

ANOVA Modellerinde Kareler Toplamı Yöntemlerinin Karşılaştırılması

Gül ERGÜN *  Serpil AKTAŞ *

* Hacettepe Üniversitesi, Fen Fakültesi, İstatistik Bölümü, Ankara - TÜRKİYE

Makale Kodu (Article Code): KVFD-2009-213

Özet

ANOVA modellerinde genel kareler toplamı ayrıştırılır ve ana etken ile etkileşim terimlerinin etkisinin anlamlılığı test edilir. Her bir ana etken ve etkileşim terimlerine ilişkin kareler toplamlarının hesaplanmasında, 1. Tür, 2. Tür, 3. Tür ve 4. Tür kareler toplamları olarak adlandırılan dört farklı yöntem mevcuttur. Araştırmalarda uygun yöntemin seçiminde model ve veri yapıları büyük önem taşımaktadır. Buna karşın, standart istatistiksel yazılımlarının çoğunda model ve veri yapıları gözetimeksizin varsayılan kareler toplamı yöntemi, 3. Tür olarak belirlenmiştir. Uygun olan yöntemin seçilmemesi F istatistiğinin değerini etkilemekte ve istatistiksel olarak hatalı sonuçlara neden olabilmektedir. Bu makalede kareler toplamına ilişkin mevcut yöntemlerin kullanıldığı durumlar vurgulanmış ve bu yöntemler arasındaki farklılıklar iki etkenli deney düzeni ile tanımlı iki örnek üzerinde tartışılmıştır.

Anahtar sözcükler: ANOVA, Çok etkenli deney düzenleri, Kareler toplamı

Comparisons of Sum of Squares Methods in ANOVA Models

Summary

The total sum of squares in ANOVA models is decomposed and the significances of the main term and interaction terms are tested. The four different methods called as Type I, II, III, and IV are used for calculating the sum of squares of the main effects and interactions. The model and data structures are crucial in the selection of the most appropriate method for applications. However, Type III method is adapted as the default sum of square method in most of the standard statistical software packages regardless of model and data structures. If the appropriate method is not selected, it directly affects the value of the F statistic and it may cause statistically invalid inferences. In this paper, the use of the sum of squares methods for various cases are indicated and the differences between the methods are discussed on two examples defined by a two-way ANOVA design.

Keywords: ANOVA, Factorial designs, Sum of squares

GİRİŞ

Bilimsel araştırmalarda bilinmeyen kitle ortalamasına ilişkin çıkarımlar yapmak büyük önem taşır. Bilindiği gibi tek etken için ikiden fazla kitle ortalaması arasındaki farkın önem kontrolü Tek yönlü Varyans Analizi (ANOVA) ile araştırılır. İki ya da daha fazla etken ve her bir deneme kombinasyonunda birden fazla tekrarın söz konusu olduğu denemelerde ise çok etkenli deney düzenleri söz konusudur.

A etkeninin düzey sayısı a , B etkeninin düzey sayısı b ve herbir deneme kombinasyonundaki tekrar sayısı k olmak üzere iki faktörlü bir ANOVA modeli,

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk} \quad i=1, \dots, a; j=1, \dots, b; k=1, \dots, k$$

eşitliği ile tanımlanır.

Burada,

y_{ijk} : A etkeninin i . düzeyi, B etkeninin j . düzeyine karşılık gelen deneme kombinasyonunda bulunan bağımlı değişken,

α_i : A etkeninin i . düzeyinin etki miktarı,

β_j : B etkeninin j . düzeyinin etki miktarı,

$(\alpha\beta)_{ij}$: A etkeninin i . düzeyinin etkisi ile B etkeninin j . düzeyinin birlikte etki miktarı,

n_{ij} : herbir deneme kombinasyonundaki tekrar sayısı ve

ε_{ijk} : hata terimidir.

Parametre kestirimleri için $\sum \alpha_i = 0$, $\sum \beta_j = 0$, $\sum \sum$



İletişim (Correspondence)



+90 312 2977907



gul@hacettepe.edu.tr

$(\alpha\beta)_{ij} = 0$, $\sum\sum(\alpha\beta)_{ij} = 0$ kısıtları söz konusudur ve ANOVA ile ana etken ve etkileşim terimleri için hipotezler test edilir ¹. Tüm deneme kombinasyonlarında n_{ij} 'ler eşit ise incelenen düzene "dengeli düzen", n_{ij} 'ler eşit değil ise "dengesiz düzen" adı verilir. Herhangi bir n_{ij} 'de gözlem olmadığında, "kayıp gözlem" varlığı söz konusudur ve tüm bu özellikler verilerin analizi aşamasında dikkate alınmalıdır.

ANOVA modelleri genel doğrusal modeller yardımıyla çözülür ve kareler toplamlarının hesaplanmasında 1. Tür, 2. Tür, 3. Tür ve 4. Tür kareler toplamları olarak adlandırılan dört farklı yöntem kullanılır ^{2,4}. Bu yöntemlerde ana etki ve etkileşimlere ait kareler toplamları farklı şekillerde elde edilir. Bu nedenle, yöntemden yonteme hesaplanan F değerleri ile hipotez testinin sonuçları farklılık gösterebilir. Makalede model ve veri yapılarına bağlı olarak, yöntemlerin sonuçlar üzerindeki olası farklılıklarının vurgulanması amaçlanmıştır. Makalenin izleyen bölümlerinde kareler toplamı hesaplama yöntemlerinin kullanım alanları kısaca gözden geçirilmiş; yöntemlerin sonuçlara etkileri örneklerle gösterilmiş ve son olarak elde edilen bulgular tartışılmıştır.

KARELER TOPLAMI HESAPLAMA YÖNTEMLERİNİN KULLANIM ALANLARI

Birinci Tür yöntem, her bir ana etki veya etkileşimin sırayla modele eklendiği bir yöntemdir. Bu nedenle bu yöntemle elde edilen etki veya etkileşimlerin kareler toplamı, ardışık kareler toplamı olarak da bilinir. Birinci Tür yöntem ile elde edilen kareler toplamı, ana etki veya etkileşimlerin doğal sırada (Ana etkilerin ikili etkileşimlerden önce yer aldığı, ikili etkileşimlerin üçlü etkileşimlerden önce yer aldığı, üçlü etkileşimlerin dördü etkileşimlerden önce yer aldığı ve böyle devam eden durum) yer aldığı modelde dengeli ANOVA düzenlerinde kullanılmalıdır.

İkinci Tür yöntem ise, kısmi ardışık kareler toplamı olarak bilinir. Bu yöntem, ana etki ve etkileşimlerin modelde bulunma önceliğinden, diğer bir ifade ile değişkenlerin sırasından etkilenmez. Modelde sadece ana etkenlerin bulunduğu durumlar için bu yöntem daha uygundur. Dengeli çok etkenli deney düzenlerinde de kullanılabilir. Buna göre, İkinci Tür yöntem özellikle sadece ana etkilerin bulunduğu modellerde en güçlü sonucu verir ¹.

Dengesiz düzenlerde 1. Tür ve 2. Tür yöntemler ile kareler toplamlarının bulunması uygun bir yaklaşım

değildir. Bu durumda 3. Tür yöntemin kullanılması önerilebilir. Üçüncü Tür yöntem ile elde edilen kareler toplamları, kayıp gözlem içermeyen ve dengeli olmayan ANOVA modelleri için uygundur. Ayrıca bu yöntem örneklem büyüklüğünden etkilenmez.

ANOVA düzenlerinde bazı hücrelerde kayıp gözlem olabilir. Dengeli ya da dengesiz düzenlerde kayıp gözlem durumunda 4. Tür yöntem hem ana etki hem de etkileşimlere ait kareler toplamlarının elde edilmesinde tercih edilmelidir.

Yukarıda belirtilen farklılıklara rağmen, SAS, SPSS, MINITAB, STATISTICA, SYSTAT, STATA, UNISTAT gibi yaygın olarak kullanılan İstatistik paket programlarında varyasyon (default) yöntem, 3. Tür kareler toplamı yöntemidir. Ancak, bu makalede belirtildiği gibi, model ve veri yapılarına göre uygun yöntemin seçilmesi ve test sürecinin buradan elde edilen sonuca dayandırılması, araştırma sonuçları açısından büyük önem taşımaktadır.

YÖNTEMLERİN SAYISAL ÖRNEKLER İLE KARŞILAŞTIRILMASI

Bu bölümde iki farklı örnek veri kümesi üzerinde dengeli, dengesiz ve kayıp veri durumu incelenmiş, 1., 2., 3. ve 4. Tür yöntemlere göre analizler yapılmış ve sonuçlar tartışılmıştır. Burada amaç bu yöntemler arasındaki olası farklılıkların vurgulanması olduğundan, makalede sadece iki etkenli deney ve sabit seçimli model ele alınmıştır.

Örnek 1

Bu örnekte patateslerin büyük bölümünün toplanıldıktan sonra dayanıklılığını kaybetmesi üzerine bir deney planlanmış ve çürümeye neden olan etkenler incelenmiştir. Patateslerin toplandıktan sonra saklama sıcaklığı (10°C ve 16°C) ve saklanan ortamın oksijen miktarına (%2, %6, %10) göre çürüme yüzdeleri hesaplanmış ve bu değerler *Tablo 1*'de verilmiştir. Burada, her bir deneme kombinasyonunda eşit sayıda gözlem mevcuttur ve dengeli bir düzen söz konusudur.

Bu düzen için analiz sonuçları *Tablo 2*'de verilmiştir. Burada sd: serbestlik derecesi; KT: kareler toplamını ifade etmektedir. Dengeli düzen için dört farklı yöntemden elde edilen sonuçlar arasında bir fark olmadığı görülmüştür.

Burada *Tablo 1*'de verilen veri setinde bazı gözlemlerin elde edilemediği varsayılmış ve *Tablo 3*'te verilen dengesiz bir deney düzeni elde edilmiştir.

Tablo 1. Patates dayanıklılığı verisi için dengeli düzen
Table 1. Balanced design for potatoes' resistance data

| Sıcaklık | Oksijen Miktarı | | |
|----------|-----------------|----|-----|
| | %2 | %6 | %10 |
| 10°C | 13 | 10 | 15 |
| | 15 | 11 | 17 |
| | 11 | 20 | 9 |
| | 12 | 21 | 10 |
| 16°C | 26 | 15 | 19 |
| | 26 | 16 | 17 |
| | 19 | 22 | 24 |
| | 20 | 23 | 25 |

Tablo 2. Dengeli deney düzeni için ANOVA sonuçları
Table 2. ANOVA results for balanced design

| Değişim Kaynağı | sd | Kareler Toplamı (1. 2. 3. 4. Tür) | F | P |
|-----------------|----|-----------------------------------|--------|-------|
| Sıcaklık (S) | 1 | 322.667 | 19.890 | 0.000 |
| Oksijen (O) | 2 | 2.333 | 0.072 | 0.931 |
| S x O | 2 | 46.333 | 1.428 | 0.266 |
| Hata | 18 | 292 | | |

Tablo 3. Patates dayanıklılığı verisi için dengesiz deney düzeni
Table 3. Unbalanced design for potatoes' resistance data

| Sıcaklık | Oksijen Miktarı | | |
|----------|-----------------|----|-----|
| | %2 | %6 | %10 |
| 10°C | 13 | 10 | 15 |
| | 15 | 11 | 17 |
| | 11 | 20 | |
| | 12 | | |
| 16°C | 26 | 15 | 19 |
| | 26 | 16 | |
| | 19 | | |
| | 20 | | |

Bu düzene ilişkin analiz sonuçları ise, *Tablo 4*'te verilmiştir. Tabloda "Sıcaklık" etkenine ait kareler toplamının, 3. Tür ve 4. Tür yöntemlerde aynı, 1. Tür ve 2. Tür yöntemlerde ise, farklı olduğu görülmektedir. Buna göre, %1 anlamlılık düzeyinde 1.Tür ve 2. Tür kareler toplamı hesaplama yöntemleriyle yokluk hipotezi reddedilirken; 3. Tür ve 4. Tür yöntemlerde yokluk hipotezi reddedilememiştir.

Tablo 4. Dengesiz deney düzeni için ANOVA sonuçları
Table 4. ANOVA results for unbalanced design

| Değişim Kaynağı | sd | Kareler Toplamı (1. 2. 3. 4. Tür) | F | P |
|-----------------|----|-----------------------------------|-------|-------|
| Sıcaklık (S) | 1 | 159.52 ^a | 13.91 | 0.004 |
| | | 151.45 ^b | 13.21 | 0.005 |
| | | 77.657 ^{c,d} | 6.77 | 0.026 |
| Oksijen (O) | 2 | 27.17 ^{a,b} | 1.18 | 0.345 |
| | | 32.02 ^{c,d} | 1.39 | 0.292 |
| S x O | 2 | 58.57 ^{a,b,c,d} | 2.55 | 0.127 |
| Hata | 10 | 114.66 | | |

^a: 1. Tür, ^b: 2. Tür, ^c: 3. Tür, ^d: 4. Tür

Makalede kayıp gözlem durumunda yöntemlerin karşılaştırılması için, *Tablo 1*'de yer alan verilerde "Sıcaklık=16°C" ve "Oksijen miktarı=%10" deneme kombinasyonuna karşılık gelen hücre boş bırakılmış ve aşağıdaki *Tablo 5* elde edilmiştir.

Tablo 5. Kayıp gözlem için deney düzeni

Table 5. Experimental design for missing observation

| Sıcaklık | Oksijen Miktarı | | |
|----------|-----------------|----|-----|
| | %2 | %6 | %10 |
| 10°C | 13 | 10 | 15 |
| | 15 | 11 | 17 |
| | 11 | 20 | 9 |
| | 12 | 21 | 10 |
| 16°C | 26 | 15 | 19 |
| | 26 | 16 | 17 |
| | 19 | 22 | 24 |
| | 20 | 23 | 25 |

Dört farklı yöntem uygulanarak elde edilen sonuçlar *Tablo 6*'da verilmiştir. Burada görüldüğü gibi "Oksijen" etkeni için 4. Tür kareler toplamı farklılık göstermiştir.

Tablo 6. Kayıp gözlem düzeni için ANOVA sonuçları

Table 6. ANOVA results for missing value design

| Değişim Kaynağı | sd | Kareler Toplamı (1. 2. 3. 4. Tür) | F | P |
|-----------------|----|-----------------------------------|-------|-------|
| Sıcaklık (S) | 1 | 249.41 ^a | 15.13 | 0.001 |
| | | 182.25 ^{b,c,d} | 11.05 | 0.005 |
| Oksijen (O) | 2 | 6.04 ^{a,b,c} | 0.18 | 0.834 |
| | | 20.16 ^d | 0.61 | 0.555 |
| S x O | 2 | 42.25 ^{a,b,c,d} | 0.18 | 0.130 |
| Hata | 10 | 247.25 | 0.61 | |

^a: 1. Tür, ^b: 2. Tür, ^c: 3. Tür, ^d: 4. Tür

Makalede dengesiz düzenlerde yöntemler arası farklılığın vurgulanması için aşağıda verilen örnek ele alınmıştır.

Örnek 2

Burada, bir ilaçtaki etken maddenin, depolama zamanı ve depolama süresinden etkilenip etkilenmediğine ilişkin bir deney planlanmış ve 16 kutu ilaç, 3 ay ve 6 ay depolama süresi ve 20°C ve 30°C ortam sıcaklığında depolanmış ve deney sonunda kaybedilen etken madde miktarı ölçülmüştür. Ancak laboratuvara taşıma sırasında 6 kutu hasar gördüğünden *Tablo 7*'de verilen dengesiz bir deney düzeni ortaya çıkmıştır.

Tablo 7'de verilen dengesiz deney düzeni için ANOVA sonuçları *Tablo 8*'de verilmiştir.

Tablo 7. İlaç depolama verisi**Table 7.** Drug storage data

| Depolama Zamanı | Depolama Sıcaklığı | |
|-----------------|--------------------|------|
| | 20°C | 30°C |
| 3 ay | 2 | 9 |
| | 5 | 12 |
| | | 15 |
| 6 ay | 6 | |
| | 6 | 16 |
| | 7 | |
| | 7 | |

Tablo 8. İlaç depolama verisi için ANOVA sonuçları**Table 8.** ANOVA results for drug storage data

| Değişim Kaynağı | sd | Kareler Toplamı (1. 2. 3. 4. Tür) | F | P |
|-----------------|----|-----------------------------------|-------|-------|
| Zaman | 1 | 0.100 ^a | 0.026 | 0.878 |
| | | 23.520 ^{b,c,d} | 6.005 | 0.049 |
| Sıcaklık | 1 | 158.42 ^{a,b} | 40.45 | 40.45 |
| | | 155.52 ^{c,d} | 39.71 | 39.71 |
| Z x S | 1 | 0.48 ^{a,b,c,d} | 0.123 | 0.738 |
| Hata | 6 | 23.5 | | |

^a: 1. Tür, ^b: 2. Tür, ^c: 3. Tür, ^d: 4. Tür

Sonuçlar incelendiğinde “Zaman” etkeni için kareler toplamının ve *F* değerlerinin büyük ölçüde farklı olduğu görülmektedir. Burada, 1. Tür yönteme göre yokluk hipotezi %5 anlamlılık düzeyinde reddedilemezken, diğer yöntemlere göre yokluk hipotezinin reddedildiği görülmektedir.

SONUÇ

Kareler toplamı hesaplama yöntemleri model ve veri yapılarına bağlı olarak farklı sonuçlar verebilir. Araştırmalarda uygun yöntemin seçilmesi gerekir. 1. Tür yöntem, ana etki ve etkileşim terimlerinin modele katılma sırasını dikkate aldığından dolayı, özellikle dengeli düzenlerde kullanılmalıdır. İkinci Tür yöntem ise, 1. Tür yöntemin aksine etkenlerin modele katılma sıralamasını dikkate almaz ve bu yöntem de dengeli düzenlerde kullanılabilir. Ancak, 2. Tür yöntem sadece ana etkilerin oluşturduğu modellerde istatistiksel açıdan en iyi sonucu verir. Üçüncü Tür yöntem ise, kayıp gözlem içermeyen dengeli ya da dengesiz düzenlerde uygundur. Dördüncü Tür yöntem ise, kayıp gözlem içeren dengeli ya da dengesiz düzenlerde kullanılmamalıdır. Bunun yanında dengeli düzenlerde dört yöntem her zaman aynı sonucu verir. Makalede örnek olarak alınan iki farklı veri seti için kayıp gözlem içermeyen dengesiz düzenlerde, 2. Tür, 3. Tür ve 4. Tür yöntemler benzer sonuçlar vermiştir. Kayıp gözlem varlığında

ise, 4. Tür kareler toplamı yöntemi diğer yöntemlere göre farklı sonuçlar vermiştir. Bu yazıda yöntemlerin karşılaştırılmasında sadece iki etkinli bir deney düzeni ele alınmıştır. Etken sayısının ikiden fazla olduğu durumlar ele alınabilir ve kapsamlı bir simülasyon çalışması ile elde edilen sonuçlar genellenebilir. Ancak, bu makalenin hedefi öncelikle yöntemlerin sonuçlar üzerindeki olası farklılığının sergilenmesi ve araştırmacıların bu konuda özen göstermeleri gerektiğinin vurgulanmasıdır.

Literatürde dengesiz durumlarda hangi yöntemin kullanılmasının gerektiğine ilişkin bir uzlaşmaya varılmamış olsa da genel kanı 3. Tür’ün kullanılmasından yanadır. Bu yüzden SAS, SPSS, JMP, Minitab, Stata, Statistica, Systat ve Unistat gibi programlar, değiştirilmediği sürece kareler toplamlarını 3. Tür’e göre; R, S-Plus, Genstat ve Mathematica programları ise, 1. Tür’e göre hesaplamaktadır.

Sonuç olarak, araştırmacılar deney planını kurup verileri topladıktan sonra kurduğu bu düzenin dengeli ya da dengesiz olduğuna, verilerde kayıp gözlem olup olmadığına analiz aşamasında dikkat etmeli ve ardından hangi kareler toplamının uygun olacağına karar vermelidir. İstatistik paket programlarının çoğunda varsayılan yöntem 3. Tür olduğu için, program kullanıcıları verinin yapısına bakmaksızın 3. Tür yönteme dayalı sonuçları kullanmaktadır. Ancak bu yazıda vurgulandığı gibi, test sonuçları yöntemlere göre farklılık gösterebilmektedir. Bu nedenle Genel Doğrusal Modeller kullanılarak elde edilen çözümlerlerde, yukarıda sözü edilen yöntemlerden en uygun olanı seçilmelidir. Burada uygun yöntemin seçimi, *F* istatistiğinin değerini etkileyeceğinden ana etki ya da etkileşimlere ait hipotez testlerinin sonucunda farklılık gösterecektir. Bu nedenle uygun yöntemin kullanımı istatistiksel olarak hata yapma olasılığını azaltacaktır.

KAYNAKLAR

- Montgomery DC:** Design and Analysis of Experiments. John Wiley & Sons. 5th ed., pp. 60-103, Singapore, 1991.
- Pendelton OJ, Tress M, Bremer V:** Interpretation of the four types of analysis of variance tables in SAS. *Commun in Statist - Theory and Methods*, 15 (9): 2785-2808, 1986.
- Tanguma J, Speed FM:** Interpreting the four types of sums of squares in SPSS. Annual meeting of the Mid-South Educational Research Association, Bowling Green, KY, November 16, 2000.
- Langsrud Ø:** ANOVA for unbalanced data: Use Type II instead of Type III sum of squares. *Statistics and Computing*, 13, 63-167, 2003.