 \end{aligned}$$

olarak elde edilir. Bu değer, Eşitlik (2.1)'den 23.22° 'ye karşılık gelmektedir.

3.5. Dairesel Saçılım

Bir diğer yayılma ölçüsü de *dairesel saçılım*dır. Dairesel saçılım;

$$\delta = \frac{1 - \rho_2}{2\bar{R}^2} \quad (3.20)$$

biçiminde tanımlanmaktadır. Eşitlikte;

$$\rho_2 = \frac{1}{n} \sum_{i=1}^n \cos 2(\theta_i - \bar{\theta}) \quad (3.21)$$

değeri merkezi ikinci trigonometrik momenti göstermektedir.

Dairesel saçılımın, ortalama yön için güven aralığının hesaplanmasında, farklı örneklem ortalama yönlerinin karşılaştırılmasında ve birleştirilmesinde önemli bir yeri vardır (Fisher 1993).

Örnek 3.7

Örnek 3.1'deki veriler için önce merkezi ikinci trigonometrik momentin bulunması gerekmektedir. Ortalama yön 18.12° olarak bulunmuştur. Buna göre;

$$\begin{aligned} \rho_2 &= \frac{1}{4} \{ \cos[2(18 - 18.12)] + \cos[2(2 - 18.12)] + \cos[2(344 - 18.12)] \\ &\quad + \cos[2(316 - 18.12)] \} \\ &= 0.41347 \end{aligned}$$

olarak bulunur. Buna göre daireysel saçılım;

$$\delta = \frac{1 - 0.41347}{2(0.92117)^2} = 0.34561$$

olarak elde edilir.

3.6. Dairesel Standart Hata

Ortalama yön tahmininin *dairesel standart hatası*;

$$\sigma = \sqrt{\frac{\delta}{n}} \quad (3.22)$$

ile hesaplanmaktadır. Burada, dairesele saçılım değeri Eşitlik (3.20)'den bulunmaktadır (Fisher 1993). Dairesel standart hata özellikle, ortalama yön için güven aralıklarının belirlenmesinde kullanılmaktadır.

Örnek 3.8

Örnek 3.1'deki veriler için ortalama yönün dairesele standart hatası;

$$\sigma = \sqrt{\frac{0.34561}{4}} = 0.29394$$

olarak bulunur.

3.7. Medyan Yönü

Dairesel veri setlerinde örneklem medyanını bulabilmek için veriyi iki eşit gruba bölen bir eksen (medyan eksen) seçilir. Örneklem sayısı tek olduğunda medyan yönü veri noktasıyla çakışır, çift olduğunda ise iki nokta arasındaki orta nokta alınır.

Birim çember üzerinde bulunan noktalardan oluşan bir veri kümesi verildiği varsayalım. Verilen bu noktalardan herhangi bir P noktası aşağıdaki özellikleri sağlıyorsa bu P noktası örneklem medyanı olarak adlandırılır.

- (i) Örneklem noktalarının yarısı P noktasından geçen PQ çapının her iki tarafında bulunmaktadır.
- (ii) Örneklem noktalarının çoğunluğu P noktasına Q noktasından daha yakındır.

Burada \overline{OP} vektörü örneklem medyan yönü olarak adlandırılmaktadır.

Medyan yönü, $\tilde{\theta}$ simgesiyle gösterilmektedir ve

$$\int_{\tilde{\theta}}^{\tilde{\theta}+\pi} f(\theta) d\theta = \int_{\tilde{\theta}+\pi}^{\tilde{\theta}+2\pi} f(\theta) d\theta = \frac{1}{2} \quad (3.23)$$

biçiminde tanımlanan integralin çözümü ile bulunmaktadır. Buradaki $f(\theta)$ yoğunluk fonksiyonu

$$f(\tilde{\theta}) > f(\tilde{\theta} + \pi)$$

eşitsizliğini sağlamaktadır. Simetrik bir dağılım için, medyan yönü simetri ekseninde olacaktır. Medyan tek bir tane olmayabilir. Fakat, bütün tek modlu dağılımlar tek medyana sahiptir (Mardia 1972).

Gruplanmış bir seri için medyan yönü;

$$\tilde{\theta} = l + \frac{\frac{n}{2} - f_0}{f_{+1} - f_0} \times h \quad (3.24)$$

ile tanımlanır. Burada;

l : medyan sınıfının alt sınırını;

f_0 : medyan sınıfının $(\theta_i - 180^\circ, \theta_i)$ aralığındaki sıklığını;

f_{+1} : medyan sınıftan bir sonraki sınıfın $(\theta_i - 180^\circ, \theta_i)$ aralığındaki sıklığını ve

h : sınıf aralığının uzunluğunu göstermektedir.

Örnek 3.9

Medyan yönünün, Örnek 2.1'de verilen veriler için elde edilmesi:

$(0^\circ, 360^\circ)$ aralığına dönüştürülen gözlem değerleri: $20^\circ, 20^\circ, 52^\circ, 120^\circ, 128^\circ, 136^\circ, 168^\circ, 172^\circ, 180^\circ, 180^\circ, 188^\circ, 192^\circ, 236^\circ, 256^\circ, 328^\circ$. Bu noktaların medyan yönü, 15 gözlem değeri olduğu için 8. gözlem olan 172° olacaktır.

Örnek 3.10

Medyan yönünün, Örnek 2.3'de verilen gruplanmış veri seti için elde edilmesi:

Medyan yönü için gerekli hesaplamalar Çizelge 3.2.'de görülmektedir. Üçüncü sütunda birikimli sıklıklar görülmektedir. Bu sütunun ilk yarısı, $\theta_i - 180^\circ$ 'nin altındaki açı değerlerinin birikimli sıklıklarını gösteren dördüncü sütunda tekrar edilmektedir. Burada θ_i ; i -inci sınıf aralığının üst sınırını göstermektedir. $(\theta_i - 180^\circ, \theta_i)$ aralığındaki sıklık değerleri, üçüncü sütundaki değerlerden dördüncü sütundaki değerler çıkartılarak elde edilmektedir. Toplam sıklık, 75'in yarısı olan 37. gözlem beşinci sütuna göre, $(20^\circ, 200^\circ)$ ve $(40^\circ, 220^\circ)$ aralıklarına karşılık gelen sıklıklar içerisinde yer almaktadır. Sıklık sayısı çemberin alt yarısında daha yüksek olduğu için, medyan yönü $(20^\circ, 40^\circ)$ aralığı yerine $(200^\circ, 220^\circ)$ aralığında olmalıdır. $(20^\circ, 200^\circ)$ aralığında 34 gözlem ve $(40^\circ, 220^\circ)$ aralığında 38 gözlem bulunmaktadır. Buna göre medyan yönü;

$$\tilde{\theta} = 200 + \frac{37.5 - 34}{38 - 34} \times 20 = 217.5^\circ$$

olarak bulunur.

Çizelge 3.2. Çizelge 2.1.'de verilen veri seti için medyan yönünün hesaplanması

θ_i İçin Üst Sınıf Sınırı	Sıklık	Birikimli Sıklık	$\theta_i - 180^\circ$ Altındaki Sıklık	$(\theta_i - 180^\circ, \theta_i)$ Aralığındaki Sıklık	$(\theta_i - 180^\circ, \theta_i)$ Aralığı
20°	4	4			
40°	3	7			
60°	3	10			
80°	4	14			
100°	8	22			
120°	5	27			
140°	2	29			
160°	2	31			
180°	4	35	0	35	(0,180)
200°	3	38	4	34	(20,200)
220°	7	45	7	38	(40,220)
240°	4	49	10	39	(60,240)
260°	3	52	14	38	(80,260)
280°	3	55	22	33	(100,280)
300°	4	59	27	32	(120,300)
320°	3	62	29	33	(140,320)
340°	5	67	31	36	(160,340)
360°	8	75	35	40	(180,360)

3.8. Mod Yönü

Örneklem mod yönü $\tilde{\theta}$; verinin en çok yoğunlaştığı yön anlamına gelmektedir. Verilen bir sıklık dağılımının mod yönü, kesim noktası seçiminden sonra doğru üzerindeki süreçle aynı süreç uygulanarak elde edilmektedir. Dolayısıyla dağılımın merkezinde en yüksek yoğunlaşma meydana gelmiş olur. Sonuç olarak, gruplanmış bir seri için mod yönü;

$$\tilde{\theta} = l + \frac{f_0 - f_{-1}}{2f_0 - f_{-1} - f_{+1}} \times h \quad (3.25)$$

eşitliği yardımıyla bulunur. Burada;

l : mod sınıfının alt sınırını;

f_0 : mod sınıfının sıklığını;

f_{-1} : mod sınıfından bir önceki sınıfın sıklığını;

f_{+1} : mod sınıfından bir sonraki sınıfın sıklığını ve

h : sınıf aralığının uzunluğunu göstermektedir (Mardia 1972).

Örnek 3.11

Mod yönünün, Örnek 2.3'de verilen gruplanmış veri seti için elde edilmesi:

Çizelge 3.2.'ye göre serinin en yüksek sıklığı olan 8 değeri, iki tane mod sınıfı oluşturmaktadır. Bunun için doğrusal verilerde olduğu gibi, sınıflar örneğin üçer üçer birleştirilmelidir. Buna göre oluşturulan yeni çizelge aşağıda verilmektedir.

Çizelge 3.3. Çizelge 2.1.'de verilen veri seti için mod yönünün hesaplanması

θ_i için Üst Sınıf Sınırı	Sıklık
60°	10
120°	17
180°	8
240°	14
300°	10
360°	16

Çizelge 3.3.'e göre;

$$l = 60 , f_0 = 17 , f_{-1} = 10 , f_{+1} = 8 , h = 60$$

değerleri elde edilir. Buna göre mod yönü;

$$\tilde{\theta} = 60 + \frac{17-10}{34-10-8} \times 60 = 86.25^\circ \text{ olarak elde edilir.}$$

4. DAİRESEL VERİLERDE OLASILIK DAĞILIMLARI

Olasılık dağılımları istatistiksel analizde çok önemli bir yere sahiptir. Eğer dağılımdaki parametrelerin doğru tahmin edilmesiyle veri setine uygun bir olasılık dağılımı uydurulabilirse, parametre tahminleriyle belirlenen olasılık dağılımının özel biçimi kullanılarak veri setine ilişkin yeterince bilgi edinilebilmektedir.

Doğrusal, dairesel ve küresel olmak üzere üç tür verinin normal dağılım, von Mises dağılımı ve Fisher dağılımına sahip olduğu varsayalım. Bu dağılımlarda iki tür parametre bulunmaktadır. Birincisinde, konum ya da referans yönü, diğerinde ise konuma ilişkin saçılım tanımlanmaktadır. Normal dağılım için saçılım miktarı σ^2 varyansı ile ölçülmektedir. Gözlem değerleri fazla miktarda bir arada toplandığında σ^2 değeri 0'a yakın bir değer almakta ve σ^2 değeri büyüdükçe artan bir yayılma göstermektedir. Von Mises ve Fisher dağılımları için saçılım miktarı, yoğunlaşma parametresi κ ile ölçülmektedir. Burada, $\kappa = 0$ ise, düzgün dağılım söz konusudur ve κ değeri arttıkça referans yönüne olan yoğunlaşma artmaktadır.

Bu bölümde, literatürde önerilen diğer dairesel olasılık dağılımları da kısaca verilecektir.

Dairesel dağılım, toplam olasılığı bir birim çemberin çevresi üzerinde toplanmış olan bir olasılık dağılımıdır. Çevre üzerindeki her bir nokta bir yönü ifade etmektedir. Radyan cinsinden ölçülen bir θ dairesel rassal değişkeninin değişim aralığı, $[0, 2\pi)$ ya da $[-\pi, \pi)$ olarak tanımlanabilir.

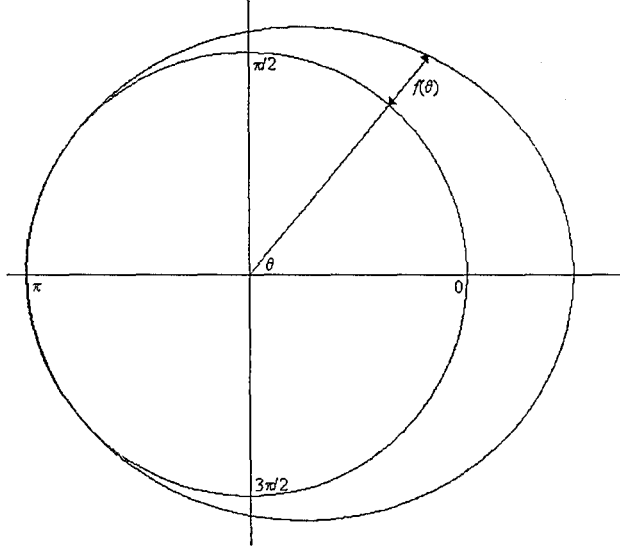
Temel olarak iki tür dairesel dağılım bulunmaktadır: *Kesikli* dağılımlarda; olasılık yığınları yalnızca sayılabilir sayıdaki yönlerle belirtilmektedir. İkinci durum olan *sürekli* dağılımlarda ise; $f(\theta)$ olasılık yoğunluk fonksiyonu bulunmaktadır ve bu fonksiyon aşağıdaki temel özelliklere sahiptir:

$$(i) f(\theta) \geq 0,$$

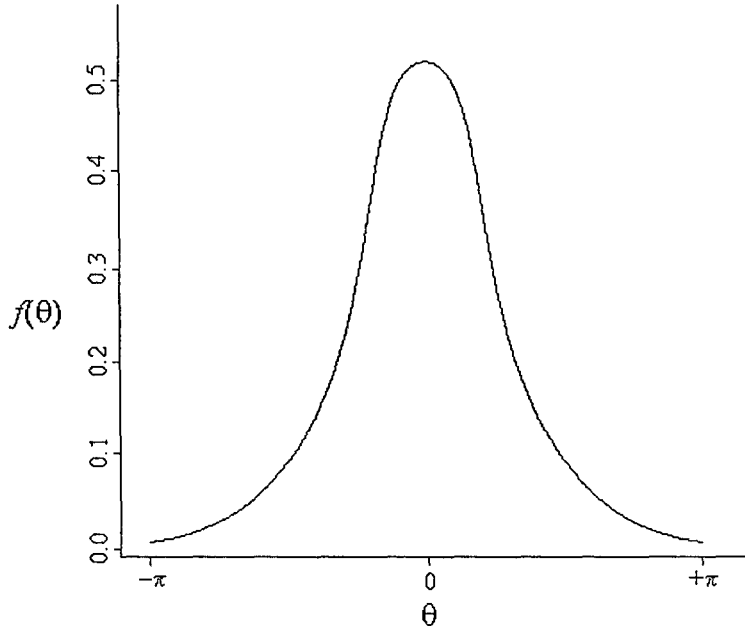
$$(ii) \int_0^{2\pi} f(\theta) d\theta = 1,$$

$$(iii) f(\theta) = f(\theta + k \cdot 2\pi), \text{ her } k \text{ tamsayısı için } (f \text{ fonksiyonu periyodiktir}).$$

Şekil 4.1. ve Şekil 4.2.'de sürekli bir dairesel dağılımın sırasıyla, dairesel ve doğrusal gösterimleri verilmektedir.



Şekil 4.1. Sürekli bir dairesel dağılımın dairesel gösterimi



Şekil 4.2. Sürekli bir dairesel dağılımın doğrusal gösterimi

Dairesel dağılımların simetri özelliği;

$$f(\mu + \theta) = f(\mu - \theta) , \forall \theta$$

biçiminde tanımlanır. θ rassal değişkeninin dağılımı, bu özelliğe sahipse, θ rassal değişkeni verilen bir μ ortalama yönüne göre simetriktir denilir. Özel olarak, sıfır yönüne göre simetri, yoğunluk fonksiyonunun, $\theta \rightarrow -\theta \pmod{2\pi}$ dönüşümü altında değişmemesi anlamına gelmektedir. Ayrıca, eğer bir dağılım $\theta = \mu$ yönüne göre simetrik ise, $\theta = \mu + \pi$ yönüne göre de simetriktir.

Dairesel bir rassal değişken ya θ , ($0 \leq \theta < 2\pi$) açısıyla, ya da iki-boyutlu birim vektörle ($X = \cos\theta, Y = \sin\theta$)' gösterilebilmektedir.

Gerçek doğru ya da düzlem üzerindeki bilinen olasılık dağılımları yardımıyla çok sayıda dairesel dağılım türetilmektedir. Bu yöntemlerden bazıları;

- (i) Bilinen bir dağılımın birim çember etrafına sarılması,
- (ii) Maksimum entropi vb. özelliklerin tanımlanması,
- (iii) İki-değişkenli doğrusal bir rassal değişkenin kendi dairesel bileşenine dönüştürülmesidir. Bu yöntemle oluşturulan dağılımlara *offset dağılımlar* adı verilmektedir.

(Jammalamadaka ve SenGupta 2001).

4.1. Çember Üzerindeki Olasılık Dağılımları

Bu kesimde, dairesel veri analizinde karşılaşılan temel olasılık dağılımları kısaca verilecek, en genel dağılım olan von Mises dağılımı ayrıntılı olarak incelenecektir.

4.1.1. Düzgün dağılım

Toplam olasılık çemberin çevresi üzerinde düzgün olarak yayılma gösteriyor ise dairesel *düzgün dağılım* söz konusudur. Bu dağılımda, 0° ile 360° arasındaki bütün yönler eşit olasılıklıdır. Dolayısıyla, dağılım izotropik ya da rassal dağılım olarak bilinmektedir. Bu dağılımın tanımlı bir ortalaması ya da tercih edilen bir yönü bulunmamaktadır. Ayrıca, herhangi bir yöne doğru yoğunlaşma olmadığından, varyans değeri 1'e eşittir.

Örneklem bileşke vektörünün uzunluğu olan R 'nin ve yönü olan $\bar{\theta}$ 'in birbirinden bağımsız olabilmesi için örneklemin düzgün dağılımdan seçilmiş olması gerekmektedir.

Dağılımın düzgün olmadığı durumda, bir ya da birden fazla tercih edilen yöne doğru yoğunlaşma olduğu düşünülebilir. Tercih edilen yön ya da mod bir tane ise, dağılım *tek-modlu* bir dağılımdır.

Olasılık yoğunluk fonksiyonu;

$$f(\theta) = \frac{1}{2\pi} \quad , \quad 0 \leq \theta < 2\pi \quad (4.1)$$

Dağılım Fonksiyonu;

$$F(\theta) = \frac{\theta}{2\pi} \quad , \quad 0 \leq \theta < 2\pi \quad (4.2)$$

Momentler;

Ortalama yön : μ , tanımsızdır

Ortalama bileşke uzunluğu : $\rho = 0$

Dairesel saçılım : $\delta = \infty$

$$\alpha_p = 0, p \geq 1$$

$$\beta_p = 0, p \geq 1$$

Burada, α_p ve β_p değerleri sırasıyla p-inci kosinüs ve sinüs momentleridir.

Dairesel veriler için düzgün dağılımın en önemli özelliği, test edilebilen çeşitli alternatif dağılım (tek-modlu, çok-modlu) türlerine karşı sıfır modeli

olarak kullanılmasıdır. Ayrıca, ortalama yön bulunmaması durumundan dolayı, düzgün dağılım, dairesel veri analizinde temel bir rol oynamaktadır.

4.1.2. Cardioid dağılım

Cardioid dağılım, iki parametrelili, tek-modlu ve μ 'ye göre simetrik bir dağılımdır. Bazen *Kosinüs dağılımı* olarak da ifade edilmektedir.

Olasılık yoğunluk fonksiyonu;

$$f(\theta) = \frac{1}{2\pi} \{1 + 2\rho \cos(\theta - \mu)\} \quad , \quad 0 \leq \theta < 2\pi \quad , \quad -1/2 < \rho < 1/2 \quad (4.3)$$

Dağılım Fonksiyonu;

$$F(\theta) = \frac{\rho}{\pi} \sin(\theta - \mu) + \frac{\theta}{2\pi} \quad , \quad 0 \leq \theta < 2\pi \quad (4.4)$$

Momentler;

Ortalama yön	:	μ
Ortalama bileşke uzunluğu	:	$\rho (< 1/2)$
Dairesel saçılım	:	$\delta = \frac{1}{2\rho^2}$
		$\alpha_p = 0, p \geq 2$
		$\beta_p = 0, p \geq 1$

$\rho \rightarrow 0$ iken, Cardioid dağılım düzgün dağılıma yakınsamaktadır.

4.1.3. Genel sarmal dağılımlar

X , gerçek doğru üzerindeki herhangi bir rassal değişken ise, $g(x)$ olasılık yoğunluk fonksiyonu ve $G(x)$ dağılım fonksiyonu olmak üzere;

$$\Theta \equiv X [\text{mod } 2\pi] \quad (4.5)$$

tanımlanarak, Θ dairesel rassal değişkeni elde edilebilmektedir. Θ 'nın olasılık yoğunluk fonksiyonu $f(\theta)$, birim yarıçaplı bir çemberin çevresi etrafına $g(x)$ fonksiyonu sarılarak elde edilmektedir;

$$f(\theta) = \sum_{k=-\infty}^{\infty} g(\theta + 2k\pi) \quad , \quad 0 \leq \theta < 2\pi \quad (4.6)$$

ve dağılım fonksiyonu;

$$F(\theta) = \sum_{k=-\infty}^{\infty} [G(\theta + 2k\pi) - G(2k\pi)] \quad , \quad 0 \leq \theta < 2\pi \quad (4.7)$$

olacaktır.

Çember üzerinde birçok sayıda dağılım bu yöntemle elde edilebilmektedir (Fisher 1993).

4.1.4. Sarmal Cauchy dağılımı

Sarmal Cauchy dağılımı, tek-modlu ve simetrik bir dağılımdır. Doğru üzerindeki Cauchy dağılımının çember etrafına sarılmasıyla elde edilmektedir.

Olasılık yoğunluk fonksiyonu;

$$f(\theta) = \frac{1}{2\pi} \frac{1 - \rho^2}{1 + \rho^2 - 2\rho \cos(\theta - \mu)} \quad , \quad 0 \leq \theta < 2\pi \quad , \quad 0 \leq \rho \leq 1 \quad (4.8)$$

Burada; $\rho = e^{-\sigma}$ 'dır.

Dağılım fonksiyonu;

$$F(\theta) = \frac{1}{2\pi} \cos^{-1} \left(\frac{(1 + \rho^2) \cos(\theta - \mu) - 2\rho}{1 + \rho^2 - 2\rho \cos(\theta - \mu)} \right) \quad , \quad 0 \leq \theta < 2\pi \quad (4.9)$$

Momentler;

Ortalama yön : μ

Ortalama bileşke uzunluğu : ρ

Dairesel saçılım : $\delta = \frac{1 - \rho^2}{2\rho^2}$

$$\alpha_p = \rho^p$$

$$\beta_p = 0, p \geq 1$$

$\rho \rightarrow 0$ iken, sarmal Cauchy dağılımı düzgün dağılıma yakınsamaktadır. $\rho \rightarrow 1$ iken, sarmal Cauchy dağılımı μ yönünde yoğunlaşan nokta dağılıma yakınsamaktadır.

Uygun olarak seçilen saçılım parametre değerleri için, sarmal Cauchy dağılımı; sarmal normal ve von Mises dağılımına oldukça benzemektedir (Fisher 1993).

4.1.5. Sarmal normal dağılım

Sarmal normal dağılım, iki parametrelili, tek-modlu ve $\theta = \mu'$ ye göre simetrik bir dağılımdır. Doğru üzerindeki normal dağılımın çember etrafına sarılmasıyla elde edilmektedir.

Olasılık yoğunluk fonksiyonu;

$$f(\theta) = \frac{1}{2\pi} \left(1 + 2 \sum_{p=1}^{\infty} \rho^p \cos p(\theta - \mu) \right), 0 \leq \theta < 2\pi, 0 \leq \rho \leq 1 \quad (4.10)$$

Dağılım fonksiyonu, olasılık yoğunluk fonksiyonunun ileriye doğru integrasyonu ile elde edilmektedir.

Momentler;

Ortalama yön : μ

Ortalama bileşke uzunluğu : ρ

$$\begin{aligned} \text{Dairesel saçılım} & : \delta = \frac{1-\rho^4}{2\rho^2} \\ & \alpha_p = \rho p^2 \\ & \beta_p = 0, p \geq 1 \end{aligned}$$

$f(\theta)$ 'nin, σ^2 varyanslı normal dağılımın sarılmasıyla elde edildiği düşünülürse,

$$\rho = e^{-\frac{1}{2}\sigma^2} \quad \text{ya da} \quad \sigma^2 = -2 \log \rho \quad (4.11)$$

olacaktır.

$\rho \rightarrow 0$ iken, sarmal normal dağılım düzgün dağılıma yakınsamaktadır. $\rho \rightarrow 1$ iken, sarmal Cauchy dağılımı μ yönünde yoğunlaşan nokta dağılıma yakınsamaktadır.

Uygun olarak seçilen saçılım parametre değerleri için, sarmal normal dağılım ve von Mises dağılımı, şekil bakımından birbirlerine oldukça benzemektedir. Uygulamalarda daha uygun olan kullanılmalıdır. Bu da, genel olarak istatistiksel çıkarsaması daha kolay olan von Mises dağılımının kullanılması anlamına gelmektedir (Fisher 1993).

4.2. Von Mises (Dairesel Normal) Dağılımı

Von Mises dağılımına sahip bir θ dairesel rassal değişkeni;

$$f(\theta; \mu, \kappa) = \frac{1}{2\pi I_0(\kappa)} e^{\kappa \cos(\theta - \mu)}, \quad 0 \leq \theta < 2\pi, \quad \kappa \geq 0, \quad 0 \leq \mu < 2\pi \quad (4.12)$$

olasılık yoğunluk fonksiyonuna sahiptir ve bu fonksiyon $VM(\mu, \kappa)$ ile gösterilmektedir. Burada, $I_0(\kappa)$; birinci tür ve sıfır sırasında dönüştürülmüş Bessel fonksiyonunu ifade etmektedir ve

$$I_0(\kappa) = \frac{1}{2\pi} \int_0^{2\pi} e^{\kappa \cos(\phi - \mu)} d\phi \quad (4.13)$$

biçiminde tanımlanır. Bessel fonksiyonlarına ilişkin bilgiler EK-1'de verilmektedir. μ parametresi ortalama yönü ifade ederken, κ parametresi yoğunlaşma parametresi olarak tanımlanmaktadır.

Von Mises dağılımı, fizik alanına ilişkin olarak 1905 yılında Langevin tarafından tartışılmış, istatistiksel bir model olarak 1918 yılında von Mises tarafından sunulmuştur. Von Mises, ölçülen atom ağırlıklarının integral değerlerinden sapmaları üzerine yapmış olduğu çalışmalar sonucunda dağılımı elde etmiştir. Dağılımın çeşitli özelliklerine ilişkin yorumlar getiren Gumbel ve ark. (1953), bu dağılımın önemini ve doğru üzerindeki normal dağılımla olan benzerliklerini vurgulamak amacıyla *Dairesel Normal dağılım* olarak adlandırmışlardır. Bu çalışmada, daha yaygın olarak kullanıldığı için von Mises dağılımı ifadesi kullanılacaktır. Von Mises dağılımı üzerine oldukça geniş çaplı çalışmalar yapılmış ve bir çok teknik geliştirilmiştir. Dolayısıyla uygulamalarda dairesel veri setleri için kullanılan dağılım, von Mises dağılımıdır. Von Mises dağılımının birikimli dağılım fonksiyonu;

$$F(\theta) = \frac{1}{2\pi I_0(\kappa)} \left[\theta I_0(\kappa) + 2 \sum_{p=1}^{\infty} \frac{I_p(\kappa) \sin p(\theta - \mu)}{p} \right], 0 \leq \theta < 2\pi \quad (4.14)$$

biçiminde tanımlanır. Burada; $I_p(\kappa)$, birinci tür p -inci sıra dönüştürülmüş Bessel fonksiyonudur. Von Mises yoğunluk fonksiyonunun çeşitli özellikleri aşağıda verilmektedir:

- *Simetri Özelliği*: Kosinüs fonksiyonunun simetrik olmasından dolayı, von Mises yoğunluk fonksiyonu μ ortalama yönüne (ve de $\mu + \pi$ yönüne) göre simetriktir.
- *μ Noktasında Mod Özelliği*: Kosinüs fonksiyonunun sıfır noktasında maksimum değere ulaşmasından dolayı, von Mises yoğunluk fonksiyonu $\theta = \mu$ noktasında maksimumdur. Dolayısıyla, μ ;

$$f(\mu) = \frac{e^{\kappa}}{2\pi I_0(\kappa)} \quad (4.15)$$

maksimum değeri ile *mod yönü*dür.

- $(\mu \pm \pi)$ *Noktasında Antimod Özelliği*: Yine, $\cos \mu = -1$, minimum değer olmasından dolayı, von Mises yoğunluk fonksiyonu $\theta = \mu \pm \pi$ noktasında minimum yoğunluğa sahiptir.

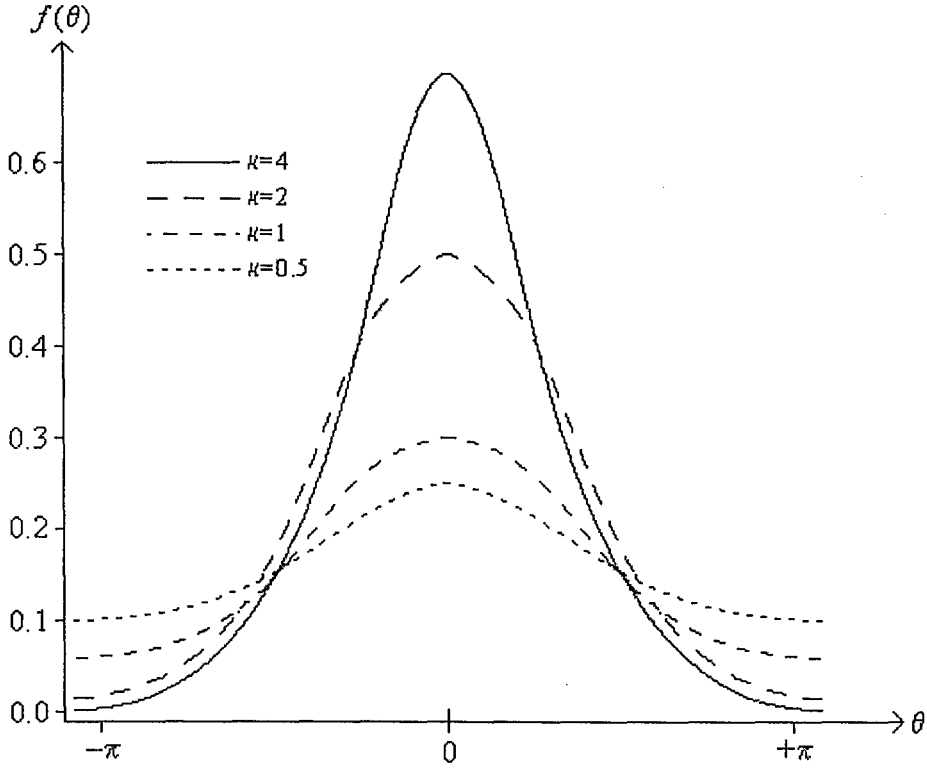
$$f(\mu \pm \pi) = \frac{e^{-\kappa}}{2\pi I_0(\kappa)} \quad (4.16)$$

Buna göre, $\mu \pm \pi$ yönü antimod yönüdür.

- κ 'nın *İşlevi*: Eşitlik (4.15) ve (4.16)'ya göre;

$$\frac{f(\mu)}{f(\mu \pm \pi)} = \frac{\frac{e^{\kappa}}{2\pi I_0(\kappa)}}{\frac{e^{-\kappa}}{2\pi I_0(\kappa)}} = e^{2\kappa}$$

olduğu görülmektedir. Burada, κ değeri ne kadar büyük olursa, $f(\mu)$ fonksiyonunun $f(\mu \pm \pi)$ fonksiyonuna oranı da o kadar büyüyecektir. Bu da, anakütle ortalama yönü ve mod yönü olan μ etrafında daha büyük bir kümelenme olduğunu göstermektedir. Buna göre κ , ortalama yön μ 'ye ilişkin yoğunlaşmayı ölçen bir parametre olmaktadır. Şekil 4.3.'te çeşitli yoğunlaşma parametresi değerleri için elde edilen von Mises yoğunluk fonksiyonları S-Plus yardımıyla çizilmiştir (Jammalamadaka ve Sen Gupta 2001).



Şekil 4.3. $\mu = 0^\circ$ ve $\kappa = 0.5, 1, 2$ ve 4 değerleri için von Mises yoğunluk fonksiyonları

• *Diğer Dağılımlarla Karşılaştırmalar:* $\kappa = 0$ için, Eşitlik (4.12)'ye göre $VM(\mu, \kappa)$, düzgün dağılıma, κ 'nın küçük değerlerinde ise cardioid dağılıma indirgenmektedir. Büyük κ değerlerinde, θ rassal değişkeninin dağılımı bilinen normal dağılıma yakınsamaktadır. Ayrıca, $\kappa \rightarrow \infty$ iken, dağılım $\theta = \mu$ noktası üzerinde kümelenmektedir. Dolayısıyla, κ genellikle yoğunlaşma parametresi olarak tanımlanmaktadır.

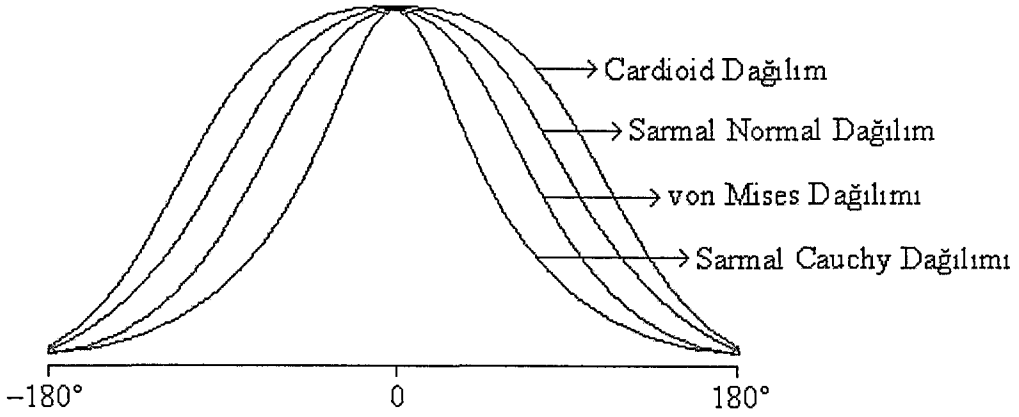
Doğrusal veri setleri için, normal dağılım, sıklıkla kullanılan bir dağılım olarak görülmektedir ve şekilsel istatistiksel analiz, σ^2 değerine bakılmadan yapılabilmektedir. Dairesel ya da küresel veri setleri için ise, şekilsel istatistiksel analiz κ değerine bakılmadan yapılamamaktadır. Yoğunlaşma parametresinden dolayı ortaya çıkan sorunlara rağmen, von Mises dağılımı dairesel veriler için çoğu zaman kullanılan bir dağılımdır. Von Mises dağılımı, özellikle $\kappa \geq 2$ durumu için sarmal normal dağılıma yakınsamaktadır. Uygulamalarda hangisi uygunsa o kullanılmaktadır. Bu dağılımlar ve de dairesel düzgün dağılım halen dairesel veri analizinde kullanılan temel dağılımlardır (Fisher 1993).

Von Mises dağılımının diğer dağılımlarla olan ilişkileri incelendiğinde, μ 'nün ortalama olarak kabul edilebileceği ve $1/\kappa$ 'nın von Mises dağılımındaki etkisinin, σ^2 'nin normal dağılımdaki etkisiyle aynı olduğu görülebilmektedir. Yoğunlaşma parametresi κ 'nın, dairesel varyans V ile olan ilişkisi;

$$V = 1 - \rho = 1 - A(\kappa) \quad (4.17)$$

biçiminde tanımlanır (Mardia 1972).

Şekil 4.4.'te, incelenen dört önemli simetrik, tek-modlu dağılımın aynı κ değerlerine sahipken grafiksel olarak karşılaştırılması verilmektedir.



Şekil 4.4. $\mu = 0^\circ$, $\kappa = 1$ iken, tek-modlu ve simetrik dağılımların karşılaştırılması

Von Mises dağılımının momentleri kısaca özetlenirse;

Ortalama yön	: μ
Ortalama bileşke uzunluğu	: $\rho = A_1(\kappa)$
Dairesel saçılım	: $\delta = \frac{1}{\kappa A_1(\kappa)}$
	$\alpha_p = A_p(\kappa)$
	$\beta_p = 0, p \geq 1$

EK-2 ve EK-3'te ρ 'nun bir fonksiyonu olarak κ değerleri ve κ 'nın bir fonksiyonu olarak ρ değerleri tabloları verilmektedir.

5. DAİRESEL VERİLERDE TEK-ÖRNEKLEM ORTALAMA YÖN TESTLERİ

Bu bölümde, dairesel verilerde ortalama yön için tek örneklem testi incelenecek ve ardından von Mises dağılımına sahip bir anakütlede ortalama yön testi verilecektir.

5.1. Tek Örneklem Ortalama Yön Testi

Belirlenen bir α anlamlılık seviyesinde $\mu = \mu_0$ hipotezinin, $\mu \neq \mu_0$ alternatif hipotezine karşı testi için iki durum söz konusudur. Birinci yöntem, ortalama yön için güven aralığının belirlenmesi yardımıyla test uygulamak, ikincisi ise, n örneklem hacmine dayalı olarak doğrudan test uygulamaktır.

5.1.1. Güven aralığının belirlenmesi yoluyla test

Bu yöntemde, ortalama yönün dairesel standart hatasından yararlanılmaktadır. Buna göre sırasıyla;

$$\rho_2 = \frac{1}{n} \sum_{i=1}^n \cos 2(\theta_i - \bar{\theta}), \quad \delta = \frac{1 - \rho_2}{2\bar{R}^2} \quad \text{ve} \quad \sigma = \sqrt{\frac{\delta}{n}}$$

değerleri hesaplanır. Bu işlemlerden sonra, ortalama yön için $(1-\alpha)$ 'lık güven aralığı;

$$(\mu_0 - \sin^{-1}(z_{\alpha/2} \sigma), \mu_0 + \sin^{-1}(z_{\alpha/2} \sigma)) \quad (5.1)$$

ile verilmektedir. Burada, $z_{\alpha/2}$; standart normal dağılım tablosundan elde edilmektedir. Kurulacak hipotezler;

$$H_0: \mu = \mu_0$$

$$H_1: \mu \neq \mu_0 \quad (5.2)$$

olduğunda, ortalama yön değeri (5.1)'de verilen aralığın içinde ise, α anlamlılık seviyesinde H_0 hipotezi kabul edilir, aralığın dışında ise H_0 hipotezi reddedilir (Fisher 1993 ve İnal ve Günay 1999).

5.1.2. Örneklem hacmine dayalı test

Uygun bir güven aralığının belirlenemediği durumda doğrudan uygulanabilen bu test örneklem hacmi 25 veya daha fazla olduğunda kullanılabilir. Bu yöntemle göre, öncelikle dairesel standart hata σ yukarıdaki testte verildiği gibi hesaplanmaktadır. Test istatistiği;

$$S = \frac{\sin(\bar{\theta} - \mu_0)}{\sigma} \quad (5.3)$$

olarak tanımlanır. Bu değer α güvenilirlik seviyesinde, standart normal dağılım tablosundan elde edilen değerle karşılaştırılır. $|S|$ değeri tablo değerinden büyük olduğunda $\mu = \mu_0$ hipotezi reddedilir, aksi halde kabul edilir (Fisher 1993).

5.2. Von Mises Dağılımına Sahip Anakütlerde Ortalama Yön Testi

İncelenen anakütlenin μ ortalama yönü ve κ yoğunluk parametresi ile von Mises dağılımına $VM(\mu, \kappa)$ sahip olduğu varsayalım. Buna göre olasılık yoğunluk fonksiyonu:

$$f(\theta) = \{2\pi I_0(\kappa)\}^{-1} \exp\{\kappa \cos(\theta - \mu)\} , 0 \leq \theta < 2\pi , \kappa \geq 0 \quad (5.4)$$

ile verilir.

Bu kesimde, von Mises dağılımına sahip bir anakütlerde ortalama yönün μ_0 gibi bir değere eşit olup olmadığı testi incelenecektir. İlk olarak, ortalama yön için güven aralığının belirlenmesi yöntemiyle uygulanan test verilecektir. Doğrudan test uygulanabilmesi için ise, yoğunlaşma parametresi κ 'nın bilindiği ve bilinmediği durum olmak üzere iki durum söz konusudur. Bu durumlar da ayrı ayrı ele alınacaktır.

5.2.1. Güven aralığının belirlenmesi yoluyla test

Güven aralığı kullanılarak yapılan testte, ortalama yön için dairesel standart hata değeri hesaplanmaktadır. Fakat, burada von Mises dağılımı için özel bir biçim alan standart hata formülü kullanılmaktadır.

$$\sigma_{VM} = \frac{1}{\sqrt{nR\hat{\kappa}}} \quad (5.5)$$

Buna göre; ortalama yön için $(1-\alpha)$ 'lık güven aralığı;

$$(\mu_0 - \sin^{-1}(z_{\alpha/2} \sigma_{VM}), \mu_0 + \sin^{-1}(z_{\alpha/2} \sigma_{VM})) \quad (5.6)$$

olacaktır. Ortalama yön değeri (5.6)'da verilen aralığın içinde yer alıyorsa, α anlamlılık seviyesinde H_0 hipotezi kabul edilir, aralığın dışında yer alıyorsa H_0 hipotezi reddedilir (Fisher 1993).

5.2.2. Örneklem hacmine ve yoğunlaşma parametresine dayalı test

Bu test, örneklem hacmi n ve yoğunlaşma parametresi $\hat{\kappa}$ tahmin değerinin büyüklüğüne bağlı olarak uygulanmaktadır.

5.2.2.1. Yoğunlaşma parametresinin bilinmediği durum

Von Mises dağılımı için ortalama yönün dairesel standart hatası σ_{VM} Eşitlik (5.5)'te tanımlandığı şekilde hesaplanır. Test istatistiği ise;

$$E_n = \frac{\sin(\bar{\theta} - \mu_0)}{\sigma_{VM}} \quad (5.7)$$

olarak tanımlanır. Bu değer α güvenilirlik seviyesinde, standart normal dağılım tablosundan elde edilen değerle karşılaştırılır. Karşılaştırmalar alternatif hipotezin değişik durumları için aşağıda verildiği gibi yapılmaktadır:

- $H_0: \mu = \mu_0, H_1: \mu \neq \mu_0$ testi için $|E_n| > z_{\alpha/2}$ ise H_0 reddedilir.
- $H_0: \mu = \mu_0, H_1: \mu < \mu_0$ testi için $\mu_0 - \pi < \bar{\theta} < \mu_0$ ve $E_n < -z_\alpha$ ise H_0 reddedilir.
- $H_0: \mu = \mu_0, H_1: \mu > \mu_0$ testi için $\bar{\theta} < \mu_0 + \pi$ ve $E_n > -z_\alpha$ ise H_0 reddedilir.

Test, n ve $\hat{\kappa}$ 'nin aldığı değerlere bağlı olarak aşağıdaki durumlarda kullanılmaktadır:

Çizelge 5.1. n ve κ değerlerine bağlı olarak testin uygulanabildiği durumlar

$\hat{\kappa}$	n
$0.4 \leq \hat{\kappa} < 1.0$	$n \geq 25$
$1.0 \leq \hat{\kappa} < 1.5$	$n \geq 15$
$1.5 \leq \hat{\kappa} < 2.0$	$n \geq 10$
$\hat{\kappa} \geq 2.0$	Bütün n 'ler

(Fisher 1993).

5.2.2.2. Yoğunlaşma parametresinin bilindiği durum

Yoğunlaşma parametresi κ 'nın κ_0 gibi bilinen bir değere eşit olduğu varsayalım.

$\kappa_0 \geq 2$ ise, Von Mises dağılımı için ortalama yönün dairesel standart hatası σ_{VM} ;

$$\sigma_{VM} = \frac{1}{\sqrt{n\rho\kappa_0}} \quad (5.8)$$

biçimindedir. Buradaki ρ değeri, EK-3'te verilen tabloda κ_0 kullanılarak hesaplanabilmektedir. Test istatistiği;

$$E_n = \frac{\sin(\bar{\theta} - \mu_0)}{\sigma_{VM}} \quad (5.9)$$

olarak tanımlanır. Bu değer α güvenilirlik seviyesinde, standart normal dağılım tablosundan elde edilen değerle karşılaştırılır.

- $H_0: \mu = \mu_0$, $H_1: \mu \neq \mu_0$ testi için $|E_n| > z_{\alpha/2}$ ise H_0 reddedilir.
- $H_0: \mu = \mu_0$, $H_1: \mu < \mu_0$ testi için $\mu_0 - \pi < \bar{\theta} < \mu_0$ ve $E_n < -z_\alpha$ ise H_0 reddedilir.
- $H_0: \mu = \mu_0$, $H_1: \mu > \mu_0$ testi için $\bar{\theta} < \mu_0 + \pi$ ve $E_n > -z_\alpha$ ise H_0 reddedilir.

(Fisher 1993).

6. OLASILIK ORANLARININ ARDIŞIK TESTİ

Ardışık testi diğer testlerden ayıran temel özellik, test için gereken örneklem hacminin daha önceden belirlenmemiş olup, bir rassal değişken olmasıdır. Bu nedenle örneklem hacminin beklenen değeri söz konusudur.

Ardışık testin her aşamasında aşağıda verilen üç karardan birine ulaşılmaktadır.

- 1- H_0 hipotezinin kabul edilmesi,
- 2- H_0 hipotezinin reddedilmesi,
- 3- Gözlemlerin yetersizliği kararı.

Eğer, ilk iki karardan birine ulaşılmış ise, test sona ermektedir. Ancak, üçüncü karara ulaşılmış ise, bir gözlem daha eklenerek teste devam edilmektedir. Ardışık testte H_0 hipotezinin kabul ya da reddedilmesi kararlarından birine ulaşılması kesindir ve olasılığı 1'dir.

Ardışık test, zaman ve para bakımından maliyetin yüksek olduğu alanlarda büyük tasarruf sağlamaktadır. Ayrıca bazı durumlarda ardışık olmayan testler için gerekli örneklem hacminden %50 daha az gözlem değeri, ardışık test için yeterli olmaktadır.

Ardışık test yardımıyla dağılımı bilinen bir kitlenin parametrelerine ilişkin kararlar verilebilmektedir. Test için gereken H_0 ve H_1 basit hipotezleri test edilmek istenen anakütle parametrelerine bağlı olarak kurulmakta ve önceden belirlenen birinci ve ikinci tip hata olasılıklarına göre test edilmektedir. Örneğin, X rassal değişkeni normal dağılım gösteriyor ise, ardışık test ile anakütle ortalaması ya da varyansı test edilebilmektedir. Ancak anakütle ortalaması test edilmek istendiğinde anakütle varyansının, anakütle varyansı test edilmek istendiğinde ise, anakütle ortalamasının bilinmesi gerekmektedir. Dolayısıyla ardışık testte test edilecek anakütle parametresi dışında tüm anakütle parametre değerlerinin bilinmesi gerekmektedir (Bacanlı 1988).

6.1. Kuramsal Tanım

Herhangi bir X rassal değişkeninin olasılık yoğunluk fonksiyonu (o.y.f.) $f(x; \Theta)$ ile gösterilmektedir.

$f(x; \Theta)$ o.y.f.'daki Θ parametresi test edilmek istendiğinde basit hipotez;

$$H_0: \Theta = \Theta_0 ,$$

$$H_1: \Theta = \Theta_1 \quad (\Theta_0 < \Theta_1) \quad (6.1)$$

biçiminde kurulmaktadır.

X rassal değişkeninin o.y.f., H_0 hipotezi doğru iken $f(x; \Theta_0)$, H_1 hipotezi doğru iken $f(x; \Theta_1)$ 'dir. X rassal değişkeninin aldığı değerler x_1, x_2, \dots, x_n olarak belirtildiğinde,

$$f(x; \Theta_0) = f(x_1, x_2, \dots, x_n; \Theta_0)$$

$$f(x; \Theta_1) = f(x_1, x_2, \dots, x_n; \Theta_1) \quad (6.2)$$

x_1, x_2, \dots, x_n değerleri birbirinden bağımsız iseler,

$$f(x_1, x_2, \dots, x_n; \Theta_0) = f(x_1; \Theta_0), \dots, f(x_n; \Theta_0)$$

$$f(x_1, x_2, \dots, x_n; \Theta_1) = f(x_1; \Theta_1), \dots, f(x_n; \Theta_1) \quad (6.3)$$

olmaktadır. Tanımlanan bileşik o.y.f.'larının birbirine oranı *genel olabilirlik oranı* olarak bilinir.

$$L_n = \prod_{i=1}^n \frac{f(x_i; \Theta_1)}{f(x_i; \Theta_0)} \quad (6.4)$$

Ardışık testte H_0 hipotezini H_1 hipotezine karşı test etmek için, önce birinci tip hata, α , (H_0 hipotezi doğru iken reddedilmesi olasılığı) ve ikinci tip hata, β , (H_1 hipotezi doğru iken reddedilmesi olasılığı) olasılıklarına bağlı olan A ve B ($B < A$) pozitif değişmezleri saptanmaktadır ve n -inci aşamada genel olabilirlik oranı (L_n) ile aşağıda belirtildiği gibi karşılaştırılmaktadır.

- 1- $L_n \leq B$ ise, H_0 hipotezi kabul edilerek sürece son verilir,
- 2- $L_n \leq A$ ise, H_0 hipotezi reddedilerek sürece son verilir,
- 3- $B < L_n < A$ ise, gözlemlerin yetersizliğine karar verilir ve bir gözlem daha eklenerek sürece devam edilir.

Burada,

$$A = \frac{1-\beta}{\alpha}, \quad B = \frac{\beta}{1-\alpha} \quad (6.5)$$

dır.

Genel olabilirlik oranı, A ve B değişmezlerine dayanarak yapılan bu teste *Olasılık Oranlarının Ardışık Testi* veya *Wald tipi ardışık test* denilmektedir.

L_n değeri belirsizlik durumunda (yani; $f(x_i; \Theta_1) = f(x_i; \Theta_0) = 0$ ise) 1 değerine eşit olarak alınmaktadır. $L_n = \infty$ olduğunda ise (yani; $f(x_i; \Theta_1) > 0$, $f(x_i; \Theta_0) = 0$ ise), H_0 hipotezi reddedilerek sürece son verilmektedir.

L_n değeri logaritma yardımıyla daha kolay bulunabileceğinden Eşitlik (6.4)'ün logaritması alınarak,

$$\ln L_n = \sum_{i=1}^n \ln \left[\frac{f(x_i; \Theta_1)}{f(x_i; \Theta_0)} \right] \quad (6.6)$$

elde edilir. Burada,

$$Z_i = \ln \left[\frac{f(x_i; \Theta_1)}{f(x_i; \Theta_0)} \right] \quad (6.7)$$

ile gösterildiğinde (6.6) eşitliği,

$$\ln L_n = \sum_{i=1}^n Z_i \quad (6.8)$$

olarak yazılabilmektedir.

Bu şekilde n -inci aşamada $\ln L_n$ bulunarak $\ln A$ ve $\ln B$ ile aşağıda belirtildiği gibi karşılaştırılmaktadır:

- 1- $\sum_{i=1}^n Z_i \leq \ln B$ ise, H_0 hipotezi kabul edilerek sürece son verilmektedir.
- 2- $\sum_{i=1}^n Z_i \geq \ln A$ ise, H_0 hipotezi reddedilerek sürece son verilmektedir.
- 3- $\ln B < \sum_{i=1}^n Z_i < \ln A$ ise, gözlemlerin yetersizliğine karar verilir ve bir gözlem daha eklenerek sürece devam edilir.

6.1.1. A ve B değişmezlerinin saptanması

Ardışık testte amaç α ve β hata olasılıklarını aşmadan en küçük örneklem hacminde H_0 hipotezi hakkında güvenilir bir karar vermektir. Bu amaçla, önceden belirtilen hata olasılıklarına bağlı olarak A ve B değişmezlerini saptamak gerekmektedir.

Bağımsız n tane gözlemden oluşan (x_1, x_2, \dots, x_n) örnekleminin n -inci aşamasında H_0 hipotezi kabul edilmiş ise, $L_n \leq B$ koşulu sağlanmıştır.

$$\prod_{i=1}^n \frac{f(x_i; \Theta_1)}{f(x_i; \Theta_0)} \leq B$$

$$f(x_1, x_2, \dots, x_n; \Theta_1) \leq B f(x_1, x_2, \dots, x_n; \Theta_0) \quad (6.9)$$

Burada, $f(x_1, x_2, \dots, x_n; \Theta_1)$ H_1 hipotezi doğru iken H_0 hipotezinin kabul edilmesi olasılığı olan β değerine, $f(x_1, x_2, \dots, x_n; \Theta_0)$ ise H_0 hipotezi doğru iken H_0 hipotezinin kabul edilmesi olasılığı olan $(1-\alpha)$ değerine eşittir. Bu değerler (6.9)'da yerine yazıldığında;

$$\beta \leq B(1-\alpha)$$

$$\frac{\beta}{1-\alpha} \leq B \quad (6.10)$$

elde edilmektedir.

(x_1, x_2, \dots, x_n) örnekleminde H_0 hipotezi reddedilmiş ise, $L_n \geq A$ koşulu sağlanmıştır.

$$\prod_{i=1}^n \frac{f(x_i; \Theta_1)}{f(x_i; \Theta_0)} \geq A$$

$$f(x_1, x_2, \dots, x_n; \Theta_1) \geq A f(x_1, x_2, \dots, x_n; \Theta_0) \quad (6.11)$$

Burada, $f(x_1, x_2, \dots, x_n; \Theta_1)$; H_1 hipotezi doğru iken H_0 hipotezinin reddedilmesi olasılığı olan $(1-\beta)$ 'ya, $f(x_1, x_2, \dots, x_n; \Theta_0)$ ise; H_0 hipotezi doğru iken H_0 hipotezinin reddedilmesi olasılığı olan α 'ya eşittir. Bu değerler (6.11)'de yerine yazıldığında;

$$(1-\beta) \geq A\alpha$$

$$\frac{1-\beta}{\alpha} \geq A \quad (6.12)$$

elde edilir.

Uygulamalarda A ve B değişmezlerini saptamak için (6.10) ve (6.12) eşitsizliklerinde yalnızca eşitlik işareti kullanılmaktadır. Dolayısıyla, $A = \frac{1-\beta}{\alpha}$

ve $B = \frac{\beta}{1-\alpha}$ olur (Çingir 1990).

6.1.2. Karakteristik işlem fonksiyonu

$H_0: \Theta = \Theta_0$ hipotezini kabul etme olasılığı Θ 'nın bir fonksiyonudur. Bu fonksiyona *karakteristik işlem fonksiyonu* denilmektedir ve $P(\Theta)$ ile gösterilmektedir.

$$P(\Theta) = P(H_0 \text{ kabul} \mid H_1 \text{ yanlış ya da } H_0 \text{ yanlış}) \quad (6.13)$$

yazılabilmektedir.

H_0 'ın kabul edilme olasılığını bulmak için Wald'ın önerdiği yöntemle göre;

$$E\{[f(x; \Theta_1) / f(x; \Theta_0)]^h\} = 1 \quad (6.14)$$

olacak şekilde Θ parametresi bulunmaktadır. Eğer X rassal değişkeni sürekli ise,

$$\int_{-\infty}^{+\infty} [f(x; \Theta_1) / f(x; \Theta_0)]^h f(x; \Theta) dx = 1 \quad (6.15)$$

eşitliğinden yararlanılır. $h = 0$ olduğunda Eşitlik (6.14)'deki beklenen değer ifadesi 1'e eşittir. Ancak h 'nin diğer değerleri için de 1 değerini almaktadır. $h \neq 0$ için $f(x; \Theta)$ fonksiyonu tanımlıdır. $h > 0$ ya da $h < 0$ olabilir. Burada $h > 0$ olduğu durum ele alınacaktır.

Bir Θ parametresine karşılık gelen X rassal değişkeninin olasılık yoğunluk fonksiyonu;

$$g(x; \Theta) = [f(x; \Theta_1) / f(x; \Theta_0)]^h f(x; \Theta) \quad (6.16)$$

olsun. Bu fonksiyona göre H_0 ve H_1 hipotezleri,

$$\begin{aligned} H_0: f &= f(x; \Theta) \\ H_1: f &= g(x; \Theta) \end{aligned} \quad (6.17)$$

Eşitlik (6.17) biçiminde kurularak ardışık test uygulanabilmektedir. Bu hipotezlere ilişkin genel olabilirlik oranı,

$$L'_n = \prod_{i=1}^n [g(x_i; \Theta) / f(x_i; \Theta)] \quad (6.18)$$

olmaktadır. $g(x_i; \Theta)$ değeri Eşitlik (6.18)'de yerine yazıldığında,

$$L'_n = \prod_{i=1}^n [f(x_i; \Theta_1) / f(x_i; \Theta_0)]^h \quad (6.19)$$

şeklini almaktadır. L'_n oranı A^h ve B^h pozitif değişmezleriyle karşılaştırılarak Kesim 6.1.'de tanımlanan üç karardan birine ulaşılmaktadır.

A^h ve B^h değerlerine karşılık gelen birinci tip ve ikinci tip hata olasılıkları α' ve β' ile gösterildiğinde,

$$A^h = (1-\beta') / \alpha' \quad (6.20)$$

$$B^h = \beta' / (1-\alpha') \quad (6.21)$$

bulunabilmektedir. Bu ifadelerden yararlanarak,

$$\alpha' = (1-B^h) / (A^h - B^h)$$

olarak bulunur. H_0 hipotezini kabul etme olasılığı ise,

$$P(\Theta) = 1-\alpha' \quad (6.22)$$

$$P(\Theta) = \frac{A^h - 1}{A^h - B^h} \quad (6.23)$$

olur.

Eşitlik (6.23) ile verilen eşitlikte Θ parametresi h ($-\infty \leq h \leq +\infty$) değerine bağlı olduğundan her bir h değeri için Θ parametresinin ve karakteristik işlem fonksiyonunun ($P(\Theta)$) alacağı değerler bulunabilmektedir (Çizelge 6.1.). h 'ın $-\infty$, 0 ve $+\infty$ noktalarında $P(\Theta)$ değerleri,

$$\lim_{h \rightarrow \infty} P(\Theta) = 1$$

$$\lim_{h \rightarrow -\infty} P(\Theta) = 0$$

$$\lim_{h \rightarrow 0} P(\Theta) = \frac{\ln A}{\ln A - \ln B}$$

olur (Wald 1947).

Çizelge 6.1. Olasılık oranlarının ardışık testi için çeşitli h değerlerinde Θ ve $P(\Theta)$ değerleri

h	$-\infty$	-1	0	1	∞
Θ		Θ_1		Θ_0	
$P(\Theta)$	0	β	$\ln A / (\ln A - \ln B)$	$1 - \alpha$	1

6.1.3. Ortalama örneklem sayısı fonksiyonu

Ardışık testte H_0 ve H_1 hipotezlerinden biri hakkında karar verebilmek için gerekli olan örneklem hacmi (n), test süresince değişen bir rassal değişkendir. Bu nedenle beklenen değeri söz konusudur.

Test edilecek parametre Θ olduğundan örneklem hacminin beklenen değeri ($E(n; \Theta)$), Θ 'nın bir fonksiyonu olduğundan *ortalama örneklem sayısı (o.ö.s.) fonksiyonu* olarak adlandırılmaktadır.

Daha önce Eşitlik (6.7) ile tanımlanan,

$$Z_i = \ln [f(x_i; \Theta_1) / f(x_i; \Theta_0)]$$

değişkeni $n < N$ varsayımı altında,

$$\sum_{i=1}^N Z_i = \sum_{i=1}^n Z_i + \sum_{i=n+1}^N Z_i \quad (6.24)$$

olarak yazılabilmektedir. Eşitliğin beklenen değeri ise,

$$N E(Z) = E\left(\sum_{i=1}^n Z_i\right) + E\left(\sum_{i=n+1}^N Z_i\right) \quad (6.25)$$

olmaktadır.

$$\begin{aligned} E\left(\sum_{i=n+1}^N Z_i\right) &= E(N - n)E(Z) = (N - E(n))E(Z) \\ &= N E(Z) - E(n)E(Z) \end{aligned} \quad (6.26)$$

olur. Bu değer Eşitlik (6.25)'de yerine yazılırsa,

$$E(n) = \frac{E\left(\sum_{i=1}^n Z_i\right)}{E(Z)} \quad E(Z) \neq 0 \quad (6.27)$$

olarak elde edilmektedir. Test edilecek parametre Θ olduğunda Eşitlik (6.27),

$$E(n; \Theta) = \frac{E\left(\sum_{i=1}^n Z_i; \Theta\right)}{E(Z; \Theta)} \quad (6.28)$$

olur.

H_0 hipotezini kabul etme olasılığı $P(\Theta)$ ve reddetme olasılığı $(1 - P(\Theta))$ olduğunda Eşitlik (6.28),

$$E(n; \Theta) = \frac{P(\Theta)E_1\left(\sum_{i=1}^n Z_i; \Theta\right) + (1 - P(\Theta))E_2\left(\sum_{i=1}^n Z_i; \Theta\right)}{E(Z; \Theta)} \quad (6.29)$$

biçiminde yazılmaktadır.

Z 'nin standart sapması ve $|E(Z)|$ küçük ise, $E_1\left(\sum_{i=1}^n Z_i; \Theta\right)$, $\ln B$ 'ye,

$E_2\left(\sum_{i=1}^n Z_i; \Theta\right)$, $\ln A$ 'ya çok yakın bir değer almaktadır. Bu değerler

Eşitlik (6.29)'da yerine yazıldığında o.ö.s. fonksiyonunun yaklaşık formülü;

$$E(n; \Theta) = \frac{P(\Theta) \ln B + (1 - P(\Theta)) \ln A}{E(Z; \Theta)} \quad , \quad E(Z; \Theta) \neq 0 \quad (6.30)$$

olarak elde edilmektedir.

Böylece örneklem hacmi bilinmediği halde, Eşitlik (6.30)'dan yararlanılarak anakütle parametre değerine göre, en az ve en çok ne kadarlık bir örneklem hacmine gerek duyulacağı hakkında ön bilgi elde edilebilmektedir.

(6.30) fonksiyonu H_0 hipotezi doğru iken,

$$E(n; \Theta_0) = \frac{(1 - \alpha) \ln B + \alpha \ln A}{E(Z; \Theta_0)} \quad (6.31)$$

H_1 hipotezi doğru iken,

$$E(n; \Theta_1) = \frac{\beta \ln B + (1 - \beta) \ln A}{E(Z; \Theta_1)} \quad (6.32)$$

olarak yazılabilmektedir.

Eşitlik (6.30)'da $E(Z; \Theta) = 0$ ise o.ö.s. fonksiyonu,

$$E(n; \Theta) = \frac{-\ln A \log B}{E(Z^2; \Theta)} \quad (6.33)$$

ile verilmektedir (Çingı 1990 ve Wald 1947).

6.2. Varyansı Bilinen Normal Dağılımın Ortalamasının Ardışık Testi

6.2.1. Testin çıkarsaması

X , varyansı bilinen (σ^2), ortalaması (μ) bilinmeyen normal dağılıma sahip bir rassal değişken olsun. Test edilecek parametre anakütle ortalaması olduğundan hipotezler;

$$\begin{aligned} H_0: \mu &= \mu_0 \\ H_1: \mu &= \mu_1, (\mu_0 < \mu_1) \end{aligned} \quad (6.34)$$

biçiminde kurulmaktadır.

X rassal değişkeni normal dağılıma sahip olduğundan o.y.f.;

$$f(x; \mu; \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-(x - \mu)^2 / 2\sigma^2\right], \quad -\infty < x < +\infty \quad (6.35)$$

Eşitlik (6.35) biçimindedir. H_0 hipotezi doğru iken, X rassal değişkeninin o.y.f.;

$$f(x; \mu_0; \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-(x - \mu_0)^2 / 2\sigma^2\right] \quad (6.36)$$

H_1 hipotezi doğru iken;

$$f(x; \mu_1; \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-(x - \mu_1)^2 / 2\sigma^2\right] \quad (6.37)$$

ile verilmektedir.

X rassal değişkeninin aldığı değerler x_1, x_2, \dots, x_n olarak belirtildiğinde en çok olabilirlik oranı;

$$L_n = \prod_{i=1}^n \frac{f(x_i; \mu_1; \sigma^2)}{f(x_i; \mu_0; \sigma^2)}$$

$$L_n = \frac{\exp\left[-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu_1)^2\right]}{\exp\left[-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu_0)^2\right]} \quad (6.38)$$

Eşitlik (6.38) biçiminde elde edilmektedir.

$$\ln L_n = \sum_{i=1}^n Z_i$$

olduğundan,

$$\sum_{i=1}^n Z_i = \frac{\sum_{i=1}^n x_i (\mu_1 - \mu_0)}{\sigma^2} + \frac{n}{2\sigma^2} (\mu_0^2 - \mu_1^2) \quad (6.39)$$

olarak elde edilir (Wald 1947). Bu değer $\ln A$ ve $\ln B$ ile karşılaştırılarak aşağıdaki üç karardan birine ulaşılmaktadır;

- 1- $\sum_{i=1}^n Z_i \leq \ln B$ ise, H_0 hipotezi kabul edilerek sürece son verilmektedir.
- 2- $\sum_{i=1}^n Z_i \geq \ln A$ ise, H_0 hipotezi reddedilerek sürece son verilmektedir.
- 3- $\ln B < \sum_{i=1}^n Z_i < \ln A$ ise, gözlemlerin yetersizliğine karar verilir ve bir gözlem daha eklenerek sürece devam edilir.

6.2.2. Karakteristik işlem fonksiyonu

Varyansı bilinen normal dağılımda, ortalamaya bağlı olarak kurulan $H_0: \mu = \mu_0$ hipotezinin kabul edilme olasılığı $P(\mu)$;

$$P(\mu) = \frac{A^h - 1}{A^h - B^h}$$

ile bulunabilmektedir. Bunun için önce Kesim 6.1.2.'de verilen yöntemle göre;

$$E \left\{ \left[\frac{f(x; \mu_1; \sigma^2)}{f(x; \mu_0; \sigma^2)} \right]^h \right\} = 1$$

olacak şekilde μ parametresinin elde edilmesi gerekmektedir. X rassal değişkeni sürekli olduğundan bu ifade;

$$\int_{-\infty}^{+\infty} \left[\frac{f(x; \mu_1; \sigma^2)}{f(x; \mu_0; \sigma^2)} \right]^h f(x; \mu; \sigma^2) dx = 1$$

biçiminde yazılabilmektedir. İlgili fonksiyonlar eşitlikte yerine yazıldığında;

$$\int_{-\infty}^{+\infty} \left[\frac{\exp\left(-\frac{1}{2\sigma^2}(x-\mu_1)^2\right)}{\exp\left(-\frac{1}{2\sigma^2}(x-\mu_0)^2\right)} \right]^h \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right) dx = 1 \quad (6.40)$$

olmaktadır. Bu integral h 'a göre çözüldüğünde;

$$h = \frac{\mu_1 + \mu_0 - 2\mu}{\mu_1 - \mu_0} \quad (6.41)$$

elde edilmektedir (Wald 1947).

Bu eşitlik yardımıyla çeşitli h değerleri için μ parametresinin alacağı değerler bulunarak $P(\mu)$ değerleri hesaplanabilmektedir.

Çizelge 6.2. Çeşitli h değerlerine karşılık gelen μ ve $P(\mu)$ değerleri

h	$-\infty$	-1	0	1	∞
μ	∞	μ_1	$(\mu_0 + \mu_1) / 2$	μ_0	$-\infty$
$P(\mu)$	0	β	$\ln A / (\ln A - \ln B)$	$1 - \alpha$	1

6.2.3. Ortalama örneklem sayısı fonksiyonu

Varyansı bilinen normal dağılıma sahip bir anakütlede, anakütle ortalaması test edilmek istendiğinde test için gerekli olan örneklem sayısı fonksiyonu Eşitlik (6.30)'de verildiği gibi;

$$E(n; \mu) = \frac{P(\mu) \ln B + (1 - P(\mu)) \ln A}{E(Z; \mu)} \quad (6.42)$$

yaklaşık eşitliği ile bulunmaktadır. Burada,

$$Z = \log \left[\frac{f(x; \mu_1; \sigma^2)}{f(x; \mu_0; \sigma^2)} \right]$$

$$Z = \log \left[\frac{\exp(-(x - \mu_1)^2 / 2\sigma^2)}{\exp(-(x - \mu_0)^2 / 2\sigma^2)} \right] \quad (6.43)$$

$$Z = \frac{2(\mu_1 - \mu_0)x + \mu_0^2 - \mu_1^2}{2\sigma^2} \quad (6.44)$$

$$E(Z; \mu) = \frac{2(\mu_1 - \mu_0)\mu + \mu_0^2 - \mu_1^2}{2\sigma^2} \quad (6.45)$$

olarak bulunmaktadır. Bu değer, Eşitlik (6.42)'de yerine yazıldığında;

$$E(n; \mu) = \frac{P(\mu) \ln B + (1 - P(\mu)) \ln A}{\frac{2(\mu_1 - \mu_0)\mu + \mu_0^2 - \mu_1^2}{2\sigma^2}} \quad (6.46)$$

bulunmaktadır (Wald 1947 ve Bacanlı 1988).

7. DAİRESEL VERİLERDE ORTALAMA YÖN İÇİN ARDIŞIK TESTLER

Bu bölümde, von Mises dağılımına sahip anakütller için ardışık test süreçleri incelenecek ve farklı bilim dallarında dairesel veriler için ardışık süreçlere ilişkin bazı örnekler verilecektir. Ayrıca, dairesel veri analizine ilişkin problemlerde kullanılan yöntemler kısaca tartışılacaktır. Von Mises anakütlesi için ortalama yönlerin karşılaştırılması testinde uygulanan iki ardışık test süreci incelenecektir. Bu süreçlerden birincisinde, yoğunlaşma parametresi κ 'nın bilindiği varsayılmakta ve testteki karakteristik işlem fonksiyonu ile ortalama örneklem sayıları tanımlanmaktadır. Diğer testte ise κ parametresinin bilinmediği durum söz konusudur.

Dairesel verilere ardışık test uygulanması üzerine ilk çalışmalar 1966 yılında Vistelius ve ardından 1975 yılında Mardia tarafından yapılmıştır. Gadsden ve Kanji (1981), von Mises anakütlesi için bazı ardışık süreçler üzerinde tartışmışlardır (Gadsden ve Kanji 1982-83).

Jeoloji alanında, fay (ya da kayalardaki çatlak) yönlerine ilişkin yapılan çalışmalarda veri toplama sorunları ortaya çıkabilmektedir. Özel olarak, eğer üzerinde inceleme yapılan alan sınırlı bir kaya katmanı ise, fayların birçoğu kısmen ya da tamamen gizlidir. Dolayısıyla, gözlem değerlerinin elde edilebilmesi için kazı çalışması yapılması gerekebilmektedir. Jeoloji alanındaki ikinci bir uygulama, okyanus yatağında bulunan kayalardaki artık manyetik vektörlerin analizidir. Burada örneklemelerin ve bunlara ilişkin verinin elde edilmesi oldukça pahalı ve zordur. Bundan dolayı, örneklemeler teker teker ele alınmalı, yani ardışık süreç uygulanmalıdır. Üçüncü uygulama ise, biyoloji alanında yapılmıştır. Bu uygulamada kuş göçü incelenmiştir. Biyologlar kuşların uçuş açılarını, ufukta gözden kayboldukları noktayı dikkate alarak ölçmüşlerdir. Burada amaç, kuşların rassal bir yöne doğru mu yoksa, göç edecekleri yerin bulunduğu yöne doğru mu uçuşa başladıklarını araştırabilmektir. Bazı kuşlar sürü ile göç etmek yerine teker teker veya ikişer ikişer hareket etmekte ve birlikte uçmak yerine, uzayan bir zaman periyodunda uçuş gerçekleştirmektedirler. Böylece, veri seti ardışık hale dönüşmüş olmaktadır. Biyoloji alanında yapılan bir diğer uygulama çalışması

örümcek anakütlesinin saptanabilmesi amacıyla örümcek ağlarına ilişkin yapılmıştır. Veri yeni bir ağ örüldüğünde ardışık olarak elde edilmektedir (Gadsden ve Kanji 1982-83).

θ değişkenine ait gözlem değerlerinin, ortalama yön μ_0 ve yoğunlaşma parametresi κ olmak üzere von Mises dağılımına sahip bir anakütleden çekildiği varsayalım. Buna göre θ 'nin olasılık yoğunluk fonksiyonu;

$$f(\theta) = \{2\pi I_0(\kappa)\}^{-1} \exp\{\kappa \cos(\theta - \mu)\} , 0 < \theta, \mu \leq 2\pi, \kappa > 0 \quad (7.1)$$

ile verilmektedir. Burada, $I_0(\kappa)$; birinci tür ve sıfır sırasında dönüştürülmüş Bessel fonksiyonudur;

$$I_0(\kappa) = \sum_{r=0}^{\infty} \left(\frac{1}{2}\kappa\right)^{2r} (r!)^{-2}$$

κ değeri sıfıra eşitken, olasılık yoğunluk fonksiyonu düzgün dağılım olacaktır.

$$f(\theta) = (2\pi)^{-1} , 0 < \theta < 2\pi \quad (7.2)$$

Von Mises dağılımı, çember üzerindeki normal dağılıma eşdeğer olarak kullanılmasına karşın, normal dağılımın tüm özelliklerine sahip değildir. Özellikle, $\theta_1 \sim VM(\mu_0, \kappa_1)$ ve $\theta_2 \sim VM(\mu_1, \kappa_2)$ iken, $\theta_1 + \theta_2$ 'nin dağılımının von Mises olduğu genellikle söylenememektedir. Bu durum, hipotez testi için dağılım teorisi çalışmalarında bir çok soruna yol açmaktadır. Fakat, bu dağılım ve bu dağılımın çok boyuta genellemeleri, dairesel veri analizinin kurulmasında temel teşkil etmektedir.

Gadsden ve Kanji (1981) tarafından, ortalama yön μ_0 için iki ardışık test geliştirilmiştir. Birinci test olan *U-testi* yoğunlaşma parametresi κ 'nın bilindiği varsayımına dayanmaktadır. U-testinde Eşitlik (7.1)'de verilen von Mises dağılımı için olasılık oranlarının ardışık testi kullanılmaktadır. Test için *karakteristik işlem fonksiyonu* ve *ortalama örneklem sayısı fonksiyonu* bulunmuştur. İkinci test olan *V-testi* ise, κ 'nın bilinmesine ilişkin herhangi bir varsayım bulunmamaktadır.

Fakat, μ_0 ve κ parametrelerinin fonksiyonu için test geliştirebilmek amacıyla *Cox teoremi* kullanılmaktadır (Wetherill 1975). Bu testte, her bir adımda değerlendirilmesi gereken karmaşık bir fonksiyon gerekmektedir.

7.1. Yoğunlaşma Parametresi Bilindiğinde Ortalama Yön İçin Ardışık Test (U-Testi)

7.1.1. Kuramsal tanım

Yoğunlaşma parametresi bilinen von Mises anakütlesinde ortalama yön test edilmek istendiğinde hipotezler;

$$H_0: \mu = \mu_0$$

$$H_1: \mu = \mu_1 \tag{7.3}$$

biçiminde kurulur. Olasılık oranlarının ardışık testi tekniğinin uygulanabilmesi için n -inci gözlem için olabilirlik fonksiyonunun incelenmesi gerekmektedir. Buradaki ardışık olarak gözlenen dairesel gözlem değerleri, sırasıyla $\theta_1, \theta_2, \dots, \theta_n$ ile ifade edildiği takdirde von Mises dağılımı için n -inci gözlemden sonra Kesim 6.1.'de tanımlanan olabilirlik oranı aşağıdaki gibi elde edilmektedir:

$$f(\theta_1, \dots, \theta_n; \mu_0) = \frac{1}{[2\pi I_0(\kappa)]^n} e^{\kappa \sum_{i=1}^n \cos(\theta_i - \mu_0)}$$

$$f(\theta_1, \dots, \theta_n; \mu_1) = \frac{1}{[2\pi I_0(\kappa)]^n} e^{\kappa \sum_{i=1}^n \cos(\theta_i - \mu_1)}$$

$$L_n = \prod_{i=1}^n \frac{f(\theta_i; \mu_1)}{f(\theta_i; \mu_0)} = \frac{\frac{1}{[2\pi I_0(\kappa)]^n} e^{\kappa \sum_{i=1}^n \cos(\theta_i - \mu_1)}}{\frac{1}{[2\pi I_0(\kappa)]^n} e^{\kappa \sum_{i=1}^n \cos(\theta_i - \mu_0)}}$$

$$\begin{aligned}
\ln L_n &= \kappa \left[\sum_{i=1}^n \cos(\theta_i - \mu_1) - \sum_{i=1}^n \cos(\theta_i - \mu_0) \right] \\
&= \kappa \left[\sum_{i=1}^n (\cos \theta_i \cos \mu_1 + \sin \theta_i \sin \mu_1) \right. \\
&\quad \left. - \sum_{i=1}^n (\cos \theta_i \cos \mu_0 + \sin \theta_i \sin \mu_0) \right] \\
&= \kappa \sum_{i=1}^n (\cos \theta_i \cos \mu_1 + \sin \theta_i \sin \mu_1 \\
&\quad - \cos \theta_i \cos \mu_0 - \sin \theta_i \sin \mu_0) \\
&= \kappa \sum_{i=1}^n [\cos \theta_i (\cos \mu_1 - \cos \mu_0) + \sin \theta_i (\sin \mu_1 - \sin \mu_0)] \\
&= \kappa \sum_{i=1}^n \left[\cos \theta_i \left\{ 2 \sin \left(\frac{\mu_1 + \mu_0}{2} \right) \sin \left(\frac{\mu_0 - \mu_1}{2} \right) \right\} \right. \\
&\quad \left. + \sin \theta_i \left\{ 2 \sin \left(\frac{\mu_1 - \mu_0}{2} \right) \cos \left(\frac{\mu_1 + \mu_0}{2} \right) \right\} \right]
\end{aligned}$$

$$\nu_1 = \frac{\mu_0 + \mu_1}{2} \text{ ve } \nu_2 = \frac{\mu_0 - \mu_1}{2} \text{ olsun.}$$

$$\begin{aligned}
\ln L_n &= 2\kappa \sum_{i=1}^n (\cos \theta_i \sin \nu_1 \sin \nu_2 + \sin \theta_i \sin(-\nu_2) \cos \nu_1) \\
&= 2\kappa \sum_{i=1}^n (\sin \theta_i \cos \nu_1 - \cos \theta_i \sin \nu_1) \sin(-\nu_2)
\end{aligned}$$

$$\ln L_n = 2\kappa \sum_{i=1}^n \sin(\theta_i - \nu_1) \sin(-\nu_2) \quad (7.4)$$

(Jeffrey 1995). Olasılık oranlarının ardışık testine göre;

$$\ln B < \ln L_n < \ln A \quad (7.5)$$

eşitsizliği kullanılarak (7.3)'de verilen hipotezlerin test edilebileceği bilinmektedir. Burada A ve B değerleri (6.5) eşitliğinde tanımlandığı gibi;

$$A = \frac{1-\beta}{\alpha}, \quad B = \frac{\beta}{1-\alpha}$$

olarak verilmektedir. Buna göre, n -inci gözlem elde edildikten sonra, (7.5) eşitsizliğinde L_n değeri yerine konulduğunda U-testi;

$$\ln B < 2\kappa \sum_{i=1}^n \sin(\theta_i - \nu_1) \sin(-\nu_2) < \ln A \quad (7.6)$$

biçiminde elde edilir.

7.1.2. Karakteristik işlem fonksiyonu

Karakteristik işlem fonksiyonu von Mises dağılımında μ test edilmek istendiğinde, (6.23) eşitliğinde verildiği gibi;

$$P(\mu) = \frac{A^h - 1}{A^h - B^h} \quad (7.7)$$

tanımlanır. Burada h ; Kesim 6.1.2.'de verilen yönteme göre,

$$E\left[\left[\frac{f(\theta: \mu_1)}{f(\theta: \mu_0)}\right]^h\right] = 1$$

olacak şekilde, μ parametresinin elde edilmesi gerekmektedir. Buna göre;

$$\left[\frac{f(\theta: \mu_1)}{f(\theta: \mu_0)}\right]^h = \left[\frac{\frac{1}{2\pi I_0(\kappa)} e^{\kappa \cos(\theta - \mu_1)}}{\frac{1}{2\pi I_0(\kappa)} e^{\kappa \cos(\theta - \mu_0)}}\right]^h = e^{\kappa h [\cos(\theta - \mu_1) - \cos(\theta - \mu_0)]}$$

$$E\left(\left[\frac{f(\theta:\mu_1)}{f(\theta:\mu_0)}\right]^h\right) = \int_0^{2\pi} \left[\frac{f(\theta:\mu_1)}{f(\theta:\mu_0)}\right]^h f(\theta) d\theta = 1$$

$$\int_0^{2\pi} \frac{1}{2\pi I_0(\kappa)} e^{\kappa \cos(\theta-\mu)} e^{\kappa h [\cos(\theta-\mu_1) - \cos(\theta-\mu_0)]} d\theta = 1 \quad (7.8)$$

integral çözümlendiğinde,

$$[1 + 2h^2 - 2h\cos(\mu - \mu_1) - 2h\cos(\mu - \mu_0) - 2h^2\cos(\mu_1 - \mu_0)]^{1/2} = 1 \quad (7.9)$$

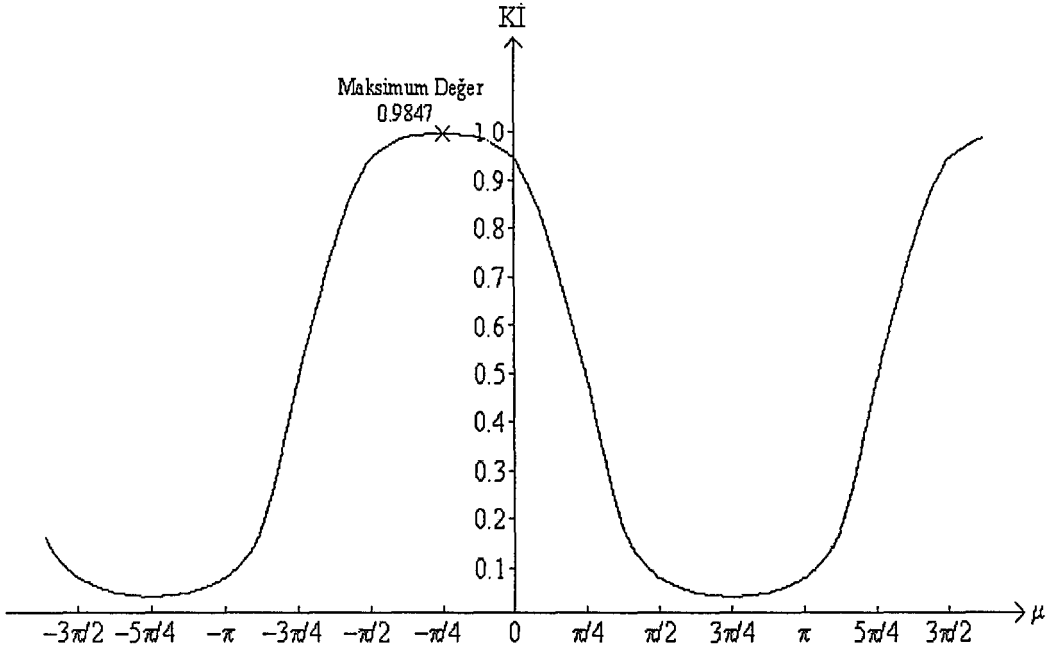
elde edilir. Burada μ , gerçek ortalama yönü göstermektedir. Bazı cebirsel ve trigonometrik işlemler sonucunda Eşitlik (7.9)'dan h değeri;

$$h = \frac{\sin(\mu - \nu_1)}{\sin \nu_2} \quad (7.10)$$

olarak elde edilmektedir (Gadsden ve Kanji 1981).

Doğrusal verilerde kabul etme olasılıkları çeşitli h değerleri için hesaplanmaktadır. Dairesel verilerde ise, doğrusal verilerden farklı olarak, karakteristik işlem fonksiyonunun minimum ve maksimum değerleri bulunmaktadır. Eşitlik (7.7)'deki karakteristik işlem fonksiyonunun μ 'ye göre türevi alındığında minimum değer $\mu = 90^\circ + \nu_1$ ve maksimum değer $\mu = 270^\circ + \nu_1$ olarak elde edilir.

Örnek olarak, Şekil 7.1.'de, $\alpha = \beta = 0.05$ hata olasılıkları ile $H_0: \mu_0 = 0^\circ$, $H_1: \mu_1 = 90^\circ$ hipotezlerinin ardışık testi için karakteristik işlem fonksiyonu eğrisi verilmektedir.



Şekil 7.1. $\mu_0 = 0$, $\mu_1 = 90^\circ$ ve $\alpha = \beta = 0.05$ iken U-testi için karakteristik işlem fonksiyonu

(Gadsden ve Kanji 1982-83).

7.1.3. U-testi için ortalama örneklem sayısı fonksiyonu

U-Testi için ortalama örneklem sayısı fonksiyonu;

$$E(n; \mu) = \frac{P(\mu) \ln B + [1 - P(\mu)] \ln A}{E(z; \mu)} \quad (7.11)$$

eşitliğinden elde edilir. Burada;

$$z = \ln \left[\frac{f(\theta; \mu_1)}{f(\theta; \mu_0)} \right] = \ln \left[\frac{\frac{1}{2\pi I_0(\kappa)} e^{\kappa \cos(\theta - \mu_1)}}{\frac{1}{2\pi I_0(\kappa)} e^{\kappa \cos(\theta - \mu_0)}} \right]$$

$$z = \kappa [\cos(\theta - \mu_1) - \cos(\theta - \mu_0)]$$

dır. Buradan;

$$E(z; \mu) = E[\kappa (\cos \theta \cos \mu_1 + \sin \theta \sin \mu_1 - \cos \theta \cos \mu_0 - \sin \theta \sin \mu_0)]$$

elde edilir. Von Mises anakütlesi için;

$$E(\cos\theta) = A(\kappa) , E(\sin\theta) = 0 , A(\kappa) = I_1(\kappa) / I_0(\kappa)$$

olarak tanımlanır (Mardia 1972). Dolayısıyla,

$$E(z; \mu) = E[\kappa(\cos\theta(\cos\mu_1 - \cos\mu_0) + \sin\theta(\sin\mu_1 - \sin\mu_0))]$$

$$E(z; \mu) = A(\kappa)(\cos\mu_1 - \cos\mu_0)$$

$$E(z; \mu) = A(\kappa) \left[2 \sin\left(\frac{\mu_1 + \mu_0}{2}\right) \sin\left(\frac{\mu_0 - \mu_1}{2}\right) \right]$$

$$E(z; \mu) = 2A(\kappa) \sin \nu_1 \sin \nu_2 \quad (7.12)$$

elde edilmektedir. Bu değer (7.11) eşitliğinde yerine yazıldığında, ortalama örneklem sayısı;

$$E(n; \mu) = \frac{P(\mu) \ln B + [1 - P(\mu)] \ln A}{2A(\kappa) \sin \nu_1 \sin \nu_2} \quad (7.13)$$

olarak bulunabilmektedir.

Dairesel verilerde ortalama örneklem sayısının maksimum ve minimum değerlerinin elde edilmesi mümkündür. Dolayısıyla doğrusal verilerde H_0 ve H_1 hipotezleri doğru iken elde edilen ortalama örneklem sayısı değerleri dairesel verilerde maksimum ve minimum değerler için bulunmaktadır. (7.13) eşitliğinin ν_2 'ye göre türevi alınır ve sıfıra eşitlenir. Sonuç olarak,

$$E(n; \mu)_{\min} = \frac{P(\mu) \ln B + [1 - P(\mu)] \ln A}{2A(\kappa) \sin \nu_1}$$

$$E(n; \mu)_{\max} = \infty$$

elde edilir (Gadsden ve Kanji 1982-83).

Örnek 7.1

Ghana Üniversitesi'nde yapılan bir çalışmada, örümcek ağlarının dünya yüzeyiyle yaptıkları açıların dik olduğu gözlenmiştir. Belli durumlarda, açılara ilişkin olarak yapılan bu çalışmanın, ağları yapan örümcek türlerinin belirlenmesinde gerekli olan sürecin çok uzun olmasını engellemesi beklenmektedir.

Isoxya cicatricosa türlerine ilişkin geçmişte yapılmış olan çok sayıda gözleme göre ağların ortalama açısı $\mu = 28.12^\circ$ ve yoğunlaşma parametresi $\kappa = 38.17$ olarak bulunmuştur. *Araneus rufipalpis* türleri için parametreler ise, $\mu = 15.66^\circ$ ve $\kappa = 37.94$ 'dür. Bir grup ağ gözlenmiş ve her bir ağın örüldüğü sırada yaptıkları açılar kaydedilmiştir;

$$25^\circ, 12^\circ, 31^\circ, 26^\circ, 17^\circ, 15^\circ, 24^\circ, 10^\circ, 16^\circ, 12^\circ$$

Bu ağları ören örümcek türlerinin yukarıda belirtilen iki türden birine ait olduğu düşünülmektedir. Buna göre test edilecek hipotezler;

$$H_0: \mu = 28.12^\circ$$

$$H_1: \mu = 15.66^\circ$$

olacaktır.

Yukarıdaki veriye, birleştirilmiş κ tahmini kullanılarak U-testi uygulanmıştır. Çünkü U-testinin varsayımına göre, yoğunlaşma parametresi her iki hipotez altında da aynıdır. Buna göre veriler birleştirildiğinde, $\kappa = 38.05$ değeri elde edilir. $\alpha = \beta = 0.05$ hata olasılıkları kullanıldığında, testin sınır değerleri;

$$-0.35635 < -\sum_{i=1}^n \sin(\theta_i - 21.89^\circ) < 0.35635$$

olarak bulunur. Hesaplamalar sonucunda;

<u>Gözlem</u>	<u>$-\Sigma \sin(\theta_i - 21.89^\circ)$</u>
1	-0.05425
2	0.11751
3	-0.04082
4	-0.11249
5	-0.02725
6	0.09271
7	0.05589
8	0.26192
9	0.36454

elde edilmiştir. Buna göre, $n = 9$ gözlemden sonra üst sınır aşıldığı için süreç sona ermektedir ve H_1 hipotezi kabul edilir. Sonuç olarak, örümcekler ikinci tür olan *Araneus rufipalpis* grubundan gelmektedir. Bu sonucun doğruluğu, daha sonra örümcek sınıflamasına uyarlanan genel yöntemlerle ispatlanmıştır (Gadsden ve Kanji 1981).

7.2. Yoğunlaşma Parametresi Bilinmediğinde Ortalama Yön İçin Ardışık Test (V-Testi)

Çoğu zaman yoğunlaşma parametresinin bilindiği varsayımı gerçekçi olmayabilmektedir. Bu durumun giderilmesi amacıyla, Gadsden ve Kanji (1981) tarafından V-testi tasarlanmıştır. V-testi yoğunlaşma parametresinin bilinmediği varsayımı altında ortalama yön için uygulanan bir ardışık testtir.

Von Mises dağılımına sahip bir anakütleden örneklem çekildiği, fakat yoğunlaşma parametresi κ 'nın bilinmediği varsayılınsın. Burada amaç, hipotezlerin kurulup test edilmesine ilişkin olarak μ ve κ parametrelerinin bir fonksiyonunu bulabilmektir. Bu fonksiyon, dağılımı bilinmeyen κ parametresinden bağımsız olmalıdır (Wetherill 1975). Dolayısıyla, testin çıkarımı Kesim 6.1.'de tanımlanan olasılık oranlarının ardışık testine dayalı olarak yapılabilmektedir.

(7.1) eşitliğinde verilen von Mises dağılımının n -inci gözlem için olabilirlik fonksiyonu (λ_n);

$$\begin{aligned}
\lambda_n &= \frac{1}{[2\pi I_0(\kappa)]^n} e^{\kappa \sum_{i=1}^n \cos(\theta_i - \mu)} \\
&= [2\pi I_0(\kappa)]^{-n} e^{\kappa \sum_{i=1}^n (\cos \theta_i \cos \mu + \sin \theta_i \sin \mu)} \\
&= [2\pi I_0(\kappa)]^{-n} e^{\kappa \cos \mu \sum_{i=1}^n \cos \theta_i + \kappa \sin \mu \sum_{i=1}^n \sin \theta_i}
\end{aligned}$$

Burada, $C = \sum_{i=1}^n \cos \theta_i$, $S = \sum_{i=1}^n \sin \theta_i$, $\tau = \cos \mu$, $\nu = \sin \mu$ 'dür. Buna göre;

$$\lambda_n = [2\pi I_0(\kappa)]^{-n} e^{\kappa \tau C + \kappa \nu S} \quad (7.14)$$

elde edilir. $C, S; \mu$ ve κ parametreleri için yeterli istatistiklerdir (Mardia 1972) ve bileşke uzunluğu $R = \sqrt{C^2 + S^2}$ 'nin dağılımı, bu dağılımın μ, κ 'dan bağımsız olmasını sağlayan iki hipotez için ardışık test bulmada kullanılabilir.

Yönsel veri dağılım teorisine göre, R 'nin dağılımı, bu bağımsızlık kriterini sağlamamaktadır. Fakat, ardışık sürecin her aşamasında C bilindiği için, $R|C$ koşullu dağılımının bulunması gerekmektedir. $R|C$ 'nin olasılık yoğunluk fonksiyonu;

$$f(R|C) = \frac{\frac{1}{2} R \phi_n(R^2) e^{\Delta \sqrt{R^2 - C^2}}}{\sqrt{R^2 - C^2} \int_0^{\infty} \cos Ct J_0^n(t^2 - \Delta^2)^{1/2} dt} \quad (7.15)$$

olarak tanımlanır. Burada, $R \phi_n(R^2)$;

$$R \phi_n(R^2) = R \int_0^{\infty} u J_0(Ru) J_0^n(u) du$$

dur. $J_p(u)$ değeri; p -inci sıra standart Bessel fonksiyonudur ve

$$J_p(u) = \sum_{r=0}^{\infty} (-1)^r [\Gamma(p+r+1)\Gamma(r+1)]^{-1} \left(\frac{1}{2}u\right)^{2r+p}$$

olarak tanımlanır. Δ ise,

$$\Delta = \kappa v = \kappa \sin \mu$$

olarak verilir (Mardia 1972).

Bu fonksiyon, Δ formu dışında, μ ve κ 'dan bağımsızdır. Dolayısıyla (7.3) eşitliğindeki hipotez,

$$H_0: \Delta = \Delta_0$$

$$H_1: \Delta = \Delta_1$$

biçiminde tanımlanır.

Ardışık test için gerekli olan olabilirlik oranı;

$$L_n = \frac{f(R|C|\Delta_1, n)}{f(R|C|\Delta_0, n)} = \frac{\frac{1}{2} R \phi_n(R^2) e^{\Delta_1 \sqrt{R^2 - C^2}}}{\sqrt{R^2 - C^2} \int_0^{\infty} \cos Ct J_0^n(t^2 - \Delta_1^2)^{1/2} dt} \cdot \frac{\frac{1}{2} R \phi_n(R^2) e^{\Delta_0 \sqrt{R^2 - C^2}}}{\sqrt{R^2 - C^2} \int_0^{\infty} \cos Ct J_0^n(t^2 - \Delta_0^2)^{1/2} dt}$$

$$L_n = \frac{e^{(\Delta_1 - \Delta_0) \sqrt{R^2 - C^2}} \int_0^{\infty} \cos Ct J_0^n(t^2 - \Delta_0^2)^{1/2} dt}{\int_0^{\infty} \cos Ct J_0^n(t^2 - \Delta_1^2)^{1/2} dt}$$

olarak elde edilir. Eşitliğin her iki tarafının logaritmaları alınırsa;

$$\ln L_n = (\Delta_1 - \Delta_0)\sqrt{R^2 - C^2} + \ln \left[\int_0^\infty \cos Ct J_0^n(t^2 - \Delta_0^2)^{1/2} dt \right] - \ln \left[\int_0^\infty \cos Ct J_0^n(t^2 - \Delta_1^2)^{1/2} dt \right] \quad (7.16)$$

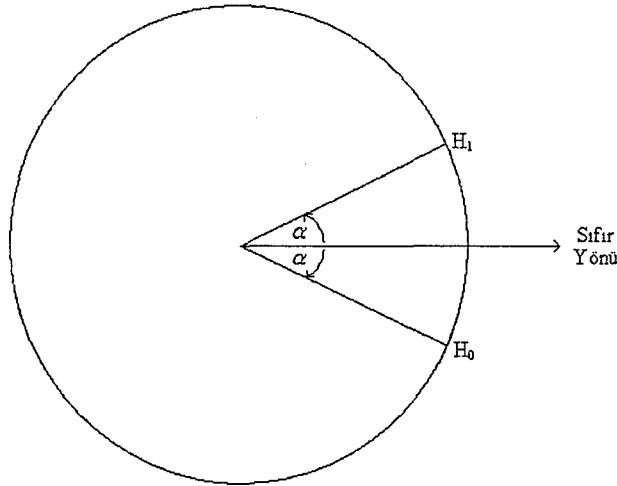
bulunur. Buna göre test;

$$\ln B < (\Delta_1 - \Delta_0)S + \ln \left[\int_0^\infty \cos Ct J_0^n(t^2 - \Delta_0^2)^{1/2} dt \right] - \ln \left[\int_0^\infty \cos Ct J_0^n(t^2 - \Delta_1^2)^{1/2} dt \right] < \ln A \quad (7.17)$$

olmaktadır. Eğer, $-\Delta_1 = \Delta_0$ ise, V-testi;

$$\ln B < (\Delta_1 - \Delta_0)S < \ln A \quad (7.18)$$

eşitsizliğine indirgenmektedir. Bu test, sıfır yönünden ve sıfır yönünün tersinden açısal olarak eşit uzaklıklardaki iki ortalama yönün testidir (bkz. Şekil 7.2.).



Şekil 7.2. Zıt yönlerde, fakat ortalama yönü açısal olarak eşit hipotezler için V-testi

Uygulamada bu testin $\Delta_0 = -\Delta_1$ durumu dışında kullanılması pek mümkün olmamaktadır (Gadsden ve Kanji 1981).

Ayrıca, bu testte, $E(S) = E\left(\sum_{i=1}^n \sin \theta_i\right) = 0$ olduğu için, ortalama örneklem sayısı tanımsızdır.

8. UYGULAMA

Bu bölümde, von Mises dağılımına sahip dairesel verilere ilişkin 2. Bölüm'de verilen önemli şematik gösterimlerin yapılması, 3. Bölüm'de verilen tanımlayıcı istatistiklerin elde edilmesi, 5. Bölüm'de verilen tek örneklem ortalama yön testlerinin yapılması ve 7. Bölüm'de verilen ortalama yön için ardışık testin kullanımı uygulamalı olarak incelenecektir. İlk olarak, von Mises dağılımından veri türetilerek yukarıda belirtilen işlemler yapılacak, daha sonra ise meteoroloji alanında elde edilen bir veri seti için uygulama yapılacaktır.

Bu amaçla, *Oriana for Windows Version 1.06* ve *S-Plus 6.1 for Windows* olmak üzere iki tane paket programdan yararlanılmıştır. Oriana paket programı dairesel veriler için hazırlanmış bir programdır. Menülerinde temel istatistiklerin hesaplanması, F-testleri, Ki-kare testleri gibi testler bulunmaktadır ve dairesel histogram grafik çeşitleri ile dağılım grafikleri çizilebilmektedir. S-Plus paket programında ise, dairesel veriler için *CircStats Package* yazılımından yararlanılmaktadır. Bu yazılım sayesinde veri dönüşümleri yapılabilmekte, dairesel grafikler çizilebilmekte, parametre değerleri hesaplanabilmekte, dairesel yoğunluk fonksiyonlarından veri türetilenilmekte, çeşitli hipotez testleri yapılabilmekte ve Bessel fonksiyonları hesaplanabilmektedir.

8.1. Veri Türetimi Uygulaması

8.1.1. Veri türetimi işlemi

Von Mises dağılımından veri türetimi işlemi için EK-4'te verilen programdan yararlanılmıştır. Burada, ortalama yön $\mu = 100^\circ$ ve yoğunlaşma parametresi $\hat{\kappa} = 15$ değerleri kullanılarak, von Mises dağılımından $N = 200$ hacimlik veri türetilmiş ve bu anakütleden $n = 50$ hacimlik örneklem seçilmiştir. Türetilen verilerden elde edilmiş olan örneklem değerleri Çizelge 8.1.'de verilmektedir:

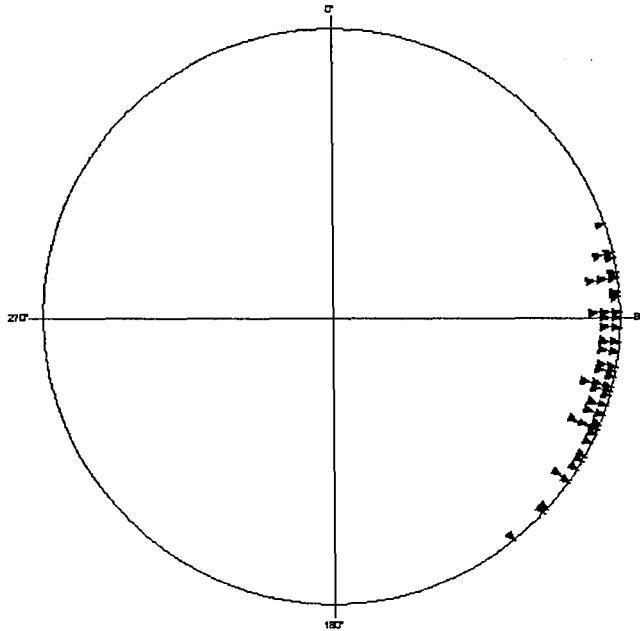
Çizelge 8.1. Von Mises dağılımından türetilen verilerden seçilen örneklem

141.89947	101.91036	114.17793	81.27442	116.84536	95.79061
82.29468	89.05286	108.41218	100.40679	100.31639	82.70915
77.53981	97.30467	105.58327	78.07436	82.48128	102.75578
119.83104	101.54473	110.72528	86.47231	92.18931	85.73791
104.25110	105.96573	120.21541	112.49771	92.46383	71.44693
89.65603	110.65147	95.67104	106.65362	113.05267	104.11281
77.82523	89.38608	122.65788	113.91254	113.63390	133.54389
125.16530	132.22788	108.29909	90.26493	125.26151	90.81131
104.10834	97.44142				

8.1.2. Şematik gösterimler

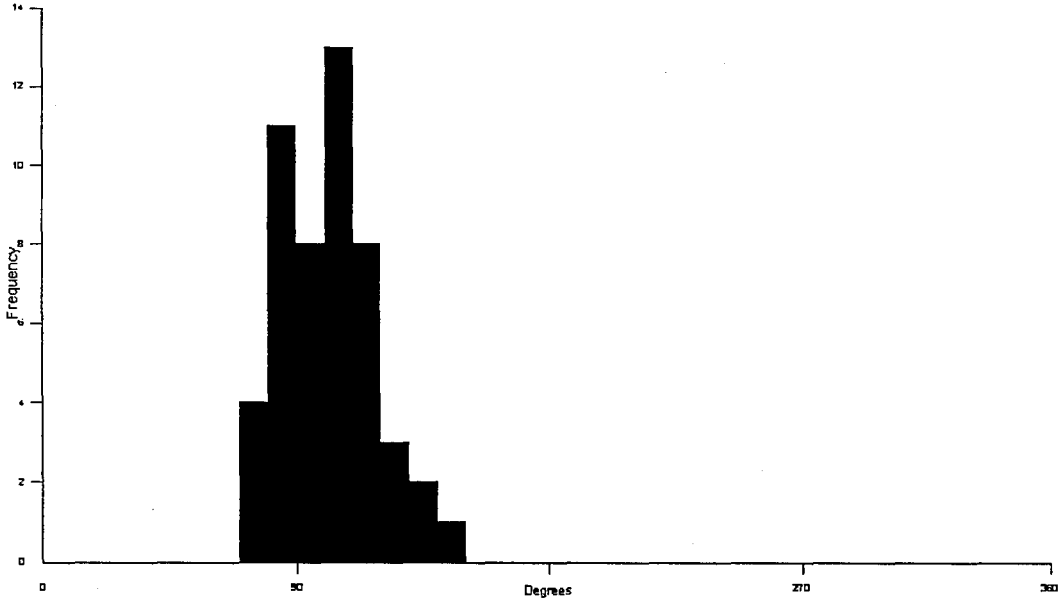
Kesim 8.1.1.'de von Mises dağılımına sahip anakütleden çekilen 50 gözlem değeri için Oriana paket programı kullanılarak elde edilen dairesel grafikler aşağıda verilmektedir.

Şekil 8.1.'de, örneklem verilerinin 1. tür ham veri grafiği görülmektedir.



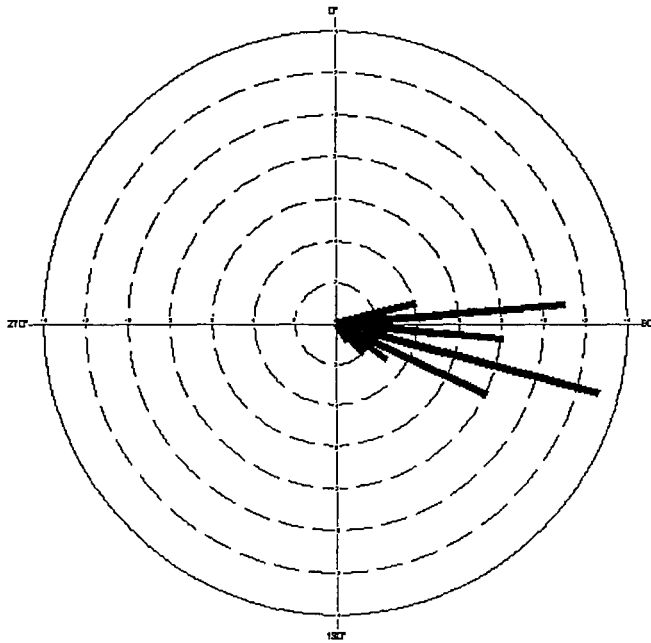
Şekil 8.1. Von Mises dağılımından türetilen verilerden seçilen örneklemin ham veri grafiği

Şekil 8.2.'de, örneklemin 10° 'lik grup genişliği seçildiği durumdaki doğrusal histogramı görülmektedir.



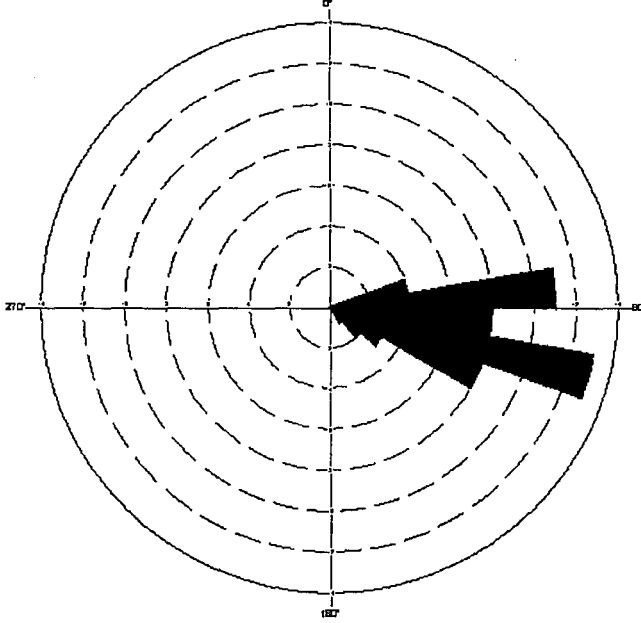
Şekil 8.2. Von Mises dağılımından türetilen verilerden seçilen örneklemin doğrusal histogramı

Şekil 8.3.'te, örneklemin 10° 'lik grup genişliği seçildiği durumdaki dairesel histogramı görülmektedir.



Şekil 8.3. Von Mises dağılımından türetilen verilerden seçilen örneklemin dairesel histogramı

Şekil 8.4.'te, örneklemin 10° 'lik grup genişliği seçildiği durumdaki gül şeması görülmektedir.



Şekil 8.4. Von Mises dağılımından türetilen verilerden seçilen örneklemin gül şeması

8.1.3. Tanımlayıcı istatistiklerin hesaplanması

Kesim 8.1.1.'de von Mises dağılımından türetilen 200 dairesel gözlem değerinden seçilmiş olan 50 hacimlik örneklem için S-Plus ve Oriana paket programları kullanılarak elde edilen tanımlayıcı istatistik değerleri;

Ortalama Yön: $\bar{\theta} = 102.0746^\circ$

Bileşke Uzunluğu: $R = 48.10479$

Ortalama Bileşke Uzunluğu: $\bar{R} = 0.9620959$

Yoğunlaşma Parametresi tahmini: $\hat{\kappa} = 13.45585$

Dairesel Varyans: $V = 0.03790414$

Dairesel Standart Sapma: $\nu = 15.91^\circ$

Dairesel Standart Hata: $\sigma = 2.25^\circ$

olarak bulunmuştur. Bu değerlerin hesaplanmasında kullanılan S-Plus programları EK-4'te verilmektedir.

8.1.4. Tek örneklem ortalama yön testleri

Kesim 8.1.3.'te verildiği gibi, örneklem 102.0746° ortalama yönü ve 13.45585 yoğunluk parametresi ile von Mises dağılımına $VM(102.07, 13.46)$ sahiptir.

Güven aralığının belirlenmesi yöntemiyle ortalama yön test edilmek istendiğinde kurulacak hipotezler;

$$H_0: \mu = 100^\circ$$

$$H_1: \mu \neq 100^\circ$$

biçiminde olsun. Eşitlik (5.5)'e göre;

$$\sigma_{VM} = \frac{1}{\sqrt{nR\hat{\kappa}}} = \frac{1}{\sqrt{50(0.9620959)(13.45585)}} = 0.0393052$$

elde edilir. Buna göre; ortalama yön μ için $(1-\alpha)$ 'lık güven aralığı (5.6)'dan yararlanılarak;

$$(102.0746 - \sin^{-1}((1.9604)(0.0393052)), 102.0746 + \sin^{-1}((1.9604)(0.0393052))) \\ = (97.65535548^\circ, 106.4938445^\circ)$$

olacaktır. Burada $\alpha = 0.05$ olarak alınmıştır. H_0 hipotezinde verilen 100° değeri verilen aralığın içinde yer aldığı için, 0.05 anlamlılık seviyesinde H_0 hipotezi kabul edilir. Örneklem ortalama yönü 100° 'ye eşittir.

Doğrudan test uygulanmak istendiğinde daha önce de belirtildiği gibi, yoğunlaşma parametresi κ 'nın bilindiği ve bilinmediği durum olmak üzere iki durum söz konusudur.

Önce κ 'nın bilinmediği durum göz önüne alınsın. Buradaki test istatistiği (5.7) eşitliğinden;

$$E_n = \frac{\sin(102.0746 - 100)}{0.0393052} = 0.9210153$$

olarak bulunur. $|0.9210153| < 1.9604$ olduğundan 0.05 anlamlılık seviyesinde H_0 hipotezi kabul edilir. Ortalama yön 100° 'dir.

κ 'nın bilindiği durumda ise, $\kappa_0 = 15$ değeri göz önüne alınsın. $\kappa_0 \geq 2$ olduğundan, von Mises dağılımı için ortalama yönün dairesel standart hatası σ_{VM} ;

$$\sigma_{VM} = \frac{1}{\sqrt{50(0.9620959)(15)}} = 0,0372272$$

olarak hesaplanır. Test istatistiği;

$$E_n = \frac{\sin(102.0746 - 100)}{0.0372272} = 0.9724258$$

bulunur. $|0.9724258| < 1.9604$ olduğundan 0.05 anlamlılık seviyesinde H_0 hipotezi kabul edilir. Örnekleme ortalama yönü 100° 'ye eşittir.

8.1.5. Ortalama yön için ardışık testler

Bu kesimde, von Mises dağılımından türetilen veriler için ilk olarak yoğunlaşma parametresinin bilindiği durumda ortalama yön için ardışık test olan U-testi, ardından yoğunlaşma parametresinin bilinmediği durumda kullanılan V-testinin uygulamaları yapılacaktır. Bu testlerin bilgisayar ortamında yapılabilmesi amacıyla, S-Plus'ta ayrı ayrı programları yazılmıştır. Bu programlar yardımıyla, araştırmacılar önceden belirledikleri birinci ve ikinci tip hata olasılıklarında, kurdukları hipotez doğrultusunda U-testi ile V-testi için sınır değerlerini ve aşama aşama test istatistiği değerlerini görebilmekte ve testi sonuçlandırabilmektedirler. Yazılan programlar EK-4'te verilmektedir. U ve V testlerine ilişkin algoritmalar aşağıda verildiği gibidir:

U-testinin algoritması:

- Adım 1: $\kappa, \mu_0, \mu_1, \alpha, \beta$ değerlerini belirle
- Adım 2: ν_1 ve ν_2 değerlerini hesaplatır
- Adım 3: Alt ve üst sınır değerlerini hesaplatır
- Adım 4: i -inci gözlem değerini gir ($i = 1, \dots, n$)
- Adım 5: Test istatistiği \leq Alt sınır ise H_0 kabul. i gözlemden sonra iterasyon durur
- Adım 6: Test istatistiği \geq Üst sınır ise H_0 red. i gözlemden sonra iterasyon durur
- Adım 7: Alt sınır $<$ Test istatistiği $<$ Üst sınır ise bir gözlem daha ekle ve Adım 4'e git

V-testinin algoritması:

- Adım 1: Δ_0, α, β değerlerini belirle
- Adım 2: Δ_1 değerini hesaplatır
- Adım 3: Alt ve üst sınır değerlerini hesaplatır
- Adım 4: i -inci gözlem değerini gir ($i = 1, \dots, n$)
- Adım 5: Test istatistiği \leq Alt sınır ise H_0 kabul. i gözlemden sonra iterasyon durur
- Adım 6: Test istatistiği \geq Üst sınır ise H_0 red. i gözlemden sonra iterasyon durur
- Adım 7: Alt sınır $<$ Test istatistiği $<$ Üst sınır ise bir gözlem daha ekle ve Adım 4'e git

Test edilecek hipotezler;

$$H_0: \mu = 105^\circ$$

$$H_1: \mu = 100^\circ$$

biçiminde kurulsun.

$\hat{\kappa} = 13.45585$ değeri bilindiğine göre; $\alpha = 0.05$ ve $\beta = 0.10$ hata olasılıkları seçildiğinde, U-testi için elde edilen bilgisayar çıktısı aşağıda verilmektedir:

lower limit---> -2.25129179860649

upper limit---> 2.89037175789616

Observation No	Theta	Test statistics
1	87.62306	0.301384459497434
2	92.97328	0.495669153855559
3	76.55710	1.00920897291002
4	141.89947	0.264125071199501
5	96.74280	0.381879843757719
6	100.34360	0.426049597743889
7	104.45389	0.386026267675253
8	97.99723	0.478183818011891
9	89.94386	0.733378961543391
10	83.82249	1.10930114679813
11	85.51894	1.45213699931688
12	101.91036	1.46421728858088
13	103.36845	1.44642532578375
14	92.89950	1.64220065658803
15	94.34748	1.80866595823194
16	114.17793	1.5710624486873
17	81.27442	1.99605192898612
18	110.32594	1.83621278118561
19	102.97693	1.82644168738494
20	117.73713	1.51793105947162
21	116.84536	1.22708542045037
22	107.73308	1.12001945101823
23	95.79061	1.25716701296877
24	114.78015	1.00749390854477
25	100.14020	1.05582765347201
26	72.45115	1.6436300891441
27	65.03213	2.35771565769417
28	82.29468	2.76315378924454
29	70.21753	3.39011130803455

Alternative hypothesis is accepted.

After 29 observations iteration is stopped.

Bu sonuçlara göre, $n = 29$ gözlemden sonra üst sınır aşıldığı için süreç sona ermektedir ve H_0 hipotezi reddedilir. Sonuç olarak U-testine göre ortalama yön 100° olarak bulunmuştur.

Yoğunlaşma parametresinin bilinmediği varsayalım. Buna göre V-testi uygulamak gerekmektedir. Buradaki hipotezler örneğin,

$$H_0: \delta = 0.145$$

$$H_1: \delta = -0.145$$

olarak kurulmuş olsun. $\alpha = 0.05$ ve $\beta = 0.10$ hata olasılıkları seçildiğinde, V-testi için elde edilen bilgisayar çıktısı aşağıda verilmektedir:

lower limit--->	-2.25129179860649	
upper limit--->	2.89037175789616	
Observation No	Theta	Test statistics
1	87.62306	-0.289750483829972
2	92.97328	-0.579360096018346
3	76.55710	-0.861414701881168
4	141.89947	-1.04035721667345
5	96.74280	-1.32835135189326
6	100.34360	-1.61363847289753
7	104.45389	-1.89445963668165
8	97.99723	-2.18163932919708
9	89.94386	-2.47163918999699
Null hypothesis is accepted.		
After 9 observations iteration is stopped.		

Bu sonuçlara göre, $n = 9$ gözlemden sonra alt sınır aşıldığı için süreç sona ermektedir ve H_0 hipotezi kabul edilir.

8.2. Meteorolojik Verilere Uygulama

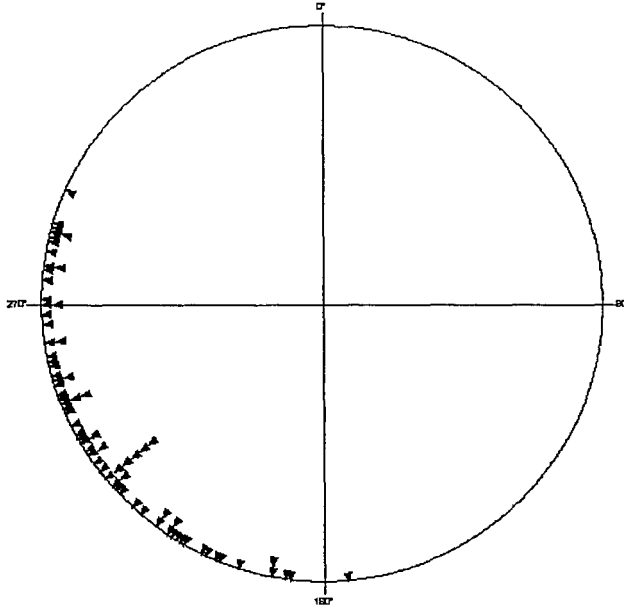
Dairesel verilerin en çok kullanıldığı alanlardan birisi meteorolojik gözlem çalışmalarıdır. Çalışmada incelenen yöntemlerin gerçek veriler üzerinde uygulanması amacıyla, Anadolu Üniversitesi Sivil Havacılık Yüksekokulu Meydan Meteoroloji İstasyon Müdürlüğü'nden, Anadolu Üniversitesi havaalanı için ölçülmüş olan *rüzgar yönlerinden* oluşan veri değerleri elde edilmiştir. Bu veriler ardışık açısız gözlemlerden oluşmaktadır. Gözlemler 15'er saniyelik zaman aralıklarında ölçülmektedir. Dolayısıyla ardışık veridir. Bu kesimde, önce Kesim 8.1.'de olduğu gibi, elde edilen veriler paket programlar yardımıyla şematik olarak verilecek, tanımlayıcı istatistikleri elde edilecek ve çalışmanın temel amacı olan daireysel verilerde ardışık çözümleme konusu içerisinde verilen U-testinden elde edilen sonuçların değerlendirilmesi ve yorumları yapılacaktır.

Uygulamada ele alınan veriler 28 Eylül 2002 tarihinde saat 12:00'dan 18:00'a kadar yapılan 5'er dakikalık rüzgar yönü ölçümlerinden oluşmaktadır. Öncelikle bu gözlem değerleri için inceleme yapılacak, ardından 28 Eylül 2002 tarihindeki gözlemler temel alınarak 29 Eylül 2002 tarihindeki ortalama yön için ardışık test uygulanacaktır.

8.2.1. Rüzgar yönü verilerinin şematik gösterimleri

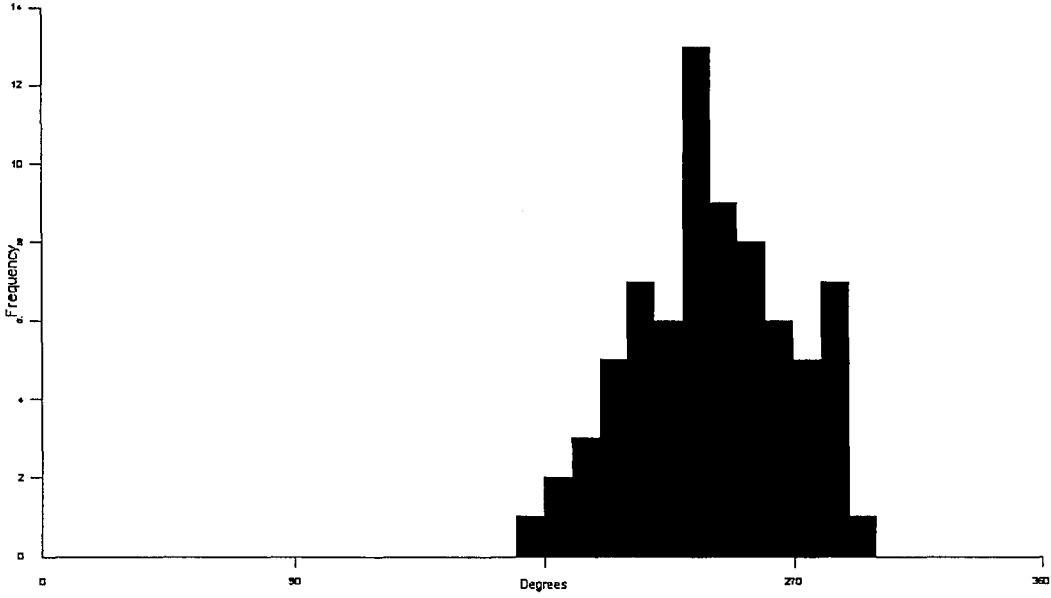
28 Eylül tarihinde saat 12:00'dan 18:00'a kadar toplam 73 rüzgar yönü elde edilmiştir. Bu gözlem değerleri için Oriana paket programı kullanılarak elde edilen dairesel grafikler aşağıda verilmektedir.

Şekil 8.5.'te, verilerin 1. tür ham veri grafiği görülmektedir.



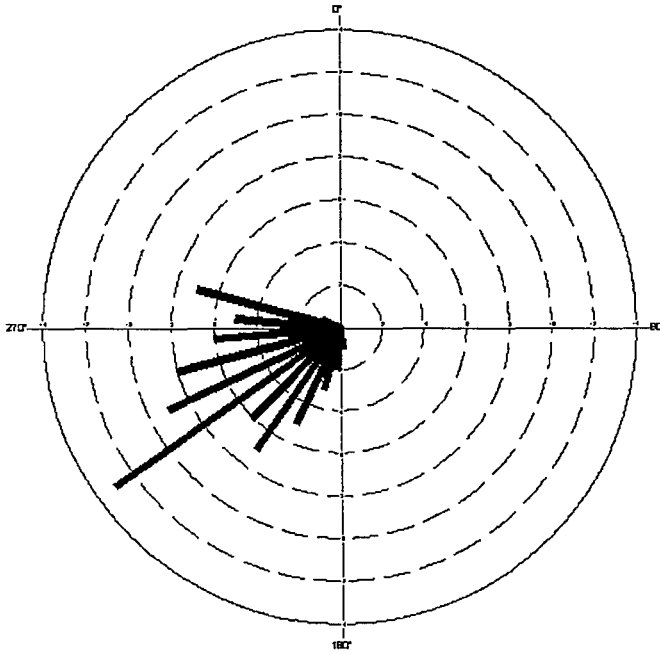
Şekil 8.5. Rüzgar yönü verilerinin ham veri grafiği

Şekil 8.6.'da, rüzgar yönü verilerinin 10°'lik grup genişliği seçildiği durumdaki doğrusal histogramı görülmektedir.



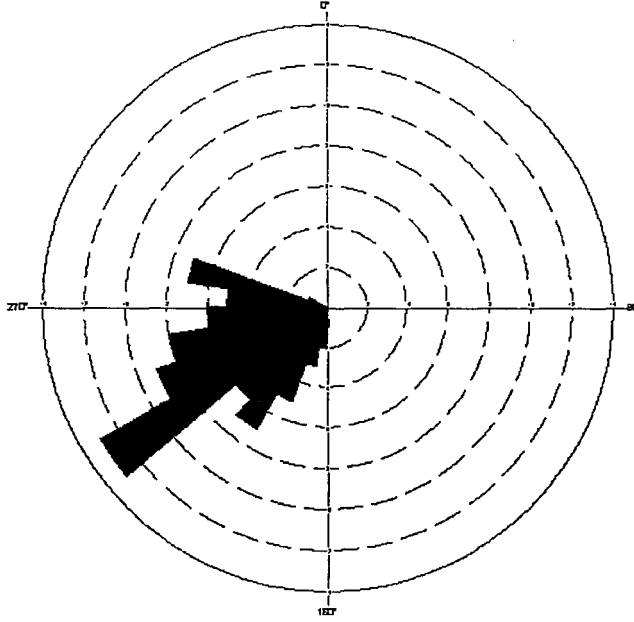
Şekil 8.6. Rüzgar yönü verilerinin doğrusal histogramı

Şekil 8.7.'te, rüzgar yönü verilerinin 10° 'lik grup genişliği seçildiği durumdaki dairesel histogramı görülmektedir.



Şekil 8.7. Rüzgar yönü verilerinin dairesel histogramı

Şekil 8.8.'de, rüzgar yönü verilerinin 10° 'lik grup genişliği seçildiği durumdaki gül şeması görülmektedir.



Şekil 8.8. Rüzgar yönü verilerinin gül şeması

Şekillerde görüldüğü gibi veriler 175° – 294° aralığında değerler almakta ve dağılımın şekli von Mises dağılımına uymaktadır.

8.2.2. Tanımlayıcı istatistiklerin hesaplanması

Belirlenen gün ve saat için seçilen 73 gözlem değeri için S-Plus ve Oriana paket programları kullanılarak elde edilen tanımlayıcı istatistik değerleri;

Ortalama Yön: $\bar{\theta} = 240.968^\circ$

Bileşke Uzunluğu: $R = 64.46779$

Ortalama Bileşke Uzunluğu: $\bar{R} = 0.8831204$

Yoğunlaşma Parametresi tahmini: $\hat{\kappa} = 4.576622$

Dairesel Varyans: $V = 0.1168796$

Dairesel Standart Sapma: $\nu = 28.57^\circ$

Dairesel Standart Hata: $\sigma = 3.34^\circ$

olarak bulunur.

8.2.3. Ortalama yön için ardışık test

Bu kesimde, 29 Eylül 2002 tarihinde saat 12:00'dan başlayarak ölçümler ardışık olarak elde edilmiştir. Veriler için yoğunlaşma parametresinin bilindiği durumda ortalama yön için ardışık test olan U-testi uygulanacaktır. Bu nedenle μ_0 değeri olarak 28 Eylül 2002 tarihinde elde edilmiş olan ortalama yön kullanılacak, 29 Eylül 2002 tarihinde de ortalama yönün bu değere eşit olup olmadığı test edilecektir. Bunun için alternatif hipotez olarak $\mu_1 = 230^\circ$ değeri alınsın.

Buna göre, test edilecek hipotezler;

$$H_0: \mu = 241^\circ$$

$$H_1: \mu = 230^\circ$$

olacaktır.

28 Eylül tarihinden $\hat{\kappa} = 4.576622$ değeri bilindiğine göre; $\alpha = 0.05$ ve $\beta = 0.10$ hata olasılıkları seçildiğinde, U-testi için bilgisayar çıktısı aşağıda verilmektedir:

lower limit---> -2.25129179860649

upper limit---> 2.89037175789616

Observation No	Theta	Test statistics
1	251	-0.234448109529761
2	261	-0.612135304982304
3	231	-0.543303174886509
4	217	-0.264931939696171
5	226	-0.120135747830682
6	240	-0.188967877926477
7	227	-0.0592947466996328
8	233	-0.0210274773922684
9	233	0.0172397919150959
10	241	-0.0668456445202896
11	250	-0.286503911705044
12	230	-0.202418475269658
13	235	-0.194762689523755
14	245	-0.339558881389244
15	251	-0.574006990919005
16	248	-0.763889368816257
17	243	-0.878399939712512
18	243	-0.992910510608767
19	243	-1.10742108150502
20	234	-1.08445605629382
21	221	-0.86479778910906
22	228	-0.750287218212805
23	255	-1.04313582867026
24	255	-1.33598443912772
25	234	-1.31301941391651
26	239	-1.3665772707391
27	246	-1.52645241692052
28	260	-1.89026261589827
29	248	-2.08014499379552
30	248	-2.27002737169277

Null hypothesis is accepted.

After 30 observations iteration is stopped.

Bu sonuçlara göre, $n = 30$ gözlemden sonra alt sınır aşıldığı için süreç sona ermektedir ve H_0 hipotezi kabul edilir. Sonuç olarak U-testine göre ortalama yön 241° olarak bulunmuştur. Buna göre 29 Eylül tarihinde de saat 12:00 ile 18:00 saatleri arasındaki ortalama rüzgar yönü 241° olacaktır. Dolayısıyla 29 Eylül tarihinde Anadolu Üniversitesi havaalanından kalkış ya da iniş yapacak olan uçaklar bu rüzgar yönü bilgisinden de yararlanabileceklerdir.

8.3. Sonular ve Tartışma

Önceki kesimlerde verilen test sonuçlarından da görüldüğü gibi, ardışık test sabit örneklem hacimli testlere göre örneklem hacmi yönünden daha avantajlıdır. Özellikle verilerin ardışık olarak geldiği arařtırmalarda bu tür ardışık testlerin kullanımı kolaylık sağlamaktadır.

Yapılan herhangi bir alıřmada elde edilen gözlem deęerleri doğrusal ya da dairesel olabilmektedir. Fakat, dairesel verilerin kullanıldığı arařtırmalarda bilinen istatistiksel yöntemlerin kullanılması arařtırmacıyı yanlış sonuçlara götürmektedir. Bu alıřmada, dairesel veriler için istatistiksel gösterim yöntemleri, tanımlayıcı istatistiklerin hesaplanması, ortalama yön için hipotez testleri ve verilerin ardışık olarak geldiği durumlarda kullanılan ortalama yön testi incelenmiştir.

Günümüzde yapılan arařtırmaların çoğunda örneklem hacmi önceden belirlenen sabit bir deęerdir. Sabit örneklem hacimli test sürecinde örneklemin önceden belirlenmesi gerekir. Bu yöntem uygulamada çoğu kez zaman kaybına ve yüksek bir maliyete neden olmaktadır. Ardışık test kullanıldığında bu zorluklar giderilebilmektedir. Test için, örneklem hacminin önceden bilinmesine gerek yoktur. Örneklem hacmi bir rassal deęişkendir. Test, tek bir gözlemlerle başlar ve gerekli istatistikle karşılaştırılarak hipotezler hakkında bir karara ulařılana kadar devam eder. Böylece örneklem hacminde büyük tasarruf sağlanmış olur (Peker 1998).

alıřmada kullanılan meteorolojik veriler ardışık olarak geldiği için, ölçüm sonuçları seri halde elde edilmektedir. Bundan dolayı, sabit örneklem hacimli testlerin kullanılması uygun deęildir. Elde edilen birikimli verileri ardışık olarak test etmek daha uygun bir yoldur. Bu amaçla alıřmada, verilerin ardışık olarak geldiği dairesel rüzgar yönü verilerinin testi için geliştirilen U ve V testleri incelenmiştir. Uygulama sonuçlarında da görüldüğü gibi, U-testi uygulanması sonucunda rüzgar yönü verileri için toplam 73 gözlemin elde edilmesini beklemek yerine 30 gözlem sonunda ortalama yönün ne olacağı bulunmuştur. Bu da havaalanından kalkış ya da iniş yapacak olan uçaklar için bir ön bilgi sağlayacak, uçuşlarını bu test sonuçlarını da kullanarak gerçekleştirebileceklerdir. Ayrıca,

hazırlanmış olan programlar yardımıyla da istenen hata olasılıklarında kaç gözlemle testin sonuçlandırılacağı görülmektedir. Uygulanan test sonucunda istenen verim elde edilememişse, hata olasılıkları üzerinde değişiklik yapılarak program sayesinde kolayca sonuca ulaşılabilmektedir.

KAYNAKLAR

AJNE, B., *A simple test for uniformity of a circular distribution*, Biometrika, **55**, 343-354 (1968).

ANDERSON-COOK, C. M., *An industrial example using one-way analysis of circular-linear data*, Computational Statistics & Data Analysis, **33**, 45-57 (2000).

ARTES, R., PAULA, G. A. ve RANVAUD, R., *Analysis of circular longitudinal data based on generalized estimating equations*, Aust. N. Z. J. Statist., **42**(3), 347-358 (2000).

BACANLI, S., *Korelasyon katsayısının ardışık testi*, Hacettepe Üniversitesi Fen Bilimleri Enstitüsü Bilim Uzmanlığı Tezi, Ankara (1988).

BOLTON, W., *Complex numbers*, Longman Group Limited, Essex, England (1995).

CHEENEY, R. F., *Statistical methods in geology*, George Allen & Unwin (Publishers) Ltd., London, UK (1983).

COBLENTZ, D. D. ve RICHARDSON, R. M., *Statistical trends in the intraplate stress field*, Journal of Geophysical Research, **100**, No. B10, 20245-20255 (1995).

COLLETT, D., *Outliers in circular data*, Appl. Statist., **29**, No. 1, 50-57 (1980).

ÇINGİ, H., *Ardışık çözümleme lisansüstü ders notları*, Hacettepe Üniversitesi Fen Fakültesi İstatistik Bölümü, Ankara (1990).

DAVIS, J. C., *Statistics and data analysis in geology*, 2nd ed., John Wiley, New York, USA (1986).

FISHER, N. I. ve LEE, A. J., *A correlation coefficient for circular data*, Biometrika, **70**, 2, 327-32 (1983).

FISHER, N. I., LEWIS, T. ve EMBLETON, B. J. J., *Statistical analysis of spherical data*, Cambridge University Press, Cambridge, Great Britain (1987).

FISHER, N. I., *Statistical analysis of circular data*, Cambridge University Press, Cambridge, Great Britain (1993).

GADSDEN, R. J. ve KANJI, G. K., *Sequential analysis for angular data*, The Statistician, **30**, No. 2, 119-29 (1981).

GADSDEN, R. J. ve KANJI, G. K., *Sequential analysis applied to circular data*, Commun. Statist.-Sequential Analysis, **1**(4), 304-314 (1982-83).

GUMBEL, E. J., GREENWOOD, J. A. ve DURAND, D., *The circular normal distribution: theory and tables*, J. Amer. Statist. Ass., **48**, 131-152 (1953).

GUMBEL, E. J., *Applications of the circular normal distribution*, J. Amer. Statist. Ass., **49**, 267-297 (1954).

İNAL, H. C. ve GÜNAY, S., *Olasılık ve matematiksel istatistik*, (4. Baskı) H.Ü. Fen Fakültesi Basımevi, Beytepe, Ankara (1999).

JAMMALAMADAKA, S. R. ve SEN GUPTA, A., *Topics in circular statistics*, World Scientific Publishing Co. Pte. Ltd., London, England (2001).

JEFFREY, A., *Handbook of mathematical formulas and integrals*, Academic Press, Inc., London, England (1995).

JUPP, P. E. ve MARDIA, K. V., *A general correlation coefficient for directional data and related regression problems*, Biometrika, **67**, 1, 163-73 (1980).

KOTZ, S. ve JOHNSON, N. L., *Encyclopedia of statistical sciences*, Vol. 2, John Wiley & Sons, Inc., New York, USA (1982).

LE PENNEC, J-L., *Identifying ash flow sources with directional data: an application to the kizilkaya ignimbrite, central anatolia*, Journal of Geophysical Research, **105**, No. B12, 28427-28441 (2000).

LUND, U., *Least circular distance regression for directional data*, Journal of Applied Statistics, **26**, No. 6, 723-733 (1999).

LUND, U. J., *Tree-based regression for a circular response*, Communications In Statistics, **31**, No. 9, 1549-1560 (2002).

MARDIA, K. V., *Statistics of directional data*, Academic Press, London, England (1972).

MARDIA, K. V., *Statistics of directional data*, J. Roy. Statist. Soc., B.37, 349-393 (1975).

PEKER, K. Ö., *İki-aşamalı düzenlerde hipotez testi*, Hacettepe Üniversitesi Fen Bilimleri Enstitüsü Bilim Uzmanlığı Tezi, Ankara (1998).

PRESNELL, B., MORRISON, S. P. ve LITTELL, R. C., *Projected multivariate linear models for directional data*, Journal of the American Statistical Association, **93**, 443, 1068-1077 (1998).

SENGUPTA, A. ve PAL, C., *On optimal tests for isotropy against the symmetric wrapped stable-circular uniform mixture family*, Journal of Applied Statistics, **28**, No. 1, 129-143 (2001).

STEPHENS, M. A., *Tests for randomness of directions against two circular alternatives*, American Statistical Association Journal, 280-289 (1969).

WALD, A., *Sequential analysis*, John Wiley & Sons Inc., New York, USA, (1947).

WETHERILL, G. B., *Sequential methods in statistics*, Chapman & Hall, London, England (1975).

EKLER

EK-1 Bessel fonksiyonlarına ilişkin formüller

1. Birinci tür, p -inci sıra dönüştürülmüş Bessel fonksiyonu

$$\begin{aligned} I_p(\kappa) &= \sum_{r=0}^{\infty} \frac{1}{\Gamma(r+p+1)\Gamma(r+1)} \left(\frac{\kappa}{2}\right)^{2r+p} \\ &= \sum_{r=0}^{\infty} \frac{1}{(r+p)!r!} \left(\frac{\kappa}{2}\right)^{2r+p}, \quad p = 1, 2, \dots \end{aligned}$$

$$2. \quad I_p(\kappa) = \frac{1}{2\pi} \int_0^{2\pi} \cos p\theta e^{\kappa \cos \theta} d\theta$$

3. Birinci tür, sıfırncı sıra dönüştürülmüş Bessel fonksiyonu

$$I_0(\kappa) = \sum_{r=0}^{\infty} \frac{1}{(r!)^2} \left(\frac{\kappa}{2}\right)^{2r}$$

$$4. \quad I_0(\kappa) = \frac{1}{2\pi} \int_0^{2\pi} e^{\kappa \cos \theta} d\theta = \frac{1}{\pi} \int_0^{\pi} e^{\kappa \cos \theta} d\theta$$

$p \geq 2$ için, $I_p(\kappa)$ ve $A_p(\kappa)$ değerlerini hesaplamak amacıyla, bu fonksiyonların $p = 0$ ve $p = 1$ için değerlendirilmesi yeterlidir. Bundan sonra, aşağıdaki ilişkiler $p = 2, 3, \dots$ için kullanılabilir.

$$5. \quad I_p(\kappa) = I_{p-2}(\kappa) - \frac{2(p-1)}{\kappa} I_{p-1}(\kappa)$$

$$6. \quad A_p(\kappa) = A_{p-2}(\kappa) - \frac{2(p-1)}{\kappa} A_{p-1}(\kappa)$$

Özel olarak,

$$7. \quad I_2(\kappa) = I_0(\kappa) - \frac{2}{\kappa} I_1(\kappa)$$

$$8. \quad I_3(\kappa) = I_1(\kappa) - \frac{4}{\kappa} I_2(\kappa)$$

$$9. \quad A_2(\kappa) = 1 - \frac{2}{\kappa} A_1(\kappa)$$

$$10. A_3(\kappa) = A_1(\kappa) - \frac{4}{\kappa} A_2(\kappa)$$

$$11. I_0'(\kappa) = I_1(\kappa)$$

$$12. \frac{d}{d\kappa} \{\kappa I_1(\kappa)\} = \kappa I_0(\kappa)$$

$$13. I_1'(\kappa) = I_0(\kappa) - \frac{I_1(\kappa)}{\kappa} = \frac{1}{2} [I_0(\kappa) + I_2(\kappa)]$$

$$14. A_p(\kappa) = \frac{I_p(\kappa)}{I_0(\kappa)}, \quad p = 0, 1, 2, \dots$$

15. p-inci sıra standart Bessel fonksiyonu

$$J_p(\kappa) = \sum_{r=0}^{\infty} (-1)^r \{\Gamma(p+r+1)\Gamma(r+1)\}^{-1} \left(\frac{\kappa}{2}\right)^{2r+p}$$

EK-2 Von Mises dağılımında verilen \bar{R} değerleri için $\hat{\kappa}$ en çok olabilirlik tahminleri

\bar{R}	$\hat{\kappa}$	\bar{R}	$\hat{\kappa}$	\bar{R}	$\hat{\kappa}$
0.00	0.00000	0.35	0.74783	0.70	2.01363
0.01	0.02000	0.36	0.77241	0.71	2.07685
0.02	0.04001	0.37	0.79730	0.72	2.14359
0.03	0.06003	0.38	0.82253	0.73	2.21425
0.04	0.08006	0.39	0.84812	0.74	2.28930
0.05	0.10013	0.40	0.87408	0.75	2.36930
0.06	0.12022	0.41	0.90043	0.76	2.45490
0.07	0.14034	0.42	0.92720	0.77	2.54686
0.08	0.16051	0.43	0.95440	0.78	2.64613
0.09	0.18073	0.44	0.98207	0.79	2.75382
0.10	0.20101	0.45	1.01022	0.80	2.87129
0.11	0.22134	0.46	1.03889	0.81	3.00020
0.12	0.24175	0.47	1.06810	0.82	3.14262
0.13	0.26223	0.48	1.09788	0.83	3.30114
0.14	0.28279	0.49	1.12828	0.84	3.47901
0.15	0.30344	0.50	1.15932	0.85	3.68041
0.16	0.32419	0.51	1.19105	0.86	3.91072
0.17	0.34503	0.52	1.22350	0.87	4.17703
0.18	0.36599	0.53	1.25672	0.88	4.48876
0.19	0.38707	0.54	1.29077	0.89	4.85871
0.20	0.40828	0.55	1.32570	0.90	5.3047
0.21	0.42962	0.56	1.36156	0.91	5.8522
0.22	0.45110	0.57	1.39842	0.92	6.5394
0.23	0.47273	0.58	1.43635	0.93	7.4257
0.24	0.49453	0.59	1.47543	0.94	8.6104
0.25	0.51649	0.60	1.51574	0.95	10.2716
0.26	0.53863	0.61	1.55738	0.96	12.7661
0.27	0.56097	0.62	1.60044	0.97	16.9266
0.28	0.58350	0.63	1.64506	0.98	25.2522
0.29	0.60625	0.64	1.69134	0.99	50.2421
0.30	0.62922	0.65	1.73945	1.00	∞
0.31	0.65242	0.66	1.78953		
0.32	0.67587	0.67	1.84177		
0.33	0.69958	0.68	1.89637		
0.34	0.72356	0.69	1.95357		

EK-3 Von Mises dağılımında verilen κ değerleri için $\rho = A(\kappa)$ bileşke uzunlukları

κ	ρ	κ	ρ	κ	ρ
0.0	0.00000	3.5	0.84110	7.0	0.92553
0.1	0.04994	3.6	0.84616	7.1	0.92663
0.2	0.09950	3.7	0.85091	7.2	0.92770
0.3	0.14834	3.8	0.85537	7.3	0.92874
0.4	0.19610	3.9	0.85956	7.4	0.92975
0.5	0.24250	4.0	0.86352	7.5	0.93072
0.6	0.28726	4.1	0.86726	7.6	0.93168
0.7	0.33018	4.2	0.87079	7.7	0.93260
0.8	0.37108	4.3	0.87414	7.8	0.93350
0.9	0.40984	4.4	0.87732	7.9	0.93438
1.0	0.44639	4.5	0.88033	8.0	0.93524
1.1	0.48070	4.6	0.88320	8.1	0.93607
1.2	0.51278	4.7	0.88593	8.2	0.93688
1.3	0.54267	4.8	0.88853	8.3	0.93767
1.4	0.57042	4.9	0.89101	8.4	0.93844
1.5	0.59613	5.0	0.89338	8.5	0.93919
1.6	0.61990	5.1	0.89565	8.6	0.93993
1.7	0.64183	5.2	0.89782	8.7	0.94064
1.8	0.66204	5.3	0.89990	8.8	0.94134
1.9	0.68065	5.4	0.90190	8.9	0.94202
2.0	0.69777	5.5	0.90382	9.0	0.94269
2.1	0.71353	5.6	0.90566	9.2	0.94398
2.2	0.72803	5.7	0.90743	9.4	0.94521
2.3	0.74138	5.8	0.90913	9.6	0.94639
2.4	0.75367	5.9	0.91078	9.8	0.94752
2.5	0.76500	6.0	0.91236	10	0.94860
2.6	0.77545	6.1	0.91389	12	0.95730
2.7	0.78511	6.2	0.91536	15	0.96607
2.8	0.79404	6.3	0.91678	20	0.97467
2.9	0.80231	6.4	0.91816	24	0.97937
3.0	0.80999	6.5	0.91949	30	0.98319
3.1	0.81711	6.6	0.92078	40	0.98739
3.2	0.82375	6.7	0.92202	60	0.99163
3.3	0.82993	6.8	0.92323	120	0.99582
3.4	0.83570	6.9	0.92440	∞	1.00000

EK-4 S-Plus Paket Programında dairesel veriler için kullanılan fonksiyonlar

1. Von Mises Dağılımından Veri Türetimi (rvm komutu)

```

function(n, mean, k)
{
  vm <- c(1:n)
  a <- 1. + (1. + 4. * (k^2.))^0.5
  b <- (a - (2. * a)^0.5)/(2. * k)
  r <- (1. + b^2.)/(2. * b)
  obs <- 1.
  while(obs <= n) {
    U1 <- runif(1., 0., 1.)
    z <- cos(pi * U1)
    f <- (1. + r * z)/(r + z)
    c <- k * (r - f)
    U2 <- runif(1., 0., 1.)
    if(c * (2. - c) - U2 > 0.) {
      U3 <- runif(1., 0., 1.)
      vm[obs] <- sign(U3 - 0.5) * acos(f) + mean
      vm[obs] <- vm[obs] %% (2. * pi)
      obs <- obs + 1.
    }
    else {
      if(logb(c/U2) + 1. - c >= 0.) {
        U3 <- runif(1., 0., 1.)
        vm[obs] <- sign(U3 - 0.5) * acos(f) +
          mean
        vm[obs] <- vm[obs] %% (2. * pi)
        obs <- obs + 1.
      }
    }
  }
  vm
}

```

EK-4 (Devam)

2. Tanımlayıcı İstatistiklerin Elde Edilmesi (circ.summary, est.kappa, circ.disp komutları)

```
function(data)
{
  n <- length(rad(data))
  sinr <- sum(sin(rad(data)))
  cosr <- sum(cos(rad(data)))
  rho <- sqrt(sinr^2. + cosr^2.)/n
  mean.dir <- deg(atan(sinr, cosr))
  result <- data.frame(n, mean.dir, rho)
  result
}
```

```
function(data, bias = F)
{
  mean.dir <- circ.mean(data)
  kappa <- A1inv(mean(cos(rad(data) - rad(mean.dir))))
  if(bias == T) {
    kappa.ml <- kappa
    n <- length(rad(data))
    if(kappa.ml < 2.)
      kappa <- max(kappa.ml - 2. * (n * kappa.ml)^
                  -1., 0.)
    if(kappa.ml >= 2.)
      kappa <- ((n - 1.)^3. * kappa.ml)/(n^3. + n)
  }
  kappa
}
```

```
function(data)
{
  n <- length(rad(data))
  c <- sum(cos(rad(data)))
  s <- sum(sin(rad(data)))
  r <- sqrt(c^2. + s^2.)
  rbar <- r/n
  var <- 1. - rbar
  data.frame(n, r, rbar, var)
}
```

EK-4 (Devam)

3. U-Testi

```

function(theta, kappa, mu0, mu1, alpha, beta)
{
  # circular library must be loaded
  nu1 <- (mu0 + mu1)/2
  nu2 <- (mu0 - mu1)/2
  A <- (1 - beta)/alpha
  B <- beta/(1 - alpha)
  alt <- log(B)
  ust <- log(A)
  cat("\n lower limit--->", alt, "\n", "upper limit--->", ust,
      "\n")
  cat("Observation No", "   Theta",
      "   Test statistics \n")
  n <- nrow(theta)
  i <- 1
  kosul <- 0
  while(i <= n) {
    kosul <- kosul + ((-2) * kappa * sum(sin(rad(theta[
      i, ]) - rad(nu1)) * sin(rad(nu2))))
    cat("   ", i, "   ", theta[i, ], "   ",
        kosul, "\n")
    if(kosul <= alt) {
      cat("\n Null hypothesis is accepted. \n")
      cat("\n After", i,
          "observations iteration is stopped. \n"
        )
      return()
    }
    if(kosul >= ust) {
      cat("\n Alternative hypothesis is accepted. \n")
      cat("\n After", i,
          "observations iteration is stopped. \n"
        )
      return()
    }
    i <- i + 1
  }
  cat("\n   ADD AN OBSERVATION")
  return()
}

```


EK-4 (Devam)

4. V-Testi

```

function(theta, delta0, alpha, beta)
{
  # circular library must be loaded
  delta1 <- -1 * delta0
  A <- (1 - beta)/alpha
  B <- beta/(1 - alpha)
  alt <- log(B)
  ust <- log(A)
  cat("\n lower limit--->", alt, "\n", "upper limit--->", ust,
      "\n\n")
  cat("Observation No", "  Theta", "      Test statistics \n"
      )
  n <- nrow(theta)
  i <- 1
  while(i <= n) {
    thetanew <- matrix(c(), ncol = 1, nrow = i)
    for(j in 1:i) {
      thetanew[j, ] <- theta[j, ]
    }
    kosul <- 2 * delta1 * sum(sin(rad(thetanew)))
    cat("      ", i, "      ", thetanew[i, ],
        "      ", kosul, "\n")
    if(kosul <= alt) {
      cat("\n Null hypothesis is accepted \n")
      cat("\n After", i,
          "observations iteration is stopped. \n"
          )
      return()
    }
    if(kosul >= ust) {
      cat("\n Alternative hypothesis is accepted \n")
      cat("\n After", i,
          "observations iteration is stopped. \n"
          )
      return()
    }
    i <- i + 1
  }
  cat("\n      ADD AN OBSERVATION  \n")
  return()
}

```