

LOJİSTİK REGRESYON ANALİZİ İLE ESKİŞEHİR’İN SIS KESTİRİMİNİN İNCELENMESİ

Cengiz AKTAŞ* , Orkun ERKUŞ**

Geliş: 12.10.2009 Kabul: 24.12.2009

ÖZET

Bu çalışmada önce, bağımlı değişkenin iki düzeyli olması durumunda demografik, davranış ve risk faktörüyle ilgili tahmin çalışmalarında oldukça sık kullanılan lojistik regresyon analizi teorik olarak kısaca incelendi. Sınıflandırma analizlerinin en çok kullanılanlarından biri olan ve birçok konuda uygulama alanı bulunan lojistik regresyonun parametre tahminine ilişkin kestirim yöntemi detaylı bir şekilde incelenerek, katsayıların yorumu için odds oranı anlatıldı. Daha sonra Eskişehir’de sis olayı ve havacılık faaliyetleri üzerine etkileri kısaca incelendi. Araştırmanın uygulama bölümünde ise, Eskişehir’de sis meydana gelmesi olasılığını verecek en uygun denklem belirlendi.

Anahtar Kelimeler : Lojistik Regresyon, Sis, Sınıflama, Odds

INVESTIGATION OF FOG FORECASTING OF ESKİŞEHİR USING LOGISTIC REGRESSION ANALYSIS

ABSTRACT

In this study, firstly, it was briefly examined which logistic regression analysis are frequently used in studies for estimating associations that demographic, behavioral, and risk factor variables have on a dichotomous outcome. Because of logistic regression have wide area of usage in classification analysis and many practical area, conjecture method used in parameter estimate have been observed in detail. Interpretation of the coefficients is explained the odds ratio. Furthermore, event of fog in Eskişehir and effects on aviation were briefly examined. As an application for this study, the most efficient equation to give the probability of happening of fog in Eskişehir were fitted.

Keywords: Logistic Regression, Fog, Classify, Odds

* Eskişehir Osmangazi Üniv. Fen-Ed. Fak. İstatistik Böl., Eskişehir, caktas@eou.edu.tr

** Eskişehir Askeri Meteoroloji İstasyon Müdürlüğü, Eskişehir.

1. GİRİŞ

Bir gözlemi birkaç küteden birine atamak, sınıflamadır. Eğer kütleler ortak varyans-kovaryans matrisine sahip ve normal dağılmışsa, diskriminant analizi kestiricileri, diskriminant analizi problemleri için lojistik regresyon kestiricilerine tercih edilebilir. Bununla birlikte pek çok diskriminant analizi uygulamasında değişkenlerden en az birinin kategorik değişken olması nedeniyle çok değişkenli normal varsayımı geçerli olmayacaktır. Böyle durumlarda bağımsız değişkenlerin kategorik ve sürekli olmaları konusunda bir kısıt getirmeyen, gözlemlerin atanması amacıyla kullanılabilen lojistik regresyon analizi önerilmektedir (Press ve Wilson, 1978, s.2).

Meteorolojik hadiselerin en önemlilerinden biri de sis olayıdır. “Stratus (St) bulutunun yer yüzeyinde meydana gelmiş hali” şeklinde kısaca tanımlanabilen sis; yatay görüş mesafesini ve eğer kalınlığı fazla ise dikey görüş mesafesini engelleyen bir hadisedir. Bu nedenle sis insan yaşamını doğrudan etkiler. Tarım ve ulaşım, özellikle yatay ve düşey görüş mesafesinin çok önemli olduğu havacılık, başta olmak üzere birçok insan faaliyeti ve çevre ve insan sağlığı, sisten önemli ölçüde etkilenir. Her yıl kara, deniz ve hava taşımacılığında sis nedeniyle pek çok can ve mal kaybı olmaktadır. Bu nedenlerle, gerek ekonomik yatırımlarda ulaşım açısından, gerekse şehirleşmede insan aktivitesi bakımından, bir bölgenin sis analizinin yapılması gereklidir.

Eskişehir’in kış aylarında uzun süreli sislerin görüldüğü birkaç şehirden biri olması ve burada yoğun uçuş faaliyetlerinin yapıldığı sivil ve askeri havaalanlarının bulunması, uçuculuğu olumsuz etkileyen sis hadisesini ön plana çıkarır.

Meteoroloji ofislerinde sis tahmini değişik hava kartları, temp diyagramları ile belli periyotların oluşabilecek meteorolojik parametrelerini hesaplayan model sonuçlarına göre yapılmaktadır. Meteoroloji balonlarıyla elde edilen atmosferin yukarı seviyelerdeki bilgileriyle hazırlanan temp diyagramları enversiyonun ve nemin hesaplanmasında faydalı olsa da gün içerisinde iki kere elde ediliyor olması ve atmosferin akışkan yapısından dolayı ilerleyen saatler için bu bilgileri kullanarak tahmin yürütmek zorlaşmaktadır. Ayrıca bu balonların yurt genelinde 8 noktadan bırakılması (Ankara, İstanbul, İzmir, Isparta, Samsun, Adana, Diyarbakır, Erzurum) Eskişehir için yaklaşık değerlerin alınmasına sebep olmaktadır.

Eskişehir’de kışın görülen kar yağışı ve sonrasında yer seviyesindeki uzun süreli buzlanma Eskişehir için çalıştırılan MM5 gibi modellerde gün içerisinde beklenen en yüksek sıcaklıkta 5-6 derecelere varan sapmalara sebep olmaktadır. Bundan dolayı görülen bir sis hadisesinde bu modeller sisin dağılacığı saatleri tam olarak hesaplayamamaktadır. Özellikle Eskişehir’deki sivil ve askeri havacılığın yoğun çalışmaları sebebiyle uçuş planlarının hazırlanması ve güvenliğinin üst seviyede sağlanabilmesi için sis gibi görüşü kısıtlayan ve uçuş için tehlike arz eden hadiselerde tahminlerin doğruluk derecesini artırmak gerekmektedir. Bu nedenle

çalışmamızın amacı, meteorolojinin ikinci derece veriler olarak kullandığı veriler yardımıyla, Eskişehir’de sis olup olmayacağını ya da sis meydana gelmesi olasılığını belirleyebilecek denklemi elde etmektir.

Lojistik regresyon modelleri, son yıllarda biyoloji, tıp, ekonomi, meteoroloji, tarım ve veterinerlik ve taşıma sahalarında yaygın olarak kullanılmaktadır. Literatürde lojistik regresyon analizi üzerine pek çok çalışma yapılmıştır. Bunlardan bazıları şunlardır:

Cornfield (1962), lojistik regresyondaki katsayı tahmin işlemlerinde diskriminant fonksiyonu yaklaşımını ilk kez kullanarak popüler hale getirmiştir. Bu dönüm noktası niteliğindeki çalışmadan sonra da Breslow ve Day (1980) epidemioloji, Abbott (1985) yaşam analizi ile ilgili uygulamalı çalışmalar yapmışlardır. Gardside ve Glueck (1995) insanlarda beslenme şekli, sigara ve alkol kullanımı, fiziksel aktivite gibi risk faktörlerinin kalp hastalığı üzerindeki etkilerini incelemişlerdir.

Türkiye’de de bu konuda çeşitli alanlarda çalışmalar yapılmıştır. Bulardan bazıları ise şunlardır: Vupa ve Çelikoğlu (2006) akciğer kanseri hastaları için lojistik regresyon modeli önermişlerdir. Ünsal ve Güler (2005) Türk bankacılık sektörünü lojistik regresyon analiziyle incelerken, Aktaş ve Yılmaz (2001) LPG kullanan özel araç sürücülerinin sınıflandırılmasını, lojistik regresyon analiziyle incelemişlerdir.

Literatürde sis ve havaalanı ile ilgili yapılan çalışmalardan bazıları da aşağıdaki gibidir:

Lewis (2004) yaptığı çalışmada regresyon analizini kullanarak deniz sisi tahminleri yapmıştır. Stern ve Parkyn (1999), Melbourne havaalanı için sis tahminleri üzerinde çalışırken, Fabbian, Dear ve Lellyett (2007) Canberra uluslararası havaalanı için sis tahminlerini geliştirmeye çalışmışlardır.

Uğur (1984) yaptığı çalışmada Yeşilköy Hava Meydanının sis ve özelliklerini incelerken, Tuncer (1991) regresyon analizi yardımıyla Esenboğa Hava Limanı için sis tahmin çalışması yapmıştır.

Çalışmamızın ikinci bölümünde lojistik regresyon analizi anlatılarak, üçüncü bölümde özellikle askeri ve sivil havacılıkta çok önemli bir yer tutan Eskişehirde sis olayı ve havacılık faaliyetleri üzerine etkileri kısaca incelenecektir. Dördüncü bölümde ise uygulamaya yer verilecektir.

2. GÖZLEMLERİN VAROLAN GRUPLARA ATANMASI

Değişkenler arası ilişkileri incelemede en çok kullanılan istatistik yöntemlerinden biri regresyon analizidir. Regresyon analizi çözümüne başlamadan yapılması gereken değişkenlerin niteliklerinin bilinmesi ve bağımlı değişken ile bağımsız değişkenin en iyi şekilde tayin edilmesidir. Genelde bilinen bağımlı değişken ölçülebilir

nitelikte olup, sürekli bir değişkendir. Ancak her zaman bağımlı değişken ölçülebilir nitelikte olmayabilir.

Lojistik regresyon analizinin kullanım amacı istatistikte kullanılan diğer model yapılandırma teknikleriyle aynıdır. En az değişkeni kullanarak en iyi uyuma sahip olacak şekilde bağımlı (sonuç) değişkeni ile bağımsız değişkenler kümesi (açıklayıcı değişkenler) arasındaki ilişkiyi tanımlayabilen ve genel olarak kabul edilebilir modeli kurmaktır.

Lojistik regresyonu, doğrusal regresyondan ayıran en belirgin özellik ise lojistik regresyonda bağımlı değişkeninin kategorik değişken olmasıdır. Lojistik regresyon ve doğrusal regresyon arasındaki bu fark hem parametrik model seçimine, hem de varsayımlara yansımaktadır. Lojistik regresyonda da, doğrusal regresyon analizinde olduğu gibi bazı değişken değerlerine dayanarak kestirim yapılmaya çalışılır, ancak iki yöntem arasında üç önemli fark vardır:

1-Doğrusal regresyon analizinde tahmin edilecek olan bağımlı değişken sürekli iken, lojistik regresyonda bağımlı değişken kesikli bir değişken olmalıdır.

2-Doğrusal regresyon analizinde bağımlı değişkenin değeri, lojistik regresyonda ise bağımlı değişkenin alabileceği değerlerden birinin gerçekleşme olasılığı kestirilir.

3-Doğrusal regresyon analizinde bağımsız değişkenlerin çoklu normal dağılım göstermesi koşulu aranırken, lojistik regresyonun uygulanabilmesi için bağımsız değişkenlerin dağılımına ilişkin hiçbir ön koşul yoktur (Çoşkun v.d, 2004: 43).

Gözlemleri verilerin yapısında bulunan olası gruplara atamak için;

i)Kümeleme analizi,

ii)Diskriminant analizi,

iii)Lojistik regresyon analizi tekniklerinden yararlanılır.

Kümeleme analizinde, verilerin yapısındaki grup sayısı bilinmemekte, gözlemler uzaklık ya da benzerlik ölçütlerine göre kümelenebilir. Burada amaç yalnızca gözlemlerin oluşturduğu kümenin yapısını bulmaktır.

Diskriminant ve lojistik regresyon analizinde ise yapısındaki grup sayısı bilinmemekte ve bu verilerden faydalanarak bir ayrım modelini elde edilmektedir. Kurulan bu model yardımı ile veri kümesine yeni alınan gözlemlerin gruplara atanması yapılmaktadır (Başarır, 1990:1). Çalışmamızda lojistik regresyon modeli uygulanacağından sadece lojistik regresyon analizi incelenecektir.

2.1. İki Düzeyli Lojistik Regresyon Analizi

k bağımsız değişken ve N gözlem olduğunda doğrusal regresyon modelinin genel formu i.gözlem için

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki} + \epsilon_i \text{ 'dir.} \quad (1)$$

Örneklem büyüklüğü n olduğunda ise doğrusal regresyon modeli

$$y_i = \hat{\beta}_0 + \hat{\beta}_1 x_{1i} + \hat{\beta}_2 x_{2i} + \dots + \hat{\beta}_k x_{ki} + \epsilon_i \quad (2)$$

şeklinde yazılır.

Bağımlı değişkenin alabileceği değerlerin 0-1 arasında olmasını sağlamak için bağımsız değişken ve bağımlı değişken arasında eğrisel bir ilişkiyi sağlayan modeli kullanmak daha uygundur. β_1 'in işaretine göre S veya ters S şeklinde olan eğrileri sağlayan

$$E(y_i) = \pi_i = \frac{\exp(\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki})}{1 + \exp(\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki})} \quad (3)$$

formundaki bu fonksiyona “Lojistik Fonksiyon” adı verilir. Bu lojistik fonksiyonlar genellikle S şeklinde fonksiyon olarak isimlendirilir. Bunlar 0 ve 1 asimtotlarına sahiptir ve böylece E(y), 0 ile 1 sınırları arasında kalır.

Lojistik fonksiyonun diğer bir özelliği de kolayca doğrusallaştırılabilir olmasıdır ve

$$\eta = \ln\left(\frac{\pi_i}{1 - \pi_i}\right) \quad (4)$$

dönüşümü yapılarak bağlantı fonksiyonu elde edilir. Eşitlik (4)'deki $\pi_i/(1-\pi_i)$ oranı ise “Odds Oranı” olarak nitelendirilir. ln odds dönüşümü ise “Lojit” olarak isimlendirilir ve ln odds için elde edilen

$$E(y_i) = \eta = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki} \quad (5)$$

modele “Lojistik (ya da lojit) Regresyon Modeli “ denir. E(y_i) ise -∞ , ∞ arasında değer almaktadır (Agresti, 1990: 106).

2.2. Lojistik Sınıflandırma ve Katsayıların Yorumlanması

Bir gözlemi birkaç gruptan birine atamak, sınıflamadır. Genellikle $P(y_i=1/x_i)$ değerini belirlemek amacıyla kullanılan lojistik regresyon modeli aynı zamanda bir sınıflandırma modeli olarak da kullanılır. Seçilen n birimlik bir örneklem sonucu elde edilen

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \dots + \hat{\beta}_k x_{ki} \quad (6)$$

lojistik regresyon denklemi yardımıyla bulunan

$$P_i = \frac{e^{\hat{y}_i}}{1 + e^{\hat{y}_i}} \quad (7)$$

değerinin ≥ 0.5 olması durumunda $y_i=1$, $\pi_i, P_i < 0.5$ ise $y_i=0$ biçiminde sınıflandırılır (Aktaş ve Yılmaz, 2001: 253).

Lojistik regresyon fonksiyonunda tahmin edilen regresyon katsayılarının yorumlanması doğrusal regresyon modelindeki kadar kolay değildir. x eksenindeki başlangıç noktasına göre hazırlanan lojistik regresyon modelinde x değişkenindeki bir birimlik artışın tesirini ölçmek zordur. β_1 katsayısı yorumlanırken x'deki bir birimlik artış için $\pi_i/(1-\pi_i)$ odds tahmini ile $\exp(\beta_1)$ çarpılarak elde edilen lojistik regresyon fonksiyonundan yararlanılır.

Lojistik modeldeki etkiler odds'a dayanır. x'in bir değerinde kestirilen odds'un, diğer değerinde kestirilen odds'a oranı olarak verilmektedir. Bu istatistik x=1 olan bireylerin x=0 olan bireylere nazaran bağımlı değişkenin kaç kat daha fazla 1 olarak görüldüğü sonucunu verir (Bircan, 2004: 29).

3. ESKİŞEHİR'DE SİS OLAYI VE HAVACILIK FAALİYETLERİ ÜZERİNE ETKİLERİ

Atmosferik su buharının çok küçük su damlaları şeklinde yoğunlaşarak, yerde yatay görüşü engelleyen bulut oluşturmasına sis adı verilir. Diğer bir ifadeyle; Stratus (St) bulutunun yer yüzeyinde meydana gelmiş haline sis denir. Sis olduğu zamanlarda yer yüzeyindeki görüş uzaklığı 1 km' den daha azdır. Genel olarak sisin tabanı yerden 300 Fit veya daha fazla yükselmesi halinde, Stratus bulutu olarak rapor edilir. Sis içinde nisbi nem % 80 ve üzerindedir.

Eskişehir'in topoğrafik yapısını, Sakarya ve Porsuk havzalarındaki düzlükler ile bunları çevreleyen dağlar oluşturur. Havza düzlüklerini kuzeyden Bozdağ, Sündiken sıradağları, batı ve güneyden ise İç Batı Anadolu eşiğinin doğu kenarında yer alan Türkmen Dağı, Yazılıkaya Yaylası ve Emirdağ kuşatır. İç Anadolu Bölgesinin karakteristik özelliğini taşıyan Eskişehir'de daha çok termik kökenli Sibirya yüksek

basıncının etkisi ile yerin aşırı radyasyon kaybına bağlı olarak oluşan radyasyon sisleri ile Akdeniz üzerinden gelen sıcak havanın soğuk kara üzerinde soğuyup yoğunlaşması ile adveksiyon sisleri görülmektedir. Coğrafi yapısı ve kışın hakim olan doğulu hafif ve orta rüzgâr, yoğun sisli günlerde, sisin dağılmasını engelleyerek uzun süre kalkmamasına yol açmaktadır. Ayrıca soğumanın çok olduğu kış günlerinde yerdeki buz tabakası da ısınmayı etkilediği için sisin ısınmayla birlikte dağılma süreci uzamaktadır.

Hava ulaşımında meteorolojinin yeri oldukça önemlidir. Meteorolojik bilginin mevcut olmadığı bir havaalanında uçakların emniyetli olarak iniş ve kalkış yapmaları mümkün değildir. Uçuş hizmetlerini yürütmek üzere her havaalanında Uluslararası Sivil Havacılık Teşkilatı (ICAO) şartlarına uygun mutlaka bir meteoroloji istasyonu bulunmaktadır. Bu istasyonlar günün her saatinde, aralıksız, uçaklara gerekli meteorolojik desteği sağlamak için çalışmalar yapar. Uçuşa gidecek pilotlar, uçuş programlanmadan önce mutlaka meteoroloji istasyonlarından iniş-kalkış yapacakları meydanlar ile uçacakları yolboyunun meteorolojik bilgilerini alırlar.

Hava alanlarında iniş ve kalkışları önemli derecede etkileyen faktör ise sistir. Sis oluşumunda meteorolojik rüyet (görüş mesafesi) ve pist görüş uzaklığı (RVR) çok dar olduğundan, iniş ve kalkış esnasında pilotlar pisti yeterince göremezler. Meydanın hava trafiğine açılabilmesi için böyle durumlarda meteoroloji istasyonları, pist görüş uzaklığını sık aralıklarla sıhhatli bir şekilde ölçmelidir (Yağan, Çamalan ve Akgün, 1991: 2).

Meteorolojik hadiseler uçakların iniş ve kalkışı sırasında çok büyük önem arz eder. Pilotların görsel yanılgılar yaşamasına sebep olur. Hava durumu ile ilgili görsel yanılgılar şu şekildedir:

- a. Hafif yağmur, sis, pus, duman veya karanlık genelde yüksek hissi verir.
- b. Alçak sisin (300 feet kalınlığındaki) hemen üzerinde meydan ve pist görülebilir, fakat sisin içine girildiğinde ileri görüş ve dikey görüş kaybedilir. Ayrıca sisin içine giriş “baş yukarıda” yanılgısına sebep olur ve pilot buna tepki olarak bası aşağı verir.
- c. Pus içerisinde uçuş pistin uzak olduğu yanılgısına sebep olur.
- d. Şiddetli yağmurlar derinlik ve mesafe algılamasını etkiler. Ön cam üzerindeki yağmur yüksekmiş hissi verir ve baş aşağı verildiğinde istenen süzülüş hattının altına inilir. Gün ışığında yağmur yaklaşma ışıklarının etkisini azaltır ve pistten uzakmış hissi verir. Gece yağmur yaklaşma ışıklarının parlaklığını artırır. Bu durum piste yakınmış hissi verir ve pist esiğinden daha önce iniş yapılacak şekilde baş aşağı durumuna sebep olur.

e. Yan rüzgar durumlarında pist ışıkları ve çevre hava aracının uçtuğu baştan farklı bir açı ile görülür. Uçuş mürettebatı düşme düzeltmesini muhafaza etmeli ve pist orta hattı uzanımında alçalmasına devam etmelidir.

f. Islak pistler küçük bir ışığı bile yansıtarak derinlik hissi algılamasını yanıltır. Bu durum genelde palyede geç kalma ve sert iniş ile sonuçlanır (Serin, 2006: 81).

Eskişehir ilini sisli illerden ayıran en önemli özelliği ise Eskişehir’de hem Anadolu Üniversitesi Sivil Havacılık Yüksekokulu ve sivil bir hava alanı olması, hem de Türk Hava Kuvvetleri’nin 1.Ana Jet Üssü’ne sahip olmasıdır. Dolayısıyla Eskişehir uçuşların yoğun olduğu iller arasında yer almaktadır. Bu nedenle bundan sonraki kısımda lojistik regresyon yardımıyla, sis kestirimlerinde kullanılabilecek denklem belirlenecektir.

4. VERİ KAYNAĞI, DEĞİŞKENLERİN BELİRLENMESİ VE AMPİRİK BULGULAR

Meteoroloji uzmanları gerek sis, gerekse diğer meteorolojik olayların kestiriminde birçok veriyle çalışmaktadır. Birinci derece veriler olarak kullanılan haritalar, diyagramlar, bilgisayar çıktı ve görüntüleri sayısal olarak ifade edilemediği için ve daha çok meteorolojik bilgiyle yorumlandığı için istatistiki analizler için uygun olmamaktadır. İkinci derecede önemli veriler havanın kapalılık durumu, rüzgarın yönü ve hızı, sıcaklık, nisbi nem, basınç vb. sayısal olarak ifade edilebilen verilerdir; ancak Eskişehir’deki meteorolojide bu verilerin kestirim amacıyla kullanılması genelde gözardı edildiği için sisle ilgili kestirimlerin isabet derecesi azalmaktadır.

Çalışmamızdaki veriler 2007 yılı Ocak, Kasım, Aralık ayları ve 2008 yılı Ocak ayına ait olup, sisin yoğun olarak görüldüğü 2 Z ve 8 Z saatleri (04.00 ve 10.00) arasındaki 861 gözlemden oluşmaktadır. Veriler, T.C. Çevre ve Orman Bakanlığı Devlet Meteoroloji İşleri Genel Müdürlüğünden alınmıştır.

Analizdeki bağımlı değişken görüş mesafesi olarak,

$$y_i = 0 \text{ Görüş mesafesi} \geq 1000 \text{ m}$$

$$y_i = 1 \text{ Görüş mesafesi} < 1000 \text{ m}$$

şeklinde kodlanarak uygulanmıştır. Bağımsız değişkenler ise aşağıdaki şekilde tanımlanmıştır:

$$\text{SICAK: Sıcaklık; } ^\circ\text{C,}$$

NEM: Havanın hacim birimi (m³) başına içerdiği su buharının gram cinsinden ağırlığı,

BASINÇ: Basınç; milibar,

RÜZHIZ: Rüzgar Hızı; knot cinsinden ölçülmektedir.

KAP: Kapalılık; 8/8 üzerinden değerlendirilmektedir. 0/8, 1/8, 2/8 “0”, 3/8, 4/8, “1”, 5/8, 6/8, 7/8 “2”, 8/8 “3” ve gökyüzünün görülmediği günler ise “4” olarak kodlanmıştır.

RÜZYÖN: Rüzgar yönü; derece ile ölçülmektedir. 0⁰ – 45⁰ “0”, 46⁰ – 90⁰ “1”, 91⁰ – 135⁰ “2”, 136⁰ – 180⁰ “3”, 181⁰ – 225⁰ “4”, 226⁰ – 270⁰ “5”, 271⁰ – 315⁰ “6”, 316⁰ – 360⁰ “7” olarak kodlanmıştır.

SAAT: Saat; 2 Z ve 8 Z saatleri arası alınmıştır.

SPSS paket programı kullanılarak, ileri doğru değişken seçme tekniğiyle, lojistik regresyon analizi sonucu elde edilen enbüyük olasılırlık katsayı kestirimleri ve diğer çıktı sonuçları Tablo 1’de verilmiştir:

Tablo 1. İleriye Doğru Değişken Seçme Tekniğine Göre Analiz Sonuçları

| Değişken | $\hat{\beta}_i$ | S.E | Wald | df | Anlamlılık | Exp($\hat{\beta}_i$) |
|----------|-----------------|---------|---------|----|------------|------------------------|
| SICAK | -0,0843 | 0,0407 | 4,2978 | 1 | 0,0382 | 0,9192 |
| NEM | 0,3863 | 0,0686 | 31,6836 | 1 | 0,0000 | 1,4715 |
| BASINÇ | 0,0625 | 0,0262 | 5,6859 | 1 | 0,0171 | 1,0645 |
| RÜZHIZ | -0,1664 | 0,0483 | 11,8797 | 1 | 0,0006 | 0,8467 |
| KAP | 0,6387 | 0,1249 | 26,1508 | 1 | 0,0000 | 1,8941 |
| Sabit | -97,855 | 25,0257 | 15,2896 | 1 | 0,0001 | |

Tablo 1’deki sonuçlara göre sis oluşumundaki etkili faktörlerin, hava sıcaklığı, basınç, nem rüzgar hızı ve gökyüzü kapalılığı olduğu sonucunu ortaya koymuştur. Çünkü Wald istatistik değerlerinin tamamı anlamlıdır.

Dolayısıyla sınıflama için kullanılacak denklem;

$$\hat{y} = -97,855 - 0,0843x_{1i} + 0,3863x_{2i} + 0,0625x_{3i} - 0,1664x_{4i} + 0,6387x_{5i} \quad (8)$$

olacaktır.

Çoklu doğrusal regresyonda, regresyon katsayılarının yorumu açıktır. Diğer bağımsız değişkenlerin değerleri aynı kalmak koşuluyla bir bağımsız değişkendir

bir birimlik değişimin bağımlı değişkende yarattığı değişim miktarını ifade eder. Oysa lojistik regresyondaki katsayı kestirimlerinin yorumu çoklu doğrusal regresyondaki gibi değildir. Sis meydana gelmesi olasılığının sis olmaması olasılığına oranı olan odds'lar ile yorum yapılmaktadır. Bu değerler de $Exp(\hat{\beta}_i)$ sütunundaki değerlerdir. Tablo 1'deki $Exp(\hat{\beta}_i)$ değerlerine göre sis oluşumundaki en önemli değişkenin, havanın kapalılık durumu olduğu tespit edilmiştir. Diğer değişkenler sabit kalmak koşuluyla, kapalılıktaki bir birimlik artış sis meydana gelmesi olasılığını 1,8941 kat artıracaktır. Yine diğer değişkenler sabit kalmak koşuluyla nem miktarındaki bir birimlik artış, sis meydana gelmesi olasılığını 1,47 kat artıracaktır.

Tablo 1'deki sonuçlara göre, sıcaklık ve rüzgar hızının sis oluşumunu azalttığı görülmektedir. Bu da meteorolojik olarak beklenen bir durumdur. Çünkü sıcaklıktaki artış yeryüzünün ısınmasını sağlayacağından dolayı sisin kalkmasına yardımcı olacaktır. Yine rüzgar hızında olan artış sisin oluşmasını engeller ve var olan sisin dağılmasına neden olur.

Çoklu doğrusal regresyonda katsayıların anlamlılığına ilişkin genel anlamlılık sınaması F testine karşılık gelebilecek benzer bir test lojistik regresyon analizi için geliştirilmiştir. L_0 sadece sabit terimden oluşan modelin olabirlik değeri, L_1 elde edilen modelin olabirlik değeri olmak üzere

$$C = -2\log(L_0/L_1) = -2(\log L_0 - \log L_1) \quad (9)$$

olarak tanımlanan ölçüt (k-1) serbestlik derecesiyle Ki-kare dağılımı göstermektedir (Tatlıdil, 1992: 232).

Denklemin anlamlılığı için $C = 292,164$ olarak bulunmuştur. $\alpha = 0,05$ ve 5 serbestlik dereceli Ki-kare tablo değeri 11,07'dan daha büyük olduğundan denklem anlamlı bulunmuştur. Bu denklem için elde edilen sınıflandırma tablosu da Tablo 2'de verilmiştir.

Tablo 2. İleriye Doğru Değişken Seçme Tekniğine Göre Sınıflandırma Sonuçları

| Gözlemlenen | Kestirim | | Doğruluk Yüzdesi |
|-------------|-------------|-------------|------------------|
| | 0 (sis yok) | 1 (sis var) | |
| 0 (sis yok) | 734 | 14 | 98,13 |
| 1 (sis var) | 49 | 64 | 56,64 |
| | | Genel | 92,68 |

Tablo 2'deki sonuçlara göre elde edilen (8) nolu denklemin doğru sınıflandırma oranı %92,68 olup oldukça yüksek bir orandır. Dolayısıyla elde edilen bu oran da denklemin anlamlı olduğunun bir göstergesidir.

5. SONUÇ

Çalışmamızda, bağımlı değişkenin iki düzeyli, bağımsız değişkenler arasında da kategorik değişken(ler)in olduğu durumlarda gözlemlerin gruplara atanmasında bir ayrımsama modeli olarak kullanılan ve son yıllarda diskriminant analizine alternatif olarak geniş bir uygulama alanı bulan, Lojistik Regresyon Analizi kısaca incelenmiştir.

Bu çalışmada, özellikle sis hadisesinin yoğun olarak görüldüğü Kasım, Aralık ve Ocak ayları değerleri ve sis oluşumunun genel olarak başladığı ve sona erdiği sabah 4 ve 10 saatleri arasındaki veriler kullanılarak, lojistik regresyon denklemi tahmin edilmiştir. Bu amaçla, önce enbüyük olabilirlik katsayı kestirimleriyle ve ileri doğru değişken seçme tekniğiyle sis oluşumundaki değişkenlerin sıcaklık, nem, basınç, havanın kapalılık durumu ve rüzgarın hızı olduğu tespit edilmiştir. Lojistik regresyon analizine göre doğru sınıflandırma oranı %93 olarak bulunmuştur. Dolayısıyla elde edilen sonuçlar teoride beklenen sonuçları desteklemektedir. Elde edilen analiz sonuçlarına göre sis oluşumundaki en etkili faktörlerin havanın kapalılık durumu, nem ve basınç olduğu belirlenirken, rüzgarın hızı ve sıcaklığın sisin dağılmasında etkili olduğu sonucuna ulaşılmıştır. Bu da meteorolojide beklenen sonuçları desteklemektedir. (8) nolu denklem için elde edilen doğru sınıflama yüzdesi oldukça yüksektir. Ayrıca Ki-kare testi sonucuna göre de denklemin anlamlı olduğu tespit edilmiştir. Dolayısıyla lojistik regresyon için belirlenen (8) nolu denklemin en uygun ayrımsama denklemi olarak da kullanılabilceği ifade edilecektir. Gelecekteki bir zaman dilimi için kapalılık, nem, sıcaklık, rüzgar hızı ve basınç değişken değerlerinin elde edilen (8) nolu denklemde yerine yazılması suretiyle sis olup olmayacağı ya da sisin meydana gelme olasılıkları hesaplanabilecektir.

Özellikle hava alanlarında sis hadisesinden dolayı uçakların kalkamadığı veya rötarlı kalktığı bilinmektedir. Uçuş planları hazırlanırken meteorolojiden tahminler alınmaktadır. Bu yüzden yapılan tahminler oldukça önem arz etmektedir. Eskişehir'deki meteoroloji istasyonu çalışanlarının şimdiye kadar kullandıkları tahmin yöntemleriyle beraber, çalışmamızda geliştirilen ve meteorolojide kestirim amacıyla kullanılacak lojistik regresyon denklemi yardımıyla, sis tahminlerindeki doğruluk derecesinin artacağı düşünülmektedir. Dolayısıyla doğruluk derecesi yüksek sis tahminleri, Anadolu Üniversitesi Sivil Havacılık Yüksekokulu, Türk Hava Yolları ve 1. Ana Jet Üssünde uçuş yapan pilotların emniyetli uçuşları için oldukça yararlı olacaktır.

KAYNAKÇA

Abbott, R.D. (1985), Logistic Regression In Survival Analysis, American Journal of Epidemiology, 121, 465-471.

Agresti, A. (1990), Analysis Of Ordinal Categorical Data, John Wiley and Sons, New York.

Aktaş, C. ve Yılmaz V. (2001), “Eskişehir’de Lpg Kullanan Özel Araç Sürücülerinin Sınıflandırılmasında Lojistik Regresyon Analizi”, İstanbul Kent İçi Ulaşım Sempozyumu, İstanbul, 251-256.

Anderson, T.W., (2003), An Introduction to Multivariate Statistical Analysis, Third Edition, New Jersey: Wiley-Interscience.

Başarır, G. (1990), Çok Değişkenli Verilerde Ayrımsama Sorunu ve Lojistik Regresyon Analizi, Doktora Tezi (Yayımlanmamış).

Bircan, H. (2004), “Lojistik Regresyon Analizi: Tıp Verileri Üzerine Bir Uygulama”, Kocaeli Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, 2, 185-208.

Breslow, N.E. and Day, N.E. (1980), Statistical Methods In Cancer Research, The Analysis Of Case-Control Studies, International Agency Of Cancer, Lyon, France.

Cornfield, J. (1962), Joint Dependence Of The Risk Of Coronary Heart Disease On Serum Cholesterol And Systolic Blood Pressure: A Diskrimant Function Analysis, Federation Proceedings, 21, 58-61.

Coşkun, S., ve diğerleri (2004), “Lojistik Regresyon Analizinin İncelenmesi Ve Diş Hekimliğinde Bir Uygulaması”, Cumhuriyet Üniversitesi Diş Hekimliği Fakültesi Dergisi, Cilt:7, Sayı: 1, 42-50.

Fabbian, D., Dear, R. and Lellyett, S. (2007), “Application of Artificial Neural Network Forecasts to Predict Fog at Canberra International Airport”, Wea.Forecasting, 22, 372–381.

Gardside, P.S. and Glueck, C.J. (1995), “The Important Role Of Modifiable Dietary And Behaviour Characteristic In The Causation And Prevention Of Coronary Heart Disease Hospitalization And Mortality”, Journal of American college of Nutrition, 14, 71-79.

Lewis, D.M. (2004), Forecasting Advective Sea Fog With The Use Of Classification And Regression Tree Analyses For Kunsan Air Base., Master's Thesis, Air Force Institute of Technology Graduate School of Engineering and Management (AFIT/ENP).

Press, S.J. and Wilson, S., (1978), "Choosing Between Logistic Regression And Discriminant Analysis", *Journal of American Statistical Association*, 73, 364, 699-705.

Serin, U. (2006), Havacılıkta Yaklaşma İniş Kazalarının Analizi ve Alınması Gereken Tedbirler, Gazi Üniversitesi, Fen Bil. Ens., Yüksek Lisans Tezi, Ankara.

Stern, H. and Parkyn, K. (1999), "Predicting the Likelihood of Fog at Melbourne Airport", 8th Conference on Aviation, Range and Aerospace Meteorology, Amer. Meteor. Soc., Dallas.

Tatlıdil, H. (1992), Uygulamalı Çok Değişkenli İstatistiksel Analiz, Ankara:Engin Yayınları.

Tuncer, İ. (1991), Esenboğa Hava Limanının Sis Etüdü, T.C. Başbakanlık Devlet Meteoroloji İşleri Genel Müdürlüğü, Yüksek Lisans Tezi, Ankara.

Uğur, H. (1984), Yeşilköy Hava Meydanında Sis ve Özellikleri, T.C. Başbakanlık Devlet Meteoroloji İşleri Genel Müdürlüğü, Uzmanlık Tezi, Ankara.

Ünsal, A. ve Güler, H. (2005), "Türk Bankacılık Sektörünün Lojistik Regresyon Ve Diskriminant Analizi ile İncelenmesi", *VII. Ulusal Ekonometri ve İstatistik Sempozyumu*, İstanbul Ün.

Vupa, Ö. ve Çelikoğlu, C. (2006), "Model Building In Logistic Regression Models About Lung Cancer Data", *Anadolu Ü. Bilim ve Teknoloji Dergisi*, cilt:7, 1, 127-141.

Yağan, S., Çamalan G. ve Akgün, N. (1991), Esenboğa Hava Limanı Sis Etüdü ve Sis Dağıtma Yöntemleri, Çevre ve Orman Bakanlığı Devlet Meteoroloji İşleri Genel Müdürlüğü, Ankara.