

Kantil Regresyon Analizi Üzerine Bir Uygulama

Bahar ARSAN AYSA¹, Hamit MİR TAGİOĞLU², Sıddık KESKİN³, Yıldırım DEMİR^{4*}

¹Bağımsız Araştırmacı, Bitlis

²Bitlis Eren Üniversitesi, Fen-Edebiyat Fakültesi İstatistik Bölümü, Bitlis

³Van Yüzüncü Yıl Üniversitesi, Tıp Fakültesi Temel Tıp Bilimleri Bölümü, Van

⁴Van Yüzüncü Yıl Üniversitesi, İktisadi ve İdari Bilimler Fakültesi Ekonometri Bölümü, Van

¹<https://orcid.org/0000-0002-3304-2616>

²<https://orcid.org/0000-0003-2952-9584>

³<https://orcid.org/0000-0001-9355-6558>

⁴<https://orcid.org/0000-0002-6350-8122>

*Sorumlu yazar: ydemir@yyu.edu.tr

Araştırma Makalesi

Makale Tarihiçesi:

Geliş tarihi: 26.02.2024

Kabul tarihi: 30.05.2024

Online Yayınlanma: 10.06.2024

Anahtar Kelimeler:

Bootstrap yöntemi

Doğrusal regresyon

En küçük kareler

Kantil regresyon

Şarap kalitesi

ÖZ

Genel olarak bir araştırma yapılmak istendiğinde, değişkenler arasındaki ilişki incelenmek ya da bir model yardımıyla tahmin yapılmak istenebilir. Bu model yardımıyla tahmin yapılmak istendiğinde, yaygın kullanılan yöntem Standart regresyon analizi yöntemidir. Ancak, bu regresyon analizinin kullanılabilmesi için bazı varsayımların yerine gelmesi gerekmektedir. Bu varsayımlar yerine gelmediği veya bazı araştırmalarda transformasyonlar yapılmasına rağmen bu varsayımlar sağlanmadığı durumlarda alternatif yöntemler tercih edilmektedir. Bu alternatif yöntemlerden birisi de Kantil regresyon yöntemidir. Kantil regresyon, heterojen yapıdaki veri setleri için oldukça kullanışlıdır. Herhangi bir varsayım gerektirmediği için esnek bir yöntemdir. Çalışmada Wine Quality veri setinin 300 birimlik kısmı ve STATA 14 paket programı kullanılmıştır. Önce şarap kalitesi veri setinden 1 bağımlı (kalite) ve 5 bağımsız değişken alınarak 100, 200 ve 300 örneklem genişliklerinde doğrusal ve 0,20, 0,25, 0,50 ve 0,75 kantil değerlerine göre Kantil regresyon analizleri gerçekleştirilmiştir. Daha sonra aynı işlemler 1 bağımlı (kalite) ve 11 bağımsız değişken için tekrarlanmıştır. Kısmi regresyon katsayısı, standart hata ve güven aralıkları temel alınarak doğrusal ve Kantil regresyon analiz sonuçları; örneklem genişliğine ve değişken sayısına göre karşılaştırılmıştır. Sonuç olarak; modele dahil edilen değişken sayısının artırılmasının, katsayılar belirgin etkisinin olmadığı gözlenmiştir. Ayrıca kantil değerlerin sonuçlar üzerindeki etkisinin, değişken sayısından daha önemli olduğu gözlenmiştir. Bu nedenle de düşük katsayı ve dar güven aralığı elde etmek için uygun kantil değerinin tercih edilmesi önemlidir.

An Application on Quantile Regression Analysis

Research Article

Article History:

Received: 26.02.2024

Accepted: 30.05.2024

Available online: 10.06.2024

Keywords:

Bootstrap method

Least squares

Linear regression

ABSTRACT

In general, when a research is desired to be conducted, the relationship between variables may be examined or prediction may be made with the help of a model. When it is desired to make predictions with the help of this model, the commonly used method is the standard regression analysis method. However, some assumptions must be fulfilled in order to use this regression analysis. Alternative methods are preferred when these assumptions are not fulfilled or when these assumptions are not met although transformations are made in some studies. One of these alternative methods is the quantile regression

Quantile regression
Wine quality

method. Quantile regression is very useful for heterogeneous data sets. It is a flexible method since it does not have any assumptions. In the study, 300 units of the Wine Quality data set and STATA 14 package program were used. Firstly, by taking 1 dependent (quality) and 5 independent variables from the wine quality data set, quantile regression analyses were performed according to 0,20, 0,25, 0,50 and 0,75 quantile values as well as standard linear regression at sample sizes of 100, 200 and 300. Then, the same procedures were repeated for 1 dependent (quality) and 11 independent variables. The results of linear and quantile regression analyses based on partial regression coefficient, standard error and confidence intervals were compared according to sample size and number of variables. As a result, it was observed that increasing the number of variables included in the model did not have much effect on the coefficients. In addition, it was determined that the effect of quantile values on the results is more important than the number of variables. Therefore, it is important to choose the appropriate quantile value to obtain low coefficient and narrow confidence interval.

To Cite: Arsan Aysal B, Mirtağođlu H, Keskin S, Demir Y., 2024. Kantil regresyon analizi üzerine bir uygulama. Kadirli Uygulamalı Bilimler Fakóltesi Dergisi, 4(2): 454-478.

Giriş

Doğadaki deđişkenler arasında var olan ilişkinin doğru belirlenebilmesi, ilgili deđişkene istenilen şekilde yön verilebilmesi açısından oldukça önemlidir. Deđişkenler arası ilişkileri belirlemek üzere farklı istatistik yaklaşımlar geliştirilmiş ve en yaygın olanı regresyon analizidir. Araştıracının cevap (sonuç) deđişkeni veya bağımlı deđişken olarak aldığı deđişkenle, açıklayıcı (bağımsız) deđişken veya deđişkenler olarak aldığı özellikler arası ilişkileri belirlemek üzere kurulan doğrusal veya doğrusal olmayan modeller genel olarak Regresyon modelleri olarak adlandırılır. Regresyon modelleri ve bu modellerle ilişkili istatistik analizler, minimum sayıdaki açıklayıcı deđişkenle, bağımlı deđişkeni en yüksek doğrulukla tahmin etmek üzere gerekli işlemleri içerir.

Deđişkenler arasındaki ilişkiler, genel olarak doğrusal ve doğrusal olmayan ilişkiler olmak üzere iki kısımda incelenmektedir. Analiz süreci ve elde edilen sonuçların yorumlanmasındaki kolaylık bakımından doğrusal modeller, doğrusal olmayan modellere göre daha çok tercih edilmektedir. Regresyon modellerine dayalı olan regresyon çözümlenmeleri için bazı varsayımlar gereklidir. Bu varsayımların sağlanmaması durumunda, ya deđişkenlere uygun transformasyonlar uygulanarak varsayımların sağlanmasına çalışılmakta ya da bu varsayımlara karşı esnek olan parametrik olmayan yöntemler tercih edilmektedir.

Çoklu doğrusal regresyon analizinde parametre tahmini, çoğunlukla En Küçük Kareler (EKK) yöntemine göre yapılmaktadır. Diđer bir ifade ile Çoklu doğrusal regresyon, gözlemlerin her bir deđeri için açıklayıcı deđişkenlere dayalı cevap deđişkeninin (koşullu) ortalamasını hesaplamak için EKK yöntemini kullanmaktadır. EKK yöntemi hesaplama işlemlerinin, diđer yöntemlere göre kısmen daha kolay olması bu yöntemin tercih

edilebilirliğini artırmaktadır. Ancak bu yöntem için bazı varsayımların sağlanması gerekmektedir. Bu varsayımların sağlanmaması durumlarında alternatif yöntemlere başvurulabilir. Bu yöntemlerden birisi de Koenker ve Bassett (1978) tarafından önerilen Kantil regresyon modelidir.

Kantil regresyon, bağımsız değişken/ler ile bağımlı değişkenin, araştırmacı tarafından belirlenmiş olan kantilleri arasındaki ilişkiyi modellemektedir. Böylece Kantil regresyon, bağımlı değişkenin koşullu medyanını veya diğer kantillerini tahmin eden ve standart doğrusal regresyonun varsayımları sağlanmadığında kullanılabilen çoklu doğrusal regresyon analizinin bir uzantısı olarak düşünülebilir (Kuşkaya, 2023). En Küçük Kareler regresyonuna göre iki önemli avantajı bulunmaktadır. Birincisi, cevap değişkeni ile ilgili herhangi bir varsayım gerektirmemesi, diğeri ise uç değerlere karşı sağlam (robust) olmasıdır. Bu nedenle başta ekoloji bilimi olmak üzere, sağlık bilimleri, ekonometri ve ziraat gibi birçok alanda kullanılabilir. Kantil regresyon, bağımsız değişken/ler ile bağımlı değişkenin, araştırmacı tarafından belirlenmiş olan kantilleri arasındaki ilişkiyi modellemektedir. Böylece Kantil regresyon, bağımlı değişkenin koşullu medyanını veya diğer kantillerini tahmin eden ve standart doğrusal regresyonun varsayımları sağlanmadığında kullanılabilen çoklu doğrusal regresyon analizinin bir uzantısı olarak düşünülebilir (Kuşkaya, 2023). En Küçük Kareler regresyonuna göre iki önemli avantajı bulunmaktadır. Birincisi, cevap değişkeni ile ilgili herhangi bir varsayım gerektirmemesi, diğeri ise uç değerlere karşı sağlam (robust) olmasıdır. Bu nedenle başta ekoloji bilimi olmak üzere, sağlık bilimleri, ekonometri ve ziraat gibi birçok alanda kullanılabilir.

Literatür incelemesinde, Kantil regresyon ile ilgili yeterli sayıda çalışmanın olduğu, ancak farklı kantil oranları, değişken sayısı ve gözlem sayısı bakımından çalışmaların sınırlı olduğu görülmüştür. Bu nedenle bu çalışmada, Kantil regresyon yöntemine ilişkin genel bilgi verilmiş ve 3 farklı örneklem genişliği, farklı değişken sayıları ile 4 farklı kantil oran kombinasyonları kullanılarak Kantil ve standart çoklu regresyon analizi sonuçları değerlendirilmiştir.

Kantil regresyon modeli ilk olarak Koenker ve Bassett (1978) tarafından önerilmiştir. Kantil regresyon modellerinin doğrusal regresyon modellerine bağlı bir yan dal olabileceği ifade edilmektedir (Koenker ve Bassett, 1978). Doğrusal olmayan tüm Kantil regresyon modellerini kapsayan yarı Bayesçi modeller Taddy ve Kottas (2010), tarafından incelenmiştir. Bu modeller, tüm değişkenlerin birleşik dağılımları üzerinde etkili olabilecek esnek modellere ulaşma imkânı tanımıştır. Bu modellerde, esnek modellere ulaşabilmek adına, bağımlı değişkenin koşullu dağılımlarından faydalanılmıştır. Bunun için farklı kantil değerlerindeki eğriler yorumlanmıştır. Doğrusal ve Kantil regresyon modelleri Kuan ve Lin (2010) tarafından, detaylı bir şekilde incelenmiş ve gerçekleştirilen uygulamalar sonucunda modelin genel olarak ki-kare dağılımına eğilim gösterdiği vurgulanmıştır.

Serbest ekonomiden yana olan bankaların işleyişini konu alan çalışmada; klasik ekonometrik modeller kullanılarak elde edilen çıkarımların tüm resmi görmeye imkân tanımayacağı Schaeck (2008) tarafından vurgulanmıştır. Kantil değerler için sonuçlar incelendiğinde, eş benzer faktörlerin bağımlı değişkeni aynı şekilde etkilemesinin mümkün olmadığı gözlenmiştir. Ayrıca değişkenler arasındaki etkilerin bağımlı değişkenin kuyruğunu etkilemesi durumunda, etkin bir tahminleme yöntemi olarak Kantil regresyona başvurulması

gerektiđi belirtilmiřtir. Kiren Grler ve ark. (2018), Hanehalkı Bte Anketi veri seti ile yaptıkları alıřmada Kantil regresyon ve EKK ynteminden elde edilen sonuları karřılařtırmıřlar. Sonu olarak, byk veri gruplarıyla alıřılırken, deđiřen varyans sorununun ortaya ıktıđı ve bu durumda katsayıların etkin tahmin edilememesi ve aykırı deđerlerin varlıđı gibi sorunlarla karřılařılabileceđi vurgulanmıřtır. Ayrıca, tm bu sorunlara rađmen Kantil regresyon ynteminin etkin bir tahminci olma zelliđini koruduđuna dikkat ekilmiřtir. ınar (2019), simlasyon yntemiyle u deđerler barındıran normal dađılmalı bir veri seti retmiřtir. Bu veri setini kullanarak Kantil regresyon model etkinliđini EKK ile karřılařtırmıřtır. Sonu olarak, Kantil regresyon model etkinliđinin EKK'dan ok daha iyi olduđu gzlenmiřtir. zelik ve ark. (2021) tarafından, Kantil regresyon ve dođrusal olmayan en kk kareler yntemiyle dođal karaam meřcereleri iin gvde apı modeli geliřtirilmiřtir. Hem tm gvde hem de farklı nisbi boy sınıfları iin Kantil regresyonla elde edilen gvde apı tahmin sonularının dođrusal olmayan en kk kareler yntemiyle elde edilen sonulara gre daha bařarılı olduđu grlmřtr. Sonu olarak, gvde apı tahminlerinde Kantil regresyon ynteminin alternatif bir yntem olarak kullanılabilmesi belirtilmiřtir. Demir (2022) tarafından, rn ve hizmet ihracatı ile ham tarım rn ihracatının ekonomik byme zerindeki etkileri; verilere %25, %50 ve %100 bootstrap uygulanarak Kantil ve dođrusal regresyon yntemleriyle belirlenmeye alıřılmıřtır. Sonu olarak, normallik varsayımının sađlanmadıđı durumlarda, Kantil regresyonun ve sađlandıđı durumlarda ise dođrusal (EKK yaklařımı) regresyonun daha iyi tahminler verdiđi ifade edilmiřtir.

Materyal ve Metot

alıřmada, <https://www.kaggle.com/datasets?search=wine+quality> internet sitesinden alınan Wine Quality veri seti kullanılmıřtır. 2019 yılında veri setinin 300 birimlik kısmı alınarak yeni bir veri seti oluřturulmuřtur. Analizlerde STATA 14 paket programı kullanılmıřtır.

İlk olarak, řarap kalitesi veri setinden 1 bađımlı (kalite) ve 5 bađımsız deđerken alınarak 100, 200 ve 300 rneklem geniřliklerinde dođrusal ve 0,20, 0,25, 0,50 ve 0,75 kantil deđerlerine gre Kantil regresyon analizleri gerekleřtirilmiřtir. Farklı rneklem byklkleri ve farklı kantil deđerlerinin yanı sıra; farklı deđerken sayısının da sonulara etkisini gzlemlemek iin daha sonra aynı iřlemler; 1 bađımlı (kalite) ve 11 bađımsız deđerken iin tekrarlanmıřtır. Kısmi regresyon katsayısı, standart hata ve gven aralıkları temel alınarak dođrusal ve Kantil regresyon analizi sonuları, rneklem geniřliđine ve deđerken sayısına gre karřılařtırılmıřtır.

Regresyon modelleri ile parametre tahminlerinde EKK yaklařımı en sık kullanılan modeldir. EKK, hata kareler toplamını minimize eden α ve β parametrelerinin tahmini

değerlerinin bulunmasını esas almaktadır. Bu yaklaşım ile elde edilecek model minimum hata payına sahip olmalı ve gerçek değerlere en yakın değerler elde eden en uygun eğriyi ortaya çıkarmalıdır. EKK ile etkin sonuçların elde edilebilmesi için gerekli varsayımlar sağlanmalı ve bu varsayımlar sağlanmadığında EKK yanlı sonuçlar verebildiğinden, alternatif bir yöntem olarak En Küçük Medyan Kareler yöntemi önerilmektedir (Erickson ve ark., 2006).

En Küçük Medyan Kareler modeli, veri setinin %50'ye kadar sapan değer içermesi halinde, EKK yaklaşımından daha etkin sonuçlar vermekte ve bu durumda oldukça güçlü bir regresyon modeli olarak kullanılabilir (Rousseeuw ve Leroy, 1987). Medyan Regresyon modelinin geliştirilmiş bir hali ise Kantil regresyon modelidir. Medyan Regresyon modelinde, veri setinin sadece ortanca değerine göre sonuçlar hesaplanırken, Kantil regresyonda farklı kantil değerlerine göre veri seti incelenerek daha kapsamlı ve daha etkin sonuçlar hesaplanmaktadır.

Doğrusal Regresyon

Regresyon analizine yönelik ilk çalışmalar Sör Francis Galton tarafından yapılmıştır. Galton, değişkenler arası ilişkileri inceleyen ve yorumlayan ilk bilim adamlarındandır. 1885 yılında kalıtım konusunda çalışmalar yaparken, babalar ve oğulların boyları konusunda yaptığı araştırmalarda, değişkenler arasında ortalamaya doğru yönelimin olduğunu göstererek regresyon kavramının temelini atmıştır. Karl Pearson ve Lee (1903) ise bu çalışmayı daha da ileriye taşıyarak, çocuklar arasında babası uzun boylu olanların babalarının aksine kısa boylu olabileceğini, uzun boylu çocukların babalarının ise kısa olabileceğini vurgulamıştır. Bunun sonucunda, uzun ve kısa çocukların boylarının benzer şekilde ortalamaya gerilediğini (regressing) ifade etmiştir (Gujarati, 2004).

İstatistik tahmin çalışmaları yapılmak istendiğinde, ilk akla gelen Regresyon modelidir. Regresyon modeli ile iki ya da daha fazla değişken arasındaki ilişkinin gösterilmesi mümkün olabilmektedir (Chatterjee ve Simonoff, 2013). İki değişken arasındaki ilişki neden sonuç ilişkisi olarak da tanımlanabilir. Sonuç değişkeni bağımlı değişken (Y) olarak, neden değişkeni ise bağımsız değişken (X) olarak ifade edilmektedir. Neden sonuç değişkenlerinin doğru belirlenmesi regresyon analizinin doğru sonuç vermesi bakımından oldukça önemlidir (Orhunbilge, 2000). X ve Y arasındaki ilişki için çizilen grafiğe serpilme diyagramı denilmektedir. Doğrusal regresyon, bir bağımlı ve bir veya daha fazla bağımsız değişken arasındaki doğrusal ilişkiyi açıklamakta ve bağımsız değişkenlerin bağımlı değişkeni nasıl etkilediğini belirlemek için kullanılmaktadır (Draper ve Smith, 1998). Doğrusal regresyon basit

ve çoklu regresyon olmak üzere ikiye ayrılmaktadır. En basit haliyle doğrusal regresyon modeli;

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_j X_{ij} + \varepsilon_i \quad i = 1, 2, \dots, n \quad j = 0, 1, \dots, k \quad (1)$$

olarak ifade edilmektedir.

Eşitlik (1)'de Y_i : bağımlı değişkeni, X_{ij} : bağımsız değişkenleri, β_0 : sabit terimi, $\beta_1, \beta_2, \dots, \beta_j$: bağımsız değişken katsayılarını ve ε_i : rassal değişkenler olup sıfır etrafında normal dağılan hata terimini göstermektedir. Modeldeki bağımsız değişkenlerin etkisi β terimleri ile belirlenmektedir (Demir, 2022). Eşitlik (1) çoklu doğrusal regresyon modeli olup, j 'nin 1 olması basit regresyon modelini ifade etmektedir. Doğrusal regresyon, hata kareler ortalamasını en iyi minimize eden β katsayılarının tahminçileri olan $\hat{\beta}$ değerlerini belirlemeyi amaçlamaktadır. EKK varsayımları sağlandığında $\hat{\beta}$ değerlerinin belirlenmesinde EKK yöntemi kullanılmaktadır (Demir, 2020). Gerekli varsayımlardan herhangi birinin sağlanmaması durumunda ise parametre tahmini için farklı bir model kullanmanın daha doğru sonuçlar vereceği savunulmaktadır (Berry, 1993). Bu varsayımlar;

- 1) Bağımsız değişkenlerin nicel veya nitel olması, bağımlı değişkenin ise nicel ve sürekli olması,
- 2) Bağımsız değişkenlerin varyansının sıfırdan farklı olması,
- 3) Bağımsız değişkenler arasında yüksek doğrusal ilişkinin olmaması,
- 4) Hata terimi ortalamasının sıfır olması,
- 5) Bağımsız değişkenler ve hata terimi arasında korelasyon olmaması,
- 6) Hata terimi varyansının sabit olması,
- 7) Hataların normal dağılması şeklinde özetlenebilir (Rawlings ve ark., 1998; Demir, 2020).

EKK yöntemiyle $\hat{\beta}$ parametre tahminleri;

$$\hat{\beta} = (X'X)^{-1} X'Y \quad (2)$$

olarak hesaplanmaktadır. Bu yöntemde amaç, ε_i 'nin tahminçisi olan e_i 'yi minimum yapmaktır.

$$e_i = Y_i - \hat{Y}_i \quad (3)$$

Eşitlik (3)'te Y_i gözlenen değerleri, \hat{Y}_i ise Y_i değerlerinin tahmini değerlerini ifade etmektedir. Burada hataların toplamı sıfır ve hata kareler toplamı ise minimum olur. EKK yönteminde amaç gerçek değerlerle tahmini değerler arasındaki farkların kareleri toplamını

minimum yapan $\hat{\beta}$ deęerlerini belirlemektir. Bylece, Eşitlik (2) yardımıyla baęımlı deęişkenin (Y_i) tahmini deęerleri (\hat{Y}_i) Eşitlik (4) ile hesaplanmaktadır.

$$\hat{Y}_i = X_i\hat{\beta}_i + e_i \quad (4)$$

$\hat{\beta}$ katsayısının, sıfırdan nemli derecede farklı olması Y deęişkeni ile X deęişkenleri arasında nemli (anlamlı) bir ilişkinin varlığını gsterirken, katsayı işaretleri ise bunlar arasındaki ilişkinin ynn gstermektedir. EKK tahmin edicilerinin doęrusal ve yansız olması durumunda, en dşk varyans ile gereęe en yakın tahminler elde edilmektedir. Bu nedenle EKK, analizlerde sıka tercih edilmektedir. Gerekli varsayımlar saęlandığında EKK, deęişkenler arasındaki ilişkiyi aıklamada etkin sonular verirken, varsayımların saęlanmaması durumunda EKK'nın etkinlięi azalmaktadır (Gujarati, 2004). Bu gibi durumlarda Kantil regresyon gibi alternatif yntemlerle daha etkin sonular elde edilebilir.

Kantil Regresyon

Kantil regresyon, doęrusal regresyon yntemiyle aynı grevi grmektedir. EKK iin gerekli varsayımlar saęlanmadığı durumlarda ya veri setine bazı dnşmler uygulanmakta ya da Kantil regresyon gibi bazı alternatif yntemlerle veriler arasındaki ilişkiler belirlenmek istenmektedir (Demir, 2022). Belirlenen kantiller iin medyan regresyon ynteminin geliřtirilmiř bir durumu olan Kantil regresyon, veri setlerinde u deęerler bulunduęunda EKK'dan daha gl tahminler vermektedir (Alakaya, 2019). Bu yntem, daęılım iin herhangi bir varsayımda bulunmadığından doęrusal regresyon varsayımları saęlanmadığı durumda, Kantil regresyonla ok daha gl sonular saęlanabilir. Kantil regresyon, tm kořullu daęılımları modelleyebileceęi gibi parametrik bir daęılım gerektirmeden tm kantiller iin doęrusal regresyon ile belirlenmesi zor olan bilgiler verebilmektedir. Doęrusal regresyon modeli tek bir kořullu ortalamaya sahip ve bu nedenle yalnızca bir eşitlikle gsterilirken, Kantil regresyon modeli ise birden fazla kořullu ortalamaya sahip olduęundan model birden fazla eşitlikle gsterilmektedir (Demir, 2022).

Kantil regresyon modeli;

$$Y_i = \beta + \varepsilon_i \quad (5)$$

olarak gsterilmektedir. Y_i ; simetrik F daęılım fonksiyonlu, baęımsız zdeř daęımlı ve β medyanlı rastgele deęişkeni gstermektedir. x_θ deęeri, populasyonun θ 'inci kantilini gsterirse, modelde θ 'inci rnek kantili Eşitlik (6)'nın β 'ya gre minimizasyonu ile bulunmaktadır.

$$\min_{\beta} \frac{1}{n} \left\{ \sum_{i:y_i \geq \beta} \theta |y_i - \beta| + \sum_{i:y_i < \beta} (1 - \theta) |y_i - \beta| \right\} \quad (6)$$

Bu ifade doğrusal regresyon modeline genelleştirilirse Eşitlik (5);

$$Y_i = x_i \beta_{\theta} + e_i \quad (7)$$

olarak yeniden yazılabilir. Burada bağımsız değişkenler vektörü olan x_i ; bağımsız değişkenlerle bağımlı değişkene ait koşullu dağılımın θ 'inci kantili arasındaki doğrusal regresyonu, β_{θ} ; θ 'inci Kantil regresyona ait parametre vektörünü ve e_i ise bağımsız ve sıfır etrafında simetrik F dağılımına sahip hatayı göstermektedir. Böylece $(0 < \theta < 1)$ aralığı için θ 'inci Kantil regresyon;

$$\min_{\beta} \frac{1}{n} \left\{ \sum_{i:y_i \geq x_i \beta} \theta |y_i - x_i \beta| + \sum_{i:y_i < x_i \beta} (1 - \theta) |y_i - x_i \beta| \right\} \quad (8)$$

eşitliğinin β 'ya göre minimize edilmesiyle tahmin edilmekte ve y 'nin θ 'inci kantili olarak tanımlanmaktadır. Parametre tahmini için Eşitlik (9)'dan faydalanılmakta ve θ değerleri için $\hat{\beta}$ tahmincileri Eşitlik (10) ile hesaplanmaktadır.

$$\min_{\beta} \frac{1}{n} \left\{ \sum_{i=1}^n \rho_{\theta}(y_i - x_i \beta) \right\} \quad (9)$$

$$\hat{\beta}(\theta) = \arg \min_{\beta \in \mathbb{R}^p} \left\{ \sum_{i=1}^n \rho_{\theta}(y_i - x_i \beta) \right\} \quad (10)$$

Bağımlı değişkenin koşullu dağılımının farklı noktalarındaki bağımsız değişkenlerde oluşan değişikliklere farklı tepkiler vermesine bağlı olarak Kantil regresyonun her bir kantili de birbirinden farklı sonuçlar vermektedir. Doğrusal regresyonda en uygun regresyon eğrisi, gözlemlerin ağırlık merkezinden veya gözlem noktalarının ortasından geçerken, Kantil regresyonda kantillerden geçmekte ve asimetric olarak belirlenmektedir. Yani kantil değeri 0,2 olarak alınır, verilerin %20'si ilk regresyon eğrisinin altında %80 ise üstünde kalacaktır. Bu durumda veri setinin birikimli dağılımı daha detaylı görüntülenebilir. Böylece ortalamanın değil de kantillerin kullanılmasının önemi;

- Normal dağılım göstermeyen veri setleri ile yapılan analizlerde daha etkili olabilmesi,
- Güvenilir tahminler yapabilmesi,
- Temsili bir değerle veya aykırı değerlerle ilgilenildiğinde EKK'ya göre daha güçlü bir yöntem olarak kullanılabilmesi,
- Değişken etkilerinin yerleri ve ölçek parametreleri açısından incelenebilmesi,

Hata terimleri yarı parametrik varsayımlara sahip olduğunda değişen varyanslı örneklerde daha etkin sonuçlar verebilmesi, şeklinde sıralanabilir (Keskin, 2012; Özyıldırım, 2019).

Bootstrap Yöntemi

Efron (1979) tarafından geliştirilen ve Yeniden örnekleme yöntemlerinden birisi olan Bootstrap yöntemi, uygulamada yaygın bir şekilde kullanılmaktadır. Bootstrap yönteminde mevcut örneklem popülasyon olarak kabul edilmekte ve bu örneklemden belirli sayıda tekrarlı örneklemeler alınarak ilgili tahminci için yeni bir örnekleme dağılımı oluşturmaktır (Çamurlu ve Erilli, 2019).

Bootstrap yöntemi, veri seti dağılımı için herhangi bir varsayıma gerek duymaz ve varsayımlarda sapmalar görülse de, Bootstrap ile yapılan tahminler bu sapmalardan etkilenmez (Okutan, 2009). Dolayısıyla, yöntem hem parametrik hem de parametrik olmayan yöntemlerle birlikte kullanılmaktadır. Bu yöntemde hata terimleri küçültülerek daha düşük standart sapmalar sağlanmakta ve bunun sonucunda da daha sağlam parametre tahminleri yapılabilmektedir (Efron ve Tibshirani, 1993).

Bu yöntem önce; örnek ortalaması, standart sapması ve güven aralığının hesaplanması için geliştirilmiştir. Ayrıca, EKK için gerekli varsayımlar yerine getirilemediği durumlarda bir düzeltme işlemi olarak kullanılabileceği gibi model seçim kriteri olarak tahmin hatasının hesaplanması, yanlılık tahmini, regresyon analizi, güven aralığının oluşturulması ve hipotez testleri gibi farklı alanlarda da kullanılabilmektedir (Gayaker, 2015).

Bulgular ve Tartışma

Tablo 1’de 11 bağımsız değişen ve bağımlı değişken olan kalite değişkeni için bazı tanımlayıcı istatistikler verilmiştir.

Tablo 1. Değişkenler için tanımlayıcı istatistikler

Değişken (N=300)	Ortalama	Standart Sapma	Standart Hata	Ranj	Minimum	Maximum
Sabit Asitlik	7,995	1,508	0,087	10,40	4,60	15,00
Uçucu Asitlik	0,553	0,181	0,010	1,15	0,18	1,33
Sitrik Asit	0,235	0,193	0,011	1,00	0,00	1,00
Kalan Şeker	2,438	1,163	0,067	9,50	1,20	10,70
Klorit	0,099	0,072	0,004	0,57	0,04	0,61
Serbest Kükürt	15,130	9,087	0,525	49,00	3,00	52,00
Toplam Kükürt	56,383	35,675	2,060	145,00	8,00	153,00
Yoğunluk	0,999	0,002	0,0001	0,01	0,99	1,00
pH	3,327	0,156	0,009	1,16	2,74	3,90
Sülfat	0,690	0,243	0,014	1,67	0,33	2,00
Alkol	9,878	0,814	0,047	5,00	9,00	14,00
Kalite	5,407	0,723	0,042	4,00	4,00	8,00

Tablo 1'e göre; en büyük (145) değişim aralığına (ranj) Toplam Kükürt değişkeni ve en küçüğüne (0,01) ise Yoğunluk değişkeni sahiptir. Zira Toplam Kükürt değişkeninin maksimum değeri 153, minimum değeri 8 ve Yoğunluk değişkeninin ise sırasıyla 1,00 ve 0,99 olarak belirlenmiştir. En yüksek ortalama (56,383), standart sapma (35,675) ve standart hataya (2,060) Toplam Kükürt değişkeninin sahip olduğu; en düşük ortalama (0,099) Klorit, standart sapma (0,002) ve standart hataya (0,0001) ise Yoğunluk değişkeninin sahip olduğu görülmektedir. Bağımsız değişkenler arasındaki ilişki için Tablo 2'de bu değişkenler arasındaki korelasyonlar verilmiştir.

Tablo 2. Bağımsız değişkenler arasındaki korelasyon katsayıları

	Sabit Asitlik	Uçucu Asitlik	Sitrik Asit	Kalan Şeker	Klorit	Serbest Kükürt	Toplam Kükürt	Yoğunluk	pH	Sülfat
Uçucu Asitlik	-0,250									
Sitrik Asit	0,498	-0,608								
Kalan Şeker	0,224	-0,072	0,157							
Klorit	0,208	0,094	0,110	0,027						
Serbest Kükürt	0,012	-0,039	0,101	0,076	0,192					
Toplam Kükürt	0,004	0,035	0,210	0,084	0,224	0,796				
Yoğunluk	0,343	-0,082	0,224	0,171	0,237	-0,101	-0,110			
pH	-0,659	0,198	-0,487	0,012	-0,354	-0,110	-0,205	-0,333		
Sülfat	0,342	-0,304	0,372	0,087	0,239	0,087	0,081	0,100	-0,351	
Alkol	0,042	-0,118	0,014	0,257	-0,136	-0,012	-0,075	-0,290	0,219	0,114

Tablo 2 incelendiğinde, en yüksek korelasyonun Toplam Kükürt ile Serbest Kükürt değişkenleri arasında (0,796) olduğu görülmektedir. Bunu sırasıyla pH ile Sabit Asitlik değişkenleri arasındaki korelasyon (-0,659) ve Sitrik Asit ile Uçucu Asitlik değişkenleri arasındaki korelasyon (-0,608) izlemektedir. Böylece Toplam Kükürt artarken buna bağlı olarak Serbest Kükürt değerinin de arttığı, pH artarken Sabit Asitlik değerinin ve Sitrik Asit artarken de Uçucu Asitlik değerinin azaldığı söylenebilir.

EKK varsayımlarının sağlanıp sağlanmadığına dair bazı test sonuçları Tablo 3'te verilmiştir.

Bağımsız değişkenler arasındaki korelasyon katsayılarının ($r \geq 0,75$) büyük olması durumunda çoklu doğrusal bağlantıdan şüphelenip diğer kriterlere bakılmalıdır (Demir, 2020). Tablo 2'de verilen Toplam Kükürt ile Serbest Kükürt değişkenleri arasındaki korelasyon katsayısı 0,796 olduğundan dolayı çoklu doğrusal bağlantı sorunu için VIF (Varyans şişirme faktörü) değerlerine bakılmıştır.

Regresyon analizinde, VIF çoklu doğrusallığın derecesini belirlemek için istatistik bir ölçüt olarak kullanılmaktadır. Çoklu doğrusallık, iki ya da daha fazla bağımsız değişkenin

birbiriyle ilişkili olması durumu olarak tanımlanmaktadır. VIF değerlerinin 10'dan büyük olması durumunda çoklu doğrusal bağlantı sorunundan söz edilebilir (Özen ve ark., 2023). Tablo 3 incelendiğinde, bağımsız değişkenlere ait VIF değerlerinden, en büyüğünün 6,286 olduğu görülmektedir. Dolayısıyla değişkenler arasında çoklu doğrusal bağlantı sorunu olmadığı söylenebilir.

Tablo 3. EKK varsayımlarının sınanması için bazı test sonuçları

Değişken	VIF	Shapiro-Wilk		Grubbs		Durbin-Watson		Breusch-Pagan		
		W-İst.	p	G-İst.	p	DW-İst.	p	BP-İst.	p	sd.
Sabit Asitlik	6,268	0,874	<0,001	4,646	<0,001					
Uçucu Asitlik	1,621	0,954	<0,001	4,316	0,0018					
Sitrik Asit	2,597	0,926	<0,001	3,961	0,0090					
Kalan Şeker	1,858	0,661	<0,001	7,104	<0,001					
Klorit	1,994	0,427	<0,001	7,091	<0,001					
Serbest Kükürt	2,315	0,910	<0,001	4,058	0,0058	2,021	0,5294			
Toplam Kükürt	2,822	0,916	<0,001	2,708	0,971			35,686	<0,001	11
Yoğunluk	5,322	0,949	<0,001	3,886	0,0126					
pH	3,872	0,985	0,0036	3,759	0,0216					
Sülfat	2,005	0,730	<0,001	5,400	<0,001					
Alkol	2,373	0,774	<0,001	5,063	<0,001					
Kalite		0,797	<0,001	3,586	0,044					

p=0,05 olarak alınmıştır, sd= Serbestlik derecesi

Hata terimlerinin normal dağılıma sahip olup olmadığı için Shapiro-Wilk test istatistiğine ve aykırı değerlerin varlığı için ise Grubbs test istatistiğine bakılmıştır. Shapiro-Wilk olasılık değerlerinin 0,05'den küçük olmasından dolayı 0,05 anlamlılık düzeyinde hiçbir değişkenin normal dağılıma sahip olmadığı söylenebilir. Grubbs testi olasılık değerine göre; Toplam Kükürt değişkeninin ($p>0,05$) aykırı değer barındırmadığı, ancak geriye kalan 11 değişkenin ($p<0,05$) aykırı değer barındırdığı belirlenmiştir.

Değişen varyans durumunda EKK yöntemi yansız ve tutarlı olma özelliğini koruyabilmekte ancak minimum varyans özelliğini kaybetmektedir. Bu nedenle regresyon analiz sonuçlarını olumsuz etkilemektedir (İşleyen ve ark., 2018). Breusch-Pagan test istatistiği olasılık değerine ($p<0,05$) göre modelde değişen varyans problemi görülmüştür.

Çoklu doğrusal regresyonda temel model varsayımları arasında otokorelasyon da yer almaktadır. Otokorelasyon, hata terimlerinin birbirlerini izleyen değerleri arasındaki ilişki durumudur. Otokorelasyonun varlığında modelden sağlıklı sonuçlar elde edilememektedir (İşleyen ve ark., 2018). Durbin-Watson test istatistiği olasılık değerine ($p>0,05$) göre modelde otokorelasyon problemi görülmemektedir.

Şarap kalitesini etkileyen değişkenleri belirlemek üzere, 5 ve 11 bağımsız değişken alınarak 300 örneklem genişliğinde standart regresyonla birlikte, 0,20, 0,25 0,50 ve 0,75 kantil

değerleri için Kantil regresyon analizi uygulanmıştır. Daha sonra bootstrap yöntemiyle 300 örneklem genişliğinden 100 ve 200 örneklem genişliğine sahip yeni örnekler çekilerek benzer işlemler bu örneklemelere de uygulanmıştır. Regresyon katsayıları ve güven aralıkları 5 bağımsız değişken için Tablo 4’te ve 11 bağımsız değişken için ise Tablo 5’te verilmiştir.

Tablo 4. Beş değişkenli model için regresyon analiz sonuçları

Değişken	n=100					n=200					n=300				
	b ± S _b		%95 güven aralığı			b ± S _b		%95 güven aralığı			b ± S _b		%95 güven aralığı		
Sabit	5,91+0,59		4,75-7,07			5,85+0,39		5,09-6,61			5,52+0,93		4,94-6,11		
Sabit Asitlik	s,r	0,10+0,07	-0,03-0,23			0,05+0,05		-0,03-0,14			0,13+0,03		0,07-0,19		
	0,20	2,82e ⁻¹⁶ +0,03	-0,06-0,06			2,82e ⁻¹⁶ +0,03		-0,06-0,06			7,22e ⁻¹⁶ +0,02		-0,03-0,03		
	0,25	1,99e ⁻¹⁶ +0,03	-0,06-0,06			-		-			-1,45e ⁻¹⁶ +1,78e ⁻¹⁶		-4,95e ⁻¹⁷ -2,06e ⁻¹⁶		
	0,50	-6,22e ⁻¹⁷ +0,04	-0,07-0,72			-6,22e ⁻¹⁷ +0,04		-0,07-0,72			0,17+0,05		0,07-0,26		
	0,75	0,05+0,11	-0,16-0,27			0,05+0,11		-0,16-0,27			0,12+0,05		0,03-0,21		
Uçucu Asitlik	s,r	-2,05+0,49	-3,01- -1,09			-1,30+0,27		-1,83-0,78			1,63+0,26		-2,14- -1,12		
	0,20	-6,26e ⁻¹⁶ +0,17	-0,33-0,33			-6,26e ⁻¹⁶ +0,17		-0,33-0,33			1,61e ⁻¹⁶ +0,14		-0,28-0,28		
	0,25	-1,20e ⁻¹⁵ +0,22	-0,43-0,43			-		-			-4,37e ⁻¹⁵ +1,51e ⁻¹⁵		-7,35e ⁻¹⁵ -1,39e ⁻¹⁵		
	0,50	1,96e ⁻¹⁵ +0,22	-0,43-0,43			1,96e ⁻¹⁵ +0,22		-0,43-0,43			-1,29+0,41		-2,09- -0,48		
	0,75	-1,41+0,66	-2,72-0,10			-1,41+0,66		-2,72-0,10			-1,53+0,40		-2,31- -0,75		
Sitrik Asit	s,r	-1,70+0,50	-2,69-0,71			-0,90+0,29		-1,47- -0,33			-0,91+0,29		-1,49- -0,33		
	0,20	2,67e ⁻¹⁵ +0,18	-0,35-0,35			2,67e ⁻¹⁵ +0,18		-0,35-0,35			-8,88e ⁻¹⁶ +0,16		-0,32-0,32		
	0,25	-2,13e ⁻¹⁵ +0,22	-0,44-0,44			-		-			-1,06e ⁻¹⁵ +1,72e ⁻¹⁵		2,33e ⁻¹⁵ -2,33e ⁻¹⁵		
	0,50	2,16e ⁻¹⁵ +0,23	-0,46-0,46			2,16e ⁻¹⁵ +0,23		-0,46-0,46			-0,93+0,47		-1,85- -0,02		
	0,75	-1,11+0,71	-2,51-0,30			-1,11+0,71		-2,51-0,30			-1,25+0,45		-2,13- -0,36		
Kalan Şeker	s,r	0,00+0,05	-0,09-0,09			-0,03+0,03		-0,09-0,04			0,01+0,03		-0,05-0,07		
	0,20	-6,60e ⁻¹⁶ +0,02	0,04-0,04			-6,60e ⁻¹⁶ +0,02		0,04-0,04			2,90e ⁻¹⁷ +0,02		-0,04-0,04		
	0,25	1,11e ⁻¹⁶ +0,02	-0,04-0,04			-		-			1,60e ⁻¹⁵ +1,93e ⁻¹⁶		5,39e ⁻¹⁴ -5,39e ⁻¹⁴		
	0,50	-4,19e ⁻¹⁷ +0,03	-0,05-0,05			-4,19e ⁻¹⁷ +0,03		-0,05-0,05			-0,01+0,05		-0,11-0,10		
	0,75	-1,24+0,09	-0,29-0,04			-1,24+0,09		-0,29-0,04			-0,02+0,05		-1,20-0,08		
Klorit	s,r	0,87+0,67	-0,46-2,21			0,04+0,53		-0,94-1,08			-0,43+0,58		-1,57-0,71		
	0,20	-3,55e ⁻¹⁵ +0,33	-0,64-0,64			-3,55e ⁻¹⁵ +0,33		-0,64-0,64			-1,89e ⁻¹⁶ +0,32		-0,62-0,62		
	0,25	3,12e ⁻¹⁵ +0,30	-0,60-0,60			-		-			6,21e ⁻¹⁵ +3,39e ⁻¹⁵		1,29e ⁻¹⁴ -1,29e ⁻¹⁴		
	0,50	-8,88e ⁻¹⁶ +0,42	-0,84-0,84			-8,88e ⁻¹⁶ +0,42		-0,84-0,84			0,30+0,92		-1,51-2,10		
	0,75	0,23+1,30	-2,32-2,79			0,23+1,30		-2,32-2,79			-0,78+0,89		-2,53-0,98		
Kalite	s,r	5,91+0,59	4,75-7,07			5,85+0,39		5,09-6,61			5,52+0,93		4,94-6,11		
	0,20	5+0,24	4,53-5,47			5+0,24		4,53-5,47			5+0,16		4,68-5,32		
	0,25	5+0,26	4,49-5,52			-		-			5+1,74e-15		5-5		
	0,50	5+ 031	4,39-5,61			5+0,32		4,38-5,62			4,78+0,47		3,85-5,70		
	0,75	6,62+0,95	4,75-8,50			6,62+0,95		4,75-8,50			6,15+0,46		5,25-7,05		
	s,r	0,20	0,25	0,50	0,75	s,r	0,20	0,25	0,50	0,75	s,r	0,20	0,25	0,50	0,75
R²	0,18	0,0	-0,0	-0,0	0,14	0,12	-0,0	-	-0,0	0,09	0,20	-0,0	-0,0	0,03	0,05
R²_{düz}	0,14	-				0,10					0,19				
RMSE	0,61					0,59					0,65				

s.r. Standart regresyon

Tablo 4 incelendiğinde; en yüksek regresyon sabiti değerinin, 5,91 ile 100 örneklem genişliğinde elde edildiği ve bunu 5,85 ile 200 ve 5,52 ile 300 örneklem genişliklerinin izlediği görülmektedir. Ayrıca $n=100$ olduğunda diğer örneklem genişliklerine göre daha geniş güven aralığı elde edildiği gözlenmiştir.

200 örneklem genişliği ve 0,25 kanilinde hem bağımlı hem de bağımsız değişkenler için herhangi bir sonuç hesaplanamamıştır. Ayrıca 0,2, 0,5 ve 0,75 kantillerinde bağımsız değişkenlere ait regresyon katsayısı, standart hata ve güven aralığı değerleri 100 ve 200 örneklem genişliklerinde aynı bulunmuştur.

Sabit Asitlik: Standart regresyon modeli için en yüksek kısmi regresyon katsayısı 0,13 ile 300 örneklem genişliğinde elde edilirken, bunu 0,10 ile 100 ve 0,05 ile 200 örneklem genişlikleri izlemektedir.

100 örneklem genişliğinde; kantil değerler için en yüksek kısmi regresyon katsayısı 0,05'lik değerle 0,75 kantilinde bunu, $2,82e^{-16}$ ile 0,20 kantili izlerken, en düşük katsayı ise $-6,22e^{-17}$ ile 0,5 kantilinde elde edilmiştir.

Örneklem genişliği 200 olması durumunda; 0,75 kantili ile standart regresyon modelindeki kısmi regresyon katsayısı aynı (0,05) ve standart hata değerleri birbirine yakın bulunmuştur. Kantiller içerisinde 0,50 kantilinin en geniş güven aralığına sahip olduğu gözlenmiştir.

Örneklem genişliği 300 olması durumunda; kısmi regresyon katsayılarında değişkenlik gözlenmekle birlikte, standart regresyon modelindeki katsayıya en yakın katsayı 0,75 kantil değerinde elde edilmiş, ancak standart hata değeri, standart modeldeki değerden bir miktar büyük bulunmuştur. En yüksek katsayı ve en geniş güven aralığı ise 0,5 kantil değerinde elde edilmiştir.

Uçucu Asitlik: Standart regresyon modeli için en yüksek katsayı 1,63 ile 300 örneklem genişliğinde elde edilirken, bunu $-1,30$ ile 200 ve $-2,05$ ile 100 örneklem genişliği izlemiştir.

$n=100$ olması durumunda; kantil değerler için en yüksek kısmi regresyon katsayısı $1,96e^{-15}$ ile 0,50 kantilinde elde edilirken bunu, $-1,20e^{-12}$ ile 0,25 kantili izlemiş, en düşük katsayı ise $-1,41$ ile 0,75 kantil değerinde elde edilmiştir. 100 ve 200 örneklem genişliği için en geniş güven aralığı 0,75 kantilinde elde edilmiştir.

Örneklem genişliği 300 olması durumunda; kısmi regresyon katsayılarında değişkenlik gözlenmekle birlikte, standart regresyon modelindeki katsayıya en yakın kısmi regresyon katsayısı 0,20 kantilinde, en geniş güven aralığı ise 0,50 kantilinde elde edilmiştir.

Sitrik Asit: Standart regresyon modeli için en yüksek kısmi regresyon katsayısı sırasıyla $-0,90$ ile 200, $-0,91$ ile 300 ve $-1,70$ ile 100 örneklem genişliklerinde elde edilmiştir.

Örneklem genişliği 100 olduğunda kantiller için en yüksek kısmi regresyon katsayısı $2,67e^{-15}$ ile 0,20 kantilinde elde edilmiş ve bunu, $2,16e^{-15}$ ile 0,50 kantili izlemiş, en düşük katsayı ise -1,11 ile 0,75 kantilinde elde edilmiştir. En geniş güven aralığı ise 0,75 kantilinde gözlenmiştir.

Örneklem genişliği 300 olduğunda; kısmi regresyon katsayılarında değişkenlik gözlenmekle birlikte, standart regresyon modelindeki katsayıya (-0,91) en yakın katsayı 0,50 kantilinde (-0,93) elde edilmiş ancak standart hata değeri, standart modeldeki değerden bir miktar büyük bulunmuştur. En geniş güven aralığı ise 0,50 kantilinde elde edilmiştir.

Kalan Şeker: Standart regresyon modeli için en yüksek kısmi regresyon katsayısı 0,01 ile 300 örneklem genişliğinde elde edilmiş, bunu 0 ile 100 ve -0,03 ile 200 örneklem genişlikleri izlemiştir.

Örneklem genişliği 100 olduğunda; kantiller için en yüksek kısmi regresyon katsayısı $1,11e^{-16}$ ile 0,25 kantilinde ve en düşük katsayı ise -0,12 ile 0,75 kantilinde elde edilmiştir. En geniş güven aralığı ise 0,75 kantilinde gözlenmiştir.

Örneklem genişliği 300 olduğunda; kısmi regresyon katsayılarında değişkenlik gözlenmekle birlikte, standart regresyon modelindeki katsayıya en yakın katsayı 0,20 kantilinde elde edilmiş, ancak standart hata değeri, standart modeldeki değerden bir miktar küçük bulunmuştur. En geniş güven aralığı ise 0,75 kantilinde elde edilmiştir.

Klorit: Standart regresyon modeli için en yüksek kısmi regresyon katsayısı 0,87 ile 100 örneklem genişliğinde elde edilirken, bunu 0,04 ile 200 ve -0,43 ile 300 örneklem genişlikleri izlemiştir.

100 örneklem genişliğinde; kantil değerler için en yüksek kısmi regresyon katsayısı 0,23 ile 0,75 kantilinde elde edilmiş ve bunu, $3,12e^{-15}$ ile 0,25 kantili izlemiş, en düşük katsayı ise - $8,88e^{-16}$ ile 0,5 kantilinde elde edilmiştir. Klorit değişkeni için 100 ve 200 örneklem genişliklerinde en geniş güven aralığı 0,75 kantilinde gözlenmiştir.

Örneklem genişliği 300 olduğunda; kısmi regresyon katsayılarında değişkenlik gözlenmekle birlikte, standart regresyon modelindeki katsayıya en yakın katsayı 0,75 kantilinde elde edilmiş ancak standart hata değeri, standart modeldeki değerden bir miktar büyük bulunmuştur. Kantil değerler için en büyük katsayı ve en geniş güven aralığı ise 0,5 kantilinde elde edilmiştir.

Kalite: Standart regresyon modeli için en yüksek kısmi regresyon katsayısı 5,91 ile 100 örneklem genişliğinde en düşüğü ise 5,52 ile 300 örneklem genişliğinde elde edilmiştir.

100 örneklem genişliğinde kantil değerler için en yüksek kısmi regresyon katsayısı 6,62 ile 0,75 kantilinde, geriye kalan kantillerde ise 5 olarak bulunmuştur. Dolayısıyla 0,75 kantili

için kısmi regresyon katsayısının diğer kantillerdeki katsayılardan farklı ve yüksek olduğu belirlenmiştir.

Örneklem genişliği 200 olması durumunda; 0,20, 0,50, 0,75 kantilleri için kısmi regresyon katsayıları, 100 örneklem genişliğindeki katsayılarla aynı ve standart hata ve güven aralığı değerleri ise ya aynı ya da birbirlerine çok yakın hesaplanmıştır. En geniş güven aralığı 0,75 kantiline aittir.

Örneklem genişliği 300 olduğunda; kısmi regresyon katsayılarında değişkenlik gözlenmekte, standart regresyon modelindeki katsayıya en yakın katsayı 5 ile 0,2 ve 0,25 kantilinde elde edilirken en geniş güven aralığı ise 0,75 kantilinde elde edilmiştir.

Tablo 5. On bir değişkenli model için regresyon analiz sonuçları

Değişken	n = 100		n = 200		n = 300	
	b ± S _b	%95 güven aralığı	b ± S _b	%95 güven aralığı	b ± S _b	%95 güven aralığı
Sabit	45,31+122,67	-198,47-289,10	0,51+85,62	-169,4-169,4	71,41+57,89	0,05-0,28
s,r	0,08+0,14	-0,20-0,36	0,12+0,10	-0,07-0,31	0,17+0,06	0,05-0,28
Sabit Asitlik	0,20	-0,07+0,20	-4,69e ⁻¹⁶ +0,07	-0,13-0,13	5,85e ⁻¹⁶ +0,03	-0,07-0,07
0,25	1,05e ⁻¹⁵ +0,05	-0,11-0,11	-8,01e ⁻¹⁷ +0,05	-0,09-0,09	4,12e ⁻¹⁶ +1,84e ⁻¹⁶	5,03e ⁻¹⁷ -7,73e ⁻¹⁶
0,50	2,19e ⁻¹⁵ +0,14	-0,27-0,27	0,29+0,10	-0,16-0,21	0,18+0,10	-0,02-0,37
0,75	0,09+0,26	-0,43-0,61	0,05+0,19	-0,32-0,41	0,13+0,08	-0,03-0,29
s,r	-1,95+0,51	-2,96- -0,93	-1,22+0,28	-1,78- -0,66	-1,32+0,25	-1,81- -0,83
Uçucu Asitlik	0,20	-0,82+0,71	4,44e ⁻¹⁶ +0,21	-0,41-0,41	-4,72e ⁻¹⁵ +0,14	-0,29-0,29
0,25	2,08e ⁻¹⁴ +0,20	-0,39-0,39	3,19e ⁻¹⁵ +0,14	-0,27-0,27	-2,47e ⁻¹⁵ +7,82e ⁻¹⁶	-4,01e ⁻¹⁵ +9,34e ⁻¹⁶
0,50	-3,80e ⁻¹⁵ +0,50	-0,99-0,99	-0,47+0,28	-1,03-0,09	-0,89+0,42	-1,72- -0,06
0,75	-1,61+0,95	-3,50-0,28	-0,85+0,55	-1,93-0,24	-0,73+0,35	-1,41- -0,06
s,r	-1,41+0,57	-2,55- -0,28	-0,71+0,35	-1,40- -0,03	-0,68+0,29	-1,25-0,10
Sitrik Asit	0,20	-0,67+0,79	1,43e ⁻¹⁵ +0,25	-0,50-0,50	-2,26e ⁻¹⁶ +0,17	-0,34-0,34
0,25	1,71e ⁻¹⁴ +0,22	-0,44-0,44	1,04e ⁻¹⁴ +0,17	-0,33-0,33	-2,77e ⁻¹⁵ +9,23e ⁻¹⁶	-4,58e ⁻¹⁵ +9,52e ⁻¹⁶
0,50	-4,14e ⁻¹⁵ +0,55	-1,10-1,10	-0,04+0,35	-0,73-0,64	-0,24+0,50	-1,22-0,74
0,75	-0,89+1,06	-2,99-1,22	-0,35+0,68	-1,68-0,99	-0,00+0,40	-0,80-0,79
s,r	0,00+0,07	-0,14-0,14	-0,03+0,05	-0,13-0,07	0,04+0,04	-0,04-0,12
Kalan Şeker	0,20	-0,05+0,10	-9,10e ⁻¹⁸ +0,04	-0,07-0,07	-8,85e ⁻¹⁷ +0,02	-0,05-0,05
0,25	3,72e ⁻¹⁶ +0,03	-0,05-0,05	4,36e ⁻¹⁶ +0,03	-0,05-0,05	-8,76e ⁻¹⁷ +1,30e ⁻¹⁶	3,43e ⁻¹⁶ -1,68e ⁻¹⁶
0,50	1,42e ⁻¹⁶ +0,07	-0,13-0,13	-0,06+0,05	-0,65-0,04	-0,02+0,07	-0,16-0,12
0,75	0,12+0,13	-0,14-0,37	-0,07+0,10	-0,27-0,12	-0,00+0,06	-0,12-0,11
s,r	-0,00+1,18	-2,36-2,35	-0,40+0,84	-2,05-1,26	-0,54+0,69	-1,89-0,81
Klorit	0,20	0,31+1,64	-3,65e ⁻¹⁵ +0,61	-1,20-1,20	-1,50e ⁻¹⁵ +0,40	-0,79-0,79
0,25	-7,55e ⁻¹⁵ +0,46	-0,91-0,91	-1,20e ⁻¹⁴ +0,41	-0,81-0,81	3,11e ⁻¹⁵ +2,16e ⁻¹⁵	-1,14e ⁻¹⁵ -7,36e ⁻¹⁵
0,50	2,68e ⁻¹⁵ +1,15	-2,28-2,28	-0,25+0,08	-1,91-1,41	0,02+1,17	-2,28-2,32
0,75	-1,80+2,20	-6,17-2,57	-2,34+1,63	-5,55-0,88	-1,84+0,95	-3,70-0,02
s,r	0,50+0,01	-0,00-0,04	0,00+0,00	-0,00-0,02	0,01+0,01	-0,00-0,01
Serbest Kükürt	0,20	0,00+0,01	1,88e ⁻¹⁷ +0,00	-0,00-0,00	-2,60e ⁻¹⁶ +0,00	-0,00-0,00
0,25	9,46e ⁻¹⁸ +0,00	-0,00-0,00	5,37e ⁻¹⁷ +0,00	-0,00-0,00	-7,64e ⁻¹⁸ +1,85e ⁻¹⁷	-4,41e ⁻¹⁷ -2,88e ⁻¹⁷
0,50	-1,07e ⁻¹⁶ +0,00	-0,02-0,02	0,00+0,00	-0,00-0,00	0,01+0,01	-0,01-0,03
0,75	0,02+0,00	-0,01-0,06	0,00+0,01	-0,01-0,04	0,01+0,00	-0,00-0,02
s,r	-0,00+0,00	-0,01-0,04	-0,00+0,00	-0,00-0,00	-0,00+0,00	-0,00- -0,00

	0,20	-0,00+0,00	-0,01-0,00	-2,33e ⁻¹⁷ +0,00	-0,00-0,00	5,99e ⁻¹⁷ +0,00	-0,00-0,00								
Toplam Kükürt	0,25	2,85e ⁻¹⁷ +0,00	-0,00-0,00	-3,14e ⁻¹⁹ +0,00	-0,00-0,00	-3,97e ⁻¹⁹ +5,21e ⁻¹⁸	-1,06e ⁻¹⁷ -9,85e ⁻¹⁸								
	0,50	4,49e ⁻¹⁷ +0,00	-0,00-0,00	-0,00+0,00	-0,00-0,00	-0,00+0,00	-0,00-0,00								
	0,75	-0,01+0,00	-0,02-0,00	-0,00+0,00	-0,01-0,00	-0,01+0,00	-0,01-0,00								
	s,r	-39,31+124,92	-287,59-2008,9	2,50+87,09	-169,30-174,31	-70,0+58,89	-185,9-45,9								
Yoğunluk	0,20	183,82+173,33	-160,65-528,30	1,89e ⁻¹³ +63,28	-124,82-124,82	-3,27e ⁻¹³ +34,39	-67,70-67,70								
	0,25	-3,13e ⁻¹² +48,25	-95,89-95,89	-1,19e ⁻¹² +42,64	-84,11-84,11	-8,97e ⁻¹⁴ +1,86e ⁻¹³	4,55e ⁻¹³ -2,76e ⁻¹³								
	0,50	-9,20e ⁻¹³ +121,40	-241,27-241,27	20,77+87,10	-151,1-192,6	-27,9+100,46	-225,6-169,8								
	0,75	-207,77+232,17	-669,16-253,62	52,34+168,97	-280,9-385,7	-52,3+81,24	-212,2-107,6								
	s,r	-0,76+0,94	-2,62-1,10	0,13+0,63	-1,11-1,37	0,39+0,44	-0,48-1,26								
Ph	0,20	-1,14+1,30	-3,73-1,44	1,10e ⁻¹⁵ +0,46	-0,90-0,90	5,99e ⁻¹⁵ +0,26	-0,51-0,51								
	0,25	2,03e ⁻¹⁴ +0,36	-0,72-0,72	-1,85e ⁻¹⁵ +0,31	-0,61-0,61	1,52e ⁻¹⁵ +1,39e ⁻¹⁵	-1,06e ⁻¹⁷ -4,26e ⁻¹⁵								
	0,50	4,41e ⁻¹⁶ +0,91	-1,81-1,81	-0,03+0,63	-1,28-1,21	0,41+0,75	1,08-1,89								
	0,75	-1,40+1,74	-4,86-2,07	0,04+1,22	-2,37-2,45	0,32+0,61	-0,88-1,52								
	s,r	0,16+0,33	-0,49-0,81	0,12+0,24	-0,34-0,60	0,44+0,20	0,04-0,84								
Sülfat	0,20	-0,23+0,45	-1,13-0,67	1,74e ⁻¹⁷ +0,17	-0,34-0,34	1,09e ⁻¹⁵ +0,12	-0,24-0,24								
	0,25	4,14e ⁻¹⁵ +0,13	-0,25-0,25	3,92e ⁻¹⁶ +0,12	-0,23-0,23	4,35e ⁻¹⁶ +6,46e ⁻¹⁶	-8,36e ⁻¹⁶ -1,71e ⁻¹⁵								
	0,50	-1,18e ⁻¹⁵ +0,32	-0,63-0,63	0,00+0,24	-0,46-0,47	-0,02+0,35	-0,70-0,67								
	0,75	0,44+0,60	-0,77-1,65	0,56+0,46	-0,36-1,47	0,75+0,28	0,19-1,30								
	s,r	0,24+0,15	-0,05-0,54	0,18+0,09	0,00-0,37	0,19+0,07	0,06-0,32								
Alkol	0,20	0,18+0,20	-0,22-0,60	-5,33e ⁻¹⁶ +0,07	-0,13-0,13	-5,75e ⁻¹⁶ +0,04	-0,08-0,08								
	0,25	-3,61e ⁻¹⁵ +0,06	-0,11-0,11	-7,01e ⁻¹⁶ +0,05	-0,09-0,09	6,75e ⁻¹⁷ +2,09e ⁻¹⁶	-3,44e ⁻¹⁶ -4,79e ⁻¹⁶								
	0,50	-9,05e ⁻¹⁶ +0,14	-0,29-0,29	0,21+0,09	0,02-0,39	0,20+0,11	-0,02-0,42								
	0,75	0,37+0,27	-0,18-0,91	0,30+0,18	-0,06-0,66	0,25+0,09	0,07-0,43								
	s,r	45,31+122,67	-198,47-289,10	0,51+85,82	-169,39-169,41	71,41+57,89	0,05-0,28								
Kalite	0,20	-174,79+179,19	-513,03-163,44	5+62,21	-117,71-127,71	5+33,81	-61,54-61,54								
	0,25	-3,61e ⁻¹⁵ +2,13	-89,15-99,15	5+41,92	-77,69-77,69	5+1,82e ⁻¹³	-5-5								
	0,50	5+47,36	-231,90-241,90	-17,30+85,63	-186,22-151,61	29,02+98,74	165,32-223,37								
	0,75	5+119,20	-239,04-667,04	-49,21+166,15	-376,9-278,5	53,87+79,85	-103,3-211,03								
	s,r	0,20	0,25	0,50	0,75	s,r	0,20	0,25	0,50	0,75	s,r	0,20	0,25	0,50	0,75
R²	0,25	0,03	-0,0	0,0	0,23	0,21	-0,0	-0,0	0,0	0,19	0,32	-0,0	-0,0	0,10	0,17
R²_{Düz}	0,16					0,16					0,30				
RMSE	0,60					0,57					0,61				

s.r. Standart regresyon

Tablo 5 incelendiğinde; en yüksek regresyon sabitinin, 71,41 ile 300 örneklem genişliğinde elde edildiği ve bunu 45,31 ile 100 ve 0,51 ile 200 örneklem genişliklerinin izlediği görülür. En geniş güven aralığının ise 100 örneklem genişliğinde elde edildiği gözlenmiştir.

Sabit Asitlik: Standart regresyon modeli için kısmi regresyon katsayıları sırasıyla 0,17 ile 300, 0,12 ile 200 ve 0,08 ile 100 örneklem genişliklerinden elde edilmiştir.

100 örneklem genişliğinde; kantiller için en yüksek kısmi regresyon katsayısı 0,09 ile 0,75 kantilinde elde edilmiş, bunu, 2,19e⁻¹⁵ ile 0,50 kantili izlemiş ve en düşük katsayı ise -0,07 ile 0,20 kantilinde elde edilmiştir.

Örneklem genişliği 200 olduğunda; kısmi regresyon katsayılarında değişkenlik gözlenmekte ve standart regresyon modelindeki katsayıya en yakın katsayı 0,75 kantilinde elde edilmiştir. Ayrıca en geniş güven aralığı da bu kantilde gözlenmiştir.

Örneklem genişliği 300 olması durumunda; standart regresyon modelindeki kısmi regresyon katsayısına en yakın katsayı 0,5 kantilinde ve en yüksek standart hata ile en geniş güven aralığı da yine bu kantilde elde edilmiştir. Ayrıca, sabit değer için güven aralığının, standart regresyon analizindeki güven aralığı ile aynı olduğu gözlenmiştir.

Uçucu Asitlik: Standart regresyon için en yüksek kısmi regresyon katsayısı -1,22 ile 200 örneklem genişliğinde elde edilmiş ve bunu 300 ve 100 (-1,32, -1,95) örneklem genişlikleri izlemiştir.

100 örneklem genişliğinde; kantiller için en yüksek kısmi regresyon katsayısı $2,08e^{-14}$ ile 0,25 kantilinde elde edilmiş, bunu $-3,80e^{-15}$ ile 0,50 kantili izlemiş ve en düşük katsayı ise -1,61 ile 0,75 kantilinde elde edilmiştir.

Örneklem genişliği 200 olduğunda; kısmi regresyon katsayılarında bir değişkenlik gözlenmiştir. Standart regresyon modelindeki katsayıya en yakın katsayı 0,75 kantilinde elde edilmiş, yine en yüksek standart hata değeri ile en geniş güven aralığı da bu kantilde gözlenmiştir.

Örneklem genişliği 300 olması durumunda; standart regresyon modelindeki kısmi regresyon katsayısına en yakın katsayı 0,50 kantilinde elde edilmiş, ayrıca en yüksek standart hata değeri ve en geniş güven aralığı da yine 0,50 kantilinde elde edilmiştir.

Sitrik Asit: Standart regresyon için en yüksek kısmi regresyon katsayısı -0,68 ile 300 örneklem genişliğinde sağlanmış ve bunu 200 ve 100 (-0,71, -1,41) örneklem genişlikleri izlemiştir.

100 örneklem genişliğinde; kantiller için en yüksek kısmi regresyon katsayısı $1,71e^{-14}$ ile 0,25 kantilinde elde edilmiş ve bunu, $-4,14e^{-15}$ ile 0,50 kantili izlemiştir. Kantillerde en düşük katsayı ise -0,89 ile 0,75 kantilinde elde edilmiştir.

Örneklem genişliği 200 olduğunda; kısmi regresyon katsayılarında bir değişkenlik gözlenmekte ve standart regresyon katsayısına en yakın katsayı 0,75 kantilinde elde edilmiştir. Ayrıca en yüksek standart hata ve en geniş güven aralığı da yine bu kantilde gözlenmiştir.

Örneklem genişliği 300 olması durumunda; standart kısmi regresyon katsayısına en yakın katsayı 0,5 kantilinde elde edilmiş, ayrıca en büyük standart hata değeri ve en geniş güven aralığı da yine 0,50 kantilinde elde edilmiştir.

Kalan Şeker: Standart regresyon için en yüksek kısmi regresyon katsayısı 0,04 ile 300 örneklem genişliğinde elde edilmiş ve bunu 100 ve 200 (0,00, -0,03) örneklem genişlikleri izlemiştir.

n=100 olduğunda; kantiller için en yüksek kısmi regresyon katsayısı 0,12 ile 0,75 kantilinde elde edilmiş ve bunu, $3,72e^{-16}$ ile 0,25 kantili izlemiştir. Ayrıca, en düşük kısmi regresyon katsayısı ise -0,05 ile 0,20 kantilinde elde edilmiştir.

Örneklem genişliği 200 olması durumunda; kısmi regresyon katsayılarında değişkenlik gözlenmiş ve standart kısmi regresyon katsayısına en yakın katsayı 0,20 kantilinde elde edilmiştir. Ayrıca en yüksek standart hata 0,75 ve en geniş güven aralığı ise 0,50 kantilinde gözlenmiştir.

Örneklem genişliği 300 olması durumunda; standart regresyon modelindeki kısmi regresyon katsayısına en yakın katsayı 0,75 kantilinde elde edilmiştir. Ayrıca en yüksek standart hata ve en geniş güven aralığı ise 0,50 kantilinde elde edilmiştir.

Klorit: Standart regresyon için en yüksek kısmi regresyon katsayısı -0,00 ile 100 örneklem genişliğinde elde edilmiş, bunu sırasıyla 200 ve 300 örneklem genişlikleri izlemiştir.

n=100 olduğunda; kantiller için en yüksek kısmi regresyon katsayısı 0,31 ile 0,20 kantilinde elde edilmiş ve bunu, $2,68e^{-15}$ ile 0,50 kantili izlemiştir. En düşük kısmi regresyon katsayısı ise -1,80 ile 0,75 kantilinde elde edilmiştir.

Örneklem genişliği 200 olması durumunda; standart kısmi regresyon katsayısına en yakın katsayı 0,50 kantilinde, en farklı katsayısı ise 0,75 kantilinde elde edilmiştir. Ayrıca en yüksek standart hata ve en geniş güven aralığı da 0,75 kantilinde gözlenmiştir.

Örneklem genişliği 300 olduğunda; kısmi regresyon katsayılarında değişkenlik gözlenmiş ve standart kısmi regresyon katsayısına en yakın katsayı 0,20 kantilinde elde edilmiştir. Ancak en büyük standart hata ve en geniş güven aralığı 0,50 kantilinde elde edilmiştir.

Serbest Kükürt: Standart regresyon için en yüksek kısmi regresyon katsayısı 0,50 ile 100 örneklem genişliğinde elde edilmiş ve bunu 300 (0,01) ve 200 (0,00) örneklem genişlikleri izlemiştir.

n=100 olduğu durumda; kantiller için en yüksek kısmi regresyon katsayısı 0,02 ile 0,75 kantininde elde edilmiş ve bunu, $9,46e^{-18}$ ile 0,25 kantili izlemiştir. En düşük kısmi regresyon katsayısı ise $-1,07e^{-16}$ ile 0,5 kantilinde elde edilmiştir.

Örneklem genişliği 200 olduğunda; kısmi regresyon katsayılarının birbirlerine çok yakın oldukları 0,5, 0,75 kantiller ve standart regresyon katsayılarının aynı olduğu belirlenmiştir. Yine 0,75 kantili hariç diğer kantiller için standart hata ve güven aralığı değerleri de aynı bulunmuş ve en yüksek standart hata ve güven aralığı ise 0,75 kantilinde gözlenmiştir.

Örneklem genişliği 300 olduğunda; kısmi regresyon katsayıları, standart hata ve güven aralığı değerlerinin 200 örneklem genişliğindeki değerlerle oldukça benzerlik gösterdiği, ancak burada en büyük standart hatanın ve en geniş güven aralığının 0,50 kantilide olduğu gözlenmiştir.

Toplam Kükürt: Standart regresyonda tüm örneklem genişlikleri için kısmi regresyon katsayıları (-0,00) aynı bulunmuştur.

Örneklem genişliği 100 olduğunda; kantiller için en yüksek kısmi regresyon katsayısı $4,49e^{-17}$ ile 0,5 kantilinde elde edilirken bunu sırasıyla, 0,25, 0,20 ve 0,75 kantilleri izlemektedir.

$n=200$ olduğu durumda; kısmi regresyon katsayılarının birbirine oldukça yakın olduğu, 0,50 ve 0,75 kantilleri için regresyon katsayıları ile standart regresyon katsayısının aynı olduğu belirlenmiştir. Ayrıca, tüm kantillerde standart hata değerinin (0,00) aynı ve 0,75 kantili hariç diğer kantillerde güven aralığının da aynı olduğu gözlenmiştir.

Örneklem genişliğinin 300 olduğu durumda; en yüksek kısmi regresyon katsayısının $5,99e^{-17}$ ile 0,2 kantilinde ve 0,5 kantili ile standart regresyon kısmi regresyon katsayılarının -0,00 olarak elde edildiği belirlenmiştir. Tüm modellerde standart hata değerlerinin hemen hemen aynı ve en geniş güven aralığının ise 0,25 kantilinde elde edildiği söylenebilir.

Yoğunluk: Standart regresyon modeli için en yüksek kısmi regresyon katsayısı 2,50 ile 200 örneklem genişliğinde elde edilmiş ve bunu sırasıyla 100 (-39,31) ve 300 (-70,0) örneklem genişlikleri izlemiştir.

$n=100$ olduğunda; kantiller için en yüksek kısmi regresyon katsayısı 183,82 ile 0,20 kantilinde ve en düşük katsayı ise -207,77 ile 0,75 kantilinde elde edilmiştir. Ayrıca standart regresyon kısmi regresyon katsayısına en yakın katsayı $-3,13e^{-12}$ ile 0,25 kantilinde gözlenmiştir.

Örneklem genişliğinin 200 olması durumunda; kantil kısmi regresyon katsayıları arasında en yüksek katsayı (52,34) 0,75 kantilinde ve en düşük katsayı ise $(-1,19e^{-12})$ 0,25 kantilinde elde edilirken, standart regresyon kısmi regresyon katsayısına en yakın katsayı ise 0,20 kantilinde elde edilmiştir. En yüksek standart hata ve en geniş güven aralığı da 0,75 kantilinde gözlenmiştir.

300 örneklem genişliğinde, standart regresyon katsayısına en yakın kısmi regresyon katsayısı 0,75 kantilide, en küçük katsayı 0,25 kantilinde ve en büyük katsayısı ise standart regresyonda saptanmıştır. Ayrıca en büyük standart hata ve en geniş güven aralığı 0,5 kantilinde gözlenmiştir.

Ph: Standart regresyon modeli için en yüksek kısmi regresyon katsayısı 0,39 ile 300 örneklem genişliğinde elde edilmiş ve bunu sırasıyla 200 ve 300 örneklem genişlikleri izlemiştir.

n=100 olduğunda; standart regresyon analizindeki kısmi regresyon katsayısına en yakın katsayı 0,2 kantilinde elde edilmiş, ayrıca kantiller için en yüksek kısmi regresyon katsayısı $2,03e^{-14}$ ile 0,25 kantilinde ve en düşük katsayı ise -1,40 ile 0,75 kantilinde elde edilmiştir.

n=200 olduğunda; kantil kısmi regresyon katsayıları arasında en yüksek katsayının 0,04 ile 0,75 kantilinde elde edildiği ve bunun standart regresyondaki kısmi regresyon katsayısına en yakın katsayı olduğu, yine en yüksek standart hata ve güven aralığının da bu kantile ait olduğu gözlenmiştir.

Örneklem genişliği 300 olduğunda; en yüksek kısmi regresyon katsayısının ve standart hatanın, en geniş güven aralığının 0,5 kantilinde elde edildiği, ayrıca standart regresyonun kısmi regresyon katsayısına en yakın katsayının da bu kantille sağlandığı gözlenmiştir.

Sülfat: Standart regresyon modeli için en yüksek kısmi regresyon katsayısı 0,44 ile 300 örneklem genişliğinde ve en küçük katsayı ise 0,12 ile 200 örneklem genişliğinde elde edilmiştir.

100 örneklem genişliğinde; en yüksek kısmi regresyon katsayısı 0,44 ile 0,75 kantilinde, sonra standart regresyonda (0,16) ve en düşük katsayı (-0,23) ise 0,20 kantilinde elde edilmiştir.

Örneklem genişliği 200 olduğunda; standart regresyonun kısmi regresyon katsayısına en yakın katsayı 0,25 kantilinde elde edilmiştir. Ayrıca en yüksek kısmi regresyon katsayısı (0,56), en büyük standart hata (0,46) ve en geniş güven aralığı 0,75 kantilinde gözlenmiştir.

Örneklem genişliği 300 olduğunda; standart regresyonun kısmi regresyon katsayısına en yakın ve en yüksek katsayı 0,75 değeriyle 0,75 kantilinde elde edilmiştir. Ancak en büyük standart hata (0,35) ve en geniş güven aralığı (-0,70-0,67) ise 0,50 kantilinde elde edilmiştir.

Alkol: Standart regresyon modeli için büyükten küçüğe doğru kısmi regresyon katsayıları sırasıyla 0,24 ile 100, 0,19 ile 300 ve 0,18 ile 200 örneklem genişliğinden elde edilmiştir.

100 örneklem genişliğinde; en yüksek kısmi regresyon katsayısı 0,37 ile 0,75 kantilinde, sonra standart regresyonda (0,24) ve en düşük katsayı ($-3,61e^{-15}$) ise 0,25 kantilinde elde edilmiştir. En geniş güven aralığı ve en büyük standart hata ise 0,75 kantilinde gözlenmiştir.

Örneklem genişliği 200 olduğunda; en yüksek kısmi regresyon katsayısı (0,30), en büyük standart hata (0,18) ve en geniş güven aralığı (-0,06-0,66) 0,75 kantilinde gözlenmiştir. Standart regresyonun kısmi regresyon katsayısına en yakın katsayı ise 0,5 kantilinde elde edilmiştir.

300 örneklem genişliğinde; standart regresyonun kısmi regresyon katsayısına en yakın katsayı (0,20) ve en geniş güven aralığı 0,50 kantilinde elde edilmiştir.

Kalite: Standart regresyon için büyükten küçüğe doğru kısmi regresyon katsayıları sırasıyla 71,41 ile 300, 45,31 ile 100 ve 0,51 ile 200 örneklem genişliğinde elde edilirken, örneklem genişliğinin 100 olması durumunda en geniş güven aralığı elde edilmiştir.

100 örneklem genişliğinde; kantiller için en yüksek kısmi regresyon katsayısı 5 ile 0,5 ve 0,75 kantillerinde ve en düşük katsayı ise -174,79 ile 0,20 kantilinde elde edilmiştir.

Örneklem genişliği 200 olduğunda; standart regresyonun kısmi regresyon katsayısına (0,51) en yakın ve en yüksek kısmi regresyon katsayısı (5) 0,20 ve 0,25 kantilleride, en yüksek standart hata (166,15) ve en geniş güven aralığı 0,75 kantilinde (-376,9-278,5) elde edilmiştir.

Örneklem genişliği 300 olması durumunda; kısmi regresyon katsayıları arasında en yüksek katsayı (71,41) standart regresyonla ve buna en yakın katsayı (53,87) ise 0,75 kantilinde elde edilmiştir. Ayrıca en geniş güven aralığı da yine 0,75 kantilinde gözlenmiştir.

Sonuç

Beş bağımsız ve bir bağımlı değişkene sahip uygulamalarda, tüm örneklem genişliklerine göre sonuçlar değerlendirildiğinde; hesaplanan sabit katsayılar birbirine yakın bulunmuş ve örneklem hacmi arttıkça bu artışa bağlı olarak katsayılarda küçük azalmalar meydana gelmiştir. Ayrıca regresyon uygulamaları için hesaplanan sabit değerlerin, standart regresyonda kalite değişkeninin katsayısına eşit olduğu görülmüştür.

Sabit regresyona göre; tüm örneklem genişliklerinde kalite değişkenine en yüksek etki Uçucu Asitlik ve en düşük etki ise Kalan Şeker değişkeni tarafından yapılmıştır. Uçucu Asitlik değişkeni, 100 ve 200 örneklem genişliklerinde kaliteyi olumsuz etkilerken, 300 örneklem genişliğinde ise olumlu etkilemiştir. Ayrıca tüm örneklem genişliklerinde kalite üzerinde Sabit Asitliğin olumlu ve Sitrik asitin ise olumsuz etkisi olduğu gözlenmiştir.

Kantil regresyon analizlerine göre; 200 örneklem genişliğinde 0,25 kantilinde değişkenler için herhangi bir değer hesaplanamamıştır. Bu nedenle bu kantil hariç diğer kantillerde tüm değişkenler için 100 ve 200 örneklem genişliklerinde aynı değerler hesaplanırken, 300 örneklem genişliğinde farklı değerler hesaplanmış ve özellikle de 0,5 kantili kısmi regresyon katsayılarında önemli değişkenlikler gözlenmiştir. Böylece örneklem hacminin 200'e çıkarılmasının, değişkenlerin kalite üzerindeki etkisini değiştirmede, ancak 300'e çıkarılması durumunda değişkenlerin kalite üzerindeki etkisini değiştirdiği ve özellikle de 0,5 kantilinin bundan çok etkilendiği belirlenmiştir.

100 ve 200 örneklem genişliklerinde 0,75 dışındaki diğer kantillerde değişkenlerin kaliteye etkilerinin çok düşük olduğu belirlenmiştir. 0,75 kantilinde, Sabit Asitlik ve Klorit değişkenleri kaliteyi olumlu ve diğer üç değişken ise olumsuz etkilemiştir. Ayrıca en yüksek

etki, -1,41 ile Uçucu Asitlik tarafından yapılmıştır. 300 örneklem genişliğinde ise 0,75 kantilinin yanı sıra, 0,50 kantilinde de değişkenlerin kaliteyi önemli ölçüde etkilediği gözlenmiştir. Bu örneklem genişliğinde sırasıyla hem 0,50 hem de 0,75 kantil regresyon analizinde kalite, üzerinde en yüksek etkinin -1,29 ve -1,53 değerleriyle Uçucu Asitlik ve en düşük etkinin ise -0,01 ve -0,02 değerleriyle Kalan Şeker tarafından yapıldığı ve bu etkilerin negatif olduğu gözlenmiştir. Ayrıca en geniş güven aralığı, 100 ve 200 örneklem genişliklerinde ve genellikle 0,75 kantilinde, 300 örneklem genişliğinde ise genellikle 0,50 kantilinde elde edilmiştir. Böylece kısmi regresyon katsayılarının büyümesinin güven aralığını artırdığı gözlenmiştir.

0,20, 0,25, 0,50 kantilleri için %5'ten küçük ve birbirine çok yakın R2 değerleri elde edilirken 0,75 kantilinde nispeten daha büyük değerler elde edilmiştir. Standart regresyonda ise en yüksek değerler elde edilmiş ve R2 değerleri maksimum %20'ye kadar çıkmıştır. Ayrıca örneklem genişliği arttıkça 0,75 kantili için bu değer küçüldüğü gözlenmiştir.

On bir bağımsız ve bir bağımlı değişkene sahip uygulamalarda tüm örneklem genişliklerine göre sonuçlar değerlendirildiğinde; hesaplanan sabit katsayılar birbirinden oldukça farklı bulunmuştur. Ayrıca regresyon uygulamaları için hesaplanan sabit değerlerin standart regresyonda kalite değişkeninin kısmi regresyon katsayısına eşit olduğu görülmüştür.

Standart regresyona göre; tüm örneklem genişliklerinde kalite değişkenine en yüksek etki Yoğunluk ve daha sonra Uçucu Asitlik tarafından gerçekleştirilmiştir. Yoğunluk 100 ve 300 örneklem genişliklerinde kaliteyi olumsuz etkilerken, 200 örneklem genişliğinde ise olumlu etkilemiştir. Uçucu Asitlik değişkeninin ise tüm örneklem hacimlerinde kalite üzerine olumsuz etkisi olmuştur. Ayrıca tüm örneklem genişliklerinde Sabit Asitlik, Serbest Kükürt, Sülfat ve Alkol kaliteyi olumlu ve Uçucu Asitlik, Sitrik Asit ve Klorit ise olumsuz etkilemiştir.

Kantil regresyon analizlerine göre; tüm örneklem genişliklerinde kalite üzerinde en yüksek etki sırasıyla Yoğunluk, Klorit ve Uçucu Asitlik tarafından gerçekleştirilmiştir. Kalite üzerinde 100 ve 300 örneklem genişliklerinde bu üç değişken olumsuz, 200 örneklem genişliğinde ise Yoğunluk olumlu iken diğer iki değişken olumsuz etkiye sahiptir. Ayrıca kaliteye en yüksek etkiyi; 100 ve 200 örneklem genişliklerinde hemen hemen tüm değişkenler 0,75 kantilinde gösterirken, 300 örneklem genişliğinde bazı değişkenler 0,50 ve bazıları da 0,75 kantilinde göstermiştir. Ayrıca 0,75 kantilinde örneklem genişliği arttıkça, genellikle değişken etkisinin azaldığı gözlenmiştir. 100 örneklem genişliğinde; 0,25 ve 0,50, 200 ve 300 örneklem genişliklerinde ise 0,20 ve 0,25 kantillerde değişkenlerin kaliteye etkilerinin çok düşük olduğu gözlenmiştir. En geniş güven aralığı, 100 ve 200 örneklem genişliklerinde genellikle 0,75 kantilinde, 300 örneklem genişliğinde ise genellikle 0,50 kantilinde elde edilmiştir. Ayrıca

kısmi regresyon katsayılarına göre bir genelleme yapılamazsa da örneklem genişliği arttıkça güven aralığının düştüğü gözlenmiştir.

Modele dahil edilen değişken sayısının artırılmasının, katsayılara çok fazla etkisinin olmadığı gözlenmiştir. Kantil değerlerin sonuçlar üzerindeki etkisinin değişken sayısından daha önemli olduğu belirlenmiştir. Bu nedenle de düşük katsayı ve dar güven aralığı elde etmek için uygun kantil değerinin tercih edilmesi önemlidir.

Örneklem genişliği 100 olduğunda 0,25 ve 0,50 kantilleri, 200 ve 300 olduğunda ise 0,20 ve 0,25 kantilleri kullanılabilir. Böylece örneklem büyüklüğü ne olursa olsun her durumda 0,25 kantilinde etkin sonuçların elde edilebileceği söylenebilir.

Teşekkür: Bu makale, Bahar Arsan Aysal'ın "Tarım ve hayvancılıkta doğrusal çıkarsama yöntemi (Kantil regresyon analizi) üzerine bir uygulama" başlıklı Yüksek Lisans tezinden özetlenmiştir.

Çıkar Çatışması Beyanı

Makale yazarları aralarında herhangi bir çıkar çatışması olmadığını beyan ederler.

Araştırmacıların Katkı Oranı Beyan Özeti

Yazarlar makaleye eşit oranda katkı sağlamış olduklarını beyan ederler

Kaynaklar

Alakaya D., 2019. Kantil regresyon ve doğrusal regresyon yöntemlerinin performansını etkileyen faktörlerin incelenmesi. Mersin Üniversitesi Sağlık Bilimleri Enstitüsü Biyoistatistik ve Tıbbi Bilişim Ana Bilim Dalı Yüksek Lisans Tezi, sayfa no:78, Mersin.

Berry WD., 1993. Understanding regression assumption. Sage Publications, London.

Chatterjee S, Simonoff JS., 2013. Handbook of regression analysis. John Wiley & Sons, New Jersey.

Çamurlu S, Erilli NA., 2019. Kantil regresyon analizinde bootstrap tahmini. Erciyes Üniversitesi Fen Bilimleri Enstitüsü Dergisi, 35(2): 16-25.

Çınar UK., 2019. En küçük kareler regresyonuna alternatif bir yöntem: kantil regresyon. Avrasya Uluslararası Araştırmalar Dergisi, 7(18): 57-71.

Demir Y., 2020. Çoklu doğrusal regresyon ve bazı cezalı tahmin yöntemlerinin incelenmesi. Sosyal ve Beşeri Bilimlerde Teori ve Araştırmalar II, Gece Akademi, Ankara, Ed: Öztürk S, 261-276.

Demir Y., 2022. Ürün ve hizmet ile ham tarım ürünleri ihracatının ekonomik büyümeye etkisinin doğrusal ve kantil regresyonla incelenmesi. İnsan ve Toplum Bilimleri Araştırmaları Dergisi, 11(4): 2311-2332.

Draper NR, Smith H., 1998. Applied regression analysis. John Wiley & Sons, New York.

Efron B, Tibshirani R., 1993. An introduction to the bootstrap. Chapman & Hall, USA.

Efron B., 1979. Bootstrap methods: another look at the jackknife. The Annals of Statistics, 7(1): 1-50.

Erickson J, Har-Peled S, Mount DM., 2006. On the least median square problem. Discrete & Computational Geometry, 36: 593-607.

Gayaker S., 2015. Durağan olmayan var sistemlerinde bootstrap yöntemi ile granger nedensellik sınaması. Gazi Üniversitesi Sosyal Bilimler Enstitüsü, Ekonometri Ana Bilim Dalı Yüksek Lisans Tezi, sayfa no:66, Ankara.

Gujarati DN., 2004. Temel ekonometri. Literatür Yayıncılık, İstanbul.

İşleyen Ş, Altun Y, Görür Ç., 2018. Hizmet ihracatı ve turizm harcamalarının ekonomik büyümeye katkısı üzerine ampirik bir analiz: 1996-2017 Türkiye örneği. Anemon Muş Alparslan Üniversitesi Sosyal Bilimler Dergisi, 6(6): 953-960.

Keskin B., 2012. Sağlam bir çıkarsama yöntemi: kantil regresyon. Muğla Sıtkı Koçman Üniversitesi Fen Bilimleri Enstitüsü İstatistik Ana Bilim Dalı Yüksek Lisans Tezi, sayfa no:99, Muğla.

Kiren Gürler Ö, Üçdoğruk Birecikli Ş, Kökçen Eryavuz A., 2018. Türkiye’de hanehalkı tüketim ve gıda harcamalarının kantil regresyon yöntemiyle araştırılması. Uluslararası İktisadi ve İdari İncelemeler Dergisi, 18(EYİ Özel Sayısı): 219-238.

Koenker R, Bassett G., 1978. Regression quantiles. The Econometric Society, 46(1): 33-50.

Kuan CM, Lin HY., 2010. An encompassing test for non-nested quantile regression models. Economics Letters, 107(2): 257-260.

Kuşkaya S., 2023. Yenilenebilir enerji ve ekonomik büyüme ilişkisinin kantil regresyon ile modellenmesi: ABD örneği. İstanbul Gelişim Üniversitesi Sosyal Bilimler Dergisi, 10(1): 234-245.

Okutan D., 2009. Bootstrap yönteminin regresyon analizinde kullanımı ve diğer yöntemlerle karşılaştırılması. Ege Üniversitesi Fen Bilimleri Enstitüsü Uygulamalı İstatistik Ana Bilim Dalı Yüksek Lisans Tezi, sayfa no:55, İzmir.

Orhunbilge N., 2000. Uygulamalı regresyon ve korelasyon analizi. Nobel Akademik Yayıncılık, İstanbul.

Özçelik R, Alkan O, Kalkanlı Ş., 2021. Kantil regresyon ile gövde çapı modelinin geliştirilmesi. Türkiye Ormancılık Dergisi, 22(3): 250-256.

Özen K, İşleyen Ş, Görür Ç., 2023. Investigation of the relationship between the use of selected energy types and agricultural area utilised by ardl bound test: the case of Turkey. Van Yüzüncü Yıl Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi, 8(16): 354-372.

Özyıldırım Y., 2019. Finansal kapsayıcılık ile ekonomik büyüme arasındaki ilişkinin kantil regresyon yöntemiyle analizi. Uşak Üniversitesi Sosyal Bilimler Enstitüsü Ekonometri Ana Bilim Dalı Yüksek Lisans Tezi, sayfa no:124, Uşak.

Pearson, K. 1903. Mathematical contributions to the theory of evolution: II. On the influence of natural selection on the variability and correlation of organs. Royal Society Philosophical Transactions, 200(Series A):1-66.

Pearson K, Lee A., 1903. On the laws of inheritance. Biometrika, 2(1903): 357-462.

Rawlings JO, Pantula SG, Dickey DA., 1998. Applied regression analysis: a research tool Second Edition. Springer, New York.

Rousseeuw PJ, Leroy AM., 1987. Robust regression and outlier detection. John Wiley & Sons, New York.

Schaeck K., 2008. Bank liability structure, FDIC loss, and time to failure: a quantile regression approach. Journal of Financial Services Research, 33: 163-179.

Taddy M, Kottas A., 2010 A bayesian nonparametric approach to inference for quantile regression. Journal of Business & Economics Statistics, 28(3): 357-369.