83669

ATATÜRK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
ZOOTEKNİ ANABİLİM DALI

NONPARAMETRİK REGRESYON METODLARININ İNCELENMESİ

Mehmet TOPAL

Yönetici:Prof.Dr. Necati YILDIZ

Yüksek Lisans Tezi

ÖZET

Günümüzde, veri analizlerinde uygulanan parametrik metodlara karşı gelen birçok nonparametrik metot geliştirilmiştir. Bu metodlardan birisi de değişkenler arasındaki ilişkinin fonksiyonel şekli olan nonparametrik regresyondur. Bu çalışmada, basit regresyon doğrusunun Y eksenini kestiği nokta olan α'yı ve doğrunun eğimi olan β'yı tahmin etmek için geliştirilmiş olan nonparametrik regresyonda, medyana göre parametre tahminleri ile smoothing tahminler incelenmiştir. Araştırılan bu metodların bazıları Atatürk Üniversitesi koyunculuk işletmesinde yetiştirilen kuzulardan elde edilen verilere uygulanmış ve medyana göre regresyon parametreleri tahmininin smoothing tahminlerden daha kolay yapılabildiği ve daha sağlam sonuç verdiği tesbit edilmiştir.

SUMMARY

Many nonparametric methods corresponding to parametric methods applied in data analysis have been developed recently. One of these methods is nonparametric regression which is a functional form of relation among variables. In this research, the estimations of α and β according to smoothing and median methods proposed in the nonparametric regression were examined to estimate the α, the point on the Y axis intersected by simple regression line, and β, the slope of the line. Some of these methods were applied to data obtained from the lambs raised at agricultural farm of Atatürk University. Estimation of regression parameter according to median was easier to apply and more robust than smoothing estimations.

TEŞEKKÜR

Bu çalışmayı yapmada; danışmanlığımı üstlenerek bu tezin planlanmasında ve yürütülmesinde katkılarını esirgemeyen , eğitici ve sabırlı kişiliğiyle benim yetişmemi sağlayan değerli hocam sayın Prof. Dr. Necati Yıldız'a teşekkürlerimi bir borç bilirim. Ayrıca, bölümümüzde uyumlu bir çalışma sistemi kurmuş bulunan ve yetişmemde büyük bir payı bulunan değerli bölüm başkanı sayın Prof. Dr. Hakkı Emsen ve Anabilim Dalı başkanı sayın Prof. Dr. Fatin Sezgin'e teşekkürü bir borç bilirim. Tez çalışmamda katkılarını esirgemeyen sayın Dr. Ömer Cevdet Bilgin'e ve gerekli fiziki şartları sağlayan dekan Prof. Dr. A.Vahap Yağanoğlu ve Enstitü Müdürü Prof. Dr. Y.Kemal Yoğurtçu'ya teşekkür ederim.

Ayrıca, beni bugünlere getiren aileme ve varlıklarından güç aldığım tüm arkadaşlarıma teşekkür ederim.

Erzurum,1999 Mehmet Topal

İÇİNDEKİLER

1. GİRİŞ

Nonparametrik istatistik metodlar 18. yüzyıldan itibaren geliştirilmeye başlamıştır. Parametrik olmayan (Nonparametrik) istatistik teknik olarak kabul edebileceğimiz ilk teknik 1710 yılında John Arbuthnot tarafından işaret testi olarak literatüre tanıtılmıştır. Nonparametrik testler 1940'lı yıllara kadar pek yaygın olarak uygulanmamıştır. Nonparametrik kelimesi, ilk kez 1942 yılında Wolfowitz tarafından yayınlanan bir makalede ifade edilmiştir. Fazla ve kuvvetli varsayımlara bağımlılığı azaltmak ve daha genel bir takım sonuçlara ulaşmak için, özellikle 1942'den sonra hızlı bir gelişme gösteren nonparametrik metodlar günümüzde pek çok araştırma sahasında uygulanmaktadır.

Griffin (1962), nonparametrik metodları, populasyon hakkında katı faraziyeler öngörmeyen testler olarak ifade etmiştir.

Yıldız ve Bircan, (1994), Nonparametrik Metodları, normal dağılış faraziyelerinin geçerli olmadığı durumlarda dağılıştan ve parametreden bağımsız olarak uygulanabilen metodlar olarak tarif etmişlerdir.

Bilinmeyen dağılım fonksiyonu sürekli olan veya onun üzerinde herhangi bir bilgiye gerek duymayan metodlar parametrik olmayan metodlardır. Ayrıca parametrik olmayan test, populasyon parametrelerinin değerleri üzerine hiç bir varsayım yapılmadığı durumlarda uygulanabilir (Aytaç, 1984).

Parametrik olmayan istatistik metodların özelliklerini aşağıdaki gibi özetleyebiliriz:

1) Parametrik olmayan istatistik metodlar, populasyon hakkında genellikle sayısı çok az olan varsayımlara bağlıdır. Bu durumlarda parametrik olmayan testlerin kullanılabilmesi için populasyon dağılımının sürekli olması yeterlidir (Fraser, 1958).

2) Parametrik olmayan metodlar, adlandırma (nominal) ve dereceleme (ordinal) ölçüm ıskalalarına göre elde edilen verilere uygulanabilirler.

3)Gözlemlerden elde edilen veriler genellikle bir sıralama veya işarete dönüştürüldükten sonra test edilir.

Bu özellikler altında parametrik olmayan istatistik metodların parametrik istatistik metodlardan üstün ve zayıf yönleri vardır. Nonparametrik metodların parametrik metodlara göre üstün yönleri;

1) Çoğu nonparametrik istatistik metodlar en az faraziyeye dayandığından yanlış uygulanma ihtimali düşüktür.

2) Örnek büyüklüğü çok küçükse (n = 6 gibi) parametrik olmayan istatistik metodlar daha sağlam sonuç verir.

3) Matematik ve istatistik konularında sınırlı düzeyde bilgiye sahip araştırmacılar parametrik olmayan metodları kolayca öğrenip uygulayabilirler.

4) Parametrik olmayan istatistik metodlar, verilerin basitce sınıflara ayırılarak incelenmesinde çok geçerlidir. Verileri sıralayıcı bir çerçeve içerisinde ele alır. Yani adlandırma ve sıralama verileri gibi zayıf bir ölçekle ölçülmüş verilere uygulanabilir.

Parametrik olmayan metodların parametrik metodlara karşı en önemli dezavantajı; parametrik metodların kullanılabilir olduğu durumlarda yani, parametrik testin uygulanmasındaki varsayımlar geçerli olduğu zaman, kullanılması durumunda testin gücünün parametrik testlerin gücünden daha küçük olmasıdır (Aytaç, 1984). Başka bir ifade ile;

$$G_{\text{parametrik olmayan}} = (1 - \beta_n) < G_{\text{parametrik}} = (1 - \beta_p)$$

şeklinde belirtilebilir. Parametrik olmayan metodların diğer bir dezavantajı ise örneklerin geniş hacimde olması ve özellikle hesaplamaları yapmak için elimizde bilgisayar bulunmaması durumunda, elle yapılması gereken hesaplamaların uzun ve yorucu olmasıdır.

Günümüzde, verilere uygulanan parametrik istatistik metodlara tekabül eden birçok nonparametrik istatistik metod geliştirilmiştir. Bu metodlardan birisi de değişkenler arasındaki ilişkinin fonksiyonel şekli olan nonparametrik regresyondur.

İki değişken arasındaki fonksiyonel ilişki bir matematik formülle ifade edilir. Eğer x değişkeni bağımsız ve y değişkeni bağımlı ise değişkenler arasındaki fonksiyonel ilişki;

$$Y = f(x)$$

şeklinde olur. Fonksiyon f, x' in verilen bir belirli değerine Y' nin karşı gelen değerini kısaca tanımlar.

Dolayısıyla bir fonksiyonel bağıntı olan regresyon; bağımsız değişkendeki değişmelerin bağımlı değişkeni hangi yönde ve ne miktarda etkilediğini belirler. Regresyonda asıl amaç değişkenler arasındaki fonksiyonel bağıntıyı en iyi ifade edecek matematik denklemi bularak, bu denklemi istatistik analizlerde ve bağımlı değişkenin değerlerini tahminde kullanmaktır. Bir regresyon denklemi bir bağımsız değişken içeriyorsa basit regresyon eşitliği, birden fazla bağımsız değişken içeriyorsa çoklu regresyon denklemi olarak adlandırılır (Ross, 1987).

Başlıca regresyon denklemleri;

1) $Y = \alpha + \beta x$	Doğrusal (linear)
2) $Y = \alpha + \beta_1 x + \beta_2 x^2$	İkinci dereceden (quadratic)
3) $Y = \alpha + \beta_1 x + \beta_2 x^2 + \beta_3 x^3$	Üçüncü dereceden (cubic)
4) $Y = \alpha + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \beta_4 x^4$	Dördüncü dereceden (quartic)

şeklinde ifade edilir (Davis, 1973). Bunlardan basit linear regresyon modeli, bağımsız değişkenin değeri ve ortalama cevap arasında linear ilişkiyi varsayar. Basit linear regresyon modeli;

$$Y=\alpha+\beta x+e$$

şeklindedir. Burada x bağımsız değişken, Y cevap değişkeni ve e; ortalaması 0, varyansı σ^2 olan bağımsız ve özdeş dağılmış şansa bağlı hataları temsil eder. Ayrıca α regresyon doğrusunun Y eksenini kestiği noktayı, β' da doğrunun eğimini ifade eder.

Hardle'e (1997) göre; bir regresyon eğrisine nonparametrik yaklaşımın dört amacı vardır. Birincisi; iki değişken arasındaki genel bir ilişkiyi inceleyen çok yönlü bir metod sağlar. İkincisi, bir sabit parametrik modelin varlığı olmaksızın gözlemlere ilişkin tahminlerin yapılmasını sağlar. Üçüncüsü, izole edilmiş noktaların etkisini hesaba katarak sahte gözlemlerin bulunması için bir araç temin eder. Dördüncüsü, ardışık x değerleri arasında interpolasyon yapmak ya da kayıp değerleri yedeklemek için esnek bir metod ortaya koyar.

Nonparametrik regresyon, verilerin gerçek doğal yapısını elde etmek için basit ve faydalı bir araçtır (Chu and Marron, 1991).

1950'den itibaren gelişmeye başlayan ve 1968-1990 tarihleri arasında yoğun bir şekilde araştırılan nonparametrik regresyon konusu henüz tam olarak gelişmesini tamamlayamamıştır. Bu konu hakkında yabancı literatürde yapılmış birçok araştırma olmasına rağmen yine de konu birçok yönüyle henüz tam olarak açıklığa kavuşmamıştır. Türkiye'de bu konu hakkında belirgin olarak yapılmış herhangi bir çalışma bulunmamaktadır. Bu sebeple bu konu hakkındaki bilgileri Türkçe literatüre kazandırmak amacıyla bu çalışma planlanmıştır.

Adichie (1967,a,b), α ve β'nın nokta tahminlerini hesaplamış ve bu tahminlerin etkinliklerini Brown-Mood medyan tahminler ile mukayese etmiş ve medyan tahminlerinin etkinliğinde bazı kayıpların olduğunu belirtmiştir. Ayrıca $H_0{:}\alpha=0$ ve $H_0{:}\beta=0$ hipotezleri için rank puan testlerinin bir sınıfını teklif etmiştir.

Cryer, et al. (1972), her bir reel sayı x'in [0,1] aralığında olduğunu farzederek, ortalama $\mu(x)$ ve medyan $m(x)$'in nonparametrik tahmin edicilerini tanımlamışlar ve $m(.)$'nın monoton olan $\hat{m}(.)$ tahmincisini tartışmışlardır.

Hill (1962), konveksliğin (dışbükeyliğin) alternatifine karşı bir medyan regresyon eğrisinin linearliğinin bir testini önermiştir. Doğrusallığın önerilen testi, Brown-Mood metodu vasıtasıyla (X_i,Y_i) noktalarının bir setinde (kümesinde) tahmin edilmiş bir doğru dayanağında tarif edilmiştir.

Hogg and Randles (1975), serbest dağılımlı istatistiği kullanarak $H_0:\beta=\beta_0$ hipotezini test etmişlerdir. Kullanılan test istatistiği $K(\beta_0)=\sum c(Q_i)a(R_i)$ dir. Burada c(.) ve a(.) artan puan (score) fonksiyonlar ve Q_i, x_i'nin R_i' de $y_i-\beta_0 x_i$'nin rankıdır.

Hussain and Sprent (1983), bir regresyon doğrusu eğiminin non-parametrik tahmincilerinin sağlam olduğunu belirtmişler ve $x_j \neq x_i$ olmak üzere $(x_i,y_i),(x_j,y_j)$ nokta çiftleri ile bileşen doğrularla ilgili bütün eğim çiftleri dizisinin $b_{ij}=\frac{(y_j-y_i)}{(x_j-x_i)}$ şeklinde olduğunu ifade etmişlerdir. Ayrıca b_{ij} eğim çiftleri takımından (kümesinden) β'nın non-parametrik tahmincilerinin birkaçını da tanımlayarak bunların birbirlerinden farklarını kısaca belirtmişlerdir.

Jaeckel (1972), kendi ismini verdiği tahmincinin $(Y_j-Y_i)/(c^j-c^i)$ eğim çiftlerinin ağırlıklı medyanı olduğunu ifade etmiştir. Aynı zamanda linear modelde regresyon katsayılarının tahmini problemine cazip bir yaklaşımla, residualların mümkün olduğu kadar küçültülmesiyle meydana gelen katsayıların değerlerini bulmuş ve bu durumu çoklu regresyonda incelemiştir.

Jureckova (1971), uygun rank istatistiklere dayandırılan regresyon parametre vektörünün bir tahmininin yapısını çoklu regresyon durumlarında araştırmış; α ve β'nın tahminini hesaplamıştır.

Kildea (1981), enküçük kareler tahmincisinin bir doğal medyan benzerinin elde edilmesi ile nonparametrik regresyon parametrelerini tahmin metodu olan Brown-Mood tahmincisinin nasıl değiştiğini göstermiştir

Rao and Gore (1982), nonparametrik prosedürlerin basit adaptasyonlarıyla test edilebilen bir-örnekli, iki-örnekli ve birkaç-örnekli linear regresyonlarda kesişim parametresi hakkındaki çeşitli hipotezleri izah etmişlerdir.

Maritz (1979), $v_i = {}^{y_i}\!/_{x_i}$ ve $u_i = {}^{1}\!/_{x_i}$ olması durumunda $v_i = \alpha u_i + \beta$ yapısında tekrar $y_i = \alpha + \beta x_i$ nin yazılmasıyla teşekkül eden bir alternatif tahminci önermiştir. Sonra Maritz u üzerine v'nin regresyonu ile α eğimini tahmin etmek için nonparametrik regresyon parametrelerini tahmin metodu olan Theil metodunu uygulamıştır.

Potthoff (1974), iki basit regresyon doğrusunun parelel olup olmadığının bir nonparametrik testini incelemiş ve uygulanan nonparametrik test iki-örnek Wilcoxon testi ile benzer olarak geliştirilmiştir.

Sen (1968), Kendall'ın tau'sundan esinlenerek β'nın basit ve sağlam tahminci üzerinde çalışmıştır. Nokta tahmincisini, $x_i \neq x_j$ ile noktaların $(y_j-y_i)/(x_j-x_i)$ eğim çiftleri dizisinin (kümesinin) medyanı olarak tarif etmiştir. Sen, ileri sürdüğü tahmincinin çeşitli özelliklerini incelemiş ve kendi ismini verdiği metodu en küçük kareler ve diğer nonparametrik tahminciler ile mukayese etmiştir.

Sievers (1978), basit linear regresyon modelinde β parametresinin tahminin ve güven sınırlarını incelemiştir. Önerdiği rank prosedürler; eğim çiftleri için kullanılan ağırlıklarla, Sen ve Theil'in prosedürlerinin genişletilmişidir.

Hardle and Mammen (1993), parametrik regresyon uygunluğuna karşı nonparametriği mukayese etmişler. "Regresyon verilerinin parametrik modelinin uygunluğu, bir nonparametrik smoothing tahmincisi ile mukayesesiyle değerlendirilebilir" şeklinde ifade etmişlerdir.

Rosenblatt (1971), $n \to \infty$ için $h(n) \downarrow 0$ bandwithların bir dizisi (sırası) ve $\int w(u)du = 1$ ile integrali alınabilir sınırlı ağırlık fonksiyonu w(u) nun seçimi vasıtasıyla tahminlerin bir basit dizisini determine etmiş ve kendi ismini verdiği Kernel tahminciyi geliştirerek bu tahmincinin özelliklerini incelemiştir. Eddy'de (1980), en çok elde edilen değerin bir kernel tahmini, f_n'nin kernel tahmininin maksimize edilmiş herhangi bir değeridir. "Bir ihtimal yoğunluk fonksiyonu olan f(t) nin ençok elde edilen bir değeri maksimize edilmiş f olan t'nin bir değeridir" şeklinde ifade etmiş, en çok elde edilen değerin optimum kernel tahmincilerini incelemiş ve bu çalışmasında Rosenblatt kernel yoğunluk tahmincisinin çeşitli özelliklerini açıklamıştır.

Parzen (1962), ihtimal yoğunluk fonksiyonunun tahminini göstermiş ve bu tahmine Parzen kernel yoğunluk tahmincisi demiştir. Ayrıca bandwith h'ın nasıl seçileceğini de kısaca açıklamıştır. Rice (1984), nonparametrik regresyon için bandwith parametre seçimi incelemiş ve bu çalışmasında Fourier seriler tahmin analizini yapmıştır. Ibragimov ve Hasminskii (1980), regresyonun nonparametrik tahmini üzerine bir araştırma yapmışlardır.

Priestley ve Chao (1972), iki değişken arasındaki uygun bir genel foksiyonel ilişkinin problemini incelemişlerdir ve kendi isimlerini verdikleri kernel smoothing tahmincisini geliştirmişlerdir. Benedetti (1977),bir bilinmeyen regresyon fonksiyonunun tahmini için Priestley ve Chao tarafından önerilen nonparametrik tekniği mütlaa etmiş ve bu tahmin metodunu ilerletmeye çalışmıştır. Ayrıca bu çalışmasında çeşitli kernellerin fonksiyonel yapısını kısaca göstermiştir. Hardle ve Gasser (1984), Priestley-Chao kernel tahmincisini ve Gasser-Müller kernel ağırlıklarını kullanarak sağlam (robust) nonparametrik fonksiyon uygunluğu incelemişlerdir. Gasser, et al (1984), regresyon fonksiyonların ve türevlerinin nonparametrik tahmini için kernellerin seçimini araştırmışlardır.

Hall ve Wehrly (1991), Nadaraya ve Watson kernel tahmincisinden faydalanarak bir kernel tahmincisi geliştirmişlerdir ve bu tahminciyi bir simulasyon çalışmasında Nadaraya-Watson ve Gasser-Müller kernel tahmincilerle mukayese etmişlerdir. Ayrıca çalışmada bandwith parametresi h'ın seçiminin nasıl yapılacağını göstermişlerdir. Hall ve Marron (1990), nonparametrik regresyon modelinde varyansın bir tahmincisini önermişlerdir.

Györfi (1981), k_n-NN regresyon tahminini işlemiş ve k_n-NN regresyon tahminlerinin yakınsama oranını göstermiştir. Devroye (1978), nearest neighbor (en yakın komşu) regresyon fonksiyon tahmincilerinin üniform yakınsamasını ve optimizasyondaki uygulanışını incelemişlerdir. Stone (1977), nearest neighbor kuralının dağılıştan bağımsız olduğunu ispat etmiş ve nearest neigbor ağırlıklarını incelemiş. Cover and Hart (1967), nearest neighbor kuralını tanıtarak kabul olunabilirliğini, ve sürekli durumlar altında bir-nearest neighbor ve k-nearest neighborların yakınsamasını incelemişlerdir.

Stute (1984), m(x)=E(Y/X=x) regresyon fonksiyonunun smooth NN tahmininin Nadaraya-Watson tahmini için gerektiğinden daha zayıf durumlar altında asimtotik olarak normal olduğunu göstermiştir. Ayrıca x'in komşuluğunda çok az gözlem olduğunda nearest neighbor tahminlerin kernel tip tahminlerden daha etkili olduğunu ifade etmiştir.

Clark (1976), bir smoothing regresyon fonksiyonunun non-parametrik tahminini incelemiştir. Silverman (1984), eğri tahmini ve nonparametrik regresyon için spline smoothing yaklaşımını mütalaa etmiştir ve spline smoothing ile kernel smoothing arasındaki ilişkiyi araştırmıştır. Reinch (1967), spline fonksiyonlarla smoothing durumunu incelemiştir. Simonoff (1998), yoğunluk tahmini ve nonparametrik regresyon arasındaki bir bağ olarak katagorik veri smoothingin durumunu araştırmıştır. Speckman (1985), $Y_i=f(x_i)+e_i$, $x_i \in (a,b)$ formunun nonparametrik regresyon modellerinde linear tahminini tartışmıştır.

2. REGRESYON PARAMETRELERİNİ TAHMİN METODLARI

2.1. PARAMETRİK METODLAR

2.1.1. En Küçük Kareler Tahmin Metodu

En küçük kareler metodu, doğruyu; noktaların kendisinden ayrılışlarının kareleri toplamını minimum yapacak şekilde tayin etme esasına dayanmaktadır. Yani hataların karelerine ilişkin toplamın en küçük olması temeline dayanır. (X_i, Y_i) gözlemlerinin

$$Y_i = \alpha + \beta X_i , \qquad i = 1,2, \ldots, n$$

regresyon modeline uygun olduğu farzedilsin.

En küçük kareler tahmincisi $\hat{\beta}$, $W_{ij} = (X_j - X_i)^2$ ağırlıkları ile b_{ij}' nin tartılı (ağırlıklı) ortalamasıdır. Medyan tahmincileri en küçük karelerde uygun ağırlık metodlarını kullanarak, kesinlikle sapan gözlemlerin (outlier) etkisinin öneminin azaltılmasının mümkün olacağını belirtmişlerdir (Hussain and Sprent, 1983).

En küçük kareler tahmincisi $\hat{\beta}$,

$$\hat{\beta} = \frac{\sum_{i=1}^{n} (Y_i - \overline{Y}_n)(X_i - \overline{X}_n)}{\sum_{i=1}^{n} (X_i - \overline{X}_n)^2}$$

şeklinde bulunur. Burada; $\overline{Y}$ ve $\overline{X}$, X ve Y'nin ortalamalarıdır ve $\overline{Y}_n = (1/n)\sum_{i=1}^{n} Y_i$, $\overline{X}_n = (1/n)\sum_{i=1}^{n} X_i$ dir. Regresyon doğrusunun Y eksenini kestiği noktanın en küçük kareler tahmincisi $\hat{\alpha}$ ise

$$\hat{\alpha} = \frac{1}{n}\left(\sum Y_i - \beta \sum X_i\right) = \overline{Y} - \beta \overline{X}_n$$

şeklinde bulunur.

En küçük kareler tahminleri $\hat{\alpha}$ ve $\hat{\beta}$, α^* ve β^*' nın özel durumları gibi elde edilebilirler. Şöyle ki, $T_1(Y) = \sum_{i=1}^{n} Y_i$ ve $T_2(Y) = \sum_{i=1}^{n}(X_i - \overline{X}_n)(Y_i - \overline{Y}_n)$, $\overline{X}_n = n^{-1}\sum_{i=1}^{n} X_i$ alınsın. $\hat{\alpha}$ ve $\hat{\beta}$' nın tahmini,

Sup {b : T_2 (Y - a – bx) > 0, bütün a' lar için}

= inf {b : T_2 (Y - a – bx) < 0, bütün a' lar için}

$$= \left\{\sum_{i=1}^{n}(X_i - \overline{X}_n)(y_i - \overline{y}_n)\right\} / \sum_{i=1}^{n}(X_i - \overline{X}_n)^2 = \hat{\beta} \ .$$

ve

$$Sup\left\{a : T_1(y - a - \hat{\beta}x) > 0\right\}$$
$$= \inf\left\{a : T_1(y - a - \hat{\beta}x) < 0\right\}$$
$$= (\overline{y}_n - \hat{\beta}\overline{X}_n) = \hat{\alpha}$$

şeklinde hesaplanır (Adichie, 1967).

$\hat{\beta}$, $W_{ij} = (X_j - X_i)^2$ ağırlıklar ile b_{ij} değişkenlerinin bir ağırlıklı ortalamasıdır, halbu ki β* değişkenlerin aynı kümesinin medyanı olur (Sen, 1968). Medyan tartılı ortalamaya göre sapan değerler tarafından daha az etkilendiği için, β*, $\hat{\beta}$' dan daha sağlamdır (Hussain and Sprent, 1983).

2.1.2. Maksimum Likelihood Metodu İle $\hat{\alpha}$ ve $\hat{\beta}$ Parametrelerinin Tahmini

$(x_i - y_i)$ n tane nokta çiftinin regresyon doğrusu aşağıdaki gibi olsun;

$$Y_i = \alpha + \beta x_i + e_i, \qquad i = 1,2, \ldots, n$$

Hata takımlarının ihtimal fonksiyonunun fonksiyonel formu spesifik olduğu (kesinlikle belirtildiği) zaman α, β ve σ^2 parametrelerinin tahmincileri maksimum likelihood metodu ile elde edilebilir. Bu metod örnek gözlemlerinin ortak ihtimal dağılışını kullanır. Bu ortak ihtimal dağılışı parametrelerin bir fonksiyonu gibi göründüğü zaman, Likelihood Fonksiyon olarak adlandırılır. Verilen $Y_1,\ldots,Y_n$ gözlemlerin normal regresyon modeli için Likelihood fonksiyonu;

$$L(\alpha,\beta,\sigma^2) = \prod_{i=1}^{n} \frac{1}{(2\pi\sigma^2)^{1/2}} \exp\left[-\frac{1}{2\sigma^2}(Y_i - \alpha - \beta X_i)^2\right]$$

$$= \frac{1}{\left(2\pi\sigma^2\right)^{n/2}} \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(Y_i - \alpha - \beta X_i\right)^2\right]$$

olur. β,α ve σ^{2}' nin değerleri Maksimum Likelihood tahminciler olur. Bunlar;

Parametre	Maksimum Likelihood Tahmincisi
α	$\hat{\alpha} = \overline{Y} - \hat{\beta}\overline{X}$
β	$\hat{\beta} = \dfrac{\sum(X_i - \overline{X})(Y_i - \overline{Y})}{\sum(X_i - \overline{X})^2}$
σ^2	$\hat{\sigma}^2 = \dfrac{\sum(Y_i - \hat{Y}_i)^2}{n}$

ve hata kareler ortalaması

$$\text{HKO} = \frac{1}{n-2}\hat{\sigma}^2 = \frac{\sum\left(Y_i - \hat{Y}_i\right)^2}{n-2} = \frac{\sum e_i^2}{n-2}$$

şeklinde hesaplanırlar (Neter, Wasserman and Kutner, 1989).

2.2. MEDYANA GÖRE PARAMETRE TAHMİNİ

Bir regresyon doğrusu eğiminin (β) medyana göre çeşitli non-parametrik tahmin metodları aşağıda verilmiştir.

2.2.1. Theil Metodu

1950 yılında Theil tarafından ileri sürülen metod araştırmacıların en çok başvurdukları eğim bulma metodlarından birisidir. Bir doğrunun eğimi tahmininde kullanılan Theil (1950)' in metodu, (x_i , y_i), (x_j , y_j) gözlem çiftlerinden hesaplanan eğim tahminlerinin medyanı hesabına dayandırılmaktadır (Hussain and Sprent, 1983). Bunun için, sahip olduğumuz (x_1 , y_1), ... ,(x_n , y_n) n tane müşahede çiftinin;

$$y_i = \alpha + \beta x_i + e_i \ , \qquad i = 1, 2, ..., n$$

regresyon modeline uyduğunu farz edelim. Burada α (kesişim) ve β (eğim) bilinmeyen regresyon parametreleridir. X_i değerleri bilinen sabitler olup $X_1 < X_2 < .....< X_n$ şeklinde sıralanarak birbirinden farklıdırlar. Bu modelde e_i' ler varyansı σ_e^2 ve sıfır medyan ile simetrik bir sürekli dağılışa sahip olan bağımsız ve özdeş dağılmış şansa bağlı hatalardan oluşurlar (Rao and Gore, 1982).

β' nın tahmini $\hat{\beta}$, $i < j$ $(x_i \neq x_j)$ olmak üzere elde edilen $N = \binom{n}{2}$ tane

$$b_{ij} = (y_j - y_i) / (x_j - x_i)$$

eğimlerinin tümünün bir ağırlıklı medyanı olur (Sievers, 1978). Yani,

$$\hat{\beta} = \text{medyan } \{ b_{ij} \}$$

şeklindedir.

Theil (1950), α'nın tahmininde, tüm $a_i = y_i - \hat{\beta} x_i$ değerlerinin medyanı olduğunu ifade etmiştir (Hussain and Sprent, 1983).

2.2.1.1. Theil Tahmininin Önemlilik Testi

$(X_1, Y_1), (X_2, Y_2), \ldots\ldots, (X_n, Y_n)$ tane müşahede çiftlerinin;

$$Y_i = \alpha + \beta X_i + e_i \;, \quad (i = 1, 2, \ldots, n)$$

modeline uyduğunu farzedelim. Ve;

$H_o : \beta=\beta_0$ hipotezini

$H_1 : \beta\neq\beta_0$ alternatif hipotezine karşı test edelim (Hogg and Randles, 1975). Theil tarafından önerilen istatistik, esas itibariyle;

$$W' = \binom{n}{2}^{-1} \sum_{i<j} u\left(\frac{Y_j - Y_i}{X_j - X_i} - \beta_0 \right)$$

dır. Bu formül daha kısa olarak;

$$W' = \binom{n}{2}^{-1} \sum_{i<j} u(V_{ij})$$

şeklinde ifade edilebilir. Burada u (V) fonksiyonu;

$$u(V) = -1 \qquad V < 0 \text{ ise,}$$
$$u(V) = 0 \qquad V = 0 \text{ ise,}$$
$$u(V) = 1 \qquad V > 0 \text{ ise}$$

şeklindedir. Theil' in işaret ettiği gibi, (2W' – 1) Kendall' ın τ istatistiği ile yakından ilişkilidir. (2W' – 1) ve τ arasında sıfır olmayan durumlar için tam bir karşılıklı tekabül söz konusu olmasa bile, (2W' – 1)' nin sıfır dağılışı ile τ' nun sıfır dağılışı özdeştir (benzerdir). Bu yüzden, H_o doğru ise

$$E(W') = 1/2$$

ve

$$Var(W') = \frac{2n+5}{18n(n-1)}$$

olur ve W' asimptotik olarak H_o altında normaldir. Bundan dolayı

$$\left[\frac{2n+5}{18n(n-1)}\right]^{-\frac{1}{2}} \left| W' - \frac{1}{2} \right| > Z_{\left(\frac{1}{2}\right)\alpha}$$

kritik bölge ile test, önem seviyesi α'ya eşit olan H_o' ın nonparametrik testini tayin eder (Potthoff, 1974).

2.2.1.2. Theil Metodu İle β Katsayısının Güven Sınırları

$X_j - X_i \neq 0$ $(1 \leq i < j \leq n)$ farklılıklarının sayısı N ise, önerilen nokta tahmincisi $X_i \neq X_j$ için N tane b_{ij} eğiminin medyanı olur (Sen, 1968).

β için bir güven aralığı çift yönlü hipotez $H_0 : \beta=\beta_0$, $H_1 : \beta\neq\beta_0$ testiyle elde edilebilir. Meydana gelen güven aralığı i < j olmak üzere düzenlenmiş (muntazam) b_{ij} eğimler setinde uygun olarak seçilmiş sınırlı noktalara sahiptir (Sievers, 1978).

i < j olmak üzere toplam $N = \binom{n}{2}$ adet b_{ij} değeri küçükten büyüğe doğru sıralanır. β' nın güven aralığının alt sınırı $\hat{\beta}_a$, küçükten büyüğe doğru sıralanan k' ncı b_{ij} değeridir. β' nın güven aralığının üst sınırı $\hat{\beta}_ü$ ise büyükten küçüğe doğru sıralanan k' ncı b_{ij} değeridir (Daniel, 1990).

$$k = \frac{N - S_{\alpha/2} - 2}{2} \text{ dir.}$$

$S_{\alpha/2}$ değeri n ve α/2' ye göre Kendall' ın Tau Test İstatistik Değerleri Cetveline bakılır, bulunan değeri formülde yerine koyarak k değeri tespit edilir..

Eğer n değeri çok büyükse (n > 40) dağılış normal dağılışa yaklaşır ve dağılışın standart sapması;

$$\sigma = \sqrt{\frac{n(n-1)(2n+5)}{18}}$$

olur(Griffin,1962). O zaman standart normal dağılım tablosu kullanılarak k değeri;

$$k = \frac{N - Z_{\alpha/2}\sqrt{\frac{n(n-1)(2n+5)}{18}}}{2}$$

formülüne göre hesaplanır.

2.1.2. Adichie Metodu

Adichie'ye (1967,b) göre; $Y_1,..,Y_n$ (aşağıdaki dağılışlar ile) bağımsız tesadüfi değişkenler olsun,

$P_{\alpha\beta}(Y_j \leq y) = F(y-\alpha-\beta x_j)$

burada x_j bilinen regresyon sabitleridir ve tümü birbirinden farklıdır. $P_{\alpha\beta}$, α ve β parametre değerleri için tahmin edilmiş ihtimali ifade eder. α ve β' nın tahminleri α^* ve β^* aşağıdaki şekilde tarif edilir.

T_1 $(Y_1 \ldots. Y_n)$ ve T_2 $(Y_1 \ldots. Y_n)$ α ve β hakkında hipotez testleri için iki istatistik olsun. T_1 ve T_2'nin aşağıdaki iki durumu yerine getirdiği kabul edilsin.

A) Her bir y ve x için, sabit b için, T_1 (y + a + bx) a' da azalmayan ve her a için T_2(y + a + bx) b' de azalmayan olur. Burada y + a + bx kısaltması $(y_1 + a + bx_1, \ldots, y_n + a + bx_n)$'i gösterir.

B) $\alpha=\beta=0$ olduğu zaman, T_2 (Y) ve T_1 (Y)'nin dağılışları μ ve ν sabit noktaları civarında simetrik olurlar. T_1 ve T_2 uygun fonksiyonları için, α ve β' nın tahminleri için α^* ve β^* aşağıdaki şekilde bulunur.

$$\beta_1^* = \sup [\, b : T_2\,(y - a - bx) > \nu, \text{ tüm a' lar için}],$$

$$\beta_2^* = \inf [\, b : T_2\,(y - a - bx) < \nu\,, \text{ tüm a' lar için}],$$

$$\beta^* = \frac{1}{2}\left(\beta_1^* + \beta_2^*\right);$$

$$\alpha_1^* = \sup [\, a : T_1\,(y - a - \beta^* x\,) > \mu\,],$$

$$\alpha_2^* = \inf [\, a : T1\,(y - a - \beta^* x\,) < \mu\,];$$

$$\alpha^* = \frac{1}{2}\left(\alpha_1^* + \alpha_2^*\right).$$

Ayrıca Adichie'de (1967,a), R_j n gözlemin $|Y_1|, \ldots, |Y_n|$ mutlak değerleri dizisinde $|Y_j|$ 'nin rankı olsun. T_1 ve T_2 istatistik çifti aşağıdaki şekilde tarif edilir;

$T_1(Y) = n^{-1}\sum_j x_j \; \Psi_n\left(R_j / n+1\right) Sign Y_j$,

$T_2(Y)=n^{-1}\sum_j x_j \Psi_n\left(R_j / n+1\right) Sign Y_j$,

burada

$$\Psi_n(u)=\Psi(j / n+1) , \qquad \frac{j-1}{n}<u \leq \frac{j}{n} ,$$

ve

$$\lim \int_0^1\left[\Psi_n(u)-\Psi(u)\right]^2 du=0 .$$

Bir simetrik 2×2 matriks $\left\|\gamma_{kl}\right\|_n$ aşağıdaki gibi olsun

$$\gamma_{11}=\int_0^1 \Psi^2(u) du ;$$

$$\gamma_{22,n}=n^{-1} \sum_j x_j^2 \int_0^1 \Psi^2(u) du ;$$

$$\gamma_{21,n}=n^{-1} \sum_j x_j \int_0^1 \Psi^2(u) du .$$

Tarif:

$$M(\Psi)=n(T_1, T_2)\left\|\gamma_{kl}\right\|_n^{-1}(T_1, T_2)'$$

burada V'nin transpozu V' ve $\left\|\gamma_{kl}\right\|$'nin inversi $\left\|\gamma_{kl}\right\|_n^{-1}$ olur ve Adichie (1967,a) tarafından M(Ψ), H:α=β=0 hipotezleri için test istatistiklerinin bir sınıfı gibi ileri sürülmüştür.

2.2.3. Sen Metodu

Sen (1968)' aşağıdaki metodu geliştirmiştir. Genel olarak $x_1 \leq x_2 \leq \ldots . \leq x_n$ olduğu farzedilir. $(x_j - x_i)$ pozitif farklarının sayısı N aşağıdaki şekilde bulunur.

$$N=\sum_{1 \leq i \leq j \leq n} c\left(x_j-x_i\right)$$

Burada c (u)' yu;

$c(u) = 1 \qquad u > 0$ ise,

$c(u) = 0 \qquad u = 0$ ise,

$c(u) = -1 \qquad u < 0$ ise

şeklinde tarif ederiz.

Herhangi bir reel b sayısı için, $Z_{i(b)} = Y_i - bx_i$, i = 1,2,3, ……. n tarif edilir. O zaman temelde x_i ve $Z_{i(b)}$, i = 1,2,3,…,n arasında Kendall' ın Tau' su ile alakalı aşağıdaki istatistik ortaya çıkar.

$$U_n(b) = \left\{N\binom{n}{2}\right\}^{-\frac{1}{2}} \sum_{1\le i<j\le n} c(X_j - X_i)c(Z_j(b) - Z_i(b)).$$

Böylece, $\left\{N\binom{n}{2}\right\}^{\frac{1}{2}} U_n(b)$ farklı işaret sayısı olur ve bir sabit b için, $(Y_i - bx_i)$ ve x_i arasındaki korelasyonun tau katsayısının payında görülür. Tüm $i < j$ için $x_j \geq x_i$' den dolayı, $Z_j(b)-Z_i(b)$ tüm $1 \leq i < j \leq n$ için b' de artmayandır. Bundan dolayı $U_n(b)$ de b' de artmayandır. Şimdi, tanıma göre, $Z_1(\beta)$, …… ,$Z_n(\beta)$ n tane bağımsız ve özdeş dağılmış tesadüfi değişkendir ve $x_n = (x_1, x_2, \ldots\ldots x_n)$' den bağımsız olan F(x-α) dağılış fonksiyonuna sahiptirler. Sonuç olarak, $U_n(\beta)$ istatistik tahmincisinin sonucu 0' a yakın olacaktır. Aslında, $U_n(\beta)$ sıfır etrafında simetrik bir dağılışa sahip, dağılıştan bağımsız bir istatistiktir. Böylece β' yı tahmin etmenin bir yolu $U_n(\beta)$' yi mümkün olduğunca sıfıra yakın yapmaktır. Çünkü, $U_n(\beta)$ b' de artmayan olduğundan $U_n(\beta)$' nin sıfıra eşit olacağı bir yarı açık aralık bulunacaktır. Bu aralığın orta noktası doğal olarak β'nın bir tahmini olacaktır. Matematiksel olarak tahminci aşağıdaki şekilde tanımlanır.

$\beta_1^* = \sup\ [\, b : U_n(b) > 0],$

$\beta_2^* = \inf\ [\, b : U_n(b) < 0].$

Buradan,

$\beta^* = \frac{1}{2}\left(\beta_1^* + \beta_2^*\right)$ elde edilir (Hogg and Randles, 1975).

Ayrıca N' in tek (2K + 1) veya çift (2K) olması durumuna göre β*,

$$\beta^* = \begin{cases} X_{(K+1)} & , \quad N = 2K+1 \\ \frac{1}{2}\left[X_{(K)} + X_{(K+1)}\right], & \quad N = 2K \end{cases}$$

şeklinde yazılabilir.

2.2.4. Maritz Metodu

Maritz (1979), Theil (1950)' in istatistiğini kullanarak, F (y_i –α-βx_i) dağılış fonksiyonuna sahip olan bağımsız sürekli tesadüfi değişkenler olan Y_i' lerdeki linear regresyon durumunu mütalaa etmiştir. $e_i = Y_i - \alpha - \beta x_i$ bağımsız ve özdeş dağılmış olsunlar, i = 1,2.....n ve $x_1 < x_2 < \ldots\ldots < x_n$ bilinir ve sabittirler. β'nın tahmini için Theil(1950) metodu her j > i için,

$$e_j - e_i = (Y_j - \alpha - \beta x_j) - (Y_i - \alpha - \beta x_I) = (Y_j - Y_i) - \beta\,(X_j - X_i),$$

ifadesinin matematiksel olarak α dan bağımsızlığına ve sıfır etrafında simetrik olarak dağılışına dayanır. Theil istatistiği;

$$B\,(Y, \beta) = \sum\nolimits_{i<j} U_{ij}$$

şansa bağlı değerinden çıkarılır. Burada,

$$U_{ij} = \text{sgn}\,[(Y_j - Y_i) - \beta\,(X_j - X_i)]$$

$$E(U_{ij})=0$$

$$Var(U_{ij})=1$$

dir. İstatistik β, α hariç tutularak (elimine edilerek) e_i, e_j çiftlerinin kullanılmasıyla elde edilebilmiştir. Benzer bir usul de β' yı da hariç tutulabilir (elimine edilebilir).

$$V_{ij} = \text{sgn}\,[x_i e_j - x_j e_i] = \text{sgn}\,[x_i y_j - x_j y_i - \alpha(x_i - x_j)]$$

olsun

$$A\,(Y, \alpha) = \sum\nolimits_{i<j} V_{ij}$$

tahmin edilir. A (Y, α) istatistiği α üzerinde netice çıkarmak için uygun olabilir.

A (Y, α) = 0 tahmin denkleminin çözümüyle α' nın elde edilen hata tahmini

$(x_i y_j - x_j y_i)/(x_i - x_j)$, i < j, i, j = 1,2, ..., n

çiftlerinin, çifter çifter kesişim tahminlerinin medyanı olur; bir tahmin aynı zamanda indislerin uygun bir şekilde seçilen alt kümeleri üzerinde bu gibi tahminlerin medyanı alınarak elde edilebilir. α' nın nokta tahmini Theil istatistiği esası üzerine dayandırılan β' nın tahmini ile direkt benzer olur, yani tüm eşli eğim tahminlerinin medyanıdır.

2.2.5. Kildea Metodu

Kildea'de (1981), en küçük kareler tahmincisinin bir doğal medyan benzeri elde edilmesi ile Brown-Mood tahmincisinin nasıl değişik olabileceğini incelemiştir.

(Y_i, X_i) $i = 1, 2, \ldots, n$ n tane nokta çiftinin her birinin $y_i = \alpha+\beta x_i + e_i$ modeline uyduğu farzedilsin ve modeldeki x_i' ler bilinen sabitler, α ve β bilinmeyen parametreler ve e_i' ler ise hata payları olur.

$\{x_i; i = 1,2, \ldots, n\}$ sayılarının medyanı x^* olsun ve $\bar{x} = \dfrac{\sum x_i}{n}$, o zaman (α , β)' nın Brown-Mood tahmincisi (α', β') aşağıdaki şekilde tarif edilir,

$$\text{med.}\left(y_i - \alpha' - \beta' x_i, 1, x_i \le x^*\right) = \text{med.}\left(y_i - \alpha' - \beta' x_i, 1, x_i > x^*\right) = 0 . \quad (5.1)$$

Bu aşağıdaki şekilde tekrar yazılabilir;

$$med.\left(y_i - \alpha' - \beta' x_i, 1, i \le n\right) = 0, \quad (5.2)\,(A)$$

$$med.\left(y_i - \beta' x_i, 1, x_i \le x*\right) = med.\left(y_i - \beta' x_i, 1, x_i > x*\right) \;. \quad (5.2)\,(B)$$

En küçük kareler tahmincisi $\left(\hat{\alpha}, \hat{\beta}\right)$ yı tanımlayan normal denklemler aşağıdaki gibi yazılabilir,

$$mean\left(y_i - \hat{\alpha} - \hat{\beta} x_i, 1, i \le n\right) = 0, \quad (5.3)\,(A)$$

$$mean\left(y_i - \hat{\beta} x_i, |x_i - \bar{x}|, x_i \le \bar{x}\right) = mean\left(y_i - \hat{\beta} x_i, x_i - \bar{x}, x_i > \bar{x}\right). \quad (5.3)\,(B)$$

(5.2) ve (5.3)'ün bir mukayesesi yapılırsa, (5.2)(A), (5.3)(A)' nın bir doğal medyan benzeri (eşi) olur, fakat (5.2)(B), (5.3)(B)' den sadece ortalama ve medyan işlemleri bakımından değil aynı zamanda residualların ağırlıklandırılması ve gruplandırılması bakımından da farklıdır. Bu sebepten Kildea (5.2)(B) eşitliğinin aşağıdaki gibi değişik olduğunu kabul etmiştir.

{1,2, .., n}' in ikiye ayrılmış boş olmayan alt kümeleri S ve U olsun. Şöyle ki;

$(i \in S \text{ ve } j \in U) \Rightarrow x_i < x_j$.

$\{S_i : i \in S\}$ ve $\{u_i : i \in U\}$ pozitif sayəların (ağırlıkların) iki kümesi olsun, $\sum s_i > 0$ ve $\sum u_i > 0$. O zaman Kildea'nın değişik Brown-Mood tahminciler sınıfları;

$$med.\left(y_i - \tilde{\alpha} - \tilde{\beta}x_i, 1, i \le n\right) = 0, \qquad (5.4)\,(A)$$

$$med.\left(y_i - \tilde{\beta}x_i, s_i, S\right) = med.\left(y_i - \tilde{\beta}x_i, u_i, U\right) \qquad (5.4)\,(B)$$

şeklinde tanımlanır. Kildea (1981) de

$g(\beta) = med(y_i - \beta x_i, s_i, S) - med(y_i - \beta x_i, u_i, U)$

şeklinde bir fonksiyon türetmiş. Bu fonksiyonda β yerine $\beta = \dfrac{y_j - y_i}{x_j - x_i}$ $(i,j \in U)$ veya $(i,j \in S)$ ye göre elde ettiği değerleri yerine koyarak S ve U kümesine göre elde ettiği medyan değerlerinim farkını alıp β değerini tahmin etmiştir.

2.2.6. Brown-Mood Metodu

$x_1 < x_2 < \ldots\ldots < x_n$ n sayının medyanı aşağıdaki şekilde tarif edilir.

$$M_n = \begin{cases} X_{(k+1)} & n = 2k+1 \quad \text{ise} \\ \dfrac{1}{2}\left(X_k + X_{(k+1)}\right) & n = 2k \quad \text{ise} \end{cases}$$

Brown ve Mood α ve β 'nın birlikte tahminini aşağıdaki iki eşitlikten önermişlerdir.

$$Medyan_{\{i: x_i \le M_n\}}\left[Y_i - \hat{\alpha} - \hat{\beta}x_i\right] = 0,$$

$$Medyan_{\{i: x_i > M_n\}}\left[Y_i - \hat{\alpha} - \hat{\beta}x_i\right] = 0$$

burada M_n $(x_1, x_2, \ldots\ldots, x_n)$ noktalarının medyanıdır (Hill, 1962).

Y - βx_i değerlerinin n / 2' sinin bir yarısı medyan (x_i)' in sağında ve diğer yarısı medyan (x_i)' nin solunda oluşturulması amacıyla β' nın seçilme zorunluluğu vardır. Bu ifade β'nın Brown-Mood tahminini sağlar (Hogg and Randles, 1975).

X_i değerlerinin genel medyanı medyan (X_i) bulunur. Bu genel medyan değeri (X_i, Y_i) müşahedelerini ikiye böler. Genel medyan değerinin altındaki ve üstündeki X_i ve Y_i değerlerinin medyanı bulunur yani iki tane X_i değerinin ve iki tanede Y_i değerlerinin olmak üzere toplam dört adet medyan değeri hesaplanır. Bu medyan değerleri;

$$\beta = \frac{Y_{1(Medyan)} - Y_{2(Medyan)}}{X_{1(Medyan)} - X_{2(Medyan)}}$$

formülünde yerine konarak nihai tahminden önceki eğim bulunmuş olur. Elde edilen doğru etrafındaki müşahedelerin medyanları her iki grupta da sıfırdan farklı olduğu zaman elde edilen doğruyu düzeltmek gerekir (Daniel, 1990). Düzeltme yapılırken X medyan değerinin sağındaki ve solundaki noktaların regresyon doğrusundan dikey sapmalarının medyanı sıfır olana kadar regresyon doğrusunun pozisyonu değiştirilir ve sonuçta nihai regresyon doğrusu bulunmuş olur ve nihai doğru üzerinde işaretlenen iki nokta değerlerini

$$\beta' = \frac{Y_1 - Y_2}{X_1 - X_2}$$

formülünde yerine koyarak nihai $\hat{\alpha}$ ve $\hat{\beta}$ değerleri bulunmuş olur.

α ve β değerlerinin önemlilik testini yaptığımızda hipotezler

$H_0: \alpha = \alpha_0$, $\beta = \beta_0$

$H_1: \alpha \neq \alpha_0$, $\beta \neq \beta_0$

şeklinde kurulur. Test istatistiğini bulmak için ilk önce n adet (X_i,Y_i) değerleri dağılma diyagramında gösterilir. Diyagramda $Y=\alpha_0+\beta_0 X$ doğrusu çizilir ve X değerleri medyanından Y eksenine paralel bir çizgi çizilir. Farazi regresyon doğrusunun üzerindeki ve X değerleri medyanından çizilen dikey çizginin sol tarafındaki noktaların sayısı n_1, farazi regresyon doğrusunun üzerindeki ve X değerleri medyanından çizilen dikey çizginin sağ tarafındaki noktaların sayısı n_2 ise, test istatistiği,

$$X^2 = \frac{8}{n}\left[\left(n_1 - \frac{n}{4}\right)^2 + \left(n_2 - \frac{n}{4}\right)^2\right]$$

şeklinde tarif edilir. Bu istatistik H_0 hipotezi doğru iken n çok küçük değilse, 2 serbestlik dereceli X^2 dağılımına yaklaşır (Daniel, 1990)

2.3-SMOOTHING TEKNİKLERİ

Smoothing metodları ve teorisi 1980 yılından sonra gelişmeye başlamıştır. Smoothing teknikler uzun bir geçmişe sahiptir. 19.yy da parametrik olmayan uygulama, analizler için temel araç olarak kullanıldı. 1857'de Saksonya'lı ekonomist Engel genelde regressogram diye bilinen Engelsches Gezets'i bulmuştur. Bu tarihten itibaren, parametrik olmayan smoothing uygulamalarıyla uzun zaman ilgilenilmemiştir. İstatistikçiler 1950 yılından sonra istatistik teorinin matematik bulguları hesaplamalardaki basitliği, model faraziyeleri ile uygunluğu ve dahası matematiksel kolaylığı için sade bir parametrik uygulama önermişlerdir (Hardle,1997).

Bir regresyon eğrisi açıklayıcı değişken X ile cevap değişkeni Y arasındaki genel ilişkiyi tarif eder. X gözlendiğinde, Y'nin ortalama değeri regresyon fonksiyonu tarafından verilir. Regresyon fonksiyonunun şekli bize Y gözlemlerinin X'in kesin değerleri için tahmin edilenler olduğunu veya iki değişken arasındaki bağımlılığın özel bir çeşidinin olup olmadığını gösterir. Şayet n tane veri noktası $\{(X_i,Y_i)\}_{i=1}^{n}$ varsa, regresyon ilişkisi

$Y_i=m(x_i)+\epsilon_i$, $i=1,2,\ldots,n$,

modeliyle ifade edilir, burada m bilinmeyen regresyon fonksiyonu ve ϵ_i'ler gözlenen hatalardır. Y_i ye karşı X_i'nin dağılma diyagramına bakıldığında açıklanması mümkün bir regresyon ilişkisini kurmak daima yeterli olmaz. Çünkü dağılma diyagramındaki merkezileşmiş noktalar dışındaki noktalar bizim görüşümüzü şaşırtabilir. Regresyon analizinin asıl amacı, bilinmeyen cevap fonksiyon m ile uygun tahmini yaklaşımın sağlanılmasıdır. Gözlenen hataların azaltılması ile eğri x üzerinde Y'nin ortalama bağlılığının önemli detaylarının merkezileştirilmesi ile uygulamaya yardımcı olur. Bu genel ortalama usulü ile yaklaşık eğri bulma durumu «Smoothing» olarak adlandırılır. Bu genel olarak aşağıdaki gibi ifade edilebilir

$$\hat{m}(x)=n^{-1}\sum_{i=1}^{n}W_{ni}(X)Y_i$$

burada $\{W_{ni}(X)\}_{i=1}^{n}$, $\{X_i\}_{i=1}^{n}$ vektörüne bağımlı olabilen ağırlık dizisidir (Hardle,1997).

Nonparametrik regresyon metodları, regresyon fonksiyonlarının tayin edilmiş bir parametrik sınıfına sahip olmaksızın bir regresyon eğrisinin bir smooth uygunluğunu elde etmek için yaygın olarak uygulanırlar (Müller, 1992).

Sağlam sonuç çıkarmak için diğer ihtimaller gözönünde tutulmasına rağmen çok basit ve çok yaygın olarak kullanılan smootherler Kernel Metodlarına dayandırılmıştır (Wu and Chu, 1993). Kernel ismi, cevap değişkenlerinin bir kernel fonksiyon esasına dayanmış lokal ağırlıklı ortalamaları olan smootherlerden gelir (Chu and Marron, 1991).

Smootherler için nonparametrik regresyonun yaygın olarak mütalaa edilen iki tip modeli vardır. Bunlardan ilki sabit model (Fixed design) olup, bu modelde dağılma diyagramındaki veri noktalarının ordinatları belirleyici değer olarak düşünülür. Bu ordinat değerler, planlanmış bir denemede olduğu gibi genellikle araştırıcı tarafından seçilirler. Uygun bir şekilde alınan örnek noktalar eşit uzaklıkta olur.

Sabit düzenlenmiş nonparametrik regresyon modeli;

$$Y_i = m(x_i) + \in_i \qquad i = 1,2, ..., n$$

şeklinde verilir. Burada m(x) genellikle [0,1] aralığında tanımlanmış regresyon fonksiyonudur, x_i' ler eşit aralıkta sabit düzenlenen noktalardır, yani $x_i = i / n$ dir. $\in_i$' ler sıfır ortalama ve σ^2 $(0 < \sigma^2 < \infty)$ varyans ile bağımsız ve özdeş dağılmış regresyon hatalarıdır, Y_i' ler x_i' de m(x)' in tahmini gözlemleridir (Wu and Chu, 1993).

İkincisi, şansa bağlı (Random design) nonparametrik regresyon modelidir. Bu modelde, dağılma diyagramındaki veri noktaları iki değişkenli ihtimal dağılışından alınmış değerler gibi düşünülür. Bu ordinatlar genelde deneyci tarafından tesadüfi olarak seçilirler. Şansa bağlı nonparametrik regresyon modeli;

$$Y_i = m\,(x_i) + \in_i, \qquad i = 1,2, ..., n$$

şeklinde verilir. Burada (x_i, y_i)' ler bağımsız ve özdeş dağılmış tesadüfi değişkenlerdir, $\in_i = Y_i - m(x_i)$ şeklinde tanımlanan $\in_i$ lerin ortalaması sıfır ve varyansı σ^2 olduğu varsayılır. Herbir durumdaki amaç, eğri m(x)'i tahmin etmek için $Y_1 .... Y_n$ 'i kullanmaktır. İlk bakışta uygulamada bu iki model arasında çok az bir fark olduğu düşünülür, çünkü regresyon fonksiyonu (yani şartlı beklenen değer) sadece şartlı dağılış üzerine bağlıdır, burada ordinat değerler verilir (Chu and Marron, 1991).

2.3.1-Kernel Smoothing

$\{W_{ni}(x)\}_{i=1}^{n}$ ağırlıklı dizisi, x civarındaki ağırlıkların şekli ve büyüklüğünü ayarlayan bir skala parametresine sahip bir yoğunluk fonksiyonu vasıtasıyla $W_{ni}(x)$ ağırlık fonksiyonunca ifade edilir. Genelde bu fonksiyon şekli kernel K olarak adlandırılır (Hardle,1997). Kernel fonksiyonu süreklidir ve sınırlıdır. Simetrik reel fonksiyon K'nın integrali bire eşit olur (Rosenblatt, 1971). Yani;

$$\int K(u)du = 1.$$

Kernel smootherler için ağırlık dizisi

$$W_{ni}(x) = K_{h_n}\left(x - X_i\right) / \hat{f}_{h_n}(x)$$

şeklinde tanımlanır.Burada kernel yoğunluk tahmincisi

$$\hat{f}_{h_n}(x) = n^{-1}\sum_{i=1}^{n} K_{h_n}(x - X_i)$$

şeklindedir (Hardle ve Mammen, 1993). Her h_n için

$$K_{h_n}(u) = h_n^{-1}K(u / h_n)$$

olur ve K(u), $-\infty<x<\infty$ için ağırlık fonksiyonunun genel bir tipi olarak tarif edilmiştir. $\{h_n\}$, $n\to\infty$ iken $nh_n\to\infty$ yaklaşmasına benzer şekilde sıfıra giden pozitif reel sayıların bir dizisidir, yani h_n'nin değişmesiyle fonksiyon $K_{h_n}(x)$'in genişliği değişebilir (Priestley ve Chao, 1972).

Bir ağırlık fonksiyonu olan kernel K;

a) $K(u)\geq 0$, tüm u'lar için

b) $\int_{-\infty}^{\infty} K(u)du = 1$

c) $\int_{-\infty}^{\infty} K^2(u)du < \infty$

d) $K(u)=K(-u)$

şartlarını sağlar (Rosenblatt, 1971; Priestley ve Chao,1972; Benedetti, 1977).

2.3.1.1-Priestley ve Chao Tahmin Metodu

Priestley ve Chao (1972), regresyon fonksiyonu m(x) için aşağıdaki tahminciyi önermişlerdir;

$Y_i=m(x_i)+\epsilon_i$, $i=1,2,....,n$, $0\leq x_1\leq x_2\leq ......\leq x_n\leq 1$,

modeline göre tanımlanmış $Y_1,Y_2,.....,Y_n$ n tane gözlem olsun. Burada m(x), [0,1] aralığında sınırlanan x üzerinde bilinmeyen bir fonksiyon olarak tarif edilmiştir (m(x) $0\leq x\leq 1$ için tanımlanmış bir bilinmeyen fonksiyon). $\{x_i\}$ bilinen değer ve ϵ_i'ler(hatalar) sıfır ortalama ve sonlu varyans σ^2 ile bağımsız ve özdeş dağılmış tesadüfi değişkenlerdir. Yani;

$E[\epsilon_i]=0$, $E\left[\epsilon_i^2\right]=\sigma_e^2$, $i=1,2,...,n.$

m(x)'in Priestley ve Chao tahmincisi

$$\hat{m}_n(x)=\sum_{i=1}^{n} Y_i \frac{x_i-x_{i-1}}{h_n} K\left(\frac{x-x_i}{h_n}\right)$$

şeklindedir. Burada K(.) bir kernel fonksiyon ve $\{h_n\}$, $n\to\infty$ iken $nh_n\to\infty$ yakınlaşmasına benzer şekilde sıfıra giden pozitif reel sayıların bir dizisidir. Ayrıca bu tahminci Benedetti(1977); Hardle ve Gasser(1984); Hardle(1997) tarafından kullanılmıştır.

2.3.1.2-Nadaraya ve Watson Tahmincisi

Regresyon modeli,

$Y_i=m(x_i)+e_i$, $i\leq 1\leq n$,

şeklinde olsun. Burada m bir smooth regresyon fonksiyonu, $x_1,...,x_n$ verilen bir aralıkta sınırlanmış bilinen sabitler ve e_i'ler ortalaması sıfır varyansı σ^2 olan bağımsız ve özdeş dağılmış hatalardır. Bandwith h ve kernel K kullanılarak m(x)'in tahmincisi $\hat{m}(x)$, Nadaraya (1964) ve Watson (1964) tarafından aşağıdaki şekilde önerilmiştir. (Stute,1984; Hall ve Wehrly,1991; Hardle,1997).

$$\hat{m}_h(x) = \frac{\sum_{i=1}^{n} Y_i K((x - X_i)/h_n)}{\sum_{i=1}^{n} K((x - X_i)/h_n)} .$$

2.3.1.3-Gasser ve Müller Tahmincisi

Nonparametrik regresyon modeli,

$Y_i = m(x_i) + e_i$, $0 \le x_1 \le .... \le x_n = 1$

şeklinde olsun.

Hardle ve Gasser (1984) tarafından, Gasser ve Müller'in (1979)

$$W_{ni}(x) = h_n^{-1} \int_{S_{i-1}}^{S_i} K((x-u)/h_n)du$$

ağırlıkları kullanılarak, m(x)'in Gasser ve Müller (1979) tahmincisi $\hat{m}_h(x)$;

$$\hat{m}_h(x) = \frac{1}{2} \sum_{i=1}^{n} Y_i \int_{S_{i-1}}^{S_i} K((x-u)/h_n)du$$

şeklinde ifade edilmiştir (Hall ve Wehrly,1991). Burada $S_i=(x_i+x_{i+1})/2$ $(i=1,..,n-1)$, $S_0=0$, $S_n=1$, h_n pozitif bandwithların dizisi ve K bir kernel fonksiyondur. (Müller,1992; Chambers, Dorfman veWehrly,1993) bu ağırlıkları kullanmışlardır.

2.3.1.4-Hall ve Wehrly Tahmincisi

Hall ve Wehrly (1991)'nin, geliştirdikleri tahmin edici Nadaraya-Watson tahmin edicisinden esinlenerek geliştirilmiştir. Alınan regresyon modeli

$Y_i = f(x_i) + e_i$, $1 \le i \le n$,

olsun. Burada x_i'ler düzenlenmiş noktalar, f bir smooth regresyon ortalama fonksiyonu, e_i'ler sıfır ortalama ile bağımsız ve özdeş dağılmış hatalardır. Gözlenmiş veri çiftleri $X = \{(x_i, Y_i),\ 1 \le i \le n\}\ [a,b](a \le x_1 \le .... \le b)$ aralığında olurlar.

Hall ve Wehrly (1991), Kernel K ve bandwidth h'ı kullanarak f'in tahmincisi $\hat{f}$'yü

$$\hat{f}_{(x)} = \left[\sum_i K\{(x - x_i)/h\}Y_i\right] / \left[\sum_i K\{(x - x_i)/h\}\right],\ a \le x \le b$$

şeklinde ifade etmişlerdir. Bu formül 0<h<b-a ve tüm $x \in [a,b]$ için iyi tarif edilmiş bir tahminciyi meydana getirir.

Hall ve Wehrly (1991), bir simulasyon çalışması yapmışlardır. Yaptıkları çalışmada kendi geliştirdikleri tahmin ediciyi, Nadaraya-Watson (NW) ve Priestley-Chao (PC) tahminedicileri ile mukayeseli olarak incelemişlerdir.

Hall ve Wehrly'de (1991), aşağıdaki modeli kullanmışlardır.

$$Y_i = \left(x_i - \frac{1}{2}\right)^2 + \epsilon_i\ , \qquad i=1,2,...,n.$$

Burada $x_1,...,x_n$ bağımsız üniform (0,1) şans değişkenleri ile ilgili gözlenen değerlerdir. $\epsilon_1,....,\epsilon_n$ ise ortalaması sıfır ve varyansı σ^2 olan bağımsız normal dağılmış şans değişkenleridir.

$x_1,...,x_n$ ve $\epsilon_1,...,\epsilon_n$ ile ilgili şansa bağlı 400 örnekten σ ve n değerleri elde edilmiştir. Herbir tahminedici için band genişliğini seçmek için bir eksiği alınmış çapraz değerlendirme kullanılmıştır. Herbir şansa bağlı örnek için t_i=i/100(i=0,1,...100) de $\hat{f}_{\hat{h}}(t_i)$ tahmini hesaplanmıştır. Burada $\hat{f}_{\hat{h}}$ parametrik olmayan regresyon tahminlerinin herhangi biri anlamına gelir. $\hat{h}$, bir eksiği alınmış çapraz değerlendirmesi kullanılarak elde edilen datalardan seçilmiş band genişliğini gösterir. 0<u<1 için L(u)=16(8-15u)K(u)/9 ve $|u| \le 1$ için ise K(u)=3(1-u^2)/4 kullanılan kernel fonksiyondur.

Doğrudan seçilen band genişliğini kullanan tahminedicilerin performansını değerlendirmek için aşağıdaki formülü kullanarak integrallenmiş hata kareler ortalaması (MISE) tahmin edilmiştir

$$MISE(\hat{f}_h) = \frac{1}{2}\sum_{i=0}^{100}\left\{\hat{f}_h(t_i) - f(t_i)\right\}^2 .$$

Bu 400 simulasyonun her biri ile ilgili tahmin edicilerin herbiri için hesaplanmıştır. Tahminedicilerin herbiriyle ilgili olarak tahmin edilen MISE'nin ortalama, medyan, standart hata ve çeyrek ayrılış oranı tablo 4.2.1.1' de verilmiştir. Önerilen tahmin edici, yaklaşık %70 genel Nadaraya-Watson tahmin edicisinden daha iyi tahmin edilmiş MISE'e sahiptir. Bu durum Gasser-Müller tahmin edicisi için de geçerlidir. Ortalama veya medyanı tahmin edilmiş MISE, önerilen tahmin edici ve Gasser-Müller tahminedici için yaklaşık olarak aynıdır, fakat Nadaraya-Watson tahminedici için biraz daha büyüktür. Şansa bağlı plan ve bu özel simulasyon bakımından önerilen tahminedici bir dereceye kadar Gasser-Müller tahminedicisine tavsiye edilebilir. Şans planını kullanan bu simulasyon bakımından önerilen tahminedici diğer tahminedicilere üstünlük sağlamıştır. En kötü olanı da Nadaraya-Watson tahminedicisidir.

Tablo 4.2.1.1. Hall ve Wehrly, Nadaraya-Watson ve Gasser-Müller tahminedicilerin her biriyle ilgili olarak tahmin edilen MISE tablosu.

	Ortalama	Medyan	Standart Sapma	Çeyrek Ayrılış
		Şansa Bağlı Model		
H-W	0.087	0.064	0.118	0.062
N-W	0.097	0.085	0.063	0.059
G-M	0.087	0.067	0.072	0.068
		Sabit Model		
H-W	0.062	0.053	0.046	0.043
N-W	0.079	0.070	0.046	0.051
G-M	0.059	0.052	0.039	0.042

H-W: Hall-Wehrly tahmincici
N-W: Nadaraya-Watson tahmincisi
G-M: Gasser-Müller tahmincisi

Bir benzer simulasyon sabit plan (fixed design) için yürütülmüştür. $Y_i = \left(x_i - \frac{1}{2}\right)^2 + \epsilon_i$, (i=1,2,........n) modeli kullanılmıştır. Burada x_i=(i-0.5)/100, i=1,........100 ve $\epsilon_1, \ldots, \epsilon_{100}$ ortalaması sıfır standart sapması 0.1 (σ=0.1) olan bağımsız normal şansa bağlı değişkenlerdir. Şansa bağlı 400 örneğin herbiri için herbir X_i'de $\hat{f}_{\hat{h}}(x_i)$ hesaplanmıştır. Burada, $\hat{f}_{\hat{h}}$ parametrik olmayan regresyon tahminlerinin herhangi biri anlamına gelmektedir ve $\hat{h}$ ise birisi hariç bırakılarak meydana getirilen çapraz değerlendirme ile elde edilen verilerden seçilen band genişliğini göstermektedir. Herbir tahmin edici için aşağıdaki formül kullanılarak tahmin edilmiş MISE hesaplanmıştır.

$$MISE(\hat{f}_{\hat{h}}) = \frac{1}{100}\sum_{i=1}^{100}\left\{\hat{f}_{\hat{h}}(x_i) - f(x_i)\right\}^2 .$$

Sonuçlar Tablo 4.2.1.1 de sabit plan kısmında verilmiştir. Sabit plan (fixed-design) için elde edilen sonuçlar, şansa bağlı plan için elde edilenlerden farklıdır. Hall ve Wehrly tarafından önerilen tahminedici Gasser-Müller tahmin edicisi ile yaklaşık olarak aynıdır. En kötü performansı ise Nadaraya-Watson tahmin edicisi göstermiştir. Hall ve Wehrly'nin sabit planı kabul eden bu simulasyon çalışmasında en iyi performansı Gasser-Müller tahminedicisi göstermiştir.

Simulasyon sonucunda Hall ve Wehrly tarafından geliştirilen tekniğin genel Gasser-Müller metodu üzerine olan ilave avantajları, bu tekniğin matematik olarak basit oluşu, formülünün integral ifadeler ihtiva etmemesi; sık sık kullanılan ve çok iyi bilinen Nadaraya-Watson tahmin edicisine eşit olması ve hem şansa bağlı hem de sabit planlar için iyi anlaşılmasından kaynaklanmaktadır.

2.3.1.5-Watson ve Leadbetter Kernel Yoğunluk Tahmincisi

Watson ve Leadbetter (1962) tarafından, bir ihtimal yoğunluğu $f(x)$ 'in $\hat{f}_n(x) = \frac{1}{n}\sum_{i=1}^{n}\delta_n(x - x_i)$ formunun tahmincileri üzerinde çalışılmıştır. Burada

Bilinmeyen ihtimal fonksiyonu $f(x)$, kesin olarak $x_1,....,x_n$ n tane gözlemin bir örneğinden

$$\hat{f}_n(x) = \frac{1}{n}\sum_{i=1}^{n} \delta_n(x - x_i)$$

formunun tahmincileri vasıtasıyla tahmin edilmiş olur. Burada $\delta_n(.)$ karesi integrallenebilir fonksiyon olarak alınmıştır.

$\hat{f}_n(x)$ tahmincisinin hata kareleri integrallanmış ortalaması (M.I.S.E.) J_n,

$$J_n = \mathrm{E}\left(\int \left(\hat{f}_n(x) - f(x)\right)^2 dx\right)$$

şeklinde tanımlanmıştır.

2.3.1.6-Rosenblatt Kernel Yoğunluk Tahmincisi

Populasyondan çekilen bağımsız gözlemlerin bir örneği $X_1,..,X_n$ vasıtasıyla $f(x)$'in tahmini $\hat{f}_n(x)$;

$$\hat{f}_n(x) = \frac{1}{nh(n)}\sum_{i=1}^{n} W\left(\frac{x - X_i}{h(n)}\right)$$

olarak tayin edilmiştir (Rosenblatt, 1971). Burada ağırlık fonksiyonu W (kernel K) sınırlı fonksiyondur ve integrali bire eşittir. Bandwith h(n) bir pozitif sabittir (Eddy,1980).

2.3.1.7-Parzen Kernel Yoğunluk Tahmincisi

Dağılış fonksiyonu $F(x) = P[X \leq x]$ sürekli olan tesadüfi değişken X gibi $X_1, X_2,...,X_n$ bağımsız ve özdeş dağılmış tesadüfi değişkenler olsun, ihtimal yoğunluk fonksiyonu $f(x)$ ile

$$F(x) = \int_{-\infty}^{x} f(x')dx'$$

olur.

Verilen belli bir x noktasında dağılış fonksiyonunun F(x) değerinin bir tahmini için $F_n(x) = \frac{i}{n}$ örnek dağılış fonksiyonunun alınmasıyla doğal olur ve $F_n(x)$ esas olarak bir binomial olarak dağılmış tesadüfi değişkendir ortalama ve varyansı sırasıyla

$E[F_n(x)]=F(x)$

$Var[F_n(x)]=\{(1/n)F(x)\}\{1-F(x)\}$ şeklinde verilmiştir.

Bir ağırlık fonksiyonu olan K(y)

$$K(y)=\frac{1}{2}, \quad |y|\leq 1,$$
$$=0, \quad |y|>1.$$

şeklinde tanımlanmış olsun. O zaman $f_n(x)$'in tahmini

$$\hat{f}_n(x) = \int_{-\infty}^{\infty} \frac{1}{h} K\left(\frac{x-y}{h}\right) dF_n(y) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x-X_i}{h}\right)$$

şeklindedir (Parzen,1962).

Parzen(1962), $\hat{f}_n(x)$'in özelliklerini incelemiş ve tahminci $\hat{f}_n(x)$'in varyansını

$$Var\left[\hat{f}_n(x)\right] = n^{-1} Var\left[h^{-1}K\left((x-X)/h\right)\right]$$

ve ortalamasını

$$E\left[\hat{f}_n(x)\right] = E\left[\frac{1}{h(n)} K\left(\frac{x-X}{h(n)}\right)\right] = \int_{-\infty}^{\infty} \frac{1}{h(n)} K\left(\frac{x-y}{h(n)}\right) f(y)dy$$ şeklinde bulmuştur.

Hata Kareler Ortalaması

$$E\left|\hat{f}_n(x) - f(x)\right|^2 = \sigma^2\left[\hat{f}_n(x)\right] + b^2\left[\hat{f}_n(x)\right]$$ olarak belirtilmiştir.

Burada

$$\sigma^2\left[\hat{f}_n(x)\right] = Var\left[\hat{f}_n(x)\right]$$

ve tahminin sapması

$$b\left[\hat{f}_n(x)\right] = E\left[\hat{f}_n(x)\right] - f(x)$$

şeklindedir.

2.3.2 (k-Enyakın Komşu Tahmini (k-Nearest Neighbor Estimate)

En yakın komşu tahminlerinin yapısı kernel tahminlerininkinden farklıdır. Kernel tahmini $\hat{m}_h(x)$, x civarındaki sabit yakın komşulukta cevap değişkenlerinin bir ağırlıklı ortalaması gibi tarif edilmiştir ve bandwith h ve kernel K ile belirtilmiştir. k-nearest neighbor(k-NN) tahmini değişen bir komşulukta bir ağırlık ortalaması olur. Bu komşuluk mesafedeki k-enyakın komşuların arasında olan x-değişkenleri vasıtasıyla tarif edilir (Hardle, 1997). Regresyonda k-NN smoother veya k-NN regresyon fonksiyon tahmini $\hat{m}$ aşağıdaki gibi tarif edilir (Györfi, 1981; Devroye, 1978),

$$\hat{m}_k(x) = n^{-1}\sum_{i=1}^{n} W_{ki}(x)Yi\,;$$

burada $\{W_{ki}(x)\}_{i=1}^{n}$, J_x indekslerinin kümesi vasıtasıyla tarif edilmiş bir ağırlık kümesidir.

J_x={i: X_i x'e en yakın gözlemlerin bir tanesi}

Komşu gözlemlerin indeksleri kullanılarak k-NN ağırlık kümesinin tertip edilişi şöyledir:

$$W_{ki}(x)=\begin{cases} n/k & i \in J_x \text{ ise} \\ 0\,, & \text{aksi takdirde (yani } i \notin J_x) \end{cases}$$

Örnek: Hardle (1997), $\{(X_i, Y_i)\}_{i=1}^{5}$ {(1,5), (7,12), (3,1), (2,0), (5,4)} olsun; k=3 ve x=4 için k-NN tahmini; $\hat{m}_k(x)$ hesaplanabilir. x ile birbirine daha çok yakın k gözlemler son üç veri noktasıdır, bu yüzden

$J_x=J_4=\{3,4,5\}$ ve böylece

$W_{k1}(4)=0$

$W_{k2}(4)=0$

$W_{k3}(4)=1/3$

$W_{k4}(4)=1/3$

$W_{k5}(4)=1/3$

olur ve sonuçta $\hat{m}_3(4) = (1+0+4)/3 = 5/3$ tür.

Stone (1977) ve Devroye(1978)'e göre, k-NN ağırlık fonksiyonları;

Üniform k-NN ağırlık fonksiyonu

$$W_{ki}(x)=\begin{cases}1/k\ , & 1\le i\le k\\ 0\ , & i>k\end{cases}\quad \text{için,}$$

üçgen (triangular) k-NN ağırlık fonksiyonu

$$W_{ki}(x)=\begin{cases}(k-i+1)/b_k\ , & 1\le \mathrm{i}\le \mathrm{k}\\ 0\ , & \mathrm{i}>\mathrm{k}\end{cases}\quad \text{için,}$$

burada b_k=k(k+1)/2 dir,

quadratik k-NN ağırlık fonksiyonu

$$W_{ki}(x)=\begin{cases}(k^2-(i-1)^2)/b_k\ , & 1\le i\le k\\ 0\ , & i>k\end{cases}\quad \text{için,}$$

burada b_k=k(k+1).(4k-1)/6

şeklindedir.

X değişkeninin eşit uzaklıklı bir gridden(ızgaradan) şansa bağlı seçildiği bir deneyde k-NN ağırlıkları kernel ağırlıklarına eşittir. k=2nh olsun ve sınıra çok yakın olmayan bir x için üniform kernel K(u)=(1/2)I(|u|≤1) için $\{W_{ki}(x)\}$ ile $\{W_{hi}(x)\}$ mukayese edilir. Gerçekten $\mathrm{i}\in \mathrm{J}_x$ için;

$$W_{ki}(x)=\frac{n}{(2nh)}=\frac{1}{2}h^{-1}=W_{hi}(x)$$

k smoothing parametresi tahmin edilen eğrinin düzgünlüğünün derecesini belirler. k-NN kernel smoethirler için bandwith ile benzer rol oynar (Hardle, 1997).

k-NN tahminin önemli bir avantajı X değişkenlerinin sınıflandırılmış düzenlemeleri boyunca x devam ettiği (ilerlediği) zaman k-NN tahmininin oldukça kolay yapılabilmesidir (Devroye,1978).

Daha önce tarif edilen $W_{ki}(x)$ üniform ağırlıklar ile tayin edilen neighborhood vasıtasıyla en küçük kareler doğrusunun α ve β parametreleri aşağıdaki şekilde hesaplanır (Green et al.,1985; Hardle,1997);

$$\alpha_{xi}=\hat{m}_k(x_i)-\beta\bar{\mu}_{xi}\quad ,$$

$$\beta_{x_i} = \frac{C_{x_i} - \overline{\mu}_{x_i} . m_k(X_i)}{V_{x_i} - \overline{\mu}_{x_i}^2},$$

$$\overline{\mu}_x = k^{-1} \sum_{i \in J_x} X_i ,$$

$$C_x = \sum_{i \in J_x} X_i Y_i ,$$

$$V_x = \sum_{i \in J_x} X_i^2 .$$

$\alpha_{x_i} = W_{ki}(x)$ ağırlıklara göre elde edilen regresyon doğrusunun y eksenini kestiği nokta,

$\beta_{xi} = W_{ki}(x)$ ağırlıklara göre elde edilen regresyon doğrusunun eğimi,

$M_x = i \in Jx$ şartına uygun x_i değerlerinin ortalaması

$Cx = i \in Jx$ şartına uygun x_i ve y_i değerlerinin çarpımlı toplamı

$V_x = i \in Jx$ şartına uygun x_i değerlerinin karelerinin toplamıdır.

Stute (1984), k-NN smoothinglerin bir çeşidi olan simetrikleştrilmiş en yakın komşu tahmincileri üzerinde çalışmıştır. Bir x örneğinin deneysel (ampirik) dağılımını F_n ile gösterilir, h'da sıfıra giden bandwith olsun. Buna göre Stute (1984) tarafından önerilen tahminci

$$\hat{m}_{k(n)}(x) = (nh)^{-1} \sum_{i=1}^{n} K\left(\frac{F_n(x_i) - F_n(x)}{h} \right) Y_i \quad \text{dir.}$$

Bu tahmin aynı zamanda bir nearest neighbor tahmini olur, fakat neighborlar $\{X_i\}_{i=1}^{n}$ nin amprik dağılış fonksiyonuna dayandırılmış uzaklığında tarif edilir. Böylece bir ağırlık sırası

$$W_{k(h)}(x) = K_h(F_n(X_i) - F(x))$$

kullanılır.

2.3.3-Spline Smoothing

Nonparametrik regresyon modeli

$Y_i = g(x_i) + e_i$, $i=1,2,.....,n$

şeklinde düzenlensin. Burada x_i değerleri $x_1<x_2<.....<x_n$ şeklinde olan bilinen değerlerdir, ve e_i'ler ortalaması 0 variyansı σ^2 olan bağımsız ve özdeş dağılmış hatalardır. Bir g eğrisi için «veriye bağlılığın» genel bir ölçüsü hata kareler toplamı

$$\sum_{i=1}^{n}(Y_i - g(x_i))^2$$

dir. g fonksiyonel formda sınırsız herhangi bir eğri olsun, o zaman bu aralık (mesafe) ölçüsü verilerin aradeğerini bulan herhangi bir g tarafından sıfıra indirgenebilir (Clark,1977)

Genel variyasyonun ölçülmesinde birkaç yol vardır. Bir tanesi kabaca ölçüleri tanımlar, mesela 1., 2., ...,m. türevlerin alınması gibi. Bu duruma ceza (roughness penalty)

$$\int (g^m(x))^2 dx$$

Denir (Clark,1977; Heckman,1986; Simonoff,1998). Hata kareler toplamı ve roughness penalty'nin

$$S_\lambda(g) = \sum_{i=1}^{n}(Y_i - g(x_i))^2 + \lambda \int (g^m(x))^2 dx$$

toplam ifadesinin minimize edilmesiyle g'nin tahmini elde edilir. Burada λ bir smoothing parametredir. Penalti fonksiyon tahminciler için λ ile gösterilen smoothing parametre, g eğrisinin roughness ve hata arasındaki değişim hızını temsil eder. Ayrıca bir nonparametrik regresyon tahmininin pürüzsüzlüğünü kontrol eder (Simonoff,1998).

Şayet m=2 ise S_λ, $\{x_i\}$ bağları(kümeleri) ile bir doğal cubic spline fonksiyon olur. Cubic spline fonksiyonunu Reinsch(1967) detaylı olarak incelemiştir. Daha sonra Silvermen(1984) cubic spline durumunu ele almıştır. $[a,b]=[X_{(1)},X_{(2)}]$ aralığında iki kere türevi alınabilen fonsiyonların sınıfları üzerinde minimize $S_\lambda(.)$ problemi(durumu) cubic spline olarak tarif edilen tek çözüm $\hat{m}_\lambda(x)$'e sahiptir (Hardle,1997). Whittaker

1923'te bu smoothing işlemini gözlemlerin ayarlanması veya derecelendirilmesi olarak tanımlamıştır (Silverman,1984). Tahmin edilen eğri $m_\lambda(.)$'nın özellikleri: $\hat{m}_\lambda(x)$ birbirini takip eden iki x-değeri arasında cubic polinomiyal olur; gözlenen x_i noktalarında eğri $\hat{m}_\lambda(.)$ ve $\hat{m}_\lambda(.)$'nin ilk iki türevi süreklidir, fakat üçüncü türevde kesiklilik olabilir. Diğer bir özellik olarak, $X_{(1)}$ ve $X_{(n)}$ sınır noktalarında $\hat{m}_\lambda(x)$ in ikinci türevi sıfırdır (Hardle,1997).

Regresyon eğrisinin cubic spline tahmincisi $\hat{g}$ aşağıdaki formülün g fonksiyonlarının minimize edilmesiyle tarif edilmiş olur,

$$\lambda\int (g''(X)^2 dx + \sum_{i=1}^{n}(Y_i - g(X_i))^2$$

Parametre $\lambda>0$, düzlük arasında değiş-tokuşun kontrolünü sağlayan smoothing parametre, Y_i ve $g(x_i)$ arasındaki sapmaların karelerinin toplamı ve $\int g''^2$ ile ölçülür (Silverman, 1967).

Hardle (1997)'nin ifade ettiğine göre, $\{Y_i\}^n_{i=1}$ veri vektörü n koordinat vektörlerinin bir linear kombinasyonu gibi yazıldığı taktirde $W_{\lambda_i}(x)$ ağırlıklarının ortaya çıktığı kolay bir şekilde anlaşılır. Şöyle ki

$$\hat{m}_\lambda(x) = n^{-1} \sum_{i=1}^{n} W_{\lambda_i}(x) Y_i \, .$$

Böylece spline Y-gözlemlerinde lineer olur. Ağırlık fonksiyonu, n örnek büyüklüğüne eşit olan bir cevap değişken seti hariç, tümü sıfır olan Y-değerlerine spline smoothig metodun uygulanmasıyla çizim yapılabilir. Bununla birlikte, $\{W_{\lambda_i}(x)\}_{i=1}^{n}$ nin fonksiyonel formunun açık bir şekilde yazılması zordur ve noktaların düzenlenmesi ve smoothing parametre λ üzerine bağlılık aşırı derecede karışıktır. $W_{\lambda i}(.)$; $(X_1,0)....(X_i,0)....(X_n,0)$ verileri için spline olur. Daha genel olarak $W_\lambda(.,t)$ için nokta t'de spline tarif edilir.

Silverman (1984), kernel smoothing ve spline smoothing arasındaki ilişkileri(bağları) belirlemek için $W_\lambda(.,t)$ formunu incelemiştir. Silverman(1984)'e göre, uygun şartlar altında, ağırlık fonksiyonu $W_\lambda(.,t)$ yaklaşık olarak aşağıda formülize edilen kernel fonksiyon K_s'nin benzeridir.

$K_s(u)=(1/2)\exp(-|u|/\sqrt{2})\sin(|u|/\sqrt{2}+\pi/4)$.

2.3.4. Bandwith Seçimi (Bandwith h'ın Bulunuşu)

Kernel smoothing için bandwith parametresinin uygun bir şekilde seçilmesi gerekir. Çünkü $f_n(x)$'in ortalaması, varyansı v.b. çeşitli özellikler bandwith parametresi h'a bağlıdır. n'nin bir fonksiyonu için seçilmesi gereken bandwith h, n→∞ için sıfıra giden, yani h→0 olur. Uygun bir şekilde bandwith seçimi çeşitli araştırmacılar tarafından aşağıdaki şekilde izah edilmiştir.

Priestley ve Chao (1972), $f_n(x)$'in ortalamasını

$\delta = x_i - x_{i-1}$

$$E[f_n(x)] = \int_{-\infty}^{\infty} \frac{1}{h_n} K\left(\frac{x-u}{h_n}\right) f(u)du = \delta \sum_{i=1}^{n} f(x_i) K\left(\frac{x-x_i}{h_n}\right)$$

ve varyansını

$$Var(f_n(x)) = \frac{\delta^2\sigma^2}{h_n^2} \sum_{i=1}^{n} K^2\left\{\frac{x-x_i}{h_n}\right\} \approx \frac{\delta\sigma^2}{h_n} \int_{-\infty}^{\infty} K^2(u)du$$

şeklinde ifade etmişlerdir.

Rosenblatt (1971), hata kareler ortalaması ve sapmayı (biası) aşağıdaki gibi ifade etmiştir

Hata kareler ortalaması;

$$E(f_n(x) - f(x))^2 = \sigma^2[f_n(x)] + b^2[f_n(x)]$$

burada

$$\sigma^2[f_n(x)] = \text{var}[f_n(x)].$$

Ve

$$b[f_n(x)] = \mathrm{E}[f_n(x)] - f(x)$$
$$= (2\pi)^{-1}\int_{-\infty}^{\infty} e^{-iux}\{k(hu) - 1\}\varphi(u)du$$

tahminin sapması olur. Burada,

$$\varphi_n(u) = \int_{-\infty}^{\infty} e^{iux} dF_n(x)$$

$$k(u) = \int_{-\infty}^{\infty} e^{-iux} K(x)dx \text{ dir.}$$

$F(x)$ ise, $f(x)$ e ait dağılış fonksiyonu olup ve $F_n(x) = i/n$ dir (Parzen, 1962).

Benedetti (1977)'de, hata kareler ortalaması için asimtotik ifadeyi

$$\left(\sigma^2/h_n\right)\delta\int_{-\infty}^{\infty} K^2(u)du + h_n^{2r}\left|k_r f^r(x)\right|^2$$

şeklinde ifade etmiştir.

Karakteristik katsayı k_r ve karakteristik üs r'ye sahip olan $f_n(x)$ formu ile kullanılmış K(x) fonksiyonunun k(u) transformunu farzeden tahmin $f_n(x)$'in hata kareler ortalaması için bir yaklaşım ifadesi

$$\mathrm{E}\left|f_n(x) - f(x)\right|^2 \approx \left[f(x)/nh\right]\int_{-\infty}^{\infty} K^2(x)dx + h^{2r}\left|k_r f^{(r)}(x)\right|^2$$

yazılabilir. Burada K(x) ağırlık fonksiyonunun Fourier transformu

$$k(u) = \int_{-\infty}^{\infty} e^{iux} K(x)dx \text{ dir (Parzen,1962).}$$

n'nin sabit bir değeri için minimize edilen hata kareler ortalaması bandwith (h) ın değerini verir. Yani h'ın değerini bulabilmek için hata kareler ortalaması minimize edilir. Böylece Benedetti (1977)'ye göre

$$h_n = \left\{ \sigma^2 \delta \int K(u)du \Big/ 2r\left|k_r f^r(x)\right|^2 \right\}^{1/(1+2r)}$$

olur. Bundan dolayı hata kareler ortalaması

$$(2r+1)\left\{\sigma^2 \int K^2(u)du\right\}^{2r/(1+2r)} (2r/\delta)^{-2r/(1+2r)} \left\{k_r f^r(x)\right\}^{2/(1+2r)}$$

olur ve $n^{-2r/(1+2r)}$ için hata kareler ortalaması sıfır olur.

Parzen (1962)'e göre

$$h = \left\{ f(x) \int_{-\infty}^{\infty} K^2(x)dx \right\} \Big/ \left\{ n2r\left|k_r f^{(r)}(x)\right|^2 \right\}^{1/(2r+1)} \text{ dir.}$$

O zaman hata kareler ortalaması;

$$\mathrm{E}\left|f_n(x) - f(x)\right|^2 \approx (2r+1)\left\{\left[f(x)\Big/n2r\right]\int_{-\infty}^{\infty} K^2(x)dx\right\}^{2r/(1+2r)} \left|k_r f^{(r)}(x)\right|^{2/(1+2r)}$$

ve $n^{-2r/(1+2r)}$ için sıfır olur. Bu özellikler Rice (1984) tarafından da incelenmiştir.

Parzen tarafından bulunan hata kareler ortalamasının asimtotik ifadesi ile Priestley ve Chao tarafından bulunan hata kareler ortalamasının asimtotik ifadesinin özdeş olduğu görülür. Ayrıca bandwith h'ın değerini Rosenblatt(1971); Benedetti(1977) yukarıdaki açıklamanın benzeri olarak bulmuşlardır.

3. MATERYAL VE METOD

Bu kısımda, nonparametrik regresyon parametrelerinin tahminini amaçlayan metodların bazılarının uygulanışı ve bu uygulamada kullanılan materyal açıklanacaktır.

3.1.Materyal

Bu çalışmada, materyal kapsamında ifade edilmesi gereken unsurlar nonparametrik regresyonun uygulanmış olduğu verilerdir. Nonparametrik regresyon metodlarının uygulamasının yapılması için, Atatürk Üniversitesi Ziraat Fakültesi Koyunculuk Şubesi'nde 1998 yılında elde edilmiş kuzuların doğum ve sütten kesim ağırlık verileri kullanılmıştır.

3.2 Metod

Nonparametrik regresyonda parametre tahminlerinin uygulaması yapılan ve daha elverişli oldukları bildirilen bazı metodlar aşağıda verilmiştir.

Regresyon parametrelerinin tahmin metodları ve uygulaması için

1-Theil Metodu

2-Brown -Mood Metodu

3-En Küçük Kareler Metodu

4-k-Nearest Neighbor Smoothing

metodların uygulaması yapılmıştır. Bu metodlarla ilgili gerekli teorik açıklamalar önceki kısımda literatüre dayalı olarak verildiği için burada tekrar üzerlerinde durulmamıştır.

4. SONUÇ VE TARTIŞMA

4.1. TAHMİN METODLARI VE KARŞILAŞTIRILMALARI

4.1.1. Theil Metodu Uygulaması

Atatürk Üniversitesi Koyunculuk İşletmesi'nden n=14 İvesi kuzusundan alınan ölçümler tablo 4.1.1'de verilmiştir.

Tablo:4.1.1. Doğum Ağırlıkları (Kg) X ve Sütten Kesim Ağırlıkları (Kg) Y

X	3.4	5.2	4.5	5.4	5.0	4.6	4.9	5.7	4.3	4.1	4.7	4.0	4.4	5.1
Y	18.0	19.5	18.0	16.0	18.5	16.0	17.0	21.0	14.0	15.5	15.0	14.0	19.0	19.5

i<j olmak üzere N=$\binom{14}{2}$=91 adet $b_{ij} = \dfrac{Y_j - Y_i}{X_j - X_i}$ eğimleri tablo 4.1.2 de verilmiştir

Tablo:4.1.2. b_{ij} Eğimlerinin Sıraya Dizilmiş Durumu

-20.00	-4.44	-0.83	0.71	1.54	3.33	4.58	6.25	11.25
-17.50	-4.00	-0.83	0.83	1.82	3.33	5.00	6.43	11.68
-15.00	-3.57	-0.67	0.88	1.88	3.33	5.00	6. 67	11.68
-15.00	-3.00	0.00	1.00	2.14	3.44	5.00	6.88	12.50
-13.33	-2.50	0.00	1.00	2.50	3.57	5.00	7.00	12.50
-11.67	-2.31	0.00	1.00	2.50	3.64	5.00	8.00	15.00
-10.00	-2.22	0.00	1.30	2.50	4.00	5.83	8.33	15.00
-10.00	-2.00	0.31	1.43	2.50	4.12	6.00	9.00	16.67
-7.50	-1.67	0.38	1.43	3.00	4.50	6.11	10.00	20.00
-6.67	-1.00	0.63	1.43	3.33	4.55	6.25	10.00	50.00
-6.25								

β tahmincisi bütün bu b_{ij} değerlerinin medyanıdır. Yani,

$\hat{\beta}$=medyan{b_{ij}}

Bu değerde 91 rakamdan,

$\hat{\beta}$=2.50

olarak bulunur.

Regresyon doğrusunun sabiti olan α değeri ise $a_i = y_i - \hat{\beta}x_i$ değerlerinin medyanı olur. Buna göre α değerini hesaplamak için gerekli olan a_i değerleri aşağıdaki gibi bulunur

a_1=18.0-2.50x3.4=9.50

a_2=19.5-2.50x5.2=6.50

a_3=18.0-2.50x4.5=6.75

a_4=16.0-2.50x5.4=2.50

a_5=18.5-2.50x5.0=6.00

a_6=16.0-2.50x4.6=4.50

a_7=17.0-2.50x4.9=4.75

a_8=21.0-2.50x5.7=6.75

a_9=14.0-2.50x4.3=3.25

a_{10}=15.5-2.50x4.1=5.25

a_{11}=15.0-2.50x4.7=3.25

a_{12}=14.0-2.50x4.0=4.00

a_{13}=19.0-2.50x4.4=8.00

a_{14}=19.5-2.50x5.1=6.75

α=medyan{a_i} den

α=5.625

olarak hesaplanır. Buna göre regresyon doğrusunun denklemi ise

Y=5.625+2.5X

olarak tesbit edilir.

β katsayısının önemlilik testi yapıldığında, Theil metodu ile H_0:β=β_0 hipotezini test etmede kullanılan prosedür Kendall'ın tau istatistiğine dayanır.

Hipotezler;

H_0:β=0, H_1:β≠0

şeklinde kurulup test edilsin. $\beta_0=0$ olduğundan, her bir çift için $Y_i-\beta_0X_i$ değeri Y_i değerine eşit olacaktır. (X_i, $Y_i-\beta_0X_i$) müşahede çiftlerini, en küçük X değerinin bulunduğu çift birinci sırada olacak şekilde X değerlerine göre sıraya konulur. Böylece X'ler tabii sıralama içerisinde olur. Herbir $Y_i-\beta_0X_i$ değeri, altında yer alan $Y_i-\beta_0X_i$ değeri ile karşılaştırılır. Alttaki $Y_i-\beta_0X_i$ değeri üstteki $Y_i-\beta_0X_i$ değerinden büyükse bu $Y_i-\beta_0X_i$ değerleri çiftinin tabii sıralama içerisinde olduğu, alttaki $Y_i-\beta_0X_i$ değeri üstteki $Y_i-\beta_0X_i$ değerinden küçükse bu $Y_i-\beta_0X_i$ değeri çiftinin de tersine bir tabii sıralama içerisinde olduğu kabul edilir. Tabii sıralamadaki X_i ve $Y_i-\beta_0X_i$ değerleri ile tabii sıralamadaki $Y_i-\beta_0X_i$ çiftleri ile tersine tabii sıralamadaki $Y_i-\beta_0X_i$ çiftleri sayısı tablo 4.1.3 de verilmiştir.

Tablo:4.1.3. Tablo 4.1.1.'deki doğum ağırlıkları ile sütten kesim ağırlıkları arasında tau değerini hesaplamak için gerekli olan hesap tablosu.

X_i	$Y_i-\beta_0X_I$	Tabii Sıralamadaki $Y_i-\beta_0X_i$ Çiftleri	Tersine Tabii Sıralamadaki $Y_i-\beta_0X_i$ Çiftleri
3.4	18.0	5	7
4.0	14.0	11	0
4.1	15.5	9	2
4.3	14.0	10	0
4.4	19.0	3	6
4.5	18.0	4	4
4.6	16.0	5	1
4.7	15.0	6	0
4.9	17.0	4	1
5.0	18.5	4	1
5.1	19.5	1	1
5.2	19.5	1	1
5.4	16.0	1	0
5.7	21.0	0	0
		P=64	Q=24

Tabii sıralamadaki Y_i-β_0X_i çiftleri sayısını P ile tersine tabii sıralamadaki Y_i-β_0X_i çiftleri sayısını Q ile gösterilirse tau hesap değeri

$\hat{\tau} = \dfrac{P-Q}{n(n-1)/2} = \dfrac{64-24}{14(14-1)/2} = 0.440$ olarak bulunur.

Bu değer n=14 ve $\alpha/2$=0.025'e göre Kendall'ın tau katsayıları cetvelinden bulunan 0.407 değerinden büyük yani $\hat{\tau} = 0.440 > \tau^* = 0.407$ olduğundan H_0 hipotezi red edilir ve $\hat{\tau}$ değerinin önemli olduğu söylenir. Sonuçta İvesi koyun populasyonunda doğum ağırlığı ile sütten kesim ağırlığı arasında önemli doğrusal bir ilişkinin olduğu söylenebilir.

Theil metoduna göre β katsayısının güven sınırları tahmin edilirse; N=$\binom{14}{2}$=91 adet b_{ij} değerlerinin sıraya dizilmiş durumu tablo 4.1.2'de verilmiştir. n=14 ve $\alpha/2$=0.025 için Kendall'ın tau katsayısı değerleri tablosuna bakıldığında,

$S_{\alpha/2}$=37 ve

$C_{\alpha/2}$=$S_{\alpha/2}$-2=37-2=35 olarak bulunur. Buna göre k değeri,

$k = \dfrac{N-C_{\alpha/2}}{2} = \dfrac{91-35}{2} = 28$ olarak bulunur.

Güven aralığının alt sınırı küçükten büğüğe doğru sıralanan 28. b_{ij} değeridir. Yani

$\hat{\beta}_A = 0.00$

ve güven aralığının üst sınırı ise büyükten küçüğe doğru sıralana 28. b_{ij} değeridir

$\hat{\beta}_Ü$ =5.00

olarak bulunur. Yani %95 güven katsayısıyla güven aralığı

C(0.00<β<5.00)=0.95

şeklinde ifade edilir.

4.1.2.Brown-Mood Metodun Uygulanışı

Tablo 1'deki veriler arasındaki lineer ilişki denkleminin sabitini ve doğrunun eğimini Brown-Mood metodu ile bulmak için; ilk önce X_i ve Y_i değerleri dağılma diyagramında gösterilir. Daha sonra X değerlerinin medyanı olan 4.65 değerinden Y eksenine parelel olarak yukarıya doğru dikey bir çizgi çizilir. Çizilen bu çizgi verileri iki gruba ayırır ve genel medyanın altındaki X ve Y medyan değerleri 4.3 ve 16.0 olarak bulunur ve genel medyanın üstündeki X ve Y medyan değerleri ise 5.1 ve 18.5 dir. X_i ve Y_i değerlerinin dağılma diyagramı ve medyan X doğrusu grafik 4.1.2.1'de gösterilmiştir. Her iki gruptaki medyan X ve medyan Y'nin kesişim noktaları işaretlenerek bir doğru ile birleştirilir. Böylece nihai tahminden önceki regresyon doğrusu elde edilir. Bu doğrunun eğimi,

$$\hat{\beta} = \frac{Y_{1(Medyan)} - Y_{2(Medyan)}}{X_{1(Medyan)} - X_{2(Medyan)}} = \frac{18.5 - 16.0}{5.1 - 4.3} = 3.125$$

nihai tahminden önceki regresyon doğrusunun denklemi ise ,

Y=2.9+3.125X

olarak bulunur.

Noktaların elde edilen Y=2.9+3.125X denklemi doğrusundan dikey sapmalarının medyanı her iki grupta da sıfırdan farklı olduğundan, her iki gruptaki sapmaların medyanı sıfır olana kadar doğrunun pozisyonu değiştirilir ve böylece nihai doğru elde edilir. Nihai doğru üzerinde iki nokta işaretleyerek ve bu nokta değerlerini aşağıdaki formülde yerine koyarak;

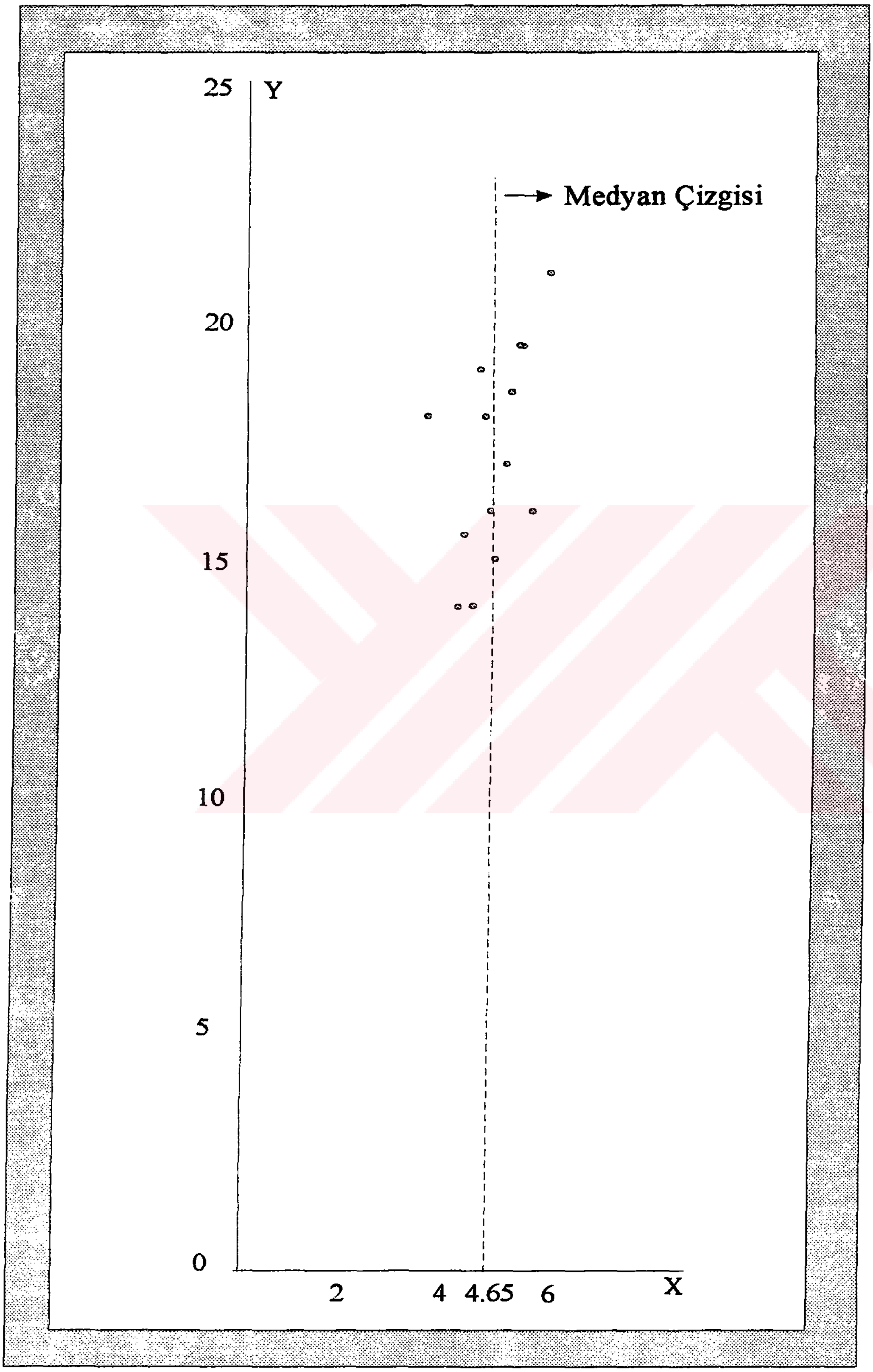
$$\beta' = \frac{Y_1 - Y_2}{X_1 - X_2} = \frac{21 - 11}{5.8 - 2.8} = \frac{10}{3} = 3.333$$

nihai doğrunun eğimi bulunmuş olur ve elde edilen regresyon doğru denklemi

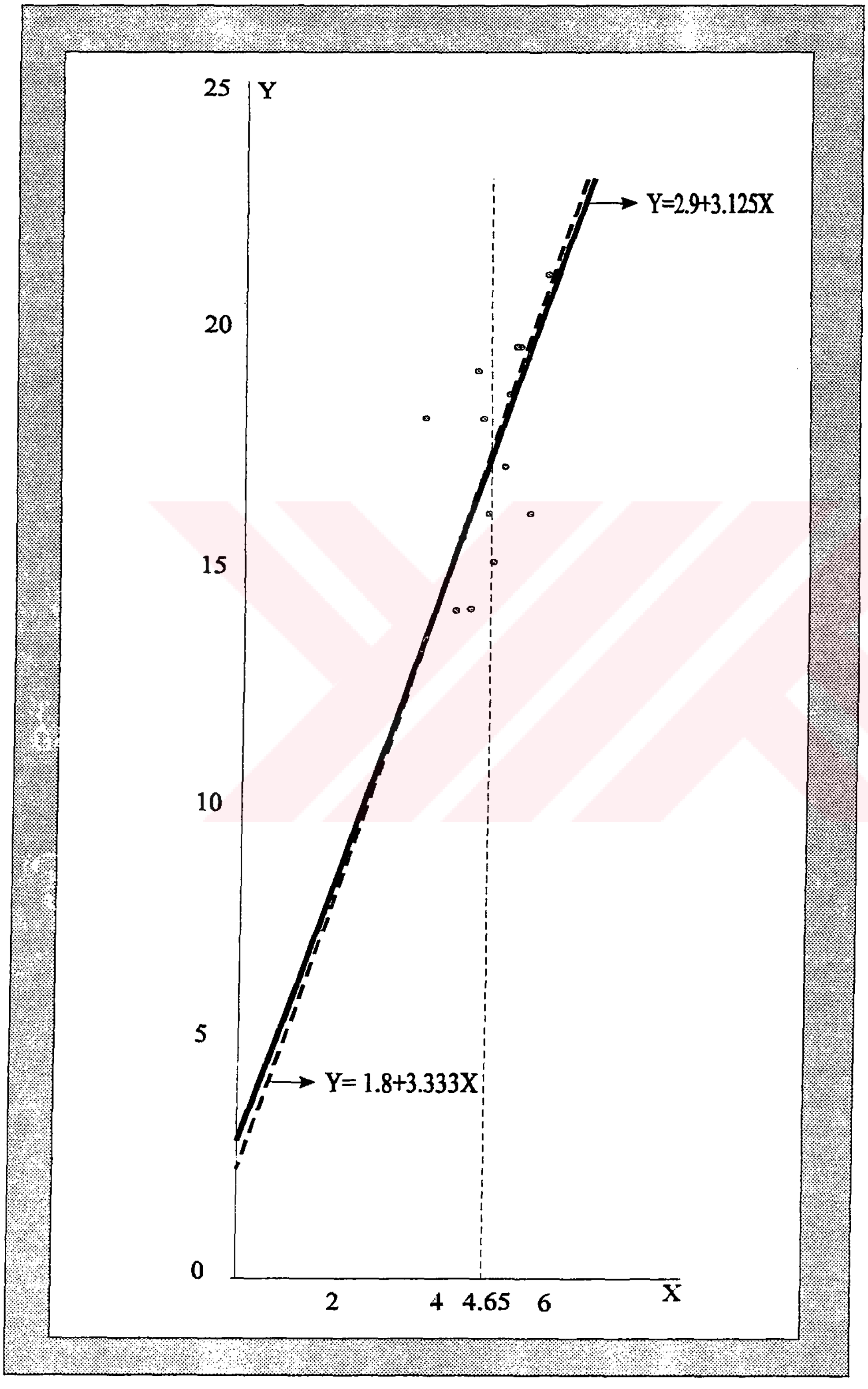
$\hat{Y}$=1.8+3.333X

yazılır. Bu eşitliğe ait doğru grafik 4.1.2.2'de nihai tahminden önceki regresyon doğrusuyla birlikte göterilmiştir. Noktaların ilk regresyon doğrusundan ve son regresyon doğrusundan dikey sapmaları Tablo 4.1.4'de verilmiştir.

Grafik 4.1.2.1 X_i ve Y_i Değerlerinin dağılma diyagramı



Grafik:4.1.2.2. Nihai tahminden önceki ve sonraki regresyon doğrularının grafiği



Tablo 4.1.4 Noktaların ilk ve son regresyon doğrularından dikey sapma değerleri

	Noktaların İlk Regresyon Doğrusundan Dikey Sapma Değerleri	Noktaların Nihai Regresyon Doğrusundan Dikey Sapma Değerleri
Genel X Medyanı Altındaki Nokta Grubu	4.7	4,9
	2.6	2,6
	1.3	1,2
	0.0	0,0
	-0.9	-0.9
	-1.1	-1,1
	-2.1	-1,9
Medyan	0.0	0.0
Genel X Medyanı Üstündeki Nokta Grubu	1,0	0.7
	0,7	0.4
	0,6	0.3
	0.3	0.0
	-0.8	-1.2
	-2,2	-2.3
	-3,2	-4.2
Medyan	0.3	0.0

α ve β katsayılarının önemlilik testi yapılırken hipotezler,

H_0:α=1.8, β=3.333

H_1=$\alpha \neq 1.8$, $\beta \neq 3.333$

şeklinde kurulur ve test istatistiği;

$$X^2 = \frac{8}{n}\left[\left(n_1 - \frac{n}{4}\right)^2 + \left(n_2 - \frac{n}{4}\right)^2\right]$$

formülüne göre hesaplanır. Formüldeki n_1 ve n_2 değerleri grafik 4.1.2.2. de dağılma diyagramına bakılarak; X değerleri medyanından çizilen dikey çizginin sol tarafındaki ve ilk regresyon doğrusunun üzerindeki noktaların sayısı n_1=3, X değerleri medyanından çizilen dikey çizginin sağ tarafındaki ve ilk regresyon doğrusunun üzerindeki noktaların n_2=4 olarak bulunup ve bu değerler yukarıdaki formülde yerine konulduğunda

$$X^2 = \frac{8}{14}\left[\left(3-\frac{14}{4}\right)^2 + \left(4-\frac{14}{4}\right)^2\right] = 0.286$$

olarak bulunur. Bu değer 0.05 hata seviyesinde $X_{0.05.2}$=5.991 değerinden küçük olduğundan H_0 hipotezi reddedilemez ve örneğin Y=1.8+3.333X regresyon doğrusuna uyan bir populasyondan çekildiği kabul edilebilir.

4.1.3.En Küçük Kareler Metodunun Uygulanışı

Tablo 4.1.1'deki verilere en küçük kareler metodu uygulanarak $Y=\alpha+\beta x_i+e_i$ regresyon modelindeki α ve β parametreleri tahmin edilecek olursa β parametre tahmincisi $\hat{\beta}$;

$$\hat{\beta} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2} = \frac{\sum x_i y_i - \frac{\sum x_i \sum y_i}{n}}{\sum x_i^2 - \frac{\left(\sum x_i\right)^2}{n}}$$ formülüne göre

=1.71

olarak bulunur ve α tahmincisi $\hat{\alpha}$ ise;

$\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x}$

=17.21-1.71x4.66

=9.24

olarak bulunur. Regresyon doğrusunun denklemi ise

Y=9.24+1.71X

olarak tesbit edilir.

β katsayısının önemlilik testi için hipotezler

$H_0:\beta=0 \quad H_1:\beta>0$

şeklinde kurulur ve regresyon katsayısının hipotez testi için yazılacak t kritik değeri

$t = \frac{\hat{\beta}-\beta}{S_b} = \frac{1.71-0}{0.91}$ =1.879 olarak bulunur. Bu değer $t_{0.05,12}$=1.782 değerinden büyük olduğundan regresyon katsayısının %95 ihtimal seviyesinde önemli olduğuna karar verilir.

4.1.4.. k- Nearest Neighbor Smoothing

Tablo 4.1.1'deki verilere k-nearest neighbor smoothing metodunu uygulayarak regresyon parametreleri şöyle tahmin edelir. n= 14 tane veri vardır ve k=9, x=4.6 için $\hat{m}_k(x)$

$$\hat{m}_k(x) = n^{-1}\sum_{i=1}^{n} W_{ki}(x)Y_i$$

$J_x = J_{4.6} = \{4.1, 4.3, 4.4, 4.5, 4.6, 4.7, 4.9, 5.0, 5.1\}$

$$W_{ki}(x) = \begin{cases} n/k & i \in J_x; \\ 0 & i \notin J_x; \end{cases}$$

buna göre $W_{ki}(x)$' ler;

$W_{k1}(4.6)=0$

$W_{k2}(4,6)=0$

$W_{k3}(4.6)=14/9$

$W_{k4}(4.6)=0$

$W_{k5}(4.6)=14/9$

$W_{k6}(4.6)=14/9$

$W_{k7}(4.6)=14/9$

$W_{k8}(4.6)=0$

$W_{k9}(4.6)=14/9$

$W_{k10}(4.6)=14/9$

$W_{k11}(4.6)=14/9$

$W_{k12}(4.6)=0$

$W_{k13}(4.6)=14/9$

$W_{k14}(4.6)=14/9$

şeklinde olur ve k-NN smoothing ise

$\hat{m}_9(4.6)$=(1/14){(0x18.0)+(0x19.5)+((14/9)x18.0)+(0x16.0)+((14/9)x18.5)+((14/9)x16)
+((14/9)x17.0)+(0x21.0)+((14/9)x14.0)+((14/9)x15.5)+((14/9)x15.0)+(0x14.0)+((14/9)
x19.0)+((14/9)x19.5)

$\hat{m}_9(4.6)$=(1/14)(0+0+28.00+28.78+24.89+26.44+21.78+24.11+23.33+29.56+30.33

=16.94

olarak bulunur. Buna göre β katsayısı;

$$\beta_{X_i} = \frac{C_{X_i} - \overline{\mu}_{X_i}.\hat{m}_k(X_i)}{V_{X_i} - \overline{\mu}^2_{X_i}}$$ formülünden

$$\overline{\mu}_x = k^{-1}\sum_{i \in J_x} X_i$$

=(1/9)(4.1+4.3+4.4+4.5+4.6+4.7+4.9+5.0+5.1)

=4.62

$$C_x = \sum_{i \in J_x} X_i Y_i$$

=(4.1x15.5)+(4.3x14.0)+(4.4x19.0)+(4.5x18.0+(4.6x16.0)+(4.7x15.0)+(4.9x17.0)

+(5.0x18.5)+(5.1x19.5)

=707.7

$$V_x = \sum_{i \in J_x} X_i^2$$

=($4.1^2+4.3^2+4.4^2+4.5^2+4.6^2+4.7^2+4.9^2+5.0^2+5.1^2$)

=193.18

$$\beta_{X_i} = \frac{707.7 - 4.62x116.94}{193.18 - 4.62^2} = \frac{629.437}{171.836} = 3.663$$

olarak bulunur.

α katsayısı ise aşağıdaki formülden;

$$\alpha_{X_i} = \hat{m}_k(X_i) - \beta\overline{\mu}_{X_i}$$

= 16.94-3.663x4.62

=0.01694

olarak bulunur ve regresyon doğru denklemi ise

Y=0.01694+3.663X

şeklinde ifade edilir.

Theil, Brown-Mood ve En Küçük Kareler tahmin metodlarına göre elde edilen regresyon parametre değerleri ve regresyon doğrusunun eğimi olan β'nın önemlilik durumu tablo 4.1.5'de özet olarak gösterilmiştir.

Tablo 4.1.5. Çeşitli metodlara göre elde edilen regresyon parametre değerleri

Metodlar	α	β
Theil Metodu	5.625	2.500*
Browm-Mood Metodu	1.800	3.333*
En Küçük Kareler Metodu	9.240	1.710*

*:β katsayısı ($P<0.05$) de önemli

Tabloya bakıldığında β katsayısının üç metod da 0.05 hata seviyesinde önemli bulunduğu görülür.

Ayrıca Theil, Brown-Mood, En Küçük Kareler ve K-Nearest Neighbor Smoothing (k-NN) Metodlarına göre elde edilen Belirleme katsayısı (R^2) ve Hata Kareler Ortalama (HKO) değerleri tablo 4.1.6'da gösterilmiştir.

Tablo 4.1.6. Çeşitli metodlara göre elde edilen R^2 ve hata kareler ortalaması değerleri

Metodlar	R^2	Hata Kareler Ortalaması (HKO)
Theil	0.488	4.2670
Brown-Mood	0.869	5.0940
En Küçük Kareler	0.228	4.0112
K-Nearest Neigbor	0.404	5.5650

Belirleme katsayısı (Determinasyon katsayısı) adı verilen R^2, regresyon denkleminin gerçek verilere uyumunu ölçmede bir kriter olarak kullanılır. Belirleme katsayısı bağımlı değişken Y'lerde ortaya çıkan variyasyonun yüzde (%) kaçının bağımsız değişken X tarafından izah edildiğini gösterir. R^2 her zaman 0 ile 1 ($0\leq R^2\leq 1$) arasında değer alır. R^2'nin 1'e yakın değerleri regresyon denkleminin gerçek verilere iyi uyduğunu, 0' yakın değerleri ise regresyon denkleminin gerçek verilere uymadığını ifade eder. Belirleme katsayısı, regresyon kareler toplamı ($\sum(\hat{Y}_i - \bar{Y})^2$)'nın genel kareler toplamı ($\sum(Y_i - \bar{Y})^2$)'na bölünmesiyle bulunur. Yani

$$R^2 = \frac{\sum_{i=1}^{n}(\hat{Y}_i - \bar{Y})^2}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}$$

dir. R^2 değerlerine göre metodların karşılaştırmada, metodlara göre elde edilen regresyon doğrularının gerçek verilere ne derecede uyumlu olduğu tespit edilmeye çalışılır. Karşılaştırma yapılırken R^2 değeri büyük çıkan metoda göre elde edilen modelin gerçek verilere daha iyi uyum sağladığı söylenebilir. Herbir metoda göre elde edilen modelin gerçek verilere uyum durumuna göre metodlar arasındaki sıralama

Brown-Mood > Theil > k-Nearest Neighbor > En Küçük Kareler

şeklinde yapılabilir. Brown-Mood ve Theil metodlarına göre elde edilen regresyon doğru denkleminin gerçek verilere uyumunun diğer metodlardan daha iyi olmasının sebebi; Brown-Mood metodunda (X_i, Y_i) nokta değerlerinin regresyon doğrusuna dikey sapmalarının medyanı sıfır olana kadar regresyon doğrusunun konumunun değiştirilmesidir, Theil metodunda ise (X_i, Y_i), (X_j, Y_j) nokta çiftlerinin eğim değerlerinin medyanı β'nın buna bağlı olarak α değerinin bulunması ile oluşturulan regresyon doğrusunda (X_i, Y_i) nokta çiftlerinin sapmalarının küçük olmasıdır. k-Nearest Neighbor metodunda ise belirlenen herhangi bir X_i değerinden uzakta olan bazı gözlemlerin işleme tabi tutulmaması ile elde edilen regresyon doğrusundan gerçek verilerin sapmalarının küçük olmasından kaynaklanabilir.

Hata Kareler Ortalaması ise gözlenen değerlerden beklenen değerlerin sapmalarının karelerinin toplamının n-2 ye bölünmesiyle bulunur. Yani,

$$HKO = \frac{\sum\left(Y_i - \hat{Y}_i\right)^2}{n-2}$$

dir. Metodlar arasındaki mukayese hata kareler ortalamasına göre yapıldığında; hata kareler ortalaması en küçük olan modelin en uygun olduğu ifade edilir. Çünkü herbir metoda göre elde edilen modelde hata kareler ortalaması en küçük modelde gözlenen değerler ile beklenen değerler arasındaki sapma da o derece küçük olur. Böylece gözlenen değerler ile beklenen değerler arasındaki sapma küçüldükçe o metoda göre uydurulan modelin gerçek verilere uyumluluğu da daha iyi olur. Hata kareler ortalamasına göre modellerin gerçek verilere uyumluluk durumuna göre metodlar arasında

En Küçük Kareler > Theil > Brown-Mood > K-Nearest Neighbor

şeklinde sıralama yapılır.

5. GENEL DEĞERLENDİRME

Bu çalışmada, non-parametrik veriler arasındaki fonksiyonel bağıntıyı ifade eden nonparametrik regresyon metodlarından birkısmı araştırılmıştır. Araştırma, tahmin metodlarının genel özellikleri ve uygulaması olmak üzere iki kısımda ele alınmıştır. Tahmin metodlarının genel özellikleri kısmında, medyana ve smoothing tekniklerine göre nonparametrik regresyon parametrelerinin (α ve β'nın) tahmin metodları ve karşılaştırma yapmak amacıyla parametrik tahmin metodlarından en küçük kareler ve maksimum likelihood metodları araştırılmıştır.

Medyana göre regresyon parametrelerinin tahmin metodları genel olarak Theil metoduna dayandırılmıştır. Bu metodların bazılarının uygulaması ve regresyon doğrusunun eğimi olan β'nın önemlilik testi yapılmıştır. Uygulaması yapılan metodlarda β'nın 0.05 hata seviyesinde önemli olduğu tesbit edilmiştir. Smoothing tekniklerinden ise Kernel smoothing teknikleri genel olarak birbirine benzemekle beraber çoğunluğu Priestley ve Chao tahmin metoduna dayandırılmıştır. Smoothing tekniklerine göre non-parametrik regresyon parametrelerini tahminde k- nearest neighbor smoothing metodu diğer smoothing metodlarından daha kolay uygulanabilmektedir. Genel olarak, nonparametrik regresyon parametrelerini tahminde medyana göre tahmin metodlarının smoothing tekniklerinden daha kolay anlaşıldığı ve uygulamasının daha kolay olduğu gözlenmiştir.

Sonuçta, herbir metoda göre elde edilen regresyon doğru denklemlerinin gerçek verilere uyumluluk durumu R^2 testi ve hata kareler ortalamsına göre incelenmiş ve herbir metoda göre elde edilen doğru denklemlerinin gerçek verilere uyumluluk sırası aşağıdaki gibi bulunmuştur.

HKO'na göre

En Küçük Kareler > Theil > Brown-Mood > K-Nearest Neighbor

R^2'ye göre

Brown-Mood > Theil > k-Nearest Neighbor > En Küçük Kareler

HKO'na göre; en küçük kareler tahmin metoduna göre elde edilen regresyon doğru denkleminin nonparametrik tahmin metodlarına göre elde edilen denklemlerden ve medyan tahmin metodlarına göre elde edilen doğru denklemlerinin de smoothing tahmin metotlarına göre elde edilenden gerçek verilere daha iyi uyumlu olduğu gözlenmiştir.

R^2'ye göre; Medyan tahmin metodlarına göre elde edilen regresyon doğru denklemlerinin smothing tahmin metoduna göre elde edilenden, smothing tahmin metodu da en küçük kareler tahmin metoduna göre elde edilen regresyon doğrusu denkleminden gerçek verilere daha iyi uyumlu olduğu gözlenmiştir.

KAYNAKLAR

Adichie, J.N.,1967a, Asymtotic efficiency of a class of non-parametric tests for regression parameters. Annals of Mathematical Statistics, 38, 884-893.

Adichie, J.N.,1967b, Estimates of regression parameters based on rank test. Annals of Mathematical Statistics, 38, 894-904.

Aytaç, M.,1984, Uygulamalı Parametrik Olmayan İstatistik Testleri. Uludağ Üniversitesi, İktisadi ve İdari Bilimler Fakültesi, Bursa, s 44-47.

Benedetti, J.K., 1977, On the nonparametric estimation of regression functions. Journal of the Royal Statistical Society, Ser.B, 39, 248-253.

Chambers, R.L., Dorfman, A.H. and Wehrly, T.E., 1993, Bias robust estimation in finite populations using nonparametric calibration. Journal of the American Statistical Association, 88, 268-277.

Chu, C.K. and Marron, J.S., 1991, Choosing a kernel regression estimator. Statistical Science, 6 (4), 404-435.

Clark, R.M., 1977, Non-prametric estimation of a smooth regression function. Journal of the Royal Statistical Society, Series B, 39, 107-113.

Cover, T.M. and Hart, P.E., 1967, Nearest neighbor pattern classification. IEEE Transactions on Information Theory, 13, 21-27.

Cryer, J.D., Robertson, T., Wright, F.T. ve Casady, R.J., 1972, Monotone median regression. Annals of Mathematical Statistics, 43, 1459-1469.

Daniel, W.W., 1990, Applied Nonparametric Statistics. Georgia State University, Boston, (2nd.ed), p18-20, 426-443.

Davis, C., 1973, Statistics and data analysis in geology. New York, p 205-206.

Devroye, L.P., 1978, The uniform convergence of nearest neighbor regression function estimators and their application in optimization. IEEE Transactions on Information Theory, 24 (2), 142- 151.

Eddy, W.F., 1980, Optimum kernel estimators of the mode. The Annals of Statistics, 8 (4), 870-882.

Fraser, D.A.S., 1958, Statistics: An Introduction. New York, p 378-381

Gasser, T., Müller, H.G. and Mammıtzsch, V., 1985, Kernels for nonparametric curve estimation. Journal of the Royal Statistical Society, Ser. B, 47 (2), 238-252.

Green, P., Jennison, C. and Seheult, A., 1985, Analysis of field experiments by least squares smoothing. Journal of the Royal Statistical Society, Ser.B, 47 (2), 299-315.

Griffin, J.I., 1962, Statistics Methods an Aplication. Rinehart and Winston, p259-273.

Györfi, L., 1981, The rate of convergence of k_n-NN regression estimates and classification rules. IEEE Transactions of Information Theory, 27, 500-509.

Hall, P. and Marron, J.S.,1990, On variance estimation in nonparametric regression. Biometrika, 77 (2), 415-419.

Hall, P. and Wehrly, T.E., 1991, A geometrical method for removing edge effects from kernel-type nonparametric regression estimators. Journal of the American Statistical Association, 86 (415), 665-672.

Hardle, W. and Gasser, T,. 1984, Robust non-parametric function fitting. Journal of the Royal Statistical Society, Ser. B, 46 (1), 42-51.

Hardle, W. and Mammen, E., 1993, Comparing nonparametric versus parametric regression fits. The Annals of Statistics, 21 (4), 1926-1947.

Hardle, W., 1997, Applied nonparametric regression. Cambridge University, UK , (6nd ed.) p 14-85.

Heckman, N.E., 1986, Spline smoothing in a partly linear model. Journal of the Royal Statistical Society, Ser. B, 48 (2), 244-248.

Hill, B.M.,1962, A test of linearity versus convexity of a median regression curve. Annals of Mathematical Statistics, 33, 1096-1123.

Hogg, R.V. and Randles, R.H., 1975, Adaptive distribution-free regression methods and their applications. Technometrics, 17 (4), 399-407.

Hussain, S.S. and Sprent, P., 1983, Non-parametrik regression. Journal of the Royal Statistical Society, Ser.A, 146, 182-191.

Ibragimov, I.A. and Hasminskii,R.Z., 1980, On nomparametric estimation of regression. Soviet Math. Dokl., 21 (3), 810-814.

Jaeckel, L.A., 1972, Estimating regression coefficients by minimizing the dispersion of the residuals. Annals of Mathematical Statistics, 43 (5), 1449-1458.

Jureckova, J., 1971, Non parametric estimate of regression coefficients. Annals of Mathematical Statistics, 42 (4), 1328-1338.

Kildea, D.G.,1981, Brown- Mood type median estimators for simple regression models. The Annals of Statistics, 9 (2),438-442.

Maritz, J.S., 1979, On Theil's method in distribution-free regression. Australian Journal of Statistics, 21, 30-35.

Müller, H.G., 1992, Change-points in nonparametric regression analysis. The Annals of Statistics, 20 (2), 737-761.

Neter, J., Wasserman, W. and Kutner, M.H.,1989, Applied Linear Regression Models. Boston, (2nd ed), p53.

Parzen, E., 1962, On estimation of a probability density function and mode. Annals of Mathemetical Statistics, 33, 1065-1076.

Potthoff, R.F.,1974, A non-parametric test of whether simple regression lines are parallel. The Annals of Statistics, 2 (2), 295-310.

Priestley, M.B. and Chao, M.T., 1972, Non-parametric function fitting. Journal of the Royal Statistical Society, Ser.B, 34, 385-392.

Rao, K.S.M. and Gore, A.P., 1982, Nonparametric tests for intercept in linear regression problems. Australian Journal of Statistics, 24 (1), 42-50.

Reinsch, C.H., 1967, Smoothing by spline functions. Num. Mathematik,10,177-183.

Rice, J., 1984, Bandwith choice for nonparametric regression. The Annals of Statistics, 12 (4), 1215-1230.

Rosenblatt, M., 1971, Curve estimates. Annals of Mathematical Statistics, 42 (6), 1815-1842.

Ross, S.M., 1987, Introduction to probability and statistics for engineers and scientists. University of California, p 245-248.

Sen, P.K., 1968, Estimates of the regression coefficient based on Kendall's tau. Journal of the American Statistical Association, 63,1379-1389.

Sievers, G.L.,1978, Weighted rank statistics for simple linear regression. Journal of the American Statistical Association, 73, 628-631.

Silverman, B.W., 1984, Spline smoothing: the equivalent variable kernel method. The Annals of Statistics, 12 (3), 898-916.

Simonoff, J.S., 1998, Three sides of smoothing: categorical data smoothing, nonparametric regression, and density estimation. International Statistical Review, 66 (2),137-156.

Speckman, P.,1985, Spline smoothing and optimal rates of convergence in nonparametric regression models. The Annals of Statistics, 13 (3), 970-983.

Stone, C.J., 1977, Consistent nonparametric regression. The Annals of Statistics, 5 (4), 595-645.

Stute, W., 1984, Asymptotic normality of nearest neighbor regression function estimates. The Annals of Statistics, 12 (3), 917-926.

Yıldız, N ve Bircan, H., 1994, Uygulamalı İstatistik. Atatürk Üniversitesi Yayınları No:704, Ziraat Fakültesi Yayınları No:308, (dördüncü baskı), s167, Erzurum.

Watson, G.S. and Leadbetter, M.R., 1962, On the estimation of the probability density I. Annals of Mathematical Statistics, 34, 480-491.

Wu, J.S. and Chu, C.K., 1993, Kernel-type estimators of jump points and values of a regression function. The Annals of Statistics, 21 (3), 1545-1566.