

YEREL ALAN KARAKTERİSTİKLERİNE İLİŞKİN DOĞRUDAN VE DOLAYLI TAHMİNLERİN ELDE EDİLMESİ

Volkan SEVİNÇ*

ÖZET

Yerel alanlara ilişkin tahminler yapmak için geliştirilmiş özel yöntemlere ihtiyaç duyulur. Bu yöntemlerin her biri incelenen konu, veri türü, problemin yapısı gibi etkenlere göre değişiklik gösterir ve farklı uygulama alanlarına sahiptir. Bu çalışmada, yerel alan kavramının açıklaması yapıldıktan sonra, yerel alanlarda kullanılan doğrudan ve dolaylı tahmin ediciler, varsayımları, yöntemleri ve gereksinim duydukları veri türü ile anlatılmıştır. Daha sonra, bir örnek uygulama olarak sentetik tahmin yöntemi ile Muğla ili işsizlik oranı tahmini elde edilmiş ve elde edilen sonuçların yorumu yapılmıştır.

Anahtar Kelimeler: Yerel alan tahminleri, Alan araştırmaları, Lokal tahmin ediciler, Küçük alanlar.

1. GİRİŞ

Bir ülke içindeki çeşitli coğrafi alanlarla ilgili istatistiksel bilgilere her zaman gereksinim duyulur. Bu tür bilgiler, yerel yönetimler ve hükümetler tarafından bölgesel politikalar oluşturmak amacıyla veya resmi olmayan kuruluşlarca, yatırım, araştırma gibi nedenlerle kullanılmaktadır. Güncel ve geleceğe dönük planlar yapmak amacıyla da söz konusu verilere ihtiyaç duyulmaktadır. Yerel alan istatistiklerinin elde edilmesinde çeşitli zorluklar vardır. Bu tür veriler nüfus sayımları veya resmi kayıtlar kullanılmadan çoğu zaman elde edilemezler. Örnekleme çalışmaları geniş alanlar için oldukça güncel ve güvenilir veriler sağlasa da, yöntemleri nedeniyle yerel alanlarda başarılı olamamaktadır. Yerel alanlardan seçilen örneklerin büyüklükleri yetersiz kalmakta, hatta seçilen örnek alanlarda, örnek elemanlarının bulunmaması durumuyla karşılaşabilmektedir. Bu nedenle yerel alan tahminleri yapmak için özel yöntemler geliştirilmiştir. Yerel alan tahmin yöntemleri gerektirdikleri veri, varsayımlar ve uygulanış şekilleri bakımından çeşitlilikler gösterirler. Bu çalışmada, önce yerel alan kavramı ve türlerinin tanımı yapılacak, daha sonra yerel alan tahmininde kullanılan doğrudan ve dolaylı tahmin yöntemleri hakkında bilgi aktarılacaktır. Sonuç bölümünde ise dolaylı tahmin yöntemlerinden sentetik tahmin yöntemine ilişkin bir örnek uygulama çalışması sunulacaktır. Çalışma konusu Muğla İli işsizlik oranının sentetik tahmin yöntemi yardımıyla elde edilmesi şeklinde seçilmiştir.

2. YEREL ALAN KAVRAMI

Yerel alan kavramını açıklayabilmek için öncelikle alan ifadesinin tanımını yapmak gerekir. Alan, bir ana kitlenin herhangi bir alt bölümü ya da alt kümesi olarak tanımlanabilir (Purcell ve Kish, 1979). Diğer bir tanıma göre ise, alan, bir ana kitlenin, örnekleme yolu ile elde edilen ve kendisi için özel tahmine ihtiyaç duyulan bir alt kümesidir. Alanların çeşitli türleri vardır (Särndal, 1993):

*Yrd. Doç. Dr., Muğla Üniversitesi, Fen Fak., İstatistik Bölümü, e-posta: volkansevinc@yahoo.com

a) Planlı Alanlar: Bu alanlarda birbirinden ayrı örnekler, alan için planlanmış tasarlanmış ve seçilmiştir. Bunların birleşimi, örneğin bütünü oluşturur. Bu tür alanlara örnek olarak ana bölgeler ve diğer ana birimlerden oluşturulmuş katmanlar verilebilir.

b) Çapraz Sınıflandırılmış Alanlar: Çapraz sınıflandırılmış alanlar, çeşitli kriterlere göre çapraz olarak sınıflandırılmış, verilerdeki, her bir birim olarak tanımlanabilir. Yaş, cinsiyet, meslek ve eğitim sınıfları bu birimlere örnek gösterilebilirler.

c) Bölgeler: İlk iki türden daha az kullanılan bir alan türü de bölgelerdir. Bölgeler alan içindeki daha küçük parçalardır. Bölgelerin seçimi, örnekleme yapılırken dikkate alınmaz. Bölgeler, seçilen örneklerin içinde dağınık halde yer alırlar.

Alanların büyüklüğü, yöntem seçimini de etkiler bu nedenle alanların büyüklüklerine göre de sınıflandırılmalarında fayda vardır. Bu sınıflandırma da aşağıdaki gibi yapılabilir (Purcell ve Kish, 1980), (Kish, 1987):

i. Ana Alanlar: Ana kitle büyüklüğünün yaklaşık 1/10'unu veya daha fazlasını kapsayan alanlardır. Bu alan türlerine örnek olarak: Ana bölgeler, 10 yaş grubu çocukları veya meslek grupları gibi daha kategorik sınıflar verilebilir.

ii. Dar Alanlar: Bu alanlar da ana kitle büyüklüğünün 1/10 ve 1/100 arasındaki kısmını kapsayan alanlardır. Eyalet nüfusları, herhangi bir yaş grubunun nüfusu, meslek-eğitim gibi iki yönlü sınıflandırmalarda bir hücreye düşen nüfus, ya da işsiz sayısı gibi kategorilerin oluşturduğu nüfus, bu alanlara örnek olarak verilebilir.

iii. Mini Alanlar: Bu alanlar ana kitle büyüklüğünün 1/100 ve 1/10000 arasındaki kısmını barındırırlar. Örneğin: ilçe nüfusları, ya da yaş-meslek-eğitim gibi üç yönlü sınıflandırmalar.

iv. Seyrek Alanlar: Seyrek alanlar ise ana kitlenin 1/10000 den daha az kısmını barındırırlar. Bunlara örnek olarak da, çeşitli etnik gruplara göre sınıflandırılan sağlık servis alanları veya küçük yerleşim birimlerine göre sınıflandırılmış belirli kronik hastalıkları olan bireyler verilebilir.

Bu sınıfların sınırları kesin değildir. Bu sınırlar örnek ve ana kitle büyüklüğüne, değişken ve istatistiklere, ölçüm ve kararlara göre değişebilir. Ancak, söz konusu sınırlar, sınıflar arasındaki uygulama farklılıklarını göstermek açısından yararlıdırlar.

Alan kavramı, planlı alanlar kavramı ile sınırlandırılıp, United Nations (1950) planlı alanlar tanımı ile son bir kez daha tanımlanacak olursa: Hakkında kesin bir rakamsal bilgi sağlamak için araştırma yapılması planlanan herhangi bir alt alan, bir çalışma alanı olarak adlandırılabilir.

3. YEREL ALANLARDA DOĞRUDAN VE DOLAYLI TAHMİN YÖNTEMLERİ

Bu bölümde, yerel alanlara ait çeşitli karakteristiklerin tahmininde kullanılan yöntemler, yöntemleri ve kullandıkları veri türü ile birlikte açıklanacaktır.

Araştırma örneklemede amaç, bilinmeyen bir ana kitle karakteristiğini tahmin etmektir. Örnekler sadece ana kitle karakteristiğini tahmin etmek için değil, alt kitle ya da alanlar için tahminler yapmak amacı ile de kullanılırlar. Bu amaçla kullanılan tahmin ediciler, eğer ilgilenilen değişken değerlerini sadece alan içindeki birimlerden ve zaman aralığından alıp kullanıyorsa, bunlar doğrudan tahmin ediciler olarak adlandırılırlar. Dolaylı tahmin ediciler ise, alandan ve alan dışındaki zaman aralığından alınan değişken değerlerini kullanırlar.

Doğrudan tahmin ediciler için, çeşitli kaynaklarda kullanılan farklı isimler bulunmaktadır. Royall (1973), bu tahmin edicileri, 'doğrudan', Gonzalez (1973), 'yansız', Holt, Smith vd. (1979), 'standart', Gonzalez (1979), 'geçerli', Kalton (1987), 'örnek - tabanlı' olarak adlandırmışlardır. Doğrudan tahmin ediciler, geleneksel tahmin yöntemleri ile elde edilen tahmin edicilerdir. Bu tür tahmin ediciler, sadece ilgilenilen alandan alınan veriye dayanırlar. Bazen bu verilere nüfus sayımı veya resmi kayıtlardan elde edilen yardımcı veriler eklenebilir. Doğrudan tahmin ediciler, yerel alan tahminlerinde genellikle yetersiz ve küçük örnek hacmi nedeniyle büyük varyanslıdırlar.

Dolaylı tahmin ediciler de çeşitli adlar altında anılmaktadır. En çok kullanılan terim, yerel alan tahmin edicileridir. Çünkü bu tahmin edicilerin birçoğu, coğrafi alanlar için tahminler yapar. Yerel alan tahmin edicileri terimindeki yerel sözcüğü iyi tanımlanmalıdır. Yerel sözcüğünden anlaşılması gereken, alanın ya da alan içindeki kitlenin küçüklüğü, örnek gözlem sayısının azlığı ve aranılan standart doğrudan tahmin edicilerin büyük varyanslı olması durumudur. Aynı şekilde alan kelimesinin de kesin sınırlayıcılığı yoktur. Bu yöntemler, isteğe bağlı herhangi bir bölgeye uygulanabilir. Bu bölgenin de her zaman coğrafi bir bölge olarak tanımlanması gerekmez. Bu tahmin edicileri tanımlamak için, Ericksen (1978), 'yerel alan', Purcell and Kish (1979), 'küçük alan', Laake (1979), 'alt alan', Holt, Smith vd. (1979), 'küçük alt gruplar', Dalenius (1987), 'dolaylı', Särndal (1984), 'model-bağımlı' terimlerini kullanmışlardır. Sentetik tahmin ediciler terimi de bu dolaylı tür tahmin edicilerin tanımı için NIDA (Ulusal Uyuşturucu Kullanımı Araştırma Enstitüsü) (1979), ve özel bir dolaylı tahmin edicinin tanımı için NCHS (Ulusal Sağlık İstatistikleri Merkezi) (1968) tarafından kullanılmıştır.

Araştırmacılar, dolaylı tahmin edicileri bazen, model temelli tahmin ediciler olarak da adlandırılırlar. Fakat bu terim seyrek kullanılır. Dolaylı tahmin edicilerin üç türü vardır:

a) Alan Dolaylı Tahmin Edici: Bu tür tahmin ediciler, ilgilenilen değişken değerlerini başka bir alandan alıp kullanırlar ama başka bir zaman aralığından almazlar.

b) Zaman Dolaylı Tahmin Edici: Bu tür tahmin ediciler ise, ilgilenilen değişken değerlerini başka bir zaman aralığından alırlar ancak, başka bir alandan alıp kullanmazlar.

c) Zaman ve Alan Dolaylı Tahmin Edici: Bu üçüncü tür dolaylı tahmin ediciler ise, ilgilenilen değişken değerlerini hem başka bir alandan, hem de başka bir zaman aralığından sağlarlar.

Bütün dolaylı tahmin ediciler, yapılması planlanan bir araştırma için eşit derecede uygun değildir. Örneğin, eğer araştırmanın amacı, verilen bir zaman aralığı içinde alanlar arasında karşılaştırma yapmak ise, alan dolaylı tahmin ediciler yerine, zaman dolaylı tahmin edicilerden birini seçmek daha uygun olur. Alan dolaylı tahmin ediciler, bazı model parametrelerine göre ilgilenilen değişkenlerin ortalamasının, alanlar arasında aynı olduğunu varsayan modellere dayanırlar. Fakat zaman dolaylı tahmin ediciler, bu değişkenlerin, zaman aralıkları boyunca aynı olduğunu varsayan modellere dayanırlar. Uygulamalarda alan ve zaman dolaylı tahmin edicilerin performansı, elde ne kadar bilgi olduğuna ve bu bilgiyle kullanılacak modelin gerçek ihtiyaca ne kadar cevap verebileceğine bağlıdır. Dolaylı tahmin yöntemlerinin belli başlı çeşitleri aşağıda verilmiştir (Rao, 2003).

3.1 Semptomatik Kayıt Yöntemleri

Bu yöntemler, kayıt sistemleri ile eski sayım verileri arasındaki istatistiksel ilişkilere dayanır. Standart semptomatik kayıt sistemleri (S.K.S.), yerel alan tahmin yöntemlerinin en eskisidir. Çeşitli kayıtları kullanan bu tür yöntemlerin çoğu, Birleşik Devletler Nüfus Sayımı Kurumu (U.S. Bureau of the Census) (1975a) tarafından önerilmiş ve test edilmiştir. Bu yöntemlerin birçoğu, yerel alandaki nüfusa ilişkin tahminlerle ilgilenir. Bu yöntemlerin çeşitleri takip eden alt başlıklarda anlatılmıştır.

3.1.1 Nüfus sayımı bileşenleri yöntemleri

Bu yöntemlerin iki çeşidi vardır ve ikisi de okula giden çocuklar arasında yapılan bir araştırmadan elde edilen göç oranının, toplam nüfus içindeki göç oranına eşit olacağı varsayımından hareket eder. Yöntemlerin farklılaştığı nokta ise, okul çocukları arasındaki göç oranının tahmininin elde edilme biçimidir. Birleşik Devletler Nüfus Sayımı Kurumu (1947), bu yöntemleri genellikle ülke içi göç miktarını hesaplamak için kullanmıştır.

3.1.2 Yaşamsal oranlar yöntemi

Yerel alan nüfus büyüklüklerinin tahmini için kullanılan bir başka yöntem, yaşamsal oranlar (vital rates) yöntemidir. İlk tanımı Bogue (1950) tarafından verilmiştir. Bu yaklaşım, geniş alan ölüm ve doğum oranlarını ve nüfus-oran işlemini kullanarak, gereken yerel alan nüfus tahminleri yapar. Bu yöntemin dayandığı varsayım, geniş alanlardaki ölüm ve doğum oranlarının yerel alanlarla aynı oranda değiştiğidir. Yöntemin uygulanması aşağıdaki gibidir:

- a) Temel alınan bir yıla ait nüfus sayımı verileri kullanılarak, ilgili yerel alan ve ana alan için kaba ölüm ve doğum oranları hesaplanır.
- b) Yerel alana ait kaba ölüm oranı, ana alana ait kaba ölüm oranına bölünür. Aynı işlem kaba doğum oranı için de tekrarlanır. Ölüm oranı;

$$\hat{x}_d = \frac{r_d}{R_d} \quad (1)$$

olmak üzere, burada r_d , yerel alan kaba ölüm oranı ve R_d , ana alan kaba ölüm oranıdır. Aynı şekilde doğum oranı;

$$\hat{x}_b = \frac{r_b}{R_b} \quad (2)$$

olmak üzere, burada, r_b , yerel alan kaba doğum oranı ve R_b , ana alan kaba doğum oranıdır.

c) Ana alan içerisinde yer alan çeşitli yerel alanlardaki kaba ölüm ve doğum oranlarının yakın geçmişte gösterdiği değişim, aynı oranların ana alan içerisinde gösterdikleri eğilim ile karşılaştırılır. Eğer herhangi bir yerel alan grubu için zamana bağlı olarak kesin bir oran değişikliği gözleniyorsa, tahmin yapılacak yıl için düzeltilmiş oranları elde etmek amacıyla, temel alınan yılın oranlarına uygulanmak üzere düzeltme faktörleri belirlenir.

d) Nüfus tahmini bulunmak istenen yıl için, ana alan içinde kaba ölüm ve doğum oranları hesaplanır. (Ana alan, kendisi için daha önce düzenli tahminler yapılan bir eyalet veya diğer bir birim olarak seçilmelidir.)

e) “d” kısmında ana alan için hesaplanmış olan ölüm kaba oranları, “c” kısmında yerel alana ait düzeltilmiş ölüm oranı ile çarpılarak, yerel alan için bir kaba ölüm oranı tahmini elde edilir. Aynı işlem doğum oranları için tekrarlanır. Ana alan için kaba ölüm veya doğum oranı,

$$\hat{x} = RF \quad (3)$$

şeklinde tanımlanır. Burada R , ana alanda istenen yıldaki kaba ölüm veya doğum oranı, F , “c” kısmında belirlenen ilgili düzeltme faktörüdür.

f) Yerel alanlar için tahmin edilmiş olan kaba ölüm oranları, yıl içerisinde kaydedilmiş ilgili yerel alana ait toplam ölüm sayısına bölünür. Aynı işlem doğum için tekrarlanır. Bu işlemler sonucunda yerel alan nüfusuna ait, ölüm ve doğum verilerine dayalı iki tahmin elde edilir. Ölüm ve doğum oranlarına dayalı tahmin;

$$\hat{y}_h = \frac{r_h}{N_h} \quad (4)$$

şeklinde tanımlanır. Burada r_h , her bir h yerel alanı için tahmin edilmiş kaba ölüm veya doğum oranı, N_h , yerel alanda ilgili yıl içinde kaydedilmiş ölüm veya doğum sayısıdır.

g) Tek bir ana kütle tahmini oluşturmak için, “f” kısmında hesaplanan ölüm ve doğum verisine dayalı iki tahminin ortalaması alınır (Rao, 2003).

3.1.3 Karma yöntem

Bu yöntem, Bogue ve Duncan (1959) tarafından yaşamsal oranlar (vital rates) yöntemine bir alternatif olarak önerilmiştir. Yöntem, yerel alan nüfusunun ayrı yaş gruplarına bölünmesi ve bu grupların her biri için ayrı nüfus tahminlerinin elde edilmesine dayanır. Tahminleri bulmak için, her bir grup için uygun yöntem ve veri kullanılır. Ayrı ayrı elde edilen tahminler bir araya getirilir ve son tahmin elde edilir.

3.1.4 Konut birimi yöntemi

Konut birimi (housing unit) yöntemi, yerel alandaki konut sayısının güncel tahminini kullanır ve her konuttaki ortalama insan sayısını hesaba katar. Dayandığı varsayım, konut sayısındaki değişimin, nüfustaki değişimi yansıtacağıdır. Birleşik Devletler Nüfus Sayımı Kurumu (1969) bu yöntemi tartışmış ve düzenleyerek bir karma yöntem oluşturmuştur. Rives (1976), konut birimi yöntemini geliştiren bir çalışma sunmuştur. Swanson (2009), nüfus tahmini yapmak amacıyla konut birimi yönteminin bir türünü önermiş ve bu yöntemin iki ana unsuru olan boş konut oranı ve konut başına düşen kişi sayısının elde edilme yollarını açıklamıştır.

3.1.5 Resmi veri kayıtları yöntemi

Bu yöntem, semptomatik kayıt yöntemlerinin en yenisidir. Birleşik Devletler Nüfus Sayımı Kurumu (1975b) tarafından kullanılmıştır. Yöntem, federal gelir paylaşımı programı için gerekli olan bütün gelir paylaşım alanlarındaki nüfusun güncellenmesi için geliştirilmiştir. Temelde nüfus sayımı bileşenleri ile aynıdır. Ayrıldıkları nokta, göç sayısının gelir servisi tarafından saptanmasıdır. Bu yöntemde, veri sınıflarından çok kişisel kayıtların kullanılması, çok küçük alanlar için bile tahmin yapılmasına olanak sağlar. Ancak, yöntem federal kullanımına sınırlıdır. Bu yöntemin detaylı bir tanımı Starsinic (1974) tarafından verilmiştir.

3.2 Sentetik Tahmin Yöntemi

Sentetik tahmin edicilerin bulunmasında kullanılan varsayım, küçük bir alanın kendisini kapsayan göreceli olarak daha büyük bir alanla benzerlik gösterdiğidir. Sentetik tahminler, nüfus sayımı ve resmi kayıtlardan yardımcı veri olarak da elde edilebilir. Varyansları genelde küçüktür. Ancak yöntemin dayandığı varsayımın ilgilene alan için geçerli olmadığı durumlarda, sentetik tahmin ediciler oldukça yanlı ve tutarsız olabilirler. Bütün dolaylı tahmin ediciler gibi, sentetik tahmin ediciler de alan dolaylı, zaman dolaylı, ya da zaman ve alan dolaylı olabilir. Sentetik isminin kullanılmasının nedeni, sentetik tahminlerin doğrudan araştırma sonuçlarından elde edilmemesidir. İstatistikte sentetik terimi, tahminlerin güvenilirliğini arttırmak için öbür benzer alanlardan veri ödünç alan yöntem anlamında kullanılmaktadır. Yerel alan h içindeki x özelliği için sentetik tahmin edici aşağıdaki biçimde ifade edilir.

$$\hat{x}_h = \sum_g \hat{x}_{hg} = \sum_g (Y_{hg} / Y_{.g}) x'_{.g} \quad (5)$$

Burada,

Y_{hg} : h . yerel alan ve g . alt grup için ilgili değişkene ait değer,
 $x'_{.g}$: Büyük alan içindeki g alt grubu için x özelliğinin örneğe dayalı tahminidir.

Burada noktalı indisler, ilgili değişkene ait değerlerin toplanmış halini göstermektedir. İlgili değişken Y , kişi sayısı toplamı olarak alınır. Bu nedenle,

$$Y_{hg} = N_{hg} \quad (6)$$

eşitliği yardımıyla (5)'de verilen sentetik tahmin edici aşağıdaki eşitliğe dönüşür:

$$\hat{x}_h = \sum_g (N_{hg} / N_{.g}) x'_{.g} \quad (7)$$

3.2.1 Regresyon ayarlı sentetik yöntem

Bu yöntem, Levy (1971) tarafından önerilmiştir. Yöntem, sentetik tahmin ile birlikte yerel düzeydeki semptomatik bilgiyi kullanır. Semptomatik veri ve sentetik veri arasındaki regresyon ilişkisine dayalı model,

$$x_h^* = \alpha + \beta Y_h + \varepsilon_h \quad (9)$$

eşitliği ile verilir. Burada,

Y_h : Semptomatik değişkenin değeri,
 α ve β : Tahmin edilmesi gereken regresyon katsayıları,
 ε_h : Rasgele hata terimleridir.

Modelde yer alan x_h^* değişkeni, sentetik tahmin ediciye dayalı olarak ise,

$$x_h^* = \{(X_h - \hat{x}_h) / \hat{x}_h\} 100 \quad (10)$$

şeklinde elde edilir. Burada,

\hat{x}_h : Sentetik tahmin,
 X_h : Gerçek değerdir.

α 'nın tahmin edicisi $\hat{\alpha}$, ve β 'nin tahmin edicisi $\hat{\beta}$ bulunursa ve hata terimi ε_h çıkarılırsa, x_h^* 'nin bir tahmin edicisi olan \hat{x}_h^* , sentetik ve semptomatik veriye dayanan aşağıdaki eşitlik ile doğrudan elde edilebilir:

$$\hat{x}_h^* = \hat{x}_h \{(\hat{\alpha} + \hat{\beta} Y_h) / 100 + 1\} \quad (11)$$

x_h^* , bilinmeyen gerçek değer X_h 'nin bir fonksiyonu olduğundan, doğrusal katsayıları bulmak için farklı bir yöntem kullanılmalıdır. Özetle böyle bir yöntem, yerel alanları tabaka oluşturacak biçimde düzenledikten sonra, en küçük kareler yoluyla α ve β 'nın tahmin değerlerini bulmalıdır. Bu tabakalar oluşturulurken, X_h 'nin yansız tahminlerini verebilecek şekilde oluşturulmalıdırlar.

3.2.2 Karma sentetik-regresyon yöntemi

Nicholls (1977) tarafından ilk olarak geliştirilen bu yöntem, temelde örnek-regresyon yöntemi ile aynı olmakla birlikte, sentetik tahmini ek bir bağımsız değişken olarak eklemesi yönünden farklıdır. Semptomatik değişkenlerin uygun seçilmesiyle, yöntem sentetik tahminler ve standart örnek-regresyon tahminlerinde gelişmeye yol açar.

Gonzalez ve Hoza (1978) bu yöntemle, yerel alandaki işsizlik miktarının tahminini konu alan bir çalışma yapmışlardır. Yöntem, sentetik tahmini, yukarıda da belirtildiği gibi, regresyon eşitliğinde bağımsız değişkenlerden herhangi biri gibi kullanmaktadır. Bağımlı değişken olarak her bir ana örnek ünitelerindeki işsizliğin güncel nüfus sayımı ile elde edilmiş tahminleri göz önüne alınmıştır. Varyansı azaltmak için ek bağımsız değişkenler de kullanılmıştır. Regresyon hata kareler ortalaması (HKO) ise aşağıdaki eşitlikle hesaplanmıştır:

$$HKO = E \left[\frac{(Y_o - \hat{Y})(Y_o - \hat{Y})}{n} \right] - \frac{(n - 2p - 2)\sigma_w^2}{n} \quad (12)$$

Burada,

- Y_o : Y 'nin gözlem değerleri,
- n : Gözlem sayısı,
- p : Bağımsız değişken sayısı,
- σ_w^2 : Ana örnek ünitesi içindeki hatadır.

σ_w^2 'nin tahmini Ericksen (1971) tarafından verilmektedir. Sentetik tahmin edicilerin yanlı olabileme tehlikesinden dolayı, değişkenlik ve yanlılıklarını yansıtabilecek bir ölçü ya da kritere ihtiyaç vardır. Bu tür bir ölçüm olarak, hata kareler ortalaması karekökü (HKOK) alınabilir.

Gonzalez ve Hoza (1978)'nin araştırmasında, en geniş alan, bir coğrafi bölgedir. Yerel alan olarak iller seçilmiş ve yanlılık ölçütü olarak göreceli hata kareler ortalaması karekökü (GHKOK) alınmıştır. Ölçüt,

$$GHKOK = \frac{HKOK}{U_m} \quad (13)$$

şeklinde tanımlanır. Burada, U_m , işsizlik aralığının orta noktasıdır. GHKOK'in değeri çok yüksek ve düşük işsizlik değerlerinde artmaktadır. Aynı çalışmada aykırı değerlerin saptanarak ayıklanmasının, GHKOK değerinin azalmasına yol açtığı gösterilmiştir.

4. MUĞLA İLİ İŞSİZLİK ORANININ SENTETİK TAHMİNİ

Bu bölümde kesim 3.2’de anlatılan sentetik tahmin yöntemine ilişkin bir uygulama çalışması sunulacaktır. Uygulama sahası olarak yerel bir alan niteliği taşıyan Muğla İli seçilmiştir. Uygulama konusu olarak ekonomik göstergeler arasında önemli bir yer tutan işsizlik konusu ele alınmıştır. Muğla İli’nin içerisindeki ilçeler küçük (alt) alanlar olarak nitelendirilebilirler. Muğla İli ise bu küçük alanları kapsayan görece daha büyük bir yerel alan olarak ele alınabilir. Yerel alan Muğla İli’nin işsizlik oranını, alt grup olarak ele alınan, bazı ilçelerinin işsizlik oranından tahmin etmek mümkündür. Uygulamalarda, ilçe sayısının ya da alt alanların belirlenmesi aşamasında, alt alanlar için mevcut veri miktarı belirleyici olmaktadır. Buna göre, Muğla İli içinde yer alan, merkez ilçe ile birlikte onbir ilçeden sekiz ilçe, rasgele örnekleme ile seçilmiştir. Bu ilçeler: Merkez, Bodrum, Datça, Dalaman, Marmaris, Milas, Kavaklıdere, Ortaca ilçeleridir.

İşsizlik oranını, ilçeleri merkez ve kırsal kesim olarak ayırarak, aynı zamanda cinsiyete dayalı olarak da hesaplamak olasıdır. Uygulamada kullanılan veriler, Devlet İstatistik Enstitüsü (D.İ.E.) tarafından yapılan 2000 genel nüfus sayımında, Muğla İli için yayınlanmış verilerden derlenmiştir. Seçilmiş ilçelerin merkez ve kırsal kesim nüfusları ile işsizlik oranları, kadın ve erkek nüfusa dayalı olarak sırasıyla Tablo 1 ve Tablo 2’de verilmiştir.

Tablo 1. İlçelerin 2000 yılına ait merkez ve kırsal kesim erkek nüfusları ile işsizlik oranları

İlçeler	İşsizlik Oranı x'_{g}	İlçe Nüfusu N_{1g}	Merkezleri	Kırsal Kesim Nüfusu N_{2g}
1- Merkez	0.09	22987		20220
2- Bodrum	0.09	18601		36910
3- Datça	0.12	4400		2967
4- Dalaman	0.12	9183		5263
5- Marmaris	0.10	16645		29971
6- Milas	0.09	19164		38413
7- Kavaklıdere	0.09	1917		4485
8- Ortaca	0.15	8538		9821

Tablo 2. İlçelerin 2000 yılına ait merkez ve kırsal kesim kadın nüfusları ile işsizlik oranları

İlçeler	İşsizlik Oranı x'_{g}	İlçe Nüfusu N_{1g}	Merkezleri	Kırsal Kesim Nüfusu N_{2g}
1- Merkez	0.19	20858		19446
2- Bodrum	0.18	13626		28698
3- Datça	0.18	3708		2839
4- Dalaman	0.25	8424		5278
5- Marmaris	0.17	12015		20671
6- Milas	0.22	18899		36332
7- Kavaklıdere	0.31	1515		4631
8- Ortaca	0.30	8385		8926

Yerel alan Muğla İli içindeki işsizlik oran için sentetik tahmin değerleri (7) eşitliği kullanılarak, ilçe merkezleri ve kırsal kesim için, cinsiyete göre ayrı ayrı hesaplanabilir. Sonuçlar aşağıda elde edilmiş ve elde edilen sonuçların özeti Tablo 3’de verilmiştir. Buna göre,

Muğla İli, ilçe merkezleri erkek işsizlik oranı,

$$\hat{x}_1 = \sum_g (N_{1g} / N_{.g}) x'_{.g} = \{[(0.09)(22987) + \dots + (0.15)(8538)] / 249485\} = 0.04$$

Muğla İli, kırsal kesim erkek işsizlik oranı,

$$\hat{x}_2 = \sum_g (N_{2g} / N_{.g}) x'_{.g} = \{[(0.09)(20220) + \dots + (0.15)(9821)] / 249485\} = 0.06$$

Muğla İli, ilçe merkezleri kadın işsizlik oranı,

$$\hat{x}_1 = \sum_g (N_{1g} / N_{.g}) x'_{.g} = \{[(0.19)(20858) + \dots + (0.30)(8385)] / 249485\} = 0.09$$

Muğla İli, kırsal kesim kadın işsizlik oranı,

$$\hat{x}_2 = \sum_g (N_{2g} / N_{.g}) x'_{.g} = \{[(0.19)(19446) + \dots + (0.30)(8926)] / 249485\} = 0.12$$

Tablo 3. Muğla İli’nde ilçe merkezleri ve kırsal kesim için, cinsiyete göre sentetik tahmine dayalı 2000 yılı işsizlik oranları

Muğla İli	İlçe Merkezleri	Kırsal Kesim
Erkek işsizlik oranı	0.04	0.06
Kadın işsizlik oranı	0.09	0.12

Elde edilen tahmin değerlerinin yanlışlıklarının belirlenmesi için (13) eşitliği ile verilen GHKOK değerleri hesaplanmış ve sonuçlar Tablo 4’de verilmiştir.

Tablo 4. Muğla İli’nde ilçe merkezleri ve kırsal kesim için, cinsiyete göre sentetik tahmine dayalı 2000 yılı işsizlik oranlarına ait yanlışlık değerleri

Muğla İli	İlçe Merkezleri	Kırsal Kesim
Erkek işsizlik oranı	0.65	0.48
Kadın işsizlik oranı	0.25	0.23

5. SONUÇ VE YORUMLAR

Yerel alan tahmin yöntemleri, genel olarak doğrudan ve dolaylı olarak iki sınıfta toplanmaktadır. Doğrudan tahmin yöntemleri, küçük örnek hacmi yüzünden büyük varyansa sahip olmaktadır. Yerel alan tahminlerinde dolaylı tahmin yöntemleri, doğrudan tahmin yöntemlerine göre daha fazla tercih edilmektedir. Sentetik tahmin dolaylı tahmin yöntemlerinden biridir. Sentetik tahmin yönteminin dayandığı varsayım, küçük bir alanın kendisini kapsayan göreceli olarak daha büyük bir alana benzer olduğudur. Sentetik tahmin yöntemi, nüfus sayımı ve resmi kayıtlardan yardımcı veri

olarak da kullanılabilir. Varyansları genelde küçüktür. Ancak, bu yöntem için ileri sürülen varsayımın ilgililenen alan için geçerli olmadığı durumlarda, sentetik tahmin ediciler oldukça yanlış ve tutarsız olabilirler.

Muğla İli'ndeki işsizlik oranının sentetik tahmin yardımıyla elde edilmesi çalışmasında erkek nüfus için işsizlik oranı, ilçe merkezlerinde 0.04 değeri ile, 0.06 değerine sahip kırsal kesime göre daha düşük oranda ortaya çıkmaktadır. Aynı durum kadın nüfus için de geçerlidir. 0.09 olarak tahmin edilen ilçe merkezleri kadın işsizlik oranı, 0.12 olan kırsal kesim oranından daha düşük değerdedir. Bu durumda Muğla İli kırsalında, hem erkek hem kadın nüfus için işsizlik oranının, ilçe merkezlerinden daha yüksek değerde olduğu sonucuna varılabilir. Genel olarak, gerek ilçe merkezleri gerek kırsal kesimde, kadın nüfus içerisindeki işsizlik oranının, erkek nüfusa ait işsizlik oranından daha yüksek değerde olduğu gözlenmektedir. Bu çalışmanın devamı olarak bu farkın nedenleri araştırılabilir.

6. KAYNAKLAR

Bogue, D. J., 1950. A Technique for Making Extensive Population Estimates. Journal of the American Statistical Association, 45, 149-163.

Bogue, D. J., Duncan, B. D., 1959. A Composite Method for Estimating Postcensal Population of Small Areas By Age, Sex and Colour, National Office of Vital Statistics, Vital-Statistics-Special Report, 47, No.6.

Dalenius, T., 1987. Panel Discussion in Small Area Statistics, John Wiley and Sons, New York.

Ericksen, E. P., 1971. A Method for Combining Sample Survey Data and Symptomatic Indicators to Obtain Estimates for Local Areas, Ph. D. Thesis, University of Michigan, Ann Arbor, Michigan.

Ericksen, E. P., 1978. A Regression Method for Estimation Population Changes for Local Areas, Journal of the American Statistical Association, 69, 867-875.

Gonzalez, M. E., 1973. Use and Evaluation of Synthetic Estimates. Proceedings at the Social Statistics Section, American Statistical Association, 87, 533-540.

Gonzalez, M. E., 1979. Case Studies on the Use and Accuracy of Synthetic Estimates: Unemployment and Housing Applications in Synthetic Estimates for Small Areas National Institute on Drug Abuse, Research Monograph 24, U.S. Government Printing Office, Washington D.C.

Gonzalez, M. E., Hoza, C., 1978. Small Area Estimation with Application to Unemployment and Housing Estimates. Journal of the American Statistical Association, 73, 7-15.

Holt, D., Smith, T. M. F., Tomberlin, T.J., 1979. A Model Based Approach to Estimation for Small Subgroups of a Population. Journal of the American Statistical Association, 74, 405-410.

Kalton, G., 1987. Panel Discussion in Small Area Statistics. John Wiley and Sons, New York.

Kish, L., 1987. Statistical Design for Research. John Wiley and Sons, New York, 1987.
Laake, P., 1979. A Prediction Approach to Subdomain Estimation in Finite Populations. Journal of the American Statistical Association, 74, 355-358.

Levy, P.S., 1971. The Use of Mortality Data in Evaluating Synthetic Estimates. 1971 Proceedings of the Social Statistics Section, American Statistical Association, 328-33.

NCHS, 1968. Synthetic State Estimates of Disability PHS Publication No. 1759, U.S. Government Printing Office, Washington D.C.

NIDA, 1979. Synthetic Estimates for Small Areas Research Monograph 24, U.S. Government Printing Office, Washington D.C.

Nicholls, A., 1977. A Regression Approach to Small Area Estimation. Australian Bureau of Statistics, Canberra, Australia, March Mimeographed.

Purcell, N. J., Kish, L., 1979. Estimation for Small Domains. Biometrics, 35, June, 365-384.

Purcell, N. J., Kish, L., 1980. Postcensal Estimates for Local Areas or Domains. International Statistical Review, 48, 3-18.

Rao, J. N. K., 2003. Small Area Estimation. John Wiley and Sons, New Jersey.

Rives, N. W., 1976. A Modified Housing Unit Method for Small Area population Estimation. Proceedings of the Social Statistics Section, American Statistical Association, 717-720.

Royall, R. A., 1973. Discussion of Papers by Gonzalez and Ericksen. Proceedings of the Social Statistics Section, American Statistical Association, 42-43.

Särndal, C. E., 1984. Design-Consistent versus Model-Dependent Estimation for Small Domains, Journal of the American Statistical Association, 79, 624-631.

Särndal, C. E., 1993. Small Area Statistics and Survey Designs, International Scientific Conference. Volume II. Contributed Papers and Panel Discussion. Central Statistical Office, Warsaw.

Starsinic, D. E., 1974. Development of Population Estimates for Revenue Sharing Areas, Census Tract Papers. Series GE 40, No.10, U.S. Bureau of the Census, U.S. Government Printing Office, Washington D.C.

Swanson, D. A., 2009. The Methods and Materials Used to Generate Two Key Elements of the Housing Unit Method of Population Estimation: Vacancy Rates (VR) and Persons Per Household (PPH). Working Paper 09-03, Blakely Center for Sustainable Suburban Development, University of California Riverside.

U.S Bureau of The Census, 1947. Population-Special Reports, Series P-47, No. 4, U.S. Government Printing Office, Washington D.C.

U.S. Bureau of The Census, 1969. Estimates of the Population of Counties and Metropolitan Areas, July 1, 1966: A Summary Report, Current Population Reports, Series P-25, No. 427, U.S. Government Printing Office, Washington D.C.

U.S. Bureau of The Census, 1975a. Population Estimates and Projection, Current Population Reports, Series P-25, No. 580, U.S. Government Printing Office, Washington D.C.

U.S. Bureau of The Census, 1975b. The Methods and Materials of Demography, Third Printing, U.S. Government Printing Office, Washington D.C.

United Nations, 1950. The Preparation of Sampling Survey Reports, Statistical Papers, Series C, No: 1.

OBTAINING THE DIRECT AND INDIRECT ESTIMATIONS FOR LOCAL AREA CHARACTERISTICS

ABSTRACT

Special methods are needed to make estimations about local areas. These methods differ from each other according to some factors such as the subject, type of data, and the structure of the problem. They also have different areas of application. In this paper, after explaining the local area concept, the direct and indirect estimators used in local areas are given with their assumptions, methods, and the type of data that they require. Later as an application, the unemployment rate in Muğla province is estimated through the synthetic estimation and the results obtained are interpreted.

Keywords: Local area estimates, Survey methods, Local estimators, Small areas.