

Çapraz Tasarımlarda Etkilerin Testi için Parametrik ve Parametrik Olmayan Testlerin Karşılaştırılması

Hilal SARI¹, Esra GÖKPINAR*¹, Hülya BAYRAK¹

¹Gazi Üniversitesi, Fen Fakültesi, İstatistik Bölümü, 06500, Ankara

(Alınış / Received: 04.07.2016, Kabul / Accepted: 22.06.2017, Online Yayınlanma / Published Online: 16.08.2017)

Anahtar Kelimeler

Çapraz tasarım,
t test,
Mann Whitney U test,
Simülasyon

Özet: Çapraz tasarımlar, psikolojik denemeler, tarım, zooloji gibi birçok alanda kullanılan tasarımlardır. Bu tasarımlar daha az örneklem kullanarak daha güçlü sonuçlar elde etmek amacıyla kullanılırlar. Bu durum, çapraz tasarımların önemli avantajlarından biridir. Çalışmada, çapraz tasarımı hakkında ayrıntılı bilgi verilmiştir. Çapraz tasarımda etkilerin analizi için kullanılan t testi ve Mann Whitney U testi simülasyon yoluyla deneysel I. tip hata oranları ve güç değerleri bakımından karşılaştırılmıştır. Bu amaçla farklı örnek çapları, farklı korelasyon katsayıları ele alınmış ve ayrıca testlerin performansları normal ve normal olmayan dağılımlar altında incelenmiştir. Simülasyon sonuçlarına göre, normal dağılım altında t testinin güç değerleri daha yüksek iken, normal olmayan dağılımlar altında Mann Whitney U testinin güç değerlerinin daha yüksek olduğu gözlenmiştir.

Comparison of Parametric and Nonparametric Tests for Testing Effects in Crossover Design

Keywords

Crossover design,
t test,
Mann Whitney U test,
Simulation

Abstract: Crossover designs are designs used in many fields such as psychological trials, agriculture, zoology. These designs are used in order to achieve stronger results by using less sample. This situation is the one of the important advantages of crossover designs. In this article, the detailed information about the crossover design were given. For testing effects in crossover design, the t test and Mann Whitney U test were compared in terms of the empirical type I error rate and power by using simulation study. For this purpose, the different sample size and different correlation coefficients were taken, also the performances of these tests were investigated under the normal and non-normal distributions. According to the results of simulation, while the power of the t test is better under the normal distribution, the power of the Mann Whitney U test is better under the non-normal distributions.

1. Giriş

Çapraz tasarım, belirli sayıdaki deneme biriminin her birine iki ya da daha fazla işlemin belirli sıra ile uygulandığı bir deney tasarım türüdür. Bu tasarımda bir işlem, belirli bir zaman periyodunda bir deney birimine uygulandıktan sonra aynı birime bir sonraki periyotta başka bir işlem uygulanmaktadır. Bu tarz tasarımlarda genellikle denekler arası farklılık, denek içi farklılıktan daha büyüktür. Çapraz tasarımın en büyük avantajı aynı deney biriminden ölçümlerin alınmasıyla her bir deney biriminin kendi kontrolü olarak kullanılabilmesidir. Böylece işlemlerin denek içinde karşılaştırılması sağlanarak deney birimlerinin farklılığından kaynaklanan hata ortadan kaldırılmış

olur. Çapraz tasarımlarda geçerli işlem karşılaştırmalarını yapabilmek için her bir periyodun başlangıcında denekler karşılaştırılabilir koşullarda olmalıdır. Örneğin hamilelik, ölüm gibi sonlanan ya da soğuk algınlığı gibi iyileşme gösteren geçici durumlarda işlem uygulanan denek başlangıçtaki haline dönmeyeceğinden bu tasarımların kullanılması uygun olmaz. Çapraz tasarımlar, semptomları durağan olan kronik vakalar için uygunken akut koşullarda uygun olmayabilir.

Çapraz tasarımlarda işlemler aynı birime uygulandığından yani işlemler birimler içinde karşılaştırıldığından, deney birimlerinin farklılığından kaynaklanan hata ortadan kaldırılarak

*İlgili yazar: eyigit@gazi.edu.tr

işlemlerin karşılaştırılmasının daha doğru, kesin ve tutarlı şekilde yapılması sağlanır. Bu durum çapraz tasarımın başlıca avantajıdır. Bu deney tasarımı örneklem sayısının az olduğu durumlarda kullanıldığı için oldukça avantaj sağlamaktadır. Böylece bir çalışma için sınırlı sayıda deney birimi olduğunda kaynakların ekonomik olarak kullanımını sağlar. Çünkü bu tasarımlar diğer tasarımlara göre aynı sayıdaki gözlem değerinin daha az sayıda deney birimi kullanarak elde edilmesini sağlamaktadır.

Çapraz tasarımlar genel olarak, klinik araştırmalar başta olmak üzere ilaç kullanımı, psikolojik denemeler, tarım ile zooloji alanlarında, otomotiv ve lastik endüstrisinde ve çeşitli alternatiflerin karşılaştırılmasını içeren diğer endüstri alanlarında kullanılmaktadır [1]. Örneğin, McNair [2] anti depresyon ilaçların insan üzerinde etkilerini inceleyen araştırmasında, çalışmalarının %68'inin çapraz tasarım yaklaşımı kullanıldığından bahsetmiştir. Jones ve Kenward [3] çalışmasında ise dişlerdeki plak miktarını ölçmek için gargara örneklerini ele almıştır. Bu çalışmada ilk grupta 18 kişi mevcutken ikinci grupta 16 kişi vardır. Deney, iki tane 6 haftalık işlem periyodu ve 3 haftalık temizlenme periyodundan oluşmaktadır. Son yıllarda da çapraz tasarım hakkında geniş bir çalışma yapılmıştır [4-11].

Çapraz tasarımlarda farkların gerçek işlem etkisine göre rastgele dağılmasını etkileyebilecek üç temel faktörün etkisi dikkate alınmalıdır. Bunlar; periyot etkisi (period effect or time effect), işlem etkisi ve aktarılmış işlem etkisidir (carryover effect). Asıl amaç bu etkilerin olup olmadığının araştırılmasıdır. Hills ve Armitage [12] ilk olarak çapraz tasarımda t testi kullanmışlardır. Brown [13] çalışmasında 2 işlem 2 periyot çapraz tasarımlar için t testini kullanmıştır. Ayrıca çalışmasının bir kısmında çapraz tasarımın geçerlilik ve etkinliğine de yer vermiştir. Bellavance ve Tardif [14] çalışmalarında biraz daha karışık olan 3 işlem 3 periyot çapraz tasarımları için işlem etkisinin olmadığı ve aktarılmış etkinin olmadığına dair yokluk hipotezlerinin testi için parametrik olmayan bir yaklaşım tanımlamışlardır. Öhrvik [15] çalışmalarında 3 veya daha fazla işlem için çapraz tasarımı parametrik olmayan yöntemlerle analiz etmiştir. İşlem sayısı 2 olduğunda Wilcoxon sıra toplamları istatistiğini önermiş ve işlem sayısı 3 veya daha fazla olduğunda Q istatistiğini kullanmak gerektiğini belirtmiştir. Putt ve Chinchilli [16] de çalışmalarında Wilcoxon sıra toplamları istatistiğini ve Q istatistiğini kullanmışlardır. Cleophas [17] ise daha farklı bir yaklaşımla kendi kontrollü klinik denemelerde t istatistiğini kullanarak işlem yanıtları arasındaki korelasyon düzeylerinin testin duyarlılığında önemli bir belirleyici olduğunu gözlemlemiştir. İşlem sayısı 2 olduğunda t testini kullanırken işlem sayısı 2'den fazla olduğunda varyans analizini kullanmıştır.

Uygulama alanlarının giderek artması sebebiyle bu

çalışmada çapraz tasarım modeli ele alınmıştır. Bu amaçla çalışma şu şekilde düzenlenmiştir. İkinci bölümde 2 işlem 2 periyot çapraz tasarım türü incelenmiş olup etkilerin olup olmadığının analizi için kullanılan parametrik t testi ve parametrik olmayan test tanıtılmıştır. Üçüncü bölümde bu testlerin farklı dağılımlar altında performansları testin birinci tip hata oranları ve güçleri bakımından karşılaştırılması yapılmıştır. Dördüncü bölümde sonuç ve yorumlara yer verilmiştir.

2. Çapraz Tasarımda Etkilerin Analizi

Y_{ijk} , rastgele değişkeninin gözlenen değeri olmak üzere çapraz tasarım için istatistiksel model aşağıdaki gibidir:

$$Y_{ijk} = \mu + s_{ik} + \pi_j + \tau_{[i,j]} + \lambda_{[i,j-1]} + \varepsilon_{ijk} \quad (1)$$

Burada y_{ijk} , i. gruptaki j. periyodunun k. biriminin yanıtı; μ , genel ortalama; s_{ik} , i. gruptaki k. birimin etkisi ($i=1, \dots, s$, $k=1, \dots, n_i$); π_j , j. periyotun etkisi ($j=1, \dots, p$); $\tau_{[i,j]}$: grup i'nin periyot j'de uygulanan işleme direkt etkisi (işlem etkisi); $\lambda_{[i,j-1]}$: periyot j'de hala devam eden aktarılmış etki (periyot j-1'de grup i'nin uygulanan işleme etkisi ölçülürken, burada $\lambda_{[i,0]} = 0$); ε_{ijk} , rassal hatadır [18]. Burada $s_{ik} \sim N(0, \sigma_s^2)$, $\varepsilon_{ijk} \sim N(0, \sigma^2$ ve $E(\varepsilon_{ijk}, s_{ik}) = 0$ ($\forall i, j, k$). Analiz işlemlerinde toplam yanıt

$$Y_{ij.} = \sum_{k=1}^{n_i} Y_{ijk}, Y_{i..} = \sum_{j=1}^p Y_{ij.}, Y_{...} = \sum_{i=1}^s Y_{i..} \quad (2)$$

ve ortalamalar

$$y_{ij.} = \frac{Y_{ij.}}{n_i}, y_{i..} = \frac{Y_{i..}}{pn_i}, y_{...} = \frac{Y_{...}}{p \sum_{i=1}^s n_i} \quad (3)$$

şeklinde elde edilir. En yaygın kullanılan çapraz tasarım, Grizzle [19] tarafından tanımlanmış olan 2x2 çapraz tasarımlar olarak adlandırılan 2 sıra, 2 periyot, 2 işlemlerli çapraz tasarımlardır. Bu tasarımlarda işlem etkilerinin karşılaştırılması amacıyla her denek farklı periyotlarda A ve B olarak etiketlendirilen iki farklı işlem alır. Deneklerin yarısına önce A sonra B işlemi uygulanırken diğer kalan deneklere ise önce B sonra A işlemi uygulanır. Buna göre 2x2 çapraz tasarım modelinde etkilerin gösterimi Tablo 1'de verildiği gibidir.

Tablo 1. 2x2 çapraz tasarım modeli

Grup	Periyot 1	Periyot 2
1 (AB)	$\mu + \pi_1 + \tau_1$ $+ s_{1k} + \varepsilon_{11k}$	$\mu + \pi_2 + \tau_2 + \lambda_1$ $+ s_{1k} + \varepsilon_{12k}$
2 (BA)	$\mu + \pi_1 + \tau_2$ $+ s_{2k} + \varepsilon_{21k}$	$\mu + \pi_2 + \tau_1 + \lambda_2$ $+ s_{2k} + \varepsilon_{22k}$

Buna göre λ_1 ve λ_2 sırasıyla A ve B işlemlerinin aktarılmış etkilerini ifade eder.

Çapraz tasarımda aktarılmış etki, işlem etkisi ve periyot etkisi olmak üzere 3 çeşit etki mevcuttur.

Çapraz tasarımda bu etkilerin olmadığı yokluk hipotezlerinin testleri için kullanılan t testi ve Mann Whitney (MW) testlerinin analizleri aşağıdaki gibidir.

2.1. İşlem etkisi ve istatistiksel analizi

Çapraz tasarımlarda işlem etkisi, periyot etkisi ve aktarılmış etkinin anlamlılığı için hipotezler kurulabilir. İşlem etkisinde, birimlere uygulanan işlemlerin hangisinin daha iyi olduğuna ya da birimlere uygulanan işlemlerin aynı etkiye sahip olup olmadıklarına bakılır. $\lambda_1 = \lambda_2 = \lambda$ olduğu varsayılarak işlem etkisinin hipotez testi için yokluk ve alternatif hipotezleri; $H_0: \tau_1 = \tau_2$ ise işlem etkisi yoktur, $H_1: \tau_1 \neq \tau_2$ ise işlem etkisi vardır.

İşlem etkisinin yokluk hipotezini test etmek için kullanılan parametrik t testi aşağıdaki gibidir. i-inci grup için d_{ik} , d_i ve s_{iD}^2 ($i=1,2$) sırasıyla

$$d_{ik} = Y_{i1k} - Y_{i2k} \quad (4)$$

$$d_i = \sum_{k=1}^{n_i} (Y_{i1k} - Y_{i2k}) \quad (5)$$

ve

$$s_{iD}^2 = \frac{1}{n_i-1} \sum_{k=1}^{n_i} (d_{ik} - d_i)^2 \quad (6)$$

olmak üzere

$$\hat{\tau}_d = \frac{1}{2} (d_{1.} - d_{2.}) \quad \text{ve} \quad s_D^2 = \frac{(n_1-1)s_{1D}^2 + (n_2-1)s_{2D}^2}{n_1+n_2-2} \quad (7)$$

şeklinde elde edilir. Test istatistiği

$$T_\tau = \frac{\hat{\tau}_d}{\frac{1}{2}s_D} \sqrt{\frac{n_1 n_2}{n_1+n_2}} \quad (8)$$

şeklinde olup H_0 hipotezinin doğruluğu altında $(n_1 + n_2 - 2)$ serbestlik dereceli t dağılımına sahiptir [18].

2.2. Periyot etkisi ve istatistiksel analizi

Birimlerin durumları birinci ve ikinci gözlem arasında değiştiğinde periyot etkileri meydana gelebilir. Herhangi bir periyot etkisi çapraz tasarımın gücünü azaltan, tasarımın genel verimini düşürmeye eğilimi olan birim içi değişebilirliği artırır [20].

$\lambda_1 + \lambda_2 = 0$ verilmişken periyot etkisinin testi aşağıdaki yokluk hipotezi kullanılarak test edilebilir.

$$\begin{aligned} H_0: \pi_1 &= \pi_2 & (\text{periyot etkisi yoktur}) \\ H_1: \pi_1 &\neq \pi_2 & (\text{periyot etkisi vardır}) \end{aligned} \quad (9)$$

İşlem etkisinin yokluk hipotezini test etmek için kullanılan parametrik t testi aşağıdaki gibidir. $c_{1k} = d_{1k}$ ve $c_{2k} = -d_{2k}$ olmak üzere

$$\hat{\pi}_d = \frac{1}{2} (c_{1.} - c_{2.}) \quad (10)$$

olarak elde edilir. Buna göre test istatistiği

$$T_\pi = \frac{\hat{\pi}_d}{\frac{1}{2}s_D} \sqrt{\frac{n_1 n_2}{n_1+n_2}} \quad (11)$$

şeklinde olup H_0 hipotezinin doğruluğu altında $(n_1 + n_2 - 2)$ serbestlik dereceli t dağılımına sahiptir [2].

2.3. Aktarılmış etki ve istatistiksel analizi

Aktarılmış etki, bir periyotta uygulanan işlemin deney birimi üzerindeki etkisinin takip eden periyotta da devam etmesidir. Aktarılmış etkiler, herhangi bir periyot ya da işlem ile tasarımı etkiler ve direk işlem etkisinin yanlı sonuç vermesine sebep olur. Bu nedenle aktarılmış etkinin varlığı ihmal edilmemeli ve mutlaka analiz edilmelidir [21]. Buna göre aktarılmış etkinin hipotez testi için yokluk ve alternatif hipotezler aşağıda verildiği gibidir.

$$\begin{aligned} H_0: \lambda_1 &= \lambda_2 & (\text{aktarılmış etki yoktur}) \\ H_1: \lambda_1 &\neq \lambda_2 & (\text{aktarılmış etki vardır}) \end{aligned} \quad (12)$$

H_0 hipotezini test etmek için parametrik test istatistiği olan t testi aşağıdaki gibi elde edilir.

$y_{1..}$ ve $y_{2..}$ sırasıyla 1. ve 2. grubun ortalaması olmak üzere

$$\hat{\lambda}_d = 2(y_{1..} - y_{2..}) \quad (13)$$

şeklinde olup. Her grubun örnek varyansı $s_i^2 = \frac{1}{n_i-1} (\sum_{k=1}^{n_i} Y_{i.k}^2 - \frac{Y_{i.}^2}{n_i})$ olmak üzere

$$s^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2} \quad (14)$$

olarak elde edilir. Buna göre test istatistiği

$$T_\lambda = \frac{\hat{\lambda}_d}{s} \sqrt{\frac{n_1 n_2}{n_1+n_2}} \quad (15)$$

dir ve H_0 hipotezinin doğruluğu altında $(n_1 + n_2 - 2)$ serbestlik dereceli Student's t dağılımına sahiptir [2].

İşlem \times periyot etkileşiminin olmaması diğer bir ifadeyle A ve B işlemleri arasındaki farkın değerlendirildikleri işlem sırasını dikkate almaksızın aynı olması ve aktarılmış işlem etkisinin olmaması klasik 2×2 çapraz tasarımların iki önemli varsayımıdır [22].

Uygulamada aktarılmış etkinin varlığı veya yokluğu nadiren bilinmektedir hatta çoğu çalışmada aktarılmış etkinin varlığı göz ardı edilip analize devam edilmektedir. Grizzle [19], işlem karşılaştırmaları için analiz edilen veriler ile aynı verileri kullanarak bazı nominal anlamlılık düzeyinde

aktarılmış etkinin yokluğu hipotezinin testini önerdi. Eğer bu hipotez reddedilirse yani aktarılmış etkinin varlığı söz konusu ise o zaman sadece ilk periyodun verileri nominal α anlamlılık düzeyinde işlem etkilerini test etmek için kullanılır. Aksi takdirde yani aktarılmış etkinin yokluğu durumunda nominal α düzeyinde her iki periyot da kullanılarak işlem etkileri test edilir.

2.4. Etkilerin analizi için kullanılan parametrik olmayan test

Çapraz tasarımların avantajlarından biri de aynı deney biriminden birden fazla ölçüm alınarak aynı sayıdaki gözlem değerine diğer tasarımlardan daha az sayıda deney birimi kullanılarak ulaşılmasıdır. Örneğin çapraz tasarımlar birinci ve ikinci tip hata riski bakımından aynı kriterleri sağlamak amacıyla klinik denemelerde sıkça kullanılan paralel grup tasarımlardan daha az denek gerektirirler. Bu durumda ise normal dağılım varsayımının sağlanması zorluğundan dolayı ya da parametrik istatistiklerin varsayımları sağlanmadığı durumda parametrik olmayan yaklaşımlar kullanılır. Aktarılmış etki, işlem etkisi ve periyot etkisi için H_0 hipotezlerini test etmek için kullanılan parametrik olmayan testlerden Mann-Whitney test istatistiği aşağıdaki gibidir. Buna göre birinci örneklemin sıra puanları toplamı R_1 , ikinci örneklemin sıra puanları toplamı R_2 olmak üzere U_1 ve U_2 istatistikleri aşağıdaki gibi hesaplanır.

$$U_1 = n_1 n_2 + \frac{n_1(n_1+1)}{2} - R_1 \quad (16)$$

$$U_2 = n_1 n_2 + \frac{n_2(n_2+1)}{2} - R_2 \quad (17)$$

Test istatistiği

$$Z = \frac{U - E(U)}{\sqrt{\text{Var}(U)}} \quad (18)$$

olmak üzere burada $U = \min(U_1, U_2)$, $E(U) = \frac{n_1 n_2}{2}$ ve $\text{Var}(U) = \frac{n_1 n_2 (n_1 + n_2 + 1)}{12}$ şeklindedir. Z test istatistiği standart normal dağılıma sahiptir.

3. Simülasyon Çalışması

Simülasyon çalışmasında, çapraz tasarımda aktarılmış etki ve işlem etkisinin olmadığına dair yokluk hipotezlerinin testleri için kullanılan Mann-Whitney (MW) ve t testlerinin performanslarını belirlemek için deneysel I. tip hata oranları ve güçleri bakımından karşılaştırılması yapılmıştır. Bu amaçla $n=3,5,10,15$ olmak üzere farklı örnek hacimleri ve $r=0,2,0,5, 0,9$ olmak üzere farklı korelasyon katsayısı değerleri ele alınmıştır. Ayrıca normal dağılım ve normal olmayan dağılım (üstel dağılım, kaydırılmış üstel dağılım, standart çifte üstel dağılım) altında testlerin performansları incelenmiştir.

Aktarılmış etkinin (carryover effect) olmadığına dair yokluk hipotezinin testi için çapraz tasarım deney

düzeninde $\mu=[0 \ 0]$ olarak alınmıştır. Buna göre testlerin deneysel I. tip hata oranları bulunurken $\lambda=[1 \ 1]$, testlerin güç değerleri bulunurken ise $\lambda=[1 \ 1.2]$, $[1 \ 1.5]$, $[1 \ 2]$, $[1 \ 2.5]$ olmak üzere λ_i 'nin farklı değerleri ele alınmıştır. $\alpha=0.05$ için bu işlemlerin her biri 5000 kez tekrarlanarak testlerin deneysel I.tip hata oranları ve güç değerleri elde edilmiştir. Simülasyon çalışmasında Matlab R2009A programı kullanılmıştır.

$\alpha=0.05$ ve normal dağılım altında aktarılmış etkinin olmadığına dair yokluk hipotezinin testi için testlerin deneysel I. tip hata oranları ve güç değerleri Tablo 2'de verilmiştir.

Tablo 2'den testlerin deneysel I.tip hata oranlarına bakıldığında tüm n ve r değerleri için t-testinin deneysel I. tip hata oranlarının nominal α değerine oldukça yakın olduğu gözlenmektedir. MW testinin deneysel I. tip hata oranları ise tüm r değerlerinde, $n=3$ iken nominal α değerinden oldukça yüksek çıkarken n değeri arttıkça nominal α değerine yaklaştığı görülmektedir.

Testlerin güç değerlerine bakıldığında ise MW testinin dağılımı asimptotik olarak standart normal dağılıma sahip olduğu için küçük örnek çaplarında testin deneysel I. tip hata oranları nominal α değerinden oldukça yüksek çıktığı gözlenmişti. Bundan dolayı küçük örnek çapları durumunda ($n=3$) MW testin güç değerleri t-testinin güç değerlerine göre daha yüksek çıkmıştır. n değerleri arttıkça testlerin güç değerleri karşılaştırıldığında genel olarak çok az farkla t testinin güç değerlerinin MW testinin güç değerlerinden daha yüksek çıktığı gözlenmiştir. Genel olarak testlerin güç değerlerini incelediğinde de beklenildiği gibi n'lerin artışı ve aktarılmış etki arasındaki fark arttıkça testlerin güç değerlerini arttırmıştır. Ayrıca Tablo 2'den r'lerin artışının testlerin güç değerlerini olumsuz yönde etkileyerek azalttığı görülmektedir.

Gruplardaki aktarılmış etkinin eşitliği altında ($\lambda_1=\lambda_2$), işlem etkisinin olmadığına dair yokluk hipotezinin testi için $\lambda=[1 \ 1]$ olarak alınmıştır. Buna göre testlerin deneysel I. tip hata oranları hesaplanırken $\mu=[1 \ 1]$ olarak alınmıştır. Testlerin güç değerleri bulunurken $\mu=[1 \ 2]$, $[1 \ 3]$, $[1 \ 4]$, $[1 \ 5]$ olmak üzere μ_i 'nin farklı değerleri altında çalışılmıştır.

$\alpha=0.05$ ve normal dağılım altında işlem etkisinin olmadığı yokluk hipotezinin testi için elde edilen testlerin deneysel I. tip hata oranları ve güç değerleri Tablo 3'de verilmiştir.

Testlerin deneysel I.tip hata oranlarına bakıldığında tüm n ve r değerleri için t-testinin deneysel I.tip hata oranlarının nominal α değerine oldukça yakın olduğu gözlenmektedir. MW testinin deneysel I.tip hata oranları ise tüm r değerlerinde, $n=3$ iken nominal α değerinden oldukça yüksek çıkarken n değeri arttıkça nominal α değerine yaklaştığı görülmektedir.

Tablo 2. Testlerin deneysel I. tip hata oranları ve güç değerleri

n	λ	r=0,2		r=0,5		r=0,9	
		T	MW	T	MW	T	MW
3	(1 1)	0,047	0,094	0,049	0,102	0,050	0,095
	(1 1,2)	0,094	0,172	0,083	0,156	0,077	0,146
	(1 1,5)	0,231	0,361	0,203	0,322	0,166	0,269
	(1 2)	0,435	0,590	0,381	0,530	0,314	0,459
	(1 2,5)	0,658	0,792	0,573	0,721	0,483	0,645
5	(1 1)	0,048	0,055	0,045	0,053	0,046	0,052
	(1 1,2)	0,146	0,156	0,123	0,127	0,107	0,116
	(1 1,5)	0,432	0,440	0,361	0,367	0,307	0,304
	(1 2)	0,769	0,762	0,673	0,662	0,569	0,568
	(1 2,5)	0,946	0,934	0,888	0,879	0,808	0,796
10	(1 1)	0,052	0,052	0,052	0,052	0,050	0,050
	(1 1,2)	0,276	0,279	0,234	0,231	0,197	0,199
	(1 1,5)	0,780	0,769	0,685	0,673	0,580	0,566
	(1 2)	0,986	0,983	0,957	0,948	0,905	0,889
	(1 2,5)	1,000	0,999	0,999	0,998	0,991	0,988
15	(1 1)	0,051	0,051	0,049	0,050	0,050	0,054
	(1 1,2)	0,399	0,388	0,340	0,328	0,270	0,261
	(1 1,5)	0,923	0,909	0,865	0,844	0,785	0,762
	(1 2)	0,999	0,999	0,996	0,994	0,984	0,977
	(1 2,5)	1,000	1,000	1,000	1,000	1,000	0,999

Tablo 3. Testlerin I. tip hata oranları ve güç değerleri

n	μ	r=0,2		r=0,5		r=0,9	
		T	MW	T	MW	T	MW
3	(1 1)	0,048	0,098	0,053	0,102	0,048	0,102
	(1 2)	0,055	0,109	0,067	0,125	0,134	0,231
	(1 3)	0,115	0,205	0,157	0,260	0,549	0,716
	(1 4)	0,319	0,470	0,467	0,626	0,981	0,996
	(1 5)	0,596	0,744	0,794	0,891	1,000	1,000
5	(1 1)	0,054	0,060	0,051	0,057	0,054	0,059
	(1 2)	0,075	0,082	0,088	0,088	0,241	0,247
	(1 3)	0,193	0,202	0,292	0,300	0,878	0,865
	(1 4)	0,584	0,580	0,787	0,785	1,000	1,000
	(1 5)	0,912	0,901	0,985	0,979	1,000	1,000
10	(1 1)	0,054	0,056	0,052	0,055	0,046	0,048
	(1 2)	0,100	0,104	0,132	0,134	0,480	0,466
	(1 3)	0,386	0,382	0,559	0,550	0,998	0,997
	(1 4)	0,914	0,902	0,989	0,986	1,000	1,000
	(1 5)	0,999	0,998	1,000	1,000	1,000	1,000
15	(1 1)	0,054	0,055	0,053	0,049	0,056	0,059
	(1 2)	0,137	0,135	0,181	0,178	0,662	0,636
	(1 3)	0,545	0,526	0,754	0,728	1,000	1,000
	(1 4)	0,985	0,980	1,000	0,999	1,000	1,000
	(1 5)	1,000	1,000	1,000	1,000	1,000	1,000

Testlerin güç değerlerine bakıldığında ise MW testinin dağılımı asimptotik olarak standart normal dağılıma sahip olduğu için küçük örnek çaplarında testin deneysel I. tip hata oranları nominal α değerinden oldukça yüksek çıktığı gözlenmiştir. Bundan dolayı küçük örnek çapları durumunda ($n=3$) MW testin güç değerleri t-testinin güç değerlerine göre daha yüksek çıkmıştır. n değerleri arttıkça testlerin güç değerlerini karşılaştırdığında da genel olarak çok az farkla t testinin güç değerlerinin MW testinin güç değerlerinden daha yüksek çıktığı gözlenmiştir. Genel olarak testlerin güç değerlerini incelediğimizde de beklenildiği gibi n 'lerin artışı ve

işlem ortalamaları arasındaki fark arttıkça testlerin güç değerlerini arttırmıştır. Ayrıca tablodan r 'lerin artışının testlerin güç değerlerini olumlu yönde etkileyerek arttırdığı görülmektedir. $\alpha=0.05$ ve normal olmayan dağılım yani üstel dağılım ($f(x) = ae^{-ax}$, $x \geq 0$), kaydırılmış üstel dağılım ($f(x) = ae^{-a(x-b)}$, $x \geq b$) ve standart çifte üstel dağılım ($f(x) = e^{-\frac{|x-\mu|}{a}}/2a$) altında aktarılmış etkinin olmadığına dair yokluk hipotezinin testi için testlerinin I. tip hata oranları ve güç değerleri aşağıdaki Tablo 4, 5 ve 6'da verilmiştir.

Tablo 4. r=0.2 için testlerin deneysel I. tip hata oranları ve güç değerleri

n	a	λ	Üstel		Kaydırılmış üstel		Standart çifte üstel	
			T	MW	T	MW	T	MW
3	1	(1 1)	0.054	0.098	0.044	0.098	0.042	0.103
		(1 2)	0.191	0.311	0.112	0.201	0.075	0.147
		(1 4)	0.908	0.972	0.567	0.714	0.322	0.474
	3	(1 1)	0.054	0.107	0.048	0.102	0.044	0.102
		(1 2)	0.073	0.131	0.054	0.108	0.051	0.110
		(1 4)	0.198	0.310	0.104	0.203	0.076	0.152
5	1	(1 1)	0.048	0.050	0.051	0.065	0.047	0.059
		(1 2)	0.373	0.367	0.182	0.205	0.115	0.132
		(1 4)	0.999	0.999	0.828	0.795	0.567	0.569
	3	(1 1)	0.054	0.058	0.041	0.050	0.045	0.055
		(1 2)	0.085	0.083	0.057	0.071	0.054	0.067
		(1 4)	0.376	0.378	0.187	0.213	0.111	0.122
10	1	(1 1)	0.052	0.053	0.049	0.054	0.051	0.055
		(1 2)	0.720	0.682	0.337	0.388	0.204	0.218
		(1 4)	1.000	1.000	0.986	0.990	0.876	0.884
	3	(1 1)	0.050	0.055	0.053	0.061	0.049	0.050
		(1 2)	0.129	0.124	0.083	0.094	0.063	0.074
		(1 4)	0.710	0.680	0.344	0.392	0.183	0.200
15	1	(1 1)	0.055	0.054	0.044	0.048	0.054	0.052
		(1 2)	0.883	0.856	0.474	0.539	0.257	0.271
		(1 4)	1.000	1.000	0.998	0.999	0.970	0.976
	3	(1 1)	0.048	0.047	0.043	0.046	0.052	0.060
		(1 2)	0.182	0.169	0.108	0.120	0.073	0.075
		(1 4)	0.885	0.856	0.482	0.553	0.257	0.277

Tablo 5. r=0.5 için testlerin deneysel I. tip hata oranları ve güç değerleri

n	a	λ	Üstel		Kaydırılmış üstel		Standart çifte üstel	
			T	MW	T	MW	T	MW
3	1	(1 1)	0.047	0.088	0.050	0.106	0.045	0.103
		(1 2)	0.180	0.281	0.104	0.190	0.072	0.138
		(1 4)	0.847	0.932	0.476	0.634	0.258	0.411
	3	(1 1)	0.051	0.098	0.046	0.106	0.041	0.096
		(1 2)	0.062	0.117	0.054	0.115	0.052	0.109
		(1 4)	0.184	0.285	0.105	0.191	0.072	0.142
5	1	(1 1)	0.051	0.057	0.047	0.064	0.052	0.064
		(1 2)	0.337	0.340	0.161	0.180	0.106	0.119
		(1 4)	0.997	0.996	0.776	0.759	0.495	0.500
	3	(1 1)	0.049	0.054	0.048	0.054	0.041	0.052
		(1 2)	0.079	0.085	0.055	0.066	0.047	0.060
		(1 4)	0.325	0.336	0.153	0.180	0.103	0.114
10	1	(1 1)	0.051	0.053	0.043	0.048	0.055	0.061
		(1 2)	0.626	0.598	0.296	0.322	0.164	0.166
		(1 4)	1.000	1.000	0.971	0.977	0.807	0.817
	3	(1 1)	0.049	0.051	0.046	0.050	0.046	0.050
		(1 2)	0.113	0.111	0.077	0.081	0.061	0.063
		(1 4)	0.637	0.605	0.286	0.329	0.170	0.178
15	1	(1 1)	0.049	0.050	0.047	0.047	0.051	0.051
		(1 2)	0.823	0.792	0.399	0.449	0.214	0.228
		(1 4)	1.000	1.000	0.997	0.998	0.943	0.946
	3	(1 1)	0.050	0.051	0.047	0.054	0.050	0.049
		(1 2)	0.167	0.155	0.090	0.094	0.075	0.077
		(1 4)	0.828	0.797	0.420	0.464	0.226	0.231

Testlerin deneysel I. tip hata oranlarına bakıldığında tüm n değerleri için t testinin deneysel I. tip hata oranlarının nominal α değerine oldukça yakın olduğu gözlenmektedir. MW testinin deneysel I. tip hata oranları küçük örnek çaplarında nominal α değerinden oldukça yüksek çıkarken n değeri arttıkça

nominal α değerine yaklaştığı görülmektedir. Bundan dolayı küçük örnek çapları durumunda testin gücü diğer teste göre daha yüksek çıkmıştır. Fakat n arttıkça testin I. tip hata oranı α 'ya yakın olduğu gözlenmiştir. Güç değerleri karşılaştırıldığında genel olarak çok az farkla t testinin güç değerleri MW

Tablo 6. $r=0.9$ için testlerin deneysel I. tip hata oranları ve güç değerleri

n	a	λ	Üstel		Kaydırılmış üstel		Standart çifte üstel	
			T	MW	T	MW	T	MW
3	1	(1 1)	0.057	0.100	0.047	0.104	0.041	0.098
		(1 2)	0.132	0.214	0.089	0.179	0.065	0.134
		(1 4)	0.676	0.842	0.426	0.584	0.220	0.365
	3	(1 1)	0.053	0.096	0.044	0.100	0.041	0.097
		(1 2)	0.061	0.108	0.044	0.108	0.043	0.109
		(1 4)	0.117	0.203	0.090	0.183	0.064	0.139
5	1	(1 1)	0.055	0.059	0.044	0.061	0.046	0.057
		(1 2)	0.228	0.233	0.138	0.170	0.084	0.098
		(1 4)	0.978	0.968	0.667	0.658	0.394	0.402
	3	(1 1)	0.054	0.055	0.046	0.062	0.046	0.060
		(1 2)	0.069	0.073	0.058	0.072	0.046	0.058
		(1 4)	0.225	0.231	0.137	0.176	0.082	0.098
10	1	(1 1)	0.052	0.054	0.044	0.052	0.057	0.057
		(1 2)	0.467	0.438	0.234	0.300	0.128	0.147
		(1 4)	1.000	1.000	0.898	0.942	0.680	0.724
	3	(1 1)	0.049	0.053	0.041	0.049	0.044	0.048
		(1 2)	0.097	0.100	0.070	0.089	0.059	0.066
		(1 4)	0.481	0.454	0.230	0.293	0.136	0.155
15	1	(1 1)	0.051	0.052	0.049	0.051	0.052	0.053
		(1 2)	0.669	0.611	0.313	0.424	0.181	0.207
		(1 4)	1.000	1.000	0.973	0.992	0.836	0.882
	3	(1 1)	0.051	0.052	0.047	0.047	0.052	0.053
		(1 2)	0.121	0.114	0.076	0.095	0.062	0.064
		(1 4)	0.662	0.611	0.324	0.423	0.176	0.200

testinin güç değerlerinden daha yüksek olduğu gözlenmiştir. Ayrıca farklı dağılımlar altında testlerin güç değerlerinin etkilendiği görülmüştür. Testlerin güç değerleri üstel dağılım altında daha yüksekken, standart çifte üstel dağılım altında testlerin güç değerleri biraz daha düşüktür. Üstel dağılımda t testinin gücü biraz daha yüksek iken kaydırılmış üstel ve standart çifte üstel dağılımlarında MW testinin gücünün biraz daha yüksek olduğu gözlenmiştir. Ayrıca dağılımların parametre değeri olan a ve r arttıkça testlerin güç değerlerinde de azalma olduğu görülmektedir.

$\alpha=0.05$ ve normal olmayan dağılım (üstel dağılım, kaydırılmış üstel dağılım ve standart çifte üstel dağılım) altında işlem etkisi testlerinin I. tip hata oranları ve güç değerleri aşağıdaki Tablo 7, 8 ve 9'da verilmiştir.

Genel olarak tablolara bakıldığında testlerin güç değerlerinin farklı dağılımlar altında bundan etkilendiği görülmüştür. Üstel dağılımda t testinin gücü biraz daha yüksek iken kaydırılmış üstel ve standart çifte üstel dağılımlarında MW testinin gücünün biraz daha yüksek olduğu gözlenmiştir. Ayrıca dağılımların parametre değeri olan a arttıkça testlerin güç değerlerinde de azalma olduğu görülmektedir. r'lerin artışının da testlerin güç değerlerini azalttığı gözlenmektedir.

4. Tartışma ve Sonuç

Çapraz tasarımların analizinde dikkate alınan en önemli etkilerden biri, bir periyotta uygulanan

işlemin deney birimi üzerindeki etkisinin, takip eden periyotta da devam eden etkisi, yani aktarılmış etkidir. Eğer bir periyotta elde edilen gözlem değeri üzerinde aktarılmış etki varsa, o periyotta uygulanan işlem etkisinin belirlenmesi imkânsız hale gelir, hesaplanırsa bile yanlış sonuçlar verir. Bunun için aktarılmış etkiler ihmal edilmemeli ve kesinlikle analiz edilmelidir. Bu yüzden de, çapraz tasarımların analizinde işlem etkisinin olup olmadığının testi için ilk olarak aktarılmış etkinin olup olmadığının testi yapılır. Eğer aktarılmış etkinin varlığı söz konusu ise o zaman sadece ilk periyodun verileri kullanılarak işlem etkisinin olup olmadığı test edilir. Aktarılmış etkinin yokluğu durumunda ise her iki periyoddaki veriler kullanılarak işlem etkisinin olup olmadığı test edilir.

Çapraz tasarımların analizinde; aktarılmış etki ve işlem etkilerinin olup olmadığının test edilmesinde, parametrik olan t testi ve parametrik olmayan MW U testi kullanılmaktadır. Çalışmada bu testler simülasyon yoluyla deneysel I. tip hata oranı ve güç bakımından normal ve normal olmayan dağılımlar altında karşılaştırmaları yapılmıştır. Buna göre işlem etkisinin ve aktarılmış etkinin olup olmadığının testinde normal ve normal olmayan dağılım altında testlerin deneysel I. tip hata oranlarına bakıldığında tüm n ve r değerleri için t testinin deneysel I. tip hata oranlarının nominal α değerine oldukça yakın olduğu gözlenmektedir. MW testinin deneysel I. tip hata oranları ise tüm r değerlerinde, n=3 iken nominal α değerinden oldukça yüksek çıkarken n değeri arttıkça nominal α değerine yaklaştığı görülmektedir.

Testlerin güç değerleri incelendiğinde ise, normal dağılım altında t testinin güç değerleri daha yüksek iken, normal olmayan dağılım altında (kaydırılmış

üstel dağılım, standart çifte üstel dağılım) MW U testinin güç değerlerinin daha yüksek olduğu gözlenmiştir.

Tablo 7. $r=0.2$ için testlerin deneysel I. tip hata oranları ve güç değerleri

n	a	μ	Üstel		Kaydırılmış üstel		Standart çifte üstel	
			T	MW	T	MW	T	MW
3	1	(1 1)	0.047	0.096	0.043	0.104	0.042	0.098
		(1 2)	0.613	0.766	0.352	0.511	0.189	0.315
		(1 4)	1.000	1.000	0.948	0.971	0.802	0.883
	3	(1 1)	0.054	0.098	0.040	0.101	0.042	0.096
		(1 2)	0.122	0.206	0.077	0.169	0.060	0.124
		(1 4)	0.616	0.767	0.356	0.510	0.179	0.310
5	1	(1 1)	0.058	0.056	0.043	0.055	0.048	0.059
		(1 2)	0.941	0.936	0.574	0.570	0.333	0.342
		(1 4)	1.000	1.000	0.997	0.991	0.976	0.959
	3	(1 1)	0.052	0.052	0.047	0.060	0.046	0.056
		(1 2)	0.197	0.199	0.113	0.134	0.076	0.087
		(1 4)	0.938	0.931	0.579	0.575	0.340	0.350
10	1	(1 1)	0.049	0.051	0.046	0.051	0.048	0.052
		(1 2)	1.000	0.999	0.853	0.896	0.587	0.616
		(1 4)	1.000	1.000	1.000	1.000	1.000	1.000
	3	(1 1)	0.057	0.060	0.048	0.055	0.040	0.049
		(1 2)	0.423	0.396	0.191	0.230	0.118	0.132
		(1 4)	1.000	0.999	0.848	0.886	0.599	0.623
15	1	(1 1)	0.049	0.052	0.050	0.053	0.044	0.049
		(1 2)	1.000	1.000	0.960	0.980	0.781	0.813
		(1 4)	1.000	1.000	1.000	1.000	1.000	1.000
	3	(1 1)	0.049	0.051	0.050	0.054	0.046	0.048
		(1 2)	0.578	0.536	0.272	0.340	0.146	0.156
		(1 4)	1.000	1.000	0.961	0.977	0.772	0.802

Tablo 8. $r=0.5$ için testlerin deneysel I. tip hata oranları ve güç değerleri

n	a	μ	Üstel		Kaydırılmış üstel		Standart çifte üstel	
			T	MW	T	MW	T	MW
3	1	(1 1)	0.050	0.094	0.043	0.101	0.041	0.101
		(1 2)	0.728	0.863	0.420	0.577	0.230	0.369
		(1 4)	1.000	1.000	0.969	0.985	0.871	0.932
	3	(1 1)	0.059	0.103	0.042	0.106	0.042	0.100
		(1 2)	0.140	0.239	0.087	0.185	0.064	0.133
		(1 4)	0.730	0.863	0.428	0.583	0.229	0.369
5	1	(1 1)	0.055	0.059	0.044	0.056	0.045	0.054
		(1 2)	0.981	0.975	0.667	0.655	0.402	0.411
		(1 4)	1.000	1.000	0.999	0.996	0.991	0.982
	3	(1 1)	0.052	0.057	0.045	0.056	0.047	0.055
		(1 2)	0.243	0.245	0.138	0.158	0.086	0.098
		(1 4)	0.983	0.979	0.671	0.658	0.396	0.414
10	1	(1 1)	0.054	0.056	0.050	0.055	0.052	0.053
		(1 2)	1.000	1.000	0.917	0.933	0.702	0.726
		(1 4)	1.000	1.000	1.000	1.000	1.000	1.000
	3	(1 1)	0.054	0.053	0.045	0.054	0.046	0.051
		(1 2)	0.506	0.474	0.246	0.292	0.132	0.140
		(1 4)	1.000	1.000	0.917	0.941	0.692	0.717
15	1	(1 1)	0.050	0.048	0.046	0.051	0.048	0.050
		(1 2)	1.000	1.000	0.984	0.993	0.861	0.881
		(1 4)	1.000	1.000	1.000	1.000	1.000	1.000
	3	(1 1)	0.045	0.049	0.045	0.049	0.050	0.052
		(1 2)	0.695	0.653	0.317	0.393	0.177	0.190
		(1 4)	1.000	1.000	0.983	0.993	0.860	0.885

Tablo 9. $r=0.9$ için testlerin deneysel I. tip hata oranları ve güç değerleri

n	a	μ	Üstel		Kaydırılmış üstel		Standart çifte üstel	
			T	MW	T	MW	T	MW
3	1	(1 1)	0.051	0.099	0.042	0.100	0.044	0.100
		(1 2)	1.000	1.000	0.837	0.912	0.606	0.750
		(1 4)	1.000	1.000	1.000	1.000	0.999	0.999
	3	(1 1)	0.062	0.106	0.041	0.103	0.043	0.103
		(1 2)	0.398	0.560	0.224	0.381	0.127	0.237
		(1 4)	1.000	1.000	0.837	0.911	0.615	0.750
5	1	(1 1)	0.055	0.057	0.040	0.053	0.047	0.054
		(1 2)	1.000	1.000	0.977	0.960	0.878	0.854
		(1 4)	1.000	1.000	1.000	1.000	1.000	1.000
	3	(1 1)	0.046	0.053	0.046	0.057	0.043	0.053
		(1 2)	0.740	0.740	0.368	0.385	0.212	0.230
		(1 4)	1.000	1.000	0.980	0.963	0.884	0.863
10	1	(1 1)	0.050	0.050	0.050	0.056	0.053	0.058
		(1 2)	1.000	1.000	1.000	1.000	0.997	0.998
		(1 4)	1.000	1.000	1.000	1.000	1.000	1.000
	3	(1 1)	0.050	0.053	0.042	0.047	0.048	0.051
		(1 2)	0.980	0.968	0.655	0.715	0.394	0.421
		(1 4)	1.000	1.000	1.000	1.000	0.996	0.997
15	1	(1 1)	0.049	0.050	0.052	0.053	0.047	0.050
		(1 2)	1.000	1.000	1.000	1.000	1.000	1.000
		(1 4)	1.000	1.000	1.000	1.000	1.000	1.000
	3	(1 1)	0.053	0.053	0.050	0.055	0.045	0.051
		(1 2)	1.000	0.998	0.812	0.876	0.534	0.571
		(1 4)	1.000	1.000	1.000	1.000	1.000	1.000

Kaynakça

- [1] Armaneri, Ö., Armaneri, S., 2006. Crossover tasarımların klinik araştırmalarda uygulaması, Fen ve Mühendislik Dergisi, 8(3):51-64.
- [2] McNair, D. M., 1971. Antianxiety drugs and human performance, Archives of General Psychiatry, 29(5), 611-617.
- [3] Jones, B., Kenward, M.G., 1989. Design and analysis of cross-over trial. USA: Chapman and Hall, 1-89.
- [4] Yılmaz, L., Bayrak, H., Gökınar, F., 2016. İki işlem üçperiyot dual dengeli tasarımların nesne düşüşlerine sağlamlığının değerlendirilmesi, SDU, Fen bilimleri Enstitüsü Dergisi, 20(1), 1-7.
- [5] Bandyopadhyay, U., Chatterjee, S., 2015. Nonparametric analysis of the two-period two treatment crossover design, Journal of Nonparametric Statistics, 27(1), 127-148.
- [6] Basu, J., Bandyopadhyay, U., Dutta, G., 2017. Nonparametric testing under crossover design for ordered categorical response, Journal of the Korean Statistical Society.
- [7] Jones, B., Kenward, M.G., 2014. Design and analysis of cross-over trials, CRC press.
- [8] Cleophas, T.J.M., Zwinderman, A.H., 2002. Crossover studies with continuous variables: power analysis, American Journal of Therapeutics, 9, 69-73.
- [9] Yıldırım, D.D., Taşdelen, B., 2013. AB/BA çapraz geçişli denemelerinde tip I hata oranı ile ilişkili faktörler, Türkiye Klinikleri J Biostat, 5(2), 61-68.
- [10] Uriarte, R.D., 2002. Incorrect analysis of crossover trials in animal behavior research, Animal Behaviour, 63, 815-822.
- [11] Elbourne, D.R., Altman, D.G., Higgins, J.P.T., Curtin, F., Worthington, H.V., Vail, A., 2002. Meta-analyses involving cross-over trials: methodological issues, International Journal of Epidemiology, 31, 140-149.
- [12] Hills, M., Armitage, P., 1979. The two-period cross-over clinical trial, British Journal of Clinical Pharmacology, 8, 7-20.
- [13] Brown, B.W., 1980. The crossover experiment for clinical trials, Biometrics, 36(1), 69-79.
- [14] Bellavance, F., Tardif, S., 1995. A nonparametric approach to the analysis of three-treatment three-period crossover design, Biometrika, 82(4), 865-875.
- [15] Öhrvik, J., 1998. Nonparametric methods in crossover trials, Biometrical Journal, 40(7), 771-789.
- [16] Putt, M.E., Chinchilli, V.M., 2004. Nonparametric approaches to the analysis of crossover studies, Statistical Science, 19(4), 712-719.
- [17] Cleophas, T.J., 1999. Clinical trials: relevance of correlation between treatment responses, Clin Research & Reg Affairs, 16(4), 193-204.

- [18] Senn, S., 2002. Cross-over trials in clinical research, second edition. England: John Wiley & Sons, 1-157.
- [19] Grizzle, J.E., 1965. The two-period change-over design and its use in clinical trials, *Biometrics*, 2, 469-480.
- [20] Reed, J.F., 2004. Analysis of two-treatment, two-period crossover trials in emergency medicine, *Annals of Emergency Medicine*, 43(1), 54-58.
- [21] Toutenburg, H., 2002. Statistical analysis of designed experiments, second edition, New York: Springer, 341-384.
- [22] Gacula, M. C., 1993. Design and Analysis of Sensory Optimization, USA: Food and Nutrition Press, 45-46.