ANADOLU ÜNİVERSİTESİ BİLİM VE TEKNOLOJİ DERGİSİ
ANADOLU UNIVERSITY JOURNAL OF SCIENCE AND TECHNOLOGY
Cilt/Vol.: 5 - Sayı/No: 1: 115-122 (2004)

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 doğ rusal verilere uygulanan temel istatistiksel yöntemlerle arasındaki farkları incelemektir. Çalışmada, dairesel veriler için temel parametre değ erlerinin hesaplanması açıklanmış ve dairesel veri analizindeki temel olasılık dağ ılımı olan von Mises dağ ılımı incelenmiştir. Ardından, dairesel veriler için ortalama yön testleri tanımlanmış ve von Mises dağ ılımı için ortalama yön testine ilişkin kuramsal bilgiler verilmiştir. Ayrıca, Anadolu Üniversitesi Sivil Havacılık Yüksekokulu Meydan Meteoroloji İstasyon Müdürlüğ ü'nden, Anadolu Üniversitesi havaalanı için ölçülmüş olan rüzgar yönlerinden oluşan veri değ erlerine incelenen yöntemler uygulanmış, elde edilen sonuçlar değ erlendirilmiştir.

Anahtar Kelimeler: Dairesel veri, Ortalama yön, von Mises dağ ılımı, 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, calculation 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 is 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. 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 kullanıldığ ı sıkça görülür. Açısal gözlemler, deneylerde farklı biçimlerde ortaya çıkarlar. Örneğ in biyolog, kaplumbağ aların 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 olduğ u 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 saattir. 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

[1] Anadolu Üniversitesi Fen Fakültesi İstatistik Bölümü, Yunusemre Kampüsü, Eskişehir;
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

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 $\theta°$ ($0°<\theta°<360°$), birim vektör olacaktır. Bu dairesel gösterimden dolayı, iki-boyutlu yönlere ilişkin gözlemler dairesel veri olarak da adlandırılır. Burada $\theta°$; 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 $\theta°$ açısal gözleminin (2.1) eşitliği yardımıyla radyan cinsinden θ açısına dönüştürülmesi tercih edilir.

$$\theta = \frac{\pi\theta°}{180} \quad , \quad 0<\theta<2\pi \tag{2.1}$$

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;

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

eşitliğinden, veya ortalama yön biliniyorsa

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

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

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

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

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

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. Örneğ in, n sayıda orjinal gözlem değ eri k tane sınıfa göre gruplandığ ında, i-inci sınıfın orta noktası θ_i ve sıklığ ı

$f_i, i = 1,\ldots,k \left(\sum_{i=1}^{k} f_i \right) = n$ olmak üzere;

$$\bar{C}_n = \frac{1}{n}\sum_{i=1}^{k} f_i \cos\theta_i \quad , \quad \bar{S}_n = \frac{1}{n}\sum_{i=1}^{k} f_i \sin\theta_i \quad , \tag{2.6}$$

$$\bar{R} = \sqrt{\bar{C}^2 + \bar{S}^2}$$

değ erleri hesaplanır.

2.2. Yoğunlaşma Parametresi

Yoğ unlaşma parametresi k'nın en çok olabilirlik tahmini $\hat{\kappa}$;

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

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

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

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

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

ile verilir (Fisher 1993).

2.3. Dairesel Varyans

Bileşke uzunluğ u R, verinin ortalama yön etrafındaki yoğ unlaşma miktarını gösterir. Bütün gözlem noktaları (birim vektörler) aynı yöne doğ ru büyük bir yoğ unlaşma gösteriyor ise, R değ erinin büyüklüğ ü n'e yakın bulunacaktır. Bunun aksine, veriler herhangi bir yoğ unlaşma göstermeksizin çember üzerinde düzgün bir dağ ılım gösteriyor ise, R değ eri sıfıra çok yakın bir değ er alacaktır.

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

ile tanımlandığ ında (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;

$$V = 1 - \bar{R} \quad , \quad 0 \le V \le 1 \tag{2.10}$$

değ erine ulaşılır. (2.10) eşitliğ i, örneklem dairesel varyansı olarak tanımlanır. Doğ rusal veri varyansında olduğ u gibi, dairesel varyans değ eri küçüldükçe, dağ ılım homojenleşir.

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

2.4. Dairesel Standart Sapma

V değ erinin (0,□) aralığ ına uygun bir dönüşümü;

$$v = \{-2\log_e(1-V)\}^{1/2} \tag{2.11}$$

biçiminde tanımlanır.

v ölçüsü, doğ ru üzerindeki bilinen standart sapmaya benzemektedir. Kuramsal araştırmalarda, V ve R değ erleri v değ erinden daha çok kullanılmaktadır.

V'nin küçük değ erlerinde, dairesel standart sapma;

$$v = (2V)^{1/2} \tag{2.12}$$

biçimindedir (Mardia 1972).

2.5. Dairesel Saçılım

Dairesel saçılım;

$$\delta = \frac{1-\rho_2}{2\bar{R}^2} \tag{2.13}$$

biçiminde tanımlanır. Eşitlikte;

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

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

2.6. Dairesel Standart Hata

Ortalama yön tahmininin dairesel standart hatası;

$$\sigma = \sqrt{\frac{\delta}{n}} \tag{2.15}$$

formülü ile hesaplanır. Burada, dairesel saçılım değ eri 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 verildiğ i varsayılsın. Bu noktalardan herhangi bir P noktası aşağ ıdaki özellikleri sağ lıyorsa, P noktası örneklemin medyanı olarak adlandırılır.

(i) Örneklem noktalarının yarısı P noktasından geçen PQ çapının her iki tarafında bulunmaktadır.

(ii) Örneklem noktalarının çoğ unluğ u P noktasına Q noktasından daha yakındır.

Bu koşulları sağ layan $\overline{OP}$ vektörüne örneklemin medyan yönü adı verilir.

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

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

biçiminde tanımlanan integralin çözümü ile bulunur. Simetrik bir dağ ılım için, medyan yönü simetri ekseninde olacaktır. Medyan tek bir tane olmayabilir. Fakat, bütün tek modlu dağ ılımlar tek medyana sahiptir (Mardia 1972).

Gruplanmış bir seri için medyan yönü;

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

ile verilir. Burada; l: medyan sınıfının alt sınırını, f_0: medyan sınıfının (θ_i—180°,θ_i) aralığ ındaki sıklığ ını, f_{+1}: medyan sınıfından bir sonraki sınıfın (θ_i—180°,θ_i) aralığ ındaki sıklığ ını ve h: sınıf aralığ ının uzunluğ unu gösterir.

2.8. Mod Yönü

Mod yönü $\breve{\theta}$; verinin en fazla yoğ unlaştığ ı yön anlamına gelir. Verilen bir sıklık dağ ılımının mod yönü, kesim noktası seçiminden sonra bilinen mod hesaplama yöntemiyle bulunur. Gruplanmış bir seride mod yönü;

$$\breve{\theta} = l + \frac{f_0 - f_{-1}}{2f_0 - f_{-1} - f_{+1}} \times h \tag{2.18}$$

eşitliğ i ile elde edilir. Burada; l: mod sınıfının alt sınırını, f_0: mod sınıfının sıklığ ını, f__1: mod sınıfından bir önceki sınıfın sıklığ ını, f_{+1}: mod sınıfından bir sonraki sınıfın sıklığ ını ve h: sınıf aralığ ının uzunluğ unu gösterir (Mardia 1972).

3. VON MISES DAĞILIMI

Bir θ dairesel raslantı değ işkeni, von Mises dağ ılımına sahipse, olasılık yoğ unluk fonksiyonu;

$$\mathrm{VM}(\mu,\kappa):\ f(\theta;\mu,\kappa) = \frac{1}{2\pi I_0(\kappa)} e^{\kappa\cos(\theta-\mu)} ,$$

$$0 \le \theta < 2\pi\ ,\ \kappa \ge 0\ ,\ 0 \le \mu < 2\pi \tag{3.1}$$

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

$$I_0(\kappa) = \frac{1}{2\pi}\int_0^{2\pi} e^{\kappa\cos(\phi-\mu)} d\phi \tag{3.2}$$

olarak verilir.

Von Mises dağ ılımında κ değ erinin büyük olması, anakütle ortalama yönü ve mod yönü etrafında daha büyük bir kümelenme olduğ unu gösterir. Buna göre κ, ortalama yöne ilişkin yoğ unlaşmayı ölçen bir parametredir (Jammalamadaka ve Sen Gupta 2001).

Doğ rusal veri setleri için, normal dağ ılım, sıkça kullanılan bir dağ ılım olmakla birlikte, şekilsel istatistiksel analiz, σ^2 değ erine bakılmaksızın yapılabilmektedir. Fakat, dairesel veri setleri için temel dağ ılım olan von Mises dağ ılımında ise, şekilsel istatistiksel analiz κ değ erine bakılmadan yapılamamaktadır.

Yoğ unlaşma parametresinin, dairesel varyans ile olan ilişkisi;

$$V = 1 - \rho = 1 - A(\kappa) \tag{3.3}$$

biçiminde tanımlanır (Mardia 1972).

Von Mises dağ ılımının momentleri;

Ortalama yön : μ

Ortalama bileşke uzunluğ u : $\rho = A_1(\kappa)$

Dairesel saçılım : $\delta = \dfrac{1}{\kappa A_1(\kappa)}$

$\alpha_p = A_p(\kappa)$

$\beta_p = 0$, p □ 1

biçimindedir.

4. DAİRESEL VERİLERDE TEK-ÖRNEK 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 aralığ ının belirlenmesi yoluyla test uygulamak, ikincisi ise, örneklem hacmine dayalı olarak doğ rudan test uygulamaktır.

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

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

değ erleri hesaplanır. Bu işlemlerden sonra, ortalama yön için $(1—\alpha)$'lık güven aralığ ı;

$$(\mu_0 - \sin^{-1}(z_{\alpha/2}\ \sigma)\ ,\ \mu_0 + \sin^{-1}(z_{\alpha/2}\ \sigma)) \tag{4.1}$$

ile bulunur. Burada, $z_{\alpha/2}$; standart normal dağ ılım tablosundan elde edilir. Kurulacak hipotezler;

$$H_0: \mu = \mu_0$$
$$H_1: \mu \neq \mu_0 \tag{4.2}$$

olduğ unda, ortalama yön değ eri (4.1)'de verilen aralığ ın 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 aralığ ının belirlenemediğ i durumda doğ rudan uygulanabilmektedir. Fakat bu test, örneklem hacmi 25 veya daha fazla olduğ unda kullanılabilir. Bu yönteme göre, öncelikle dairesel standart hata s yukarıda verilen testteki gibi hesaplanır. Test istatistiğ i;

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

olarak verilir. Bu değ er α güvenilirlik seviyesinde, standart normal dağ ılım tablosundan elde edilen değ erle karşılaştırılır. $|S|$ değ eri tablo değ erinden büyük olduğ unda 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 yoğ unluk parametresi ile von Mises dağ ılımına $VM(\mu,\kappa)$ sahip olduğ u varsayımı altında, ortalama yönün μ_0 gibi bir değ ere eşit olup olmadığ ı testi incelenecektir. İlk olarak, ortalama yön için güven aralığ ının belirlenmesi yöntemiyle uygulanan test incelenecektir. Doğ rudan test uygulamasında ise, yoğ unlaşma parametresinin bilindiğ i ve bilinmediğ i durum olmak üzere iki durum söz konusudur.

Güven aralığ ı belirlenerek uygulanan testte, ortalama yön için dairesel standart hata değ eri hesaplanır. Fakat, burada von Mises dağ ılımı için özel bir hal alan standart hata formülü kullanılmaktadır.

$$\sigma_{VM} = \frac{1}{\sqrt{n\bar{R}\hat{\kappa}}} \tag{4.4}$$

Buna göre; ortalama yön için $(1—\alpha)$'lık güven aralığ ı;

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

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

Güven aralığ ı belirlenmeden doğ rudan yapılan test, örneklem hacmi n ve yoğ unlaşma parametresi $\hat{\kappa}$ tahmin değ erinin büyüklüğ üne bağ lı olarak uygulanır.

Yoğ unlaşma parametresinin bilinmediğ i durumda, von Mises dağ ılımı için ortalama yönün dairesel standart hatası σ_{VM} Eşitlik (4.4)'teki gibi hesaplanır. Test istatistiğ i ise;

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

olarak verilir. Bu değ er α güvenilirlik seviyesinde, standart normal dağ ılım tablosundan elde edilen değ erle karşılaştırılır. Karşılaştırmalar alternatif hipotezin değ işik durumları için aşağ ıdaki 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 aldığ ı değ erlere bağ lı olarak aşağ ıdaki durumlarda kullanılır:

Çizelge 1. n ve κ Değ erlerine Bağ lı Olarak Testin

$\hat{\kappa}$	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

Uygulanabildiğ i Durumlar

Yoğ unlaşma parametresi κ'nın κ_0 gibi bilinen bir değ ere eşit olduğ u varsayıldığ ında;

$\kappa_0 \square 2$ ise, Von Mises dağ ılımı için ortalama yönün dairesel standart hatası σ_{VM};

$$\sigma_{VM} = \frac{1}{\sqrt{n\rho\kappa_0}} \tag{4.7}$$

biçimindedir. Test istatistiğ i ise;

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

olarak verilir. Bu değ er α güvenilirlik seviyesinde, standart normal dağ ılım tablosundan elde edilen değ erle 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.

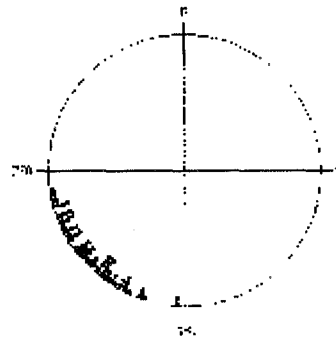
(Fisher 1993).

5. UYGULAMA

Dairesel verilerin en çok kullanıldığ ı alanlardan birisi meteorolojik gözlem çalışmalarıdır. Yukarıda incelenen yöntemlerin gerçek veriler üzerinde uygulanması amacıyla, Anadolu Üniversitesi Sivil Havacılık Yüksekokulu Meydan Meteoroloji İstasyon Müdürlüğ ü'nden, Anadolu Üniversitesi havaalanı için ölçülen ve rüzgar yönleri ve rüzgar hızlarından 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 aralığ ında 73'er rüzgar yönü ve hızı verisi elde edilmiştir. Bu dairesel gözlem değ erleri için elde edilen dairesel grafikler aşağ ıda verilmektedir.

Şekil-1'de, verilerin ham veri grafiğ i görülmektedir.



Şekil 1. Rüzgar Yönü Verilerinin Ham Veri Grafiğ i

Şekil-2'de ise, rüzgar yönü verilerinin 10°'lik grup genişliğ i seçildiğ i durumdaki dairesel histogramı görülmektedir.

$$\theta = \frac{\pi\theta^{\circ}}{180}, \qquad 0 < \theta < 2\pi$$

Şekil 2. Rüzgar Yönü Verilerinin Dairesel Histogramı

Şekil-3'te, rüzgar yönü verilerinin 10°'lik grup genişliğ i seçildiğ i durumdaki gül şeması görülmektedir.

Şekil 3. Rüzgar yönü verilerinin gül şeması

Şekillerden de görüldüğ ü gibi, veriler 184°—262° aralığ ında değ er almakta ve dağ ılımın şekli von Mises dağ ılımına uymaktadır.

Belirlenen gün ve saat için seçilen 73 gözlem değ eri için S-Plus ve Oriana paket programları kullanılarak elde edilen tanımlayıcı istatistik değ erleri; ortalama rüzgar yönü: $\bar{\theta}$ = 229.364°, ortalama rüzgar hızı: 13.62

km/saat, bileşke uzunluğ u: R = 69.496, ortalama bileşke uzunluğ u: $\bar{R}$ = 0.952, yoğ unlaşma 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 yoğ unluk parametresi ile von Mises dağ ılımına VM(229.36, 10.65) sahiptir.

Güven aralığ ının belirlenmesi yöntemiyle ortalama yön test edilmek istendiğ inde kurulacak hipotezler;

H_0: $\mu = 229°$

H_1: $\mu \neq 229°$

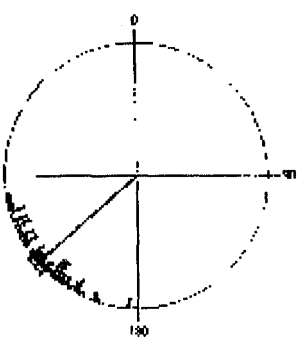
biçiminde olsun. Eşitlik (4.4)'e göre;

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

elde edilir. Buna göre; ortalama yön μ için (1—α)'lık güven aralığ ı (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 aralığ ının grafik gösterimi Şekil-4'te verilmektedir.



Şekil 4. Rüzgar Yönü Verilerinde Ortalama Rüzgar Yönü İçin %95'lik Güven Aralığ ı

Burada $\alpha = 0.05$ olarak alınmıştır. H_0 hipotezinde verilen 229° değ eri verilen aralığ ın içinde yer aldığ ı için, %5 anlamlılık seviyesinde H_0 hipotezi kabul edilir. Ortalama yön 229°'dir.

Doğ rudan test uygulanmak istendiğ inde daha önce de belirtildiğ i gibi, yoğ unlaşma parametresi κ'nın bilindiğ i ve bilinmediğ i durum olmak üzere iki durum söz konusudur.

Önce κ'nın bilinmediğ i durum göz önüne alınsın. Buradaki test istatistiğ i (4.6) eşitliğ inden;

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

olarak bulunur. $|0.172| < 1.9604$ olduğ undan 0.05 anlamlılık seviyesinde H_0 hipotezi kabul edilir. Ortalama yön 229°'ye eşittir.

κ'nın bilindiğ i durumda ise, $\kappa_0 = 10$ değ eri göz önüne alınsın. $\kappa_0 \geq 2$ olduğ undan, von Mises dağ ılımı için ortalama yönün dairesel standart hatası σ_{VM} (4.7) eşitliğ inden;

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

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

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

bulunur. $|0.167| < 1.9604$ olduğ undan %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 değ erleri doğ rusal ya da dairesel olabilir. Fakat, dairesel verilerin kullanıldığ ı 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 bağ lı yapıdaki verilerin analizi üzerinde durulmakla birlikte, yönsel veri çalışmaları, araştırmacılara çok geniş bir alanda ilerleme olanağ ı vermekte ve yeni istatistiksel yöntemler geliştirmede çok verimli bir alan olduğ u görülmektedir. Ayrıca doğ al, fiziksel, tıbbi ve de sosyal bilimlerde ortaya çıkan problemler için yeni ve farklı uygulamalar geliştirilebilmektedir (Peker 2002).

KAYNAKÇA

KAYNAKÇA

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

Cheeney, R. F. (1983). Statistical methods in geology, George Allen & 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 istatistik, (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 öğ renimini Hacettepe Üniversitesi İstatistik Bölümü'nde 1994'de, yüksek lisans öğ - renimini yine aynı bölümde, 1998'de, doktorasını da Anadolu Üniversitesi İstatistik Anabilim dalında 2002'de tamamlamıştır. 2002'den bu yana aynı bölümde yardımcı doçent olarak çalışmaktadır.

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