ANADOLU ÜNİVERSİTESİ BİLİM VE TEKNOLOJİ DERGİSİ
ANADOLU UNIVERSITY JOURNAL OF SCIENCE AND TECHNOLOGY
Cilt/Vol.:5-Sayı/No: 2 : 343-349 (2004)

DÜZELTME/**ERRATUM**

Dergimizde Cilt 5, Sayı 1'de, Sayfa 115'de yer alan yazının tümü üzerinde, şekil ve formüllerin dizilişinde hatalar bulunduğu için, makalenin tümünün yeniden basılması zorunluluğu doğmuştur.

In Volume 5, Number 1, p.115 of our journal, this paper was misprinted due to printing formulas and figures. From this point of view, this paper is reprinted.

ARAŞTIRMA MAKALESİ/**RESEARCH ARTICLE**

DAİRESEL VERİLERE UYGULANAN TANIMLAYICI İSTATİSTİKSEL YÖNTEMLER VE METEOROLOJİK BİR UYGULAMA

K. Özgür PEKER [1], Sevil BACANLI[2]

ÖZ

Bu çalışmanın amacı, dairesel verilerin özelliklerini ve doğrusal verilere uygulanan temel istatistiksel yöntemlerle arasındaki farkları incelemektir. Çalışmada, dairesel veriler için temel parametre değerlerinin hesaplanması açıklanmış ve dairesel veri analizindeki temel olasılık dağılımı olan von Mises dağılımı incelenmiştir. Ardından, dairesel veriler için ortalama yön testleri tanımlanmış ve von Mises dağılımı için ortalama yön testine ilişkin kuramsal bilgiler verilmiştir. Ayrıca, Anadolu Üniversitesi Sivil Havacılık Yüksekokulu Meydan Meteoroloji İstasyon Müdürlüğü'nden, Anadolu Üniversitesi havaalanı için ölçülmüş olan rüzgar yönlerinden oluşan veri değerlerine incelenen yöntemler uygulanmış, elde edilen sonuçlar değerlendirilmiştir.

Anahtar Kelimeler: Dairesel veri, Ortalama yön, von Mises dağılımı, Rüzgar yönü.

DESCRIPTIVE STATISTICAL METHODS APPLIED TO CIRCULAR DATA AND A METEOROLOGICAL APPLICATION

ABSTRACT

The aim of this study is to investigate the properties of circular data and the differences from principal statistical techniques of linear data. In the study, calculations of descriptive statistics for circular data are explained and von Mises distribution which is the main probability distribution in circular data analysis is examined. Then, mean direction tests for circular data are described and the theoretical informations about testing of mean direction for von Mises distribution is given. An application is also shown, using the wind directions on Anadolu University airport, obtained from Anadolu University School of Civil Aviation Airfield Meteorology Station Administration and the obtained results are evaluated.

Key Words: Circular data, Mean direction, von Mises distribution, Wind direction.

[1] Anadolu Üniversitesi Fen Fakültesi İstatistik Bölümü, Yunusemre Kampusu, Eskeşehir; **Faks:** 0222-3204910; **E-Posta:** opeker@anadolu.edu.tr

[2] Hacettepe Üniversitesi Fen Fakültesi İstatistik Bölümü Beytepe, Ankara; **E-Posta:** sevil@hacettepe.edu.tr

Geliş: 27 Ekim 2003; **Düzeltme:** 06 Şubat 2004; **Kabul**: 22 Mart 2004

1. GİRİŞ

Bir çok bilim dalı için, yapılan herhangi bir araştırmada veri toplanması aşamasında ölçümler açısal olarak elde edilmektedir. Bu tür yönsel verilerin özellikle jeoloji, meteoroloji, biyoloji, fizik, psikoloji, tıp, astronomi gibi alanlarda kullanıldığı sıkça görülür. Açısal gözlemler, deneylerde farklı biçimlerde ortaya çıkarlar. Örneğin biyolog, kaplumbağaların hareket yönünü incelerken, jeolog da fay hatlarına ilişkin bir araştırma yapabilir. İlk örnekteki yönsel araştırma iki boyutlu olarak incelenirken, ikinci örnekteki araştırma dünya yüzeyi yaklaşık olarak bir küre şeklinde olduğu için üç boyutta incelenir. Dolayısıyla, yönlere ilişkin herhangi bir gözlem kümesi, *yönsel veri* olarak adlandırılır. Bu gözlemlerin elde edilmesinde kullanılan iki temel dairesel ölçüm aracı *pusula* ve *saat*tir. Pusula ile yapılabilen gözlemler arasında, rüzgar yönleri ve kuşların uçuş yönleri sayılabilir. Benzer türdeki verilere ilişkin ölçümler bir açıölçer yardımıyla da yapılabilir. Saatle yapılabilecek gözlemlere örnek olarak da, bir hastanedeki acil servis birimine gelen hastaların 24 saat içerisindeki servise geliş zamanlarının dağılımı verilebilir. Bu türdeki veriler ay ya da yıl cinsinden de elde edilebilir.

İki-boyutlu yönler, uygun biçimde seçilmesi gereken bir *sıfır yönü* ve *dönüş doğrultusu*'na göre ölçülen açılar biçiminde gösterilebilirler. Burada, sıfır yönü başlangıç noktasını ve dönüş doğrultusu da pozitif yön olarak saat yönünün mü yoksa saat yönünün tersinin mi kullanılacağını belirtmektedir. Yön kavramında herhangi bir büyüklük söz konusu olmadığından, açısal bir gözlem değeri, 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^o (0°<θ^o?360°), birim vektör olacaktır. Bu dairesel gösterimden dolayı, iki-boyutlu yönlere ilişkin gözlemler *dairesel veri* olarak da adlandırılır. Burada θ^o; vektör ile pozitif x- ekseninin, saat yönünün tersi yönünde yaptığı açıyı gösterir.

İki-boyutlu bir yönün, açı ya da birim vektör şeklindeki sayısal gösteriminin tek bir tane olması beklenmemelidir. Çünkü dairesel gözlemin değeri, sıfır yönüne ve dönüş doğrultusunun saat yönünde olup olmaması seçimine bağlıdır. Elde edilen sonuçlar verilen gözlem değerlerinin bir fonksiyonudur ve bunlara verilen keyfi değerlere bağlı değildir. Bu özelliklerinden dolayı dairesel veri analizi, bilinen istatistiksel analizden oldukça farklıdır. Keyfi sıfır yönü ve dönüş doğrultusu seçimine ilişkin ölçülere duyulan ihtiyaç, bilinen birçok istatistiksel tekniği ve ölçüleri tamamen anlamsız olmasa da çoğu zaman hatalı kılar.

Konuyla ilgili ilk önemli çalışmalar, Mardia (1972) tarafından yapılmıştır. Mardia bu çalışmalarında istatistik teorisinin dairesel veri analizine nasıl uygulanabileceği üzerinde yoğunlaşmıştır. Bu alanda yapılan diğer çalışmalar arasında Fisher (1993), Mardia ve Jupp (2000) ve Jammalamadaka ve Sen Gupta (2001) bulunmaktadır.

2. DAİRESEL VERİLER İÇİN TANIMLAYICI İSTATİSTİKLER

2.1. Ortalama Yön, Bileşke Uzunluğu ve Ortalama Bileşke Uzunluğu

Dairesel verilere ilişkin yapılan ölçümler genellikle derece cinsinden yapılır. Fakat, bazı durumlarda, özellikle teoride, derece cinsinden ölçülen bir θ^o açısal gözleminin (2.1) eşitliği yardımıyla radyan cinsinden θ açısına dönüştürülmesi tercih edilir.

$$\theta = \frac{\pi\theta^o}{180}, \qquad 0 < \theta < 2\pi \tag{2.1}$$

Tek bir yöne doğru kümelenme gösteren bir dairesel veri seti için ortalama yön ölçüsünün belirlenebilmesi için, gözlem değerleri birim vektörler olarak düşünülür ve bu vektörlerin bileşke vektörünün yönü kullanılır.

P_i; birim çember üzerinde θ_i (i = 1,...,n) açısına bağlı olarak belirlenen herhangi bir nokta olmak üzere, $\theta_1,...,\theta_n$ açılarının ortalama yönü $\bar{\theta}$; $\overline{OP_1},...,\overline{OP_n}$ birim vektörlerinin bileşkesinin yönü olarak tanımlanır. P_i noktalarının ağırlık merkezi (C,S); $C = \sum_{i=1}^{n} \cos\theta_i$, $S = \sum_{i=1}^{n} \sin\theta_i$ olmak üzere, *bileşke uzunluğu* R;

$$R = \sqrt{C^2 + S^2}, \qquad 0 \le R \le n \tag{2.2}$$

eşitliğinden, veya ortalama yön biliniyorsa

$$R = \frac{C}{\cos\bar{\theta}} = \frac{S}{\sin\bar{\theta}}, \qquad 0 \le R \le n \tag{2.3}$$

eşitliği yardımıyla bulunur. Burada, θ_1, ..., θ_n açılarının vektör bileşkesinin yönü $\bar{\theta}$;

$$\bar{\theta} = \begin{cases} \tan^{-1}(S/C) & , \ S \ge 0 \ , \ C > 0 \\ \tan^{-1}(S/C) + \pi & , \ C < 0 \\ \tan^{-1}(S/C) + 2\pi & , \ S < 0 \ , \ C \ge 0 \\ \pi/2 & , \ S > 0 \ , \ C = 0 \\ tanımsız & , \ S = 0 \ , \ C = 0 \end{cases} \tag{2.4}$$

biçiminde tanımlanır ve *ortalama yön* olarak adlandırılır. *Ortalama bileşke uzunluğu* $\bar{R}$ ise;

$$\bar{R} = R/n, \qquad 0 \le \bar{R} \le 1 \tag{2.5}$$

olarak verilir.

Dolayısıyla, verilen bir açısal gözlem kümesi için, öncelikle bu değerler dik koordinatlara dönüştürülür ve bileşke vektörün bulunabilmesi için toplanır. Böylece bu bileşke vektörün yönü olan $\bar{\theta}$, (2.4) eşitliği kullanılarak elde edilir. Burada, R = 0 durumunda (C = 0 ve S = 0), dairesel ortalama tanımsızdır. Bileşke vektörün sıfır uzunluğa sahip olması, verinin çember üzerinde herhangi bir yöne doğru yoğunlaşma göstermediğini, yani *düzgün* dağıldığını gösterir. Bu durumda, verinin herhangi bir tercih edilen yönü ya da ortalama yönü bulunmamaktadır.

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

$$\bar{C}_n = \frac{1}{n}\sum_{i=1}^{k} f_i \cos\theta_i \quad , \quad \bar{S}_n = \frac{1}{n}\sum_{i=1}^{k} f_i \sin\theta_i \quad ,$$

$$\bar{R} = \sqrt{\bar{C}^2 + \bar{S}^2} \qquad \textbf{(2.6)}$$

değerleri hesaplanır.

2.2. Yoğunlaşma Parametresi

Yoğunlaşma parametresi κ'nın en çok olabilirlik tahmini $\hat{\kappa}$;

$$A_1(\hat{\kappa}) = R/n = \bar{R} \qquad \textbf{(2.7)}$$

Eşitlik (2.7)'nın çözümüdür. Burada $\bar{R}$;

$$A_1(x) = I_1(x) / I_0(x) \qquad \textbf{(2.8)}$$

dönüştürülmüş iki Bessel fonksiyonunun oranıdır. Eşitlik (2.7)'nin çözümü için uygun bir yaklaşım ise;

$$\hat{\kappa} = \begin{cases} 2\bar{R} + \bar{R}^3 + 5\bar{R}^5/6 & \bar{R} < 0.53 \\ -0.4 + 1.39\bar{R} + 0.43/(1-\bar{R}) & 0.53 \le \bar{R} < 0.85 \\ 1/(\bar{R}^3 - 4\bar{R}^2 + 3\bar{R}) & \bar{R} \ge 0.85 \end{cases} \qquad \textbf{(2.9)}$$

ile verilir (Fisher 1993).

2.3. Dairesel Varyans

Bileşke uzunluğu R, verinin ortalama yön etrafındaki yoğunlaşma miktarını gösterir. Bütün gözlem noktaları (birim vektörler) aynı yöne doğru büyük bir yoğunlaşma gösteriyor ise, R değerinin büyüklüğü n'e yakın bulunacaktır. Bunun aksine, veriler herhangi bir yoğunlaşma göstermeksizin çember üzerinde düzgün bir dağılım gösteriyor ise, R değeri sıfıra çok yakın bir değer alacaktır.

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

ile tanımlandığında (Mardia 1972);

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

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

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

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

olur ve sonuç olarak;

$$V = 1 - \bar{R} \quad , \quad 0 \le V \le 1 \qquad \textbf{(2.10)}$$

değerine ulaşılır. (2.10) eşitliği, *örneklem dairesel varyansı* olarak tanımlanır. Doğrusal veri varyansında olduğu gibi, dairesel varyans değeri küçüldükçe, dağılım homojenleşir.

n sayıda dairesel gözlem $\bar{\theta}$ ortalama yönü etrafında kümelenmiş ise $\bar{R}$, 1'e yakın bir değer alır ve dairesel varyans sıfıra yakın çıkar. Yönler geniş bir dağılım göstermiş ise, $\bar{R}$ değeri küçük ve dairesel varyans 1'e yakın olacaktır.

2.4. Dairesel Standart Sapma

V değerinin $(0,\infty)$ aralığına uygun bir dönüşümü;

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

(2.11)

biçiminde tanımlanır.

ν ölçüsü, doğru üzerindeki bilinen standart sapmaya benzemektedir. Kuramsal araştırmalarda, V ve R değerleri ν değerinden daha çok kullanılmaktadır.

V'nin küçük değerlerinde, dairesel standart sapma;

$$\nu = (2V)^{1/2} \qquad \textbf{(2.12)}$$

biçimindedir (Mardia 1972).

2.5. Dairesel Saçılım

Dairesel saçılım;

$$\delta = \frac{1-\rho_2}{2\bar{R}^2} \qquad \textbf{(2.13)}$$

biçiminde tanımlanır. Eşitlikte;

$$\rho_2 = \frac{1}{n}\sum_{i=1}^{n}\cos 2(\theta_i - \bar{\theta}) \qquad \textbf{(2.14)}$$

değeri merkezi ikinci trigonometrik momenti göstermektedir (Fisher 1993). Dairesel saçılım daha çok, ortalama yön için güven aralığının hesaplanmasında, ortalama yönlerin karşılaştırılmasında ve birleştirilmesinde kullanılmaktadır.

2.6. Dairesel Standart Hata

Ortalama yön tahmininin *dairesel standart hata*sı;

$$\sigma = \sqrt{\frac{\delta}{n}} \qquad \textbf{(2.15)}$$

formülü ile hesaplanır. Burada, dairesel saçılım değeri Eşitlik (2.13)'den bulunur (Fisher 1993). Dairesel standart hata özellikle, ortalama yön için güven aralıklarının belirlenmesinde kullanılmaktadır.

2.7. Medyan Yönü

Birim çember üzerindeki noktalardan oluşan bir veri kümesinin verildiği varsayılsın. Bu noktalardan herhangi bir P noktası aşağıdaki özellikleri sağlıyorsa, P noktası örneklemin 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.

Bu koşulları sağlayan $\overline{OP}$ vektörüne örneklemin *medyan yönü* adı verilir.
Medyan yönü, $\tilde{\theta}$ simgesiyle gösterilir ve

$$\int_{\tilde{\theta}}^{\tilde{\theta}+\pi} f(\theta)d\theta = \int_{\tilde{\theta}+\pi}^{\tilde{\theta}+2\pi} f(\theta)d\theta = \frac{1}{2} \quad \textbf{(2.16)}$$

biçiminde tanımlanan integralin çözümü ile bulunur. 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;

$$\tilde{\theta} = l + \frac{\frac{n}{2} - f_0}{f_{+1} - f_0} \times h \quad (2.17)$$

ile verilir. Burada; l: medyan sınıfının alt sınırını, f_0: medyan sınıfının $(\theta_i\text{–}180°, \theta_i)$ aralığındaki sıklığını, f_{+1}: medyan sınıfından bir sonraki sınıfın $(\theta_i\text{–}180°, \theta_i)$ aralığındaki sıklığını ve h: sınıf aralığının uzunluğunu gösterir.

2.8. Mod Yönü

Mod yönü $\breve{\theta}$*;* verinin en fazla yoğunlaştığı yön anlamına gelir. Verilen bir sıklık dağılımının mod yönü, kesim noktası seçiminden sonra bilinen mod hesaplama yöntemiyle bulunur. Gruplanmış bir seride mod yönü;

$$\breve{\theta} = l + \frac{f_0 - f_{-1}}{2f_0 - f_{-1} - f_{+1}} \times h \quad \textbf{(2.18)}$$

eşitliği ile elde edilir. 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 gösterir (Mardia 1972).

3. VON MISES DAĞILIMI

Bir θ dairesel raslantı değişkeni, von Mises dağılımına sahipse, olasılık yoğunluk fonksiyonu;

$$\text{VM}(\mu,\kappa)\text{:}\ f(\theta;\mu,\kappa) = \frac{1}{2\pi\,\text{I}_0(\kappa)} e^{\kappa\cos(\theta-\mu)} ,$$

$$0 \le \theta < 2\pi \ , \ \kappa \ge 0 \ , \ 0 \le \mu < 2\pi \quad \textbf{(3.1)}$$

biçiminde tanımlanır. Burada, $\text{I}_0(\kappa)$; birinci tür ve sıfır sırasında dönüştürülmüş Bessel fonksiyonudur ve

$$\text{I}_0(\kappa) = \frac{1}{2\pi}\int_0^{2\pi} e^{\kappa\cos(\phi-\mu)} \text{d}\phi \quad \textbf{(3.2)}$$

olarak verilir.

Von Mises dağılımında κ değerinin büyük olması, anakütle ortalama yönü ve mod yönü etrafında daha büyük bir kümelenme olduğunu gösterir. Buna göre κ, ortalama yöne ilişkin yoğunlaşmayı ölçen bir parametredir (Jammalamadaka ve Sen Gupta 2001).

Doğrusal veri setleri için, normal dağılım, sıkça kullanılan bir dağılım olmakla birlikte, şekilsel istatistiksel analiz, σ^2 değerine bakılmaksızın yapılabilmektedir. Fakat, dairesel veri setleri için temel dağılım olan von Mises dağılımında ise, şekilsel istatistiksel analiz κ değerine bakılmadan yapılamamaktadır.

Yoğunlaşma parametresinin, dairesel varyans ile olan ilişkisi;
$V = 1 - \rho = 1 - \text{A}(\kappa)$ **(3.3)**

biçiminde tanımlanır (Mardia 1972).

Von Mises dağılımının momentleri;
Ortalama yön : μ
Ortalama bileşke uzunluğu: $\rho = A_1(\kappa)$
Dairesel saçılım : $\delta = \dfrac{1}{\kappa A_1(\kappa)}$

$\alpha_p = A_p(\kappa)$
$\beta_p = 0$, $p \ge 1$
biçimindedir.

4. DAİRESEL VERİLERDE TEK ÖRNEKLEM ORTALAMA YÖN TESTLERİ

4.1. Tek Örneklem Ortalama Yön Testi

Belirlenen bir α anlamlılık seviyesinde H_0: $\mu = \mu_0$ hipotezinin, H_1: $\mu \neq \mu_0$ alternatif hipotezine karşı testi için iki durum söz konusudur. Birinci yöntem, ortalama yön için güven aralığının belirlenmesi yoluyla test uygulamak, ikincisi ise, örneklem hacmine dayalı olarak doğrudan test uygulamaktır.

Güven aralığının belirlenmesi yönteminde, ortalama yönün dairesel standart hatasından yararlanılır. Buna göre sırasıyla;

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

değerleri hesaplanır. Bu işlemlerden sonra, ortalama yön için $(1-\alpha)$'lık güven aralığı;
$(\mu_0 - \sin^{-1}(z_{\alpha/2}\ \sigma)\ ,\ \mu_0 + \sin^{-1}(z_{\alpha/2}\ \sigma))$
(4.1)
ile bulunur. Burada, $z_{\alpha/2}$; standart normal dağılım tablosundan elde edilir. Kurulacak hipotezler;

H_0: $\mu = \mu_0$
H_1: $\mu \neq \mu_0$ **(4.2)**
olduğunda, ortalama yön değeri (4.1)'de verilen aralığın içinde ise, α anlamlılık seviyesinde H_0 hipotezi kabul edilir, aksi halde reddedilir (Fisher 1993 ve İnal ve Günay 1999).

Örneklem hacmine dayalı test ise, uygun bir güven aralığının belirlenemediği durumda doğrudan uygulanabilmektedir. Fakat bu test, örneklem hacmi 25 veya daha fazla olduğunda kullanılabilir. Bu yönteme göre, öncelikle dairesel standart hata σ yukarıda verilen testteki gibi hesaplanır. Test istatistiği;

$$S = \frac{\sin(\bar{\theta} - \mu_0)}{\sigma} \qquad \textbf{(4.3)}$$

olarak verilir. Bu değer α güvenilirlik seviyesinde, standart normal dağılım tablosundan elde edilen değerle karşılaştırılır. $|S|$ değeri tablo değerinden büyük olduğunda H_0 hipotezi reddedilir, aksi halde kabul edilir (Fisher 1993).

4.2. Von Mises Dağılımına Sahip Anakütlelerde Ortalama Yön Testi

Bu kesimde, incelenen anakütlenin μ ortalama yönü ve κ yoğunluk parametresi ile von Mises dağılımına VM(μ,κ) sahip olduğu varsayımı altında, ortalama yönün μ_0 gibi bir değere eşit olup olmadığı testi incelenecektir. İlk olarak, ortalama yön için güven aralığının belirlenmesi yöntemiyle uygulanan test incelenecektir. Doğrudan test uygulamasında ise, yoğunlaşma parametresinin bilindiği ve bilinmediği durum olmak üzere iki durum söz konusudur.

Güven aralığı belirlenerek uygulanan testte, ortalama yön için dairesel standart hata değeri hesaplanır. Fakat, burada von Mises dağılımı için özel bir hal alan standart hata formülü kullanılmaktadır.

$$\sigma_{\text{VM}} = \frac{1}{\sqrt{n\bar{R}\hat{\kappa}}} \qquad (4.4)$$

Buna göre; ortalama yön için (1–α)'lık güven aralığı;

$(\mu_0 - \sin^{-1}(z_{\alpha/2}\ \sigma_{\text{VM}})\ ,\ \mu_0 + \sin^{-1}(z_{\alpha/2}\ \sigma_{\text{VM}}))$ (4.5)

olur. Ortalama yön değeri (4.5)'te verilen aralığın içinde yer alıyorsa, α anlamlılık seviyesinde H_0 hipotezi kabul edilir, aksi halde reddedilir.

Güven aralığı belirlenmeden doğrudan yapılan test, örneklem hacmi n ve yoğunlaşma parametresi $\hat{\kappa}$ tahmin değerinin büyüklüğüne bağlı olarak uygulanır.

Yoğunlaşma parametresinin bilinmediği durumda, von Mises dağılımı için ortalama yönün dairesel standart hatası σ_{VM} Eşitlik (4.4)'teki gibi hesaplanır. Test istatistiği ise;

$$E_n = \frac{\sin(\bar{\theta} - \mu_0)}{\sigma_{\text{VM}}} \qquad (4.6)$$

olarak verilir. Bu değer α güvenilirlik seviyesinde, standart normal dağılım tablosundan elde edilen değerle karşılaştırılır. Karşılaştırmalar alternatif hipotezin değişik durumları için aşağıdaki gibi yapılır:

- H_0: $\mu = \mu_0$, H_1: $\mu \neq \mu_0$ testi için $|E_n| > z_{\alpha/2}$ ise H_0 red,
- 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 red,
- 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 red.

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

Çizelge 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

Yoğunlaşma parametresi κ'nın κ_0 gibi bilinen bir değere eşit olduğu varsayıldığında;
$\kappa_0 \geq 2$ ise, Von Mises dağılımı için ortalama yönün dairesel standart hatası σ_{VM};

$$\sigma_{\text{VM}} = \frac{1}{\sqrt{n\rho\kappa_0}} \qquad (4.7)$$

biçimindedir. Test istatistiği ise;

$$E_n = \frac{\sin(\bar{\theta} - \mu_0)}{\sigma_{\text{VM}}} \qquad (4.8)$$

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

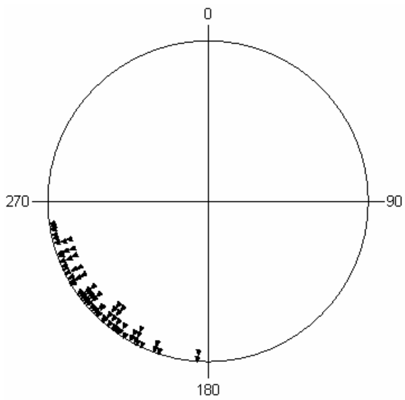
- H_0: $\mu = \mu_0$, H_1: $\mu \neq \mu_0$ testi için $|E_n| > z_{\alpha/2}$ ise H_0 red,
- H_0: $\mu = \mu_0$, H_1: $\mu < \mu_0$ testi için $\mu_0 - \pi < \bar{\theta} < \mu_0$ ve $E_n < -z_\alpha$ ise H_0 red,
- H_0: $\mu = \mu_0$, H_1: $\mu > \mu_0$ testi için $\bar{\theta} < \mu_0 + \pi$ ve $E_n > -z_\alpha$ ise H_0 red.
(Fisher 1993).

5. UYGULAMA

Dairesel verilerin en çok kullanıldığı alanlardan birisi meteorolojik gözlem çalışmalarıdır. Yukarıda incelenen yöntemlerin gerçek veriler üzerinde uygulanması amacıyla, Anadolu Üniversitesi Sivil Havacılık Yüksekokulu Meydan Meteoroloji İstasyon Müdürlüğü'nden, Anadolu Üniversitesi havaalanı için ölçülen ve *rüzgar yönleri* ve *rüzgar hızları*ndan oluşan veri seti elde edilmiştir. Bu rüzgar yönü verileri açısal gözlemlerden oluşmaktadır. Gözlemler 15'er saniyelik zaman aralıklarında ölçülmektedir. Burada, öncelikle elde edilen veriler şematik olarak verilecek, tanımlayıcı istatistikleri elde edilecek ve ardından tek örneklem ortalama yön testi uygulanacaktır.

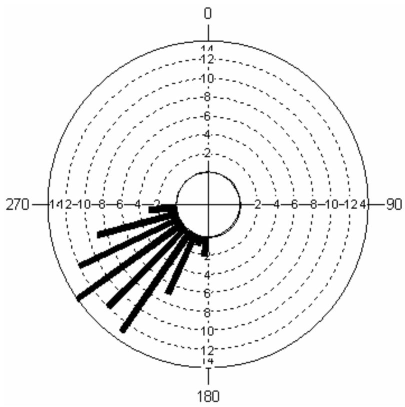
Uygulamada ele alınan veriler 29 Eylül 2002 tarihinde saat 12:00'dan 18:00'a kadar yapılan 5'er dakikalık rüzgar yönü ve hızı ölçümlerinden oluşmaktadır. Bu tarihte, verilen saat aralığında 73'er rüzgar yönü ve

hızı verisi elde edilmiştir. Bu dairesel gözlem değerleri için elde edilen dairesel grafikler aşağıda verilmektedir.
Şekil-1'de, verilerin ham veri grafiği görülmektedir.



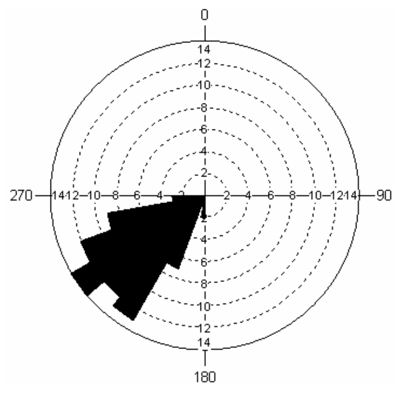
Şekil-1 Rüzgar yönü verilerinin ham veri grafiği

Şekil-2'se, rüzgar yönü verilerinin 10°'lik grup genişliği seçildiği durumdaki dairesel histogramı görülmektedir.



Şekil-2 Rüzgar yönü verilerinin dairesel histogramı

Şekil-3'te, rüzgar yönü verilerinin 10°'lik grup genişliği seçildiği durumdaki gül şeması görülmektedir.



Şekil-3 Rüzgar yönü verilerinin gül şeması

Şekillerden de görüldüğü gibi, veriler 184°–262° aralığında değer almakta ve dağılımın şekli von Mises dağılımına uymaktadır.

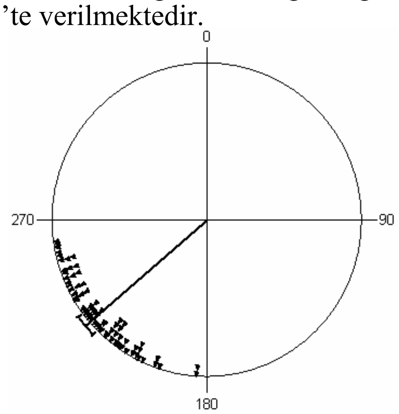
Belirlenen gün ve saat için seçilen 73 gözlem değeri için *S-Plus* ve *Oriana* paket programları kullanılarak elde edilen tanımlayıcı istatistik değerleri; ortalama rüzgar yönü: $\bar{\theta} = 229.364°$, ortalama rüzgar hızı: 13.62 km/saat, bileşke uzunluğu: $R = 69.496$, ortalama bileşke uzunluğu: $\bar{R} = 0.952$, yoğunlaşma parametresi tahmini: $\hat{\kappa} = 10.65$, dairesel varyans: $V = 0.048$, dairesel standart sapma: $\nu = 18.003°$ ve dairesel standart hata: $\sigma = 2.106°$ olarak bulunur.

Seçilen örneklem 229.364°'lük ortalama yönü ve 10.65'lik yoğunluk parametresi ile von Mises dağılımına VM(229.36, 10.65) sahiptir.

Güven aralığının belirlenmesi yöntemiyle ortalama yön test edilmek istendiğinde kurulacak hipotezler;
$H_0: \mu = 229°$
$H_1: \mu \neq 229°$
biçiminde olsun. Eşitlik (4.4)'e göre;

$$\sigma_{\mathrm{VM}} = \frac{1}{\sqrt{73(0.952)(10.65)}} = 0.0367574$$

elde edilir. Buna göre; ortalama yön μ için $(1-\alpha)$'lık güven aralığı (4.5)'ten yararlanılarak;
$(229.364 - \sin^{-1}((1.9604)(0.0367574))\ ,\ 229.364 + \sin^{-1}((1.9604)(0.0367574)) = (225.23°, 233.50°)$
olacaktır. % 95'lik bu güven aralığının grafik gösterimi Şekil-4'te verilmektedir.



Şekil-4 Rüzgar yönü verilerinde ortalama rüzgar yönü için % 95'lik güven aralığı

Burada $\alpha = 0.05$ olarak alınmıştır. H_0 hipotezinde verilen 229° değeri verilen aralığın içinde yer aldığı için, %5 anlamlılık seviyesinde H_0 hipotezi kabul edilir. Ortalama yön 229°'dir.

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 (4.6) eşitliğinden;

$$E_n = \frac{\sin(229.364 - 229)}{0.0367574} = 0.1728$$

olarak bulunur. $|0.1728| < 1.9604$ olduğundan 0.05 anlamlılık seviyesinde H_0 hipotezi kabul edilir. Ortalama yön 229°'ye eşittir.

κ'nın bilindiği durumda ise, $\kappa_0 = 10$ değeri göz önüne alınsın. $\kappa_0 \geq 2$ olduğundan, von Mises dağılımı için ortalama yönün dairesel standart hatası σ_{VM} (4.7) eşitliğinden;

$$\sigma_{VM} = \frac{1}{\sqrt{73(0.952)(10)}} = 0.0379333$$

olarak elde edilir. Test istatistiği Eşitlik (4.8)'den;

$$E_n = \frac{\sin(229.364 - 229)}{0.0379333} = 0.1675$$

bulunur. $|0.1675| < 1.9604$ olduğundan %5 anlamlılık seviyesinde H_0 hipotezi kabul edilir. Buna göre 29 Eylül tarihinde de saat 12:00 ile 18:00 saatleri arasındaki ortalama rüzgar yönü 229° ve ortalama rüzgar hızı 13.62 km/saat olacaktır. Dolayısıyla 30 Eylül tarihinde Anadolu Üniversitesi havaalanından kalkış ya da iniş yapacak olan uçaklar bu rüzgar yönü ve hızı bilgisinden yararlanabilirler.

6. TARTIŞMA VE SONUÇ

Araştırmalarda, elde edilen gözlem değerleri doğrusal ya da dairesel olabilir. Fakat, dairesel verilerin kullanıldığı araştırmalarda bilinen istatistiksel yöntemlerin kullanılması araştırmacıyı yanlış sonuçlara götürmektedir. Bu çalışmada, dairesel veriler için istatistiksel gösterim yöntemleri, tanımlayıcı istatistiklerin hesaplanması ve ortalama yön için hipotez testleri incelenmiştir.

Son yirmi yılda veri gösterimi, korelasyon, regresyon ve zamana ya da konuma bağlı yapıdaki verilerin analizi üzerinde durulmakla birlikte, yönsel veri çalışmaları, araştırmacılara çok geniş bir alanda ilerleme olanağı vermekte ve yeni istatistiksel yöntemler geliştirmede çok verimli bir alan olduğu görülmektedir. Ayrıca doğal, fiziksel, tıbbi ve de sosyal bilimlerde ortaya çıkan problemler için yeni ve farklı uygulamalar geliştirilebilmektedir (Peker 2002).

KAYNAKÇA

Ajne, B. (1968). A simple Test For Uni-formity of a Circular Distribution, *Biometrika*, **55**, 343-354.

Cheeney, R. F. (1983). *Statistical Methods in Geology*, George Allen and Unwin (Publishers) Ltd., London, UK.

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

Er, F. (2001). Dairesel (Circular) Veri Analizinde Kullanılan İstatistiksel Tekniklere Bir Bakış, 2. *İstatistik Kongresi*, Antalya, Türkiye.

Fisher, N. I. (1993). *Statistical Analysis of Circular Data*, Cambridge University Press, Cambridge, Great Britain.

Gumbel, E. J., Greenwood, J. A. ve Durand, D. (1953). The Circular Normal Distribution: Theory and Tables, *J. Amer. Statist. Ass.* 48, 131-152.

İnal, H. C. ve Günay, S. (1999). *Olasılık ve Matematiksel İstatistik,* (4. Baskı) H.Ü. Fen Fakültesi Basımevi, Beytepe, Ankara.

Jammalamadaka, S. R. ve Sen Gupta, A. (2001). *Topics in Circular Statistics*, World Scientific Publishing Co. Pte. Ltd., London, England.

Mardia, K. V. (1972). *Statistics of Directional Data*, Academic Press, London, England.

Mardia, K. V. ve Jupp, P. E. (2000). *Directional Statistics*, John Wiley & Sons Ltd., Chichester, England.

Peker, K. Ö. (2002). Dairesel Veriler ve Ardışık Testlerde Kullanımı, *Anadolu Üniversitesi Fen Bilimleri Enstitüsü Doktora Tezi*, Eskişehir.

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

Kadir Özgür Peker, lisans öğrenimini Hacettepe Üniversitesi İstatistik Bölümü'nde 1994'de, yüksek lisans öğrenimini yine aynı bölümde, 1998'de, doktorasını da Anadolu Üniversitesi İstatistik Anabilim da-lında 2002'de tamamlamıştır. 2002'den bu yana aynı bölümde yardımcı doçent olarak çalışmaktadır.

Sevil Bacanlı, lisans öğrenimini Hacettepe Üniversitesi İstatistik Bölümü'nde 1986'da, yüksek lisans öğrenimini 1988-de, doktorasını 1995'de yine aynı bölümde tamamlamıştır. 1995'ten bu yana aynı bölümde yardımcı doçent olarak çalışmaktadır.