

Örüntülü Seçenek Hipotezler İçin Bağımsız k -Örneklem Testleri

Independent k -Sample Tests for Patterned Alternatives: Review

İsmet DOĞAN,^a
Nurhan DOĞAN^a

^aBiyostatistik ve Tıbbi Bilişim AD,
Afyon Kocatepe Üniversitesi
Tıp Fakültesi, Afyonkarahisar

Geliş Tarihi/Received: 02.10.2015
Kabul Tarihi/Accepted: 27.10.2015

Yazışma Adresi/Correspondence:
Nurhan DOĞAN
Afyon Kocatepe Üniversitesi
Tıp Fakültesi,
Biyostatistik ve Tıbbi Bilişim AD,
Afyonkarahisar,
TÜRKİYE/TURKEY
nurhandogan@hotmail.com

ÖZET Bir araştırmacı bir grup denemeye ait ortalamaların, örüntü içerikli seçenek olarak sunulan ortalamalar ile eşit olup olmadığı ile ilgilenilebilir. Bu tür oluşumlar biyomedikal araştırmalarda oldukça yaygındır. Örneğin, laboratuvar fareleri üzerinde yapılan bir çalışma ile bir ilacın çeşitli düzeylerdeki dozlarının etkinliği karşılaştırılmak istenebilir. Çok sayıda denemeyi birbiri ile karşılaştıran araştırmacı testten önce her bir denemenin etki şiddetinin büyüklüğüne göre denemelerini sıralamaktadır. Tahmini olarak yapılan bu sıralama çoğunlukla araştırmacının tecrübesine dayalı olarak gerçekleşmektedir. Tek yönlü varyans analizinde veya Kruskal-Wallis testinde, araştırmacı $H_0: \mu_1 = \mu_2 = \dots = \mu_k$ biçimindeki yokluk hipotezindeki sapmaların belirlenmesi ile ilgilenmektedir. Yokluk hipotezindeki sapmalar, deneme etkisi vardır ($H_1: \mu_1 \neq \mu_2 \neq \mu_3 \neq \dots \neq \mu_k$), sıralı seçenekli deneme etkisi vardır, ($H_1: \mu_1 \geq \mu_2 \geq \mu_3 \geq \dots \geq \mu_k$ veya $H_1: \mu_1 \leq \mu_2 \leq \mu_3 \leq \dots \leq \mu_k$), ya da şemsiye seçenekli deneme etkisi vardır ($H_1: \mu_1 \geq \mu_2 \geq \dots \geq \mu_{p-1} \geq \mu_p \leq \mu_{p+1} \leq \dots \leq \mu_k$ veya $H_1: \mu_1 \leq \mu_2 \leq \dots \leq \mu_{p-1} \leq \mu_p \geq \mu_{p+1} \geq \dots \geq \mu_k$) biçiminde farklı seçenek hipotezler ile ifade edilebilmektedir. Bir araştırmada bu seçenek hipotezlerden hangisinin dikkate alınacağı araştırmacının amacına bağlı olarak değişmektedir. Bilinen tek yönlü varyans analizi veya Kruskal-Wallis testi sıralı seçenekli veya şemsiye seçenekli hipotezleri test etmek için uygun, iyi bir test değildirdir. Literatürde sıralı seçenekli ve şemsiye seçenekli hipotezlerin söz konusu olduğu durumlar için geliştirilmiş çok sayıda yöntem bulunmaktadır. Bu çalışmada, örüntü içeren seçenek hipotezlerin değerlendirilmesinde kullanılan on beş farklı yöntem tanıtılmış ve literatürden elde edilen bilgiler kullanılarak bu yöntemler birbirleri ile karşılaştırılmıştır. Literatürden elde edilen sonuçlar genel olarak dikkate alındığında sıralı seçenек hipotezin söz konusu olduğu durumlar için Jonckheere-Terpstra testinin, şemsiye seçenekli hipotezin söz konusu olduğu durumlarda ise Hettmansperger-Norton testinin tercih edilmesi gerektiği söylenebilir.

Anahtar Kelimeler: Bağımsız k -örneklem testleri; doz-cevap çalışmaları;
sıralı etkili seçenek hipotez; şemsiye etkili seçenek hipotez

ABSTRACT A researcher may be interested in knowing whether a set of treatment means are equal versus they following some pattern of ordering. Occurrences like these are common in fields like biomedical research. For example, a study may involve comparing the effectiveness of various levels of a drug dosage on laboratory mice. The researcher comparing multiple treatments in a one way setting is often able to rank the treatments according to the order of magnitude of the effect of each treatment prior to testing. The guess of this order may be based on the experience of the researcher. Covered by one-way ANOVA or Kruskal-Wallis test, the researcher is often concerned with detecting deviations from the null hypothesis of no treatment effect ($H_0: \mu_1 = \mu_2 = \dots = \mu_k$). Particular deviations of interest have included the omnibus alternative ($H_1: \mu_1 \neq \mu_2 \neq \mu_3 \neq \dots \neq \mu_k$), the ordered alternative ($H_1: \mu_1 \geq \mu_2 \geq \mu_3 \geq \dots \geq \mu_k$ or $H_1: \mu_1 \leq \mu_2 \leq \mu_3 \leq \dots \leq \mu_k$) and the umbrella alternative ($H_1: \mu_1 \geq \mu_2 \geq \dots \geq \mu_{p-1} \geq \mu_p \leq \mu_{p+1} \leq \dots \leq \mu_k$ or $H_1: \mu_1 \leq \mu_2 \leq \dots \leq \mu_{p-1} \leq \mu_p \geq \mu_{p+1} \geq \dots \geq \mu_k$). The conventional one way analysis of variance or Kruskal-Wallis test is not a good test for testing the ordered and umbrella alternatives. In this study, fifteen different testing methods improved for patterned alternatives have been introduced and these methods were compared with each other by using the information obtained from the literature. Considering the overall results obtained from the literature, it is said that for ordered alternative the Jonckheere-Terpstra test, for umbrella alternative the Hettmansperger-Norton test should be preferred.

Key Words: Independent k -sample tests; dose-response studies;
ordered alternative; umbrella alternative

doi: 10.5336/biostatic.2015-48169

Copyright © 2016 by Türkiye Klinikleri

Türkiye Klinikleri J Biostat 2016;8(1):94-102

Bilimsel araştırmalar içerisinde doz-yanıt denemeleri oldukça yaygındır. Özellikle tıbbi çalışmalarda çeşitli tedavilerin bir kontrol veya plasebo grubu ile karşılaştırılması sıklıkla karşılaşılan bir durumdur. Bu tür çalışmalarda temel amaç, test edilen ilacın istenen tıbbi bir etkiye sahip olup olmadığını belirlemek ve etkili en düşük dozunu tespit etmektir.¹ Çalışmalarda genellikle güçlü ya da en azından eşit deneme etkileri üretebilmek için doz seviyelerini artırmak gerekmektedir. Doz-yanıt ilişkisinin karakteri genellikle bilinmediğinden, farklı doz seviyelerinin karşılaştırılması ilgi konusudur. Dolayısıyla araştırmacılar, belirlenen farklı doz seviyelerinin çoklu karşılaştırma problemi ile karşı karşıya kalmaktadır. Özellikle, doz seviyesi artırıldığında bu artışın ilave bir etki oluşturup oluşturmayacağına karar verebilmek için dikkate alınan bir doz seviyesi ile bu doz seviyesine yakın bir doz seviyesinin karşılaştırılması ile ilgilenilmektedir.²

Bir maddenin biyolojik aktivitesi (zehirlenme vs.) o maddenin monoton artan veya azalan doz seviyelerinin kullanıldığı bir dizi deneyler ile araştırılabilir. Araştırmacılar, ilaçlar, gıda katkı maddeleri, suni gıdalar ve böcek ilaçları gibi maddelerin istenen bir etkiye sahip olup olmadığını, eğer istenilen etki söz konusu ise hangi doz miktarının buna sebep olduğunu bilmek istemektedirler. Bu tür araştırmalar genellikle hayvan deneyleri ile gerçekleştirilmektedir.³

Doz seviyesinin azalması veya artması ile birlikte etkisinin de monoton azaldığı veya monoton arttığı düşünülen çalışmalarda seçenek hipotez $H_1: \mu_1 \geq \mu_2 \geq \mu_3 \geq \dots \geq \mu_k$ veya $H_1: \mu_1 \leq \mu_2 \leq \mu_3 \leq \dots \leq \mu_k$ biçimlerinde oluşturulabilir. Bu durum literatürde sıralı seçenek (ordered alternative) olarak ifade edilmektedir.

Doz seviyesinin belirli bir değerine kadar monoton olarak azalması ya da artmasına bağlı olarak etkisi de azalıyor ya da artıyorsa bu durumda seçenek hipotez $H_1: \mu_1 \geq \mu_2 \geq \dots \geq \mu_{p-1} \geq \mu_p \leq \mu_{p+1} \leq \dots \leq \mu_k$ veya $H_1: \mu_1 \leq \mu_2 \leq \dots \leq \mu_{p-1} \leq \mu_p \geq \mu_{p+1} \geq \dots \geq \mu_k$ biçimlerinde oluşturulabilir. Bu durum ise literatürde şemsiye seçenek (umbrella alternative) olarak

ifade edilmektedir. Şemsiye seçenekli hipotezlerde yer alan $\mu_{p-1} \geq \mu_p \leq \mu_{p+1}$ ile $\mu_{p-1} \leq \mu_p \geq \mu_{p+1}$ ifadeleri şemsiye etkisi (umbrella effect), μ_p ise şemsiye noktası (umbrella point = peak) olarak ifade edilmektedir. Çalışmada hangi seçenek hipotezin kullanılacağı ise çoğunlukla araştırmacının tecrübesine dayalı olarak belirlenmektedir.

Bu çalışmanın amacı, özellikle doz-cevap çalışmalarında kullanılan örüntü içerikli seçenek hipotezlerin değerlendirilmesi için geliştirilmiş bağımsız k -örneklem testlerini tanıtmak ve literatürden elde edilen bilgilerden yararlanarak bu testleri birbirleri ile karşılaştırmaktır.

JONCKHEERE-TERPSTRA TESTİ^{4,5}

Jonckheere ve Terpstra tarafından önerilen yöntem göre test istatistiğinin hesaplanmasında kullanılan J değeri,

$$J = \sum_{i=1}^{k-1} \sum_{j=i+1}^k U_{ij} \quad \text{tüm } i < j,$$

U_{ij} : i 'inci gruptaki gözlem değerlerinden büyük j 'nci gruptaki gözlem sayısı,

formülü kullanılarak hesaplanmaktadır. Yöntemin test ettiği hipotezler;

$$H_0: \mu_1 = \mu_2 = \mu_3 = \dots = \mu_k$$

$$H_1: \mu_1 \leq \mu_2 \leq \mu_3 \leq \dots \leq \mu_k$$

şeklinde. Test istatistiğinin değeri;

$$z = \frac{J - \frac{N^2 - \sum_{i=1}^k n_i^2}{4}}{\sqrt{\frac{N^2(2N+3) - \sum_{i=1}^k [n_i^2(2n_i+3)]}{72}}}$$

k : Toplam grup sayısı,

n_i : i 'nci gruptaki gözlem sayısı,

N : Toplam gözlem sayısı,

formülü kullanılarak hesaplanmaktadır. Elde edilen z test istatistiğinin değeri, standart normal dağılım tablo değeri olan ∓ 1.96 değeri ile karşılaştırılır. $z \geq \mp 1.96$ ise H_0 hipotezi reddedilir.

UYARLANMIŞ JONCKHEERE-TERPSTRA TESTİ⁶

Uyarlanmış Jonckheere-Terpstra testinin Jonckheere-Terpstra testinden farkı J değerinin

hesaplanmasındaki farklılıktan kaynaklanmaktadır. Uyarlanmış Jonckheere-Terpstra testinde J değeri,

$$J = \sum_{i=1}^{k-1} \sum_{j=i+1}^k (j-i)U_{ij} \quad \text{tüm } i < j,$$

biçiminde hesaplanmaktadır. Bu farklılığın dışında Jonckheere-Terpstra Testi için önerilen kural-
ların tamamı burada da geçerlidir.

BARTHOLOMEW TESTİ^{7,8}

Yöntemin test ettiği hipotezler;

$$H_0: \mu_1 = \mu_2 = \mu_3 = \dots = \mu_k$$

$$H_1: \mu_1 \geq \mu_2 \geq \mu_3 \geq \dots \geq \mu_k$$

şeklinindedir. Bartholomew testinde test istatistiğinin değeri,

$$\bar{x} = \frac{\sum_{i=1}^k n_i \bar{x}_i}{\sum_{i=1}^k n_i}$$

$$\chi^2 = \frac{\sum_{i=1}^k n_i (\bar{x}_i - \bar{x})^2}{s_e^2}$$

\bar{x}_i : i 'inci gruba ait aritmetik ortalama,

n_i : i 'nci gruptaki gözlem sayısı,

s_e^2 : Varyans Analizi sonucunda elde edilen hata kareler ortalaması,

biçiminde hesaplanmaktadır. Bu yöntemde grup ortalamaları küçükten büyüğe doğru sıralanır. Dikkate alınan grupta, elde edilen gözlem değerlerine ait ortalama değerler arasında büyüklük sıralaması bozuluyorsa bu durumda gruplar birleştirilerek büyüklük sıralaması sağlanır. Test istatistiğinin hesaplanması işlemi grup birleştirmeleri yapıldıktan sonra oluşan grup ve ortalamalar dikkate alınarak gerçekleştirilir.

Elde edilen χ^2 değeri, Ki-kare testi için geliştirilen $k-1$ serbestlik dereceli α yanılma düzeyindeki Ki-kare tablo değeri ile karşılaştırılır. $\chi^2 \geq \chi_{k-1, \alpha}^2$ ise H_0 hipotezi reddedilir.

CHACKO TESTİ⁹⁻¹¹

Gruplardaki gözlem sayılarının eşit olduğu durumlar için Kruskal-Wallis testine benzeyen test Chacko (1963) tarafından geliştirilmiştir. Test istatistiğinin değeri aşağıda verilen eşitlik kullanılarak hesaplanmaktadır.

$$\bar{H}_k = \frac{12n}{N(N+1)} \sum_{j=1}^k t_j \left[\bar{R}_{|t_j|} - \frac{N+1}{2} \right]^2$$

n : Her bir gruptaki gözlem sayısı,

$\bar{R}_{|t_j|}$: Her bir gruba ait sıra sayılarının aritmetik ortalaması,

$$N = kn$$

Formülde yer alan t_j değeri $\bar{X}_1 \leq \bar{X}_2 \leq \dots \leq \bar{X}_k$ olacak şekilde birleştirme yapılan grup sayısını göstermektedir. Örneğin $\bar{X}_1 = 10, \bar{X}_2 = 16, \bar{X}_3 = 14, \bar{X}_4 = 12, \bar{X}_5 = 18$ olsun. Bu durumda $\bar{X}_1 \leq \bar{X}_2 \leq \bar{X}_3 \leq \bar{X}_4 \leq \bar{X}_5$ koşulu sağlanmamaktadır. Dolayısıyla koşulun sağlanabilmesi için bazı grupların birleştirilmesi gerekmektedir. Buna göre $\bar{X}_1 = 10, \bar{X}_2 = \frac{16+14+12}{3} = 14$ ve $\bar{X}_3 = 18$ olur. Yapılan bu düzenlemeye göre t_i değerleri sırasıyla $t_1 = 1, t_2 = 3$ ve $t_3 = 1$ olarak belirlenir. Yöntemin test ettiği hipotezler;

$$H_0: \mu_1 = \mu_2 = \mu_3 = \dots = \mu_k$$

$$H_1: \mu_1 \leq \mu_2 \leq \mu_3 \leq \dots \leq \mu_k$$

şeklinindedir. Formülden elde edilen değer serbestlik derecesi k olan α yanılma düzeyindeki Chacko test istatistiğine ait tablo değeri ($\bar{H}_{k, \alpha}$) ile karşılaştırılır. $\bar{H}_k \geq \bar{H}_{k, \alpha}$ ise H_0 hipotezi reddedilir.

SHORACK TESTİ¹²

Gruplardaki gözlem sayılarının eşit olduğu durumlar için geliştirilen ve Chacko testine oldukça benzeyen Shorack testinde, test istatistiğinin değeri;

$$\bar{\chi}_{rank}^2 = \frac{12n_i}{k(k+1)} \sum_{j=1}^k t_j \left[\bar{R}_{|t_j|} - \frac{k+1}{2} \right]^2$$

k : grup sayısı,

n_i : i 'inci gruptaki gözlem sayısı,

eşitliği kullanılarak hesaplanmaktadır. Friedman Testi'nin modifiye edilmesi ile elde edilmiş olan yöntem Chacko Testi'ne olan benzerliğinden dolayı literatürde Chacko-Shorack Testi olarak da bilinmektedir. Formülde yer alan t_j ve $\bar{R}_{|t_j|}$ değerleri Chacko Testi'nde olduğu gibi hesaplanmaktadır. Yöntemin test ettiği hipotezler;

$$H_0: \mu_1 = \mu_2 = \mu_3 = \dots = \mu_k$$

$$H_1: \mu_1 \leq \mu_2 \leq \mu_3 \leq \dots \leq \mu_k$$

şeklindedir. Formülden elde edilen değer serbestlik derecesi k olan α yanılma düzeyindeki Ki-kare test istatistiğine ait tablo değeri ($\chi_{k,\alpha}^2$) ile karşılaştırılır. $\bar{\chi}_r^2 \geq \chi_{k,\alpha}^2$ ise H_0 hipotezi reddedilir.

WILLIAMS TESTİ I¹³

Student t-Testi'nin, grup ortalamalarının küçükten büyüğe doğru sıralandığı uyarlanmış durumudur. Bu yöntemde grup ortalamaları $\bar{X}_0 \leq \bar{X}_1 \leq \bar{X}_2 \leq \dots \leq \bar{X}_k$ olacak şekilde küçükten büyüğe doğru sıralanır. Artan doz seviyelerinin dikkate alındığı gruplarda, elde edilen gözlem değerlerine ait ortalama değerler arasında sıralama bozuluyorsa, gruplar birleştirilerek sıralama sağlanır. Test istatistiğinin hesaplanması işlemi grup birleştirmeleri yapıldıktan sonra oluşan grup ve ortalamalar dikkate alınarak gerçekleştirilir. Karşılaştırmalara kontrol grubu ile ortalaması en büyük olan grubun karşılaştırılması ile başlanır. H_0 hipotezi kabul edilinceye kadar her defasında bir önceki doz seviyesine en yakın doz seviyesi dikkate alınarak kontrol grubu ile diğer tüm gruplar arasında ikili karşılaştırmalar yapılır. Yöntemin test ettiği hipotezler;

$$H_0: \mu_0 = \mu_i$$

$$H_1: \mu_0 \neq \mu_i$$

şeklindedir. Gruplardaki gözlem sayılarının (n) eşit olması durumunda test istatistiğinin değeri;

$$N = kn$$

$$\bar{t}_i = \frac{(\bar{X}_i - \bar{X}_0)}{\sqrt{2s^2/n}}$$

\bar{X}_i : i 'nci grup ortalaması,

X_0 : Kontrol grubunun ortalaması,

$$s^2 = \frac{\sum_i \sum_j (y_{ij} - \bar{y}_i)^2}{N - k - 1}$$

n_i : i 'nci gruba ait gözlem sayısı,

ile elde edilir. Formülden elde edilen değer Williams (1971) tarafından geliştirilen $\bar{t}_{(n-k),k,\alpha}$ tablo değeri ile karşılaştırılır. $\bar{t}_i \geq \bar{t}_{(n-k),k,\alpha}$ koşulunu sağlayan en düşük doz seviyesi LC_{50} değeri olarak kabul edilir.

Doz gruplardaki gözlem sayılarının eşit ancak kontrol grubundaki gözlem sayısının büyük olması durumunda ise test istatistiğinin değeri;

$$\bar{t}_i = \frac{(\bar{X}_i - \bar{X}_0)}{\sqrt{s^2/n_i + s^2/n_0}}$$

\bar{X}_i : i 'nci grup ortalaması,

X_0 : Kontrol grubunun ortalaması,

s^2 : Hata Kareler Ortalaması,

n_i : i 'nci grubun gözlem sayısı,

n_0 : Kontrol grubuna ait gözlem sayısı,

formülü kullanılarak hesaplanır.

WILLIAMS TESTİ II¹⁴

Williams tarafından önerilen ikinci yöntemde göre test istatistiğinin değeri,

$$\bar{M}_i = \max_{1 \leq u \leq i} \min_{i \leq v \leq k} \sum_{i=u}^v \frac{\bar{X}_i}{(v - u + 1)}$$

$$\bar{t}_i = \frac{(\bar{M}_i - \bar{X}_0)}{\sqrt{2s^2/n}}$$

formülleri kullanılarak hesaplanmaktadır. Bu farklılığın dışında Williams (1971) tarafından önerilen kuralların tamamı burada da geçerlidir.

SHIRLEY TESTİ^{3,15}

Shirley tarafından önerilen yöntemde göre test istatistiğinin değeri gruplardaki gözlem sayıları eşit ise,

$$t_i = \frac{\max_{1 \leq u \leq k} \sum_{j=u}^k \bar{R}_j / (k - u + 1) - \bar{R}_0}{\sqrt{\frac{(k+1)(N+1)}{6}}}$$

kontrol grubu dışındaki grupların gözlem sayıları birbirine eşit ancak kontrol grubundaki gözlem sayısı diğer gruplardaki gözlem sayısına eşit değilse;

$$t_k = \frac{\max_{1 \leq u \leq k} (\sum_{j=u}^k n_j \bar{R}_j / \sum_{j=u}^k n_j) - \bar{R}_0}{\sqrt{\frac{N(N+1)}{12} \left(\frac{1}{n_k} + \frac{1}{C} \right)}}$$

\bar{R}_j : j 'nci gruba ait sıra sayılar m aritmetik ortalaması,

n_j : j 'nci gruba ait gözlem sayısı,

\bar{R}_0 : Kontrol grubuna ait sıra sayılar m

aritmetik ortalaması,

k : Doz sayısı,

C : Kontrol grubuna ait gözlem sayısı,

N : Toplam gözlem sayısı,

formülleri kullanılarak hesaplanmaktadır. Eğer gruplarda tekrar eden gözlem değerleri (tie) varsa tekrar eden her bir gözlem değerine ait tekrar sayıları (l_i) kullanılarak, $\frac{N(N+1)}{12}$ yerine $\frac{N(N+1)}{12} - \sum_s \frac{(l_i^3 - l_i)}{12(N-1)}$ kullanılmalıdır. Karşılaştırmalara kontrol grubu ile ortalaması en büyük olan grubun da dahil edildiği tüm grupların dikkate alındığı karşılaştırma ile başlanır. H_0 hipotezi kabul edilinceye kadar her defasında bir önceki doz seviyesine en yakın doz seviyesi dikkate alınarak kontrol grubu ile diğer tüm gruplar arasında karşılaştırmalar yapılır. Yöntemin test ettiği hipotezler;

$$H_0: \mu_1 = \mu_2 = \mu_3 = \dots = \mu_k$$

$$H_1: \mu_1 \leq \mu_2 \leq \mu_3 \leq \dots \leq \mu_k$$

şekindedir. Formülden elde edilen değer Williams (1971) tarafından geliştirilen $\bar{t}_{(n-k),k,\alpha}$ tablo değeri ile karşılaştırılır. $t_i \geq \bar{t}_{(n-k),k,\alpha}$ koşulunu sağlayan en düşük doz seviyesi LC_{50} değeri olarak kabul edilir.

LE TESTİ^{16,17}

Kruskal-Wallis testine benzeyen ve monoton artan sıralı seçeneklerin söz konusu olduğu durumlar için Le (1988) tarafından geliştirilmiştir. Yöntemin test ettiği hipotezler;

$$H_0: \mu_1 = \mu_2 = \mu_3 = \dots = \mu_k$$

$$H_1: \mu_1 \leq \mu_2 \leq \mu_3 \leq \dots \leq \mu_k$$

şekindedir. Test istatistiğinin değeri ise aşağıda verilen eşitlik kullanılarak hesaplanmaktadır.

$$z = \frac{\sum_{i=1}^k n_i (L_i - M_i) \bar{R}_i}{\sqrt{(N(N+1)/12) \sum_{i=1}^k n_i (L_i - M_i)^2}}$$

n_i : i inci gruptaki gözlem sayısı,

\bar{R}_i : i inci gruba ait sıra sayılar m aritmetik ortalaması,

L_i : i inci grubun solunda kalan diğer tüm gruplardaki toplam gözlem sayısı,

M_i : i inci grubun sağında kalan diğer tüm gruplardaki toplam gözlem sayısı,

N : Toplam gözlem sayısı,

Elde edilen z test istatistiğinin değeri, standart normal dağılım tablo değeri olan ∓ 1.96 değeri ile karşılaştırılır. $z \geq \mp 1.96$ ise H_0 hipotezi reddedilir.

MACK-WOLFE TESTİ^{18,19}

Şemsiye noktasının bilindiği durumlar için Mack ve Wolfe tarafından önerilen yöntemle göre, yöntemin test ettiği hipotezler;

$$H_0: \mu_1 = \mu_2 = \mu_3 = \dots = \mu_k$$

$$H_1: \mu_1 \geq \mu_2 \geq \dots \geq \mu_{p-1} \geq \mu_p \leq \mu_{p+1} \leq \dots \leq \mu_k$$

şekindedir. Her bir grup için,

$$A_l = \sum_{i=1}^{l-1} \sum_{j=i+1}^l U_{ij} + \sum_{i=l}^{k-1} \sum_{j=i+1}^k U_{ji}$$

$$E(A_l) = \frac{N_1^2 + N_2^2 - \sum_{i=1}^k n_i^2 - n_l^2}{4}$$

$$\sigma^2(A_l) = \frac{2(N_1^3 + N_2^3) + 3(N_1^2 + N_2^2) - \sum_{i=1}^k n_i^2 (2n_i + 3) - n_l(2n_l + 3) + 12n_l N_1 N_2 - 12n_l^2 N}{72}$$

$$A_l^* = \frac{A_l - E(A_l)}{\sqrt{\sigma^2(A_l)}}$$

k : Toplam grup sayısı,

U_{ij} : i inci gruptaki gözlem değerlerinden büyük j inci gruptaki gözlem sayısı,

n_i : i inci gruptaki gözlem sayısı,

n_l : l inci gruptaki gözlem sayısı,

$$N_1 = \sum_{i=1}^l n_i \quad N_2 = \sum_{i=l}^k n_i$$

N : Toplam gözlem sayısı,

değerleri hesaplanır. Test istatistiğinin değeri $A_{max}^* = \max(A_1^*, A_2^*, \dots, A_k^*)$ olarak belirlenir. Elde edilen A_{max}^* değeri, Mack-Wolfe testi için geliştirilen k serbestlik dereceli α yanılma düzeyindeki tablo değeri ile karşılaştırılır. $A_{max}^* \geq A_{k,\alpha}$ ise H_0 hipotezi reddedilir.

SIMPSON-MARGOLIN TESTİ²⁰

Şemsiye noktasının bilinmediği durumlar için Simpson ve Margolin tarafından önerilen yöntemle göre, F_i ($i = 1, 2, \dots, k$) i inci gruba ait birikimli dağılım fonksiyonu olmak üzere yöntemin test ettiği hipotezler;

$$H_0: F_1(x) = F_2(x) = \dots = F_k(x)$$

$$H_1: F_1(x) \geq \dots \geq F_h(x) \leq \dots \leq F_k(x)$$

şeklindedir. Test istatistiğinin değeri için,

$$S_k = \sum_{i=1}^{k-1} \sum_{j=i+1}^k U_{ij}$$

$$R_k = \sum_{j=1}^{k-1} U_{jk}$$

$$c_i = \frac{1}{2} n_i N_{i-1}$$

$$N_i = n_1 + n_2 + \dots + n_i$$

k : Toplam grup sayısı,

U_{ij} : i 'inci gruptaki gözlem değerlerinden büyük j 'nci gruptaki gözlem sayısı,

n_i : i 'inci gruptaki gözlem sayısı,

S_k : Jonckheere – Terpstra istatistiği,

R_k : Mann – Whitney istatistiği,

değerleri hesaplanır. Test istatistiğinin değeri $M = \max(i: R_i \geq c_i)$ olmak üzere $S_M = R_2 + \dots + R_M$ olarak belirlenir. Yönteme göre $R_2 = S_2$ 'dir. Elde edilen S_M değeri, Jonckheere-Terpstra testi için geliştirilen k ve n_i serbestlik dereceli α yanılma düzeyindeki tablo değeri ²¹ ile karşılaştırılır. $S_M \geq$ Jonckheere – Terpstra $_{k,n_i,\alpha}$ ise H_0 hipotezi reddedilir.

HETTMANSPERGER-NORTON TESTİ²²

Hettmansperger ve Norton tarafından önerilen yöntem göre, X_{ij} ($j = 1, \dots, k$ ve $i = 1, \dots, n_j$) gözlem değerlerine ait birikimli dağılım fonksiyonu $F(X - \theta_j)$ olmak üzere yöntemin test ettiği hipotezler;

$$H_0: \theta_1 = \theta_2 = \theta_3 = \dots = \theta_k \quad (\theta_j, j. \text{inci gruba ait medyan}),$$

$$H_1: \theta_j = \theta_0 + \theta c_j \quad (\theta > 0; j = 1, \dots, k)$$

şeklindedir. Test istatistiğinin değeri için,

$$V = \sum \lambda_j (c_j - \bar{c}_w) \bar{R}_j$$

$$c_j = \begin{cases} j, & j = 1, \dots, p \\ 2p - j, & j = p + 1, \dots, k \end{cases}$$

$$\lambda_j = \frac{n_j}{N}$$

$$\bar{R}_j = \sum_{i=1}^{n_j} \frac{R_{ij}}{n_j}$$

$$N = \sum n_j$$

$$\bar{c}_w = \sum \lambda_j c_j$$

n_j : j 'inci gruptaki gözlem sayısı,

R_{ij} : i 'inci gruptaki j 'inci gözlem değerine karşılık gelen sıra sayısı,

değerleri hesaplanır. R_{ij} değerleri belirlenirken gruplara ait gözlem değerlerinin tamamı aynı anda dikkate alınmaktadır. V istatistiği, beklenen değeri sıfır ve varyansı,

$$\sigma^2(V) = \left[\frac{N+1}{12} \right] \sum \lambda_j (c_j - \bar{c}_w)^2$$

olan standart normal dağılıma sahiptir. Dolayısıyla, test istatistiğinin değeri,

$$V^* = \frac{V}{\sqrt{\sigma^2(V)}}$$

formülü ile hesaplanır. Elde edilen V^* değeri, standart normal dağılım tablo değeri olan ∓ 1.96 değeri ile karşılaştırılır. $V^* \geq \mp 1.96$ ise H_0 hipotezi reddedilir.

Şemsiye noktasının bilinmediği durumlarda ise test istatistiğinin değeri,

$$V_{max}^* = \max_{1 \leq t \leq k} \{V_t^*\}$$

olarak belirlenir.

SHİ TESTİ²³

v değeri $1 \leq v \leq k$ olacak şekilde belirlenen şemsiye noktasını göstermek üzere test istatistiğinin değeri,

$$T_v = \left(\frac{12}{N+1} \right)^{1/2} \sum_{i=1}^k a_{vi} \lambda_i \bar{R}_i$$

formülü kullanılarak hesaplanmaktadır. Formülde,

$$N = \sum_{i=1}^k n_i$$

$$\lambda_i = \frac{n_i}{N}$$

$$S_i = \lambda_1 + \lambda_2 + \dots + \lambda_i$$

$$\bar{R}_i = \sum_{j=1}^{n_i} \frac{R_{ij}}{n_i}$$

$$a_{vi} = \begin{cases} \beta_v \left[\sqrt{S_{i-1}(S_k - S_{i-1})} - \sqrt{S_i(S_k - S_i)} \right] / \lambda_i & \text{eğer } i < v \text{ ise} \\ \beta_v \left[\sqrt{S_{v-1}(S_k - S_{v-1})} + \sqrt{S_v(S_k - S_v)} \right] / \lambda_i & \text{eğer } i = v \text{ ise} \\ \beta_v \left[\sqrt{S_i(S_k - S_i)} - \sqrt{S_{i-1}(S_k - S_{i-1})} \right] / \lambda_i & \text{eğer } i > v \text{ ise} \end{cases}$$

n_i : i inci gruptaki gözlem sayısı,

R_{ij} : i 'inci gruptaki j 'inci gözlem değerine karşılık gelen sıra sayısı,

göstermektedir. Yöntemin test ettiği hipotezler;

$$B_0 = \{z \in R^k; z_1 = \dots = z_k\},$$

$$B_v = \{z \in R^k; z_1 \leq \dots \leq z_v \geq \dots \geq z_k\},$$

$$C_v = B_v \cap \left\{ z \in R^k; \sum \lambda_i z_i = 0, \sum \lambda_i z_i^2 = 1 \right\}$$

olmak üzere,

$$H_0: \theta_1 = \theta_2 = \theta_3 = \dots = \theta_k$$

$$H_1: \theta \in B_0^c \cap B_v \quad (\theta = \theta_1, \dots, \theta_k)$$

şekindedir. β_v değeri $a_v = (a_{v1}, \dots, a_{vk}) \in C_v$ olacak şekilde seçilen normalleştirme sabitidir. Elde edilen T_v değeri, α yanılma düzeyinde seçilen standart normal dağılım tablo değeri U_α ile karşılaştırılır. $T_v \geq U_\alpha$ ise H_0 hipotezi reddedilir. Şemsiye noktasının bilinmediği durumlarda ise test istatistiğinin değeri,

$$T_{rank}^* = \max\{T_1, T_2, \dots, T_{k-1}\}$$

olarak belirlenir.

CHEN-WOLFE TESTİ²⁴

Hettmansperger ve Norton (1987) tarafından önerilen yöntem ile benzer görüşler dikkate alınarak şemsiye noktasının bilinmediği durumlar için Chen ve Wolfe tarafından önerilmiştir. Yöntemin test ettiği hipotezler;

$$H_0: \mu_1 = \mu_2 = \mu_3 = \dots = \mu_k$$

$$H_1: \mu_1 \geq \mu_2 \geq \dots \geq \mu_{p-1} \geq \mu_p \leq \mu_{p+1} \leq \dots \leq \mu_k$$

şekindedir. Test istatistiğinin değeri,

$$(\bar{\chi}_p^2)^{1/2} = \max \left\{ \left(\frac{12}{N+1} \right)^{1/2} \sum_{i=1}^k c_i \lambda_i \bar{R}_i \right\}$$

formülü kullanılarak hesaplanmaktadır. $N = \sum_{i=1}^k n_i$ ve $\lambda_i = n_i/N$ olmak üzere maximum koşulunun sağlanabilmesi için c_i değerlerinin;

$$a. \quad \sum \lambda_i c_i = 0$$

$$b. \quad \sum \lambda_i c_i^2 = 1$$

$$c. \quad c_1 \geq c_2 \geq \dots \geq c_{p-1} \geq c_p \leq c_{p+1} \leq \dots \leq c_k$$

olacak şekilde belirlenmesi gerekmektedir. Bu kısıtlardan da anlaşılacağı üzere yöntem, kısıtlı optimizasyon tekniklerinin kullanılmasını gerektirmektedir. Hettmansperger-Norton Testi için geçerli olan tüm argümanlar bu test için de geçerlidir.

MILLEN-WOLFE TESTİ²⁵

Şemsiye noktasının bilindiği ve gruplardaki gözlem sayılarının eşit olduğu durumlar için Millen ve Wolfe tarafından önerilen yöntemle göre test istatistiğinin değeri,

$$T_b = \sum_{l=1}^k b_l A_l$$

formülü kullanılarak hesaplanmaktadır. Formülde,

$$A_l = \sum_{i=1}^{l-1} \sum_{j=i+1}^l U_{ij} + \sum_{i=l}^{k-1} \sum_{j=i+1}^k U_{ji}$$

$$b_l = \frac{U_{+l}}{\sum_{i=1}^k U_{+i}}, \quad l = 1, 2, 3, \dots, k$$

göstermektedir. A_l değerleri Mack-Wolfe test istatistiği değeri, U_{+l} değerleri ise l inci grup ile geri kalan $k-1$ tane grup arasındaki Mann-Whitney U istatistiği değerleridir. Yöntemin test ettiği hipotezler θ_l 'ler grup medyanlarını göstermek üzere,

$$H_0: \theta_1 = \theta_2 = \theta_3 = \dots = \theta_k$$

$$H_1: \theta_1 \leq \theta_2 \leq \dots \leq \theta_{p-1} \leq \theta_p \geq \theta_{p+1} \geq \dots \geq \theta_k$$

şekindedir. Elde edilen T_b değeri, gözlem sayısı (n_l), grup sayısı (k) ve α yanılma düzeyi dikkate alınarak yöntem için geliştirilen tablo değeri $T_{n_l, k, \alpha}$ ile karşılaştırılır. $T_b \geq T_{n_l, k, \alpha}$ ise H_0 hipotezi reddedilir.

SONUÇLAR

Literatürde, çalışmada dikkate alınan yöntemlerin birbirleri ile karşılaştırıldığı sınırlı sayıda çalışma bulunmaktadır. Williams (1971) çalışmasında, Bartholomew testinin gücünün kendisi ta-

rafından önerilen testten daha güçlü olduğunu göstermiştir.¹³ Mack ve Wolfe (1981) kendileri tarafından önerilen testin gücünün Jonckheere-Terpstra testinden daha iyi olduğunu bildirmişlerdir.¹⁸ Simpson ve Margolin (1986) kendileri tarafından önerilen testin gücünün bazı durumlarda, Jonckheere-Terpstra ve Chacko tarafından önerilen testlere göre düşük olmasına rağmen önerdikleri testin birçok durumda Jonckheere-Terpstra ile Chacko testleri arasında iyi bir uzlaşık güce sahip olduğunu ifade etmişlerdir.²⁰ Shi (1988) tarafından yapılan çalışmada Shi Testi ile Shorack Testi'nin eşit güçlere sahip oldukları belirtilmiştir.²³ Simpson ve Margolin (1990), Monte Carlo ve asimtotik güç bakımından birçok durumda Jonckheere-Terpstra testinin kendileri tarafından önerilen testten daha güçlü olduğunu ifade etmişlerdir.²⁶ Chen ve Wolfe (1990) tarafından sıralı seçenek hipotezin söz konusu olduğu durumlarda Jonckheere-Terpstra testinin, şemsiye noktasının bilindiği şemsiye seçenekli hipotezin söz konusu olduğu durumlarda Mack-Wolfe ve Hettmansperger-Norton testlerinin Chacko testinden daha güçlü olduğu, şemsiye noktasının bilinmediği durumlarda ise Simpson-Margolin testinin en güçlü test olduğu bildirilmiştir.²⁴ Chen (1992) tarafından yapılan çalışmada Hettmansperger-Norton, Chen-Wolfe ve Mack-Wolfe testleri karşılaştırılmış sonuç olarak Hettmansperger-Norton testinin kullanılması önerilmiştir. Yöntemler güçleri dikkate alındığında iyiden kötüye doğru Hettmansperger-Norton, Mack-Wolfe ve Chen-Wolfe şeklinde sıralanmıştır.²⁷ Chen ve Wolfe (1993) tarafından yapılan çalışmada Mack-Wolfe testinin Simpson-Margolin testinden, Simpson-Margolin testinin

de Shirley testinden daha güçlü olduğu belirlenmiştir.²⁸ Neuhauser ve ark., (1998) tarafından yapılan çalışmada, uyarlanmış Jonckheere-Terpstra testinin Jonckheere-Terpstra testinden daha güçlü olduğu ifade edilmiştir.²⁹ Magel ve Qin (2003) tarafından yapılan çalışmada şemsiye noktasının bilinmediği birçok durumda Chen-Wolfe testinin Mack-Wolfe testinden daha güçlü olduğu ifade edilmiştir.³⁰ Millen ve Wolfe (2005) tarafından yapılan çalışmada, Shi Testi'nin uygulama zorluğu ve tutucu özelliklerinden dolayı pratikte daha az kullanıldığı ifade edilmiş olup, kendileri tarafından önerilen test ile Hettmansperger-Norton ve Mack-Wolfe testleri karşılaştırılmıştır. Güç değerleri dikkate alınarak yapılan karşılaştırma sonucunda şemsiye noktasının bilindiği dengeli durumlar için yöntemlerin iyiden kötüye doğru Millen-Wolfe, Hettmansperger-Norton, Mack-Wolfe şeklinde sıralandıklarını belirtmişlerdir.²⁵ Kössler (2006) tarafından yapılan çalışmada ise Hettmansperger-Norton, Chen-Wolfe ve Shi tarafından önerilen testler güçleri dikkate alınarak birbirleri ile karşılaştırılmıştır. Şemsiye noktasının bilinmediği, farklı örneklem büyüklükleri dikkate alınarak yapılan karşılaştırmalar sonucunda yöntemler iyiden kötüye doğru Hettmansperger-Norton, Chen-Wolfe, Shi şeklinde sıralanmıştır.³¹

Literatürden elde edilen sonuçlar genel olarak dikkate alındığında sıralı seçenek hipotezin söz konusu olduğu durumlar için Jonckheere-Terpstra testinin, şemsiye seçenekli hipotezin söz konusu olduğu durumlarda ise Hettmansperger-Norton testinin tercih edilmesi gerektiği söylenebilir.

KAYNAKLAR

1. Jessica YC, Hongshik A, James JC. On sequential closed testing dose groups with a control. *Communications in Statistics-Theory and Methods* 2000;29(5-6):941-56.
2. Budde M, Bauer P. Multiple test procedures in clinical dose finding studies. *J Am Stat Assoc* 1989;84(407):792-6.
3. Shirley E. A non-parametric equivalent of Williams' test for contrasting increasing dose levels of a treatment. *Biometrics* 1977;33(2):386-9.
4. Terpstra TJ. The asymptotic normality and consistency of Kendall's test against trend when ties are present. *Indagationes Mathematicae* 1952;14(3):327-33.
5. Jonckheere AR. A distribution-free k-sample test against ordered alternatives. *Biometrika* 1954;41(1-2):133-45.
6. Tryon PV, Hettmansperger TP. A class of non-parametric tests for homogeneity against ordered alternatives. *Annals of Statistics* 1973;1(6):1061-70.
7. Bartholomew DJ. A test of homogeneity for ordered alternatives. *Biometrika* 1959;46(1-2):36-48.
8. Bartholomew DJ. A test of homogeneity of means under restricted alternatives. *Biometrika* 1961;23(2):239-81.
9. Chacko VJ. Testing homogeneity against ordered alternatives. *Annals of Mathematical Statistics* 1963;34(3):945-56.

10. Chacko VJ. Modified Chi-square test for ordered alternatives. *Sankhyā: The Indian Journal of Statistics, Series B* 1966;28(3-4): 185-90.
11. Roth GL, Daniel WW. Critical values for Chacko's homogeneity test against ordered alternatives. *Educational and Psychological Measurement* 1978;38(4):889-91.
12. Shorack GR. Testing against ordered alternatives in model I analysis of variance: normal theory and nonparametric. *Annals of Mathematical Statistics* 1967;38(6):1740-52.
13. Williams DA. A test for differences between treatment means when several dose levels with a zero-dose control. *Biometrics* 1971;27(1):103-17.
14. Williams DA. The comparison of several dose levels with a zero-dose control. *Biometrics* 1972;28(2):519-31.
15. Williams DA. A note on Shirley's nonparametric test for comparing several dose levels with a zero-dose control. *Biometrics* 1986;42(1): 183-6.
16. Le CT. A new rank test against ordered alternatives k -sample problems. *Biometrical Journal* 1988;30(1):87-92.
17. Alonzo TA, Nakas CT, Yiannoutsos CT, Bucher S. A comparison of tests for restricted orderings in the three-class case. *Stat Med* 2009;28(7):1144-58.
18. Mack GA, Wolfe DA. K -sample rank tests for umbrella alternatives. *J Am Stat Assoc* 1981;76(373):175-81.
19. Chen YI. Notes on the Mack-Wolfe and Chen-Wolfe tests for umbrella alternatives. *Biometrical J* 1991;33(3):281-90.
20. Simpson DG, Margolin BH. Recursive nonparametric testing for dose-response relationships subject to downturns at high doses. *Biometrika* 1986;73(3):589-96.
21. van de Wiel MA, Di Bucchianico A, van der Laan P. Symbolic computation and exact distributions of nonparametric test statistics. *Journal of the Royal Statistical Society. Series D (The Statistician)* 1999;48(4):507-16.
22. Hettmansperger TP, Norton RM. Tests for patterned alternatives in k -sample problems. *J Am Stat Assoc* 1987;82(397):292-9.
23. Shi NZ. Rank test statistics for umbrella alternatives. *Communications in Statistics: Theory and Methods* 1988;17(6):2059-73.
24. Chen YI, Wolfe DA. A study of distribution-free tests for umbrella alternatives. *Biometric* 1990;32(1):47-57.
25. Millen BA, Wolfe DA. A class of nonparametric tests for umbrella alternatives. *Journal of Statistical Research* 2005;39(1):1-18.
26. Simpson DG, Margolin BH. Nonparametric testing for dose-response curves subject to downturns: asymptotic power considerations. *Annals of Statistics* 1990;18(1):373-90.
27. Chen YI. Robust umbrella tests for a generalized Behrens-Fisher problem. *Biometrical J* 1992;34(3):347-57.
28. Chen YI, Wolfe DA. Nonparametric procedures for comparing umbrella pattern treatment effects with a control in a one-way layout. *Biometrics* 1993;49(2):455-65.
29. Neuhauser M, Liu PY, Hothorn L. Nonparametric tests for trend: Jonckheere's test, a modification and a maximum test. *Biometrical J* 1998;40(8):899-909.
30. Magel RC, Qin L. A non-parametric test for umbrella alternatives based on ranked-set sampling. *Journal of Applied Statistics* 2003;30(8):925-37.
31. Kössler W. Some c -sample rank tests of homogeneity against umbrella alternatives with unknown peak. *Journal of Statistical Computation and Simulation* 2006;76(1):57-74.