



Bir-Yönlü Kovaryans Analizinde Parametrelerin RAMML Yöntemi ile Tahmini

Şükrü ACITAŞ

How to cite: Acitaş, Ş. (2025). Bir-yönlü kovaryans analizinde parametrelerin RAMML yöntemi ile tahmini. *Sinop Üniversitesi Fen Bilimleri Dergisi*, 10(1), 188-199. <https://doi.org/10.33484/sinopfbid.1644300>

Araştırma Makalesi

Sorumlu Yazar

Şükrü ACITAŞ
 sacitas@eskisehir.edu.tr

Yazarlara ait ORCID

Ş.A: 0000-0002-4131-0086

Received: 21.02.2025

Accepted: 26.05.2025

Öz

Deney tasarımı temel amaç, ilgilenilen faktörün düzeyleri olarak tanımlanan denemelerin ortalamaları arasında istatistiksel olarak anlamlı bir farklılık olup olmadığının belirlenmesidir. Yanıt değişkeni (y) üzerinde kontrol edilebilen faktörün etkisinin yanı sıra bazı durumlarda kontrol edilemeyen ortak değişkenin de etkisi söz konusu olabilir. Bu durumda, deneysel hatayı azaltmak amacıyla kontrol edilemeyen ortak değişkenin yanıt değişkeni üzerindeki etkisi arındırılarak analiz yapılması gerekir. Bu amaçla, kontrol edilemeyen ortak değişken (x) modele dahil edilir. Bu yöntem, deney tasarımı kovaryans analizi (analysis of covariance – *ANCOVA*) olarak adlandırılır. Varyans analizi ile regresyon analizi tekniklerinin bir birleşimi olarak ifade edilen *ANCOVA* hem y -yönlü hem de x -yönlü aykırı değerlere karşı duyarlıdır. Literatürde y -yönlü aykırı değerlere karşı dayanıklı *ANCOVA* ile ilgili çalışmalar bulunmasına rağmen bilindiği kadarıyla x -yönlü aykırı değerlere karşı dayanıklı adaptif uyarlanmış en çok olabilirlik (robust adaptive modified maximum likelihood – *RAMML*) tahmin edicilerine dayanan *ANCOVA* bağlamında herhangi bir çalışma bulunmamaktadır. Bu çalışmada, y – yönlü aykırı değerlere ek olarak ortak değişkende aykırı değerler olması, bir başka ifade ile x – yönlü aykırı değerler olması durumu ele alınmış ve bir-yönlü *ANCOVA*’da model parametrelerinin *RAMML* tahmin edicileri elde edilmiştir.

Anahtar Kelimeler: Kovaryans analizi, x -yönlü ve y – yönlü aykırı değerler, *RAMML* metodu, dayanıklılık

Estimation of the Parameters in One-Way Analysis of Covariance with the RAMML Method

Eskişehir Teknik Üniversitesi,
 Fen Fakültesi, İstatistik Bölümü,
 Eskişehir, Türkiye

Abstract

The main purpose of the experimental design is to determine whether there is a statistically significant difference among the means of the treatments defined as the levels of the factor of interest. In addition to the effect of the controllable factor on the response variable (y), in some cases, there may also be an effect of the uncontrolled covariate. In this case, the analysis should be performed by eliminating the effect of the uncontrolled covariate on the response variable in order to reduce the experimental error. For this purpose, the uncontrolled covariate (x) is included in the model. This method is called analysis of covariance (*ANCOVA*) in experimental design. *ANCOVA*, which is a combination of analysis of variance and regression analysis techniques, is sensitive to both y - and x -outliers. Although there are studies on *ANCOVA* models robust to y -direction outliers in the literature, to the best of our knowledge, there is no

Giriş

Bir-yönlü kovaryans analizi modeli (analysis of covariance – *ANCOVA*)

$$y_{ij} = \mu + \tau_i + \beta_i x_{ij} + \varepsilon_{ij}; \quad i = 1, 2, \dots, a; \quad j = 1, 2, \dots, n \quad (1)$$

olarak ifade edilir. Burada, y_{ij} yanıt değişkeni ve x_{ij} ortak değişken (covariate) olup μ genel ortalamayı, τ_i i –inci denemenin etkisini, β_i ortak değişken ile yanıt değişkeni arasındaki regresyon doğrularının eğimini gösteren model parametreleri, ε_{ij} ise rastgele hata terimleridir. ANCOVA, yanıt değişkeninin farklı çevresel koşullara maruz kalması durumunda kullanılmaktadır [1-5]. Bu çevresel koşullar modelde ortak değişken olarak ifade edilmektedir. Açık ki, bir-yönlü ANCOVA modeli, varyans analizi ile regresyon analizinin bir birleşimidir. Eşitlik (1)’de verilen modelde $\mu_i = \mu + \tau_i$ eşitliği kullanılarak yeniden parametrelendirme yapılırsa bir-yönlü ANCOVA modeli alternatif olarak

$$y_{ij} = \mu_i + \beta_i x_{ij} + \varepsilon_{ij}; \quad i = 1, 2, \dots, a; \quad j = 1, 2, \dots, n \quad (2)$$

şeklinde de ifade edilebilir. Çalışmanın geri kalan kısmında Eşitlik (2) kullanılmıştır. Bir-yönlü ANCOVA modelinde istatistiksel çıkarım yapabilmek için hata terimlerinin bağımsız ve aynı dağılımlı $N(0, \sigma^2)$ olduğu, yanıt değişkeni ile ortak değişken arasındaki ilişkinin doğrusal olduğu, modelde ölçüm hatası olmadığı ve ortak değişkenin sabit bir başka ifade ile rastgele değişken olmadığı varsayılır, örneğin bkz. Şenoğlu ve Acıtaş [6]. Eşitlik (2)’de verilen model parametrelerinin en küçük kareler (least squares – *LS*) tahmin edicileri,

$$\hat{\mu}_{i_{LS}} = \bar{y}_{i.} - \hat{\beta}_{i_{LS}} \bar{x}_{i.}, \quad \hat{\beta}_{i_{LS}} = \frac{E_{xy_i}}{E_{xx_i}},$$

$$\hat{\sigma}_{i_{LS}}^2 = \frac{1}{n-2} \sum_{j=1}^n [y_{ij} - \bar{y}_{i.} - \hat{\beta}_{i_{LS}} (x_{ij} - \bar{x}_{i.})]^2 \quad \text{ve} \quad \hat{\sigma}_{LS}^2 = \sqrt{\frac{1}{a} \sum_{i=1}^a \hat{\sigma}_{i_{LS}}^2}$$

olarak elde edilir. Burada,

$$\bar{y}_{i.} = \frac{1}{n} \sum_{j=1}^n y_{ij}, \quad \bar{x}_{i.} = \frac{1}{n} \sum_{j=1}^n x_{ij},$$

$$E_{xx_i} = \sum_{j=1}^n (x_{ij} - \bar{x}_{i.})^2 \quad \text{ve} \quad E_{xy_i} = \sum_{j=1}^n (x_{ij} - \bar{x}_{i.})(y_{ij} - \bar{y}_{i.}), \quad i = 1, 2, \dots, a$$

dir. Normallik varsayımı altında model parametrelerinin *LS* tahmin edicilerinin optimal olduğu bilinmektedir. Bunun bir sonucu olarak, deneme ortalamaları arasındaki farkın anlamlı olup olmadığını sınamak için kullanılan test istatistiği de düzgün en güçlü testtir. Ancak, birçok durumda gerçek hayat problemlerinden elde edilen verileri modellemek için normal olmayan dağılımları kullanmak normal dağılım kullanmaya göre çok daha uygundur, örneğin bkz. Şenoğlu [7], Şenoğlu ve Avcıoğlu [8]. Veri setinde aykırı değerler olması, normallik varsayımının bozulmasının en önemli etkenlerinden biridir. Literatürde bir yönlü ANCOVA modelinde normallik varsayımı sağlanmadığında parametrelerin dayanıklı (robust) tahmin edicileri ve bu tahmin edicilere dayalı test istatistikleri ile ilgili çalışmalar mevcuttur. Örneğin, Birch ve Myers [9], bir-yönlü ANCOVA’da model parametrelerinin M-tahmin edicilerini elde etmiştir. Şenoğlu [7], bir-yönlü ANCOVA modelinde hata terimlerinin kısa kuyruklu simetrik dağılıma sahip olması durumunda model parametrelerinin uyarlanmış en çok olabilirlik (modified maximum likelihood – *MML*) tahmin edicilerinin kullanılmasını önermiştir. Şenoğlu ve Avcıoğlu [8], Model (2)’de hata terimlerinin genelleştirilmiş lojistik dağıldığı varsayımı altında model parametrelerinin *MML* tahmin edicilerini elde etmiş, bu tahmin edicilere dayalı olarak deneme ortalamaları arasında anlamlı bir farklılık olup olmadığını sınamak için dayanıklı bir test istatistiği önermiştir. Bu çalışmalarda, önerilen tahmin edicilerin etkinlikleri ve dayanıklılıkları ortak değişkende aykırı değer olması durumunda (*x*-yönlü aykırı değer) incelenmemiştir. Yukarıda da değinildiği gibi normallik varsayımı birçok gerçek hayat problemi için makul bir varsayım olmadığından, Model (2)’de hata terimlerinin dağılımının olasılık yoğunluk fonksiyonu aşağıda verilen uzun kuyruklu simetrik (long-tailed symmetric – *LTS*) olduğu varsayılmıştır:

$$f(e; p, \sigma) = \frac{1}{\sqrt{q}B(0.5, p-0.5)\sigma} \left(1 + \frac{e^2}{q\sigma^2}\right)^{-p}, \quad q = 2p - 3, -\infty < e < \infty. \quad (3)$$

Burada, *p* *LTS* dağılımının şekil parametresi, *B*(·, ·) beta fonksiyonu ve σ da ölçek parametresidir. Şekil parametresi *p* sonsuza giderken *LTS* dağılımı normal dağılıma yakınsar. *p*’nin tüm değerleri için *LTS* dağılımının sivrililiği 3’ten büyüktür. Bu nedenle, *LTS* dağılımı literatürde aykırı değerleri modellemek için sıklıkla kullanılır, örneğin bkz. Tiku ve Suresh [10]. Tiku vd. [11] ile İslam ve Tiku [12] sırasıyla basit ve çoklu doğrusal regresyon modelinde hata terimlerinin *LTS* dağılımına sahip olduğunu varsaymış ve model parametrelerinin *MML* tahmin edicilerini elde etmişlerdir. Acıtaş ve Şenoğlu [13] faktöriyel ANCOVA modelinde hata terimlerinin dağılımının *LTS* olduğu varsayımı altında model parametrelerinin *MML* tahmin edicilerini bulmuşlardır. *MML* tahmin edicileri, olabilirlik denklemlerindeki parametrelerin doğrusal olmayan fonksiyonunun yerine bu fonksiyonun birinci dereceden Taylor serisi açılımı kullanılarak elde edilir [14, 15]. Bunun bir sonucu olarak, *MML* tahmin edicileri açık olarak formüle edilebilirler, asimptotik olarak en çok olabilirlik (maximum likelihood – *ML*) tahmin edicilerine denktirler, bkz. Bhattacharyya [16], Vaughan ve Tiku [17]. *MML* tahmin edicileri, şekil parametresi *p*’nin bilindiği varsayımı altında elde edilmektedir. Tiku ve Sürücü [18],

MML tahmin edicileri elde edilirken şekil parametresinin bilindiği varsayımını gevşeterek yeni bir versiyon önermişlerdir. Dönmez [19], bu versiyonu varyans analizi, basit ve çoklu doğrusal regresyon analizlerinde kullanmıştır. MML tahmin yönteminin önerilen versiyonu adaptif uyarlanmış en çok olabilirlik (adaptive modified maximum likelihood – *AMML*) yöntemi olarak adlandırılmıştır [19]. Acıtaş vd. [20, 21], basit ve çoklu doğrusal regresyon model parametrelerinin x -yönlü aykırı değerlerin varlığı durumunda dayanıklı tahmin edicilerini elde etmek için *AMML* yöntemini geliştirerek *RAMML* yöntemini önermiştir. Bu çalışmanın amacı, y –yönlü aykırı değerlere ek olarak ortak değişkende aykırı değerler olması durumunda, Eşitlik (2)’de verilen ANCOVA modelinin parametrelerinin dayanıklı tahmin edicilerini elde etmektir. Bu kapsamda, Acıtaş vd. [20, 21] tarafından önerilen *RAMML* tahmin edicileri kullanılmıştır. *RAMML* tahmin edicilerinin dayandığı düşünce, x -yönlü aykırı değerlerin etkisini azaltacak yeni ağırlıkların tahmin sürecine dahil edilmesidir. Bilindiği kadarıyla, x -yönlü aykırı değerlere karşı *RAMML* tahmin edicilerine dayanan ANCOVA bağlamında herhangi bir çalışma bulunmamaktadır.

Parametrelerin *RAMML* Tahmin Edicileri

Eşitlik (2)’de verilen ANCOVA modelinde hata terimlerinin dağılımının LTS olduğu varsayımı altında model parametrelerinin ML tahmin edicileri aşağıda verilen olabilirlik denklemlerinin çözümünden elde edilir:

$$\frac{\partial \log L}{\partial \mu_i} = \frac{2p}{q\sigma} \sum_{j=1}^n g(z_{ij}) = 0, \quad (4)$$

$$\frac{\partial \log L}{\partial \beta_i} = \frac{2p}{q\sigma} \sum_{j=1}^n g(z_{ij}) x_{ij} = 0, \quad (5)$$

$$\frac{\partial \log L}{\partial \sigma} = -\frac{N}{\sigma} + \frac{2p}{q\sigma} \sum_{i=1}^a \sum_{j=1}^n g(z_{ij}) z_{ij} = 0. \quad (6)$$

Burada,

$$g(z_{ij}) = \frac{z_{ij}}{1 + \frac{1}{q} z_{ij}^2}, \quad z_{ij} = \frac{y_{ij} - \mu_i - \beta_i x_{ij}}{\sigma}, \quad i = 1, 2, \dots, a, \quad j = 1, 2, \dots, n \quad \text{ve} \quad N = an$$

dir. Açık ki, $g(z_{ij})$ fonksiyonu parametrelerin doğrusal olmayan bir fonksiyonudur; bu nedenle, olabilirlik denklemlerinin açık çözümleri elde edilemez. Nümerik yöntemler kullanılarak ML tahmin edicileri elde edilir ancak nümerik yöntemler kullanmak hiç yakınsamama, yanlış köke yakınsama ve birden fazla köke yakınsama gibi dezavantajlara sahiptir [22, 23]. Bu durumda, Acıtaş vd. [20, 21] tarafından önerilen *RAMML* tahmin edicileri alternatif olarak kullanılabilir. Bir-yönlü ANCOVA’da

model parametrelerinin RAMML tahmin edicileri, aşağıdaki adımlar takip edilerek elde edilir. Burada belirtmek gerekir ki, Adım 1 ve Adım 2, MML, AMML ve RAMML yöntemleri için ortaktır.

Adım 1. $g(z_{ij})$ fonksiyonunun

$$t_{ij} = \frac{y_{ij} - \mu_i - \beta_i x_{ij}}{\sigma}, \quad i = 1, 2, \dots, a, \quad j = 1, 2, \dots, n \quad (7)$$

etrafında birinci dereceden Taylor serisi açılımı yapılarak $g(z_{ij}) \cong \alpha_{ij} + \delta_{ij} z_{ij}$ şeklinde ifade edilir. Burada,

$$\alpha_{ij} = \frac{\frac{1}{q} t_{ij}}{\left(1 + \frac{1}{q} t_{ij}^2\right)^2} \quad \text{ve} \quad \delta_{ij} = \frac{1}{\left(1 + \frac{1}{q} t_{ij}^2\right)^2}, \quad i = 1, 2, \dots, a, \quad j = 1, 2, \dots, n \quad (8)$$

dir.

Adım 2. Olabilirlik denklemlerinde $g(z_{ij})$ fonksiyonun yerine Adım 1’de elde edilen $g(z_{ij}) \cong \alpha_{ij} + \delta_{ij} z_{ij}$ kullanılarak

$$\begin{aligned} \frac{\partial \log L^*}{\partial \mu_i} &= \frac{2p}{q\sigma} \sum_{j=1}^n (\alpha_{ij} + \delta_{ij} z_{ij}) = 0, \\ \frac{\partial \log L^*}{\partial \beta_i} &= \frac{2p}{q\sigma} \sum_{j=1}^n (\alpha_{ij} + \delta_{ij} z_{ij}) x_{ij} = 0, \\ \frac{\partial \log L^*}{\partial \sigma} &= -\frac{N}{\sigma} + \frac{2p}{q\sigma} \sum_{i=1}^a \sum_{j=1}^n (\alpha_{ij} + \delta_{ij} z_{ij}) z_{ij} = 0 \end{aligned}$$

şeklinde gösterilen uyarlanmış olabilirlik denklemleri elde edilir.

Adım 3. Uyarlanmış olabilirlik denklemlerine, x -yönlü aykırı değerlerin etkisini azaltmak amacıyla δ_{ij}^x ağırlıkları dahil edilir:

$$\frac{\partial \log L^*}{\partial \mu_i} = \frac{2p}{q\sigma} \sum_{j=1}^n \delta_{ij}^x (\alpha_{ij} + \delta_{ij} z_{ij}) = 0, \quad (9)$$

$$\frac{\partial \log L^*}{\partial \beta_i} = \frac{2p}{q\sigma} \sum_{i=1}^a \sum_{j=1}^n \delta_{ij}^x (\alpha_{ij} + \delta_{ij} z_{ij}) x_{ij} = 0, \quad (10)$$

$$\frac{\partial \log L^*}{\partial \sigma} = -\frac{N}{\sigma} + \frac{2p}{\sigma} \sum_{i=1}^a \sum_{j=1}^n \delta_{ij}^x (\alpha_{ij} + \delta_{ij} z_{ij}) z_{ij} = 0. \quad (11)$$

Burada,

$$\delta_{ij}^x = \frac{1}{\left(1 + \frac{1}{q} \tilde{x}_{ij}^2\right)^4} \quad \text{ve} \quad \tilde{x}_{ij} = \frac{\|x_{ij} - \text{med}_{L_1}(x_{ij})\|}{\text{medyan}_{ij} \|x_{ij} - \text{med}_{L_1}(x_{ij})\|}$$

dir. $\text{med}_{L_1}(\cdot)$ L_1 medyan olarak adlandırılır, örneğin bkz. Fritz vd. [24].

Adım 4. Eşitlik (9)-(11)'de verilen denklem sisteminin çözümü bir-yönlü ANCOVA'da model parametrelerinin RAMML tahmin edicilerini verir. Bu tahmin ediciler aşağıda verildiği gibi kapalı formda ifade edilebilirler:

$$\hat{\mu}_{i_{RAMML}} = \bar{y}_{i[\cdot]} - \hat{\beta}_{i_{RAMML}} \bar{x}_{i[\cdot]} + \frac{\Delta_i}{m_i} \hat{\sigma}_{RAMML}, \quad (12)$$

$$\hat{\beta}_{i_{RAMML}} = K_i + L_i \hat{\sigma}_{RAMML}, \quad (13)$$

$$\hat{\sigma}_{RAMML} = \frac{B + \sqrt{B^2 + 4NC}}{2\sqrt{N(N-2a)}}. \quad (14)$$

Burada,

$$\bar{y}_{i[\cdot]} = \frac{1}{m_i} \sum_{j=1}^n \delta_{ij}^x \delta_{ij} y_{ij}, \quad \bar{x}_{i[\cdot]} = \frac{1}{m_i} \sum_{j=1}^n \delta_{ij}^x \delta_{ij} x_{ij},$$

$$\Delta_i = \sum_{j=1}^n \delta_{ij}^x \alpha_{ij}, \quad m_i = \sum_{j=1}^n \delta_{ij}^x \delta_{ij},$$

$$K_i = \frac{\sum \delta_{ij}^x \delta_{ij} (x_{ij} - \bar{x}_{i[\cdot]})(y_{ij} - \bar{y}_{i[\cdot]})}{\sum \delta_{ij}^x \delta_{ij} (x_{ij} - \bar{x}_{i[\cdot]})^2}, \quad L_i = \frac{\sum \delta_{ij}^x \alpha_{ij} (x_{ij} - \bar{x}_{i[\cdot]})}{\sum \delta_{ij}^x \delta_{ij} (x_{ij} - \bar{x}_{i[\cdot]})^2},$$

$$B_i = \frac{2p}{q} \sum_{j=1}^n \delta_{ij}^x \alpha_{ij} [y_{ij} - \bar{y}_{i[\cdot]} - K_i(x_{ij} - \bar{x}_{i[\cdot]})], \quad B = \sum_{i=1}^a B_i$$

$$C_i = \frac{2p}{q} \sum_{j=1}^n \delta_{ij}^x \delta_{ij} [y_{ij} - \bar{y}_{i[\cdot]} - K_i(x_{ij} - \bar{x}_{i[\cdot]})]^2, \quad C = \sum_{i=1}^a C_i$$

dir. Burada not edilmelidir ki, ölççek parametresi σ 'nın RAMML tahmin edicisi için alternatif bir form

$$\hat{\sigma}_{RAMML} = \sqrt{\frac{1}{a} \sum_{i=1}^a \hat{\sigma}_{i_{RAMML}}^2} \quad (15)$$

şeklinde, bkz. örneğin Şenoğlu ve Avcıoğlu [8], Dönmez [19]. Burada,

$$\hat{\sigma}_{i_{RAMML}} = \frac{B_i + \sqrt{B_i^2 + 4nC_i}}{2\sqrt{n(n-2)}}, \quad i = 1, 2, \dots, a \quad (16)$$

dır. Bir başka deyişle, Eşitlik (15) $\hat{\sigma}_{RAMML}$ 'nin $\hat{\sigma}_{iRAMML}^2$ 'lerin ortalamalarının karekökü olduğunu gösterir.

RAMML Tahmin Edicilerinin Hesaplanması

RAMML tahmin edicilerinin hesaplanması için öncelikle Eşitlik (7)'de verilen t_{ij} 'lerin hesaplanması gerekmektedir. t_{ij} ifadesi, bilinmeyen μ_i , β_i ve σ parametrelerini içerdiğinden dolayı RAMML tahminlerinin hesaplanabilmesi için μ_i , β_i ve σ parametrelerinin başlangıç değerlerinin belirlenmesine ihtiyaç vardır. Burada, Eşitlik (7)'de verilen ifadenin paydasında σ yerine her bir denemeye ait σ_i 'lerin kullanılması daha uygun olduğundan t_{ij} 'lerin hesaplanmasında

$$t_{ij} = \frac{y_{ij} - \mu_i - \beta_i x_{ij}}{\sigma_i}, \quad i = 1, 2, \dots, a, \quad j = 1, 2, \dots, n \quad (17)$$

eşitliği kullanılacaktır, örneğin bkz. Dönmez [19]. RAMML tahminleri, iki iterasyon sonucunda elde edilir. Birinci iterasyonda her bir parametreye ilişkin başlangıç değerleri ilgili deneme baz alınarak belirlenir. Bu durumda, $\hat{\mu}_i^0$, $\hat{\beta}_i^0$ ve $\hat{\sigma}_i^0$ başlangıç değerleri kullanılarak

$$t_{ij}^0 = \frac{y_{ij} - \hat{\mu}_i^0 - \hat{\beta}_i^0 x_{ij}}{\hat{\sigma}_i^0}, \quad i = 1, 2, \dots, a, \quad j = 1, 2, \dots, n$$

değerleri hesaplanır. Burada, Dönmez [19] çalışmasına paralel olarak

$$\hat{\mu}_i^0 = \text{medyan} \{y_{ij} - \hat{\beta}_i^0 x_{ij}\}, \quad \hat{\beta}_i^0 = \text{medyan} \left\{ \frac{y_{i\ell+1} - y_{i\ell}}{x_{i\ell+1} - x_{i\ell}} \right\}, \quad \ell = 1, 2, \dots, n-1, \quad i = 1, 2, \dots, a$$

ve

$$\hat{\sigma}_i^0 = 1.483 \text{ medyan} \{ |y_{ij} - \hat{\mu}_i^0 - \hat{\beta}_i^0 x_{ij}| \}$$

şeklinde alınmıştır. t_{ij}^0 değerleri ve Eşitlik (12) - (13) ve (15) - (16) kullanılarak μ_i , β_i ve σ_i parametrelerinin RAMML tahminleri $\hat{\mu}_{iRAMML}^0$, $\hat{\beta}_{iRAMML}^0$ ve $\hat{\sigma}_{iRAMML}^0$ elde edilir. Bu başlangıç değerleri yerine alternatif olarak Acıtaş vd. [21] çalışmasında ele alınan başlangıç değerleri de kullanılabilir. İkinci iterasyonda ise t_{ij} değerleri, birinci iterasyon sonucunda elde edilen $\hat{\mu}_{iRAMML}^0$, $\hat{\beta}_{iRAMML}^0$ ve $\hat{\sigma}_{iRAMML}^0$ RAMML tahminleri kullanılarak bulunur. Bir başka ifade ile,

$$t_{ij}^1 = \frac{y_{ij} - \hat{\mu}_{iRAMML}^0 - \hat{\beta}_{iRAMML}^0 x_{ij}}{\hat{\sigma}_{iRAMML}^0}, \quad i = 1, 2, \dots, a, \quad j = 1, 2, \dots, n$$

değerleri hesaplanır ve iterasyon durdurulur. t_{ij}^1 değerleri ve Eşitlik (12) - (13) ve (15) - (16) kullanılarak μ_i , β_i , σ_i ve σ parametrelerinin RAMML tahminleri elde edilmiş olur. Burada belirtmek gerekir ki, dayanıklı tahmin edicilerde aranan en önemli özelliklerden bir tanesi, normallik varsayımının

sağlanması durumunda LS tahmin edicisi kadar bir etkinliğe sahip olmasıdır, örneğin bkz. Maronna vd. [25]. Bu nedenle, RAMML tahmin edicileri hesaplanırken $p = 16.5$ olarak alınır [18, 19]. Acıtaş vd. [21] basit ve doğrusal regresyon modelleri için RAMML tahmin edicilerinin performansını bilinen bazı dayanıklı regresyon tahmin edicilerinin performansları ile karşılaştırmak amacıyla kapsamlı bir simülasyon çalışması yapmıştır. Simülasyon çalışmasının sonucunda, veri setinde hem y - hem de x - yönlü aykırı değerler olması durumunda, RAMML tahmin edicilerinin daha iyi bir performansa sahip olduğu, normallik varsayımının sağlanması durumunda ise RAMML tahmin edicilerinin LS tahmin edicileri kadar etkin olduğu gösterilmiştir.

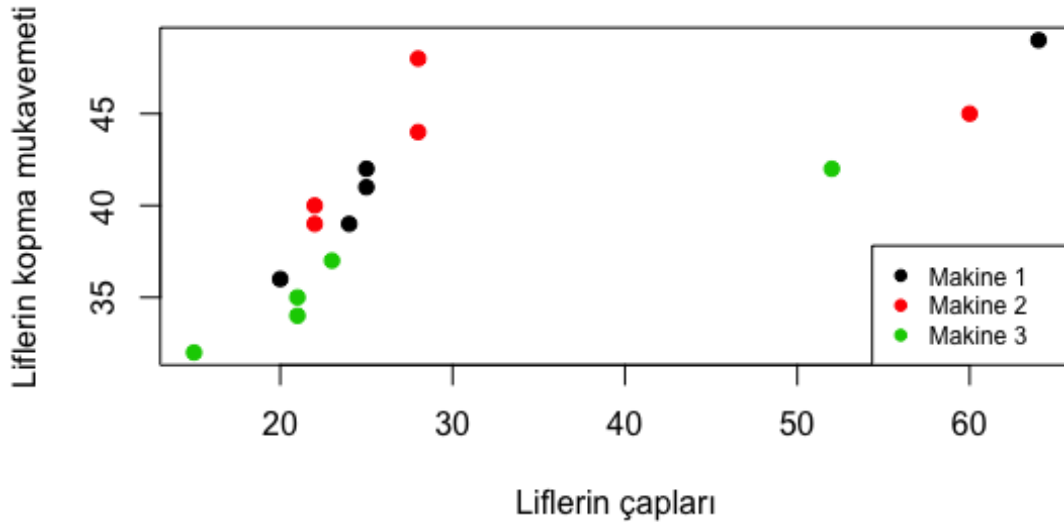
Uygulama

Silkknitter vd. [5] tarafından ele alınan veri seti, bu çalışmada geliştirilen dayanıklı ANCOVA yöntemi kullanılarak analiz edilmiştir. Söz konusu veri seti, üç farklı makinede üretilen liflerin kopma mukavemeti (y) ile liflerin çaplarını (x) içermektedir, bakınız Tablo 1.

Tablo 1. Liflerin kopma mukavemetleri ve çapları.

#	Makine 1		Makine 2		Makine 3	
	x	y	x	y	x	y
1	20	36	22	40	21	35
2	25	41	28	48	23	37
3	24	39	22	39	26	42
4	25	42	30	45	21	34
5	32	49	28	44	15	32

Silkknitter vd. [5], normallik varsayımı altında bir-yönlü ANCOVA modelini kullanarak bu veri setini analiz etmiştir, ayrıca bkz. Montgomery [26]. Şenoğlu ve Avcıoğlu [8], aynı veri seti için hata terimlerinin dağılımının genelleştirilmiş lojistik olduğunu göstermiş ve bu varsayım altında model parametrelerinin MML tahmin edicilerini elde etmiştir. Bu çalışmada, Tablo 1’deki veri seti kullanılarak Eşitlik (2)’de verilen ANCOVA modelinin parametrelerinin RAMML tahminleri elde edilmiştir, bkz. Tablo 2. Ayrıca, veri setinde yapay olarak aykırı değerler oluşturulmuştur. Bu amaçla, her bir makine için ortak değişkenin aldığı en büyük değer 2 ile çarpılmıştır. Bir başka ifade ile, $x_{15} = 32 \cdot 2 = 64$, $x_{24} = 30 \cdot 2 = 60$ ve $x_{33} = 26 \cdot 2 = 52$ olarak değiştirilmiş olup yanıt değişkeninin gözlem değerlerinde herhangi bir değişiklik yapılmamıştır. Aykırı değer içeren veri setinin serpilme grafiği Şekil 1’de verilmiştir. Şekil 1’de, oluşturulan üç aykırı değerden biri tanesinin (kırmızı renkli) x –yönlü aykırı değer olduğu görülmektedir. Aykırı değer eklenmiş veri seti kullanılarak elde edilen LS ve RAMML tahminleri de Tablo 2’de gösterilmiştir.



Şekil 1. Aykırı değer eklenmiş veri setinin serpilme grafiği

Tablo 2. Parametrelerin LS ve RAMML tahmin değerleri

Parametre	Orijinal veri		Aykırı değer içeren veri	
	LS	RAMML	LS	RAMML
μ_1	13.5722	8.0060	33.6545	4.7136
μ_2	20.9143	22.5001	39.9874	17.0094
μ_3	17.6790	12.6714	29.4556	22.2649
β_1	1.1043	1.3265	0.2451	1.4604
β_2	0.8571	0.8014	0.1004	1.0232
β_3	0.8642	1.0658	0.2479	0.6056
σ_1	0.8138	0.7315	2.1090	0.7101
σ_2	2.1336	2.0777	3.8540	1.6374
σ_3	1.7893	1.2196	1.3474	0.7618
σ	1.6749	1.4536	2.6531	1.1204

Tablo 2’den açıktır ki, veri setine aykırı değerler eklenmesi beklenildiği gibi LS tahminlerini önemli ölçüde etkilemiştir. Bir başka ifade ile parametrelerin LS tahminleri, veri setinin aykırı değer içerdiği ve içermediği durumlar için oldukça farklı çıkmışlardır. Örneğin, $\hat{\beta}_{1LS}$, $\hat{\beta}_{2LS}$ ve $\hat{\beta}_{3LS}$ aykırı değerlerden etkilendikleri için azalmışlardır. Bunun bir sonucu olarak, denemelere ilişkin regresyon doğruları aykırı değerlere yaklaşmıştır. Bununla birlikte, aykırı değer olsun ya da olmasın RAMML tahminleri büyük bir farklılık göstermemiş; hatta veri setinin aykırı değer içermesi durumunda $\hat{\sigma}_{RAMML}$ ’nin aldığı değer küçülmüştür. Bunun sebebi, RAMML yönteminin aykırı değerlere küçük ağırlık vermesidir. Tablo 3’te RAMML yönteminin gözlemlere verdiği $\delta_{ij}^x \delta_{ij}$ ağırlıkları gösterilmiştir. Tablo 3’ten açıktır ki, RAMML yöntemi Makine 1 için beşinci, Makine 2 için dördüncü ve Makine 3 için üçüncü gözleme çok küçük ağırlıklar vererek aykırı değerlerin etkisini azaltmıştır.

Tablo 3. Gözlemlere verilen ağırlıklar

#	Makine 1	Makine 2	Makine 3
1	6.014965e-02	8.679476e-01	9.997494e-01
2	9.942693e-01	6.866704e-01	8.094136e-01
3	8.116465e-01	8.673678e-01	2.212038e-06
4	9.283698e-01	1.952206e-04	9.057853e-01
5	7.194479e-12	9.651116e-01	3.355018e-01

Sonuç

Bu çalışmada, bir yönlü ANCOVA modelinde x –yönlü aykırı değer olması durumunda, model parametrelerinin RAMML tahmin edicileri elde edilmiştir. Model parametrelerinin RAMML tahmin edicileri analitik olarak ifade edilebildiğinden RAMML tahmin edicileri hesaplanma açısından kolaylık sağlamaktadır. RAMML tahmin edicilerinin performansını incelemek amacıyla, literatürde yer alan liflerin kopma kuvvetleri ve çapları ile ilgili veri setine yapay olarak aykırı değerler eklenmiş ve model parametrelerinin RAMML tahmin edicileri elde edilmiştir. RAMML tahmin edicileri aykırı değerlerin etkisini onlara küçük ağırlıklar vererek azalttığı gözlenmiştir. Ayrıca, hatanın varyansının RAMML tahmini, LS tahmininden daha küçük olarak elde edilmiştir.

Teşekkür Yazar, editörlere ve anonim hakemlere yorum ve katkıları için teşekkür eder.

Fon/Finansman Bilgileri Çalışma için herhangi bir mali sorumluluk yoktur.

Etik Kurul Onayı ve İzinler Çalışma etik kurul izni veya herhangi bir özel izin gerektirmemektedir.

Çıkar Çatışmaları/Çatışan Çıkarlar Makale için herhangi bir çıkar çatışması yoktur.

Yazarların Katkısı Yazar makalenin son halini okumuş ve onaylamıştır.

Kaynaklar

- [1] Fisher, R. A. (1932). *Statistical Methods for Research Workers*, 4th ed. Edinburgh: Oliver and Boyd Ltd.
- [2] Cochran, W. G. (1957). Analysis of covariance: Its nature and uses. *Biometrics*, 13, 261–281. <https://doi.org/10.2307/2527916>
- [3] Cox, D. R., & McCullagh, P. (1982). Some aspects of analysis of covariance. *Biometrics*, 38, 541–561. <https://doi.org/10.2307/2530040>
- [4] Smith, J. H., & Lewis, T. O. (1982). Effects of intraclass correlation on covariance analysis. *Communications in Statistics-Theory and Methods*, 11(1), 71–80.

- [5] Silknitter, K. O., Wisnowski, J. W., & Montgomery, D. C. (1999). The analysis of covariance: A useful technique for analysing quality improvement experiments. *Quality and Reliability Engineering International*, 15, 303–316. [https://doi.org/10.1002/\(SICI\)1099-1638\(199907/08\)15:4<303::AID-QRE253>3.0.CO;2-G](https://doi.org/10.1002/(SICI)1099-1638(199907/08)15:4<303::AID-QRE253>3.0.CO;2-G)
- [6] Şenoğlu, B., & Acıtaş, Ş. (2020). *İstatistiksel Deney Tasarımı: Sabit Etkili Modeller*, 4. Basım. Ankara: Nobel Yayın Dağıtım.
- [7] Şenoğlu, B. (2007). Estimating parameters in one-way analysis of covariance model with short-tailed symmetric error distributions. *Journal of Computational and Applied Mathematics*, 201, 275–283. <https://doi.org/10.1016/j.cam.2006.02.019>
- [8] Şenoğlu, B., & Avcıoğlu, M. D. (2009). Analysis of covariance with non-normal errors. *International Statistical Review*, 77(3), 366–377. <https://doi.org/10.1111/j.1751-5823.2009.00090.x>
- [9] Birch, J. B., & Myers, R. H. (1982). Robust analysis of covariance. *Biometrics*, 38, 699–713. <https://doi.org/10.2307/2530050>
- [10] Tiku, M. L., & Suresh, R. P. (1992). A new method of estimation for location and scale parameters. *Journal Statistical Planning and Inference*, 30, 281–292. [https://doi.org/10.1016/0378-3758\(92\)90088-A](https://doi.org/10.1016/0378-3758(92)90088-A)
- [11] Tiku, M. L., Islam, M. Q., & Selçuk, A. S. (2001). Nonnormal regression. II. Symmetric distributions. *Communications in Statistics Theory and Methods*, 30(6), 1021–1045. <https://doi.org/10.1081/STA-100104348>
- [12] Islam, M. Q., & Tiku, M. L. (2005). Multiple linear regression model under non-normality. *Communications in Statistics-Theory and Methods*, 33, 2443–2467. <https://doi.org/10.1081/STA-200031519>
- [13] Acıtaş, Ş., & Şenoğlu, B. (2018). Robust factorial ANCOVA with LTS error distributions. *Hacettepe Journal of Mathematics and Statistics*, 47(2), 347–363.
- [14] Tiku, M. L. (1967). Estimating the mean and standard deviation from a censored normal sample. *Biometrika*, 54, 155–165. <https://doi.org/10.1093/biomet/54.1-2.155>
- [15] Tiku, M. L. (1968). Estimating the parameters of normal and logistic distributions from censored samples. *Australian Journal of Statistics*, 10, 64–74. <https://doi.org/10.1111/j.1467-842X.1968.tb00216.x>
- [16] Bhattacharyya, G. K. (1985). The asymptotics of maximum likelihood and related estimators based on type II censored data. *Journal of the American Statistical Association*, 80, 398–404. <https://doi.org/10.1080/01621459.1985.10478130>
- [17] Vaughan, D. C., & Tiku, M. L. (2000). Estimation and hypothesis testing for a non-normal bivariate distribution and applications. *Mathematical and Computer Modelling*, 32, 53–67. [https://doi.org/10.1016/S0895-7177\(00\)00119-9](https://doi.org/10.1016/S0895-7177(00)00119-9)
- [18] Tiku, M. L., & Sürücü, B. (2009). MMLEs are as good as M-estimators or better. *Statistics and Probability Letters*, 79, 984–989. <https://doi.org/10.1016/j.spl.2008.12.001>
- [19] Dönmez, A. (2010). *Adaptive estimation and hypothesis testing methods*, (Tez No: 269583). [Doktora Tezi, Ortadoğu Teknik Üniversitesi]

- [20] Acitaş, Ş., Filzmoser, P., & Şenoğlu, B. (2019). *A robust adaptive modified maximum likelihood estimator for the linear regression model* [Conference presentation]. Olomoucian Days Applied Mathematics. Olomouc, Czech Republic; 2019.
- [21] Acitas, S., Filzmoser, P., & Senoglu, B. (2021). A robust adaptive modified maximum likelihood estimator for the linear regression model. *Journal of Statistical Computation and Simulation*, 91(7), 1394-1414. <https://doi.org/10.1080/00949655.2020.1856847>
- [22] Puthenpura, S., & Sinha, N. K. (1986). Modified maximum likelihood method for the robust estimation of system parameters from very noisy data. *Automatica*, 22, 231–235. [https://doi.org/10.1016/0005-1098\(86\)90085-3](https://doi.org/10.1016/0005-1098(86)90085-3)
- [23] Vaughan, D. C. (2002). The generalized secant hyperbolic distribution and its properties. *Communications in Statistics-Theory and Methods*, 31(2), 219-238. <https://doi.org/10.1081/STA-120002647>
- [24] Fritz H, Filzmoser P., & Croux C. (2012). A comparison of algorithms for the multivariate L1-median. *Computational Statistics*, 27, 393–410. <https://doi.org/10.1007/s00180-011-0262-4>
- [25] Maronna, R. A., Martin, R. D., & Yohai, V. J. (2006). *Robust Statistics: Theory and Methods*. Chichester: Wiley.
- [26] Montgomery, D. C. (2013). *Design and Analysis of Experiments*, 8th Edition. John Wiley & Sons, Inc.