

Web Tabanlı İstatistiksel Analiz: Faktöriyel Varyans Analizi ve Etkileşim Etkisinin Çoklu Karşılaştırma Yöntemi ile İncelenmesi

Mehmet Ali SUNGUR¹, Handan ANKARALI, Şengül CANGÜR

Düzce Üniversitesi, Tıp Fakültesi, Biyoistatistik ve Tıbbi Bilişim ABD, Düzce
malisungur@yahoo.com, handanankarali@gmail.com, sengulcangur@duzce.edu.tr
 (Geliş/Received:27.01.2016; Kabul/Accepted:07.10.2016)
 DOI: 10.17671/btd.66022

Özet—Yaygın kullanılan istatistik paket programlarının birçoğunda yer almadığı için elle hesaplamak zorunda kalınan bazı istatistiksel yöntemler bulunmaktadır. Faktöriyel denemelerde etkileşim etkisini değerlendirmek için gereken çoklu karşılaştırmalar bunlardan biridir. Bu çalışmanın amacı, iki yönlü faktöriyel denemenin analizini ve etkileşime ait sıfır hipotezi reddedildiğinde çoklu karşılaştırma yapan bir modül geliştirmektir. Geliştirilen modül, PHP ile oluşturulmuş ve JavaScript ile işlevsellik kazandırılmıştır. Dengeli tasarımlarda özet istatistiklerden faktöriyel varyans analizi yapan ek bir modül daha geliştirilmiştir. İnternet bağlantısı olan her yerden ve cihazdan modüle kolayca ulaşılabilen ve tüm tarayıcılar aracılığıyla görüntülenebilmektedir. Kullanıcı istenen işlemleri tamamladığında, anlaşılması kolay çıktılar ve kısa yorumlar verilmektedir. Tanımlayıcı istatistikler, varyans analiz tablosu ve çoklu karşılaştırmalar tek bir pencerede gösterilmektedir. Elle hesaplamak için uzman düzeyde formül ve yöntem bilinmesi gereken, hesaplaması zor ve zaman alan istatistiksel yöntemlerin, web tabanlı olarak kolay ve hızlı bir şekilde hesaplanması ve yorumlanması mümkündür. Modülün web tabanlı olması kolay erişilebilirlik gibi önemli bir avantaj sağlayacaktır. Matematiksel modeli bilinen bir istatistiksel yöntemin web tabanlı ve kullanımı kolay ara yüzlerle uygulanabilmesi mümkündür.

Anahtar Kelimeler—Web tabanlı istatistik hesaplamalar, iki yönlü varyans analizi, etkileşim, çoklu karşılaştırma, PHP

Web-Based Statistical Analysis: Factorial Analysis of Variance and Investigation of Interaction Effect by Multiple Comparison Method

Abstract—There are some statistical methods have to calculate manually, since not included in many of the commonly used statistical packages. Multiple comparisons required to evaluate the effect of interaction in factorial experiments is one of these methods. The aim of this study is to develop a module that makes analysis of a two-way factorial experiment and multiple comparison in case of the null hypothesis of interaction is rejected. The developed module was established with PHP and functionality was gained with JavaScript. An additional module that makes factorial analysis of variance through summary statistics in balanced designs was developed. The module can be easily accessed from anywhere and device with internet connection, and viewed via all browsers. When user completed requested operations, simple to understand outputs and short reviews are yielded. Descriptive statistics, analysis of variance table, and multiple comparisons are shown in a single window. It is possible to calculate and review the statistical methods that require knowing expert-level formulas and methods, difficult and time-consuming to calculate, easily and quickly as web-based. The fact that the module is web-based will provide a significant advantage as easy accessibility. It is possible to make any statistical method that has known mathematical model, web-based and with interfaces easy to use.

Keywords—Web-based statistical calculations, two-way analysis of variance, interaction, multiple comparison, PHP

1. GİRİŞ (INTRODUCTION)

Bilgisayar ve internet kullanımının oldukça yaygın olduğu, günlük hayatta ve iş hayatındaki neredeyse tüm işlemlerin web üzerinden yapıldığı çağımızda, verinin olduğu tüm bilimsel çalışmalar için gerekli olan istatistiksel analizlerin de web üzerinden yapılabilmesi önemli bir kolaylık sağlayacaktır. Ayrıca internet üzerinden hızlı erişim sağlanması, web sayfasının hızlı yüklenebilen ve herhangi bir web tarayıcısı aracılığıyla görüntülenebilen bir sayfa olması da tercih ve pratik kullanım açısından oldukça önemli bir konudur. Bu şekilde, erişim kolaylığı yanında program kurulumundan ve zamandan bağımsız işlem yapabilme olanağı da cazip bir özellik olmaktadır. Bunun yanında, istatistiksel analiz yapmak için kullanılacak olan web sayfasının temel bazı analizler için istatistik uzmanı olmayan kişilerin de rahat bir şekilde kullanabileceği ve sonuçları basitçe yorumlayabileceği bir şekilde düzenlenmesi de kullanım kolaylığı ve tercih edirlilik açısından ek katkı sağlayacaktır. Bu doğrultuda, programlama bilgisi gerektirmeyen, bilgisayarda kurulu olan platformdan bağımsız olarak çalışabilen ve tüm web tarayıcılar aracılığıyla sorunsuz olarak görüntülenebilen web sayfaları üzerinden veri girişlerinin yapılmasının ardından bir ya da birkaç basit aşamada istatistiksel analizlerin sonuçlandırılabilmesi kullanıcılar için öncelikli tercih sebebi olacaktır.

Ayrıca teknolojik gelişmeler ve mevcut paket programların sayısının çokluğuna rağmen halen istatistiksel analiz paket programlarının birçoğunda yer almadığı için elle hesaplanmak zorunda kalınan istatistiksel yöntemler bulunmaktadır. Bu yöntemlerden birisi de çalışmada incelenen bağımlı değişkenle ilgili olabilecek birden fazla bağımsız değişkenin (faktörün), bağımlı değişken üzerindeki birlikte etkilerinin araştırıldığı faktöriyel tasarımlarda, bağımsız değişkenlerin etkileşim (interaksiyon) etkisine ait çoklu karşılaştırmalardır.

İstatistiksel modellerde genel olarak bağımlı değişken üzerine bir ya da daha fazla bağımsız değişkenin etkisi araştırılmaktadır. Tek yönlü (modelde bir bağımlı değişken ve bir de bağımsız değişkenin olduğu) modellerde, bağımlı değişkendeki değişimin yeterince açıklanması mümkün olamayabilir ya da modele alınmayan başka bir değişkenin, modeldeki bağımlı değişken üzerindeki etkisi nedeniyle farklı bir sonuca ulaşılabilir. Bu nedenle, özellikle biyolojik verilerin analizinde, incelenen bağımlı değişkenle ilgili olabilecek birden fazla bağımsız değişkenin birlikte etkisinin araştırıldığı faktöriyel tasarımların kullanılması daha uygun olmaktadır [1-6]. Faktöriyel tasarım modellerinde, bağımsız değişkenlerin ayrı ayrı esas (ana) etkilerinin belirlenmesinin yanında, bağımlı değişken üzerine birlikte etkilerini gösteren etkileşim (interaksiyon) etkisi de ortaya konabilmektedir. Etkileşime ait kontrol hipotezinin reddedilmesi durumunda, çoklu karşılaştırma testleri

kullanılarak faktör seviyeleri arasındaki basit etkiler karşılaştırılmaktadır [1,3]. Yaygın kullanılan istatistiksel analiz paket programlarında, çok faktörlü denemeler analiz edilebilmesine rağmen etkileşim etkisinin anlamlı bulunması durumunda, birçok programda bu etkinin değerlendirilmesi yapılamamaktadır.

Bu çalışmada, faktöriyel bir denemenin esas etkileri ve etkileşim etkisinin incelenemediği, anlamlı bulunan etki/etkiler için gerekli çoklu karşılaştırmaların yapılabildiği web tabanlı bir modül geliştirilmesi amaçlanmıştır. Bunun yanında, tanımlayıcı istatistiklerin hesaplanması ve tanımlayıcı istatistiklerle temel düzeyde grup karşılaştırmalarının yapılması da ele alınmıştır. Böylece hem birçok paket programlarda yer almayan istatistiksel analiz yöntemlerine web tabanlı çözüm getirilebileceği hem de web tabanlı hesaplama araçlarıyla interaktif istatistiksel değerlendirmeler yapılabileceği tartışılmıştır.

2. GEREÇ VE YÖNTEM (MATERIAL AND METHOD)

2.1. 2² Faktöriyel Tasarımlar (2² Factorial Designs)

Sürekli yapıdaki bir tane bağımlı değişkenin üzerine, kategorik yapıdaki iki bağımsız değişkenin (faktörün) etkisinin incelendiği denemeler iki faktörlü denemeler olarak adlandırılır [1-7]. İki faktörlü denemelerde faktörlerin seviye sayılarına bağlı olarak çeşitli kombinasyonlarda veri matrisi ortaya çıkabilir. İki faktörlü bir tasarım için en basit düzen, her iki faktörün de iki seviyesinin olduğu (2x2) dört kombinasyonlu durumdur. Tablo 1'de 2x2 düzeninde bir faktöriyel tasarımın seviye kombinasyonları görülmektedir.

Tablo 1. İki faktörlü (2x2) faktöriyel tasarım (Factorial design with two factor (2x2))

A	B	Y
1	1	Y ₁₁
1	2	Y ₁₂
2	1	Y ₂₁
2	2	Y ₂₂

Faktör seviyelerinin her bir kombinasyonu için denek sayısının eşit olup olmamasına göre dengeli tasarımlar ve dengesiz tasarımlar söz konusudur. Dengeli tasarımlarda Tip I kareler toplamı hesaplanabilirken dengesiz tasarımlarda ise regresyon yaklaşımı ile Tip III kareler toplamı hesaplanması önerilmektedir [3,4,8,9]. Tip I kareler toplamının hesaplanmasında kullanılan formüller Eşitlik (1)-(5) arasında verilmiştir [1,3,6,7]. Sırasıyla genel kareler toplamı (1), A faktörüne (2) ve B faktörüne (3) ait kareler toplamı, hata kareler toplamı (4) ve etkileşim kareler toplamının (5) hesaplanmasında kullanılan formüller gösterilmiştir.

$$GnKT = \sum X^2 - \frac{(\sum x)^2}{N} \quad (1)$$

$$AKT = \left[\frac{(\sum a_1)^2}{n_{a1}} + \frac{(\sum a_2)^2}{n_{a2}} \right] - \frac{(\sum x)^2}{N} \quad (2)$$

$$BKT = \left[\frac{(\sum b_1)^2}{n_{b1}} + \frac{(\sum b_2)^2}{n_{b2}} \right] - \frac{(\sum x)^2}{N} \quad (3)$$

$$HKT = \left(\sum a_1 b_1^2 - \frac{(\sum a_1 b_1)^2}{n_{a1 b_1}} \right) + \left(\sum a_1 b_2^2 - \frac{(\sum a_1 b_2)^2}{n_{a1 b_2}} \right) + \left(\sum a_2 b_1^2 - \frac{(\sum a_2 b_1)^2}{n_{a2 b_1}} \right) + \left(\sum a_2 b_2^2 - \frac{(\sum a_2 b_2)^2}{n_{a2 b_2}} \right) \quad (4)$$

$$EKT = GnKT - HKT - AKT - BKT \quad (5)$$

Tip III kareler toplamının hesaplanmasında ise dizayn matrisi oluşturularak kareler toplamları hesaplanmaktadır. İki faktörlü 2x2 düzeninde bir faktöriyel tasarımın regresyon modeli Eşitlik (6) ile ve dizayn matris yapısı Tablo 2'de görülmektedir [1-3,7]. Dizayn matrisi oluşturulduktan sonra X ve Y matrislerinin oluşturulması Eşitlik (7) ve (8) ile gösterilmiş, Tip III kareler toplamının hesaplanmasına yönelik matris işlemleri ve hesaplama adımları ise Eşitlik (9)-(13) arasında verilmiştir [10].

Tablo 2. 2x2 faktöriyel tasarım için dizayn matrisi
(Design matrix for 2x2 factorial design)

M	A	B	AB	Y
1	1	1	1	Y ₁₁
1	1	-1	-1	Y ₁₂
1	-1	1	-1	Y ₂₁
1	-1	-1	1	Y ₂₂

$$Y = \mu + b_0 + b_1 X_1 + b_2 X_2 + b_{12} X_1 X_2 + e \quad (6)$$

$$X = \begin{bmatrix} 1 & X_{11} & X_{12} & \dots & X_{1k} \\ 1 & X_{21} & X_{22} & \dots & X_{2k} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & X_{n1} & X_{n2} & \dots & X_{nk} \end{bmatrix} \quad (7)$$

$$Y = \begin{bmatrix} Y_1 \\ Y_2 \\ \dots \\ Y_n \end{bmatrix} \quad (8)$$

Tip III kareler toplamı hesaplanırken, Eşitlik (7) ve (8) ile oluşturulan X ve Y matrisleri kullanılarak yapılacak matris işlemlerinde, ayrı ayrı dört model kurularak dört ayrı hata kareler toplamı hesaplanır. Bunlar, A ve B faktörlerinin sadece esas etkilerinin dikkate alındığı modelde hata kareler toplamı (HKT_{A,B}), A faktörünün esas etkisi ile AxB etkileşim etkisinin dikkate alındığı modelde hata kareler toplamı (HKT_{A,AB}), B faktörünün esas etkisi ile AxB etkileşim etkisinin dikkate alındığı modelde hata kareler toplamı (HKT_{B,AB}) ve son olarak tam model olarak da adlandırılan A ve B faktörlerinin esas etkileri ile AxB etkileşim etkisinin dikkate alındığı

modelde hata kareler toplamıdır (HKT_{A,B,AB}). Bu dört farklı hata kareler toplamının hesaplanması için kurulacak modellerde matris işlemleri Eşitlik (9) ile verilmiştir.

$$HKT = \{Y - X[(X'X)^{-1}X'Y]\}'\{Y - X[(X'X)^{-1}X'Y]\} \quad (9)$$

Matris işlemleri yardımıyla Eşitlik (9) ile verilen işlemlerle elde edilecek hata kareler toplamlarındaki farklılık, modele alınmayan etkilerden kaynaklanmakta, bu etkiye ait kareler toplamı da hata kareler toplamına eklenmektedir. Dolayısıyla her bir modele ait hata kareler toplamı ile tam modele ait hata kareler toplamının farkı alınarak model dışı bırakılan etkinin kareler toplamı bulunabilmektedir. Bu şekilde varyans analizi tablosunda kullanılmak üzere sırasıyla A faktörüne ait kareler toplamı (10), B faktörüne ait kareler toplamı (11), etkileşim kareler toplamı (12) ve hata kareler toplamının (13) hesaplanmasında kullanılan formüller gösterilmiştir. Genel kareler toplamı ise Tip I kareler toplamı için Eşitlik (1) ile verilen formülle aynı şekilde verilerden hesaplanır. Hata kareler toplamı tam model için Eşitlik (9) ile matris işlemleriyle elde edilen hata kareler toplamı olup istenirse yine Tip I kareler toplamı için Eşitlik (4) ile verilen formülle aynı şekilde verilerden de hesaplanabilir.

$$AKT = HKT_{B,AB} - HKT_{A,B,AB} \quad (10)$$

$$BKT = HKT_{A,AB} - HKT_{A,B,AB} \quad (11)$$

$$EKT = HKT_{A,B} - HKT_{A,B,AB} \quad (12)$$

$$HKT = HKT_{A,B,AB} \quad (13)$$

Faktöriyel tasarımlarda dizayn matrisi ve X ve Y matrisleri oluşturulduktan sonra, Eşitlik (9) ile verilen matris işlemleri ve Eşitlik (10)-(13) arasında verilen kareler toplamlarının hesaplanması, faktörlerin seviye sayılarından bağımsız olarak aynı şekilde yapılmaktadır. Tablo 3'te 2x3 düzeninde ve Tablo 4'te 2x4 düzeninde bir faktöriyel tasarımın dizayn matrisi yapısı verilmiştir.

Tablo 3. 2x3 faktöriyel tasarım için dizayn matrisi
(Design matrix for 2x3 factorial design)

A	B ₁ B ₂	B ₁ B ₃	AxB ₁ B ₂	AxB ₁ B ₃	Y
1	1	1	1	1	Y ₁
1	-1	0	-1	0	Y ₂
1	0	-1	0	-1	Y ₃
-1	1	1	-1	-1	Y ₄
-1	-1	0	1	0	Y ₅
-1	0	-1	0	1	Y ₆

Tablo 4. 2x4 faktöriyel tasarım için dizayn matrisi
(Design matrix for 2x4 factorial design)

A	B ₁ B ₄	B ₂ B ₄	B ₃ B ₄	AxB ₁ B ₄	AxB ₂ B ₄	AxB ₃ B ₄	Y
1	1	0	0	1	0	0	Y1
1	0	1	0	0	1	0	Y2
1	0	0	1	0	0	1	Y3
1	-1	-1	-1	-1	-1	-1	Y4
-1	1	0	0	-1	0	0	Y5
-1	0	1	0	0	-1	0	Y6
-1	0	0	1	0	0	-1	Y7
-1	-1	-1	-1	1	1	1	Y8

2.2. Çoklu Karşılaştırmalar (Multiple Comparisons)

Varyans analizinde kontrol hipotezi reddedildiğinde, hangi gruplar arasında farklılık olduğunu belirlemek için çoklu karşılaştırma testleri kullanılmaktadır. Literatürdeki çoklu karşılaştırma ve çoklu aralık testlerinden sık karşılaşılan ve farklı koşullara uygun olan testler LSD, Tukey, Bonferoni, Hochberg's GT2, Gabriel, Sidak, Dunnett t ve Scheffe testleridir. LSD tüm grupların ikili olarak eşleştirilip tek bir kritik değer ile karşılaştırıldığı bir testtir. Karşılaşacak ortalama sayısı arttıkça deneme başına hata miktarı da artmaktadır. LSD gibi tek bir kritik değer kullanan Tukey testinin güven aralığı LSD testine göre daha dardır ve gruplardaki denek sayıları eşit olmadığında Tukey-Kramer testi kullanılabilir. Bonferroni testi, t istatistiği üzerine kurulu olup eşit örneklem sayısı gerektirmeyen ve F testi önemli olmasa da uygulanabilen bir testtir. Hochberg's GT2, Tukey'e benzeyen ancak genişletilmiş t modülü (student maximum modulus) tabanında çalışan bir testtir ve Tukey kadar güçlü olmadığı kabul edilmektedir. Hochberg's GT2 gibi genişletilmiş t modülü tabanında çalışan Gabriel testi de gruplardaki örneklem sayısının eşit olmasını gerektirir. Sidak testi, özellikle LSD'ye nazaran hata miktarlarına karşı daha sıkı sınırlamalar getirebilmektedir. Dunnett t, kontrol grubunun deney gruplarıyla karşılaştırıldığı durumlarda önerilmektedir ve tek bir kritik değer kullanılmaktadır. Scheffe gruplardaki denek sayıları farklı olduğunda da uygulanabilen bir testtir. Farklı tekrür durumlarında farklı test istatistiği kullanır [1,3,6,7,11-13].

Faktöriyel tasarımlarda, faktör seviyelerinin her bir kombinasyonu için denek sayısının eşit olup olmasına göre adlandırılan dengeli tasarımlar ve dengesiz tasarımlar çoklu karşılaştırma için kullanılacak test seçiminde de önem arz etmektedir. Yapılacak varyans analizi sonucunda etkileşim etkisi ve/veya esas etkilerin anlamlı bulunması durumunda, çoklu karşılaştırma ile anlamlı farklılık gösteren alt grupların belirlenmesinde, dengeli tasarımlar söz konusu olduğunda, yani faktör kombinasyonlarının oluşturduğu alt gruplardaki denek sayıları eşit olduğunda Tukey HSD, dengesiz tasarımlar söz konusu olduğunda, yani faktör kombinasyonlarının oluşturduğu alt gruplardaki denek sayıları farklı

olduğunda ise Tukey-Kramer testinin kullanılması önerilmektedir [3,6,7,11-14].

Tukey HSD test istatistiği aşağıda verilen Eşitlik (14), Tukey-Kramer test istatistiği ise Eşitlik (15) ile hesaplanabilmektedir [3,6,7,11,13,14].

$$W = q_{\alpha,k,v} \sqrt{\frac{HKO}{n}} \quad (14)$$

$$W = q_{\alpha,k,v} \sqrt{\frac{HKO}{2} \left(\frac{1}{n_A} + \frac{1}{n_B} \right)} \quad (15)$$

Bu eşitliklerde, α :önemlilik düzeyi, k :grup sayısı ve v :serbestlik derecesi olmak üzere, $q_{\alpha,k,v}$ ifadesi Tukey tablo değerini, HKO hata kareler ortalamasını ve A ve B karşılaştırılacak gruplar olmak üzere n ise her bir grup için denek sayısını ifade etmektedir. Hesaplanan test istatistiği, iki grup ortalaması arasındaki fark ile karşılaştırılarak, ortalamalar arası farklılık hesaplanan kritik değerden büyük ya da eşit olduğunda ($d \geq W$) ilgili önemlilik düzeyinde iki ortalama arasında anlamlı bir farklılık olduğuna karar verilmektedir [3,6,7,11,13,14].

2.3. Programlama Dilleri (Programming Languages)

Bu çalışmada geliştirilen web sayfası, PHP (Hypertext Preprocessor = Zengin Metin Ön işlecisi) ve JavaScript kodlarıyla oluşturulmuştur.

PHP, çok sayıda platform ile uyumlu olarak çalışabilen, veri tabanlarıyla güvenli iletişim kurabilen ve sunucu tarafı çalışan açık kaynak bir programlama dilidir. Önerilen yapı Linux üzerinde, Apache Server ve MySQL (My Structured Query Language = Yapısal Sorgulama Dili) veri tabanı ile kullanılmasıdır [15-18]. PHP kodları, HTML (Hypertext Markup Language = Zengin Metin İşaretleme Dili) sayfalarının içine gömülü olarak kullanılabilirdiği gibi, tamamen bağımsız olarak tek başına da kullanılabilir ve web tabanlı olarak çalıştırılabilir. Sunucu tarafı çalıştığından kodlar

web sayfası üzerinde görünmemekte, kullanıcı ekranına sonuçlar yansıtılmaktadır [15,16,18,19].

JavaScript, yine farklı platformlarda çalışabilen, sunucu ve istemci uygulamaları için geliştirilmiş, istemci tarafı çalışan, nesne tabanlı küçük ve sade bir programlama dilidir. Sadece JavaScript desteği veren web tarayıcıları ile çalışmaktadır. HTML içine ve uygulamalara kolayca eklenebilen Javascript yardımı ile HTML kaynak kodları değiştirilebilmekte, böylece dinamik web sayfaları hazırlanabilmekte ve sunucudan cevap beklemeksizin anlık işlem yapılabilmektedir. İstemci tarafı çalıştığından kodlar web sayfası üzerinde görünmekte ve kullanıcı ekranına kodlar yansıtılmaktadır [19-23].

2.4. Uygulama (Practice)

Bu çalışmada, iki yönlü faktöriyel varyans analizi ve ardından hem etkileşim hem de esas etkiler için çoklu karşılaştırmaların yapılabildiği web tabanlı bir uygulama geliştirilmiştir. Uygulamada kullanıcıdan faktör seviye sayıları ve faktör kombinasyonları için en büyük denek sayısı istenmekte, buna göre oluşturulan veri matrisinde veri girişlerinin yapılmasının ardından varyans analizi tablosu elde edilmektedir. Varyans analizinde, etkileşim etkisinin istatistiksel olarak önemli olması durumunda etkileşime ait çoklu karşılaştırmalar ve/veya esas etkilerin istatistiksel olarak anlamlı olması durumunda ise esas etkilere ait çoklu karşılaştırmalar otomatik olarak yapılmaktadır. Uygulamanın algoritması şu şekildedir:

1. Kullanıcı tarafından girilen faktör seviye sayıları (sevA, sevB) ve kombinasyonlardaki en büyük n değeri değişkenlere aktarılır
2. $sevA \times sevB$ boyutlu çapraz tablo ve bu tablonun her hücresinde n adet veri giriş alanı oluşturulur
3. Kullanıcıya veri girişi ekranı yansıtılır
4. Girilen veriler dizi değişkenlerine aktarılır (örneğin 2×3 boyutlu tasarım için 6 dizi)
5. Diziler için sayısal fonksiyonlar tanımlanır
6. Matris işlemleri için fonksiyonlar tanımlanır
7. Dizinin n sayısına göre dizayn matrisi oluşturulur
8. HKT hesaplamak için matrisleri ayrıştırılır
 - a. A ve B etkileri için X_{AB} matrisi
 - b. A ve $A \times B$ etkileri için $X_{A \times B}$ matrisi
 - c. B ve $A \times B$ etkileri için $X_{B \times A}$ matrisi
 - d. A, B, $A \times B$ etkileri için $X_{A \times B \times A}$ matrisi
 - e. Veriler için Y matrisi
9. Eşitlik 9'da tanımlanan matris işlemleri yapılır
10. Eşitlik 10-13'te tanımlanan HKT hesaplanır
11. Serbestlik dereceleri, kareler ortalamaları ve F istatistikleri hesaplanır
12. Tanımlayıcı istatistikler hesaplanır
13. Çoklu karşılaştırma testi gerekip gerekmediği belirlenir, gerekmiyorsa işlem sonlandırılır
14. Gerekiyorsa, etkileşim veya esas etkilerin hangisi için çoklu karşılaştırma yapılacağı ve tasarımın dengeli ya da dengesiz oluşuna göre hangi çoklu karşılaştırma testinin kullanılacağı belirlenir

15. Seçilen çoklu karşılaştırma testine bağlı olarak Eşitlik 14 veya 15'teki W değeri hesaplanır
16. Kullanıcı ekranına sonuçlar yansıtılır

Ayrıca ham veriler mevcut olmadığında özet istatistikler (ortalama, standart sapma, n) ile iki yönlü faktöriyel varyans analizinin yapılabildiği bir uygulama daha geliştirilmiştir. Uygulamanın algoritması şu şekildedir:

1. Kullanıcı tarafından girilen faktör seviye sayıları (sevA, sevB) değişkenlere aktarılır
2. $sevA \times sevB$ boyutlu çapraz tablo ve bu tablonun her hücresinde 3 adet veri giriş alanı oluşturulur
3. Kullanıcıya veri girişi ekranı yansıtılır
4. Girilen verilerden yararlanarak Eşitlik 1-5'te tanımlanan HKT hesaplanır
5. Serbestlik dereceleri, kareler ortalamaları ve F istatistikleri hesaplanır
6. Tanımlayıcı istatistikler hesaplanır
7. Çoklu karşılaştırma testi gerekip gerekmediği belirlenir, gerekmiyorsa işlem sonlandırılır
8. Gerekiyorsa, etkileşim veya esas etkilerin hangisi için çoklu karşılaştırma yapılacağı belirlenir
9. Seçilen çoklu karşılaştırma testine bağlı olarak Eşitlik 14 veya 15'teki W istatistikleri hesaplanır
10. Kullanıcı ekranına sonuçlar yansıtılır

Ham veriler girilerek yapılan analizlerde Tip III kareler toplamı hesaplanmaktadır, bu nedenle dengeli ve dengesiz tüm faktöriyel tasarımlarda kullanılabilir. Ancak, özet istatistiklerden dizayn matrisi oluşturulamayacağından sadece Tip I kareler toplamı hesaplanabilmektedir ve bu nedenle özet istatistiklere dayalı geliştirilen uygulama sadece dengeli faktöriyel tasarımlarda kullanılabilir.

3. BULGULAR (RESULTS)

İki faktörlü ve faktör seviyelerinin iki ve üç olduğu bir örnek uygulama ve ekran görüntüleri aşağıda verilmiştir. Faktör seviye sayılarının ve faktör kombinasyonlarındaki en büyük denek sayısının belirtilip "Veri Matrisi Oluştur" butonuna (Şekil 1) basıldıktan sonra belirtilen özelliklerde veri matrisi oluşturulmaktadır.

Faktör 1 Seviye Sayısı	2 ▼
Faktör 2 Seviye Sayısı	3 ▼
n (max)*	6
Veri Matrisi Oluştur	

Şekil 1. Veri matrisi oluşturma (Creating data matrix)

Özellikle dengesiz tasarımlar söz konusu olduğunda her bir faktör kombinasyonu için girilecek veri sayısını tek tek belirtmek hem pratik olmayacağı hem de karışıklığa sebep olacağı için kullanıcıdan sadece bir n değeri girmesi

istenmekte ve bu durum açıklama olarak “Alt gruplardaki (faktör kombinasyonlarındaki) n sayıları eşit değil ise en büyük n sayısını, eşit ise sadece birinin n sayısını giriniz” şeklinde verilmektedir. Örneğin bu çalışmada, faktör seviyelerine (2x3) göre toplam altı kombinasyon olup bu kombinasyonlardaki n sayıları 4, 5 veya 6 şeklinde olduğu için n (max) değerine 6 girilmesi gerekmekte ve tüm kombinasyonlar için 6 veri giriş alanı oluşturulmaktadır. Denek sayısının 6’dan az olduğu kombinasyonlarda, boş kalan veri giriş alanları için gerekli kontroller yapılarak her bir kombinasyonun gerçek n değeri saptanmaktadır.

	B1	B2	B3
A1	1 79	1 58	1 78
	2 73	2 76	2 74
	3 72	3 72	3 80
	4 77	4 66	4 97
	5 97	5 79	5 81
	6 81	6	6
A2	1 73	1 87	1 88
	2 65	2 98	2 75
	3 89	3 86	3 77
	4 71	4 82	4 80
	5	5 81	5
	6	6	6
Hesapla			

Şekil 2. Veri girişi (Data entry)

Oluşturulan veri matrisinde doğrudan klavye aracılığıyla veri girişi yapıldıktan sonra “Hesapla” butonuna (Şekil 2) basılmasıyla kullanıcı işlemleri tamamlanmış olmakta, bu iki aşama dışında kullanıcının herhangi bir işlem ya da seçim yapması gerekmemektedir. Tüm analiz işlemleri otomatik olarak tamamlanmakta ve kullanıcı ekranına doğrudan analiz sonuçları (Şekil 3) yansıtılmaktadır.

Analiz sonuçları tek ekranda, tablolar alt alta düzenlenmiş olarak, kullanıcı açısından karmaşıklığı en aza indirmek için sadece gerekli olan sonuçlar şeklinde verilmektedir. Analiz sonuçlarında, ilk olarak faktör seviyelerinde ve tüm kombinasyonlarda özet istatistiklerin gösterildiği bir tanımlayıcı istatistik tablosu (Şekil 3a) yer almaktadır. Ardından varyans analizi tablosu (Şekil 3b) ile faktörlerin esas etkilerinin ve etkileşim etkisinin önemlilik durumları belirlenmektedir. İstatistiksel önemlilik gösteren etkiye ait p değerleri kullanıcının dikkatini çekmesi açısından kırmızı renkle vurgulanmaktadır. Yine kullanıcı açısından hem karmaşıklığı en aza indirmek hem de yorumlamayı kolaylaştırmak amacıyla varyans analizinin sonuçlarına ilişkin bir cümle ile yorum (Şekil 3c) yapıldıktan sonra, gerekliyse anlamlı bulunan etkilere ait çoklu karşılaştırma tablosu (Şekil 3d) otomatik olarak oluşturulmaktadır. Etkileşime ait çoklu karşılaştırma tablosu oluşturulması durumunda, tüm ikili karşılaştırmalara ilişkin p değerleri hesaplanmakta, ancak sadece yorumlamalarda kullanması gereken p değerleri ekrana yansıtılmaktadır. Örneğin bu çalışmada, a1b1/a2b2 veya a1b1/a2b3 karşılaştırmalarına

ilişkin p değerleri hesaplanabilmekte, ancak iki faktörün birden seviyesi değiştiği için böyle bir karşılaştırma uygun olmayacağından ekrana yansıtılmamaktadır. Çoklu karşılaştırma tablosuna ilişkin olarak kullanıcı, aşağıda verilen beş farklı yorumdan uygun olanı ile bilgilendirilmektedir:

- Çalışmanızda **etkileşim etkisi önemlidir**. Bu nedenle etkileşime ait çoklu karşılaştırma testi yapılmıştır.
- Çalışmanızda etkileşim etkisi önemli değildir. Bu nedenle etkileşime ait çoklu karşılaştırma testine gerek yoktur. Ancak **faktörlerin her ikisinin de esas etkileri önemli** olduğundan faktörlere ait çoklu karşılaştırma testleri yapılmıştır.
- Çalışmanızda etkileşim etkisi önemli değildir. Bu nedenle etkileşime ait çoklu karşılaştırma testine gerek yoktur. Ancak **ikinci faktörün esas etkisi önemli** olduğundan bu faktöre ait çoklu karşılaştırma testi yapılmıştır.
- Çalışmanızda etkileşim etkisi önemli değildir. Bu nedenle etkileşime ait çoklu karşılaştırma testine gerek yoktur. Ancak **birinci faktörün esas etkisi önemli** olduğundan bu faktöre ait çoklu karşılaştırma testi yapılmıştır.
- Çalışmanızda etkileşim etkisi önemli değildir. Bu nedenle etkileşime ait çoklu karşılaştırma testine gerek yoktur. Ayrıca faktörlerin her ikisinin de esas etkileri önemli olmadığından faktörlere ait çoklu karşılaştırma testlerine de gerek yoktur.

TANIMLAYICI İSTATİSTİKLER							
		n	Ortalama	St. Sapma	St. Hata	Min	Max
A1	B1	6	79.833	9.087	3.710	72.000	97.000
	B2	5	70.200	8.379	3.747	58.000	79.000
	B3	5	82.000	8.803	3.937	74.000	97.000
	B	16	77.500	9.675	2.419	58.000	97.000
A2	B1	4	74.500	10.247	5.123	65.000	89.000
	B2	5	86.800	6.760	3.023	81.000	98.000
	B3	4	80.000	5.715	2.858	75.000	88.000
	B	13	80.923	8.836	2.451	65.000	98.000
A	B1	10	77.700	9.405	2.974	65.000	97.000
	B2	10	78.500	11.316	3.578	58.000	98.000
	B3	9	81.111	7.219	2.406	74.000	97.000
	B	29	79.034	9.306	1.728	58.000	98.000

VARYANS ANALİZİ TABLOSU						
VK	KT	SD	KO	F	p	
Faktör A	67.793		1	67.793	0.975	0.334
Faktör B	69.506		2	34.753	0.500	0.613
Etkileşim (AxB)	683.450		2	341.725	4.914	0.017
Hata	1.599.433		23	69.541		
Genel	2.424.966		28			

Çalışmanızda etkileşim etkisi önemlidir.
Bu nedenle etkileşime ait çoklu karşılaştırma testi yapılmıştır.

FAKTÖR A x FAKTÖR B						
	A1B1	A1B2	A1B3	A2B1	A2B2	A2B3
A1B1		0.423	0.998	0.916		
A1B2	0.423		0.260		0.046	
A1B3	0.998	0.260				0.999
A2B1	0.916				0.276	0.934
A2B2		0.046		0.276		0.825
A2B3			0.999	0.934	0.825	

Şekil 3. (a) Tanımlayıcı istatistikler, (b) Varyans analiz tablosu, (c) Yorum ve (d) Çoklu karşılaştırma tablosu
(a) Descriptive statistics, (b) Analysis of variance table, (c) Comment and (d) Multiple comparison table)

İşlemlerin nasıl yapıldığının ve uygulamanın kullanımının gösterilmesi amacıyla, iki faktörlü (2x3) bir uygulama gösterilmiştir. İki faktörlü ve faktörlerin ikişer seviyesi olduğu durumdan, faktörlerin beşer seviyesinin olduğu duruma kadar (2x2, 2x3, 2x4, ..., 4x5, 5x5) tüm alternatif uygulamalar hazırlanmış olup uygulamaların tamamına “www.masungur.com/webstat.php” web sayfasında yer alan “*istatistik analiz ve hesaplama araçları*” yoluyla erişilebilmektedir. Toplam denek sayısı veya her bir faktör kombinasyonundaki denek sayısı bakımından ise herhangi bir kısıtlama bulunmamaktadır.

4. TARTIŞMA (DISCUSSION)

İstatistiksel analizlerde elle hesaplama yapmak zor ve zaman alıcı bir işlemdir. Ayrıca hatalı hesaplama yapma olasılığı da bulunmaktadır. Formüllerin karmaşıklığı ve hesaplama işlemlerinin uzunluğu kişilerin alternatif çözüm yolları aramasına neden olmaktadır. Günümüzde en çok başvurulan yöntem paket programlardır. Ancak herkesin her koşulda paket programlara erişim şansı olmayacağı gibi, bazı istatistiksel analizler her paket programda yer almayabilmektedir.

Bu çalışma kapsamında ele alınan faktöriyel tasarımlar açısından bakıldığında, birçok istatistiksel analiz paket programında faktöriyel tasarımlara ait esas etkiler için çoklu karşılaştırmalar mevcutken etkileşimlere ait çoklu karşılaştırma sonuçları yer almamakta ve dolayısıyla bu durum paket program kullanıcılarını elle hesaplamaya götürerek belki de sonucun hatalı olarak elde edilmesine yol açmaktadır. Faktöriyel tasarımlar bir örnek olmakla birlikte, farklı yöntem ve durumlarda veya paket programa erişim şansı dikkate alındığında birçok açıdan bu sorun söz konusu olabilmektedir.

Hesaplama işlemlerinin kolaylaştırmanın ve hatalı hesaplama ihtimalinden kurtulmanın bir yolu da hesaplama işlemlerini web tabanlı programlama dilleri aracılığıyla yapmaktır. Kod yapılarının karmaşık olması, kodların iyi bilinmesi ve uzmanlık gerektirmesi dezavantaj gibi görünmesine karşın bu kısmın kullanıcıyı ilgilendirmiyor olması, kullanıcının sadece veri girişi yapması gerekmesi, hesaplamaları herkes için oldukça

pratikleştirilmektedir. Yaygın olarak kullanılan programlama dillerinden olan PHP dili de bu amaç için oldukça uygundur. Kod esnekliği, açık kod mantığı ile geliştirilebilir olması ve ücretsiz bir dil olması gibi avantajlara sahip olan PHP dili ile formüller yazılabilmekte, bu formüller fonksiyona çevrilebilmekte ve bir kez oluşturulan fonksiyonlar tekrar tekrar kullanılabilir [15,18]. PHP sunucu taraflı çalıştığından, çıktı alabilmek için girdilerin sunucuya iletilmiş olması gerekmektedir [15,18,19]. Anlık tepkisel ihtiyaçların olması durumunda, bu durum JavaScript desteği ile kolayca çözümlenebilmektedir. İstemci taraflı çalışan ve böylece anlık tepki veren JavaScript sayesinde hazırlanan uygulamalara dinamiklik kazandırılabilen ve kullanıcının anlık olarak uyarılması gereken durumlar çözümlenebilmektedir [19-22].

PHP ve JavaScript ile hazırlanmış olan uygulamanın hazırlık aşamaları oldukça uzun ve zahmetli olmakla birlikte, kullanmaya başladıktan sonra kazandırdığı pratiklik ve zaman kazancı dikkate alındığında bu dezavantaj ortadan kalkmaktadır. Ayrıca uygulamanın internet erişimi olan her yerde ve JavaScript desteği veren tüm tarayıcılarla çalışabiliyor olmasının verdiği erişim kolaylığı da bir diğer avantaj olarak ortaya çıkmaktadır.

Bu çalışmada oluşturulan uygulamaya internet bağlantısı olan ve web tarayıcısı bulunan her türlü bilgisayar, mobil cihaz vb. araçlardan kolaylıkla ulaşılabilmekte ve gerekli olan veri girişi yapıldıktan sonra sonuçlar sadece bir tuş aracılığıyla elde edilebilmektedir. Dolayısıyla, elle uygulaması oldukça karmaşık ve zor olan, ayrıca hesaplama işlemleri için zaman gerektiren bir istatistiksel yöntem, kolaylıkla erişilebilir, uzmanlık alanı olmayan bir kişi tarafından da uygulanabilir ve yorumlanabilir hale gelmektedir.

Günümüzde web tabanlı ya da interaktif istatistiksel analiz ifadeleri oldukça fazla tercih edilmekte ve kullanılmaktadır. Basitçe, en bilinen arama motorları aracılığıyla “interactive statistical analysis”, “web based statistical analysis”, “online statistical calculations” gibi terimlerle yapılacak aramalarda çıkan milyonlarca sonuç ve web sayfası da bunu ortaya koymaktadır. Ancak buradaki en önemli sorunlardan biri, bu web sayfalarının ya da hesaplama araçlarının büyük çoğunluğunun bilimsel içerikli web siteleri olmaması, kişisel olarak oluşturulmuş bilgi paylaşım siteleri, forum veya blog tarzı siteler olmasıdır. Dolayısıyla bu sayfaların hazırlanmasında meslek profesyonellerinin etkin olup olmadıkları, hesaplama araçlarının doğru ve uygun şekilde kullanılıp kullanılmadığı sorunu da ortaya çıkmaktadır. Ayrıca, bu şekildeki web sayfaları ile yapılacak analizlerde, bu sayfaların referans olarak gösterilememesi ya da kabul edilmemesi sorunu da bulunmaktadır.

Sonuç olarak, daha önce ifade edilen erişim kolaylığı, zaman tasarrufu, kullanım kolaylığı gibi birçok nedenden dolayı tercih edilebilirliği yüksek olan web tabanlı

uygulamaların, bilimsel yönden geçerli ve alan uzmanları tarafından hazırlanmış olması da tercih edilirliliği üzerinde oldukça önemli bir etkiye sahip olacaktır. Bu çalışma kapsamında hazırlanan web tabanlı uygulamanın etkinliği, kullanıma ve tercih edilme oranlarının incelenmesi amacıyla gerek veri tabanı üzerinden kullanım istatistikleri incelenerek gerekse bir anket çalışması ile araştırılması planlanmaktadır.

Hangi durumda hangi analizin yapılması gerektiği, veriye uygun yöntemin hangisi olduğu, hangi yöntemin nasıl kullanılacağı ve yorumlanacağı gibi konular istatistik uzmanına başvurulması gereken konulardır. Bu uygulamayı kullanacak olan kişilerin temel düzeyde istatistik bilgisine sahip oldukları varsayılmıştır.

5. SONUÇ (CONCLUSSION)

Bu çalışmada, iki yönlü faktöriyel varyans analizinin ve çoklu karşılaştırmalarının bir web uygulaması aracılığıyla hesaplanması üzerinde durulmuş ve elle hesaplama yerine web tabanlı analiz yapmanın pratikliği gösterilmeye çalışılmıştır. Bu çalışmada hazırlanan web sayfaları ve uygulamaların ışığında, teorik temelleri bilinen ancak henüz yaygın kullanılan paket programlarda yer almayan herhangi bir istatistiksel yöntemin, programlama dilleri sayesinde web tabanlı olarak basit ve anlaşılır bir şekilde yapılabileceği ortaya konmuştur.

Günümüzde, gerek araştırmacılar gerek öğrenciler tarafından kullanılan öğrenme ve uygulama araçlarının, öncelikli olarak her yerden erişim imkanı olan kaynaklar arasından tercih edildiği, hatta mobil araçlarla dahi ulaşılabilen kaynak ve uygulamaların pratiklik ve kolay ulaşım bakımından tercih edilirliliğinin daha yüksek olacağı dikkate alındığında, bu kullanıcılara da ulaşabilmek adına, bu ve benzeri web tabanlı uygulamaların ne kadar önemli olduğu ve bu önemin giderek artacağı açıktır. Bu şekildeki uygulamaların yaygınlaşması, gerek eğitim-öğretim ve gerekse bilimsel araştırmacılar için birçok yönden faydalı olacaktır.

KAYNAKLAR (REFERENCES)

- [1] D.C. Montgomery, "Introduction to Factorial Designs", **Design and Analysis of Experiments**, Wiley, New York, 170-217, 2000.
- [2] G.E.P. Box, J.S. Hunter, W.G. Hunter, "Factorial Designs at Two Levels", **Statistics for Experimenters: Design, Innovation, and Discovery**, Wiley, New Jersey, 173-233, 2005.
- [3] G.W. Oehlert, "Factorial Treatment Structure", **A First Course in Design and Analysis of Experiments**, W.H. Freeman and Company, New York, 165-202, 2010.
- [4] H.J. Seltman, "Two-Way ANOVA", **Experimental Design and Analysis**, Carnegie Mellon University, Pittsburgh, 267-292, 2014.

- [5] L.M. Collins, J.J. Dziak, K.C. Kugler, J.B. Trail, "Factorial Experiments: Efficient Tools for Evaluation of Intervention Components", *American Journal of Preventive Medicine*, 47(4), 498-504, 2014.
- [6] K. Sümbüloğlu, V. Sümbüloğlu, "İki Yönlü Varyans Analizi", **Biyoistatistik**, Ankara: Hatiboğlu Yayınları, Ankara, 87-102, 2012.
- [7] E. Efe, Y. Bek, M. Şahin, "Faktöriyel Düzenlenmiş Denemeler", **SPSS'te Çözümleri ile İstatistik Yöntemler II**, T.C. Kahramanmaraş Sütçü İmam Üniversitesi Rektörlüğü Yayın No:10, Kahramanmaraş, p.87-100, 2000.
- [8] B.H. Cohen, "Calculating a Factorial ANOVA from Means and Standard Deviations", *Understanding Statistics*, 1(3), 191-203, 2002.
- [9] J.A. Landsheer, G. Wittenboer, "Unbalanced 2x2 Factorial Designs and the Interaction Effect: A Troublesome Combination", *Plos One*, 10(3):e0121412, 2015.
- [10] G.W. Snedecor, W.G. Cochran, **Statistical Methods**, Iowa State University Press, Iowa, 1-459, 1989.
- [11] J.A. Rafter, M.L. Abell, J.P. Braselton, "Multiple Comparison Methods for Means", *Society for Industrial and Applied Mathematics*, 44(2), 259-278, 2002.
- [12] A.J. Klockars, G. Sax, "Multiple Comparisons in Factorial Designs", **Multiple Comparisons**, SAGE Publications, Thousand Oaks, CA, 66-77, 1986.
- [13] L.E. Toothaker, "Multiple Comparisons for the Two-Way Anova: Factorial or Randomized Blocks", **Multiple Comparison Procedures**, SAGE Publications, Thousand Oaks, CA, 67-80, 1993.
- [14] A.J. Hayter, "A Proof of the Conjecture that the Tukey-Kramer Multiple Comparisons Procedure is Conservative", *The Annals of Statistics*, 12(1), 61-75, 1984.
- [15] E. Naramore, J. Gerner, Y.L. Scouarnec, J. Stolz, M.K. Glass, **Beginning PHP5, Apache, and MySQL Web Development**, Wiley, Indianapolis, 1-660, 2005.
- [16] K. Yank, "Introducing PHP", **Build Your Own Database Driven Web Site Using PHP & MySQL**, SitePoint Pty.Ltd., Victoria, Australia, 73-113, 2009.
- [17] M. Rochkind, "Platforms and Tools", **Expert PHP and MySQL**, Apress, New York, 45-81, 2013.
- [18] S. Suehring, T. Converse, J. Park, **PHP 6 and MySQL 6 Bible**, Wiley, Indianapolis, 1-793, 2009.
- [19] M.A. Alan, M. Gülmez M, "İnternet Üzerinde Bazı Veri Tabanı Yönetim Dillerinin Karşılaştırılması", *C.Ü. İktisadi ve İdari Bilimler Dergisi*, 2(2), 61-65, 2001.
- [20] S. Stefanov, **JavaScript for PHP Developers**, O'Reilly, Sebastopol, CA, 1-136, 2013.
- [21] N.C. Zakas, **Professional JavaScript for Web Developers**, Wiley, Indianapolis, 1-856, 2012.
- [22] M. Morrison, "The Interactive Web: Reacting to the Virtual World", **Head First JavaScript**, O'Reilly, Sebastopol, CA, 1-31, 2008.
- [23] D.S. McFarland, "Writing Your First JavaScript Program", **JavaScript & jQuery, The Missing Manual**, O'Reilly, Sebastopol, CA, 21-40, 2012.