**TAŞINMAZ DEĞERLEME VE GELİŞTİRME**

**TEZSİZ YÜKSEK LİSANS PROGRAMI**

# TAG 611 Taşınmaz Değerlemede İstatistiksel Analiz

**Prof.Dr. Mehmet Ali CENGİZ**

**Doç.Dr.Yüksel TERZİ**

**ÜNİTE: IX**

**KORELASYON ANALİZİ - BASİT DOĞRUSAL REGRESYON ANALİZİ**

**İÇİNDEKİLER**

|  |  |
| --- | --- |
| 9.1. Korelasyon  9.2. Kısmi Korelasyon Katsayısı  9.3. Spearman Sıra Korelasyon Katsayısı  9.4. Basit Doğrusal Regresyon Analizi  9.4.1. Parametrelerin En Küçük Kareler Tahmini (EKK)  9.5. Eğrisel Regresyon Modelleri | 3  6  8  10  11  21 |

## 9.1. KORELASYON

Korelasyon analizinde iki veya daha çok sayıda değişken arasında bir ilişki bulunup bulunmadığı, eğer varsa bu ilişkinin derecesi ve fonksiyonel şekli belirlenmeye çalışılır. Örneğin reklamların satışı arttırdığı şeklinde bir düşünce yaygındır. Ancak satışların artışı sadece reklamlar ile açıklanamaz. Nüfus artışı, moda, fiyat rakiplerle rekabet satışları etkileyen diğer nedenler olarak düşünülebilir. Öyle ise reklamlar ile satış arasında ilişkinin olup olmadığı incelenmelidir.

Bir değişkenin değeri artarken diğer değişkenin değeri düzenli artıyor veya eksiliyorsa iki değişkenarasındaki ilişki doğrusaldır. İlişki grafik üzerinden de incelenebilir.

|  |  |  |
| --- | --- | --- |
| istati41 | istati42 | istati43 |
| Korelasyon=+1 | Korelasyon=-1 | Korelasyon=0 |

Doğrusal korelasyonun hesaplanmasında Pearson Momentler Çarpımı korelasyonu kullanılır. Bu formülün uygulanabilmesi için veriler en az aralıklı ölçekle toplanmalı ve süreklilik gösteren nicel bir değişken olmalıdır.



Korelasyon katsayısının değeri -1 ile +1 arasında değişir. Sonucun +1 çıkması iki değişken arasında kuvvetli olumlu ilişkinin bulunduğunu, -1 ise kuvvetli olumsuz ilişkinin bulunduğunu gösterir. Korelasyon katsayısı 0 'a yaklaştıkça ilişkinin kuvveti zayıflar, sıfır ise iki değişken arasında ilişkinin olmadığını gösterir.

**Korelasyon katsayısının önem denetimi:**

Hesaplanmış olan korelasyon katsayısının tesadüfi mi yoksa gerçek bir ilişkiyi mi gösterdiğinin belirlenmesi için denetlenmesi gerekir. Denetim için kurulan hipotezler aşağıdaki gibidir.

H0 : ρ=0 (doğrusal ilişki yoktur);

H1 : ρ> 0 (doğrusal ilişki vardır)

Test istatistiği şu formüle göre hesaplanır,



Serbestlik derecesi (n-2) olup, n çok büyük olduğunda r’nin yaklaşacağı değeri ρ ile gösterelim Bu durumda ρ=0 etrafında korelasyon katsayısının dağılımı simetrik olup, t dağılımına yaklaşır.

**ÖRNEK 9.1.**Aşağıda bir işletmede gün olarak kullanılan izin (X) ile performans puanları (Y) verilmiştir. Bu iki değişken arasında ilişkinin anlamlı olup olmadığını test ediniz. Tablo=1.73

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **X** | **Y** | **X2** | **Y2** | **XY** |
| 1 | 14 | 1 | 196 | 14 |
| 2 | 13 | 4 | 169 | 26 |
| 3 | 12 | 9 | 144 | 36 |
| 3 | 13 | 9 | 169 | 39 |
| 2 | 11 | 4 | 121 | 22 |
| 1 | 12 | 1 | 144 | 12 |
| 4 | 12 | 16 | 144 | 48 |
| 5 | 11 | 25 | 121 | 55 |
| 4 | 14 | 16 | 196 | 56 |
| 3 | 13 | 9 | 169 | 39 |
| 6 | 12 | 36 | 144 | 72 |
| 5 | 12 | 25 | 144 | 60 |
| 10 | 10 | 100 | 100 | 100 |
| 9 | 11 | 81 | 121 | 99 |
| 1 | 14 | 1 | 196 | 14 |
| 8 | 11 | 64 | 121 | 88 |
| 9 | 10 | 81 | 100 | 90 |
| 7 | 9 | 49 | 81 | 63 |
| 6 | 12 | 36 | 144 | 72 |
| 7 | 10 | 49 | 100 | 70 |
| ∑x 96 | ∑y  236 | ∑x2   616 | ∑y2 2824 | ∑xy   1075 |

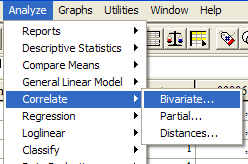
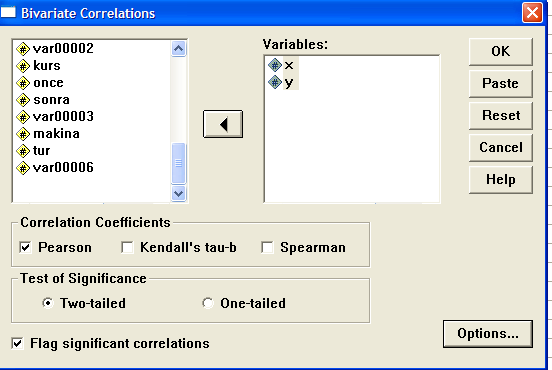


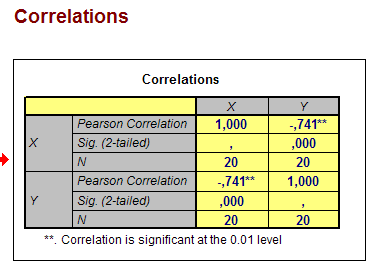
Elde edilen sonuca göre kullanılan izin miktarı ile performans puanları arasında negatif yönlü kuvvetli ilişki vardır. Kullanılan izin miktarı arttıkça performans puanları düşmektedir. Bulunan korelasyonun gerçekten önemli olup olmadığı incelenirse



H0 : ρ =0 ; H1 : ρ > 0 th=> t(20-2), 0.05 =1,73 olduğundan H0 reddedilir. Bulunan korelasyon doğrusaldır-önemlidir ve tesadüfi değildir.

**SPSS ÇÖZÜM:**



**9.2. KISMİ KORELASYON KATSAYISI (Partial Correlation)**

İki değişken arasında korelasyon katsayısının yüksek çıkması muhakkak o değişkenler arasında bir neden-sonuç ilişkisi bulunduğu anlamına gelmez. İki değişken arasında neden-sonuç ilişkisi olmadığı halde korelasyon katsayısı eğer yüksek çıkıyorsa, bu iki değişkenin üçüncü bir değişkenden etkilenmesi söz konusu olabilir.

Kısmi korelasyon katsayısı, iki değişken arasındaki ilişkinin bir yada daha çok değişkenin kontrol edilmesiyle hesaplanan bir değerdir.



rXY,Z : Z değişkeni sabitken (etkisi giderildiğinde), X ile Y arasındaki kısmi korelasyon

rXY : X ile Y arasındaki korelasyon

rXZ : X ile Z arasındaki korelasyon

rYZ : Y ile Z arasındaki korelasyon

Kısmi korelasyon katsayısının anlamlı olup olmadığı F testi ile incelenebilir. Bu testte n gözlem sayısını, k ise tahmin edilen parametre sayısını gösterir.



H0 : ρ =0 (z sabitken x ile y arasında doğrusal ilişki yoktur.)

H1 : ρ > 0 (z sabitken x ile y arasında doğrusal ilişki vardır.)

**Örnek 9.2.** Talep (Y), gelir (X) ve fiyat (Z) serileri arasındaki ilişki araştırılıyor. Fiyatı sabit alarak talep ile gelir arasındaki kısmi korelasyonu bulup ve doğrusal ilişki olup olmadığını araştırınız? Tablo=6.6

Yıllar Gelir(X) Talep (Y) Fiyat(Z)

1. 3 8 4
2. 2 9 7
3. 5 13 11
4. 6 22 18

∑ : 16 52 40

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Yıllar | X | Y | Z | X2 | Y2 | Z2 | XY | XZ | YZ |
| 1995 | 3 | 8 | 4 | 9 | 64 | 16 | 24 | 12 | 32 |
| 1996 | 2 | 9 | 7 | 4 | 81 | 49 | 18 | 14 | 63 |
| 1997 | 5 | 13 | 11 | 25 | 169 | 121 | 65 | 55 | 143 |
| 1998 | 6 | 22 | 18 | 36 | 484 | 324 | 132 | 108 | 396 |
| Toplam | 16 | 52 | 40 | 74 | 798 | 510 | 239 | 189 | 634 |



Z(fiyat) sabitken x (gelir) ile y(talep) arasında ki ilişki aynı yönlü fakat düşük bir ilişki söz konusudur.

H0 : ρ =0 (z sabitken x ile y arasında doğrusal ilişki yoktur.)

H1 : ρ > 0 (z sabitken x ile y arasında doğrusal ilişki vardır.)



Fh=0,2<Ftablo=6,6 Ho kabul edilir. Yani z sabitken x ile y arasında doğrusal ilişki yoktur.

**9.3. SPEARMAN SIRA KORELASYON KATSAYISI**

Doğrusal korelasyonda ilişkisi araştırılan değişkenlerin nicel ve normal olması gerekir. Bu varsayımlar sağlanmadığında Spearman Sıra Korelasyon Katsayısı kullanılır.

Sıra korelasyon katsayısının hesaplanmasında önce gözlem değerleri büyükten küçüğe (veya küçükten büyüğe) doğru sıralanır ve bun sıralamaya göre sıra numarası verilir. Aynı olan gözlemlerin sıra numaraları verilirken sıra no ortalaması alınır.



Di : X ve Y’nin sıra numaraları arasındaki fark

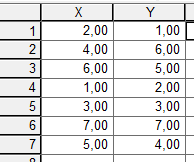
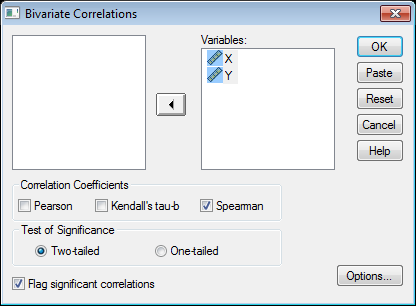
n : Gözlem sayısı

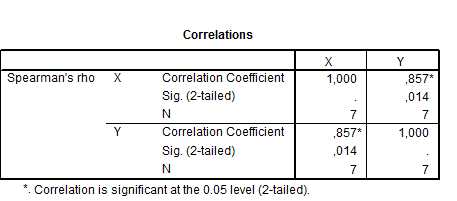
**Örnek 9.3.** 7 öğrencinin boy ve ağırlıklarının büyükten küçüğe doğru sıra puanları aşağıda gösterilmiştir. Veriler normal dağılım göstermediğine göre, boy ve ağırlıklar arasındaki ilişkiyi hesaplayınız?

|  |  |  |  |
| --- | --- | --- | --- |
| X | Y | D | D2 |
| 2 | 1 | 1 | 1 |
| 4 | 6 | -2 | 4 |
| 6 | 5 | 1 | 1 |
| 1 | 2 | -1 | 1 |
| 3 | 3 | 0 | 0 |
| 7 | 7 | 0 | 0 |
| 5 | 4 | 1 | 1 |

Öğrencilerin boy uzunlukları ile ağırlıkları arasında aynı yönde önemli bir ilişki vardır.

**SPSS ÇÖZÜM:**



## 9.4. Basit Doğrusal Regresyon Analizi

## Regresyon analizinin temelinde; gözlenen bir olayın değerlendirilirken, hangi olayların etkisi içinde olduğunun araştırılması yatmaktadır. Bu olaylar bir veya birden çok olacağı gibi dolaylı veya direkt etkileniyor da olabilirler.

## Regresyon analizi yapılırken, gözlem değerlerinin ve etkilenilen olayların bir matematiksel gösterimle yani bir fonksiyon yardımıyla ifadesi gerekmektedir. Kurulan bu modele regresyon modeli denilmektedir. Regresyon analizi incelenirken, genellikle konusunu oluşturan, etkilendiği olaylara değişkenler adı verilir ve bu değişkenlerin yer alacağı matematiksel model incelenir.

İstatistikte değişkenler arasındaki ilişkinin derecesini gösteren katsayıya korelasyon, değişkenler arasındaki ilişkinin fonksiyonel şeklini belirleyen denkleme ise regresyon denklemi denir.

Regresyon modelinin kullanılması, ilgilenilen olayla ilgili olarak, bir sebep-sonuç ilişkisi bulunması gerekmektedir. Örneğin 1990-1997 yılları arasındaki hisse senedi fiyatlarını incelersek, seçilen zaman aralığında bir matematiksel model kurma gereği vardır ve bu modelde bir sebep, sonuç ilişkisi aranmaktadır. Sebep, hisse senedinin fiyatını yükselten veya düşüren unsurlardır. Faiz oranları, ekonomik nedenler, enflasyon oranları vs. olarak incelenebilir. Sonuç ise hisse senedinin fiyatının değişmesidir.

Sebep-sonuç ilişkisi, regresyon modeli kurulurken, bağımlı ve bağımsız değişkenler olarak anlatılmaktadır. Yukarıdaki hisse senedi fiyatı sonuç olan bağımlı değişken, faiz oranları, ekonomik nedenler, enflasyon oranları vs. sebep olan bağımsız değişkenlerdir.  
Regresyon analizi yapılırken kurulan matematiksel modelde yer alan değişkenler bir bağımlı değişken ve bir veya birden çok bağımsız değişkenden oluşmaktadır.

Bağımsız değişkenler kurulacak modelde bir değişkenli olarak ele alınırsa basit doğrusal regresyon, birden fazla bağımsız değişkenli olarak alınırsa çoklu regresyon modeli konusunu oluşturmaktadır.

**Sebep (X) Sonuç (Y)**

Gelir Harcama

Yaş Boy

Gübre Verim

Yem miktarı Süt miktarı

Çalışma süresi Alınan not



Regresyon modelinde Y ve ε rasgele değişkenler, X ise rasgele değişken değildir.

Regresyon modelinde Xi’lerin hatasız ölçüldüğü, Y’nin ise belli bir hata (εi) miktarı ölçüldüğü varsayılır.

X ile Y arasında doğrusal ilişkinin olması regresyon hattının bir doğru şeklinde olmasından çok parametrelerin modelde doğrusal bir ilişki içinde olması ile ilgilidir.





β0: Doğrunun Y eksenini kestiği nokta. X’in sıfır değerine karşılık Y’nin alacağı değer.

β1: Eğim, hız, regresyon katsayısı. X’in 1 birim değişmesi halinde Y’nin kaç birim ve ne yönde değişeceğini gösterir.



Doğrunun merkezden (orjinden) geçmesini test eder.

Doğrunun X eksenine paralel olup olmadığını (eğimin önemli olup olmadığını) test eder.

**9.4.1. Parametrelerin En Küçük Kareler Tahmini (EKK)**

Basit Doğrusal Regresyon modelinin parametreleri β0 ve β1 bilinmeyenlerdir. Gözlenen veri yardımıyla bunların tahmin edilmesi gerekir.

   Parametre tahmini en küçük kareler metodu (Least Squares Method) kullanılarak yapılır. Bu yöntemde esas, dağılım diyagramında (scatter diagram) görülen tüm noktalar için doğruya uzaklıklarının bulunması ve bunların toplamının minimize edilmesidir. Ancak regresyon analizinde bu toplam fonksiyonu daima sıfır olacağından β0 ve β1 değerlerini bulmada kullanılamaz. Gerçek değerler Yi, tahmin değerleri ve bu değerler arasındaki fark artıklar ( e) bulunur.

,



Normal denklemler



Varyans Analizi Tablosu

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Değişim Kaynağı | Serbestlik Derecesi  (sd) | Kareler  Toplamı  (KT) | Kareler  Ortalaması  (KO) | F  Testi |
| Regresyon | 1 | RKT= | RKO= RKT/1 | RKO/AKO |
| Artık | n-2 | AKT= | AKO=AKT/(n-2) |  |
| Genel | n-1 | GKT= |  |  |



: Tahmin edilen regresyon modelinin standart hatası



## *BASİT DOĞRUSAL REGRESYON MODELİ VARSAYIMLARI*

1. Bağımsız değişkenin değerleri sabit kabul edilir. Bağımlı değişkenin değerleri ise rasgeledir. Fakat korelasyon analizinde her iki değişkende tesadüfidir.

2. Değişkenler hatasız ölçülmüştür.

3. Her Xi değeri için;

Yi değerleri birbirinden bağımsızdır,

Yi gözlemlerinin tüm dağılımları normaldir,

Yi gözlemlerinin tüm dağılımları aynı varyansa sahiptir.

4. Bu alt küme değerlerinin (Y’lerin) varyansları eşittir.

5. Bağımlı ve bağımsız değişkenler arasındaki ilişki doğrusaldır.

**Hata Terimi (εi) Varsayımları**

1. Rassal hataların beklenen değeri sıfırdır. E(ε)=0

2. ε’ların olasılık dağılımının varyansı sabittir.

3. Hata değerleri birbirinden bağımsızdır.

4. Rassal hataların dağılımı normaldir.

5. Hatalar ile bağımlı değişken arasında korelasyon yoktur.

Cov (ε,Y)=0

6. Hatalar ile bağımsız değişkenler birbirinden bağımsızdır.

Cov (ε,Xi)=0

**R2 (coefficient of determination):**

Determinasyon veya belirlilik katsayısıdır. Bağımsız değişken X’in regresyon modeli ile bağımlı değişken Y’yi ne kadar açıklayabildiğini görmek için bir ölçüt olan karar katsayısı, R*²* kullanılır. R2 örneğin açıklanabilen değişkenliğin toplam değişkenlik içinde oranı olarak tanımlanır. R2 değeri verilerdeki değişkenliğin açıklanmasında regresyon denkleminin başarısının bir ölçüsü olarak kullanılır. Regresyon Modelinin performansı R2, karar katsayısı ile ölçülür.



R2 1’e ne kadar yakınsa, regresyon o kadar anlamlıdır, belirleyicidir. Bu konuda bilimsel bir karar verebilmek için hipotez testi yapılır. Belirlilik katsayısı bağımsız değişkenin bağımlı değişkenin davranışını açıklamadaki başarı ölçüsünü veren tanımlayıcı istatistiktir. Bağımsız değişken sayısı k, n gözlem sayısına oranı büyükse regresyon katsayısının temsil gücü azalır. Bu tür durumlarda düzeltilmiş regresyon katsayısı kullanılır.

Tahmin edilen regresyon modelininin katsayıların güven aralığı



Tahmin edilen regresyon modelinin ve katsayıların önemliliği testi



H0 : Regresyon modeli anlamlı değildir.

H1 : Regresyon modeli anlamlıdır.



Ho red edilir. Tahmin edilen regresyon modeli önemlidir.

**Örnek 9.4.** Bir ürüne ait reklam harcamaları ve yapılan satış miktarları aşağıdaki gibi bulunmuştur.

a) Veri seti için matematiksel modeli tahmin ediniz?

b) ANOVA tablosunu oluşturunuz?

c) Tahminin standart hatasını bulunuz?

d) Belirtme katsayısını bulup, sonucu yorumlayınız?

e) 60 (bin TL) reklam harcaması yapılırsa tahmini satış miktarını tahmin ediniz?

Reklam harcama (bin TL): 10 20 30 40 50

Satış (bin adet) : 3 4 6 7 10



Çözüm:

a)



b)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Değişim Kaynağı | Serbestlik Derecesi  (sd) | Kareler  Toplamı  (KT) | Kareler  Ortalaması  (KO) | F  Testi |
| Regresyon | 1 | RKT= =28,9 | RKO= RKT/1 =28,9 | F=RKO/AKO =78,75 |
| Artık | n-2=5-2=3 | AKT= =1,10 | AKO=AKT/(n-2) =0,367 |
| Genel | n-1=5-1=4 | GKT= =30,0 |  |



c) Tahminin standart hatası= =0.61

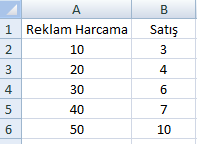
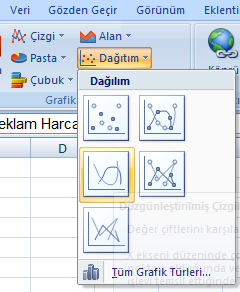


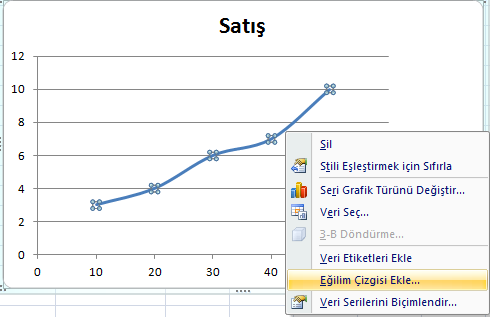
d)

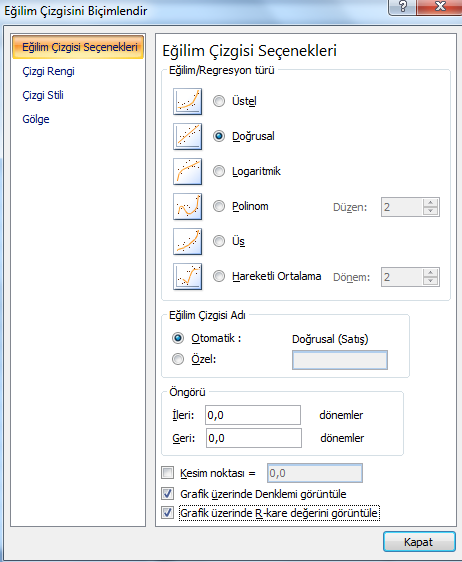


e)

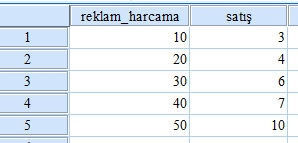
**EXCEL ÇÖZÜM**

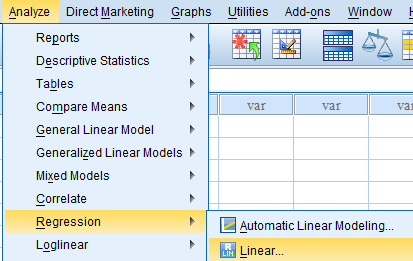
 

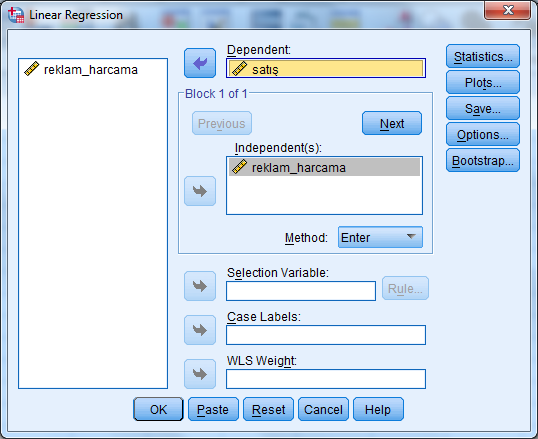


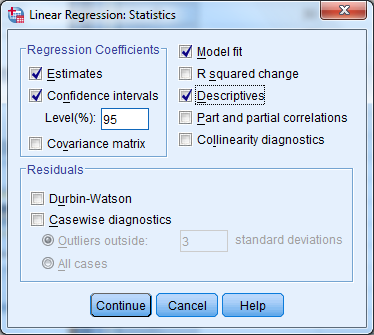


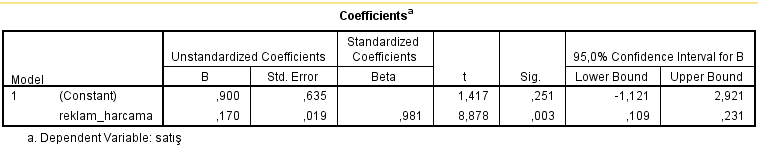
**SPSS ÇÖZÜM:**

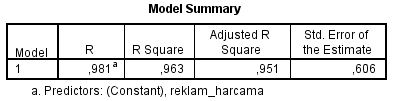


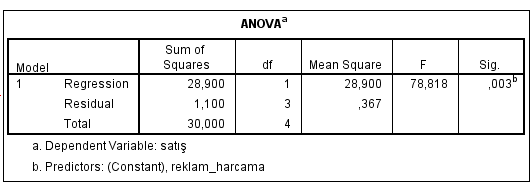


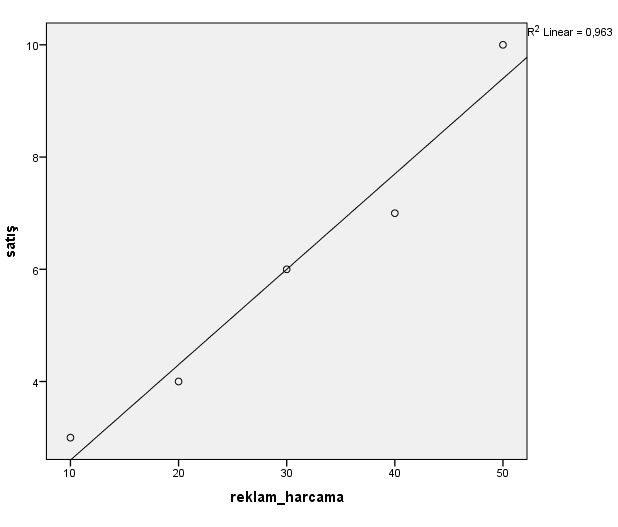












**9.5. Eğrisel Regresyon Modelleri**

* Doğrusal (Linear) model : 
* Kuadratik (Quadratic) model: 
* Kübik (Cubic) model : 
* Compound model : 
* Power model : 
* Inverse model : 
* Exponential model : 
* Growth model : 
* S Model : 
* Logarithmic model : 

**Örnek 9.5.** Bir bölgedeki konutların büyüklüklerine göre fiyatları aşağıdaki gibi ölçülmüştür. Bu veri setine en uygun regresyon modelini tahmin ediniz.

