İstatistik Araştırma Dergisi Cilt 04 No 01 Sayfa 75-91 Nisan 2005

BASİT DOĞRUSAL OTOREGRESİF MODELLER SİSTEMİNDE PARAMETRE TAHMİNİ VE HİPOTEZ TESTİ: SİMETRİK İNOVASYONLAR

Özlem TÜRKER* Ayşen D. AKKAYA**

ÖZET

Bu çalışmada, simetrik dağılıma sahip hata terimli basit otoregresif modeller sistemi incelenerek normal dağılım varsayımının geçersiz olduğu durumlardaki metodoloji birden fazla bağımsız bilgi kaynağı olduğu durumlara genellenmiş, uyarlanmış en çok olabilirlik yöntemi ile tahmin ediciler elde edilmiş ve parametre vektörünün tüm kaynaklar için değişip değişmediğini test edecek sağlam ve etkin test istatistikleri geliştirilmiştir. Elde edilen tahmin ediciler ve test istatistikleri bu alandaki uygulamalarda sıkça kullanılan en küçük kareler yöntemi ile elde edilen tahmin edici ve test istatistikleri ile karşılaştırılmış ve daha iyi sonuçlar verdiği görülmüştür.

Anahtar Kelimeler: *En çok olabilirlik tahmin edicileri, En küçük kareler tahmin edicileri, Uyarlanmış en çok olabilirlik tahmin edicileri, Otoregresyon, Sağlamlık, Simetrik dağılımlar.*

1. GİRİŞ

Özilintili inovasyonlara sahip basit doğrusal regresyonların ana problemi parametrelerin tahminidir ve bu alandaki çalışmaların çoğu normallik varsayımı altında yapılmıştır (Durbin, 1960; Schäffler, 1991). Oysaki artık uygulamada normal olmayan durumlarda daha çok karşılaşıldığı yönünde bir görüş vardır (Pearson, 1932; Elveback ve diğerleri, 1970; ve Tse, 1991). En çok olabilirlik (ML) yöntemi teoride normal olmayan dağılımlara genelleştirilebilmekle beraber pratikte sorunlar içermektedir. Matematiksel karmaşası nedeni ile normal olmayan durumlara genellenmesi mümkün olmamıştır. Dolayısı ile, bu alan uygulamalarında çoğunlukla en küçük kareler (LS) tahmin ediciler yöntemi kullanılmaktadır. Oysaki Tiku ve diğerleri (1999) ve Akkaya ve Tiku (2001a, b) uyarlanmış en çok olabilirlik (MML) yöntemi ile normal olmayan durumlara genellemeleri yapmışlar ve LS tahmin edicilerinin etkinliğinin düşük olduğunu ve örneklem sayısı, n, arttıkça MML tahmin edicilerine göre etkinliklerinin gittikçe azaldığını göstermişlerdir.

Ancak bu çalışmalarda tek bir bilgi kaynağı olduğu varsayılmış ve tek bir model kullanılmıştır. Oysaki günümüzde bilgi teknolojilerinin ilerlemesi ile artık birden fazla kaynaktan bilgi alınabilmektedir. Bu ise bu bilgilerin tutarlılığı ve bilgilerin birleştirilip

---

* Çankaya Üniversitesi, Mühendislik ve Mimarlık Fakültesi, Endüstri Mühendisliği Bölümü, Ankara, Türkiye email: ozlemt@cankaya.edu.tr

** ODTÜ, Fen Edebiyat Fakültesi, İstatistik Bölümü, Ankara, Türkiye.

birleştirilemeyeceği sorularını gündeme getirmektedir. Bunun için metodolojinin birden fazla bilgi kaynağı olduğu durumlara genellenmesi ve parametre vektörünün tüm kaynaklar için sabit kalıp kalmadığının test edilmesine olanak verecek şekilde geliştirilmesi önem kazanmaktadır.

Bu çalışmada, bahsedildiği gibi metodoloji genişletilmiş ve simetrik dağılımlar olan Student t ailesi için MML yöntemi kullanılarak parametre tahminleri yapılmıştır. MML tahmin edicilerinin hesaplanması kolay ve LS tahmin edicilerinden çok daha etkili olduğu görülmüştür. Ayrıca herbir bilgi kaynağının parametre vektörünün aynı olup olmadığını test etmek için ilgili test istatistikleri geliştirilmiş ve yine MML tahmin edicilerine dayanan test istatistiklerinin LS tahmin edicilerine dayanan test istatistiklerine göre daha güçlü ve sağlam olduğu görülmüştür.

2. PARAMETRELERİN TAHMİNİ

Aşağıda verildiği gibi k tane birbirinden bağımsız basit doğrusal otoregresif modeli ele alalım:

$$Y_{ij} - \phi_i y_{ij-1} = \mu_i + \delta_i (x_{ij} - \phi_i x_{ij-1}) + \varepsilon_{ij} \qquad (|\phi_i| < 1,\ 1 \le j \le n_i,\ 1 \le i \le k). \qquad (1)$$

Burada ε_{it} aşağıdaki tanımlanan uzun kuyruklu simetrik aileden gelen birbirine eşit veya farklı varyansa sahip bağımsız ve özdeş rassal inovasyonlardır:

$$f(\varepsilon) = \frac{1}{\sigma\sqrt{k}\beta(1/2, p-1/2)}\left(1 + \frac{\varepsilon^2}{k\sigma^2}\right)^{-p} \qquad -\infty < \varepsilon < \infty; \quad k = 2p-3. \qquad (2)$$

Burada $\beta(a,b)$ beta fonksiyonunu ifade etmektedir. Bu ailenin basıklığı, μ_4 / μ_2^2, sonlu p değeri için her zaman 3'den büyüktür. Öte yandan $p < 2$ için basıklık sonsuzdur ve $p = \infty$ olduğunda ise çok iyi bilinen normal dağılıma, $N(0, \sigma^2)$, dönüşmektedir. Ayrıca $t = \sqrt{(v/k)}(\varepsilon/\sigma)$, $v = 2p - 1$ serbestlik dereceli Student t dağılımına sahiptir.

Anlatımda kolaylık olması açısından $k = 2$ ve $n_1 = n_2$ olarak alınmıştır. Ayrıca p değerinin bilindiği varsayılmaktadır. Uygulamada p değeri bilinmemekle birlikte elde edilen tahmin ediciler ileride gösterileceği üzere p değerindeki olası sapmalara karşı sağlam olduğundan Q-Q grafiği ile belirlenen değer kullanılabilinir. Başlangıç değerleri için Vinod ve Shenton (1996) iki model önermektedir:

Model A: Başlangıç değeri sabit olarak alınır. Özellikle $Y_0 = 0$.

Model B: Rassaldır; $Y_0 = \varepsilon_0 / \sqrt{1-\phi^2}$.

Bu çalışmada daha genel olduğundan B Modeli kullanılmıştır.

Başlangıç değerlerine (y_{i0} ve x_{i0}) koşullu olabilirlik fonksiyonu aşağıdaki gibidir:

$$L \propto \left(\frac{1}{\sigma_1\sigma_2}\right)^n \left[\prod_{t=1}^{n}\left(1 + \frac{z_{1j}^2}{k}\right)^{-p} \prod_{t=1}^{n}\left(1 + \frac{z_{2j}^2}{k}\right)^{-p}\right]. \qquad (3)$$

Burada $z_{ij} = \{(y_{itj} - \phi_i y_{ij-1}) - \mu_i - \delta_i (x_{ij} - \phi_i x_{itj-1})\}/\sigma_i$.

Buradan elde edilecek olabilirlik fonksiyonlarının $p < \infty$ için çok köklü olduğu bilinmekle birlikte kök sayısı bilinmemekte ve n sayısı arttıkça kök sayısı da artmaktadır (Vaughan, 1992). Dolayısı ile ML tahmin edicileri açık bir formda bulunamamakta ve olabilirlik denklemlerinin iteratif yöntemlerle çözülmesi gerekmektedir. Oysa yapılan çalışmalar bu yöntemlerin de yanlış değerlere yakınsama veya hiç yakınsamama gibi problemleri olduğunu göstermektedir (Tiku ve diğerleri, 1986; Puthenpura ve Sinha, 1986; Tiku ve Suresh, 1992; Tiku ve diğerleri, 1999). Bu sorunlar $\phi_i = 0$ durumunda yani modelin basit doğrusal regresyon olması durumunda dahi Islam ve diğerleri (2001), Tiku ve diğerleri (2001) ve Islam ve Tiku (2004) tarafından da gözlemlenmiştir. Bahsi geçen sorunların aşılması için bu çalışmada uyarlanmış en çok olabilirlik (MML) tahmin yöntemi kullanılmıştır. Bu yöntem Tiku (1967) ve Tiku ve Suresh (1992) çalışmalarına dayanmakta olup adımları şöyledir:

(i) olabilirlik denklemleri sıralı istatistikler ile ifade edilir,

(ii) Taylor serileri açılımının ilk iki terimi kullanılarak olabilirlik denklemlerindeki sorun yaratan (doğrusal olmayan) terimler doğrusallaştırılır,

(iii) elde edilen yeni denklemler MML tahmin edicilerini elde etmek için çözülür.

Bu yöntem doğrultusunda hata teriminin, ε_{ij}, (verilen μ_i, δ_i ve ϕ_i değerleri için) küçükten büyüğe sıralanması ile elde edilen sıralanmış z_{ij} değerlerini aşağıdaki gibi tanımlayalım:

$$z_{i(j)} = \varepsilon_{i(j)}/\sigma = \{w_{i(j)} - \mu_i\} / \sigma; \qquad w_{i(j)} = (y_{i[j]} - \phi_i\, y_{i[j]-1}) - \delta_i(x_{i[j]} - \phi_i\, x_{i[j]-1}),$$

$$z_{i(1)} \le z_{i(2)} \le \ldots \le z_{i(n)} \qquad (i = 1, 2). \tag{4}$$

Burada $\left(y_{i[j]}, y_{i[j]-1}, x_{i[j]}, x_{i[j]-1}\right)$, j'inci sıralı hata terimine, $\varepsilon_{i(j)}$, ait $\left(y_{ij}, y_{ij-1}, x_{ij}, x_{ij-1}\right)$ gözlemlerini ifade etmektedir. Tüm örneklem toplamları sıralamadan etkilenmeyeceği için olabilirlik denklemlerinde z_{ij} yerine $z_{i(j)}$ yazılır. Ardından adım (ii)'de belirtildiği gibi $g_i\{z_{i(j)}\}$ terimi $t_{i(j)} = E(z_{i(j)})$ etrafında doğrusallaştırılır (Tiku ve Suresh, 1992):

$$g_i(z_{i(j)}) \cong g_i(t_{i(j)}) + \left[z_{i(j)} - t_{i(j)}\right]\left\{\frac{\partial g_i(z_{i(j)})}{\partial z_{i(j)}}\right\}_{z_{i(j)} = t_{i(j)}} = \alpha_{ij} + \beta_{ij} z_{i(j)} \qquad 1 \le j \le n; \quad i = 1,2$$

$$\alpha_{ij} = \frac{(2/k)t_{i(j)}^3}{\left[1 + (1/k)t_{i(j)}^2\right]^2} \qquad \text{ve} \qquad \beta_{ij} = \frac{1 - (1/k)t_{i(j)}^2}{\left[1 + (1/k)t_{i(j)}^2\right]^2}. \tag{5}$$

Buradaki $t_{i(j)}$ değerleri $n \le 20$ için Tiku ve Kumra (1981) çalışmasında mevcuttur. $n \ge 10$ için ise parametrelerin etkinliğinde önemli bir etkisi olmayan aşağıdaki denklemden elde edilen yaklaşık değerler kullanılır:

$$\frac{1}{\sqrt{k}\beta(1/2, p - 1/2)} \int_{-\infty}^{t_{i(j)}} \left(1 + \frac{z^2}{k}\right)^{-p} dz = \frac{j}{n+1} \qquad 1 \le j \le n\,; \quad i = 1, 2. \tag{6}$$

Dağılımın simetrik olması nedeniyle $t_{i(j)} = - t_{i(n-j+1)}$ olur ve $\alpha_{ij} = -\alpha_{i(n-j+1)}$ olacağından $\sum_{i=1}^{n} \alpha_{ij} = 0$ ve $\beta_{ij} = \beta_{i(n-j+1)}$ olur. Olabilirlik fonksiyonlarındaki $g_i(z_{i(j)})$ teriminin doğrusallaştırılmış ifadesi ile yer değiştirilmesi sonucu elde edilen uyarlanmış olabilirlik denklemlerinin ($\partial \ln L^* / \partial \mu_i$, $\partial \ln L^* / \partial \delta_i$, $\partial \ln L^* / \partial \phi_i$ ve $\partial \ln L^* / \partial \sigma$) çözülmesi ile aşağıda verilen MML tahmin edicileri elde edilir:

$$\hat{\mu}_i = \bar{v}_{i[.]} - \hat{\delta}_i \bar{u}_{i[.]}, \quad \hat{\delta}_i = G_i + H_i \hat{\sigma}, \quad \hat{\phi}_i = K_i + D_i \hat{\sigma}, \quad \hat{\sigma} = \left(B + \sqrt{B^2 + 8nC}\right) \Big/ 4n . \quad (7)$$

Burada
$$G_i = \frac{\sum_{j=1}^{n} \beta_{ij} \left(u_{i[j]} - \bar{u}_{i[.]}\right) v_{i[j]}}{\sum_{j=1}^{n} \beta_{ij} \left(u_{i[j]} - \bar{u}_{i[.]}\right)^2}, \quad H_i = \frac{\sum_{j=1}^{n} \alpha_{ij} u_{i[j]}}{\sum_{j=1}^{n} \beta_{ij} \left(u_{i[j]} - \bar{u}_{i[.]}\right)^2}, \quad (8)$$

$$K_i = \frac{\left[\begin{array}{l} \sum_{j=1}^{n} \beta_{ij} \left(y_{i[j]} - \hat{\delta}_i x_{i[j]}\right)\left(y_{i[j]-1} - \hat{\delta}_i x_{i[j]-1}\right) \\ - \frac{1}{m_i} \sum_{j=1}^{n} \beta_{ij} \left(y_{i[j]-1} - \hat{\delta}_i x_{i[j]-1}\right) \sum_{j=1}^{n} \beta_{ij} \left(y_{i[j]} - \hat{\delta}_i x_{i[j]}\right) \end{array}\right]}{\sum_{j=1}^{n} \beta_{ij} \left(y_{i[j]-1} - \hat{\delta}_i x_{i[j]-1}\right)^2 - \frac{1}{m_i} \left\{\sum_{j=1}^{n} \beta_{ij} \left(y_{i[j]-1} - \hat{\delta}_i x_{i[j]-1}\right)\right\}^2} \quad (9)$$

$$D_i = \frac{\sum_{j=1}^{n} \alpha_{ij} \left(y_{i[j]-1} - \hat{\delta}_i x_{i[j]-1}\right)}{\sum_{j=1}^{n} \beta_{ij} \left(y_{i[j]-1} - \hat{\delta}_i x_{i[j]-1}\right)^2 - \frac{1}{m_i} \left\{\sum_{j=1}^{n} \beta_{ij} \left(y_{i[j]-1} - \hat{\delta}_i x_{i[j]-1}\right)\right\}^2} \quad (i=1, 2), \quad (10)$$

$$B = \frac{2p}{k} \left[\sum_{j=1}^{n} \alpha_{1j} \left\{v_{1[j]} - G_1 u_{1[j]}\right\} + \sum_{j=1}^{n} \alpha_{2j} \left\{v_{2[j]} - G_2 u_{2[j]}\right\}\right], \quad (11)$$

$$C = \frac{2p}{k} \left[\begin{array}{l} \left\{\sum_{j=1}^{n} \beta_{1j} \left(v_{1[j]} - \bar{v}_{1[.]}\right)^2 - G_1 \sum_{j=1}^{n} \beta_{1j} v_{1[j]} \left(u_{1[j]} - \bar{u}_{1[.]}\right)\right\} \\ + \left\{\sum_{j=1}^{n} \beta_{2j} \left(v_{2[j]} - \bar{v}_{2[.]}\right)^2 - G_2 \sum_{j=1}^{n} \beta_{2j} v_{2[j]} \left(u_{2[j]} - \bar{u}_{2[.]}\right)\right\} \end{array}\right]. \quad (12)$$

Burada $m_i = \sum_{j=1}^{n} \beta_{ij}$ (i=1, 2), $v_{i[j]} = y_{i[j]} - \phi_i\, y_{i[j]-1}$, $u_{i[j]} = x_{i[j]} - \phi_i\, x_{i[j]-1}$ $(1 \le j \le n, i=1,2)$,

$$\bar{v}_{i[.]} = \frac{1}{m_i} \sum_{j=1}^{n} \beta_{ij} v_{i[j]} , \quad \bar{u}_{i[.]} = \frac{1}{m_i} \sum_{j=1}^{n} \beta_{ij} u_{i[j]} \quad (13)$$

olarak tanımlanmaktadır.

Hesaplamalar: Akkaya ve Tiku (2001a, b)'de önerilen yöntemdeki gibi $\gamma_i = -\delta_i \phi_i$ (i = 1, 2) şeklinde tanımlanır ve başlangıç değerleri parametrelerin LS tahmin edicileri ile hesaplanır. Elde edilen başlangıç değerleri kullanılarak denklem (7)'deki gibi $\hat{\delta}_i$ (i = 1, 2) ve $\hat{\sigma}$ hesaplanır. Ardından yine aynı denklemdeki gibi $\hat{\phi}_i$ (i = 1, 2) elde edilir. Daha sonra başlangıç değerlerini hesaplanan MML tahmin değerleri, $\hat{\delta}_i$, $\hat{\phi}_i$ ve $-\hat{\delta}_i \hat{\phi}_i$ (i = 1, 2), ile değiştirerek ikinci bir iterasyon yapılır. Son MML tahmin değerleri, $\hat{\delta}_i$, $\hat{\phi}_i$ (i = 1, 2) ve $\hat{\sigma}$, üçüncü iterasyonun sonunda elde edilir. En son olarak yine denklem (7)'den $\hat{\mu}_i$ (i = 1, 2), MML tahmini değerleri, elde edilir. Yapılan analizler sonucu 3 iterasyonun parametrelerin sabitlenmesi için yeterli olduğu görülmüştür.

Hatırlatma: β_i $(1 \le i \le n)$ katsayıları orta noktaya kadar artan daha sonra azalan simetrik bir yapı göstermektedir. Dolayısı ile eğer ilk değeri pozitif ise tüm değerleri pozitif olmakta ve bu da $\hat{\sigma}$'in olması gerektiği gibi reel ve pozitif olmasını sağlamaktadır. Oysaki küçük p değerleri ve büyük örneklem sayısı için ilk birkaç değer negatif olabilmekte ve bu da $\hat{\sigma}$'in reel ve pozitif olmasını engellemektedir. Bunu önlemek için eğer denklem (12)'de verilen C değeri negatif ise α_i ve β_i sırasıyla $\alpha_i^* = 0$ ve $\beta_i^* = 1 / \left\{1 + (1/k)t_{(i)}^2\right\}$ şeklinde tanımlanır ve bu yeni değerler ile MML tahmin edicileri hesaplanır. Böylece $\hat{\sigma}$'in her zaman reel ve pozitif olması sağlanmış olunur. $z_{(i)} - t_{(i)} = 0$ ve $\alpha_i + \beta_i z_{(i)} \cong \alpha_i^* + \beta_i^* z_{(i)}$ olduğundan bu değişim MML tahmin edicilerinin asimptotik özelliklerini değiştirmemektedir (Tiku ve diğerleri, 2001).

Hatırlatma: LS tahmin edicileri MML tahmin edicilerinin özel durumudur. MML tahmin edicilerinde $\alpha_{ij} = 0$ ve $\beta_{ij} = 1$ olduğunda LS tahmin edicileri elde edilmektedir.

3. ASİMPTOTİK ÖZELLİKLER

Örneklem sayısı sonsuza yaklaştıkça $[g_i(z_{i(j)}) - (\alpha_{ij} + \beta_{ij} z_{i(j)})]$farkları sıfıra yakınsar. Bunun sonucunda $(1/n)[(\partial \ln L/\partial\mu_i) - (\partial \ln L^*/\partial\mu_i)]$ ve $(1/n)[(\partial \ln L/\partial\delta_i) - (\partial \ln L^*/\partial\delta_i)]$ (i = 1, 2) asimptotik olarak sıfıra eşit olur (Vaughan ve Tiku, 2000). Dolayısı ile MML tahmin edicileri asimptotik olarak ML tahmin edicilerine denktir (Bhattacharya, 1985; Vaughan ve Tiku, 2000).

1) $\partial \ln L^* / \partial\mu_i$ (i = 1, 2) aşağıdaki formda yazılabildiğinden

$$\frac{\partial \ln L}{\partial \mu_i} \cong \frac{\partial \ln L^*}{\partial \mu_i} = \frac{2m_i p}{\sigma^2}\{\hat{\mu}_i(\delta_i, \phi_i) - \mu_i\} \quad (i = 1, 2); \tag{14}$$

$\hat{\mu}_i(\delta_i, \phi_i) = \bar{v}_{i[.]} - \delta_i \bar{u}_{i[.]}$, MML tahmin edicileri, $\hat{\mu}_i(\delta_i, \phi_i)$ (i = 1, 2), δ_i ve ϕ_i biliniyorken μ_i (i = 1, 2) için minimum varyanslı, $Var[\hat{\mu}_i(\delta_i, \phi_i)] = k\sigma^2 / 2m_i p$ (i = 1, 2), tahmin edicilerdir.

2) Benzer şekilde

$$\frac{\partial \ln L}{\partial \delta_i} \cong \frac{\partial \ln L^*}{\partial \delta_i} = \frac{2m_i p}{k\sigma^2}\left[\frac{1}{m_i}\sum_{j=1}^{n}\beta_{ij}\left(u_{i[j]}-\bar{u}_{i[.]}\right)^2\right]\left(\hat{\delta}_i(\phi_i,\sigma)-\delta_i\right), \; (i=1,2) \qquad (15)$$

$\hat{\delta}_i(\phi_i, \sigma) = G_i + H_i\,\sigma$, olarak yazılabildiğinden MML tahmin edicileri, $\hat{\delta}_i(\phi_i, \sigma)$ (i=1,2), δ_i ve ϕ_i biliniyorken minimum varyanslı, $\mathrm{Var}[\hat{\delta}_i(\phi_i,\sigma)] = \frac{k}{2p}\frac{\sigma^2}{\sum_{j=1}^{n}\beta_{ij}\left(u_{i[j]}-\bar{u}_{i[.]}\right)^2}$ (i = 1, 2), tahmin edicilerdir.

4. TAHMİN EDİCİLERİN ETKİNLİKLERİ

Uygulamada yaygın olarak LS tahmin edicileri kullanılmaktadır (Velu and Gregory 1987; Maller 1989; Nagaraja ve Fuller 1992). Ancak bu bölümde elde edilen sonuçlar LS tahmin edicilerinin MML tahmin edicilerine göre daha az etkili olduğunu göstermektedir.

Tablo 1. (1) Beklenen Değer, (2) Varyans, ve LS tahmin edicilerinin göreli etkinlik değerleri, RE, ($\phi_1 = \phi_2 = 0.2$)

n = 30									
p		δ_1	δ_2	ϕ_1	ϕ_2	σ_1	σ_2	μ_1	μ_2
2.0	LS	0.9999	0.9971	0.1496	0.1622	0.8835	1.3312	0.0026	-.0099
(1)	MML	1.0027	0.9975	0.1610	0.1705	1.1126	1.6799	0.0027	-.0049
	LS	0.0309	0.0289	0.0291	0.0299	0.1109	0.3304	0.0396	0.0917
(2)	MML	0.0191	0.0183	0.0253	0.0246	0.1309	0.3784	0.0251	0.0567
	RE	**61.8**	**63.3**	**86.9**	**82.3**	**118.0**	**114.5**	**63.4**	**61.8**
3.5	LS	1.0030	1.0001	0.1648	0.1511	0.9369	1.4016	0.0002	-.0020
(1)	MML	1.0036	1.0011	0.1685	0.1559	0.9746	1.4624	0.0000	-.0018
	LS	0.0284	0.0388	0.0312	0.0331	0.0348	0.0715	0.0386	0.0882
(2)	MML	0.0249	0.0343	0.0298	0.0318	0.0268	0.0616	0.0339	0.0773
	RE	**87.7**	**88.4**	**95.5**	**96.1**	**77.0**	**86.2**	**87.8**	**87.6**
5.0	LS	1.0001	1.0002	0.1548	0.1521	0.9350	1.4069	0.0011	0.0045
(1)	MML	1.0000	1.0003	0.1558	0.1541	0.9630	1.4488	0.0009	0.0051
	LS	0.0252	0.0341	0.0336	0.0337	0.0236	0.0519	0.0386	0.0858
(2)	MML	0.0238	0.0324	0.0327	0.0330	0.0230	0.0508	0.0367	0.0811
	RE	**94.4**	**95.0**	**97.3**	**97.9**	**97.5**	**97.9**	**95.1**	**94.5**
n = 100									
2.0	LS	0.9978	0.9968	0.1876	0.1843	0.9284	1.4298	-.0020	-.0013
(1)	MML	0.9973	0.9969	0.1930	0.1890	1.0609	1.6157	-.0017	-.0028
	LS	0.0100	0.0089	0.0084	0.0091	0.0398	0.1766	0.0102	0.0251
(2)	MML	0.0060	0.0049	0.0058	0.0066	0.0212	0.0853	0.0062	0.0143
	RE	**60.0**	**55.1**	**69.0**	**72.5**	**53.3**	**48.3**	**60.8**	**57.0**
3.5	LS	0.9990	0.9999	0.1889	0.1907	0.9743	1.4702	-.0010	-.0028
(1)	MML	0.9959	1.0013	0.1898	0.1915	0.9879	1.4844	-.0016	0.0001
	LS	0.0116	0.0097	0.0093	0.0094	0.0094	0.0223	0.0096	0.0231
(2)	MML	0.0099	0.0087	0.0083	0.0085	0.0069	0.0166	0.0083	0.0202
	RE	**85.3**	**89.7**	**89.2**	**90.4**	**73.4**	**74.4**	**86.5**	**87.4**
5.0	LS	0.9962	0.9990	0.1864	0.1868	0.9773	1.4715	-.0003	0.0049
(1)	MML	0.9956	0.9997	0.1870	0.1881	0.9880	1.4868	0.0003	0.0053
	LS	0.0092	0.0085	0.0095	0.0089	0.0077	0.0181	0.0098	0.0239
(2)	MML	0.0090	0.0082	0.0091	0.0084	0.0068	0.0161	0.0092	0.0223
	RE	**97.8**	**96.5**	**95.8**	**94.4**	**88.3**	**89.0**	**93.9**	**93.3**

Tablo 1'de [100000/n] Monte Carlo tekrarından elde edilen parametre tahmin edicilerinin beklenen değerlerinin, varyanslarının ve aşağıdaki şekilde tanımlanan LS tahmin edicilerinin göreli etkinliğinin simüle değerleri verilmektedir:

RE = 100(MML tahmin edicilerinin varyansı / LS tahmin edicilerinin varyansı). (16)

Simülasyonlarda genelleştirmede herhangibir kayıp olmaksızın $\mu_1 = \mu_2 = 0$, $\delta_1 = 1$, $\sigma_1 = 1$ ve $\sigma_2 = 1.5$ olarak alınmıştır. Sonuçlar diğer ϕ_i değerleri (negatif, pozitif ve/veya birbirinden farklı) ve p değerleri için benzer olduğundan burada sadece $\phi_1 = \phi_2 = 0.2$ ve $p = 2.0$, 3.5 ve 5.0 değerleri raporlanmıştır. X-değerleri düzgün dağılımdan, U(0, 1), elde edilmiştir. Ayrıca X-değerlerinin normal dağılımdan elde edilmesi durumundaki sonuçlar da incelenmiş ve benzer sonuçlar elde edilmiştir.

Tablo 1'den görüleceği gibi MML tahmin edicileri LS tahmin edicilerine göre daha etkilidirler. Hatta dikkat çeken genel olarak örneklem sayısı arttıkça LS tahmin edicilerinin etkinliğinin daha çok azalmakta olmasıdır.

5. HİPOTEZ TESTİ

Buradaki ana ilgi değişik bilgi kaynaklarından elde edilen bilgilerde açıklayıcı değişkenin bağımlı değişkene etkisinin aynı olup olmadığı yani $H_0 : \delta_1 = \delta_2$ hipotezinin test edilmesidir. Bunun için iki durum incelenmiştir: 1) $\sigma_1 = \sigma_2$, ve 2) $\sigma_1 \neq \sigma_2$.

Durum 1) H_0: $\delta_1 = \delta_2$ hipotezini test etmek için,

$$t = \frac{\hat{\delta}_1 - \hat{\delta}_2}{\hat{\sigma}\sqrt{\dfrac{p-3/2}{p\sum_{j=1}^{n}\beta_{1j}\left(\hat{u}_{1[j]} - \bar{\hat{u}}_{1[.]}\right)^2} + \dfrac{p-3/2}{p\sum_{j=1}^{n}\beta_{2j}\left(\hat{u}_{2[j]} - \bar{\hat{u}}_{2[.]}\right)^2}}}; \tag{17}$$

$\hat{u}_{ij} = x_{ij} - \hat{\phi}_i x_{ij-1}$ $(i = 1, 2)$, test istatistiği kullanılır. Burada $\hat{\sigma}$, $\hat{\sigma} = \sqrt{\left(\hat{\sigma}_1^2 + \hat{\sigma}_2^2\right)/2}$ şeklinde hesaplanan σ'nın birleştirilmiş tahmin edicisidir.

$\hat{\phi}_i$ $(i = 1, 2)$ ve $\hat{\sigma}$, n sonsuza yaklaştıkça sırasıyla ϕ_i $(i = 1, 2)$ ve σ'ya yakınsadığından test istatistiği, t, asimptotik olarak normal, N(0, 1), dağılır ve asimptotik güç fonksiyonu $P\{Z \geq z_\alpha - |\lambda|\}$ şeklindedir. Burada z_α standart normal değişkeninin, Z, %100(1 - α) noktası ve

$$\lambda^2 = \frac{(\delta_1 - \delta_2)^2}{\sigma^2\left(\dfrac{p-3/2}{p\sum_{j=1}^{n}\beta_{1j}\left(u_{1[j]} - \bar{u}_{1[.]}\right)^2} + \dfrac{p-3/2}{p\sum_{j=1}^{n}\beta_{2j}\left(u_{2[j]} - \bar{u}_{2[.]}\right)^2}\right)} \tag{18}$$

merkezsizlik parametresidir.

Bu sonuçlar Slutsky teoremine ve MML denklemlerinin asimptotik olarak ML denklemlerine denk olmasına dayanmaktadır (Bhattacharya, 1985; Vaughan ve Tiku, 2000).

Durum 2) H_0: $\delta_1 = \delta_2$ hipotezini test etmek için,

$$t = \frac{\hat{\delta}_1 - \hat{\delta}_2}{\sqrt{\dfrac{\hat{\sigma}_1^2(p-3/2)}{p\sum_{j=1}^{n}\beta_{1j}\left(\hat{u}_{1[j]} - \bar{\hat{u}}_{1[.]}\right)^2} + \dfrac{\hat{\sigma}_2^2(p-3/2)}{p\sum_{j=1}^{n}\beta_{2j}\left(\hat{u}_{2[j]} - \bar{\hat{u}}_{2[.]}\right)^2}}}; \qquad (19)$$

$\hat{u}_{ij} = x_{ij} - \hat{\phi}_i\, x_{ij-1} (i=1,2)$, test istatistiği kullanılır. Durum 1'deki aynı sebepten, t test istatistiği asimptotik olarak N(0, 1) dağılır ve asimptotik güç fonksiyonu Prob$\{Z \geq z_\alpha - |\lambda|\}$ şeklinde olur. Burada z_α standart normal değişkeninin, Z, %100(1-α) noktası ve

$$\lambda^2 = \frac{(\delta_1 - \delta_2)^2}{\left(\dfrac{\sigma_1^2(p-3/2)}{p\sum_{j=1}^{n}\beta_{1j}\left(u_{1[j]} - \bar{u}_{1[.]}\right)^2} + \dfrac{\sigma_2^2(p-3/2)}{p\sum_{j=1}^{n}\beta_{2j}\left(u_{2[j]} - \bar{u}_{2[.]}\right)^2}\right)} \qquad (20)$$

merkezsizlik parametresidir.

6. SAĞLAMLIK

Genel olarak tahmin yöntemlerinde belli bir dağılım varsayılır ve tüm istatistiksel analizler bu dağılım varsayımı altında optimal hale getirilir. Ancak uygulamada bu varsayılan dağılım tamamen doğru olmayabilir ve bu dağılımdan olası farklılıklar görülmesi olağandır. Üstelik veri aykırı değerler de içerebilir. Tüm bunlar sağlamlık kavramının gündeme gelmesine neden olmuştur. θ'nın herhangi bir tahmin edicisi, $\hat{\theta}$, varsayılan dağılım altında tamamen etkili (yansız ve varyansı Cramer-Rao minimum varyans sınırına yakınsıyor) ve olası alternatif dağılımlar için de etkililiği yüksek ise bu tahmin edici sağlam olarak tanımlanır. Herhangi bir testin, $H_0 : \theta = 0$ gibi, birinci tip hatası olası alternatifler için daha önceden belirlenmiş seviyeyi aşmıyor ise bu test ölçüt sağlam olarak tanımlanır ve eğer hem varsayılan dağılım hem de olası alternatifleri için testin gücü yüksek ise test etkin sağlam olarak tanımlanır (Tiku ve diğerleri, 1986 önsöz).

Önceki bölümlerde elde edilen MML tahmin edicilerinin sağlamlığını göstermek için varsayılan dağılım olarak p = 3.5 alınmış ve sonuçlar LS tahmin edicileri ile karşılaştırılmıştır. Oldukça geniş bir alternatif kümesi için inceleme yapılmıştır:

(a) Dağılımın yanlış tanımlanması:

(1) p = 2.0, **(2)** p = 2.5, **(3)** p = 5 ve **(4)** p = ∞ (normal dağılım)

(b) Dixon'ın aykırı değer modeli

(n-r) gözlem f(3.5, σ) ve r (hangileri olduğu bilinmiyor) gözlem

(5) f(3.5, 2σ), **(6)** f(3.5, 4σ), ve **(7)** f(∞, 4σ), r = [0.1n + 0.5] (tam sayı) dağılımdan geliyor.

(c) Karışık model:

(8) 0.90 f(3.5, σ) + 0.10 f(3.5, 4σ)

(d) Kontaminasyon modeli:

(9) 0.90 f(3.5, σ) + 0.10 U(-0.5, 0.5).

Tablo 2. (1) Beklenen Değer, (2) LS ve MML Tahmin edicilerinin varyansı

		δ_1	δ_2	ϕ_1	ϕ_2	σ_1	σ_2	μ_1	μ_2
		Model (1)							
(1)	LS	1.0014	1.0016	0.4382	0.4212	0.8777	1.2995	0.0053	0.0018
	MML	1.0006	1.0043	0.4501	0.4348	0.8496	1.2618	0.0049	0.0051
(2)	LS	0.0190	0.0264	0.0247	0.0283	0.1120	0.2222	0.0429	0.1097
	MML	0.0112	0.0170	0.0195	0.0240	0.0464	0.0981	0.0263	0.0680
		Model (2)							
(1)	LS	1.0035	0.9977	0.4354	0.4301	0.9176	1.3730	0.0051	-.0041
	MML	1.0024	0.9989	0.4440	0.4382	0.9261	1.3876	0.0041	-.0056
(2)	LS	0.0195	0.0225	0.0260	0.0269	0.0604	0.1350	0.0466	0.1081
	MML	0.0147	0.0167	0.0223	0.0237	0.0343	0.0800	0.0347	0.0813
		Model (3)							
(1)	LS	1.0009	0.9956	0.4251	0.4188	0.9342	1.4003	-.0073	0.0123
	MML	1.0007	0.9951	0.4294	0.4234	0.9908	1.4851	-.0064	0.0105
(2)	LS	0.0197	0.0289	0.0273	0.0297	0.0240	0.0538	0.0473	0.1121
	MML	0.0186	0.0284	0.0266	0.0288	0.0243	0.0555	0.0442	0.1062
		Model (4)							
(1)	LS	1.0061	0.9978	0.4212	0.4318	0.9428	1.4080	-.0081	0.0049
	MML	1.0056	0.9966	0.4216	0.4331	1.0229	1.5272	-.0085	0.0023
(2)	LS	0.0175	0.0204	0.0289	0.0276	0.0160	0.0381	0.0454	0.1025
	MML	0.0185	0.0217	0.0302	0.0290	0.0199	0.0476	0.0475	0.1059
		Model (5)							
(1)	LS	1.0006	1.0075	0.4264	0.4305	0.9285	1.4020	0.0084	-.0041
	MML	1.0003	1.0056	0.4310	0.4374	0.9692	1.4575	0.0079	-.0030
(2)	LS	0.0260	0.0301	0.0303	0.0271	0.0328	0.0838	0.0468	0.1136
	MML	0.0225	0.0260	0.0289	0.0252	0.0272	0.0639	0.0412	0.0995
		Model (6)							
(1)	LS	1.0001	1.0055	0.4245	0.4241	0.9299	1.3987	0.0033	0.0068
	MML	1.0016	1.0049	0.4305	0.4313	0.9718	1.4584	0.0028	0.0089
(2)	LS	0.0218	0.0292	0.0303	0.0285	0.0326	0.0760	0.0443	0.1112
	MML	0.0192	0.0259	0.0287	0.0269	0.0278	0.0623	0.0387	0.0971
		Model (7)							
(1)	LS	1.0040	1.0008	0.4281	0.4235	0.9441	1.4125	-.0080	-.0009
	MML	1.0048	1.0008	0.4289	0.4249	1.0250	1.5340	-.0069	-.0023
(2)	LS	0.0135	0.0226	0.0273	0.0306	0.0167	0.0375	0.0482	0.1099
	MML	0.0144	0.0240	0.0287	0.0325	0.0209	0.0468	0.0495	0.1146
		Model (8)							
(1)	LS	0.9980	1.0026	0.4286	0.4244	1.4321	2.1094	0.0002	-.0127
	MML	0.9990	0.9996	0.4456	0.4427	1.3250	1.9688	0.0007	-.0090
(2)	LS	0.0459	0.0766	0.0262	0.0287	0.2963	0.5759	0.1248	0.2575
	MML	0.0226	0.0408	0.0201	0.0222	0.1388	0.2827	0.0633	0.1360
		Model (9)							
(1)	LS	0.9951	0.9978	0.4243	0.4327	0.8880	1.3332	-.0056	0.0051
	MML	0.9949	0.9977	0.4329	0.4393	0.9144	1.3701	-.0048	0.0065
(2)	LS	0.0194	0.0131	0.0274	0.0271	0.0343	0.0786	0.0428	0.0941
	MML	0.0165	0.0109	0.0247	0.0243	0.0282	0.0610	0.0349	0.0774

Tablo 2'de n=30, ϕ_1 = ϕ_2 = 0.5 için simulasyon sonuçları verilmektedir. Simulasyonlarda genelleştirmede herhangibir kayıp olmaksızın $\mu_1 = \mu_2 = 0$, $\delta_1 = 1$, $\sigma_1 = 1$ ve $\sigma_2 = 1.5$ olarak alınmıştır. Yine diğer tüm değerlerde de benzer sonuçlar elde edildiğinden burada raporlanmamıştır. Görüleceği üzere MML tahmin edicileri oldukça sağlamdır.

Yine aynı alternatif modeller için MML tahmin edicilerine dayalı test istatistiklerinin sağlamlığı araştırılmış ve LS tahmin edicilerine dayalı test istatistikleri ile karşılaştırılmıştır. Örnek teşkil etmesi bakımından n = 30 için simule edilmiş varyanslarla hesaplanmış test istatistiklerinin 1. tip hataları Tablo 3'de verilmiştir.

Tablo 3. LS ve MML tahmin edicilerine dayalı test istatistiklerinin birinci tip hataları; Varsayılan dağılım p = 3.5 olan t-ailesi.

	$\phi_1 = \phi_2 = 0.2$		$\phi_1 = \phi_2 = 0.5$		$\phi_1 = \phi_2 = 0.9$	
Modeller	**LS**	**MML**	**LS**	**MML**	**LS**	**MML**
(1)	0.044	0.045	0.047	0.048	0.046	0.047
(2)	0.047	0.045	0.057	0.058	0.051	0.053
(3)	0.050	0.052	0.056	0.055	0.055	0.059
(4)	0.048	0.049	0.050	0.050	0.059	0.056
(5)	0.047	0.045	0.048	0.045	0.050	0.054
(6)	0.045	0.044	0.044	0.047	0.052	0.051
(7)	0.050	0.052	0.047	0.045	0.048	0.046
(8)	0.048	0.048	0.050	0.055	0.049	0.051
(9)	0.045	0.048	0.048	0.049	0.053	0.053

Tablo 3'den her iki testin de ölçüt sağlam olduğu görülmektedir. Seçilmiş modeller için testlerin güçlerini veren Tablo 4 incelendiğinde MML tahmin edicilerine dayanan testlerin LS tahmin edicilerine dayalı testlerden daha güçlü olduğu görülmektedir.

Tablo 4. Simule edilmiş varyanslarla hesaplanan test istatistiklerinin güç değerleri

	Varsayılan Model		**Model (2)**		**Model (6)**		**Model (8)**	
λ	**LS**	**MML**	**LS**	**MML**	**LS**	**MML**	**LS**	**MML**
0.00	0.050	0.052	0.057	0.058	0.044	0.047	0.050	0.055
0.16	0.195	0.214	0.207	0.246	0.167	0.174	0.109	0.151
0.32	0.485	0.535	0.491	0.579	0.400	0.437	0.233	0.363
0.48	0.761	0.814	0.771	0.856	0.680	0.721	0.395	0.629
0.64	0.938	0.962	0.939	0.974	0.880	0.906	0.600	0.842
0.80	0.992	0.997	0.987	0.997	0.968	0.980	0.779	0.946

Diğer taraftan yine seçilmiş modeller için asimptotik varyanslar kullanılarak hesaplanan testlerin güç değerlerinin verildiği Tablo 5'e bakıldığında LS tahmin edicilerine dayanan test istatistiklerinin önceden belirlenen 0.05 düzeyinden daha yüksek 1. tip hataya sahip olduğu görülmektedir. MML tahmin edicilerine dayanan testlerde böyle bir sorun görülmezken alternatif modeller altındaki güç değerleri ile varsayılan modelin güç değerleri karşılaştırıldığında testin etkin sağlam olduğu da görülmektedir.

Tablo 5. Asimtotik varyanslarla hesaplanan test istatistiklerinin güç değerleri

	Varsayılan Model		Model (2)		Model (6)		Model (8)	
λ	LS	MML	LS	MML	LS	MML	LS	MML
0.00	0.059	0.054	0.072	0.051	0.058	0.053	0.070	0.044
0.16	0.225	0.211	0.246	0.236	0.206	0.187	0.155	0.127
0.32	0.530	0.532	0.539	0.558	0.446	0.449	0.318	0.313
0.48	0.789	0.809	0.804	0.842	0.719	0.731	0.497	0.569
0.64	0.937	0.955	0.938	0.965	0.892	0.905	0.664	0.771
0.80	0.988	0.995	0.983	0.995	0.969	0.979	0.815	0.907

7. K (> 2) DENKLEME GENELLEME

K (> 2) tane bağımsız bilgi kaynağı olduğu durumlarda tahmin ediciler aynen bölüm 2'deki gibidir. Dolayısı ile burada hipotez testlerinin genellemesi yapılacaktır. Burada ilgilenilen ana hipotezler şöyledir:

1) Tüm regresyon katsayılarının eşitliği, $\delta_1 = \delta_2 = \ldots = \delta_k = \delta$;
2) Tüm regresyon katsayılarının sıfıra eşit olduğu;
3) Regresyon katsayılarının bir kısmının, örneğin ilk $q < k$ tanesinin, sıfıra eşitliği;
4) Regresyon katsayılarının bir kısmının, δ_i $(1 \le i \le q)$, eşitliği.

7.1. δ_i $(1 \le i \le k)$ Eşitliğinin Testi

Test istatistiği şu şekilde tanımlanır:

$$\chi_1^2 = \left(\hat{\delta}_1 - \hat{\delta}, \ldots, \hat{\delta}_k - \hat{\delta}\right) \begin{pmatrix} \hat{V}_1 & 0 & \ldots & 0 \\ . & \hat{V}_2 & \ldots & 0 \\ . & . & \ldots & . \\ . & . & . & \hat{V}_k \end{pmatrix}^{-1} \begin{pmatrix} \hat{\delta}_1 - \hat{\delta} \\ . \\ . \\ \hat{\delta}_k - \hat{\delta} \end{pmatrix} ;$$

$$\hat{V}_i = \frac{k}{2p} \frac{\hat{\sigma}_i^2}{\sum_{j=1}^{n_i} \beta_{ij} \left(u_{i[j]} - \bar{u}_{i[.]}\right)^2} . \tag{21}$$

$(n_i)^{1/2} \hat{\delta}_i$ normal dağıldığı ve n_i büyüdükçe $\hat{\sigma}_i$, σ_i'ye yakınsadığından büyük n_i için χ_1^2 test istatistiğinin H_0 altında dağılımı k-1 serbestlik dereceli Ki-Kare olur. Alternatif hipotez altında ise k-1 serbestlik dereceli merkezi olmayan Ki-Kare dağılama sahip olur. Merkezsizlik parametresi ise

$$\lambda_1^2 = \left(\delta_1 - \delta, \ldots, \delta_k - \delta\right) \begin{pmatrix} V_1 & 0 & \ldots & 0 \\ . & V_2 & \ldots & 0 \\ . & . & \ldots & . \\ . & . & . & V_k \end{pmatrix}^{-1} \begin{pmatrix} \delta_1 - \delta \\ . \\ . \\ \delta_k - \delta \end{pmatrix} \tag{22}$$

olur. Burada V_i denklem (21)'deki $\hat{\sigma}_i^2$'nin σ_i^2 $(1 \le i \le k)$ ile değiştirilmiş halidir.

LS tahmin edicilerini kullanırsak test istatistiği aşağıdaki gibi olur:

$$\chi_2^2 = \left(\tilde{\delta}_1 - \tilde{\delta}, \ldots, \tilde{\delta}_k - \tilde{\delta}\right) \begin{pmatrix} \tilde{V}_{1o} & 0 & \ldots & 0 \\ . & \tilde{V}_{2o} & \ldots & 0 \\ . & . & \ldots & . \\ . & . & . & \tilde{V}_{ko} \end{pmatrix}^{-1} \begin{pmatrix} \tilde{\delta}_1 - \tilde{\delta} \\ . \\ . \\ \tilde{\delta}_k - \tilde{\delta} \end{pmatrix}. \qquad (23)$$

Burada $\tilde{\delta}$ aşağıda tanımlandığı gibi birleştirilmiş LS tahmin edicisidir:

$$\tilde{\delta} = \sum_{i=1}^{k} w_i \tilde{\delta}_i \Big/ \sum_{i=1}^{k} w_i \ , \ w_i = 1 / V_{io}, \ \tilde{V}_{io} = \frac{\tilde{\sigma}_i^2}{\sum_{j=1}^{n_i} \beta_{ij} \left(u_{i[j]} - \bar{u}_{i[.]}\right)^2} \quad (1 \le i \le k). \qquad (24)$$

H_0 altında büyük n_i değerleri için χ_2^2, k-1 serbestlik dereceli Ki-Kare dağılımına sahiptir. Alternatif hipotez altında ise k-1 serbestlik dereceli, aşağıda verilen merkezsizlik parametresine sahip merkezi olmayan Ki-Kare dağılır:

$$\lambda_2^2 = \left(\delta_1 - \delta, \ldots, \delta_k - \delta\right) \begin{pmatrix} V_{1o} & 0 & \ldots & 0 \\ . & V_{2o} & \ldots & 0 \\ . & . & \ldots & . \\ . & . & . & V_{ko} \end{pmatrix}^{-1} \begin{pmatrix} \delta_1 - \delta \\ . \\ . \\ \delta_k - \delta \end{pmatrix} \qquad (25)$$

Burada V_{i0} denklem (24)'deki $\tilde{\sigma}_i^2$'nin σ_i^2 $(1 \le i \le k)$ ile değiştirilmiş halidir.

χ_1^2'nin χ_2^2'den daha güçlü olduğunu ispatlamak için

$$\lim_{n_i \to \infty} \frac{1}{n_i} \frac{2p}{k} \sum_{j=1}^{n_i} \beta_{ij} \left(u_{i[j]} - \bar{u}_{i[.]}\right)^2 > \frac{1}{n_i} \sum_{j=1}^{n_i} \left(u_{ij} - \bar{u}_i\right)^2 \qquad (26)$$

olduğunu göstermek yeterlidir.

Soldaki limit $\left[\frac{1}{n_i} \sum_{j=1}^{n_i} \left(u_{ij} - \bar{u}_i\right)^2\right] \frac{2p}{k} E\left[\frac{1 - z^2/k}{\left(1 + z^2/k\right)^2}\right]$, k = 2p – 3 (p ≥ 2) şeklinde yazılabilir ve $\int_{-\infty}^{\infty} (1 + z^2/k)^{-j} dz = \sqrt{k}\beta(1/2, j - 1/2)$ $j \ge 1$ olduğundan buradaki beklenen değer (p – ½) / (p + 1)'dir ve $\frac{2p}{k} \frac{(p - 1/2)}{(p+1)} = \frac{p(p - 1/2)}{(p+1)(p - 3/2)}$ tüm p (≥ 2) değerleri için birden büyük olduğundan denklem (26) sağlanmaktadır.

Örnek olarak, k = 3 ve p = 3.5 için χ_1^2 ve χ_2^2 testlerinin simule edilmiş 1. tip hata ve güç değerleri Tablo 6'da verilmektedir. MML tahmin edicilerine dayanan χ_1^2 testinin üstünlüğü açıktır.

Hatırlatma: Eğer σ_i^2 $(1 \le i \le k)$ değerleri birbirine eşit olursa $\hat{\sigma}_i^2$ birleştirilmiş $\hat{\sigma}^2$ ile değiştirilir. Elde edilen test istatistiği alternatif hipotez altında denklem (24)'de σ_i^2'nin

σ^2 ile değiştirilmesi ile elde edilen merkezsizlik parametresine sahip (k-1, N-k), $N = \sum_{i=1}^{k} n_i$, serbestlik dereceli merkezi olmayan F dağılır.

Tablo 6. (1) $H_0 : \delta_1 = \delta_2 = \delta_3$ (2) $H_0 : \delta_1 = \delta_2 = \delta_3 = 0$ hipotezleri için güç değerleri

	(1)		(2)	
λ	**LS**	**MML**	**LS**	**MML**
0.0	0.063	0.054	0.057	0.054
0.1	0.095	0.104	0.095	0.101
0.2	0.242	0.271	0.269	0.300
0.3	0.553	0.591	0.552	0.600
0.4	0.783	0.826	0.835	0.863
0.5	0.932	0.955	0.957	0.971

7.2. $\delta_i = 0$ $(1 \le i \le k)$ Testi

$H_0{:}\delta_i = 0$ $(1 \le i \le k)$ hipotezini $H_1{:}$ $\delta_i \neq 0$ (bazı i için) alternatif hipotezine karşın test etmek için ilgili test istatistiği

$$\chi^2 = \left(\hat{\delta}_1, \ldots, \hat{\delta}_k\right) \begin{pmatrix} \hat{V}_1 & 0 & \ldots & 0 \\ . & \hat{V}_2 & \ldots & 0 \\ . & . & \ldots & . \\ . & . & . & \hat{V}_k \end{pmatrix}^{-1} \begin{pmatrix} \hat{\delta}_1 \\ . \\ . \\ \hat{\delta}_k \end{pmatrix} \tag{27}$$

şeklindedir. Burada $\hat{V}_i$ denklem (21)'deki gibidir.

Dağılımlar ise aynen bölüm 7.1'deki gibidir. Sadece burada serbestlik derecesi artık k'dır. Eğer σ_i^2 $(1 \le i \le k)$ değerleri birbirine eşit olursa $\hat{\sigma}_i^2$ birleştirilmiş $\hat{\sigma}^2$ ile değiştirilir. Elde edilen F istatistiğinin serbestlik derecesi (k, N-k) olur.

Örnek olarak, k = 3 ve p = 3.5 için MML ve LS tahmin edicilerine dayanan χ^2 testlerinin simule edilmiş 1. tip hata ve güç değerleri Tablo 6'da verilmektedir. MML tahmin edicilerine dayanan testin üstünlüğü açıktır.

7.3. Regresyon Katsayılarının Bir Kısmının Testi

$H_0{:}\delta_1 = \delta_2 = \ldots = \delta_q = 0$ hipotezini $H_1{:}\delta_i \neq 0$ (bazı i = 1, 2, ..., q değerleri için) alternatif hipotezine karşın test etmek için ilgili test istatistiği denklem (27)'de verilen test istatistiğinin parçalanması ile aşağıdaki şekilde elde edilir:

$$\chi^2 = \left(\hat{\delta}_1, \ldots, \hat{\delta}_q\right) \begin{pmatrix} \hat{V}_1 & 0 & \ldots & 0 \\ . & \hat{V}_2 & \ldots & 0 \\ . & . & \ldots & . \\ . & . & . & \hat{V}_q \end{pmatrix}^{-1} \begin{pmatrix} \hat{\delta}_1 \\ . \\ . \\ \hat{\delta}_q \end{pmatrix} \tag{28}$$

Dağılımlar ise aynen bölüm 7.2'deki gibidir. Sadece burada serbestlik derecesi artık q'dur.

Eğer σ_i^2 $(1 \le i \le k)$ değerleri birbirine eşit olursa (28)'deki $\hat{V}_i$ $(1 \le i \le q)$,

$$\hat{V}_i = \frac{k}{2p} \frac{\hat{\sigma}^2}{\sum_{j=1}^{n_i} \beta_{ij}\left(u_{i[j]} - \bar{u}_{i[.]}\right)^2} \; ; \; \hat{\sigma}^2 = \sum_{i=1}^{q} (n_i - 1)\hat{\sigma}_i^2 \Big/ (N-k), \; N = \sum_{i=1}^{q} n_i$$ ile değiştirilir. Elde

edilen test istatistiğinin H_0 altında dağılımı (q-1, N-k) serbestlik dereceli F olur.

Öte yandan varyansların eşitliği durumunda $H_0{:}\delta_1 = \delta_2 = \ldots = \delta_q$ hipotezi $H_1{:}\delta_i \ne \delta_j$ (bazı $i \ne j = 1, 2, \ldots, q$ değerleri için) alternatif hipotezine karşın test edilmek istenir ise test istatistiği

$$\chi^2 = \left(\hat{\delta}_1 - \hat{\delta}, \ldots, \hat{\delta}_q - \hat{\delta}\right) \begin{pmatrix} \hat{V}_1 & 0 & \ldots & 0 \\ . & \hat{V}_2 & \ldots & 0 \\ . & . & \ldots & . \\ . & . & . & \hat{V}_q \end{pmatrix}^{-1} \begin{pmatrix} \hat{\delta}_1 - \hat{\delta} \\ . \\ . \\ \hat{\delta}_q - \hat{\delta} \end{pmatrix} \tag{29}$$

olur. Burada $\hat{\delta}$, denklem (24)'deki gibi $\hat{\delta}_i$'lerin $(1 \le i \le q)$ birleştirilmesi ile elde edilir. H_0 altında dağılımı (q-1, N-k) serbestlik dereceli F olur.

7. SONUÇLAR

Uygulamada sık rastlanan normal olmayan inovasyonlu bağımsız basit doğrusal otoregresif modeller sisteminde ML tahmin edicilerinin hesaplanmasında sorunlar yaşanmaktadır. Dolayısı ile çoğunlukla pratikte LS tahmin edicileri kullanılmaktadır. Ancak LS tahmin edicilerinin kullanılması etkinliğin kaybına sebep olmaktadır.

Bu çalışmada uzun-kuyruklu simetrik aile incelenmiştir. MML tahmin edicileri bulunmuş ve etkili ve sağlam oldukları gösterilmiştir. Ayrıca LS tahmin edicilerinin etkinliğinin MML tahmin edicilerine göre daha düşük olduğu görülmüştür. Bunların yanısıra MML tahmin edicilerine dayalı test istatistikleri geliştirilmiş ve bunların güçlü ve sağlam oldukları gösterilmiştir. Yine LS tahmin edicilerine dayanan test istatistikleri ile karşılaştırıldığında bulunan test istatistiklerinin daha güçlü olduğu görülmüştür. Dolayısı ile normal olmayan inovasyonlu otoregresif modellerde LS yerine MML tahmin edicilerinin kullanılması tavsiye edilmektedir.

KAYNAKLAR

AKKAYA, A. D. ve TIKU, M. L. (2001a), *Estimating parameters in autoregressive models in non-normal situations: asymmetric innovations,* Commun. Statist. Theory Meth., 30(3), 517-536.

AKKAYA, A. D. ve TIKU, M. L. (2001b), *Corrigendum: time series models with asymmetric innovation,* Commun. Statist. Theory Meth., 30(10), 2227-2230.

BHATTACHARYA, G. K. (1985), *The asymptotics of maximum likelihood and related estimators based on Type II censored data,* J. Amer. Statist. Assoc., 80, 398-404.

DURBIN, J. (1960), *Estimation of parameters in time series regression model*, JRSS B, 22, 139-153.

ELVEBACK, L. R., GUILLIER, C. L. ve KERTING, F. R. (1970), *Health, normality and the ghost of Gauss*, J. American Medical Assoc., 211, 69-75.

ISLAM, M. Q., TIKU, M. L. ve YILDIRIM, F (2001), *Non-normal regression I*, Commun. Statist. Theory Meth., 30, 993-1020.

ISLAM, M. Q. ve TIKU, M. L. (2004), *Multiple linear regression models under non-normality,* Commun. Stat. – Theory Meth.

MALLER, R. A. (1989), *Regression with autoregressive errors-some asymptotic results*, Statistics, 20, 23-39.

NAGARAJ, N. K. ve FULLER, W. A. (1992), *Least squares estimation of the linear model with autoregressive errors*, New Directions in time series analysis, Part I, IMA Vol. Math. Appl. 45, New York: Springer, 215-225.

PEARSON, E. S. (1932), *The analysis of variance in cases of nonnormal variation,* Biometrika, 23, 114-133.

PUTHENPURA, S. ve SINHA, N. K. (1986), *Modified maximum likelihood method for the robust estimation of system parameters from very noisy data,* Automatica, 22, 231-235.

SCHÄFFLER, S. (1991), *Maximum likelihood estimation for linear regression model with autoregressive errors*, Statistics, 22, 191-198.

TIKU, M. L. (1967), *Estimating the mean and standard deviation from censored normal samples,* Biometrika, 54, 155-165.

TIKU, M. L. ve KUMRA, S. (1981), *Expected values and variances and covariances of order statistics for a family of symmetrical distributions (Student's t)*, Selected Tables in Mathematical Statistics 8, Providence, RI: American Mathematical Society, 141-270.

TIKU, M. L., TAN, W. Y. ve BALAKRISHNAN, N. (1986), *Robust Inference*, New York, Marcel Dekker.

TIKU, M. L. ve SURESH, R. P. (1992), *A new method of estimation for location and scale parameters*, J. Stat. Plann. And Inf., 30, 281-292.

TIKU, M. L., WONG, W. K. ve BIAN, G. (1999), *Estimating parameters in autoregressive models in non-normal situations; symmetric innovations*, Commun. Statist. Theory Meth., 28(2), 315-341.

TIKU, M. L., ISLAM M. Q., and SELÇUK, A. S. (2001), *Non-normal regression II*, Commun. Statist. Theory Meth., 30, 1021-1045.

TSE, Y. K. (1991), *Price and volume in the Tokyo stock exchange: an explanatory study*, Japanese Financial Market Research (W. T. Ziemba, W. Bailey and Y. Hamano, Eds.), New York: Elsevier Science Publishers B. V., 91-119.

VAUGHAN, D. C. (1992), *On the Tiku-Suresh method of estimation*, Commun. Statist. Theory Meth., 21, 451-469.

VAUGHAN, D. C. ve TIKU, M. L. (2000), *Estimation and hypothesis testing for a non-normal bivariate distribution and applications*, J. Mathematical and Computer Modelling, 32, 53-67.

VELU, R. ve GREGORY, C. (1987), *Reduced rank regression with autoregressive errors*, Econometrics, 35, 317-335.

VINOD, H. D. ve SHENTON, L. R. (1996) *Exact moments for autoregressive and random walk models for a zero or stationary initail value*, Econometric Theory, 21, 391-404.

ESTIMATION OF PARAMETERS AND HYPOTHESIS TESTING IN THE SYSTEM OF SIMPLE AUTOREGRESSIVE MODELS: SYMMETRIC INNOVATIONS

ABSTRACT

In this study, the simple autoregressive models system having symmetric innovations have been investigated and the methodology under non-normality has been extended to various independent sources of information. Modified likelihood estimators are obtained; robust and efficient statistics for testing whether the parameter vector remains the same from one source to another are developed. The estimators and the test statistics obtained are compared to the corresponding least squares statistics which are widely used in this context in applications, and found to give more accurate results.

Key Words: *Maximum Likelihood Estimators, Least Squares Estimators, Modified Likelihood Estimators, Autoregression, Robustness, Non-normality.*