



Bezostaja buğdayının fiyatına etki eden bazı donelerin alternatif regresyon yöntemleriyle karşılaştırılması

Fatma ÇİFTÇİ

Aşır GENÇ

Önerilen atf: Çiftci F & Genç A. (2020). Bezostaja buğdayının fiyatına etki eden bazı donelerin alternatif regresyon yöntemleriyle karşılaştırılması, *Uluslararası Sosyal Bilimler Eğitimi Dergisi*, 6(2), 215-237. DOI: 10.47615/issej.837784

Makale linki: <https://doi.org/10.47615/issej.837784>



© 2020 Yazar(lar). Baskılar ve izinler:
USBED/Uluslararası Sosyal Bilimler Eğitimi
Dergisi

USBED/Uluslararası Sosyal Bilimler Eğitimi Dergisi hakemli bir çevrimiçi dergidir. Bu makale araştırma, öğretim ve özel çalışmalar amacıyla kullanılabilir. Makalenin içeriğinden yalnızca yazarlar sorumludur Dergi makalelerin telif hakkına sahiptir. Yayıncı, araştırma materyalinin kullanımıyla bağlantılı veya doğrudan veya dolaylı olarak ortaya çıkan herhangi bir kayıp, işlem, talep veya masraf veya zarardan sorumlu tutulamaz.

Bezostaja Buğdayının Fiyatına Etki Eden Bazı Donelerin Alternatif Regresyon Yöntemleriyle Karşılaştırılması

Fatma ÇİFTÇİ 

PTT Başmüdürlüğü, Konya, 42080, Türkiye

Aşır GENÇ 

Necmettin Erbakan Üniversitesi, Konya, 42080, Türkiye

ÖZ

Regresyon analizi iki ya da daha fazla değişken arasındaki sebep sonuç ilişkisini inceleyen, bu konu ile ilgili tahminler yapabilmek için matematiksel model kuran tekniktir. Bu çalışmada, fiyata etki eden değişkenler ile fiyat arasındaki doğrusal bağıntı incelenmiştir. 12 aylık Bezostaja buğday çeşidi verisini modelleyen çoklu regresyon analizi ile parametre tahmin yöntemleri karşılaştırılmaya çalışılmıştır. Model parametrelerinin tahmin edicilerinden; En küçük kareler yöntemi (EKK), En küçük mutlak sapma regresyon (LAD) yöntemi, M-Regresyon yöntemi, Parametrik olmayan regresyon yöntemi, Bayes regresyon yöntemi ele alınmıştır. Çoklu lineer regresyon modeline mevcut verilere söz konusu yöntemler uygulanarak tahminler yapılmış ve sonuçlar karşılaştırılarak uygun tahmin yöntemi önerilmiştir. Bu çalışmada kullanılan veri seti için Parametrik Olmayan Regresyon tahmin yöntemi uygun bir tahmin edici olarak görülse de HKO ve AIC kriterine göre EKK tahmin yöntemi bu veri seti için en uygun yöntemi olarak öne çıkmaktadır.

MAKALE TARİHİ

Geliş tarihi 08 Aralık 2020

Kabul tarihi 09 Aralık 2020

Yayın tarihi 31 Aralık 2020

ANAHTAR KELİMER

çoklu regresyon, buğday, alternatif tahmin yöntemleri

Makale Türü

Araştırma makalesi

İLETİŞİM Fatma ÇİFTÇİ  fatma_ifc@hotmail.com  PTT Başmüdürlüğü, 42080, Konya, Türkiye

Fatma ÇİFTÇİ, Aşır GENÇ © 2020 The Author(s).

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

The Comparison of Some Datum that Affects the Price of Bezostaja Wheat with Alternative Regression Methods

Fatma ÇİFTÇİ 

General Directorate of Turkish Post (PTT), Konya, 42080, Turkey

Aşır GENÇ 

Necmettin Erbakan University, Konya, 42080, Turkey

ABSTRACT

Regression analysis is a technique which examines the reason and result between two or more variables and creates a mathematical model in order to make a prediction on the subject. In this study, the linear relationship between the variables that affect the price and the price itself has been examined. Multiple regression analysis modeling for 12 months Bezostaja wheat varieties data and parameter estimation methods were tried to be compared. Estimators of model parameters; Least squares method (LS), Smallest absolute deviation regression method (LAD), non-parametric regression method, M-regression method and Bayes regression method. In the multiple linear regression model, estimates were made by applying these methods to the available data, and the appropriate method was proposed by comparing the results. Under these circumstances, although non-parametric regression estimation method seems to be a convenient estimator, according to the MSE and AIC criteria, LS estimation method steps forward as the most convenient method for this data set.

ARTICLE HISTORY

Received 08 December 2020

Accepted 09 December 2020

KEYWORDS

multiple regressions, wheat, alternative methods

Type of the Paper

Research article

CONTACT Fatma ÇİFTÇİ  fatma_ifc@hotmail.com  General Directorate of Turkish Post (PTT), 42080 Konya, Turkey.

Fatma ÇİFTÇİ, Aşır GENÇ © 2020 The Author(s).

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way

Giriş

Buğday üretiminde üretim masrafları (tohum, sulama, ilaçlama, toprak işleme vb.) yaparak elde ettikleri ürünün fiyatını etkileyen bazı faktörler bulunmaktadır. Buğday fiyatını etkileyen unsurlar Toprak Mahsulleri Ofisi (TMO) tarafından; hektolitre ağırlığı, kırık dane oranı, kusurlu dane oranı, süne/kıvımlı tahribatına uğramış dane oranı, çimlenmiş/filizlenmiş dane oranı, diğer muhtelif madde oranı, dönmeli dane oranı, protein oranı ve rutubet oranı şeklinde belirlenmiştir (Anonim 2010a).

İstatistiksel analizlerin içinde çok kullanılan regresyon analizi tepki (bağımlı) değişkeni denen bir değişken ile açıklayıcı (bağımsız) değişkenler arasındaki bağıntının belirlenmesinde ve bu bağıntı yardımıyla çıkarılacak istatistiksel sonuçların elde edilmesinde kullanılan yöntemlerden oluşmaktadır. Regresyon analizi yapılmasındaki amaç; tepki değişkenini açıklayıcı değişkenlerin bir fonksiyonu olarak ifade etmek ve bu fonksiyon yardımıyla tepki değişkeninin değerlerini tahmin etmek, öngörmek, açıklayıcı değişkenlerin tepki değişkeni üzerindeki etkilerini tahmin etmek, tepki veya açıklayıcı değişkenlerin etkileri ile ilgili öne sürülen hipotezleri test etmek olabilir. Fiyat oluşumunda etkili olan hektolitre ağırlığı, kırık dane oranı, kusurlu dane oranı, süne/kıvımlı, rutubet oranı gibi çok sayıda faktöre bağlı olarak sebep-sonuç bağıntısı ortaya konulabilir. Bu bağıntı çoklu regresyon analizi ile istatistiksel olarak araştırılabilir. Fiyat, bağımlı değişken olup diğer faktörler bağımsız değişken olarak ele alınmıştır. Çoklu regresyon yönteminde bağımlı değişkenin gerçek ölçüleri ile bağımsız değişkenlerden elde edilen kestirim ölçüleri arasındaki uzaklığı en küçük yapan regresyon katsayılarının tahmini En Küçük Kareler kestirimi (EKK) ile hesaplanmaktadır (Akdeniz ve Öztürk 1996, Karadavut, 2005).

Regresyon analizinde elde edilen denklem, parametre tahmini için kullanılmak isteniyorsa; hata terimlerinin beklenen değerlerinin ortalamasının 0 (sıfır) olması, varyansların homojen olup σ^2 'ye eşit olması, hata terimleri ile açıklayıcı değişkenler arasında korelasyon bulunmaması gibi bazı varsayımların sağlanması gerekmektedir (Alma ve Vupa, 2008). Ancak hata terimlerinin dağılımı normal dağılım olduğu zaman EKK yöntemi en çok olabilirlik yöntemiyle (MLE) aynı olmaktadır.

Hataların dağılımı normal dağılım olmadığı durumlarda, gözlenen verilerde aykırı gözlemler ya da uç noktalar veya bozulma noktaları olması durumunda EKK yöntemi ile bulunan tahminlerde varsayım bozulmalarıyla karşı karşıya kalınabilmektedir. Bu nedenle buğday fiyatını etki eden faktörlere ait verilere lineer regresyonda En Küçük Kareler yöntemine alternatif olan Lad Regresyon, Parametrik Olmayan Regresyon, Bayes Regresyon, Robust Regresyon ve benzeri tahmin yöntemleri uygulanabilir. Bu çalışmada, En Küçük Kareler yöntemine alternatif olan LAD Regresyon, M-regresyon, Parametrik Olmayan Regresyon ve Bayes Regresyon yöntemlerinden en uygun yöntem bulunmaya çalışılmıştır. Tahmin yöntemlerinin seçiminde Çoklu belirtme katsayısı (R^2), Akaike bilgi kriteri (AIC) ve Hata kareler ortalaması (HKO) dikkate alınmıştır.

Parametrelere göre lineer olan çoklu lineer regresyon modelinde, Y_i ; bağımlı değişken, X_{ij} ; bağımsız değişkenler, β_j ; j . regresyon katsayısı ve e_i ; i . hata terimi ve model fonksiyonu,

$$Y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik} + \varepsilon_i, i = 1, 2, \dots, n \quad (1)$$

regresyon denklemi ile ilgili;

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \dots \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{21} & \dots & x_{1k} \\ 1 & x_{12} & x_{22} & \dots & x_{2k} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_{1k} & x_{2k} & \dots & x_{nk} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \dots \\ \beta_k \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \dots \\ \varepsilon_n \end{bmatrix} \quad (2)$$

gösterimleri altında regresyon denklemi;

$$\underline{Y} = X \underline{\beta} + \underline{\varepsilon} \quad (3)$$

biçiminde yazılabilir. Hata terimi ε nun dağılımına bağlı olarak ε_i ler normal dağılımlı olduğunda modellere "hipotez modeli" denir.

Amaç

Bu araştırmanın amacı bezostaja buğdayının üretiminde üretim masrafları yaparak elde ettikleri ürünün fiyatını etkileyen bazı faktörlerin, buğdayın fiyatına olan etkileri belirlenmek istenmiştir.

Bu amaç için şu sorulara cevaplar verilecektir.

- 1- Buğday fiyatına hektolitre ağırlığının etkisi var mıdır?
- 2- Buğday fiyatına süne tahribatının etkisi var mıdır?
- 3- Buğday fiyatına rutubetin etkisi var mıdır?

Yöntem

Bezostaja Buğday çeşidinin kullanıldığı bu çalışma Konya Ticaret Borsasından alınan 2004 yılına ait 12 aylık veri incelenmiştir. Fiyatı diğer değişkenlerin (süne, hektolitre ve rutubet gibi) ne şekilde etki ettiğini doğru bir şekilde belirleyebilmek için, EKK, LAD, M-Regresyon, Parametrik Olmayan Regresyon ve Bayes Regresyon tahmin yöntemleri uygulanarak en uygun tahmin yöntemi tercih edilmek istenmiştir. Eşitlik (1) ile belirtilen lineer regresyon modelinde aykırı değerler olduğunda EKK tahmin yönteminin bu aykırı değerlerden etkilendiğini ileri sürerek LAD regresyon tahmin yöntemi elde edilmiştir. Ayrıca Huber 1964 yılında aykırı değerlerin etkisini azaltmak amacıyla M regresyon tahmin yöntemini geliştirmiştir.

En Küçük Kareler Regresyon Yöntemi

Lineer regresyon modelinin matris gösterimiyle,

$$\underline{Y}_{n \times 1} = X_{n \times p} \underline{\beta}_{n \times 1} + \underline{\varepsilon} \quad (4)$$

$\underline{\varepsilon}$ ile ilgili klasik varsayımlar $\underline{\varepsilon} \rightarrow N(0, \sigma^2)$ modelinde $\underline{\beta} \in \mathfrak{R}^p$ parametresinin tahmin edilmesi problemi ele alınsın. Yöntem olarak en küçük kareler yöntemi düşünülürse,

$$\min_{\underline{\beta}} (\underline{Y} - X \underline{\beta})' (\underline{Y} - X \underline{\beta}) = \min_{\underline{\beta}} (\underline{Y}' \underline{Y} - 2 \underline{\beta}' X' \underline{Y} + \underline{\beta}' X' X \underline{\beta}) \quad (5)$$

optimizasyon problemini çözmek için $\underline{\beta}$ ya göre türev alınıp sıfıra eşitlendiğinde,

$$X' X \hat{\underline{\beta}} = X' \underline{Y} \quad (6)$$

normal denklemlerine ulaşılır. X tam sütun ranklı yani,

$$\text{rank}(X) = \text{rank}(X' X) = p \quad (7)$$

olduğundan,

$$\hat{\underline{\beta}} = (X' X)^{-1} X' \underline{Y} \quad (8)$$

çözümü bulunur (Birkes ve Dodge, 1993). \underline{Y} örneklem olmak üzere \underline{Y} nin bir fonksiyonu olan $\hat{\underline{\beta}}$, $\underline{\beta}$ için bir tahmin edici olmak üzere bu tahmin ediciye en küçük kareler tahmin edicisi denir.

Lad Regresyon

Mutlak en küçük sapmalar (LAD) regresyonu için eşitlik (1) deki lineer regresyon modeli göz önüne alındığında,

$$\sum |Y_i - \beta_0 + \beta_1 X_{i1} + \dots + \beta_n X_{in}| \quad (9)$$

şeklinde elde edilir. Bir noktadan geçen en iyi doğru algoritma üretilerek bulunur. Algoritma için vektör gösterimi;

$$\hat{\beta} = \begin{bmatrix} \hat{\beta}_0 \\ \vdots \\ \hat{\beta}_k \end{bmatrix}, \quad X_i = \begin{bmatrix} X_{i1} \\ \vdots \\ X_{ik} \end{bmatrix}$$

biçimindedir. Eşitlik (9), $\sum |Y_i - \hat{\beta}' X_i|$ şeklinde elde edilir. Burada amaç bu değeri minimize eden $\hat{\beta}'$ vektörünü bulmaktır. Algoritmaya $\hat{\beta}'$ vektörüyle başlanılır. Sonra $\sum |Y_i - \hat{\beta}' X_i|$ 'nin minimum değeri veren daha iyi bir vektör bulunarak, $\hat{\beta}'$ 'nin en iyi vektörü hesaplanır. Her adımda $\hat{\beta}$ tahminleri vektöründen daha iyi bir $\hat{\beta}^*$ vektörü,

$$\hat{\beta}^* = \hat{\beta} + td \quad (10)$$

biçiminde elde edilir. Eşitlik (10) daki vektörün bulunabilmesi için yön vektörü d ve t değerlerinin hesaplanması gerekmektedir. t değerini bulmak için eşitlik (11) minimize edilir.

$$\sum \left| Y_i - (\hat{\beta} + td)' X_i \right| \quad (11)$$

Burada; $Z_i = Y_i - \hat{\beta}' X_i$, $W_i = d' X_i$ dönüşümü yapılarak,

$$|Z_i - tW_i|$$

elde edilir (Birkes ve Dodge, 1993). Elde edilen bu değer $\sum |(Y_i - Y_0) - \hat{\beta}'(X_i - X_0)|$, u minimize edecek $\hat{\beta}'$ yı bulmak içindir. Daha sonra Z_i/W_i oranları hesaplanarak, küçükten büyüğe sıralanır. Z ve W bu sıraya göre sıralanarak k indeksi hesaplanır.

$$|W_1| + \dots + |W_{k-1}| < \frac{1}{2} T$$

$$|W_1| + \dots + |W_k| > \frac{1}{2} T$$

Burada, $T = \sum |W_i|$, dir. t , yı minimum yapan değer Z_k/W_k , dir. Algoritmanın her bir adımında p bağımsız değişken sayısından bir fazla $(p+1)$ yön vektörü vardır. Her bir d_j vektörü için (+) pozitif yön olduğu gibi (-) negatif yönde mevcuttur. Bu nedenle bağımsız değişken sayısının bir fazlasının iki katı sayıda yön vektörü olacaktır. $\sum |Y_i - (\hat{\beta} + td)' X_i|$ değerinin mümkün olduğunca $t=0$ değerine yaklaştıran yön vektörü seçilir.

M-Regresyon Regresyon

M – Regresyon için çoklu regresyon modelindeki Huber M tahminleri;

$$\sum \rho\left(Y_i - \left(\hat{\beta}_0 - \hat{\beta}_1 X_{i1} + \dots + \hat{\beta}_n X_{in}\right)\right) \quad (12)$$

değerini minimize eden değerler olup

$$\rho(e) = \begin{cases} -k \leq e \leq k & e^2 \\ e < -k \text{ ya da } k < e & 2k|e| - k^2 \end{cases} \quad (13)$$

biçiminde elde edilir. Huber M-regresyon tahmini $\hat{\beta}$ vektörü, $\sum \rho\left(Y_i - \left(\hat{\beta}_0 - \hat{\beta}' X_i\right)\right)$, yi minimize eden $\hat{\beta}$ vektörü olarak tanımlanır (Birkes ve Dodge, 1993). Regresyon katsayısı vektörü $\hat{\beta}$ ilk olarak EKK ile tahmin edilir. β 'nin ilk tahminlerine ulaşılan kadar algoritma bu şekilde devam edilir.

Parametrik Olmayan Regresyon

Parametrik olmayan regresyon tahmin yöntemi aykırı gözlemlerin bulunduğu veri setleri için güçlü bir analiz yöntemidir. Aykırı gözlemlerden dolayı parametreler bozulduğu için birçok yöntem en uygun sonucu vermemektedir. Bu nedenle parametrik olmayan regresyon tahmin yöntemine ihtiyaç duyulmaktadır (Hardle, 1990).

$Y_i - \left(\hat{\beta}_1 X_{i1} + \dots + \hat{\beta}_n X_{in}\right)$ farkının medyanı $\hat{\beta}_0$ 'ın parametrik olmayan tahminini belirler. Eşitlik (14) deki değer minimize edilerek bir algoritma üretilir. Vektör gösterimi,

$$g(b) = \sum \left[\text{rank}(Y_i - b'X_i) - \frac{n+1}{2} \right] (Y_i - b'X_i) \quad (14)$$

olup burada, $b = (b_1, \dots, b_k)$, $X_i = (X_{i1}, \dots, X_{ik})$ dir (Birkes ve Dodge, 1993).

$\hat{\beta}$ vektörü, $g(b)$, yi minimize ederek elde edildiğinden EKK yöntemi ile iterasyona başlanır. $g(b)$ fonksiyonun daha küçük değerlerini veren vektörler bulunur. Başlangıç vektörü $\hat{\beta}^0$ bulunur, $\hat{\beta}^* = \hat{\beta}^0 + t^* d$ elde edilir. Burada t^* ve d , $Z_i = Y_i - (\hat{\beta}^0)' X_i$, $u_i^0 = \text{rank}(Z_i) - \frac{1}{2}(n+1)$, $u^0 = u_i^0$ ile $n \times 1$ vektördür. $X_c = X_{ij} - \bar{X}_j$, lardan oluşan $n \times p$, lık matristir. $d = (X_c' X_c)^{-1} X_c' u^0$, $w_i = d' X_i$, t^* , da $\frac{|w_i - w_j|}{\sum |w_i - w_j|}$ ağırlıkları ile $\frac{|z_i - z_j|}{|w_i - w_j|}$ oranlarının ağırlıklı medyanıdır. Bu şekilde iterasyon

tamamlanarak $\hat{\beta}^*$ geçerli vektör olur ve iterasyon tekrarlanır (Birkes ve Dodge, 1993).

Bayes Regresyon

Y bağımlı değişken vektörü ve θ modelin bilinmeyen parametre vektörü olsun. Bayes yaklaşımında θ için dağılım belirtilmemekte ve veri setine bakmadan önce parametre hakkında ne bilindiğine bakılmaktadır. Böylelikle ilk olarak veri seti vektörü için olasılık dağılımı belirlenir. Genellikle kullanılan dağılım normal dağılımdır. Çoklu lineer Bayes Regresyon modeli,

$$f(\beta|y, \sigma) = cf(\beta|\sigma)f(y|\beta, \sigma) \quad (15)$$

olup burada c sabit değer, $f(y|\beta, \sigma)$, y 'nin bir fonksiyonudur. Çoklu normal dağılımın olasılık yoğunluk fonksiyonunun ortalama vektörü $X\beta$ ve varyans-covaryans matrisi $\sigma^2 I$ (I birim matristir.) olup,

$$f(y|\beta, \sigma) = c \exp\left[-\frac{1}{2\sigma^2}\|y - X\beta\|^2\right] \quad (16)$$

$$\|y - X\beta\|^2 = \|y - X\hat{\beta}_{LS}\|^2 + \|X\hat{\beta}_{LS} - X\beta\|^2 \quad (\text{Pisagor teoremi})$$

$$f(y|\beta, \sigma) = c \exp\left[-\frac{1}{2\sigma^2}\|X\hat{\beta}_{LS} - X\beta\|^2\right]$$

Her bir β 'nin sonsal dağılımı EKK tahmini $\hat{\beta}_{EKK}$ dir.

$$f(\beta|\sigma) = c \exp\left[-\frac{1}{2\sigma^2}(\beta - b)'V^{-1}(\beta - b)\right] \quad (17)$$

b vektörünün varyans - covaryans matrisi $\sigma^2 V$ dir. Önsel bilgiyi elde etmek için b ve V seçilerek $\hat{\beta}$,

$$\hat{\beta} = V_* V^{-1} b + V_* X'X \hat{\beta}_{EKK} \quad (18)$$

biçiminde elde edilir. Burada, $V_* = (V^{-1} + X'X)^{-1}$ dir (Birkes ve Dodge, 1993).

Seçme Kriterleri

Uygun yöntemin seçiminde Hata Kareler Ortalaması (HKO), Akaike Bilgi Kriteri (AIC) ve Çoklu belirtme katsayısından (R^2) yararlanılacak olup en büyük R^2 değeri, en küçük AIC ve HKO değerlerini veren yöntem tercih edilecektir.

Çoklu Belirtme Katsayısı

Çoklu belirtme katsayısı (R^2), X bağımsız değişkenlerinin Y bağımlı değişkenindeki toplam varyasyonu açıklayabilme oran ve ($0 < R^2 < 1$) olup,

$$R^2 = \frac{b'X'Y - (\sum Y)^2/n}{\sum Y^2 - (\sum Y)^2/n} \quad (19)$$

(Newton ve ark 1996). Ancak, R^2 , nin 1'e yakın çıkması modelin iyi olduğu anlamına gelmez. Modele herhangi bir parametre eklendiğinde ya da çıkartıldığında R^2 , nin değerini değiştirecektir. Bu nedenle R^2 , si büyük olan modeller her zaman en iyi modeller olmayabilir (Montgomery ve Peck, 1992), (Şahin, 2000).

Akaike Bilgi Kriteri (AIC)

Akaike bilgi kriteri (AIC), farklı modeller arasında en uygun modeli seçmek için kullanılmakta olup,

$$AIC = \ln \left(\frac{\sum e_i^2}{n} \right) + \frac{2k}{n} \quad (20)$$

Eşitlik (20)' de k sabit terim dahil parametre sayısı ve n gözlem sayısını, $\sum e_i^2$ ise hata kareler toplamının vermektedir. Model karşılaştırılmalarında en küçük AIC değerini veren model tercih edilmektedir.

Hata Kareler Ortalaması Kriteri (HKO)

En uygun modeli seçmede kullanılan bir diğer kriter de Hata kareler ortalaması kriteridir. Hata kareler ortalaması (HKO) ne kadar küçük olursa, beklenen değerle gözlenen değer arasındaki sapma o kadar küçük olmaktadır.

$$HKO = \frac{1}{n} \sum (Y_i - \hat{Y}_i)^2 \quad (21)$$

formülü ile hesaplanır (Ekinci ve ark., 2016).

Bulgular

Bu çalışmada, buğdayın fiyatının oluşumunda Toprak Mahsulleri Ofisi (TMO) tarafından belirlenen kriterler arasında süne, hektolitre ve rutubet gibi faktörler baz alınarak uygulama yapılmıştır. Çalışmada kullanılan veri, Konya Ticaret Borsasından Konya bölgesi 2004 yılına ait Bezostaja buğdayı fiyat oluşumunda etkili olan hektolitre ağırlığı, süne/kıymı, rutubet oranı değerlerini içerir.

Tablo 1. Konya bölgesindeki Bezostaja buğdayı fiyatı (Y), hektolitire (X_1), süne (X_2) ve rutubet (X_3)

Fiyat	Hektolitire	Süne	Rutubet
4.000	79.1	3.7	8.1
3.880	78.1	4.7	8.9
3.700	78.6	6.4	12.6
4.000	79.2	3.1	10.1
3.700	77.5	6.0	9.0
3.860	76.8	1.5	8.9
3.770	80.4	3.4	9.9
3.780	72.5	1.5	10.8
4.070	78.6	1.5	11.8
3.860	72.1	1.8	9.2
4.040	81.1	3.5	10.1
3.900	77.8	3.8	10.7
3.980	75.2	0.9	9.6
4.000	77.8	1.5	10.1
3.750	79.0	1.3	10.9
3.970	79.3	2.5	10.6
3.600	71.4	1.9	11.1
3.490	74.2	0.6	9.9
3.870	77.7	4.4	9.0
3.850	77.8	2.8	11.8

Tablo 1.'deki veri seti kullanılarak,

$$\text{Fiyat} = \beta_0 + \beta_1 \times \text{Hektolitire} + \beta_2 \times \text{Süne} + \beta_3 \times \text{Rutubet} + \varepsilon$$

regresyon modeli elde edilmiştir.

EKK tahmin yöntemine alternatif olarak LAD, M-Regresyon, Parametrik Olmayan Regresyon, Bayes Regresyon tahmin yöntemleri SPSS ve Excel paket programları kullanılarak modelin Çoklu belirtme katsayısı (R^2), Akaike bilgi kriteri (AIC) ve Hata kareler ortalaması (HKO) elde edilmiş olup veri setine en uygun model seçilmek istenmiştir. Yapılan hesaplamalar Tablo 2. de özetlenmiştir.

Tablo 2. Bezostaja buğday verisi için elde edilen Alternatif Regresyon Modelleri

Tahmin Yöntemi	Regresyon Modeli	R^2	AIC	HKO
EKK	$\hat{Y} = 1432 + 35.4X_1 - 29.3X_2 - 22.3X_3$	0.344	9.99	14.634,76
LAD	$\hat{Y} = 1220 + 27.36X_1 - 69.18X_2 - 22.12X_3$	0.192	14.11	907.998,28
M-REGRESYON	$\hat{Y} = 1259.576 + 39.001X_1 - 48.74X_2 - 24.499X_3$	0.502	10.01	16.342,28

PARAMETRİK OLMAYAN	$\hat{Y}=1259.576+32.95X_1-36.51X_2-5.41X_3$	0.697	12.13	125.338,2
BAYES	$\hat{Y}=1436+34.81X_1-44.62X_2-13.750X_3$	0.359	10.04	15.395,76

Tablo 2. deki sonuçlar ele alındığından EKK tahmin yönteminin HKO değeri 14.634,76 olarak elde edilmiş olup diğer yöntemlere göre daha küçük bulunmuştur. Benzer biçimde EKK tahmin yönteminin AIC değeri 9.99 olarak elde edilmiş olup diğer yöntemlere göre daha küçük bulunmuştur. Ancak Parametrik Olmayan Regresyon tahmin yönteminin R^2 değeri 0.697 olup diğer yöntemlerin R^2 değerinden daha büyük çıkmıştır.

Tartışma ve Sonuç

Çalışmada EKK tahmin yöntemine alternatif bazı regresyon modelleri verilmeye çalışılmıştır. Burada LAD, M-regresyon, Parametrik olmayan regresyon ve Bayes regresyon tahmin yöntemleri incelenmiştir. İncelenen örneğe ait belirtme katsayısı en yüksek Parametrik olmayan regresyonda $R^2 = 0.697$ olarak bulunmuştur. Kullanılan bağımsız değişkenlerin, bağımlı değişken (Y)'deki toplam varyasyonu açıklayabilme oranı nispi olarak 0.697 tespit edilmiştir. Kalan 0.303'lük kısım çalışılmayan diğer faktörler nedeniyle oluşmuştur. R^2 'nin yüksek çıkmaması, gözlemler içerisinde aykırı gözlemler olabileceğini ifade etmektedir. Bütün bu açıklamaların ışığında Parametrik Olmayan Regresyon tahmin yöntemi uygun bir tahmin edici olarak görülse de HKO ve AIC kriterine göre EKK tahmin yöntemi bu veri seti için en uygun yöntem öne çıkmaktadır.

Öneriler

Parametrik Olmayan Regresyon tahmin yöntemi uygun bir tahmin edici olarak görülse de HKO ve AIC kriterine göre EKK tahmin yöntemi bu veri seti için en uygun yöntem öne çıkmaktadır.

Bu tür analizlerde R^2 'nin yüksek çıkmaması, gözlemler içerisinde aykırı gözlemler olabileceğini ifade etmektedir.

Bu tür araştırmalarda etki faktörlerinin sınırları genişletilerek daha fazla il verisi ile sınır genişletilebilir.

Araştırmanın Sınırlılıkları ve Gelecek Araştırmalar

Yapılan araştırma,

- Araştırma, Konya ilinde bulunan Ticaret Borsası 2004 verileri ile sınırlıdır.
- Çalışma grubunu 2004 yılı 12 aylık buğday verileri ile sınırlıdır.
- Verilerin karşılaştırılmasında fiyatı etkileyen unsurlar fiziksel unsurlar ile sınırlıdır.

Bu sınırlılıklar göz önünde bulundurularak gelecek araştırmalarda farklı illerde veriler farklı karşılaştırma türleri ile uygulamalar yapılması ve farklı ürün çeşitlerinde uygulama yapılması önerilebilir.

Yazar Katkıları

Bu çalışma yüksek lisans tez çalışmasından üretilmiştir. Çalışma yazarları yüksek lisans öğrencisi ve danışmanını içermektedir. Çalışmanın her aşamasında her iki yazar katkı sağlamıştır. Tüm yazarlar makalenin yayınlanmış halini okuyup kabul etmişlerdir.

Yayın Etiği

Araştırma ve Yayın Etiği'ne uyuldu

ORCID

Fatma ÇİFTÇİ  <http://orcid.org/0000-0001-6362-1777>

Aşır GENÇ  <http://orcid.org/0000-0002-0339-6050>

Kaynakça

- Akaike, H. (1970). "Statistical Predictor Identification", *Annals of the Institute of Statistical Mathematics*, 22, pp. 203-217.
- Akdeniz, F., ve Öztürk, F. (1996). *Lineer Modeller*, A.Ü.F.F. Döner Sermaye İşletmesi Yayınları, No:38, Ankara.
- Alma, Ö.G., ve Vupa, Ö. (2008). Regresyon Analizinde Kullanılan En Küçük Kareler ve En Küçük Medyan Kareler Yöntemlerinin Karşılaştırılması. *Sdü Fen Edebiyat Fakültesi Fen Dergisi (E-Dergi)*. 3(2) 219-229.
- Anonim (2010a). *Hububat alım ve satış esaslarına ilişkin uygulama yönetmeliği*, TMO.
- Birkes, D., ve Dodge Y. (1993). *Alternative Methods Of Regression*. USA. A Wiley Interscience Publication. John Wiley & Sons, INC.
- Ekinci, E.B.M., Alhan A., Ergör Z.B. (2016). Parametrik Olmayan Regresyon Analizi: Faiz oranı, Enflasyon ve Döviz Kuru Arasındaki İlişkinin İncelenmesi Örneği. *Bankacılık ve Sigortacılık Araştırmaları Dergisi Cilt: 2, Sayı: 9, Ankara*
- Hardle, W. (1990). *Applied Nonparametric Regression*, Cambridge University Press, s. 4-5.
- Huber, P. (1964). Robust estimation of a location parameter. *Annals of Mathmatical Statistics*, vol.35, pp.73-101.
- Jabr, R. (2005). "Power system Huber M-estimation with equality and inequality constraints," *Electric Power System Research*, 74, 239-246.
- Karadavut, U., Genç, A., Tozluca, A., Kınacı, İ., Aksoyak, Ş., Palta, Ç., Pekgör, A. (2005). Nohut (*Cicer arietinum L.*) Bitkisinde Verime Etki Eden Bazı Karakterlerin Alternatif Regresyon Yöntemleriyle Karşılaştırılması. *Tarım Bilimleri Dergisi* 2005, 11 (3) 328-333.
- Montgomery, D. C. and E. A. Peck. (1992). *Intoduction to Lineer Regression Analysis*, 2nd ed., John Wiley and Sons, New York.
- Newton, H.J., Carroll, J.H., Wang, N., Whiting, D. (1996). *Statistics 30x Class Notes*. University of Leeds. U.K.
- Şahinler, S. (2000). En Küçük Kareler Yöntemi ile Doğrusal Regresyon Modeli Oluşturmanın Temel Prensipleri. *MKÜ Ziraat Fakültesi Dergisi* 5 (1-2): 5773.

Extended Abstract

Introduction

In wheat production, there have been some factors that affect the price of the production by making manufacturing costs such as seed, irrigation, spraying, tillage etc. The factors affecting wheat price have been determined as hectoliter weight, broken sweet rate, defective grain rate, sunn pest and insect damaged grain rate, germinated/sprouted grain rate, other various item rates, rotary grain rate, protein rate and humidity rate by Turkish Grain Board (abbreviation for TMO) (Anonymous 2010a).

Regression analysis, widely used in statistical analysis, consists of methods used in determining the correlation between a variable called reaction (dependent) variable and explanatory (independent) variables and obtaining statistical results with the help of this relation. The purpose of regression may be to express the response variable as a function of explanatory variables and to guess and foresee the values of reaction variable with the help of this function, estimating the effects of explanatory variables on the response variable, to test the hypotheses proposed regarding the effects of response or explanatory variables. The cause and effect relation can be determined depending upon many factors such as hectoliter weight, broken grain rate, defective grain rate, sunn pest and insect damaged grain rate, germinated/sprouted grain rate and humidity rate which are effective in the formation of price. This relationship can be investigated statistically via multiple regression analysis. The price is the dependent variable and other factors can be considered as independent variables. In multiple regression method, the estimation of regression coefficients that minimize the distance between the actual measures of the dependent variable and the predictive measures obtained from independent variables is calculated by the Least Squares Estimation (abbreviation for LS in Turkish) (Akdeniz and Öztürk 1996), (Karadavut, 2005).

If the equation obtained in the regression analysis is desired to be used for parameter estimation, some assumptions are needed to be made. For instance, error terms are required to come from normal distribution, average of expected values of errors need to be 0 (zero), variations should be homogeneous and equal to σ^2 and there has to be a correlation between error terms and explanatory variables. (Alma ve Vupa, 2008). However, when the distribution of error terms is a normal distribution, the LS method is the same as the most likelihood method (Faraway, 2005).

For this reason; in linear regression, LAD regression, Non-parametric regression, Bayes regression, Robust regression and such methods that are alternative to the least squares method can be applied to the data that belongs to the factors which affect the price of the wheat. In this study, it is tried to find the most convenient method among LAD regression, M-regression, Non parametric regression and Bayes regression which are alternative methods to the Least Squares method. In the selection of estimation methods, multiple specification coefficient (R^2), Akaike information criteria (AIC) and error mean squared (EMS) are taken into consideration.

In multiple linear regression model which is linear according to the parameters, Y_i ; dependent variable, X_{ij} ; independent variables, β_j ; J . regression coefficient and ε_i ; i error term and the model function is,

$$Y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik} + \varepsilon_i, i = 1, 2, \dots, n \tag{1}$$

related to the regression equation,

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \dots \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{21} & \dots & x_{1k} \\ 1 & x_{12} & x_{22} & \dots & x_{2k} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_{1k} & x_{2k} & \dots & x_{nk} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \dots \\ \beta_k \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \dots \\ \varepsilon_n \end{bmatrix} \tag{2}$$

regression equation under notations,

$$\underline{Y} = X \underline{\beta} + \underline{\varepsilon} \tag{3}$$

Model are called “hypothesis models” when the ε is normally distributed depending on the distribution of error term ε_i .

Purpose

The aim of this research is to determine the effect of the product price factor on the price of wheat for the production and production of bezostaja wheat.

For this purpose, the following questions will be answered.

- Does the hectolitre weight affect the price of wheat?
- Does sunflower damage have an effect on the price of wheat?
- Does moisture have an effect on wheat price?
-

Method

In this study, in which the Bezostaja wheat was used, the annual data of 2004 for 12 months that has been taken from Konya Commodity Exchange Market is examined.

In order to define correctly how other variables (like sunn pest, hectoliter and humidity) affect the price by applying LS, LAD, M regression, non-parametric regression and Bayes regression estimation methods, the most suitable model is desired to be preferred.

When there are outliers in the Linear regression model indicated by Equation 1, LAD regression estimation method has been obtained by assuming that the LS estimation method is affected by these outliers. In addition, in 1964, Huber has improved M regression estimation method in order to reduce the impact of outliers.

Least Squares Regression

With the matrix representation of the linear regression model,

$$\underline{Y}_{n \times 1} = X_{n \times p} \underline{\beta}_{n \times 1} + \underline{\varepsilon} \quad (4)$$

Let's consider the problem of estimating the $\underline{\beta} \in \mathbb{R}^p$ parameter in the classical assumptions about $\underline{\varepsilon}$ in model $\underline{\varepsilon} \rightarrow N(0, \sigma^2)$.

$$\min_{\underline{\beta}} (\underline{Y} - X \underline{\beta})' (\underline{Y} - X \underline{\beta}) = \min_{\underline{\beta}} (\underline{Y}' \underline{Y} - 2 \underline{\beta}' X' \underline{Y} + \underline{\beta}' X' X \underline{\beta}) \quad (5)$$

To solve the optimization problem, when $\underline{\beta}$ derivative is taken and equal to zero

$$X' X \hat{\underline{\beta}} = X' \underline{Y} \quad (6)$$

normal equations are had. X is full column rank,

$$\text{rank}(X) = \text{rank}(X'X) = p \quad (7)$$

because of,

$$\hat{\underline{\beta}} = (X'X)^{-1} X' \underline{Y} \quad (8)$$

solution is found (Birkes and Dodge, 1993). \underline{Y} is a sample, $\hat{\underline{\beta}}$ is a function of \underline{Y} , if $\underline{\beta}$ is a predictor this estimator is called a least squares estimator.

Lad Regression

When linear regression model in equation (1) is taken into consideration for the absolute least deviations (LAD) regression,

$$\sum |Y_i - \beta_0 + \beta_1 X_{i1} + \dots + \beta_n X_{in}| \quad (9)$$

The best line passing through a point is found out by developing algorithm. The vector illustration for the algorithm is;

$$\hat{\underline{\beta}} = \begin{bmatrix} \hat{\beta}_0 \\ \vdots \\ \hat{\beta}_k \end{bmatrix}, \quad X_i = \begin{bmatrix} X_{i1} \\ \vdots \\ X_{ik} \end{bmatrix}$$

Equation (9) is obtained as;

$$\sum |Y_i - \hat{\underline{\beta}}' X_i|$$

It aims to find the vector $\hat{\underline{\beta}}'$ which minimizes this value. The algorithm starts with the vector $\hat{\underline{\beta}}'$. Then the best vector of $\hat{\underline{\beta}}'$ is calculated by finding a better vector which

gives the minimum value in $\sum |Y_i - \hat{\beta}' X_i|$. $\hat{\beta}$ estimations are obtained as a better $\hat{\beta}^*$ vector in every step.

$$\hat{\beta}^* = \hat{\beta} + td \tag{10}$$

In order to find the vector in equation (10), the direction vector d and t value are needed to be calculated. Equation (11) is minimized in order to find the t value.

$$\sum \left| Y_i - (\hat{\beta} + td)' X_i \right| \tag{11}$$

Here; by transforming $Z_i = Y_i - \hat{\beta}' X_i$, $W_i = d' X_i$

$$|Z_i - tW_i|$$

is obtained (Birkes and Dodge, 1993). This obtained value is to find out $\hat{\beta}$ which will minimize $\sum |(Y_i - Y_0) - \hat{\beta}'(X_i - X_0)|$. Then by calculating Z_i/W_i rates, they are sorted ascendingly. k index is calculated by sorting Z and W according to this line.

$$|W_1| + \dots + |W_{k-1}| < \frac{1}{2} T$$

$$|W_1| + \dots + |W_k| > \frac{1}{2} T$$

Here, $T = \sum |W_i|$. The value which minimizes t is Z_k/W_k . There is one more $(p + 1)$ direction vector than P independent variable number in each step of algorithm. There is a positive direction (+) for each d_j vector as well as a negative direction (-). Therefore, the number of direction vectors must be $2 \times n$ where n is the number of independent variables. A vector is chosen such that it converges to the value of the $\sum |Y_i - (\hat{\beta} + td)' X_i|$ at $t = 0$.

M-Regression

In multiple M Regression model, Huber M estimations are the values which minimize;

$$\sum \rho \left(Y_i - (\hat{\beta}_0 - \hat{\beta}_1 X_{i1} + \dots + \hat{\beta}_n X_{in}) \right) \tag{12}$$

and are obtained as:

$$\rho(e) = \begin{cases} -k \leq e \leq k & e^2 \\ e < -k \text{ or } e > k & 2k|e| - k^2 \end{cases} \quad (13)$$

Huber M-regression estimation vector $\hat{\beta}$ is defined as $\hat{\beta}$ vector which minimize $\sum \rho(Y_i - (\hat{\beta}_0 - \hat{\beta}' X_i))$ (Birkes and Dodge, 1993). The regression coefficient vector $\hat{\beta}$ is firstly estimated via LS. The algorithm continues in this way until the first estimations of $\hat{\beta}$, are reached.

Non Parametric Regression

Non parametric regression estimation is a strong analysis method for the data set which includes outlier observation. Due to the outlier observations, parameters are corrupted and many methods are unable to give the most convenient result. For this reason, non-parametric regression estimation is required (Hardle, 1990).

The median of $Y_i - (\hat{\beta}_1 X_{i1} + \dots + \hat{\beta}_n X_{in})$ difference determines the non-parametric estimation of $\hat{\beta}^0$. By minimizing the value in equation (14), an algorithm is produced. The vector illustration is,

$$g(b) = \sum \left[\text{rank}(Y_i - b'X_i) - \frac{n+1}{2} \right] (Y_i - b'X_i) \quad (14)$$

and, $b = (b_1, \dots, b_k)$, $X_i = (X_{i1}, \dots, X_{ik})$ (Birkes and Dodge, 1993).

As the vector $\hat{\beta}$ is acquired by minimizing $g(b)$ with LS method, iteration is started. The vectors which give smaller values of the function $g(b)$ are found. The initial vector $\hat{\beta}^0$ is found and $\hat{\beta}^* = \hat{\beta}^0 + t^* d$ is acquired. Here is, t^* and d , $Z_i = Y_i - (\hat{\beta}^0)' X_i$, $u_i^0 = \text{rank}(Z_i) - \frac{1}{2}(n+1)$, $u^0 = u_i^0$ and the $n \times 1$ vector. The $n \times p$ matrix consists of $X_c = X_{ij} - \bar{X}_j$, $d = (X_c' X_c)^{-1} X_c' u^0$, $w_i = d' X_i$, t^* is the weighted median of $|w_i - w_j| / \sum |w_i - w_j|$ and $|z_i - z_j| / |w_i - w_j|$. By this way, the iteration is completed, and $\hat{\beta}^*$ becomes the valid vector, and the iteration is repeated (Birkes and Dodge, 1993).

Bayes Regression

Y is the dependent variable vector and θ is the unknown parameter vector of the model. In Bayes approach, for θ the range is not specified and before having a look at the data set, we look at what is known about the parameter. By this way, probability distribution for the data set vector is determined initially. The generally used distribution is normal distribution.

The multiple linear Bayes regression model is;

$$f(\beta|y, \sigma) = cf(\beta|\sigma)f(y|\beta, \sigma) \tag{15}$$

Here, c fixed value, $f(y|\beta, \sigma)$, is a function of y . Multiple normal distribution probability average vector is $X\beta$ and the variance-covariance matrix is σ^2I (I is unit matrix).

$$f(y|\beta, \sigma) = c \exp\left[-\frac{1}{2\sigma^2}\|y - X\beta\|^2\right] \tag{16}$$

$$\|y - X\beta\|^2 = \|y - X\hat{\beta}_{LS}\|^2 + \|X\hat{\beta}_{LS} - X\beta\|^2 \quad (\text{the theorem of pythagoras})$$

$$f(y|\beta, \sigma) = c \exp\left[-\frac{1}{2\sigma^2}\|X\hat{\beta}_{LS} - X\beta\|^2\right]$$

LS estimation of posterior distribution of per β is $\hat{\beta}_{LS}$

$$f(\beta|\sigma) = c \exp\left[-\frac{1}{2\sigma^2}(\beta - b)'V^{-1}(\beta - b)\right] \tag{17}$$

The variance-covariance matrix of the vector b is σ^2V . b and V are chosen to obtain a priori knowledge and $\hat{\beta}$ is obtained as;

$$\hat{\beta} = V_*V^{-1}b + V_*X'X\hat{\beta}_{LS} \tag{18}$$

Here the result is, $V_* = (V^{-1} + X'X)^{-1}$ (Birkes and Dodge, 1993).

Selection Criteria

In the selection of the appropriate method, Error mean squared (MSE), Akaike information criteria (AIC) and multiple specification coefficient (R^2) are going to be used and the method which gives the biggest R^2 value and the smallest AIC and MSE values are going to be preferred.

Multiple Designation Coefficient

Multiple specification coefficient (R^2), the rate of X independent variables to explain Y dependent variables' total variation is ($0 < R^2 < 1$).

$$R^2 = \frac{b'XY - (\sum Y)^2/n}{\sum Y^2 - (\sum Y)^2/n} \quad (19)$$

(Newton and et. al. 1996). However, the fact that R^2 is close to 1 does not mean that the model is good. It will change the value of R^2 when any parameters are added to or removed from the model. Therefore, the models which have the biggest R^2 value might not always be the best models (Montgomery and Peck, 1992), (Şahin, 2000).

Akaike information criteria

Akaike information criteria (AIC) are used to choose the most appropriate one among different models.

$$AIC = \ln \left(\frac{\sum e_i^2}{n} \right) + \frac{2k}{n} \quad (20)$$

In equation (20), k represents the number of parameters including fixed term and n represents the number of observations. Also, $\sum e_i^2$ represents the total of error squares. In model comparisons, the model which gives the smallest AIC value is preferred.

Error mean squared (MSE)

Another criteria used in choosing the most convenient model is Error mean squared method. The smaller the error square mean, the smaller the deviation between the expected value and the observed value is. Error mean squared (MSE) criteria;

$$MSE = \frac{1}{n} \sum (Y_i - \hat{Y}_i)^2 \quad (21)$$

It is calculated by this formula (Ekinci and et. al., 2016).

Findings

During the formation of the price of wheat, there have been several criteria determined by TMO (abbreviation for Turkish Grain Board). In this study, sunn pest, hectoliter and humidity are chosen from these criteria in practice. The data used in this study includes the values of hectoliter weight, sunn pest/aelia and the humidity rate which were effective in the price formation of Bezostaja wheat in Konya, in 2004 according to the Konya Commodity Exchange Market.

Table 1. The price of Bezostaja wheat^(*y*), hectoliter^(*X*₁), sunn pest^(*X*₂) and humidity^(*X*₃) in Konya.

Price	Hectoliter	Sunn Pest	Humidity
4.000	79.1	3.7	8.1
3.880	78.1	4.7	8.9
3.700	78.6	6.4	12.6
4.000	79.2	3.1	10.1
3.700	77.5	6.0	9.0
3.860	76.8	1.5	8.9
3.770	80.4	3.4	9.9
3.780	72.5	1.5	10.8
4.070	78.6	1.5	11.8
3.860	72.1	1.8	9.2
4.040	81.1	3.5	10.1
3.900	77.8	3.8	10.7
3.980	75.2	0.9	9.6
4.000	77.8	1.5	10.1
3.750	79.0	1.3	10.9
3.970	79.3	2.5	10.6
3.600	71.4	1.9	11.1
3.490	74.2	0.6	9.9
3.870	77.7	4.4	9.0
3.850	77.8	2.8	11.8

By using the data set in Table.1, the regression model is handled as,

$$\text{Price}=\beta_0+\beta_1\times\text{Hectoliter}+\beta_2\times\text{Sunpest}+\beta_3\times\text{Humidity}+\varepsilon$$

As alternatives for the LS estimation method, LAD, M-regression, non-parametric regression, Bayes regression estimation methods, SPSS and Excel packaged software are used. By this way, the multiple coefficient of determination (R^2), Akaiki information criteria (AIC) and error mean squared (MSE) of the model are obtained and the best model for the data set is tried to be chosen. The calculations are summarized in Table 2.

Table 2. The alternative regression models obtained for the datum of Bezostaja wheat

Forecasting Method	Regression Model	R^2	AIC	MSE
LS	$\hat{Y}=1432+35.4X_1-29.3X_2-22.3X_3$	0.344	9.99	14.634,76
LAD	$\hat{Y}=1220+27.36X_1-69.18X_2-22.12X_3$	0.192	14.11	907.998,28
M-Regression	$\hat{Y}=1259.576+39.001X_1-48.74X_2-24.499X_3$	0.502	10.01	16.342,28
Non-Parametric Regression	$\hat{Y}=1259.576+32.95X_1-36.51X_2-5.41X_3$	0.697	12.13	125.338,2
BAYES	$\hat{Y}=1436+34.81X_1-44.62X_2-13.750X_3$	0.359	10.04	15.395,76

When we have a look at Table.2, the EMS value of LS estimation method is 14.634,76 and this value is found to be smaller than other methods. Similarly, AIC value of LS estimation method is 9,99 and this value is also found to be smaller than other methods. However, the R^2 value of non-parametric regression estimation method is 0,697 and this value is found to be bigger than the R^2 values of other methods.

Discussion and Conclusion

In this study, it is tried to be given some alternative regression models for the LS estimation method. LAD, M-regression, non-parametric regression and Bayes regression estimation methods are examined. The coefficient of determination that belongs to the examined sample is found to be the biggest as $R^2 = 0,697$ in non-parametric regression.

The defining rate of the used independent variables for the total variation of the dependent variable (Y) is found relatively 0,697. The remaining 0,303 part is formed due to the unexamined other factors. It can be stated that there might be outlier observations among observations due to the high R^2 values. Under these circumstances, although non-parametric regression estimation method seems to be a convenient estimator, according to the EMS and AIC criteria, LS estimation method steps forward as the most convenient

method for this data set.

Implication and Suggestions

Although the Nonparametric Regression estimation method is considered to be a suitable estimator, the most suitable method for this data set stands out according to the HKO and AIC criteria.

The fact that this type of analysis is not high indicates that there may be contradictory observations within the observations.

In such studies, the boundaries of impact factors can be expanded and the boundaries can be expanded with more province data.

Research Limitations and Future Research

Research done,

- The research is limited to the Commodity Exchange 2004 data in Konya.
- The study group is limited to the 12-month wheat data of 2004.
- In comparison of data, the factors affecting the price are limited to physical factors.

Considering these limitations, it may be suggested to make applications with different types of data comparisons in different provinces and to apply in different product types in future studies.

Author Contributions

This study is produced from master thesis. Study authors include graduate students and advisors. Both authors contributed jointly at each stage of the study. All authors have read and accepted the published version of the article.

Publication Ethics

Research and Publication Ethics was complied