



Estimation of ill-posed linear deterministic regression model: generalized maximum entropy and bayesian approach

Sibel Örk Özel^{1*} , Selin Çabuk² 

¹Department of Econometrics, Cukurova University, Adana, 01330, Turkey

²Department of Industrial Engineering, Cukurova University, Adana, 01330, Turkey

Highlights:

- Estimation of the ill-posed regression model
- Model estimation with generalized maximum entropy
- Using the alternative bayes approach to generalized maximum entropy

Keywords:

- Ill-posed regression model
- Generalized maximum entropy
- Alternative bayes approach

Article Info:

Research Article
Received: 15.03.2021
Accepted: 20.08.2021

DOI:

10.17341/gazimmfd.897120

Correspondence:

Author: Sibel Örk Özel
e-mail: sork@cu.edu.tr
phone: +90 322 338 7265 / 6166

Graphical/Tabular Abstract

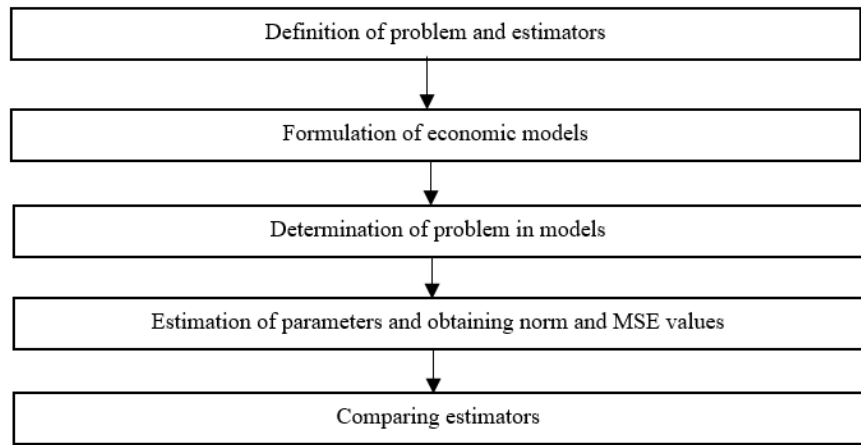


Figure A. Steps of solution process for model

Purpose: There are many econometric models that can be established in engineering, social sciences and many scientific branches. The aim of the study is to show the methods of obtaining the solutions of econometric models shown with ill-posed equation systems, which is one of these models, and the efficiency of the methods.

Theory and Methods:

Ill-posed regression model cannot be estimated by methods such as least squares. For the estimation of these models, the Generalized Maximum Entropy (GME) method and the alternative Bayesian approach to the Generalized Cross Entropy (GCE) methods and GME are used.

Results:

In the application part of the study, GME and Bayesian approach have been handled to the models with ill-posed regression model. The parameter estimates, norm and MSE values obtained are given. The norm and MSE of the GME estimator and the norm and MSE of the Bayesian approach are very close to each other. This result indicate that the Bayesian approach can be used as an alternative to the GME.

Conclusion:

It can be said that the GME estimator, which has a slightly smaller norm and MSE than Bayesian approach, is more effective than the Bayesian approach.



Eksik sunumlu doğrusal deterministik regresyon modelinin tahmini: genelleştirilmiş maksimum entropi ve bayesçi yaklaşım

Sibel Örk Özel^{1*}, Selin Çabuk²

¹Çukurova Üniversitesi, İktisadi ve İdari Bilimler Fakültesi, Ekonometri Bölümü, 01330 Adana, Türkiye

²Çukurova Üniversitesi, Mühendislik Fakültesi, Endüstri Mühendisliği Bölümü, 01330 Adana, Türkiye

Ö N E Ç İ K A N L A R

- Eksik sunumlu regresyon modelinin tahmini
- Genelleştirilmiş maksimum entropi ile model tahmini
- Genelleştirilmiş maksimum entropiye alternatif bayes yaklaşımının kullanılması

Makale Bilgileri

Araştırma Makalesi

Geliş: 15.03.2021

Kabul: 20.08.2021

DOI:

10.17341/gazimmfd.897120

Anahtar Kelimeler:

Eksik sunumlu regresyon modeli,
genelleştirilmiş maksimum entropi,
alternatif bayes yaklaşımı

ÖZ

Regresyon modelleri; mühendislik, sosyal bilimler ve birçok bilim dalında kullanılmaktadır. Bu çalışmada, eksik sunumlu denklem sistemleriyle gösterilen regresyon modellerinin çözümlerinin elde edilişi ele alınmıştır. Bu amaçla çalışmada, Genelleştirilmiş Maksimum Entropi (GME) ile GME ve Genelleştirilmiş Çapraz Entropi (GCE) yöntemlerine alternatif Bayes yaklaşımı kullanılmıştır. Gerçek veri kümesi üzerinde yapılan uygulama sonucunda elde edilen normlara ve hata kareleri ortalamalarına (HKO) göre tahmin ediciler karşılaştırılmıştır. GME'ye alternatif olarak kullanılacak olan Bayes yaklaşımı ile GME karşılaştırıldığında, yapılan uygulama sonucunda GME tahmin edicinin Bayes yaklaşımına göre daha etkin olduğu sonucuna ulaşılmıştır.

Estimation of ill-posed linear deterministic regression model: generalized maximum entropy and bayesian approach

H I G H L I G H T S

- Estimation of the ill-posed regression model
- Model estimation with generalized maximum entropy
- Using the alternative bayes approach to generalized maximum entropy

Article Info

Research Article

Received: 15.03.2021

Accepted: 20.08.2021

DOI:

10.17341/gazimmfd.897120

Keywords:

Ill-posed regression model,
generalized maximum entropy,
alternative bayes approach

ABSTRACT

Regression models are used in engineering, social sciences and many scientific branches. In this study, obtaining the solutions of regression models represented by equation systems with ill-posed problem is discussed. For this purpose, an alternative Bayesian approach to Generalized Maximum Entropy (GME) and Generalized Cross Entropy (GCE) methods, and GME estimator are used in the study. The estimators are compared according to the norms and mean squared errors (MSE) obtained as a result of the application made on the real data set. When the Bayes approach, which can be used as an alternative to GME, is compared with the GME, it is concluded that the GME estimator is more effective than the Bayesian approach.

1. GİRİŞ (INTRODUCTION)

Regresyon analizi; mühendislik, sosyal bilimler, biyoloji, veri madenciliği, örüntü tanıma vb. alanlarda en yaygın kullanılan veri analizi tekniklerinden biri olarak düşünülebilir. Genel olarak amacı, bağımlı bir değişken ile bir dizi bağımsız değişken arasındaki ilişkiyi tahmin etmek, bağımsız değişkenlerin bağımlı değişkeni tahmin edip etmediğini ve bağımsız değişkenlerin modellenen deneysel verileri nasıl açıkladığını anlamak için uygun bir matematiksel modeli (örneğin, doğrusal, polinom, ikinci dereceden) elde etmektir [1].

Birçok sektörde olabileceği gibi endüstriyel verilerin kullanımı ile üretim ve fiyat arasındaki ilişkinin de belirlenmesinde kullanılabilen gecikmesi dağıtılmış modellerde dahil doğrusal bir regresyon modelinde karşımıza iki tür problem çıkmaktadır. Bunlardan birisi kötü koşulluluk, diğeri ise eksik sunumluluk problemidir. Çok değişkenli regresyon modellerinde yaygın olarak karşılaşılan kötü koşulluluk problemi, önemli derecede çoklu iç ilişki (çoklu doğrusal bağlantı) problemi olarak da bilinmektedir. Uygulamalarda, önemli derecede çoklu iç ilişki probleminin varlığı durumunda klasik tahmin teknikleri ile modelin parametreleri kararlı olarak tahmin edilememektedir. Bu problemin varlığı durumunda kararlı tahminler elde etmek adına yanlış tahmin ediciler olarak bilinen Ridge, Liu, Genelleştirilmiş Maksimum Entropi (GME) gibi teknikler önerilmiştir.

GME, bir modelin yeniden parametrelendirilmesine dayanır. GME’de bilinmeyen parametreler ve hata terimleri, bazı destekler üzerinde tanımlanarak uygun olasılık dağılımlarının beklenen değeri olarak görülürler. Daha sonra, gözlemlenmeyen her bir değerle kullanılan verilere bağlı ortak entropileri maksimize eder ve uygun olasılık dağılımları için daha az varsayım ile geleneksel tahmin ediciler kullanılarak hesaplanan tahminlerden daha iyi tahminlere ulaşılabilir [2].

Al-Nasser [2], çalışmasında basit doğrusal yapısal ölçüm hata modeli parametrelerinin belirsizliğini tam olarak ölçmek için iki adımda GME tahmin yöntemini kullanan bir prosedür sunmaktadır. Önerilen tahmin prosedürü, her adımda GME sistemini formüle ederken bilinmeyen parametrelerin sayısını en aza indirme ve dolayısıyla tahminlerin değişkenliğini azaltma yeteneğine sahiptir. Bu çalışmada, Monte Carlo simülasyon karşılaştırması maksimum olasılık tahmin edicileriyle deneyler ve tek adımlı GME tahmin prosedürü sunulmuştur. Grünwald ve Dawid [3], geleneksel olarak ayrı kabul edilen entropiyi maksimize ederek ve en kötü durumda beklenen kaybı asgariye indirerek iki problem arasında yakın bir ilişki tanımlamışlar ve bunu geliştirmişlerdir. “Karar verici” ve “Doğa” arasındaki sıfır toplamli oyunların denge teorisine dayanan bir formülasyon kullanılarak, bu iki sorunun birbirleriyle ilişkili olduğu ve her birinin çözümünün diğeri için

fayda sağladığı gösterilmiştir. Suhartanto vd. [4] çalışmalarında, turist davranışını tahmin etmek için çoklu regresyon ile karşılaştırıldığında yapısal eşitlik modellemesinin (SEM) uygulanabilirliğini değerlendirmişlerdir. Bu çalışmada 403 turistin katıldığı bir anket kullanılmıştır. SEM için AMOS ve çoklu regresyon için SPSS programları kullanılarak karşılaştırma yapılmıştır. Çalışmada, SEM uygulamasının çoklu regresyona kıyasla turist davranışı konusunda daha iyi bir tahmin sunduğu ortaya konulmuştur.

Ciavolino ve Calcagni [1]; çalışmalarında eksik sunumluluk problemine sahip bulanık regresyon modellerini tahmin etme sorununu ele almışlardır. Çalışmada, GME tahmin yöntemine dayalı yeni bir bulanık regresyon modeli sunulmuştur. Önerilen yöntemin bazı özelliklerini daha iyi vurgulamak için iki tane Monte Carlo deneyi yapılmış olup, gerçek bir vaka çalışması analiz edilmiştir. Ciavolino ve Calcagni [5] çalışmalarında, GME yönteminin mantığına dayanan net girdi / bulanık çıktı regresyon modeli sunmuşlardır. Yaklaşım; küçük örnekler, hatalı tasarım matrisi (örn. çoklu bağlantıdan dolayı), eşitsizlik kısıtlamaları olan tahmin problemleri gibi belirli problemlerle başa çıkılması gereken birkaç durumda kullanılabilir. GME’yi tanımladıktan sonra bulanık regresyon modeli olarak, GME yaklaşımından sağlanan özelliklerin değerlendirildiği iktisadi verilere sahip bir vaka çalışmasını ele almışlardır. Ayrıca modelin bazı özelliklerini daha iyi değerlendirmek için vaka çalışmasının ana sonuçları üzerinde bir duyarlılık analizi yapmışlardır. Son olarak, bazı kritik noktalar daha sonraki çalışmalar için önerilerle birlikte tartışılmıştır.

Macedo [6] çalışmasında, Ridge Regresyonunu ve GME tahmin edicisini ridge trace analizinde özneliği ortadan kaldırmak için kullanmıştır. Bir simülasyon çalışması ve iki ampirik uygulama, geliştirilmiş tahmin edicinin performansını göstermek için kullanılmıştır. Tamamlayıcı materyal olarak bir MATLAB kodu verilmiştir. Kamar ve Msallam [7] çalışmalarında, dört parametrelili Weibull büyüme modelini tahmin etmek için GME, Bayes ve maksimum ardıl (maximum a posteriori) olmak üzere üç yöntem sunmuş ve karşılaştırmışlardır. Bu amaca ulaşmak için, bir simülasyon tekniği kullanılmıştır. Hesaplama sonuçlarından Bayes yönteminin en iyi tahminleri verdiği gösterilmiştir. Heckelevi vd. [8] çalışmalarında, gecikmesi dağıtılmış modellerde GME ve Genelleştirilmiş çapraz entropi (GCE)’ye alternatif Bayes yaklaşımını ele almışlardır. Bilgi verici ve bilgi vermeyen önsel ortalamalar olduğunda bu tekniklerin sunduğu sonuçları ifade ederek alternatif Bayes yaklaşımının GME ve GCE teknikleri yerine kullanılabilirliğini göstermişlerdir. Chinnakum ve Boonyasana [9] çalışmalarında, panel veri kullanılan regresyon modelleri için ikili bir GME tahmin edicisi kullanılarak Tayland’da gelen uluslararası turistlerin davranışını etkileyen faktörleri incelemişlerdir. Entropi yaklaşımının avantajı olan panel verileri için eksik

sunumluluk problemi ve entropi yaklaşımı ile başa çıkma yeteneği henüz turizm literatüründe araştırılmadığı için çalışmada uluslararası turizm talebiyle ilgili bir dizi önemli ekonomik faktör, gelir, fiyat, döviz kuru ve nüfus sayısı ele alınarak model tahmin edilmiştir. Çalışma, iki yöntemin sonuçlarını, yani en küçük kareler (EKK) tahmin edicisi ve GME tahmin edicisini kıyaslamaktadır. Elde edilen bulgulara göre GME tahmin edicisi, EKK tahmin edicisinden daha düşük hata kareleri ortalamasına (HKO) sahiptir. Böylece çalışmada, GME tahmin edicisinin EKK tahmin edicisinden daha iyi performans gösterdiği sonucuna ulaşılmıştır. Turizm talep tahmin sonuçlarına göre, binlerce turistin geliri, döviz kurları ve menşei ülkelerdeki nüfus sayısındaki artış Tayland'a gelen uluslararası ziyaretçi girişleri üzerinde olumlu etkiye sahipken, görece fiyatın uluslararası ziyaretçi girişlerini olumsuz etkilediği görülmüştür. Çalışma ayrıca kişi başına milli gelirin Tayland turizm talebi için güçlü tahmin gücüne sahip olduğunu bulmuştur.

Tarkhamtham ve Yamaka [10]; Shannon entropi ölçüsünü değiştirmek amacıyla doğrusal olmayan kink regresyon bağlamında Renyi GME ve Tsallis GME isimli yüksek dereceli GME tahmin edicilerinin performansı üzerinde araştırma yapmışlardır. Performans karşılaştırması için, çeşitli hata dağılımları ile iki farklı örnek boyutunu kullanarak Monte Carlo simülasyonu kullanmışlardır. Daha sonra elde edilen model gerçek verilere uygulanmıştır. Gerçek veri uygulamasında, yüksek dereceli GME tahmin edicilerinin, Shannon GME tahmin edicisinden çok farklı olmadığı sonucuna ve simülasyon çalışmasında ise yüksek dereceli GME tahmin edicilerinin Shannon GME'den tamamen üstün olmadığı sonucuna ulaşılmıştır. Bununla birlikte ortalama mutlak hata kriterine göre, Renyi ve Tsallis GME'nin Shannon GME'den daha iyi performans gösterdiği çalışmanın sonuçlarıdır. Böylece, yüksek dereceli GME tahmin edicisinin doğrusal olmayan ekonometrik çerçevede alternatif bir araç olarak kullanılabilmesi sonucuna varılmıştır. Maneejuk vd. [11] çalışmasında, bağımlı ve bağımsız değişkenler arasındaki doğrusal olmayan ilişkiyi yakalamak için kink regresyon modelini önermiştir. Model, sürekli regresyonun herhangi bir eşik veya bükülme noktasında düzeltilmesine izin vermek için kullanılır. Böylece her bağımsız değişken ile bağımlı değişken arasındaki ilişkiler için kink etkilerinin farklı olmasına izin verilmiş olur. Ayrıca bazı durumlarda, regresyon modelinde ortaya çıkan problemler olabilir (bilinmeyen parametrelerin sayısı gözlem sayısını aşarsa veya temeldeki dağılım bilinmiyorsa). Bu nedenle, modeli tahmin etmek için GME tahmin edicisi uygulanmıştır. Bu çalışma, standart EKK, Bayesçi ve Maksimum Olabilirlik dahil olmak üzere birçok geleneksel tahmin ediciyle karşılaştırmalı olarak hem simülasyona hem de gerçek veri setine dayalı deneyler yürütmektedir. Deneysel sonuçlar, GME tahmin edicisinin parametre tahminleri için yararlı bir araç olduğunu göstermektedir. Simülasyonlar ayrıca, verilerin sınırlı olduğu ve normal olmayan dağılımın tutulduğu, önerilen tahmin yönteminin mükemmel sonuçlu örnek özelliklerini ortaya çıkarmıştır.

Uluslararası literatür incelendiğinde 2000'li yıllarla birlikte GME tahmin edicisinin kullanılmaya başlandığı ve diğer tahmin edicilere göre etkinliğinin tespit edilmeye çalışıldığı anlaşılmaktadır. Ancak literatürde doğrusal regresyon modellerinde açıklayıcı değişken sayısının fazla olması durumunda ortaya çıkan eksik sunumluluk probleminin [12] regresyon modellerinde uygulandığı çalışmalar oldukça azdır. Eksik sunumluluk probleminin varlığı durumunda klasik tahmin tekniklerini kullanmak mümkün değildir [13]. Çünkü EKK yöntemi gibi klasik tahmin teknikleri kullanıldığında tek (biricik) tahmin değerleri elde edilememektedir. Bu durumda yanlı tahmin ediciler kullanılmaktadır. Bu yanlı tahmin edicilerden biri olan Bayes ise gelişen teknoloji ile birlikte özellikle mühendislik çalışmalarında çok önemli bir kavram olarak karşımıza çıkmaya başlamış ve teknolojinin hız kesmeden gelişmeye devam etmesiyle de bundan sonraki çalışmalarda sıklıkla kullanılacaktır.

Çalışmamızda amaç, eksik sunumlu modellerin tahmininde kullanılabilir olan yanlı tahmin ediciler olan GME, GCE ve alternatif Bayes yaklaşımlarını ele almak ve eksik sunumlu doğrusal deterministik bir model olan gecikmesi dağıtılmış bir modele GME ve alternatif Bayes yaklaşımını uygulayarak sonuçları tartışmaktır. Bu amaçla bahsedilen tahmin ediciler normlarına ve hata kareleri ortalamalarına (HKO) göre karşılaştırılmış ve etkinlikleri açısından elde edilen sonuçlar tartışılmıştır. Yapılan literatür incelemesi sonucunda bahsedilen bu tahmin edicileri hem norm hem de HKO bazında etkinlikleri açısından karşılaştıran bir çalışma ile karşılaşılmamıştır. Bu çalışmada, GME ve GME'ye alternatif Bayes yaklaşımının hem norm hem de HKO değeri bazında karşılaştırılarak etkinliklerinin değerlendirilmesinin uluslararası literatüre katkı sağlayacağı düşünülmektedir.

2. TEORİK METOT (THEORETICAL METHOD)

2.1. Çalışmada Kullanılan Model (Model used in the Study)

Çoklu doğrusal regresyon modeli ile tanımlanmış bir modelde parametre sayısının gözlem sayısından fazla olması durumu ile karşılaşılabılır. Bu durum araştırmacıyı eksik sunumluluk problemi ile karşı karşıya getirir. Böyle bir durumda araştırmacı modeli tahmin etmek için yeterli örneklem çapına ulaşamamışsa eksik sunumlu modellerin tahmininde kullanılan teknikleri uygulamak zorundadır. Bu durum açıklanan ve açıklayan değişkenlerin modelde mutlaka yer alması gerekliliğinden ortaya çıkabilir.

Diğer bir durum ise gecikmesi dağıtılmış modellerin gecikme uzunluğunun belirlenmesinde kullanılan kriterlerin farklı gecikme uzunluğunu işaret etmeleri durumunda da ortaya çıkabilir. Gecikme uzunluğunun belirlenmesinde; son tahmin hata kriteri (FPE), Akaika bilgi kriteri (AIC), Hannan-Quinn bilgi kriteri (HQ), Schwarz bilgi kriteri (SC), genelleştirilmiş çapraz geçerlilik yöntemi (GCV) gibi kriterler kullanılmaktadır [14].

Eksik sunumluluk probleminin tanımlandığı bu durum, çalışmamızda gecikmesi dağıtılmış bir model tahmin edilerek tartışılacaktır. Gecikme sayısı k olan model Eş. 1 şeklindedir.

$$y_t = \alpha + \beta_0 x_t + \beta_1 x_{t-1} + \beta_2 x_{t-2} + \dots + \beta_k x_{t-k} + e_t \quad (1)$$

2.2. Kullanılan Tahmin Tekniklerinin Tanıtımı (Presentation of used Estimation Techniques)

2.2.1. Genelleştirilmiş maksimum entropi (GME) tahmin edicisi

(Generalized maximum entropy (GME) estimator)

GME tahmin edicisi Golan, Judge ve Miller [13] tarafından 1996 yılında önerilmiş olup, bir bilgilendirme ölçütü olarak Shannon [15] tarafından ortaya atılan ve Jaynes [16, 17] tarafından işlevsel hale getirilen maksimum entropi (ME) tahmin edicisinin genel lineer modeller için genelleştirilmiş halidir. ME yöntemi sadece eksik-sunumlu inverse problemler için biricik (unique) tahmin elde etmek için kullanılırken GME tahmin edicisi hem eksik-sunumlu hem de kötü-koşullu inverse problemler için kullanılabilir [13].

Genel lineer modeller;

$$\begin{aligned} \mathbf{y} &: T \times 1 \text{ boyutlu açıklanan değişken vektörü,} \\ \mathbf{X} &: T \times K \text{ boyutlu açıklayıcı değişkenler matrisi,} \\ \boldsymbol{\beta} &: K \times 1 \text{ boyutlu bilinmeyen parametreler vektörü,} \\ \mathbf{u} &: T \times 1 \text{ boyutlu hata vektörü} \end{aligned}$$

olmak üzere Eş. 2 ile ifade edilir.

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u} \quad (2)$$

Golan vd. [13] tıkmaz (compact) desteklerle birlikte bilinmeyen parametreleri ve hataları GME ile tahmin etmek için (2) numaralı denklemde verilen model parametrelerini ve hata terimlerini yeniden tanımlamışlardır. $\mathbf{p}_k = [p_{k1}, p_{k2}, \dots, p_{kM}]'$ ve $\mathbf{z}_k = [z_{k1}, z_{k2}, \dots, z_{kM}]'$ olmak üzere ($M \geq 2$) bilinmeyen parametreler vektörü Eş. 3 olarak ifade edilir.

$$\boldsymbol{\beta} = \mathbf{Z}\mathbf{p} \quad (3)$$

Burada; \mathbf{Z} : $K \times KM$ boyutlu kompakt destek matrisi ve \mathbf{p} : $KM \times 1$ boyutlu ağırlıklar vektörüdür. Ayrıca, \mathbf{V} : $T \times TJ$ boyutlu destek noktaları matrisi ve \mathbf{w} : $TJ \times 1$ boyutlu bilinmeyen ağırlıklar vektörü ($J \geq 2$) kullanılarak bilinmeyen hata vektörü Eş. 4 ile yazılabilir.

$$\mathbf{u} = \mathbf{V}\mathbf{w} \quad (4)$$

Sonuç olarak yeniden parametrelendirilen model Eş. 5 ile ifade edilebilir.

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u} = \mathbf{XZ}\mathbf{p} + \mathbf{V}\mathbf{w} \quad (5)$$

Yeniden parametrelendirilen model GME problemi olarak Eş. 6 ile ifade edilebilir.

$$\begin{aligned} \max H(\mathbf{p}, \mathbf{w}) = & \\ - \sum_{k=1}^K \sum_{m=1}^M p_{km} \ln p_{km} - \sum_{t=1}^T \sum_{j=1}^J w_{tj} \ln w_{tj} & \quad (6) \end{aligned}$$

Buradan hareketle Lagrange denklemi Eş. 7 şeklinde kurulup çözüldüğünde Eş. 8. ve Eş. 9 olduğu gösterilebilir.

$$\begin{aligned} L = -p' \ln p - w' \ln w + \lambda'(y - XZp - Vw) + \\ \gamma'[1_K - (I_K \otimes 1'_M)p] + \delta[1_N - (I_N \otimes 1'_J)w] \end{aligned} \quad (7)$$

$$\hat{p}_{km} = \frac{\exp(-\sum_{t=1}^T \hat{\lambda}_t z_{tkm} x_{tk})}{\sum_{m=1}^M \exp(-\sum_{t=1}^T \hat{\lambda}_t z_{tkm} x_{tk})} \quad (8)$$

$$\hat{w}_{tj} = \frac{\exp(-\hat{\lambda}_t v_{tj})}{\sum_{j=1}^J \exp(-\hat{\lambda}_t v_{tj})} \quad (9)$$

Elde edilen tahmin ediciler yerlerine yazıldığında Eş. 10 ve Eş. 11 ile gösterilen GME tahmin edicilerine ulaşılır.

$$\hat{\boldsymbol{\beta}} = \mathbf{Z}\hat{\mathbf{p}} \quad (10)$$

$$\hat{\mathbf{u}} = \mathbf{V}\hat{\mathbf{w}} \quad (11)$$

2.2.2. Genelleştirilmiş çapraz entropi (GCE) tahmin edicisi (Generalized cross entropy (GCE) estimator)

ME yönteminin aksine çapraz entropi (CE) yönteminde gerçek olasılık \mathbf{p} ile belli olayların olasılıklarının önsel tahmini olan \mathbf{q} arasındaki entropi mesafesi en aza indirgenmek üzere model yeniden formüle edilebilir. Burada p değerleri bilinmemektedir. Ancak olasılıklar hakkında ön örneklem bilgisi veya örneklem dışı bilgi olmasa bile X rasgele değişkeninin beklenen değeri bilinir [18]. Eş. 12 ve Eş. 13 ile gösterilen kısıtlar verildiğinde CE, Eş. 13 şeklinde yazılır [13].

$$\mathbf{y} = \mathbf{X}\mathbf{p} \quad (12)$$

$$\sum_{i=1}^K p_i = 1 \quad (13)$$

$$\begin{aligned} I(\mathbf{p}, \mathbf{q}) = \sum_{i=1}^K p_i \ln(p_i/q_i) = \sum_{i=1}^K p_i \ln p_i \\ - \sum_{i=1}^K p_i \ln q_i = \mathbf{p}' \ln \mathbf{p} - \mathbf{p}' \ln \mathbf{q} \end{aligned} \quad (14)$$

CE formülasyonu \mathbf{Z} ve \mathbf{V} destek matrisleri kullanılarak GCE, Eş. 15 ile yeniden parametrelendirilerek formüle edilir.

$$I(\mathbf{p}, \mathbf{q}, \mathbf{w}, \mathbf{u}) = \mathbf{p}' \ln(\mathbf{p}/\mathbf{q}) - \mathbf{w}' \ln(\mathbf{w}/\mathbf{u}) \quad (15)$$

Burada kısıtlar Eş. 16, Eş. 17 ve Eş. 18 ile verildiği gibidir [13].

$$\boldsymbol{\alpha} = \boldsymbol{\Gamma}\mathbf{Z}\mathbf{p} + \mathbf{V}\mathbf{w} \quad (16)$$

$$\mathbf{1}_K = (\mathbf{I}_K \otimes \mathbf{1}'_M)\mathbf{p} \quad (17)$$

$$\mathbf{1}_T = (\mathbf{I}_T \otimes \mathbf{1}'_J)\mathbf{w} \quad (18)$$

2.2.3. GME ve GCE Tahmin Edicilerine Alternatif Bayes Yaklaşımı

(Alternative Bayesian Approach to GME ve GCE Estimators)

Eksik sunumlu sistemlerin çözümünde ise GME ve GCE yöntemlerine alternatif olarak Bayes yaklaşımı kullanılabilir [8]. Bayes yaklaşımında; model parametreleri stokastik değişkenler olarak ele alınmaktadır. Bu bağlamda yöntem, önsel yoğunluk $[p(\beta)]$, olabilirlik fonksiyonu $[L(\beta|y)]$ ve ardıl yoğunluk $[h(\beta|y)]$ arasında ayrımları dikkate alarak seçim yapar. Bu üç önemli bileşenden önsel yoğunluk olan $p(\beta)$, bünyesinde modelin parametreleri ile ilgili önsel bilgiler içerir. $L(\beta|y)$, olabilirlik fonksiyonu, varsayılan modele bağlı olarak veriden elde edilen bilgiyi ve ardıl yoğunluk olan $h(\beta|y)$ ise önsel bilgiyi ve veri bilgisini birlikte ele alarak Bayes kuramı bağlamında sonuçlandırır. Bu üç eleman arasındaki ilişki (19) nolu ilişki ile açıklanabilir [19, 8]. Bu ilişkiye göre, ardıl yoğunluk, önsel yoğunluk ve olabilirlik fonksiyonunun çarpımı ile orantılı olduğu bilgisini Eş. 19 ifadesi özetler.

$$h(\beta|y) \propto p(\beta)L(\beta|y) \tag{19}$$

Ardıl yoğunluk olasılık tanımlarının kullanımıyla β hakkında istatistiksel sonuçların elde edilmesine veya bir kayıp ölçütüne göre optimal olan nokta tahminlerinin elde edilmesine olanak tanır. Ardıl olasılık ortalaması, karesel kayıp fonksiyonunu minimize eden değerdir.

Bu bileşenlerin uygun yorumlanması ile eksik sunumlu doğrusal modelin parametrelerinin tahminine GME yaklaşımını, Bayesçi yaklaşım kapsamında ele almak mümkündür. GCE/GME beklenen değer işlemcisi, β ile temsil edilirse, GME optimizasyon problemi (Eş. 20) ile ifade edilir.

$$\max_{\beta} \{h(\beta|y) \propto p(\beta)L(\beta|y) \propto [\sum_{k=1}^K H(\beta_k)]I_{\{\beta: y=X\beta\}}(\beta)\} \tag{20}$$

Burada $I_A(\beta)$, standart gösterge fonksiyonu olup $\beta \in A$ ise 1 değerini ve $\beta \notin A$ ise 0 değerini alır. $H(\beta_k)$ ise β 'nin optimal değeri GME çözümü ile elde edilen optimal değerine eşit olacaktır.

Bayes bağlamında, amaç fonksiyonu $h(\beta|y)$, model parametrelerinin birleşik ardıl yoğunluğu şeklinde yorumlanır. Burada $h(\beta|y)$, yani birleşik ardıl yoğunluk $p(\beta) \propto \sum_{k=1}^K H(\beta_k)$ ile tanımlanan önsel olasılığın olabilirlik fonksiyonu ile çarpımı şeklinde tanımlanır. Olabilirlik fonksiyonu ise doğrusal model olan $y=X\beta+\varepsilon$ ilişkisinde kısıtı sağlamayan β değerlerine sıfır ağırlığını veren ancak doğrusal model ve veri ile uyumlu olan β değerlerine pozitif ağırlık veren fonksiyondur. Bu durum, kısıtları sağlamayan β değerleri için sıfır değerli ardıl yoğunluk ağırlıklarına ve β 'nin diğer tüm değerleri için önsel ağırlığa göre farklı ardıl ağırlıklar anlamına gelir. $h(\beta|y)$ ifadesini maksimum yapan β değeri, β 'nin ardıl dağılımının modudur. Bu ardıl dağılıma β tahmininin “en yüksek değerli ardıl yoğunluğu” denir [8].

Heckeley vd. [8]'nin de ifade ettiği gibi Bayes formülasyonu içinde ele alınan GME/GCE yukarıdaki şekilde yorumlandığında, entropi yaklaşımına alternatif olarak Bayes ile yaklaşımın üç yararlı özellik sağladığı görülür. Bu özellikler şu şekildedir:

- Destek noktaları seçimi ve entropi ölçütü tarafından yapılan ağırlıklandırmalar uygun tanımlanırsa, Bayes yaklaşımı GME/GCE yaklaşımına tamamen eşit formüle edilebilir.
- Bilinmeyenlere ilişkin önsel bilgi, uygun önsel yoğunluk saptanarak açık bir şekilde formüle edilebilir.
- Optimizasyon problemi daha az sayıda değişken içerir ve uygun önsel yoğunluk fonksiyonları seçiminde daha az hesaplama işlemine gereksinim duyar.

Tablo 1. Çalışmada kullanılan veri seti (Data Set used in the Study)

t	y	x	xt-1	xt-2	xt-3	xt-4	xt-5	xt-6	xt-7	xt-8
1	2072	1767	-	-	-	-	-	-	-	-
2	2077	2061	1767	-	-	-	-	-	-	-
3	2078	2289	2061	1767	-	-	-	-	-	-
4	2043	2047	2289	2061	1767	-	-	-	-	-
5	2062	1856	2047	2289	2061	1767	-	-	-	-
6	2067	1842	1856	2047	2289	2061	1767	-	-	-
7	1964	1866	1842	1856	2047	2289	2061	1767	-	-
8	1981	2279	1866	1842	1856	2047	2289	2061	1767	-
9	1914	2688	2279	1866	1842	1856	2047	2289	2061	1767
10	1991	3264	2688	2279	1866	1842	1856	2047	2289	2061
11	2129	3896	3264	2688	2279	1866	1842	1856	2047	2289
12	2309	4014	3896	3264	2688	2279	1866	1842	1856	2047
13	2614	4041	4014	3896	3264	2688	2279	1866	1842	1856
14	2896	3710	4041	4014	3896	3264	2688	2279	1866	1842
15	3058	3383	3710	4041	4014	3896	3264	2688	2279	1866
16	3309	3431	3383	3710	4041	4014	3896	3264	2688	2279
17	3446	3613	3431	3383	3710	4041	4014	3896	3264	2688

3. UYGULAMA (APPLICATION)

Çalışmada kullanılan veri; Hill, Griffiths ve Judge'ın [20] kaynaklarında gecikmesi dağıtılmış model için kullandıkları gözlemlerden seçilmiştir. İki farklı örneklem sayısı ve iki farklı gecikme uzunluğu kullanılarak model tahmini için GME ve GME'ye alternatif Bayes tahmin edicileri kullanılmıştır. İlk önce örneklem büyüklüğü $n=4$ ve gecikme uzunluğu $lag=3$, daha sonra $n=9$ ve $lag=8$ olarak belirlenmiştir. Bu uygulamada yapılan tüm analizler MATLAB programında kod yazılarak yapılmıştır.(Tablo 1)

Bu çalışmada kullanılan gözlemlerle elde edilecek sonuç iki nedenden dolayı önem arz etmektedir. Birincisi $y=X\beta$ ilişkisinde bulunan X gözlemler matrisi ilk durumda 4×5 , ikinci durumda 9×10 boyutludur. Yani β vektörünün elemanlarını tahmin etmek için yeterli gözlem olmadığından problem, eksik sunum problemidir. Çoklu iç ilişki problemine neden olan bir diğer sorun ise kötü koşulluluk problemidir.

Kötü koşulluluk probleminin varlığının sınanması amacı ile koşul sayıları belirlenmiştir. Koşul sayıları belirlenirken öncelikle özdeğerler hesaplanmakta ardından, λ_{max} en büyük özdeğer ve λ_{min} en küçük özdeğer olmak üzere koşul sayısı $\sqrt{\lambda_{max}/\lambda_{min}}$ formülünden hesaplanmaktadır. Çalışmada her iki durum için de hesaplanan özdeğerler sırasıyla Tablo 2 ve Tablo 3'te sunulmuştur.

Tablo 2. İlk duruma ait özdeğerler (Eigenvalues of the first state)

Boyut	Özdeğerler
1	2,7546e-15
2	751,14
3	59965,27
4	26728,71
5	64370780,88

Tablo 3. İkinci duruma ait özdeğerler (Eigenvalues of the second state)

Boyut	Özdeğerler
1	-3,5458e-13
2	391,8460
3	1646,18
4	18818,48
5	37407,85
6	82677,25
7	101442,81
8	5814045,54
9	10311634,46
10	692053673,59

Elde edilen özdeğerler yardımı ile birinci durum için koşul sayısının 292,7413 ve ikinci durum için koşul sayısının 1329 olduğu görülür ki her iki durum için de sonuçlar kötü koşulluluk probleminin varlığını işaret eder.

Çoklu iç ilişki problemi tespit edildiğinde model EKK ile tahmin edilemez. Bunun yerine yanlış tahmin edicilerin kullanılması gerekmektedir. Bu doğrultuda model sırasıyla

GME'ye alternatif Bayes tahmin edicisi ve GME tahmin edicisi ile tahmin edilmiştir. Alternatif Bayes tahmin edicisi kullanılırken iki durum baz alınmıştır. Parametre tahminleri elde edildikten sonra tahmin edicilerin norm ve HKO değerleri sırasıyla Eş. 21 ve Eş. 22 eşitlikleri kullanılarak hesaplanmıştır.

$$norm(\hat{\beta}) = \hat{\beta}'\hat{\beta} \quad (21)$$

$$HKO(\hat{\beta}) = (\hat{\beta} - \beta)'(\hat{\beta} - \beta) \quad (22)$$

Alternatif Bayes tahmin edicisinde ilk olarak modelin hata terimi içermediği durum ele alınmıştır. β 'lar için önsel yoğunluklar aşağıdaki gibi belirlenmiş ve çalışma süresince sabit tutulmuştur. Önsel yoğunluklar belirlenirken tüm durumlar için Pukelsheim [21] tarafından önerilen $\pm 3\sigma$ kuralı uygulanmıştır.

- Birinci model için;

$$u=[-1,028 \ 1,068; -0,378 \ 0,534; -0,130 \ 0,560; -0,423 \ 0,508]$$

- İkinci model için;

$$u=[-1,210 \ 1,770; -0,412 \ 1,424; 0,046 \ 1,083; -0,370 \ 1,556; -0,5 \ 1,065; 0,377 \ 0,861; 0,113 \ 1,189; -0,291 \ 1,852; -1,078 \ 2,843]$$

- Birinci modele ait parametre tahminleri:

$$\hat{\beta} = \begin{bmatrix} 0,4351 \\ 0,1353 \\ 0,1615 \\ 0,3335 \end{bmatrix} \text{norm}(\hat{\beta})= 0,3450 \quad HKO(\hat{\beta})= 0,0085$$

- İkinci modele ait parametre tahminleri:

$$\hat{\beta} = \begin{bmatrix} -0,4530 \\ 0,3270 \\ 0,4740 \\ 0,1087 \\ -0,4528 \\ 0,4075 \\ 0,2301 \\ 0,1082 \\ 0,3700 \end{bmatrix} \text{norm}(\hat{\beta})= 1,1221 \quad HKO(\hat{\beta})= 0,2112$$

Alternatif Bayes tahmin edicisinin kullanıldığı ikinci durumda ise birinci durumda olduğu gibi β 'lar için aynı önsel yoğunluklar kullanılmıştır. Bu durumda parametrelerin gerçekleşecekleri değer aralığının orta noktasının seçimine olanak sağlayan simetrik üçgensel dağılım kullanılarak, tek ardil mod değerinin belirlenmesinde en yüksek ardil yoğunluk tahmin edicisi kullanılmıştır. Simetrik üçgensel dağılımının kullanılmasının nedeni, bilindiği gibi, verinin kısıtlı ve değişkenler arasındaki ilişkinin önceden bilindiği durumlarda kullanılmasıdır.

En yüksek ardıl yoğunluk tahmin edicisi, β 'lar üçgensel dağılıma sahip olmak üzere amaç fonksiyonu ve kısıtlar altında uygulandığında tamamen aynı sonuçlar elde edilmiştir.

Son olarak model GME tahmin edicisi ile tahmin edilmiştir. Burada Bayes'teki önsel yoğunluk destek matrislerinde bulunan değerler orta nokta olmak üzere hataların destek matrisleri (Z) ve y'lerin standart sapması baz alınarak $\pm 3\sigma$ kuralına göre ise V destek matrisi belirlenmiştir.

- Birinci modele ait parametre tahminleri:

$$\hat{\beta} = \begin{bmatrix} -0,3996 \\ 0,4561 \\ 0,0769 \\ 0,2198 \\ 0,2736 \end{bmatrix} \text{norm}(\hat{\beta}) = 0,3371 \quad \text{HKO}(\hat{\beta}) = 0,0320$$

- İkinci modele ait parametre tahminleri:

$$\hat{\beta} = \begin{bmatrix} -0,0016 \\ -0,4252 \\ 0,3408 \\ 0,4749 \\ -0,0166 \\ -0,3928 \\ 0,5090 \\ 0,2212 \\ -0,0125 \\ 0,3664 \end{bmatrix} \text{norm}(\hat{\beta}) = 1,1195 \quad \text{HKO}(\hat{\beta}) = 0,1924$$

Elde edilen bulgular incelendiğinde birinci model için GME'ye alternatif Bayes tahmin edicisinin norm değerinin 0,3450; GME tahmin edicinin norm değerinin ise 0,3371 olduğu görülmektedir. Bununla birlikte GME'ye alternatif Bayes tahmin edicisinin HKO değeri 0,0085; GME tahmin edicinin HKO değeri ise 0,0320 olarak elde edilmiştir. İkinci durumda ise GME'ye alternatif Bayes tahmin edicisinin norm değerinin 1,1221 ve GME tahmin edicinin norm değerinin ise 1,1195 olduğu, bununla birlikte GME'ye alternatif Bayes tahmin edicisinin HKO değerinin 0,2112 ve GME tahmin edicinin HKO değerinin ise 0,1924 olduğu görülmektedir.

Elde edilen değerlerin birbirlerine yakın olması Bayes tahmin edicinin GME tahmin ediciye alternatif olarak kullanılabileceğini doğrulamaktadır. Bununla birlikte az da olsa değerler arasında etkinlik anlamında kıyaslama yapabilecek farklılıklar bulunmaktadır. Daha az gözleme sahip ilk durumda Bayes tahmin edicisinin HKO değeri GME tahmin ediciden daha düşük çıksa da gözlem sayısı yani örneklem büyüklüğü arttıkça çoklu iç ilişki probleminin daha da artması ile birlikte her iki tahmin edicinin HKO değerleri artmıştır. Ancak GME tahmin edicinin HKO değeri alternatif Bayes tahmin ediciye göre daha az artmıştır ve ikinci durumda HKO değerleri kıyaslamasında GME tahmin edicinin HKO değerinin alternatif Bayes tahmin edicinin HKO değerinden daha küçük olduğu görülmüştür.

Bu sonuç çoklu iç ilişki probleminin derecesinin artmasına karşı GME tahmin edicinin daha dirençli olduğunu ortaya çıkarmaktadır. Her iki durumda da daha düşük norm değerine sahip GME tahmin edicinin daha büyük örneklemelerde daha düşük HKO değerine sahip olması ile GME tahmin edicinin alternatif Bayes tahmin ediciden daha etkin olduğu sonucuna ulaşılmaktadır.

4. SİMGELER (SYMBOLS)

y	: Açıklanan (bağımlı) değişken
x	: Açıklayıcı (bağımlı) değişken
B	: Model parametresi
$\hat{\beta}$: Parametre tahmini
y	: $T \times I$ boyutlu açıklanan değişken vektörü
X	: $T \times K$ boyutlu açıklayıcı değişkenler matrisi
β	: $K \times I$ boyutlu bilinmeyen parametreler vektörü
u	: $T \times I$ boyutlu hata vektörü
Z	: $K \times KM$ boyutlu kompakt destek matrisi
p	: $KM \times I$ boyutlu ağırlıklar vektörü
V	: $T \times TJ$ boyutlu destek noktaları matrisi
w	: $TJ \times I$ boyutlu bilinmeyen ağırlıklar vektörü
\otimes	: Kronecker çarpımı
$p(\beta)$: Önsel yoğunluk
$L(\beta y)$: Olabilirlik fonksiyonu
$h(\beta y)$: Ardıl yoğunluk
σ	: Standart hata
λ_{max}	: En büyük özdeğer
λ_{min}	: En küçük özdeğer

5. SONUÇLAR VE TARTIŞMALAR (RESULTS AND DISCUSSIONS)

GME yöntemi 2000'li yıllardan itibaren uygulanmaya başlanmış olup diğer yanlı tahmin edicilere göre etkinliği de gösterilmiştir. Ancak farklı regresyon modellerinde uygulanabilen GME yönteminin eksik sunumlu regresyon modellerinde uygulandığı çalışmalar oldukça azdır. Ayrıca Bayes ile ilgili yapılan çalışmalar, teknolojinin hızlı bir şekilde gelişmesi ile birlikte tüm mühendislik alanlarında kullanılmaya başlanmış olup teknolojiadaki gelişmeler ile birlikte kullanımı daha da artacaktır. Çalışmamızda, GME ve GME'ye alternatif Bayes yaklaşımının norm ve HKO açısından karşılaştırmalı olarak birlikte kullanılması ve bu tahmin tekniklerinin eksik sunumlu regresyon modellerinde uygulanmasının hem ulusal hem de uluslararası literatüre katkı sağlayacağı düşünülmektedir.

Bu çalışmada eksik sunumlu doğrusal deterministik regresyon modelinin tahminine GME ve Bayesçi yöntemle yaklaşılmış ve parametreler tahmin edilmiştir. Çalışmanın başlangıcında, hem eksik sunumluluk, hem de kötü koşulluluk problemlerinin yer aldığı regresyon modellerinin parametrelerinin GME ve GCE yaklaşımı ile tahmin edilebilirliği tartışılmış ve bunun için uygun önsel bilginin destek noktaları, referans olasılıklar ve entropi ölçütünün birlikte oluşturduğu etkiyle elde edilebileceği gösterilmiştir.

Çalışmanın uygulama bölümünde eksik sunumlu modellere GME ve Bayesçi yaklaşım ele alınmış, elde edilen parametre

tahminleri, norm değerleri ve HKO değerleri verilmiştir. EKK gibi klasik tahmin teknikleri kullanılarak tahmin edilemeyen bir model GME ve Bayesçi yaklaşım ile tahmin edilebilmiştir. Ele alınan her iki durum için de hem GME tahmin edicisinin normu ile Bayesçi yaklaşımın normu hem de GME tahmin edicisinin HKO değeri ile Bayesçi yaklaşımın HKO değeri birbirine oldukça yakın çıkmış olup elde edilen bulgular GME'ye alternatif olarak Bayesçi yaklaşımın kullanılabileceğini göstermektedir.

Ayrıca örneklem büyüklüğü arttıkça HKO değerinin daha az arttığı GME tahmin edicisinin her iki durumda da daha küçük norm değerine de sahip olması sonucu ile GME tahmin edicisinin alternatif Bayes yaklaşımına göre daha etkin olduğu söylenilebilir.

KAYNAKLAR (REFERENCES)

1. Ciavolino E. ve Calcagni A.A., Generalized Maximum Entropy (GME) estimation approach to fuzzy regression model, *Applied Soft Computing*, 38, 51–63, 2016.
2. Al-Nasser A.D., Two steps generalized maximum entropy estimation procedure for fitting linear regression when both covariates are subject to error, *Journal of Applied Statistics*, 41 (8), 1708–1720, 2014.
3. Grünwald P.D. ve Dawid A.P., Game theory, maximum entropy, minimum discrepancy and robust bayesian decision theory, *The Annals of Statistics*, 32 (4), 1367–1433, 2004.
4. Suhartanto D., Kusdiyol L., Chen B., Dean D., Setiawati L., Predicting consumer behaviour in tourism industry: comparing structural equation modelling (SEM) and multiple regression, *Materials Science and Engineering*, 830, 1-4, 2019.
5. Ciavolino E. ve Calcagni A.A., Generalized maximum entropy (GME) approach for crisp-input/fuzzy-output regression model, *Quality & Quantity*, 48, 3401–3414, 2014.
6. Macedo P. Ridge regression and generalized maximum entropy: an improved version of the ridge-gme parameter estimator, *Communications in Statistics - Simulation and Computation*, 46 (5), 3527-3539, 2017.
7. Kamar S.H.ve Msallam B.S., Comparative study between generalized maximum entropy and bayes methods to estimate the four parameter weibull growth model, *Journal of Probability and Statistics*, 2020, 1-7, 2020.
8. Heckelei T., Mittelhammer R., Jansson T., A Bayesian alternative to generalized cross entropy solutions for underdetermined econometric models. Discussion paper, Institute for Food and Resource Economics, University of Bonn, 2008.
9. Chinnakum W. ve Boonyasana P., Modelling thailand tourism demand: a dual generalized maximum entropy, *Thai Journal of Mathematics, Special Issue on Entropy in Econometrics*, 67–78, 2017.
10. Tarkhamtham P., Yamaka W., High-order generalized maximum entropy estimator in kink regression model, *Thai Journal of Mathematics, Special Issue: Structural Change Modeling and Optimization in Econometrics* 2018, 185-200, 2019.
11. Maneejuk P., Yamaka W., Sriboonchitta S., Entropy inference in smooth transition kink regression, *Communications in Statistics - Simulation and Computation*, 1-24, 2020.
12. Paris Q. ve Howitt R.E., An analysis of ill-posed production problems using maximum entropy, *American Journal of Agricultural Economics*, 80 (1), 124-138, 1998.
13. Golan A., Judge G., Miller D. Maximum Entropy Econometrics: Robust Estimation with Limited Data. John Wiley & Sons, New York, A.B.D., 1996.
14. Ramanathan R., *Introductory Econometrics with Applications*. Fourth Edition, Harcourt Brace College Publishers, 1998.
15. Shannon C.E., A mathematical theory of communication, *The Bell System Technical Journal*, 27 (3), 379-423, 1948.
16. Jaynes E.T. Information theory and statistical mechanics, *Physical Review*, 106 (4), 620-630, 1957.
17. Jaynes E.T. Information theory and statistical mechanics II, *Physics Review*, 108 (2), 171-190, 1957.
18. Campbell L.L., Minimum cross entropy estimation with inaccurate side information, *IEEE Transactions on Information Theory*, 45 (7), 2650-2652, 1999.
19. Zellner A., *Introduction to Bayesian Inference in Econometrics*. John Wiley & Sons, New York, A.B.D., 1971.
20. Hill R.C., Griffiths W.E., Judge G.G., *Undergraduate Econometrics*, John Wiley & Sons. New York, A.B.D., 2001.
21. Pukelsheim, F., The three sigma rule, *The American Statistician*, 48 (2), 88-91, 1994.

