SİTE İÇERİSİNDEKİ DAİRE FİYATLARINI ETKİLEYEN UNSURLARIN SEMİPARAMETRİK REGRESYON ANALİZİ İLE BELİRLENMESİ

Münevver TURANLI[57*], Seda BAĞDATLI KALKAN[58*]

ÖZET

Günümüzde bilgisayar ve teknolojinin gelişmesi nedeniyle, sayısal analizlerde kullanılan istatistiksel tekniklerde gelişmeler meydana gelmiştir. Bu tekniklerden birisi de parametrik ve parametrik olmayan regresyon modellerinin kullanılanamadığı durumlarda; analizlerin yapılabilmesi için geliştirilen semiparametrik regresyon modelleridir.

Semiparametrik regresyon modeli ile ilgili ilk çalışmalar 1985 yılında başlamış ve o yıldan günümüze kadar yapılan çeşitli çalışmalarla geliştirilmiş ve çeşitli alanlarda uygulanmıştır.

Bu çalışmada da önce semiparametrik regresyon modeli teorik olarak anlatılmaya çalışılmış ve daha sonra günümüzde inşaat sektöründe çok önemli bir yeri olan siteler ve bu siteler içerisindeki dairelerin fiyatını etkileyen unsurlar semiparametrik resresyon analizi ile incelenmiştir.

Anahtar Kelimeler: Semiparametrik regresyon, daire fiyatları

DETERMINING THE EFFECTIVE FACTORS OF APARTMENT PRICES IN HOUSING ESTATE WITH SEMIPARAMETRIC REGRESSION

ABSTRACT

At the president day, due to the development of computers and technology statistical tecniques which used in numerical analysis has been developed. One of these tecniques is semiparametric regression which can be used when parametric and nonparametric regression models are not suitable for the analysis. Semiparametric regression has been developed to perform analysis in such cases.

The first studies about the semiparametric regression model was initiated in 1985 and from that day forward it has been developed in various studies and applied in various fields.

İn this study semiparametric regression model has explained theoretical framework and after that housing estates which have very important place in the construction industry today and the effective factors of apartment prices in this places were examined by semiparametric regression.

Keywords: Semiparametric regression, apartment prices.

57 * İstanbul Ticaret Üniversitesi, Fen Edebiyat Fakültesi, Prof.Dr.
58 * İstanbul Ticaret Üniversitesi, Fen Edebiyat Fakültesi, Arş.Gör.

GİRİŞ

Klasik regresyon teknikleri, bağımlı değişkenin bağımsız değişkenlerle doğrusal bir ilişki içerisinde olduğunu ve ilişkinin şeklinin biliniyor olduğunu varsayar. Bu varsayımların sağlanamaması durumunda ise parametre tahminleri güvenilir olmamaktadır. İlişkinin şeklinin bilinmediği ya da bilinen parametrik matematiksel kalıplara uymadığı durumlarda parametrik olmayan regresyon teknikleri kullanılmaktadır. Ancak bu teknikler birden fazla bağımsız değişken olması durumunda çok boyutluluğun yarattığı sıkıntı nedeniyle özellikle yorumlama aşamasında sorunlara neden olmaktadır. Birden fazla bağımsız değişken söz konusu olduğunda, bağımsız değişkenlerin bazıları bağımlı değişkenle doğrusal ilişki içerisinde bulunabilirken, bazıları doğrusal olmayan ilişki içerisinde bulunabirler. Bu tür ilişkilerin modellenebilmesi için, parametrik ve parametrik olmayan regresyon fonksiyonunun toplamsal olarak birleşiminden oluşan semiparametrik regresyon modellerinden yararlanılmaktadır.

Bu çalışmada öncelikle sitelerdeki dairelerin satış fiyatını etkileyen özellikler açıklanmış, daha sonra uygulamada kullanılacak semiparametrik regresyon modelleri teorik olarak incelenmiştir. Uygulama aşamasında ise öncelikle uygulamada kullanılan değişkenler belirlenmiş ve değişkenlerin orijinal grafikleri incelenmiştir. Bu aşamadan sonra istatistiksel olarak anlamsız olan değişkenler belirlenmiş ve semiparametrik regresyon modelleri ile doğrusal, logaritmik ve karesel dönüşümle oluşturulan modeller karşılaştırılmıştır. Son olarak semiparametrik regresyon modeli seçilmiş ve modelin istatistiksel anlamlılığı incelenmiştir.

SİTELERDEKİ DAİRELERİN SATIŞ FİYATINI ETKİLEYEN ÖZELLİKLER

Son yıllarda site içerisinde konut alımı ve yaşantısı önem kazanmaktadır. Ancak daire satın alınırken sadece site içerisinde olması değil diğer bir çok özelliklerde dikkate alınmaktadır. Öncelikle daire satın alırken fiyatı etkileyen ve dikkat edilmesi gereken unsurlardan bahsetmek gerekmektedir. Yapılan çeşitli araştırmalara göre daire satın alınırken müşterilerin 40 farklı özelliği dikkate aldıkları belirlenmiştir. (www.satilikdaire-ariyorum.blogcu.com; 29.03.2010). Bu özellikler, inşaat kalitesi, arsa alanı, İnşaat alanı, toplam kullanım alanı, net alan (oda, salon, koridor, balkon), sosyal tesisler, yeşil alan, mevki, doğa manzarası, güvenlik sistemi, spor kompleksi, kapıcı dairesi, asansör, hidrafor ve su deposu, açık otopark, kapalı otopark, şömine ve barbekü, balkon, oda sayısı, satış kabiliyeti, havuz, ulaşım, semt özelliği, tapu durumu, zemin durumu, ısınma, konutun bulunduğu kat, prim getiri potansiyeli, deniz manzarası, alış veriş merkezine yakınlığı, aidat, malzeme kalitesi, yapım yılı, bina özelliği (apartman ya da bağımsız ev olması), güneş alma durumu, özel dekorasyon, kira geliri, depreme dayanıklılık, net alanla brüt alan arasındaki fark gibi unsurlardır.

Görüldüğü gibi dairelerin satış fiyatlarını etkileyen bir çok unsur bulunmaktadır. Dairelerin satış fiyatı üzerinde, kuşkusuz sitenin bulunduğu mevkinin önemi çok büyüktür. Ancak bu çalışmada mevki gibi özelliklerin değil daha çok fiziksel özelliklerin satış fiyatı üzerinde etkisi araştırılmak istenmiştir. Ayrıca incelenen bütün daireler için açık havuz, kapalı havuz, açık otopark, kapalı otopark, spor salonu, koşu alanı, kafeterya, güvenlik ve çocuk parkı gibi özellikler bulunmaktadır. Böylelikle site özelliklerinin satış fiyatı üzerindeki etkisi ortadan kaldırılmıştır (Bağdatlı, 2010: 47).

SEMİPARAMETRİK REGRESYON MODELLERİ

Semiparametrik regresyon modelleri standart regresyon tekniklerini genelleştiren ve her bir değişkenin etkisinin açık bir şekilde yorumlanmasını sağlayan toplamsal modellerin özel bir durumudur (Aydın, 2005: 48). Bu durumda, semiparametrik regresyon modelleri toplamsal modellere parametrik bileşen eklenerek oluşturulan modellerdir. Semiparametrik regresyon modelleri çok boyutluluğun yarattığı sıkıntı nedeniyle parametrik olmayan modellere tercih edilmektedir.

Semiparametrik regresyon modeli,

$$y_i = \alpha + f_1(x_1) + \ldots + f_j(x_j) + \beta_1 x_{j+1} + \ldots + \beta_k x_k + \varepsilon \qquad (1)$$

Biçiminde ifade edilir. (1) modelindeki j tane değişkenin y bağımlı değişkeni üzerinde doğrusal olmayan etkisi bulunmakta ve modelin parametrik olmayan bölümünü oluşturmaktadır. Diğer değişkenlerin ise, y bağımlı değişkeni üzerinde doğrusal etkisi bulunmakta ve modelin parametrik bölümünü oluşturmaktadır. Modelin parametrik bölümünde kukla değişkenler gibi kesikli değişkenlere yer verilebilmektedir. Aşağıda görülen (2) modeli incelendiğinde

$$y_i = \alpha + f_1(x_1) + f_2(x_2) + \beta_3 x_3 + \beta_4 x_4 + \varepsilon \qquad (2)$$

x_3 değişkeni kukla değişken, x_1, x_2, x_4 değişkenleri ise sürekli değişkenlerdir. Böyle bir semiparametrik regresyon modelinde bir çok etkileşim modele dahil edilebilir. Örneğin, x_1 ve x_2 değişkenleri arasındaki doğrusal olmayan ilişki tahmin edilebilir. Bu durumda, çoklu parametrik olmayan regresyon modellerinde olduğu gibi üç boyutlu bir grafik elde edilecektir. Araştırmanın konusuna ve içeriğine bağlı olarak x_3 ve x_4 değişkenleri arasındaki ilişkiler de incelenebilir. Ayrıca modelin parametrik ve parametrik olmayan bölümündeki değişkenlerin birbirleriyle etkileşimi de analiz edilebilmektedir. Semiparametrik regresyon modelinin parametrik kısmı doğrusal ilişkiyi, modelin parametrik olmayan kısmı ise doğrusal olmayan ilişkiyi ifade etmektedir. Bu nedenle semiparametrik regresyon modellerine yarı doğrusal modeller adı da verilmektedir (Lee, 1990: 203).

Bir model kurma aşamasında ilk olarak değişkenler belirlenmektedir. Değişkenler belirlendikten sonra ise modelin fonksiyonel şeklinin veya matematiksel kalıbının belirlenmesi gerekmektedir. Matematiksel kalıp oluşturken öncelikli olarak yapılması gereken grafiklerin incelenmesidir. Bağımlı değişken ile herbir bağımsız değişkenin ayrı ayrı çizilecek grafikleri incelenerek, bağımlı değişken ile bağımsız değişkenler arasındaki ilişkinin yapısı hakkında fikir sahibi olunabilir. Matematiksel kalıp ile ilgili tereddütler söz konusu olduğunda, farklı şekillerin denenmesi en uygun sonucu elde etmek için yararlı olacaktır. Herbir değişkenin ilişkisine tek tek bakıldıktan sonra bağımsız değişkenlerin bir veya bir kaçı için parametrik olmayan, diğerleri için parametrik ilişki uygun ise semiparametrik regresyon modeli en uygun model olarak tercih edilecektir (Çağlayan, 2002: 75).

Semiparametrik regresyon modellerinde parametrik kısım, doğrusal olabileceği gibi dönüşüm yöntemleri (Logaritmik, karesel dönüştürme vs), uygulanarak doğrusallaştırılabilinen yapıda da olabilir. Semiparametrik modelin parametrik kısmının belirlenmesinde farklı modeller tahmin edilebilir. Tahmin edilen bu modellerden artık kareler toplamını minimum yapan model semiparametrik modelin parametrik kısmı olarak tahmin edilebilir.

Değişkenler arasındaki ilişkiyi en iyi şekilde açıklayacak model çeşitli denemeler sonucunda da bulunabilir. Özellikle değişkenler arasında şekli tam belirlenemeyen ilişkiler varsa farklı matematiksel kalıplar veya farklı değişkenler için parametrik olmayan ilişkileri kapsayan modellerin denenmesi ve en uygun olanının seçilmesi gerekmektedir (Çağlayan, 2002: 76).

3.1. Semiparametrik Regresyon Modellerinin Tahmini: Backfitting Algoritması

Toplamsal modellerin ve semiparametrik regresyon modellerinin tahmininde tekrarlı (iterative) algoritmalara ihtiyaç duyulmaktadır. Bu modellerin tahmini için geliştirilen birçok algoritma bulunmakta ve bu algoritmalar birçok değişik bilgisayar programında yer almaktadır. Özellikle, R programı birçok algoritmayı desteklemektedir. Bu algoritmalardan en çok kullanılanlar Newton- Raphson algoritması ve backfitting algoritmasıdır. Backfitting algoritması Hastie ve Tibshirani tarafından 1990 yılında tanıtılmıştır. Bu algoritma parametrik olmayan ve parametrik bileşenleri tahmin edebilen en kolay yöntem olarak bilinmektedir.

Model tahmin etme aşamasında, x değişkenleri birbirlerine dik iseler modelin parametrik kısmı iki değişkenli modeller serisi olarak sıradan en küçük kareler yöntemi kullanılarak tahmin edilebilir. Parametrik olmayan bileşenlerin tahmininde ise lowess ya da splaynlar kullanılabilir. Bağımsız değişkenler arasında korelasyon bulunmaması durumuna genellikle rastlanmamaktadır. Bu durumda toplamsal modelleri ya da semiparametrik modelleri tahmin ederken bağımsız değişkenler arasındaki ilişkiyi dikkate alacak yöntemlere ihtiyaç duyulmaktadır. Backfitting algoritması parametrik ve parametrik olmayan bileşenleri tahmin ederken bağımsız değişkenler arasındaki korelasyonu dikkate almak üzere tasarlanmıştır.

Backfitting algoritması kısmi regresyon fonksiyonları fikrini önermektedir. Eşitlik (3)'de iki bağımsız değişkenli toplamsal bir model görülmektedir.

$$y = \alpha + f_1(x_1) + f_2(x_2) + \varepsilon \qquad (3)$$

Bu modelde f_2'nin gerçek fonksiyonel formunun bilindiği ancak f_1'in bilinmediği varsayılsın. Bu durumda f_1'in tahmini için (3) modeli kısmi regresyon fonksiyonu olarak eşitlik (4)'deki gibi yeniden düzenlenmelidir.

$$y - \alpha - f_2(x_2) = f_1(x_1) + \varepsilon \qquad (4)$$

(4) eşitliğinde x_1'e karşı $y - \alpha - f_2(x_2)$ 'nin düzgünleştirilmesi $f_1(x_1)$'in tahminini elde etmeyi sağlamaktadır. Bu nedenle, bir kısmi regresyon fonksiyonunu bilmek

diğer kismi regresyon fonksiyonunu tahmin etmeye olanak sağlamaktadır. Gerçek durumda, hiçbir regresyon fonksiyonunu bilmek mümkün olmamaktadır. Ancak f 'lerden herhangi biri için bir başlangıç değeri belirlenirse, toplamsal modellerin tahmini için kismi regresyon fonksiyonları tekrarlı yöntemler ile çözümlenir. Model (5) tahmin edilmek istensin:

$$y_i = \alpha + f_1(x_1) + ... + f_k(x_k) + \varepsilon \qquad (5)$$

(5) eşitliğinde S_j, sütunları X tahminlerinden oluşan bir matrisi ifade etmektedir. X ise, kolonları x değişkenlerinden oluşan model matrisini ifade etmektedir. Toplamsal modellerin tahmini için backfitting algoritması aşağıdaki adımlardan oluşmaktadır (Hastie ve Tibshirani, 1999: 118-119).

$\alpha = \bar{y}$ ve $S_j = X\ (j = 1,...,m)$ başlangıç değerleri olarak seçilir.

Her x değişkeni için kismi artıklar hesaplanır. x_1 değişkeni için tahmin edilen kismi artıklar eşitlik (6)'da görüldüğü gibidir.

$$\hat{e}_p^j = y_i - \sum_{i=2}^{k} S_j - \alpha \qquad (6)$$

x_1 değişkeni civarında e_p^j düzgünleştirilir. Bu aşama için parametrik olmayan regresyon modeli seçilmelidir (splaynların özelliklerinden dolayı birçok bilgisayar yazılımı backfitting algoritmasının üçüncü aşaması için splaynları kullanmaktadır).

S_j'deki x_1 değişkeni, x_i'nin düzgünleştirilmiş tahminleri ile değiştirilir.

2'den k'ya kadar olan her x değişkeni için yukarıdaki adımlar tekrarlanır.

Eşitlik (7)'de görülen model ile artık kareler toplamı hesaplanır.

$$RSS = \sum_{i=1}^{n}\left[\left(yi - \sum_{i=1}^{k} S_j\right)^2\right] \qquad (7)$$

Artık kareler toplamındaki değişim belirli bir tolerans seviyesinde ise model yakınsar ve algoritma durur. Eğer değilse, bu işlem artık kareler toplamındaki değişim belirli bir tolerans seviyesine gelene kadar devam eder.

Backfitting algoritması durduğunda S_j'nin her sütünü x değişkeninin parametrik olmayan tahminini içerir. Bu tahminler x değişkenleri arasındaki ilişkiyi dikkate alır. Dolayısıyla, üç x değişkenine sahip bir toplamsal model tahmin edildiğinde $\hat{f}_1$'in

grafiği, x_2 ve x_3 değişkenlerinin etkisi sabit tutulduğunda x_1'in y_i üzerindeki etkisi olarak yorumlanabilir. Backfitting algoritmasının bir çok varyasyonu bulunmaktadır. Bu varyasyonlardan en çok kullanılanlardan biri başlangıç değeri olarak sıradan en küçük kareler tahmincilerini kullanmaktır.

Backfitting algoritması semiparametrik regresyon modelleri için aynı adımları içermektedir. Öncelikle, modeldeki her bir bağımsız değişken için kısmi artıklar oluşturulur. Eğer seçilen değişkenin doğrusal olmayan uyumu söz konusu ise bu değişken için kısmi artıklar, aynı bağımsız değişkene karşı düzeltilebilirler. Eğer seçilen değişkenin doğrusal uyumu söz konusu ise düzgünleştirme yöntemi yerine sıradan en küçük kareler yöntemi kullanılır. Bacfkitting algoritması, algoritmada yapılabilen değişikliklerden dolayı birçok regresyon modelinin tahmininde kullanılmaktadır.

Semiparametrik Regresyon Modellerinde Çıkarım

Semiparametrik regresyon modellerinde çıkarım, doğrusal modellerde çıkarım ile parametrik olmayan modellerde çıkarımın birleşiminden oluşmaktadır. Modeldeki parametrik olmayan değişkenler için güven bantları hesaplanır. Modeldeki doğrusal bileşenler için ise güven aralıklarını oluşturmak ve hipotez testlerini uygulamak için standart hatalar hesaplanır.

Parametrik olmayan değişken için oluşturulacak olan güven bantları ve standart hataların hesaplanması için varyans-kovaryans matrisinin tahminine ihtiyaç duyulmaktadır. Semiparametrik regresyon modellerinde varyans-kovaryans matrisinin tahmini parametrik olmayan regresyon modellerindeki tahmin ile çok benzer ancak daha karmaşıktır.

Semiparametrik modellerde parametrik kısımda bulunan ε_i 'nin klasik doğrusal regresyon varsayımlarını sağlaması gerekmektedir. Bu durumda semiparametrik regresyon modellerinin hata teriminin (ε_i) klasik doğrusal regresyonun tüm varsayımlarına sahip olması gerekir (Fox, 2000: 35). Bu varsayımların gerçekleşmesi tutarlı tahminlerin elde edilmesini sağlayacaktır. Literatürde, varsayımların geçerliliğini incelemek için bazı testler bulunmaktadır. Ancak, semiparametrik regresyon modellerinde varsayımlar çoğunlukla artık grafikleri yardımıyla incelenmektedir.

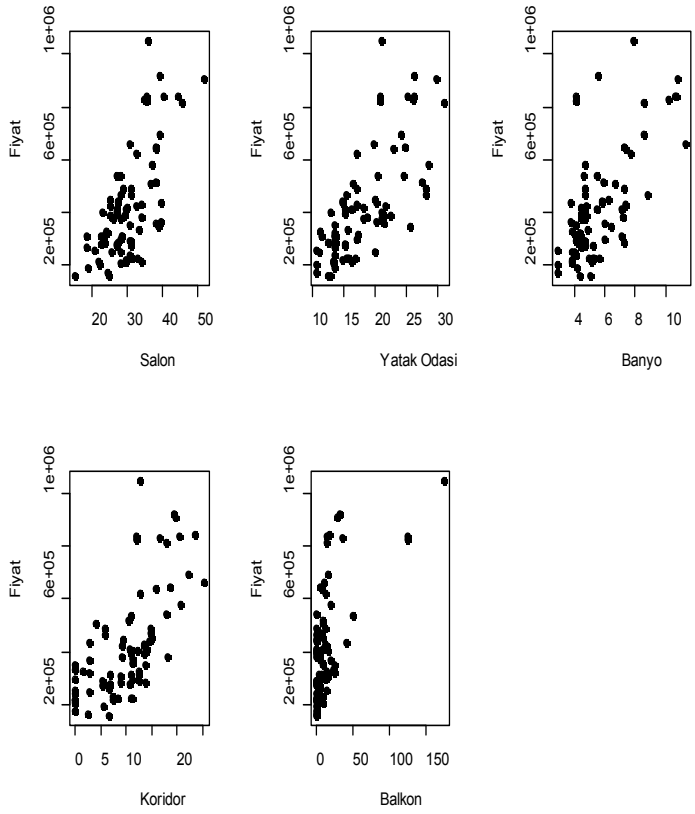
UYGULAMA

4.1.Uygulamada Kullanılan Değişkenler ve Değişkenlerin Orijinal Grafiklerinin İncelenmesi

Uygulamada kullanılan veriler, İstanbul'un Çekmeköy İlçesinde, birbiriyle yakın mesafede bulunan siteler içerisindeki 81 daireden elde edilmiştir. Daha önce bahsedildiği üzere dairelerin satış fiyatı üzerinde, sitenin bulunduğu mevkinin önemi çok büyüktür. Ancak bu çalışmada mevki gibi özelliklerin değil daha çok fiziksel özelliklerinin satış fiyatı üzerinde etkisi araştırılmak istenmiştir. Ayrıca incelenen bütün dairelerin içinde bulunduğu sitelerde açık havuz, kapalı havuz, açık otopark, kapalı otopark, spor salonu, koşu alanı, kafeterya, güvenlik ve çocuk parkı bulunmaktadır. Böylelikle site özelliklerin satış fiyatı üzerindeki etkisi ortadan kaldırılmış ve sadece aşağıda adı

geçen ev özelliklerinin etkisini ortaya çıkarmak hedeflenmiştir. Salon büyüklüğü, yatak odası büyüklüğü, banyo büyüklüğü, koridor büyüklüğü, balkon büyüklüğü, oda sayısı (oda sayısı için kukla değişken : 2 odalı, 3 odalı, 4 odalı, 5 odalı, 6 odalı), cephe (cephe durumu için kukla değişken : site içi cepheli, site dışı cepheli, site içi ve site dışı cepheli), su deposu (su deposu için kukla değişken: su deposu var, su deposu yok) gibi özellikler uygulamanın bağımsız değişkenleri olarak ele alınmıştır. Görüldüğü gibi incelenen değişkenler site içerisindeki dairelerin fiziksel özellikleridir. Dairenin site içerisinde bulunması ya da bağımsız olması dairelerin satış fiyatını etkileyecektir. Ayrıca sitenin diğerlerinden çok farklı ve ayrıcalık yaratacak özellikleri de olabilmektedir. Ancak bu uygulama için seçilen siteler bölge olarak ve site özelliği olarak çok büyük benzerlikler göstermektedir.

İncelenen değişkenlerin fiyat bağımlı değişkeni ile olan ilişkisini belirlemek amacıyla Şekil 1'de görülen grafikler oluşturulmuştur. Bu grafikler incelendiğinde, bütün değişkenler için fiyat ile ilişkinin kesinlikle doğrusal ya da kesinlikle eğrisel olduğunu söylemek mümkün değildir. Dolayısıyla, hangi değişkenlerin modele düzgünleştirilerek dahil edilmesi gerektiğini saptamak için sadece grafikler yeterli olmamaktadır. Bu nedenle, fiyat ile doğrusal olmayan ilişki içerisinde bulunan değişkenlerin istatistiksel olarak test edilmesi gerekmektedir.



Şekil **1: Değişkenlerin Orijinal Grafikleri**

4.2. Fiyat İle Doğrusal Olmayan İlişki İçerisinde Bulunan Değişkenlerin Belirlenmesi

Fiyat ile doğrusal olmayan ilişki içerisinde bulunan değişkenlerin belirlenmesi amacıyla sürekli olan bütün değişkenler modele pürüzsüz fonksiyon olarak dahil edilmiştir. Bu modeldeki öncelikli amaç, değişkenlerin hangilerinin dairelerin satış fiyatı üzerinde doğrusal olmayan etki içerisinde olduklarını ortaya çıkartmaktır. Ancak, bir diğer amaç ise, fiyat bağımlı değişkeni üzerinde istatistiksel olarak anlamsız olan değişkenleri de belirlemektir. Tablo 1'de bu amaçla oluşturulan modelin sonuçları görülmektedir. Bilindiği üzere, uygulamada kukla değişkenler söz konusudur. Semiparametrik regresyon yönteminde kukla değişkenler modele parametrik olarak dahil edilmektedir (Keele, 2008: 87). Oluşturulan semiparametrik regresyon modeli sabit terim içerdiğinden kukla değişkenler içerisindeki bir kategori, referans kategori olarak seçilir. Sabit terim içeren bir modelde her kategori için oluşturulan kukla değişkenlerin modele dahil edilmesi durumunda tam çoklu doğrusal bağıntı problemi ortaya çıkacağından regresyonun tahmini gerçekleşemeyecektir. Semiparametrik regresyonda da çoklu doğrusal bağıntı olmaması varsayımının sağlanması gerekmektedir (Bağdatlı, 2010: 50).

Tablo 1. Fiyat İle Doğrusal Olmayan İlişki İçerisinde Bulunan Değişkenlerin Belirlenmesi İçin Oluşturulan Model Sonuçları

Değişkenler	*Parametre Tahmini*	*Standart Hata*	t -İstatistiği	Olasılık
Sabit	333669	23561	14.162	< 2e-16 ***
Oda Sayısı 3	64292	15124	4.251	7.47e-05 ***
Oda Sayısı 4	142767	20948	6.815	4.95e-09 ***
Oda Sayısı 5	266921	32429	8.231	1.85e-11 ***
Su Deposu	-37489	19209	-1.952	0.0556
İç Cephe Manzaralı	20429	20188	1.012	0.3156
İç ve Dış Cephe Manzarası	20133	22835	0.882	0.3814
Bileşen	F - İstatistiği	Olasılık		
Salon	6.673	0.01218 *		
Yatak Odası	4.697	0.00036 ***		
Banyo	4.113	0.00113 **		
Koridor	4.151	0.00784 **		
Balkon	59.922	< 2e-16 ***		
R^2(adj)= 0.954	Açıklanan Sapma= %96.5	GCV Değeri = 2.5682e+09		

*: Parametrelerin 0.05 anlam düzeyinde istatistiksel olarak anlamlı olduğunu ifade etmektedir.

**: Parametrelerin 0.01 anlam düzeyinde istatistiksel olarak anlamlı olduğunu ifade etmektedir.

***: Parametrelerin 0.001 anlam düzeyinde istatistiksel olarak anlamlı olduğunu ifade etmektedir.

Uygulamadaki değişkenler incelendiğinde, cephe değişkenin üç kategoriden oluştuğu gözlenmektedir. Ancak Tablo 1'de görüldüğü gibi, iç cephe ve iç-dış cephe manzara kategorileri için tahmin yapılmış ve referans kategori olarak dış cephe seçilmiştir. Referans kategori araştırıcının istediğine bağlı olarak belirlenebilmekte ve değiştirilebilmektedir. Kukla değişkenlerin yorumu referans kategoriye göre yapılmaktadır.

Örneğin; diğer değişkenlerin etkisi sabit tutulduğunda, dairenin iç cephe manzaralı olması dış cephe manzaralı olmasına göre daire fiyatını ortalama 20429 TL arttırmaktadır. Ancak, Tablo 1'de görüldüğü üzere, modele kukla değişken olarak katılan cephe değişkenin her iki kategorisi de dış cephe manzaralı olma kategorisine göre istatistiksel olarak anlamsız bulunmaktadır. Ancak referans kategori, bütün kategoriler için denenmiş ve istatistiksel olarak anlamlı sonuçlara ulaşılamamıştır. Dolayısıyla genel olarak cephe faktörünün dairelerin satış fiyatları üzerinde etkili olmadıkları sonucuna varılmakta ve yukarıda yapılan yorum gerçeği yansıtmamaktadır. Bu nedenle cephe değişkeni modelden çıkartılmış ve aynı yöntemle yeni bir model kurulmuştur.

Tablo 1'de düzgünleştirilen bileşenlerin (Pürüzsüz Fonksiyonların) F istatistikleri ve yaklaşık olasılık değerleri görülmektedir. Bu değerler incelendiğinde düzgünleştirilen değişkenlerden yatak odası, banyo, koridor, balkon ve salon değişkenlerinin istatistiksel olarak anlamlı olduğu görülmektedir. Sonuç olarak, bu değişkenlerin modele pürüzsüz bir fonksiyon olarak dahil olması anlamlıdır sonucuna ulaşılmaktadır. Ancak, her bir değişken için yapılacak ayrı analizler sonucunda uygulanması gereken son modele karar verilecektir.

Tablo 2: Fiyat İle Doğrusal Olmayan İlişki İçerisinde Bulunan Değişkenlerin Belirlenmesi İçin Oluşturulan 2. Model (Temel Model) Sonuçları

Değişkenler	*Parametre Tahmini*	*Standart Hata*	t -İstatistiği	Olasılık
Sabit	345372	20387	16.941	< 2e-16 ***
Oda Sayısı 3	64221	14645	4.385	4.54e-05 ***
Oda Sayısı 4	146501	19984	7.331	5.55e-10 ***
Oda Sayısı 5	268979	317773	8.466	5.89e-12 ***
Su Deposu	-30941	16733	-1.849	0.0692
Bileşen		F - İstatistiği	Olasılık	
Salon		6.625	0.01245 *	
Yatak Odası		4.829	0.00055 ***	
Banyo		4.760	0.00137**	
Koridor		4.181	0.01161*	
Balkon		65.728	< 2e-16 ***	
R^2(adj)= 0.955		Açıklanan Sapma= %96.5	GCV Değeri = 2.4452e+09	

*: Parametrelerin 0.05 anlam düzeyinde istatistiksel olarak anlamlı olduğunu ifade etmektedir.

**: Parametrelerin 0.01 anlam düzeyinde istatistiksel olarak anlamlı olduğunu ifade etmektedir.

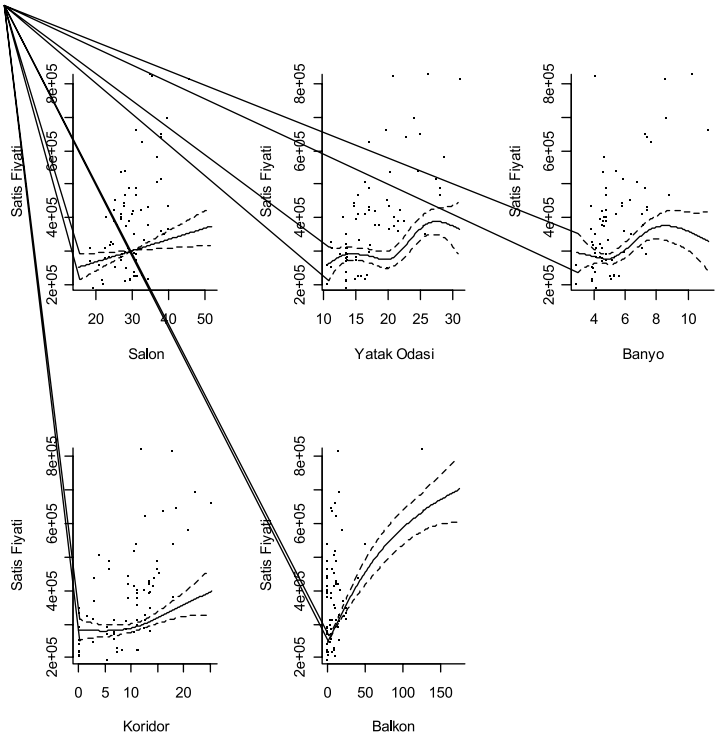
***: Parametrelerin 0.001 anlam düzeyinde istatistiksel olarak anlamlı olduğunu ifade etmektedir.

Fiyat ile doğrusal olmayan ilişki içerisinde bulunan değişkenlerin belirlenmesi amacıyla kurulan ilk modelden cephe değişkeninin istatistiksel olarak anlamlı olmadığı sonucuna ulaşıldığından cephe değişkeni modelden çıkartılarak yeni bir model oluşturulmuştur. Bu modelin sonuçları Tablo 2'de gösterilmektedir.

Tablo 2 incelendiğinde kukla değişkenlerden su deposu değişkenin 0.10 anlam düzeyinde istatistiksel olarak anlamlı olduğu diğer değişkenlerin ise çok düşük anlam düzeyinde bile istatistiksel olarak anlamlı olduğu görülmektedir. Su deposu değişkeni-

nin modelden çıkartılması ileride otokorelasyon problemine neden olabileceğinden, bu değişken modele dahil edilmektedir. Ancak, uygulanacak en son modele karar verme aşamasında bu değişken yeniden değerlendirilecektir.

Düzgünleştirilen bileşenlerin *F* istatikleri ve olasıkları incelendiğinde bu değişkenlerin modele pürüzsüz bir fonksiyon olarak dahil olması anlamlıdır sonucuna ulaşılmaktadır. Bilindiği üzere bu modellerin oluşturulmasındaki temel amaç fiyat ile doğrusal olmayan ilişkisi olan değişkenlerin belirlenmesidir. Bu amaçla Tablo 2'deki modelin grafikleri incelenmiştir.



Şekil **2. Fiyat İle Doğrusal Olmayan İlişki İçerisinde Bulunan Değişkenlerin Belirlenmesi İçin Oluşturulan 2. Model (Temel Model)'in Grafikleri**

Şekil 2 incelendiğinde salon değişkeni hariç bütün sürekli değişkenlerin fiyat bağımlı değişkeni ile doğrusal olmayan bir ilişki içerisinde olduğu görülmektedir. Sonuç olarak, Tablo 2'de oluşturulan model temel model olarak ele alınacaktır. Değişkenlerin modele pürüzsüz bir fonksiyon olarak, doğrusal olarak veya üstel dönüştürme yöntemlerini kullanarak dahil edilmesi aşamasına, temel model ile yapılan karşılaştırmalar sonucunda kesin olarak karar verilecektir.

4.3. Değişkenlerin Fiyat Değişkeni ile Doğrusal ya da Doğrusal Olmayan İlişkisinin Test Edilmesi

Bu aşamada kurulan bütün modeller, tüm sürekli değişkenlerin pürüzsüz bir fonksiyon

olarak modele[59] dahil edildikleri ve sonuçları Tablo 2'de gösterilen model (temel model) ile karşılaştırılmıştır. Modellerin karşılaştırılması semiparametrik regresyon modelleri için oluşturulan sapma analizi (Analysis of Deviance) ile gerçekleştirilmiştir. Her bir sürekli değişken için doğrusal model oluşturulmuş ve temel model ile karşılaştırılmıştır. Yapılan değerlendirmelerde sadece ilgili değişken modele doğrusal bir fonksiyon olarak dahil edilmiş ve temel model ile karşılaştırılmıştır. Örneğin, Yatak odası değişkeni için yapılan değerlendirmede, sadece yatak odası değişkeni modele doğrusal bir fonksiyon olarak dahil edilmiş ve temel model ile karşılaştırılmıştır. Bu karşılaştırma için aşağıda görülen hipotezler oluşturulmuştur. Görüldüğü gibi H_1 hipotezi temel modeli ifade etmektedir.

H_0 : E(Fiyat) =β_0 + s(Salon)[1] +β1 Oda Sayısı + β_2 Yatak Odası + s(Banyo) + s(Koridor) + s(Balkon) + β_3 Su Deposu

H_0 : E(Fiyat) =α_0 + s(Salon)[2] +α_1 Oda Sayısı + s(Yatak Odası + s(Banyo) + s(Koridor) + s(Balkon) + α_3 Su Deposu +ε

H_0 hipotezinin kabul edilmesi durumunda, yatak odası değişkeninin fiyat bağımlı değişkeni ile olan ilişkisinin doğrusal olduğu kabul edilecektir. H_0 hipotezinin red edilmesi durumunda ise (olasılık değeri 0.05'ten küçük ise H_0 hipotezi red edilir) ilişkinin doğrusal olmadığı ve yatak odası değişkeninin modele düzgünleştirilerek yani pürüzsüz bir fonksiyon olarak dahil edilmesi gerektiği sonucuna ulaşılacaktır. Sapma analizi bütün değişkenler için yapılmış ve elde edilen analiz sonuçları Tablo 3'de gösterilmiştir.

Tablo 3. Değişkenlerin Fiyat Değişkeni ile Doğrusal ya da Doğrusal Olmayan İlişkisinin Sapma Analizi Sonuçları

Değişkenler	Sapma Değeri	Olasılık Değeri	*Karar*
Yatak Odası (Büyüklüğü)	2.4037e+10	0.003089	H_0 hipotezi reddedilir
Banyo (Büyüklüğü)	-1.7336e+10	0.2042	H_0 hipotezi reddedilemez[1].
Koridor (Büyüklüğü)	8928534475	0.01325	H_0 hipotezi reddedilir.
Balkon (Büyüklüğü)	1.8216e+10	0.001987	H_0 hipotezi reddedilir.

59 $s_i(x_i)$ olarak ifade edilen fonksiyon splayn düzeltme ile düzgünleştirilmiş bir fonksiyonu ifade etmektedir. Bu fonksiyon önceki bölümlerde $f_i(x_i)$ olarak gösterilmiştir. Buradaki amaç, splayn düzeltmenin kullanıldığını vurgulamaktır.

Tablo 3'den elde edilen sonuçlara göre; yatak odası, koridor ve balkon değişkenlerinin modele düzgünleştirilerek yani pürüzsüz fonksiyonlar olarak, banyo değişkeninin ise modele doğrusal bir fonksiyon olarak dahil edilmesine karar verilmiştir. Salon değişkenine ise sapma analizi uygulanmamıştır. Şekil 2'de salon değişkeninin fiyat değişkeni ile doğrusal bir ilişkisi olduğu açıkça görülmektedir. Böyle bir değişkene sapma analizi uygulandığında olasılık değeri oluşmayabilir. Bu durumun nedeni, belirgin bir doğrusallık söz konusu olduğunda test edilmesi gereken bir durumun olmamasıdır. Özellikle, düzeltme parametresi otomatik yöntemle seçildiğinde bu durumla karşılaşılabilmektedir. (Çevrimiçi: http://r.789695.n4.nabble.com; 10.03.2010).

4.4. Logaritmik Modeller İle Temel Modelin Karşılaştırılması

Bu aşamada bütün sürekli değişkenler için logaritmik modeller oluşturulacak ve sapma analizi araclığıyla temel model ile karşılaştıralacaktır. Dönüştürme yöntemleri kullanılarak düzgünleştirilebilecek bir değişken pürüzsüz fonksiyon olarak modele dahil edildiğinde cimrilik prensibinden dolayı (parsimony) olumsuz sonuçlar yaratacaktır. Cimrilik prensibi (rule of parsimony), verinin en iyi şekilde yansıtılabilmesi için en az sayıda parametre yani değişken kullanılmasını önerir (Box ve Jenkins, 1970: 17). Modeldeki parametre sayısı arttıkça yanlılık azalır ancak varyans büyür. Fakat bir modelde çıkarımlar açısından varyansın büyük olması sorunlara yol açmaktadır. Cimrilik prensibi eğilim ve varyans arasındaki uygun dengeyi başarmaya çalışmaktadır.

Böyle bir sorunla karşılaşmamak için bütün karşılaştırmalar yapılacaktır. Örneğin, yatak odası değişkeni için yapılan değerlendirmede, yatak odası değişkeninin logaritması alınarak semiparametrik regresyon modeli oluşturulmuş ve H_1 hipotezinde görülen temel model ile karşılaştırılmıştır. Bu karşılaştırma için aşağıda görülen hipotezler oluşturulmuştur.

$$H_0 : \text{E(Fiyat)} = \beta_0 + \text{s(Salon)} + \beta_1 \text{Oda Sayısı} + \beta_2 \text{Yatak Odası} + \beta_3 (\text{Yatak Odası})^2 + \text{s(Banyo)} + \text{s(Koridor)} + \text{s(Balkon)} + \beta_4 \text{Su Deposu} + \varepsilon$$

$$H_1 : \text{E(Fiyat)} = \alpha_0 + \text{s(Salon)} + \alpha_1 \text{Oda Sayısı} + \text{s(Yatak Odası)} + \text{s(Banyo)} + \text{s(Koridor)} + \text{s(Balkon)} + \alpha_2 \text{Su Deposu} + \varepsilon$$

H_0 hipotezinin kabul edilmesi durumunda, yatak odası değişkeninin fiyat bağımlı değişkeni ile olan ilişkisinin, yatak odası değişkeninin logaritması alınarak doğrusallaştırılabildiği kabul edilecektir. H_0 hipotezinin red edilmesi durumunda ise ilişkinin logaritma alınarak doğrusallaştırılamadığı ve yatak odası değişkeninin modele düzgünleştirilerek yani pürüzsüz bir fonksiyon olarak dahil edilmesi gerektiği sonucuna ulaşılacaktır. Sapma analizi bütün değişkenler için yapılmış ve elde edilen analiz sonuçları Tablo 4'de gösterilmiştir.

Tablo 4. Logaritmik Modeller İle Temel Modelin Karşılaştırılması İçin Yapılan Sapma Analizi Sonuçları

Değişkenler	Sapma Değeri	Olasılık Değeri	*Karar*
Yatak Odası (Büyüklüğü)	2.4756e+10	0.002708	H_0 hipotezi reddedilir.
Banyo (Büyüklüğü)	-1.7023e+10	0.262	H_0 hipotezi reddedilemez.[1]
Koridor (Büyüklüğü)	860788472	Hesaplanamamıştır[2].	
Balkon (Büyüklüğü)	-3009506410	Hesaplanamamıştır[3].	

Tablo 4'de görülen sapma analizi sonuçlarına göre; Yatak odası, koridor ve balkon değişkenlerinin modele düzgünleştirilerek yani pürüzsüz fonksiyonlar olarak dahil edilmesine, banyo değişkeninin ise modele doğrusal bir fonksiyon olarak dahil edilmesine karar verilmiştir. Diğer bir ifade ile, değişkenlerin logaritmalarının alınması fiyat ile olan ilişkilerini doğrusallaştıramamıştır. Böylece ilk yapılan değerlendirmelerle aynı sonuca ulaşılmıştır.

4.5. Karesel Modeller İle Temel Modelin Karşılaştırılması

Bu aşamada bütün sürekli değişkenler için karesel modeller oluşturulacak ve sapma analizi aracılığıyla temel model ile karşılaştıralacaktır. Örneğin, yatak odası değişkeni için yapılan değerlendirmede, yatak odası değişkeninin, kendisi ve karesi semi-parametrik regresyon modeline dahil edilmiş ve H_1 hipotezinde görülen temel model ile karşılaştırılmıştır. Bu karşılaştırma için aşağıda görülen hipotezler oluşturulmuştur.

$$H_0 : \text{E(Fiyat)} = \beta_0 + \text{s(Salon)} + \beta_1 \text{Oda Sayısı} + \beta_2 \text{Yatak Odası} + \beta_3 (\text{Yatak Odası})^2 + \text{s(Banyo)} + \text{s(Koridor)} + \text{s(Balkon)} + \beta_4 \text{Su Deposu} + \varepsilon$$

$$H_1 : \text{E(Fiyat)} = \alpha_0 + \text{s(Salon)} + \alpha_1 \text{Oda Sayısı} + \text{s(Yatak Odası)} + \text{s(Banyo)} + \text{s(Koridor)} + \text{s(Balkon)} + \alpha_2 \text{Su Deposu} + \varepsilon$$

Sapma analizi sonucunda olasılık değeri 0.05 anlamlılık düzeyinden küçük olduğundan H_0 hipotezi reddedilmiş ve yatak odası değişkeninin modele karesel olarak değil, pürüzsüz bir fonksiyon olarak dahil edilmesine karar verilmiştir. Sapma analizi bütün değişkenler için yapılmış ve elde edilen analiz sonuçları Tablo 5'de gösterilmiştir.

Tablo 5. Karesel Modeller İle Temel Modelin Karşılaştırılması İçin Yapılan Sapma Analizi Sonuçları

Değişkenler	Sapma Değeri	Olasılık Değeri	*Karar*
Yatak Odası (Büyüklüğü)	2.4756e+10	0.002708	H_0 hipotezi reddedilir.
Banyo (Büyüklüğü)	-1.7023e+10	0.262	H_0 hipotezi reddedilemez[4].
Koridor (Büyüklüğü)	860788472	Hesaplanamamıştır[5].	
Balkon (Büyüklüğü)	-3009506410	Hesaplanamamıştır[6].	

Sonuç olarak yapılan bütün karşılaştırmalarda yatak odası, koridor ve balkon değişkenlerinin modele düzgünleştirilerek yani pürüzsüz fonksiyonlar olarak dahil edilmesine, banyo değişkeninin ise modele doğrusal bir fonksiyon olarak dahil edilmesine karar verilmiştir. Diğer bir değişle, değişkenlerin logaritmalarının ya da karelerinin alınması fiyat ile olan ilişkilerini doğrusallaştıramamıştır.

4.6. Uygun Semiparametrik Regresyon Modelinin Belirlenmesi

Buraya kadar yapılan analizlerde uygulanacak modele karar verilmesi amacıyla, ilk olarak istatistiksel olarak anlamlı olan değişkenler ve fiyat değişkeni ile doğrusal olmayan ilişkisi bulunan değişkenler belirlenmiştir. Daha sonra ise, ilgili değişkenler için logaritmik modeller ve karesel modellerin düzgünleştirmeye karşı üstün olup olmadığı test edilmiştir. Bütün işlemler sonucunda, cephe değişkeninin istatistiksel olarak anlamsız olduğu, yatak odası, koridor ve balkon değişkenlerinin düzgünleştirilerek diğer bir ifade ile pürüzsüz bir fonksiyon olarak modele dahil edilmesi gerektiği, banyo ve salon değişkenin fiyat değişkeni ile doğrusal bir ilişki içerisinde olduğu saptanmıştır. Tablo 2'de görüldüğü üzere su deposu değişkeninin 0.10 anlam düzeyinde istatistiksel olarak anlamlı olduğu belirlenmiştir. Uygulanacak en son model, bu değişkenin modelde bulunup bulunmaması konusunda yardımcı olacaktır. Bu model sonuçları Tablo 6'da görüldüğü gibidir.

Tablo 6: Uygulanmasına Karar verilen Semiparametrik Regresyon Modeli

Değişkenler	*Parametre Tahmini*	*Standart Hata*	*t* -İstatistiği	Olasılık
Sabit	146406	48744	3.004	0.00387 **
Salon	3328	1353	2.460	0.01678 *
Oda Sayısı 3	76260	15648	4.873	8.24e-06 ***
Oda Sayısı 4	154163	20892	7.379	5.31e-10 ***
Oda Sayısı 5	261428	341466	8.308	1.35e-11 ***
Banyo	160072	5719	2.810	0.00665 **
Su Deposu	-25603	20215	-1.267	0.21014
Bileşen		*F* -İstatistiği	Olasılık	
Yatak Odası		3.385	0.00622 **	
Koridor		3.777	0.00530 **	
Balkon		35.782	< 2e-16 ***	
R^2(adj) = 0.953		Açıklanan Sapma=%96.4	GCV Değeri = 2.633e+09	

*: Parametrelerin 0.05 anlam düzeyinde istatistiksel olarak anlamlı olduğunu ifade etmektedir.

**: Parametrelerin 0.01 anlam düzeyinde istatistiksel olarak anlamlı olduğunu ifade etmektedir.

***: Parametrelerin 0.001 anlam düzeyinde istatistiksel olarak anlamlı olduğunu ifade etmektedir.

Bu aşamadan sonra modellerin gösterimi için değişkenler, aşağıda gösterilen kısaltmalarla ifade edilecektir.

Salon: SL, Oda Sayısı 3: OD3, Oda Sayısı 4: OD4, Oda Sayısı 5: OD5, Banyo: BY, Su deposu: SD, Yatak Odası: YO, Koridor: KR, Balkon: BN, Fiyat: FY

Tablo 6'da sonuçları görülen semiparametrik regresyon modeli eşitlik (8)'de gösterilmiştir.

$$FY = 146406 + 76260OD3 + 154163OD4 + 261428OD5 + 3328SL + 160072BY - 25603SD + s(YO) + s(KR) + s(BN) + \varepsilon \quad (8)$$

Eşitlik (8), parametrik ve parametrik olmayan olarak iki bölümden oluşmaktadır. Bu bölümler için katsayı yorumları ve çıkarımlar ayrı yöntemlerle yapılmaktadır. Semiparametrik regresyon modelinin parametrik bölümü için yorumlar ve çıkarımlar doğrusal regresyon modelleri ile aynı olmaktadır. Parametrik olmayan bölüm için yorumlar grafik yardımı ile çıkarımlar ise F testi yardımı ile yapılmaktadır. Tablo 1 incelenirken su deposu değişkeninin modele dahil edilip edilemeyeceği en son modele bırakılmıştı. Tablo 6 incelendiğinde su deposu değişkenine ilişkin olasılık değeri 0.21014 olarak bulunmuştur. Bu olasılık değerine göre su deposu değişkeni istatistiksel olarak anlamlı bulunmamaktadır. Bu nedenden dolayı su deposu değişkeni modelden çıkartılarak yeni bir model oluşturmalıdır.

Tablo 7: Su Deposu Değişkeni Çıkartıldıktan Sonra Oluşturulan Semiparametrik Regresyon Modeli (Son Model)

Değişkenler	*Parametre Tahmini*	*Standart Hata*	t -İstatistiği	Olasılık
Sabit	127819	45808	2.790	0.00709 **
Salon	3305	1327	2.490	0.01561 *
Oda Sayısı 3	815503	15648	5.209	2.55e-06 ***
Oda Sayısı 4	163031	19439	8.387	1.24e-11 ***
Oda Sayısı 5	262737	31620	8.309	1.67e-11 ***
Banyo	150084	5190	2.642	0.01054 *
Bileşen		F - İstatistiği	Olasılık	
Yatak Odası		4.856	0.000544 ***	
Koridor		3.750	0.005942 **	
Balkon		26.954	< 2e-16 ***	
R^2(adj) = 0.955		Açıklanan Sapma= %96.7	GCV Değeri = 2.5802e+09	

*: Parametrelerin 0.05 anlam düzeyinde istatistiksel olarak anlamlı olduğunu ifade etmektedir.

**: Parametrelerin 0.01 anlam düzeyinde istatistiksel olarak anlamlı olduğunu ifade etmektedir.

***: Parametrelerin 0.001 anlam düzeyinde istatistiksel olarak anlamlı olduğunu ifade etmektedir.

Tablo 7'de, (8) modelinden su deposu değişkeninin çıkartılmasıyla oluşturulan semiparametrik regresyon modelinin sonuçları görülmektedir. Tablo 7'de sonuçları görülen semiparametrik regresyon modeli eşitlik (9)'da gösterildiği gibidir.

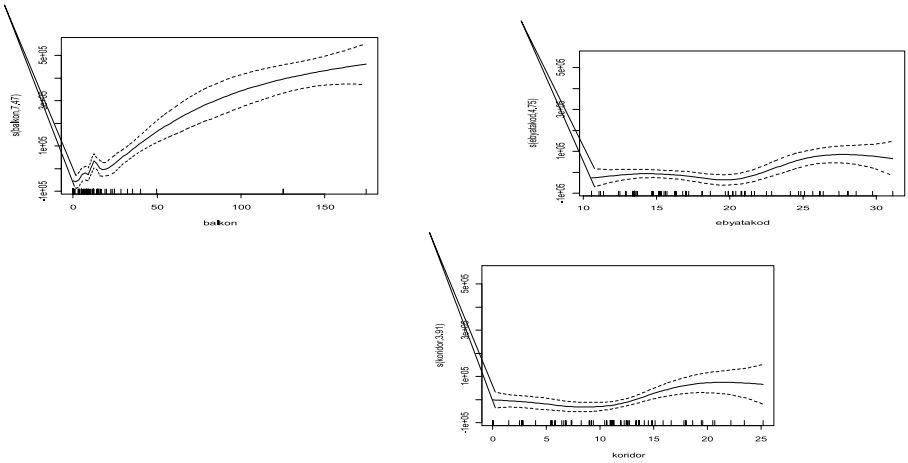
$$FY = 127819 + 815503OD3 + 1630310OD4 + 262737OD5 + 3305SL + 150084BY + s(YO) + s(KR) + s(BN) + \varepsilon \quad (9)$$

Eşitlik (9)'da oda sayısı, salon ve banyo parametrik değişkenler olarak ele alınmış, diğer değişkenler ise parametrik olmayan değişkenler olarak ele alınmıştır. Bilindiği üzere oda sayısı kukla değişken olarak ele alınmıştır. Oda sayısı kukla değişkeni için referans kategori ise, oda sayısı 2 olarak alınmıştır. Tablo 7 veya eşitlik (9) incelendiğinde, değişkenlerin dairelerin satış fiyatı üzerindeki etkileri şöyledir: Diğer değişkenlerin etkisi sabit tutulduğunda, salon metrekaresindeki bir birimlik artış, site içerisindeki dairelerin satış fiyatını ortalama 3305 TL arttırmaktadır. Diğer değişkenlerin etkisi sabit tutulduğunda, Banyo metrekaresindeki bir birimlik artış, site içerisindeki dairelerin satış fiyatını ortalama 150084 TL arttırmaktadır. Diğer değişkenlerin etkisi sabit tutulduğunda, evin üç odalı olması iki odalı olmasına göre, site içerisindeki dairelerin satış fiyatını ortalama 815503 TL arttırmaktadır. Diğer değişkenlerin etkisi sabit tutulduğunda, evin dört odalı olması iki odalı olmasına göre, site içerisindeki dairelerin satış fiyatını ortalama 163031 TL arttırmaktadır. Diğer değişkenlerin etkisi sabit tutulduğunda, evin beş odalı olması iki odalı olmasına göre, site içerisindeki dairelerin satış fiyatını ortalama 262737 TL arttırmaktadır.

Görüldüğü gibi, tüm değişkenlerle site içerisindeki dairelerin satış fiyatları arasında pozitif yönlü bir ilişki vardır. Ek olarak, Parametrik regresyon modellerinde olduğu gibi semiparametrik regresyon modellerinde de sabit terimin yorumlanması büyük önem taşımamaktadır.

(9) eşitliğindeki modelin parametrik olmayan bölümünde bulunan değişkenlerin yorumları ise, model sonucunda oluşan grafikler yardımıyla yapılmaktadır. Düzgünleştirilen fonksiyonların grafikleri Şekil 3'de görülmektedir.

Şekil 3'de sırasıyla balkon, yatak odası ve koridor ve değişkenlerinin (9) modelinden elde edilen grafikleri görülmektedir. Semiparametrik regresyon modellerinde parametrik olmayan bölümün yorumlanması sadece grafik yoluyla yapılmaktadır. Bu durumda, sonuçlar yaklaşık olarak elde edilmektedir. Örneğin, yatak odası 20 metrekare civarında olduğunda, site içerisindeki dairelerin satış fiyatı en düşük, 27-28 metrekare civarında olduğunda ise en yüksek olmaktadır. Bu durumda, yatak odası metrekaresi 17 olduğunda, satış fiyatının ne olacağı, yaklaşık olarak bilinmektedir. Koridor ve balkon metrekareleri içinde grafikler aynı şekilde yorumlanır. Aynı değişkenlerle, parametrik bir regresyon modeli olan çoklu doğrusal regresyon modeli kurulduğunda, Şekil 3'de grafikleri gösterilen değişkenlerin dairelerin satış fiyatları ile doğrusal bir ilişkisi olduğu varsayılacaktı. Bu durumda, yatak odası değişkeninin metrekaresi artıkça satış fiyatının da artması beklenecekti. Şekil 3'de görüldüğü gibi üç değişken içinde dalgalanmalar söz konusu olmaktadır. Yani, bu ilişkilerin doğrusal olduğunu kabul etmek, yanlış sonuçlara ulaşılmasına ve dolayısıyla yanlış kararlar verilmesine yol açacaktır.



Şekil 3: Eşitlik 9'da Görülen Semiparametrik Regresyon Modelinin Parametrik Olmayan Bileşenlerinin Grafikleri

4.6. Son Semiparametrik Regresyon Modeline İlişkin Çıkarımlar

Semiparametrik regresyon modeli parametrik ve parametrik olmayan, olmak üzere iki bölümden oluşmaktadır. Semiparametrik regresyon modelleri için yapılan çıkarımlar iki bölüm için farklı yöntemle yapılmaktadır.

Eşitlik 9 ve sonuçları Tablo 7'de gösterilen model, uygulanmasına karar verilen son modeli ifade etmektedir. Öncelikle, bu modelin R ortamında yazılan bir programla tahmin edilen parametrik katsayılarının anlamlı olup olmadığı incelenmelidir. Tablo 7 incelendiğinde, parametrik katsayılara ilişkin tahminler, t-istatistikleri ve olasılık değerleri görülmektedir. Parametrik regresyonda olduğu gibi t-istatistikleri, parametre tahminlerinin standart hatalara bölünmesi ile elde edilmektedir. t-istatistiklerinin anlamlı olması, incelenen bağımsız değişkenin site içerisindeki dairelerin satış fiyatları üzerindeki etkisinin anlamlı olduğunu ifade eder. Bu testi gerçekleştirmek için sabit katsayı dahil bütün katsayılar için,

$$H_0 : \beta_i = 0$$
$$H_a : \beta_i \neq 0$$

hipotezleri test edilmelidir. H_0 hipotezinin kabul edilmesi, katsayının istatistiksel açıdan anlamlı olmadığını ifade eder. Bu durumda değişkenlerin etkisinin yorumlanması anlamsız olacaktır. Örneğin, daha önce yorumlandığı gibi diğer değişkenlerin etkisi sabit tutulduğunda, salon metrekaresindeki bir birimlik artış, site içerisindeki dairelerin satış fiyatını ortalama 3305 TL arttırmaktadır. Bu şekildeki bir yorum, salon değişkenine ait katsayının istatistiksel olarak anlamsız çıkması durumunda yapılamaz.

Tablo 7 incelendiğinde, Tüm parametrik bağımsız değişkenlerin bağımlı değişken üzerindeki etkisini gösteren parametrik katsayıların istatistiksel olarak anlamlı oldukları görülmektedir. Bu durumda, semiparametrik regresyon modelinin parametrik bölümü için yapılmış olan bütün yorumların geçerli olduğu sonucuna varılır.

Semiparametrik regresyon modelinin parametrik olmayan bileşenine ait çıkarımlar, F -testi ve güven bantları yardımıyla yapılır. F -testi, semiparametrik modele dahil edilecek fonksiyonun, doğrusal olarak ya da pürüzsüz bir fonksiyon olarak mı dahil edilmesinin uygunluğunu her değişken için,

H_0 : Dogrusal Fonksiyon

H_a : Pürüzsüz Fonksiyon$\left[s(x_i\right]$

hipotezleri ile test eder. H_0 hipotezinin kabul edilmesi, ilgili değişkenin doğrusal bir fonksiyon olarak modele dahil edilmesi gerektiğini ifade eder. Bu doğrultuda Tablo 7 incelendiğinde; yatak odası, balkon ve koridor (metrekareleri) değişkenleri istatistiksel olarak önemli bulunmuştur. Dolayısıyla, söz konusu değişkenlerin semiparametrik regresyon modeline düzgünleştirilerek (pürüzsüz bir fonksiyon olarak) dahil edilmesi gerektiğini ifade etmektedir.

Pürüzsüz fonksiyonların anlamlılığı güven bantları ile de incelenebilmektedir. İlgili değişkenler için güven bantları Şekil 3'de gösterilmektedir. İncelendiği gibi bu grafikler, site içerisindeki dairelerin satış fiyatlarının ilgili değişkene göre değişimini göstermektedir. Şekil 3'de üç değişken için kesikli çizgiler güven bantlarını ifade etmektedir. Güven bantlarının, geniş olmayan bir aralıkta ve düz çizgi ile neredeyse aynı doğrultuda olması istatistiksel olarak anlamlılığı ve yapılan tahminlerin güvenilirliğini ifade eder.

Şekil 3'de yatak odası metrekaresinin değişim eğrisi ile güven bantları yüksek metrekarelere kadar aynı doğrultuda ve dar bir aralıktadır. Yüksek metrekarelere gelindiğinde üst sınır yön değiştirmektedir. Ancak bu durum fonksiyonun anlamlılığını bozmamaktadır. Sadece düşük metrekarelerde yapılan tahminlerin çok daha güvenilir olduğunu ifade etmektedir. Koridor metrekaresinin ve balkon metrekaresinin grafiği incelendiğinde aynı durum söz konusu olmaktadır. Ancak, balkon metrekaresi grafiğinde diğer değişkenlerin güven bantlarına göre değişim eğrisinden daha az bir sapma görülmektedir.

Yapılan incelemeler sonucunda semiparametrik modelin, parametrik bölümündeki değişkenlerin ve parametrik olmayan bölümündeki pürüzsüz fonksiyonların istatistiksel olarak anlamlı olduğu sonucuna ulaşılır.

Bu aşamadan sonra, semiparametrik regresyon modelindeki, bağımlı değişkenin bağımsız değişkenlerle yeterince açıklanıp açıklanamadığı incelenmiştir. Bu değişimin ölçüsü, R^2(adj) olarak ifade edilen düzeltilmiş belirlilik katsayısı ve açıklanan sapmadır. Tablo 7 incelendiğinde, düzeltilmiş belirlilik katsayısının %95 olduğu görülmektedir. Bu durumda, site içerisindeki dairelerin satış fiyatındaki değişimlerin %95'i

modeldeki bağımsız değişkenlerce açıklanabildiği sonucuna ulaşılır. Açıklanan sapma değeri %96.7 olarak hesaplanmıştır. Açıklanan sapma değerinin yüksek olması modelin tahmin güvenilirliğinin oldukça yüksek olduğunu ifade etmektedir.

SONUÇ

Uygulamada öncelikle, modele düzgünleştirilerek dahil edilmesi gereken değişkenler belirlenmiştir. Modelde kukla değişkenler olduğundan semiparametrik regresyon analizinin uygulanması uygun bulunmuştur. Semiparametrik regresyon analizinin önemli bir özelliği, bağımlı değişken ile bağımsız değişkenler arasındaki ilişkilerin şeklini istatistiksel testlerle belirleyebilmesidir. Yani, bir değişkenin düzgünleştirilerek, doğrusal ya da dönüştürme yöntemleri ile doğrusallaştırarak modele dahil edileceğine karar verir. Modellerin birbirlerine olan üstünlüğünü test eder. Semiparametrik regresyon analizi ile modelleme yapılmasa bile, bu analiz tekniğini kullanarak değişkenlerin yapılarının belirlenmesi en iyi tahminlere ulaşılmasını sağlayacaktır. Uygulamada, istatistiksel olarak anlamsız olan değişkenler belirlenmiş, semiparametrik regresyon modelleri ile doğrusal, logaritmik ve karesel dönüşümle oluşturulan modeller karşılaştırılmıştır. Uygulanacak son modele karar verilmesi aşamasında su deposu değişkenin istatistiksel olarak anlamsız olduğu tespit edilmiş ve modelden çıkartılarak yeni bir model oluşturulmuştur. Sonuç olarak, salon, banyo metrekaresi değişkenlerinin ve oda sayısı değişkenlerinin semiparametrik modelin parametrik bölümünde, diğer değişkenlerin ise modelin parametrik olmayan bölümünde yer aldığı model en iyi model olarak seçilmiştir. Son semiparametrik modelin, parametrik bölümündeki değişkenlerin anlamlılığı t testleri ile, parametrik olmayan bölümdeki değişkenlerin yani düzgünleştirilerek modele dahil edilmiş değişkenlerin anlamlılıkları ise, F testleri ve güven bantlarıyla incelenmiştir.

KAYNAKÇA

Aydın, D. (2005). Semiparametrik Regresyon Modellemede Splayn Düzeltme Yaklaşımı İle Tahmin ve Çıkarsamalar. Yayınlanmamış Doktora Tezi, Anadolu Üniversitesi, Fen Bilimleri Enstitüsü, Eskişehir.

Bağdatlı, S. (2010). Semiparametrik Regresyon ve Bir Uygulama. Yayınlanmamış Yüksek Lisans Tezi, İstanbul Ticaret Üniversitesi, Fen Bilimleri Enstitüsü, İstanbul.

Box, G.E.P. ve Jenkins, G.M., (1970). "Time Series Analysis: Forecasting and Control", Holden-Day, London.

Çağlayan, E. (2002). Yarı Parametrik Regresyon Modelleri ile Yaşam Boyu Sürekli Gelir Hipotezinin Türkiye Uygulaması. Yayınlanmamış Doktora Tezi, Marmara Üniversitesi, Sosyal Bilimler Enstitüsü, İstanbul.

Fox, J. (2000). Multiple and Generalized Nonparametric Regression. Thousand Oaks: A Sage University Paper.

Hastie, T. & Tibshirani, RJ. (1999). Generalized Additive Models. London: Chapman & Hall.

Keele, L. (2008). Semiparametric Regression For The Social Sciences. Chichester: John Wiley & Sons.

Lee, K.C. (1990). Avoiding Misspecifications and Improving Efficiency in Hedonic and, Consumption Models: Applications of Semiparametric Method. PhD. Thesis London School of Economics and Political Sciences, London. www.r.789695.n4.nabble.com (10.03.2010). www. satilikdairearіyorum.blogcu.com/konut-ev-daire-alirken-dikkat-edilmesi-gerekenler/5189921 (10.03.2010).