



**T.C.
ULUDAĞ ÜNİVERSİTESİ
SAĞLIK BİLİMLERİ ENSTİTÜSÜ
BİYOİSTATİSTİK ANABİLİM DALI**

**ÇOKLU KARŞILAŞTIRMA TESTLERİNİN GRUP SAYISI VE DENEK
SAYILARINA GÖRE KARŞILAŞTIRILMASI**

Güven ÖZKAYA

(DOKTORA TEZİ)

Bursa-2011



T.C.
ULUDAĞ ÜNİVERSİTESİ
SAĞLIK BİLİMLERİ ENSTİTÜSÜ
BİYOİSTATİSTİK ANABİLİM DALI

ÇOKLU KARŞILAŞTIRMA TESTLERİNİN GRUP SAYISI VE DENEK SAYILARINA
GÖRE KARŞILAŞTIRILMASI

Güven ÖZKAYA

(DOKTORA TEZİ)

Danışman: Doç. Dr. İlker ERCAN

Bursa-2011

Sağlık Bilimleri Enstitüsü Müdürlüğüne,

Bu tez, jürimiz tarafından doktora tezi olarak kabul edilmiştir.

Adı ve Soyadı

İmza

Tez Danışmanı Doç.Dr. İlker ERCAN

Üye Prof.Dr. Ahmet DİRİCAN

Üye Doç.Dr. Berna YAZICI

Üye Doç.Dr. Nural BEKİROĞLU

Üye Yrd.Doç.Dr. Bülent EDİZ

Bu tez, Enstitü Yönetim Kurulunun tarih, sayılı toplantısında alınan numaralı kararı ile kabul edilmiştir.

Prof.Dr. Gürsel SÖNMEZ
UÜ. Sağlık Bilimleri Enstitü Müdürü

İÇİNDEKİLER

TABLolar DİZİNİ	I
ÖZET	V
SUMMARY	VI
1.GİRİŞ	1
2.GENEL BİLGİLER	3
2.1.Tek Yönlü Varyans Analizi	3
2.2. Sabit Etkili Modelin Analizi	3
2.3.Çoklu Karşılaştırma Testleri	4
2.4. Ortalama Çiftlerinin Karşılaştırılmasında Kullanılan Çoklu Karşılaştırma Testleri...7	
2.4.1. Fisher'in En Küçük Anlamlı Fark Yöntemi (LSD)	7
2.4.2. Bonferroni ve Dunn-Sidak Testleri	8
2.4.3. Scheffe Testi.....	10
2.4.4. Student-Newman-Keuls Testi (SNK)	10
2.4.5. Tukey'in a testi.....	12
2.4.6. Tukey'in b testi	13
2.4.7. Ryan, Einot, Gabriel ve Welsch'in Metodu	13
2.4.8. Duncan Testi	15
2.4.9. Hochberg's GT2 Testi	16
2.4.10. Gabriel'in Testi.....	17
2.5. Testlerin Karşılaştırılmasında Kullanılacak Kriterler	17
2.5.1. Deneysel Hata Oranı (Experimentwise Error Rate)	19
2.5.2. Karşılaştırma Başına Hata Oranı (Comparisonwise Error Rate)	19
2.5.3. Yanlış Bulgu Oranı (False Discovery Rate)	19
2.5.4. İstatistiksel Güç Tanımları	21
3. GEREÇ VE YÖNTEM	22
4. BULGULAR.....	27
4.1. Deneysel Hata Oranı ve Karşılaştırma Başına Hata Oranı Sonuçları	27
4.2. İstatistiksel Güç ve Yanlış Bulgu Oranı Bakımından Karşılaştırılması	40
5. TARTIŞMA VE SONUÇ.....	65
6. KAYNAKLAR	73
7. TEŞEKKÜR	76
8. ÖZGEÇMİŞ.....	77

TABLolar DİZİNİ

Tablo-1: Tek yönlü varyans analiz tablosu	3
Tablo-2: $\alpha=0.05$ düzeyinde gruplararası karşılaştırmalarda q istatistiği için kritik değerler	13
Tablo-3: İstatistiksel hata tablosu	18
Tablo-4: M tane H_0 hipotezinde hesaplanan hata sayısı	20
Tablo-5: Örneklem genişliklerinin eşit olduğu durumda Tip I hatanın incelendiği simülasyon senaryoları ($\mu=40$)	23
Tablo-6: Örneklem genişliklerinin eşit olmadığı durumda Tip I hatanın incelendiği simülasyon senaryoları ($\mu=40$)	24
Tablo-7: Örneklem genişliklerinin eşit olduğu durumunda testlerin gücünün karşılaştırılmasında kullanılan simülasyon senaryoları	25
Tablo-8: Örneklem genişliklerinin eşit olmadığı durumunda testlerin gücünün karşılaştırılmasında kullanılan simülasyon senaryoları	26
Tablo-9: Grup sayısının 3 ve örneklem büyüklüklerinin eşit olduğu durum için deneysel hata oranı	28
Tablo-10: Grup sayısının 5 ve örneklem büyüklüklerinin eşit olduğu durum için deneysel hata oranı	29
Tablo-11: Grup sayısının 7 ve örneklem büyüklüklerinin eşit olduğu durum için deneysel hata oranı	30
Tablo-12: Grup sayısının 3 ve örneklem büyüklüklerinin farklı olduğu durum için deneysel hata oranı	31
Tablo-13: Grup sayısının 5 ve örneklem büyüklüklerinin farklı olduğu durum için deneysel hata oranı	32
Tablo-14: Grup sayısının 7 ve örneklem büyüklüklerinin farklı olduğu durum için deneysel hata oranı	33
Tablo-15: Grup sayısının 3 ve örneklem büyüklüklerinin eşit olduğu durum için karşılaştırma başına hata oranı	34
Tablo-16: Grup sayısının 5 ve örneklem büyüklüklerinin eşit olduğu durum için karşılaştırma başına hata oranı	35
Tablo-17: Grup sayısının 7 ve örneklem büyüklüklerinin eşit olduğu durum için karşılaştırma başına hata oranı	36
Tablo-18: Grup sayısının 3 ve örneklem büyüklüklerinin farklı olduğu durum için karşılaştırma başına hata oranı	37

Tablo-19: Grup sayısının 5 ve örneklem büyüklüklerinin farklı olduğu durum için karşılaştırma başına hata oranı	38
Tablo-20: Grup sayısının 7 ve örneklem büyüklüklerinin farklı olduğu durum için karşılaştırma başına hata oranı	39
Tablo-21: Grup sayısının 3 ve örneklem büyüklüklerinin eşit olduğu durum için her bir çift için güç değerleri	41
Tablo-22: Grup sayısının 5 ve örneklem büyüklüklerinin eşit olduğu durum için her bir çift için güç değerleri	42
Tablo-23: Grup sayısının 7 ve örneklem büyüklüklerinin eşit olduğu durum için her bir çift için güç değerleri	43
Tablo-24: Grup sayısının 3 ve örneklem büyüklüklerinin farklı olduğu durum için her bir çift için güç değerleri	44
Tablo-25: Grup sayısının 5 ve örneklem büyüklüklerinin farklı olduğu durum için her bir çift için güç değerleri	45
Tablo-26: Grup sayısının 7 ve örneklem büyüklüklerinin farklı olduğu durum için her bir çift için güç değerleri	46
Tablo-27: Grup sayısının 3 ve örneklem büyüklüklerinin eşit olduğu durum için herhangi bir çift için güç değerleri	47
Tablo-28: Grup sayısının 5 ve örneklem büyüklüklerinin eşit olduğu durum için herhangi bir çift için güç değerleri	48
Tablo-29: Grup sayısının 7 ve örneklem büyüklüklerinin eşit olduğu durum için herhangi bir çift için güç değerleri	49
Tablo-30: Grup sayısının 3 ve örneklem büyüklüklerinin farklı olduğu durum için herhangi bir çift için güç değerleri	50
Tablo-31: Grup sayısının 5 ve örneklem büyüklüklerinin farklı olduğu durum için herhangi bir çift için güç değerleri	51
Tablo-32: Grup sayısının 7 ve örneklem büyüklüklerinin farklı olduğu durum için herhangi bir çift için güç değerleri	52
Tablo-33: Grup sayısının 3 ve örneklem büyüklüklerinin eşit olduğu durum için bütün çift için güç değerleri	53
Tablo-34: Grup sayısının 5 ve örneklem büyüklüklerinin eşit olduğu durum için bütün çift için güç değerleri	54
Tablo-35: Grup sayısının 7 ve örneklem büyüklüklerinin eşit olduğu durum için bütün çift için güç değerleri	55

Tablo-36: Grup sayısının 3 ve örneklem büyüklüklerinin farklı olduğu durum için bütün çift için güç değerleri	56
Tablo-37: Grup sayısının 5 ve örneklem büyüklüklerinin farklı olduğu durum için bütün çift için güç değerleri	57
Tablo-38: Grup sayısının 7 ve örneklem büyüklüklerinin farklı olduğu durum için bütün çift için güç değerleri	58
Tablo-39: Grup sayısının 3 ve örneklem büyüklüklerinin eşit olduğu durum için yanlış bulgu oranı.....	59
Tablo-40: Grup sayısının 5 ve örneklem büyüklüklerinin eşit olduğu durum için yanlış bulgu oranı.....	60
Tablo-41: Grup sayısının 7 ve örneklem büyüklüklerinin eşit olduğu durum için yanlış bulgu oranı.....	61
Tablo-42: Grup sayısının 3 ve örneklem büyüklüklerinin farklı olduğu durum için yanlış bulgu oranı.....	62
Tablo-43: Grup sayısının 5 ve örneklem büyüklüklerinin farklı olduğu durum için yanlış bulgu oranı.....	63
Tablo-44: Grup sayısının 7 ve örneklem büyüklüklerinin farklı olduğu durum için yanlış bulgu oranı.....	64

ÖZET

Çoklu karşılaştırma testleri, varyans analizinin gerçekleştirilmesi sonrasında anlamlı olan grup veya grupları belirlemede kullanılan testlerdir. Literatürde birçok çoklu karşılaştırma testi yer almaktadır. Çalışmamızda yaygın olarak kullanılan ve istatistiksel yazılımlarda yer alan 12 çoklu karşılaştırma testi belirlenmiştir. Bu testlerin temellendiği istatistikler ve dayandığı varsayımlar farklılık göstermektedir. Farklı grup sayıları, homojen varyans yapısı, eşit ve farklı örneklem büyüklükleri için simülasyon çalışması gerçekleştirilmiştir. Testlerin birbiriyle karşılaştırılmasında deneysel hata oranı, karşılaştırma başına hata oranı, istatistiksel güç ve yanlış bulgu oranı dikkate alınmıştır. Simülasyon sonucunda LSD ve Duncan testlerinin hata oranları yüksek bulunurken Scheffe testinin düşük bulunmuştur. Buna bağlı olarak LSD ve Duncan testlerinin gücü yüksek iken Scheffe testinin düşük bulunmuştur. Çalışma sonucunda LSD, Duncan ve Scheffe testleri önerilmemektedir.

Anahtar Kelimeler: Çoklu Karşılaştırma Testleri, Tip I Hata, Güç, Yanlış Bulgu Oranı

SUMMARY

**COMPARISON OF MULTIPLE COMPARISON TESTS ACCORDING TO
NUMBER OF GROUPS AND SUBJECTS**

Multiple comparison tests are used after variance analysis to determine the different group or groups. Many multiple comparison tests are available in the literature. In our study we compare twelve multiple comparison tests which are so popular and available in statistical softwares. These tests are different according to statistical theory and assumptions. Simulation is performed according to different number of groups, homogeneity of variance, equal and different sample size. These tests are compared according to experimentwise error rate, comparisonwise error rate, statistical power and false discovery rate. In simulation result, the error rates of LSD and Duncan tests are higher than others. On the other hand the error rates of Scheffe is lower than others. Depending on this result, the power LSD and Duncan tests are higher than others. On the other hand the power of Scheffe is lower than others. LSD, Duncan and Scheffe tests are not suggested in our study.

Keywords: Multiple Comparison Tests, Type I Error, Power, False Discovery Rate

1.GİRİŞ

Deneysel arařtırmalarda, karřılařtırmalı alıřmalar olduka yaygın Őekilde gerekleřtirilir. Varyans analizi, bu alıřmalarda en sık kullanılan yntemlerden biridir. Sir Ronald A. Fisher tarafından varyans analizi adı verilen ve geliřtirilen bu yntem 1950’li yıllardan sonra arařtırmacılar tarafından sıklıkla kullanılmaya bařlanmıřtır (1). Varyans analizi, bağımsız deęiřkenlerin bağımlı deęiřken zerindeki etkisini incelemek amacıyla kullanılır.

Varyans analizindeki problem, gruplar arasındaki karřılařtırmalardır. Analiz sonucunda farklılık bulunduęunu belirtmek yeterli olmayabilir. Ayrıca farklılık yaratan grup veya gruplar belirlenmek istenebilir. Varyans analizindeki F testi bu sorulara doęrudan cevap veremez. oklu karřılařtırmalar olarak bilinen test prosedleri bu sorulara cevap vermek iin geliřtirilmiřtir.

İkinci Dnya Savařını takip eden on yılda zellikle Amerika Birleřik Devletleri’nde oklu karřılařtırmalara olan ilgide ok byk bir artıř gzlemlenmiřtir. Bu sre boyunca, tek deęiřkenli durumlar iin David B. Duncan, Charles W. Dunnett, HO. Hartley, Henry Scheffe, SN. Roy ve John W. Tukey’in nemli katkıları olmuřtur (2). Ancak bunlar Fisher, Student ve dięerleri tarafından daha nce yapılan alıřmalardan temellenmektedir (3). Tukey’in, 1953 yılında hazırladıęı zengin kaynaklardan oluřan sonuları ieren notları sınırlı kiřiye ulařmıřtır (3). Miller’in, 1966 yılında yayımladıęı kitap oklu karřılařtırma testlerini popler hale getirmiř ve yeni bir arařtırma alanına vurgu yapmıřtır (3). 1970’li yıllardan sonra bu alanda birok makale yayımlandıęı grlmektedir.

oklu karřılařtırma testleri,  veya daha fazla grup olduęunda farklılıęı yaratan grup veya grupları bulmak iin uygulanan testlerdir ve sıklıkla kullanılmaktadır. Ancak kullanılan oklu karřılařtırma testleri oęunlukla yanlış kullanılmaktadır (4-5). oklu karřılařtırma testlerinde hangi testin seileceęi nemli bir problemdir.

nceki yıllarda yapılan alıřmalarda oklu karřılařtırma testlerinden bazıları n plana ıkmıř ve nerilmiř iken bazı testlerin ise uygulanmasının sakıncalı olduęu ifade edilmiřtir. rneęin, Ryan (6), Einot ve Gabriel (7), Student-Newman-Keuls ve Duncan’ın testlerini nermemiřlerdir. Son yıllarda yapılan bir alıřmada oklu karřılařtırma testlerinden Duncan, Student-Newman-Keuls, Tukey ve Dunnett testlerinin Bonferroni ve Sidak testlerinden daha sıklıkla kullanıldıęı belirtilmiřtir (8).

Çoklu karşılaştırmalarda kullanılan testlerin temellendiği istatistikler ve dayandığı varsayımlar farklılık göstermektedir (9). Grup varyanslarının homojen olup olmaması ve grupların özelliklerine göre özel çoklu karşılaştırma testleri seçilebilir. Çoklu karşılaştırma testlerinin seçimini ve etkinliklerini inceleyen birçok çalışma yapılmıştır. Bununla birlikte yeni yöntem önerilmesi ve bilgisayar teknolojisinin zamanla gelişmesi beraberinde simülasyon kapasitesindeki artış çalışmaların yenilenmesini sağlamaktadır.

Çalışmamızda çoklu karşılaştırma testleri, varyansların homojen, örneklem büyüklüklerinin eşit ve farklı olduğu durumlarda, çeşitli grup sayıları için simülasyon yardımıyla incelenmesi amaçlanmıştır. Testler birbirleriyle kıyaslanırken deneysel hata oranı, karşılaştırma başına hata oranı, güç ve yanlış bulgu oranı gibi istatistiksel kriterler göz önüne alınmıştır. Çalışmamızın sonucunda varyans, örneklem büyüklüğü ve grup sayılarının aldığı değerlere göre istatistiksel kriterler bakımından incelenerek testlerin hangi durumlarda kullanılması gerektiği tartışılacaktır.

2.GENEL BİLGİLER

2.1.Tek Yönlü Varyans Analizi

k tane grubu veya tek bir faktörün farklı düzeylerini karşılaştırmak istediğimizi varsayalım. k gruptan elde edilen yanıtlar raslantı değişkenidir. Bu düzen Tablo-1'de gösterilmiştir. Tablo-1'de

y_{ij} : i.grupta yer alan j.deneğe ait gözlem,

y_i : i.gruba ait gözlemlerin toplamı,

\bar{y}_i : i.gruba ait gözlemlerin ortalamasını göstermektedir.

Tablo-1: Tek yönlü varyans analiz tablosu

Gözlemler						
Grup	1	2	...	n	Toplam	Ortalama
1	y_{11}	y_{12}	...	y_{1n}	$y_{1.}$	\bar{y}_1
2	y_{21}	y_{22}	...	y_{2n}	$y_{2.}$	\bar{y}_2
⋮	⋮	⋮	...	⋮	⋮	⋮
k	y_{k1}	y_{k2}	...	y_{kn}	$y_{k.}$	\bar{y}_k
Genel					$y_{..}$	$\bar{y}_{..}$

2.2. Sabit Etkili Modelin Analizi

Yukarıda belirtilen veriler ile bir model oluşturulmak istenirse bu model

$$y_{ij} = \mu_i + \varepsilon_{ij} \begin{cases} i = 1, 2, \dots, k \\ j = 1, 2, \dots, n \end{cases} \quad (1)$$

şeklinde olur (10). y_{ij} , i.grupta yer alan j.deneğe ait gözlemi, μ_i i.grubun veya faktör düzeyinin ortalamasını ve ε_{ij} rastgele hatayı göstermektedir. Denklem (1) ortalama model olarak bilinir (10). Bu modelin belirtilen veri için alternatif gösterimi aşağıdaki gibi tanımlanabilir:

$$\mu_i = \mu + \tau_i \quad i=1, 2, \dots, k \quad (2)$$

Böylece denklem (1) aşağıdaki gibi oluşur.

$$y_{ij} = \mu + \tau_i + \varepsilon_{ij} \begin{cases} i = 1, 2, \dots, k \\ j = 1, 2, \dots, n \end{cases} \quad (3)$$

Bu model formunda, μ bütün gruplar için ortak olan genel ortalama parametresi olarak adlandırılır. τ_i , i.grubun etkisini gösteren parametredir. Hem ortalama modeli hem de etki modeli doğrusal istatistiksel modellerdir. Yanıt değişkeni y_{ij} , model parametrelerinin doğrusal fonksiyonudur. Her iki modelin de kullanışlı olmasına rağmen etki modelinin kullanımı daha yaygındır. Denklem (3), ayrıca tek faktörlü varyans analizi olarak da bilinir. Hipotez testi için modelin hata terimi sıfır ortalamalı ve σ^2 varyanslı normal dağılım göstermeli ve bağımsız olmalıdır. Faktörün bütün düzeyleri için σ^2 varyansının sabit olduğu varsayılır. Gözlemler $y_{ij} \sim N(\mu + \tau_i, \sigma^2)$ dağılım gösterir ve birbirinden bağımsızdır. Faktör düzeylerinin seçimine göre incelenen model farklı isimler alır. Faktör düzeyleri araştırmacı tarafından özel olarak seçildiyse bu model sabit etkili model adını alır. Eğer araştırmacı faktör düzeylerinin bulunduğu anakütleden rastgele olarak seçerse oluşturulan model rastgele model olarak bilinir.

Sabit etkili modelin analizinde kullanılan hipotez aşağıdaki gibidir:

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

$$H_1: \mu_i \neq \mu_j \text{ En az bir çift } (i, j) \text{ için}$$

Ancak bu hipotezin test edilebilmesi için öncelikle tek yönlü varyans analizinin varsayımını olan varyansların homojenliği hipotezinin test edilmesi gereklidir (10).

$$H_0: \sigma_1^2 = \sigma_2^2 = \dots = \sigma_k^2 = 0$$

$$H_1: \text{ (En az bir grup için)}$$

Hipotezin kabul edilmesi durumunda sabit etkili model için hipotez testi gerçekleştirilebilir. Bu hipotez testi sonucunda istatistiksel anlamlılık bulunursa çoklu karşılaştırma testleri kullanılır.

2.3.Çoklu Karşılaştırma Testleri

Tek yönlü varyans analizinde gerçekleştirilen genel test sonucunda anlamlılığı yaratan grup veya grupları bulmak için çoklu karşılaştırma testlerinin kullanılması gerekir.

Çoklu karşılaştırma testinin seçimi deney sonuçlarını önemli ölçüde etkileyebilmektedir. Örneğin, birçok çoklu karşılaştırma testi normal dağılım ve homojen varyans varsayımlarının sağlanmadığı durumda kullanılması uygun değildir (11).

Birçok çoklu karşılaştırma testi, bağıntıları kullanmaktadır (10). k ortalamının karşılaştırılması, ortalamalar içerisinde ikisinin karşılaştırılması şeklinde olacağı gibi bir ortalamının geriye kalan ortalamalarla karşılaştırılması şeklinde de olabilir. Bu tanım, sembollerle aşağıdaki gibi ifade edilebilir:

$$\psi = c_1\mu_1 + c_2\mu_2 + \dots + c_k\mu_k = \sum_{j=1}^k c_j\mu_j \quad (4)$$

Ortalamaların (μ_k) önünde yer alan ağırlıkların (c_k) toplamı sıfırdır. Fakat bütün c_k 'ler sıfıra eşit olmayabilir. Yukarıda verilen tanım çoklu karşılaştırmanın anakütle ortalamaları (μ_k) türündendir. Örneklem için aşağıdaki tanım kullanılır.

$$\hat{\psi} = c_1\bar{x}_1 + c_2\bar{x}_2 + \dots + c_k\bar{x}_k = \sum_{j=1}^k c_j\bar{x}_j \quad (5)$$

Örneğin karşılaştırılacak grup sayısı $k=3$ olsun. Çoklu karşılaştırma gösterimleri için aşağıdaki örnekler verilebilir.

Birinci değişkenin ortalaması ile ikinci değişkenin ortalaması arasındaki farkı incelemek için;

$$\hat{\psi}_1 = \bar{x}_1 - \bar{x}_2 = (1)\bar{x}_1 + (-1)\bar{x}_2 + (0)\bar{x}_3 \quad (6)$$

Birinci değişkenin ortalaması ile üçüncü değişkenin ortalaması arasındaki fark,

$$\hat{\psi}_2 = \bar{x}_1 - \bar{x}_3 = (1)\bar{x}_1 + (0)\bar{x}_2 + (-1)\bar{x}_3 \quad (7)$$

Birinci değişkenin ortalaması ile diğer iki değişkenin ortalamalarının genel ortalaması arasındaki fark

$$\hat{\psi}_3 = \bar{x}_1 - \frac{\bar{x}_2 + \bar{x}_3}{2} = (1)\bar{x}_1 + (-1/2)\bar{x}_2 + (-1/2)\bar{x}_3 \quad (8)$$

şeklinde ifade edilebilir.

Ağırlıklar incelendiğinde, ilk eşitlik için ağırlık katsayıları $c_1=1$, $c_2=-1$ ve $c_3=0$ şeklindedir. İkinci eşitlik için ağırlık katsayıları $c_1=1$, $c_2=0$ ve $c_3=-1$ şeklindedir. Son karşılaştırma için ise $c_1=1$, $c_2=-1/2$ ve $c_3=-1/2$ şeklindedir.

Bağıntı türleri 4 kategoride incelenebilir (12).

- a) Ortalama çiftleri için karşılaştırma
- b) Kompleks karşılaştırma
- c) Kontrol ile karşılaştırma
- d) En iyi ile karşılaştırma

Ortalama çiftleri için yapılan karşılaştırmalarda iki ortalamanın ağırlık katsayıları 1 ve -1 olarak alınmaktadır (Örneğin, 7 no'lu eşitlikte olduğu gibi). Kompleks karşılaştırmalarda kullanılan değişken sayısı 2'den fazladır ve ağırlık katsayıları kesirli değerlerden oluşabilir (Örneğin, 8 no'lu eşitlikte olduğu gibi). Kontrol ile yapılan karşılaştırmalarda k tane grup mevcut iken bunlardan bir tanesi kontrol grubudur. k-1 grubun ortalaması kontrol grubunun ortalaması ile karşılaştırılır. Uygulamada ilgilenilen temel soru hangi grubun en iyi olduğu veya hangi grubun ortalaması en yüksek veya en düşük olduğudur. Bu durumda yapılacak olan karşılaştırma türü en iyi ile karşılaştırmadır.

Çoklu karşılaştırma testlerinin sınıflandırılması testlerin özelliklerine göre birkaç kategoride gerçekleştirilebilir. Örneğin, kurulan sıfır hipoteze bağlı olarak sadece ortalamaların ikili olarak test edildiği çoklu karşılaştırma testleri ve bunun dışında yer alan hipotezler olarak ikiye ayrılır. Bu ayrım ortalama çifti olan ve ortalama çifti olmayan çoklu karşılaştırma testleri olarak adlandırılabilir.

Ayrıca hipotezler “eşzamanlı” ve “ardışık” olarak incelenebilir. Bir eş zamanlı çoklu karşılaştırma testinde, tek yönlü varyans analizindeki F testi veya herhangi bir karşılaştırma testinin anlamlılık durumuna bakılmaksızın sabit bir kritik değer kullanılarak bütün karşılaştırmalar yapılmaktadır (11, 13). Bu prosedürler eş zamanlı test prosedürleri olarak da bilinir. Ardışık çoklu karşılaştırma prosedürleri belirli bir karşılaştırmaların anlamlılığını test ederken hem tek yönlü varyans analizindeki F testinin hem de diğer karşılaştırmaların anlamlılığını dikkate alır. İkili karşılaştırma yapılabilmesi için anlamlı bir F testi gerektiren çoklu karşılaştırma prosedürleri tutucu testler olarak adlandırılır (11, 13).

Ayrıca çoklu karşılaştırma testleri incelenecek hipotezlerin planlanmış olup olmamasına göre de sınıflandırılabilir. Eğer analiz öncesinde farklılığı incelenmek istenen ortalama grupları var ise bunun için yapılacak çoklu karşılaştırma testi “planlanmış” olarak adlandırılır. Eğer analiz sonuçları incelenerek çoklu karşılaştırma testleri gerçekleştirilirse bu testler “planlanmamış” çoklu karşılaştırma testleri olarak adlandırılır (13).

Çoklu karşılaştırma testleri, bağıntının ortogonal olup olmamasına göre iki ayrı kategoride sınıflandırılabilir. Oluşturulan bağıntının katsayılarının çarpımının toplamı sıfır ise ortogondur. Eğer bu çarpımın toplamı sıfırdan farklı ise ortogonal değildir.

Ortogonalliğin tanımı için iki bağıntının katsayıları c_i ve d_i olsun. Eğer $\sum_{i=1}^k c_i d_i = 0$ koşulu sağlanırsa iki bağıntı ortogondur. Bağıntıların birbirilerine ortogonal olması istenir. Ancak bütün ortalamaların ikili olarak karşılaştırılması istenirse ortogonallik sağlanamaz (13).

2.4. Ortalama Çiftlerinin Karşılaştırılmasında Kullanılan Çoklu Karşılaştırma Testleri

Çoklu karşılaştırma testleri, bağıntı türlerine bağlı olarak 4 kategoride incelenebilir.

- a) Ortalama çiftlerinin karşılaştırılması,
- b) Kompleks karşılaştırma,
- c) Kontrol grubu ile karşılaştırma,
- d) En iyi ile karşılaştırma,

Bu kategorilerden, ortalama çiftlerinin karşılaştırılmasında kullanılan testler incelenmiştir. Ortalama çiftlerinin karşılaştırılması için istatistiksel yazılımlar ve literatür incelendiğinde yaygın olarak kullanılan testler aşağıdaki gibi belirlenmiştir.

- En Küçük Anlamlı Fark
- Bonferroni
- Dunn-Sidak
- Scheffe
- REGW - F
- REGW - Q
- Student Neuman Keuls
- Tukey'in a
- Tukey'in b
- Duncan
- Hochberg'in GT2
- Gabriel

2.4.1. Fisher'in En Küçük Anlamlı Fark Yöntemi (LSD)

RA. Fisher, 1935 yılında k anakütle ortalamasının ikili karşılaştırılmasında kullanılan bir prosedür geliştirmiştir (2). Bu prosedür Fisher'in en küçük anlamlı fark yöntemi olarak adlandırılır. Varyans analizi sonucunda anlamlı bulunan F testi sonrasında önerildiğinden, "tutucu en küçük anlamlı fark testi" olarak da adlandırılmaktadır (14). Fisher'in en küçük anlamlı fark yöntemiyle yapılan ortalamaların ikili karşılaştırılmasında t-testi kullanılır (1). Bu test iki adımda gerçekleştirilir.

- a. Öncelikle gruplar arasında farklılık olup olmadığına karar vermek için α düzeyinde genel F testi gerçekleştirilir.
- b. İlk adım anlamlı ise gruplar arasında α düzeyinde t testi kullanılarak ikili karşılaştırmalar gerçekleştirilir.

Grupların örneklem büyüklükleri eşit olmadığı durumda Fisher'in en küçük anlamlı fark yönteminde kullanılan test istatistiği aşağıdaki gibidir:

Test istatistiği:

$$t = \frac{\bar{y}_i - \bar{y}_j}{\sqrt{\text{HKO} \left(\frac{1}{n_i} + \frac{1}{n_j} \right)}} \quad (9)$$

\bar{y}_i : i.grubun ortalaması

\bar{y}_j : j.grubun ortalaması

n_i : i.grubun örneklem büyüklüğü

n_j : j.grubun örneklem büyüklüğü

HKO: Hata kareler ortalaması

Eğer $|t| > t_{\frac{\alpha}{2}; N-k}$ ise sıfır hipotezi reddedilir. Diğer bir gösterim;

$$|\bar{y}_i - \bar{y}_j| \geq t_{\frac{\alpha}{2}; N-k} \sqrt{\text{HKO} \left(\frac{1}{n_i} + \frac{1}{n_j} \right)} \text{ ise } H_0 \text{ ret.} \quad (10)$$

Grupların örneklem büyüklüğü eşit olması durumunda ise test istatistiği aşağıdaki gibidir.

$$|\bar{y}_i - \bar{y}_j| \geq t_{\frac{\alpha}{2}; N-k} \sqrt{\frac{2\text{HKO}}{n}} \text{ ise } H_0 \text{ ret.} \quad (11)$$

2.4.2. Bonferroni ve Dunn-Sidak Testleri

Bu test eğer k tane H_0 hipotezi varsa, hedeflenen genel tip I hata oranını en çok α olacak şekilde belirlemek amacıyla her bir H_0 hipotezi α/k düzeyinde test eder. Buna eş olarak, eğer k güven aralığının her biri %100(1- α/k) güven düzeyinde hesaplanırsa bunların her biri eş zamanlı olarak en azından %100(1- α) güven düzeyini sağlamaktadır.

Her bir H_0 hipotezinin α^* düzeyinde test edildiğini varsayalım. E_i , i. H_0 hipotezinin reddilmesi olayını gösterebilir. Tip I hatanın genel olasılığı aşağıdaki gibi verilebilir (1).

$$\begin{aligned} \alpha &= P \left[E_1 \cup E_2 \cup \dots \cup E_k \right] \\ &\leq P \left[E_1 \right] + P \left[E_2 \right] + \dots + P \left[E_k \right] \\ &= k\alpha^* \end{aligned} \quad (12)$$

Böylece eğer k tane H_0 hipotezinden her biri α/k düzeyinde reddedilirse genel tip I hata oranı en fazla α olur. Bu prosedür Bonferroni metodu olarak adlandırılır. Çünkü Bonferroni

veya Boole eşitsizliğine dayanmaktadır. Bazı literatürlerde Dunn'ın çoklu karşılaştırma prosedürü olarak adlandırılır. Dunn, 1961 yılında bu prosedürün özelliklerini incelemiş ve metodun kullanımı için tablolar hazırlamıştır (1).

Bonferroni veya Dunn-Sidak testine göre verilen k adet karşılaştırmanın her biri α^* düzeyinde test edilir. Genel hata oranı (α), $k\alpha^*$ aşmaz. α 'nın küçük değerleri için üst sınıra mükemmel bir yaklaşım sağlar (1).

Fakat üst sınıra daha iyi bir yaklaşım Sidak tarafından verilen çarpımsal eşitsizlik tarafından elde edilir. Bu eşitsizlik aşağıdaki gibidir (1).

$$\alpha \leq 1 - (1 - \alpha^*)^k \leq k\alpha^* \quad (13)$$

Böylece her bir testi α/k anlamlılık düzeyinde test etmek yerine her bir karşılaştırma $1 - (1 - \alpha)^{1/k}$ anlamlılık düzeyinde test edilir.

Bonferroni ve Dunn-Sidak testleri tip I hatanın bölünmesi olarak belirtilebilir. Eğer k adet karşılaştırma belirli bir α düzeyinde yapılacaksa olası en çok genel hata oranı $k\alpha$ olur. Bu sebeple her bir karşılaştırma için düzey α/k olarak belirlenmelidir. Böylece istenen genel hata oranı α 'ya eşit olacaktır. Bonferroni düzeltmesinin hesaplanmasında Fisher'in en küçük anlamlı fark yönteminden yararlanır.

Test istatistiği:

$$t = \frac{\bar{y}_i - \bar{y}_j}{\sqrt{\text{HKO} \left(\frac{1}{n_i} + \frac{1}{n_j} \right)}} \quad (14)$$

\bar{y}_i : i.grubun ortalaması

\bar{y}_j : j.grubun ortalaması

n_i : i.grubun örneklem büyüklüğü

n_j : j.grubun örneklem büyüklüğü

HKO: Hata kareler ortalaması

Eğer $|t| > t_{\frac{\alpha^*}{2}; N-k}$ ise H_0 hipotezi reddedilir.

Ancak hipotezin sonucuna karar verilirken karşılaştırma sayısına bağlı olarak uygun $\alpha^* = \alpha/k$ düzeyi belirlenmelidir.

Dunn-Sidak testinde bu düzeltme için $\alpha^* = 1 - (1 - \alpha)^{1/k}$ formülü kullanılır.

2.4.3. Scheffe Testi

Scheffe testi doğrusal bağıntıları kullanır. Ancak bunların ortogonal olması gerekmez. Gerçekte bütün akla uygun doğrusal bağıntılar önemlilik için test edilebilir. İki ortalamayı karşılaştırırken özel bağıntılar da kullanılabilir (15). Scheffe testinde diğer testlerle önemli bulunan farklar önemsiz bulunabilir. Bu testte aşağıdaki adımlar izlenir:

- Denemede ilgilenilen tüm doğrusal bağıntılar alınır ve sayısal değerleri bulunur.

İlgilenilen doğrusal bağıntı aşağıdaki gibi ifade edilir:

$$C_m = c_{1m}\bar{y}_1 + c_{2m}\bar{y}_2 + \dots + c_{km}\bar{y}_k \quad m=1,2,\dots,u \quad (15)$$

(m: Bağıntı sayısı; k: Grup sayısı; u: Maksimum bağıntı sayısı)

- (k-1) ve (N-k) serbestlik derecesinde α güven düzeyinde F değeri saptanır.
- Bir önceki adımdaki F değeri kullanılarak $A = \sqrt{(k-1)F}$ değeri elde edilir.
- Test edilecek her bir doğrusal bağıntının standart hatası bulunur. Bu standart hata formülü

$$s_{C_m} = \sqrt{HKO \sum_i (c_{im}^2 / n_i)} \quad (16)$$

- C_m doğrusal bağıntısı, sayısal olarak As_{C_m} 'den daha büyük ise bağıntı anlamlıdır ya da ise doğrusal bağıntının sıfır olduğu hipotezi reddedilir.

2.4.4. Student-Newman-Keuls Testi (SNK)

Bu test ilk olarak Newman tarafından 1939 geliştirilmiştir (2). Newman, bir öğrencisinin önerisine uyarak standartlaştırılmış aralığa dayanan ilk çoklu karşılaştırma testini bulmuştur. Keuls, 1952 yılında Newman'ın testini biraz değiştirerek, sentez olan Newman-Keuls testini önermiştir (16). Bu test aynı zamanda Newman tarafından öğrencisinin önerisi dikkate alındığı için Student-Newman-Keuls test (SNK) olarak da bilinmektedir. SNK testi için aşağıda sıralanan adımlar izlenir:

- k sayıdaki ortalamalar küçükten büyüğe doğru sıralanır.
- Varyans çözümlemesi tablosundan hata kareler ortalaması ile serbestlik dereceleri alınır.
- Her bir deneme için ortalamaların standart hataları bulunur.

Örnekleme büyüklükleri eşit ise;

$$s_{\bar{y}_i} = \sqrt{\frac{HKO}{n}} \quad (17)$$

Örnekleme büyüklükleri eşit değilse;

$$s_{\bar{y}_i} = \sqrt{\frac{HKO}{n_h}} \quad (18)$$

$$n_h = \frac{k}{\sum_{i=1}^k \frac{1}{n_i}} \quad (19)$$

- d. HKO, serbestlik derecesi v ve k ortalamaların küçükten büyüğe doğru sıralanması sonrasında karşılaştırılmak istenen sıralı ortalama çiftlerinin içerdiği ortalama sayısı p olmak üzere α düzeyindeki standartlaştırılmış aralık istatistiğinin kritik değeri $q_\alpha(p, v)$ şeklinde gösterilir. Bu değer standartlaştırılmış aralık tablosundan elde edilir. Standartlaştırılmış aralık istatistiğinin kritik değerleri $p=2, 3, 4, \dots, k$ kullanılarak $k-1$ genişlik sıralanır.
- e. Hesaplanan kritik değerler ile $S_{\bar{y}_i}$ çarpılarak her bir p değeri için K_p ($K_p = q_\alpha(p, v) S_{\bar{y}_i}$) değeri bulunur.
- f. En büyük ortalama değerden başlayarak sıra ile bu değerle diğerleri arasındaki farklar bulunur. Bu farklar $k-1$ tanedir. Bu işlem $k(k-1)/2$ sayıda karşılaştırma oluncaya dek sürer. Bu farklar K_p değeri ile karşılaştırılarak ortalama çiftleri arasında farklılık olup olmadığı hipotezi test edilir.

Test istatistiği:

$$|\bar{y}_{(k)} - \bar{y}_{(1)}| > K_p \text{ ise } H_0 \text{ hipotezi reddedilir.}$$

$\bar{y}_{(k)}$: k grubun ortalamaları büyükten küçüğe doğru sıralandığında en büyük ortalama değeri gösterir.

$\bar{y}_{(1)}$: k grubun ortalamaları büyükten küçüğe doğru sıralandığında en küçük ortalama değeri gösterir.

$$K_p = q_\alpha(p, v) S_{\bar{y}_i}$$

2.4.5. Tukey'in a testi

Tukey'in 1953 yılında hazırladığı yayınlanmamış notlarında önerdiği çoklu karşılaştırma testi q istatistiğinin yani "standartlaştırılmış aralığa" dayalıdır (3). Tukey'in kullandığı standartlaştırılmış aralık istatistiği aşağıdaki gibidir:

$$q = \frac{\bar{y}_{\max} - \bar{y}_{\min}}{\sqrt{\text{HKO}/n}} \quad (20)$$

\bar{y}_{\max} : En büyük örneklem ortalaması

\bar{y}_{\min} : En küçük örneklem ortalaması

HKO: Hata kareler ortalaması

n: Örneklem büyüklüğü

HKO'nun serbestlik derecesi v ve karşılaştırılacak örneklem sayısı p olmak üzere α düzeyindeki standartlaştırılmış aralık istatistiğinin kiritik değeri $q_{\alpha}(p, v)$ şeklinde gösterilir. Bu değer standartlaştırılmış aralık tablosundan elde edilir. Tukey test istatistiği, örneklem büyüklüklerinin eşit olma ve olmama durumlarına göre aşağıdaki gibi hesaplanır.

$n_i=n_j$ olduğu durumda;

$$T_{\alpha} = q_{\alpha}(p, v) \sqrt{\frac{\text{HKO}}{n}} \quad (21)$$

$n_i \neq n_j$ olduğu durumda;

$$T_{\alpha} = \frac{q_{\alpha}(p, v)}{\sqrt{2}} \sqrt{\text{HKO} \left(\frac{1}{n_i} + \frac{1}{n_j} \right)} \quad (22)$$

Örneklem ortalamaları arasındaki farkın mutlak değeri hesaplanan test istatistiğinden büyük ise ortalamalar arasında fark olduğu yorumu yapılır. Örneklem büyüklüğünün eşit olmadığı durumda kullanılan eşitlik aynı zamanda Tukey-Kramer testi olarak da bilinmektedir (10, 17).

2.4.6. Tukey'in b testi

Tukey a testi, H_0 hipotezinin reddedilmesinde karşılaştırmalar boyunca aynı kritik değeri kullanır. Buna karşın SNK testinin kritik değeri kısmi H_0 hipotezi içerisinde yer alan grup sayısına bağlı olarak değişir. Tablo-2'de de görüldüğü gibi Tukey b testi, Tukey a ve SNK testinin kritik değerlerinin ortalamasını kullanır (18).

Tablo-2: $\alpha=0.05$ düzeyinde gruplararası karşılaştırmalarda q istatistiği için kritik değerler

Test	Grup Sayısı			
	5	4	3	2
Tukey'in a	3.92	3.92	3.92	3.92
Tukey'in b	3.92	3.80	3.64	3.36
SNK	3.92	3.68	3.36	2.80

2.4.7. Ryan, Einot, Gabriel ve Welsch'in Metodu

Deneysel tip I hata oranını sınırlayan ve düzeltilmiş anlamlılık seviyesine dayanan adimsal çoklu karşılaştırma prosedürü ilk olarak Ryan tarafından 1960 yılında önerilmiştir (19). SNK testinin modifikasyonu olarak da düşünülebilir (13). Maksimum hata oranını α düzeyinde sabitlemek için Ryan, k gruba ait ortalamaların karşılaştırılmasında p tane sıralı ortalama için aşağıdaki düzeltmeyi önermiştir. Karşılaştırılmak istenen sıralı ortalamalar ile bu ortalamalar arasında kalan grup sayısı toplamı p genişliğini vermektedir. Buna göre p genişliğindeki testler için α_p aşağıdaki gibi hesaplanır:

$$\alpha_p = \left(\frac{p}{k} \right) \alpha \quad (23)$$

Einot ve Gabriel (7), 1975 yılında Ryan prosedürünü biraz daha geliştirerek daha güçlü bir yöntem önermiştir (19). Buna göre;

$$\alpha_p = 1 - (1 - \alpha)^{p/k} \quad (24)$$

Welsch (20-21) ise yine Ryan'ın metoduna bir modifikasyon yapmış ve Ryan'ın metodundan biraz daha güçlü bir metot elde etmiştir. Welsch'e göre:

$$\alpha_{k-1} = \alpha \quad (25)$$

olarak alınmaktadır(19).

Ryan, Einot ve Gabriel ve Welsch'in metodu ise bu üç metodun bir kombinasyonu şeklindedir ve q istatistiğine dayalıdır (REGW-Q). REGW-Q testi için α_p değerleri aşağıdaki gibi kullanılır:

$$\alpha_p = \begin{cases} 1 - (1 - \alpha)^{p/k} & p < k - 1 \\ \alpha & p \geq k - 1 \end{cases} \quad (26)$$

REGW-Q testi için test istatistiği aşağıdaki gibidir:

Ortalama çiftleri arasında farklılık olup olmadığının hipotez testinde

$$\frac{|\bar{y}_i - \bar{y}_j|}{\sqrt{\text{HKO} \left(\frac{1}{n_i} + \frac{1}{n_j} \right)}} \geq \frac{q_{\alpha_p}(p, v)}{\sqrt{2}} \quad (27)$$

koşulu sağlanırsa H_0 hipotezi reddedilir.

\bar{y}_i : i. örneklemin ortalaması

\bar{y}_j : j. örneklemin ortalaması

HKO: Hata kareler ortalaması

v: HKO'nun serbestlik derecesi

p: k ortalamanın küçükten büyüğe doğru sıralanması sonrasında karşılaştırılmak istenen sıralı ortalama çiftlerinin içerdiği ortalama sayısı

$q_{\alpha_p}(p, v)$: p ve v parametreleriyle standartlaştırılmış aralık istatistiği

Q istatistiğine bağlı bu testin bir başka şekli F istatistiğine bağlı olarak geliştirilmiştir.

Hesaplanan test istatistiği q istatistiği yerine F istatistiği ile karşılaştırılır.

REGW-F testi için aşağıdaki adımlar izlenir:

a) Ortalama çiftleri küçükten büyüğe doğru sıralanır.

b) En büyük ortalama ile en küçük ortalama çifti arasında fark olup olmadığı aşağıdaki gibi test edilir. Anlamlı fark bulunmaması durumunda işlem sonlandırılır ve tüm ortalamalar arasında fark olmadığı yorumu yapılır. Fark bulunması halinde en küçük ortalama ile ikinci en büyük ortalama arasında fark olup olmadığı ve en büyük ortalama ile ikinci en küçük ortalama arasında fark olup olmadığı test edilir. İşlem bu şekilde devam eder.

$$\frac{\left(\sum_{i \in R} n_i \bar{y}_i^2 - \left(\sum_{i \in R} n_i \bar{y}_i \right)^2 / \sum_{i \in R} n_i \right)}{(p-1) \text{ HKO}} \geq F_{\alpha_p; p-1; v} \quad R = \{i, \dots, j\} \quad (28)$$

\bar{y}_i : i. örneklemin ortalaması

n_i : i. Örneklem büyüklüğü

HKO: Hata kareler ortalaması

p: k ortalamamın küçükten büyüğe doğru sıralanması sonrasında karşılaştırılmak istenen sıralı ortalama çiftlerinin içerdiği ortalama sayısı

α_p : REGW Q yönteminden elde edilen değer

v: HKO'nun serbestlik derecesi

$F_{\alpha_p; p-1; v}$: α_p düzeyinde p ve v serbestlik dereceleriyle F istatistiği

2.4.8. Duncan Testi

Duncan (22) tarafından 1955 yılında bütün ortalama çiftlerinin karşılaştırıldığı bir test geliştirilmiştir. Duncan'ın testi standartlaştırılmış aralık istatistiğini kullanan çoklu aralık testidir. Bu sebeple Duncan çoklu aralık testi olarak da adlandırılmaktadır. Örneklem büyüklükleri eşit olan k adet ortalamaya Duncan çoklu aralık testi uygulanmasında, ortalamalar artan bir şekilde sıralanır ve her bir ortalamaya ait standart hatalar hesaplanır.

$$s_{\bar{y}_i} = \sqrt{\frac{\text{HKO}}{n}} \quad (29)$$

Eğer örneklem büyüklükleri eşit değilse 26 no'lu denklemde n yerine bütün n_i 'lerin harmonik ortalaması n_h ile yer değiştirecektir.

$$n_h = \frac{k}{\sum_{i=1}^k \frac{1}{n_i}} \quad (30)$$

Duncan'ın anlamlı aralık tablosundan $p=2,3,\dots,k$ için $r_\alpha(p, v)$ değeri elde edilir. α , değeri anlamlılık düzeyi ve v ise hatanın serbestlik derecesidir. Tablodan elde edilen bu değerler aşağıdaki hesaplamada kullanılarak k-1 adet en küçük anlamlı aralıklar kümesine çevrilir.

$$R_p = r_\alpha(p, v) s_{\bar{y}_i} \quad p=2,3,\dots,k \text{ için} \quad (31)$$

Gözlenen en büyük ortalama ile en küçük ortalama arasındaki farkı R_k en küçük anlamlı aralık değeri ile karşılaştırılır. Bir sonraki adımda en büyük ortalama ile ikinci en küçük ortalama arasındaki fark R_{k-1} ile karşılaştırılır. Bu karşılaştırmalar diğer ortalamaların en büyük ortalama ile karşılaştırılması yapılanaya kadar sürer. Daha sonra ikinci en büyük ortalama ile diğer ortalamalar arasındaki fark R_{k-1} en küçük aralık değeri ile karşılaştırılır. Bu süreç bütün olası $k(k-1)/2$ adet karşılaştırma yapılan dek sürer. Eğer gözlenen fark ona karşılık gelen en küçük anlamlı fark değerinden büyükse ortalama çiftleri arasında anlamlı fark olduğu yorumu yapılır.

2.4.9. Hochberg's GT2 Testi

Hochberg, standartlaştırılmış maksimum modülüse dayanan bir yöntem önermiştir (23). Bu testin yürütülmesinde aşağıdaki adımlar izlenir:

- Ortalama çiftleri küçükten büyüğe doğru sıralanır.
- En büyük ortalama ile en küçük ortalama çifti arasında fark olup olmadığı aşağıdaki gibi test edilir. Anlamlı fark bulunmaması durumunda işlem sonlandırılır. Tüm ortalamalar arasında fark olmadığı yorumu yapılır. Fark bulunması halinde en küçük ortalama ile ikinci en büyük ortalama arasında fark olup olmadığı ve en büyük ortalama ile ikinci en küçük ortalama arasında fark olup olmadığı test edilir. Bu işlem böylece devam eder.

Test istatistiği:

$$|\bar{y}_{(k)} - \bar{y}_{(1)}| > \sqrt{\frac{2 \text{ HKO}}{n}} m(\alpha; \binom{k}{2}, \nu) \quad (32)$$

koşulu sağlanırsa ortalamalar arasında fark olduğu şeklinde yorumlanır.

$\bar{y}_{(k)}$: k grubun ortalamaları büyükten küçüğe doğru sıralandığında en büyük ortalama değeri gösterir.

$\bar{y}_{(1)}$: k grubun ortalamaları büyükten küçüğe doğru sıralandığında en küçük ortalama değeri gösterir.

$m(\alpha; \binom{k}{2}, \nu)$: α düzeyinde, ν serbestlik derecesi ve $\binom{k}{2}$ karşılaştırma sayısı olmak üzere standartlaştırılmış maksimum modülüs dağılım değeri

Grupların örneklem büyüklüklerinin birbirinden farklı olması durumunda harmonik ortalama kullanılır.

n_h : k tane grubun örneklem büyüklüklerinin harmonik ortalamasıdır.

$$n_h = \frac{k}{\sum_{1 \leq i \leq k} n_i^{-1}} \quad (33)$$

2.4.10. Gabriel'in Testi

Gabriel (24), 1978 yılında Hochberg gibi standartlaştırılmış maksimum modülüse dayanan bir yöntem önermiştir. Bu yöntem sadece aritmetik ortalamalar için uygundur. Eğer aşağıdaki koşul sağlanırsa iki ortalama arasında farklılık olduğu söylenir (25).

Test istatistiği:

Ortalama çiftleri arasında farklılık olup olmadığının hipotez testinde

$$|\bar{y}_i - \bar{y}_j| \geq m(\alpha; \binom{k}{2}, \nu) \sqrt{\text{HKO} \left(\frac{1}{\sqrt{2n_i}} + \frac{1}{\sqrt{2n_j}} \right)} \quad (34)$$

ise H_0 hipotezi reddedilir.

\bar{y}_i : i. örneklemin ortalaması

\bar{y}_j : j. örneklemin ortalaması

HKO: Hata kareler ortalaması

ν : HKO'nun serbestlik derecesi

$m(\alpha; \binom{k}{2}, \nu)$: α düzeyinde, ν serbestlik derecesi ve $\binom{k}{2}$ karşılaştırma sayısı olmak üzere

standartlaştırılmış maksimum modülüs dağılım değeri

n_i : i. örneklem büyüklüğü

n_j : j. örneklem büyüklüğü

2.5. Testlerin Karşılaştırılmasında Kullanılacak Kriterler

Çoklu karşılaştırma testleri kullanılırken çeşitli konular göz önünde bulundurulmalıdır (11). Bu konular,

- Anlamlılık testlerinin incelenmesindeki çeşitlilik etkisi
- Uygun düzeyde anlamlılık derecesi seçilmesi
- Uygun çoklu karşılaştırma testinin seçilmesi.

Araştırmanın hedefleri bu kararları yönlendirmelidir (11). Bu kararları vermek durumunda olan araştırmacılar çoğunlukla geleneksel seçimlere yönelmişlerdir (11). Yapılan bir araştırmaya göre eğitim ve psikoloji alanında araştırma yapan kişilerin istatistiksel uygulamaları üzerine, mevcut birçok çoklu karşılaştırma yönteminden Tukey ve Scheffé metotlarının en çok tercih edildiği belirlenmiştir (26). Çoklu karşılaştırma testlerinin kullanımından kaynaklanan istatistiksel problem, testin gerçekleştirilmesinden sonra aynı veride alt hipotez testlerinin gerçekleştirilmesidir. Bu işlem kontrol edilemeyen tip I hata ile sonuçlanabilir.

Bir hipotez testinde H_0 hipotezini kabul ya da reddederken iki tür hata yapılabilir. Bu hatalardan birincisi tip I hata olarak tanımlanan H_0 hipotezi gerçekten doğru iken bu hipotezin yanlışlıkla reddedilmesidir ve α ile gösterilir. Diğer hata türü ise tip II hata olarak tanımlanan H_0 hipotezi gerçekte yanlış iken kabul edilmesidir ve β ile gösterilir (Tablo-3). α ve β türü hatalar arasında ilişki vardır ve α azalırken β artar. Ancak α sabit tutulurken denek sayısının artması β türü hata olasılığını azaltır.

Tablo-3: İstatistiksel hata tablosu

Hipotez testi sonucundaki durum	Gerçek Durum	
	H_0 hipotezi gerçekte doğru ise	H_0 hipotezi gerçekte yanlış ise
H_0 kabul	Doğru Karar	Tip II Hata (β)
H_0 ret	Tip I Hata (α)	Doğru Karar

Çoklu karşılaştırma testlerinin anlamlılıkları hesaplanırken tip I hatanın nasıl kontrol edileceğine karar vermek oldukça zordur.

İstatistiksel çıkarsamada çeşitlilik problemi, araştırma hipotezini desteklemek veya reddetmek amacıyla büyük sonuçlar kümesinden istatistiksel olarak anlamlı sonuçların seçilmesine karşılık gelmektedir (11). İstatistiksel olarak anlamlı sonuçların, anlamlı olmayan sonuçların yer aldığı büyük bir sonuç havuzundan seçilmesi problemdir. Çünkü çoklu karşılaştırma testlerinin anlamlılıkları hesaplanırken test sayısı arttıkça en az birinin anlamlı olma olasılığı artar.

Değişken sayısının çok olması durumunda kurulacak çeşitli hipotezlerin tek değişkenli yöntemlerle test edilmesi ilk akla gelebilecek çözümdür (27). Ancak bu çözüm öngörülen hata payını çok arttıracığı için pek uygun değildir. Örneğin, $p=5$ iken, değişken ortalamaları arasında 5 kez aynı test uygulanarak farklılık ortaya çıkartılabilmektedir. Bu durumda;

$\alpha=0.01$ düzeyi için testin güvenilirliği $0.99^5=0.95099$, $\alpha=0.05$ düzeyi için ise $0.95^5=0.77378$ olmaktadır. Bir başka deyişle, toplam hata miktarı sırasıyla 0.049 ve 0.2262 gibi yüksek değerlere ulaşabilmektedir (27).

Çıkarımların çeşitliliğini dikkate alan istatistiksel düzeltmelerin gerekli olup olmadığına dair birçok tartışma ve anlaşmazlık vardır (28, 29). Bir düşünce bu tip düzeltmelerin bir yanlış çıkarıma yapma olasılığının α düzeyinde veya daha altında kontrol edilmesi koşulunu sağlaması şeklindedir (28, 30, 31). Buna “**deneysel hatayı**” (familywise error) kontrol etme yaklaşımı denir (11, 13, 18, 28, 30, 31). Buna karşıt diğer bir düşünce bu tip düzeltmelerin gerekli olmadığı ve her bir çıkarımla ayrı şekilde ilgilenilmesi gerektiğidir (32-35). Buna da “**karşılaştırma başına hatayı**” (comparisonwise error) kontrol yaklaşımı denilmektedir (11, 13, 18, 28, 31). İstatistikçiler arasında yaygın görüş deneysel hatanın kontrolüdür (3, 6, 13, 36, 37). Daha sonraları geliştirilen yeni yaklaşım ise “**yanlış bulgu oranının**” (false discovery rate) kontrolüdür (11, 30, 31).

2.5.1. Deneysel Hata Oranı (Experimentwise Error Rate)

Deneysel hata oranı hipotez ailesi içerisindeki hipotezlerden en az birinin yanlışlıkla reddedilme olasılığıdır. Kusursuz çoklu karşılaştırma testi deneysel hata oranını (yanlış pozitif çıkarıma riski) kontrol eder ve böylece belirlenen α -düzeyini aşmaz (13, 18, 30, 31).

Aynı zamanda tip II hatayı (yanlış negatif çıkarıma riski) minimize eder. Ancak bu iki amacın aynı anda gerçekleşmesi mümkün değildir.

Eğer çalışmada 3 grup varsa (A, B ve C) yapılacak çoklu karşılaştırma testleri sonucunda bu test kümesi için genel α -düzeyinin sürdürülmesi deneysel hatanın kontrolü olarak adlandırılır.

2.5.2. Karşılaştırma Başına Hata Oranı (Comparisonwise Error Rate)

Her bir test için ayrı α -düzeyinin belirlenmesi karşılaştırma başına hata kontrol prosedürü olarak adlandırılır (13, 18, 31). Eğer çalışmada 3 grup varsa (A,B ve C) yapılacak karşılaştırmalar şu şekilde olur: A-B, A-C, B-C, AB-C, AC-B, BC-A (3’ü ikili karşılaştırma olmak üzere toplam 6 karşılaştırma). Global hipotez testinin reddedilmesinden sonra A, B ve C grupları için yapılan toplam 6 karşılaştırmada α -düzeyi her bir test için 0.05 olarak seçilir.

2.5.3. Yanlış Bulgu Oranı (False Discovery Rate)

Literatürde hata oranı ile ilgili tartışmaların büyük çoğunluğu deneysel ve karşılaştırma başına hata oranları ile ilgilidir. Fakat son yıllarda çok popüler hale gelen Benjamini ve

Hochberg (38) tarafından önerilen yanlış bulgu oranıdır (YBO). YBO, yazarlar tarafından hatalı retlerin toplam retlere oranının beklenen değeri olarak tanımlanmıştır. Benjamini ve Hochberg (38), YBO'nun kontrol edilmesinin deneysel veya karşılaştırma başına hatanın kontrol edilmesinden daha makul olduğunu belirtmişlerdir.

Tablo-4: M tane H_0 hipotezinde hesaplanan hata sayısı

	Anlamlı Bulunan	Anlamlı Bulunmayan	Toplam
Doğru H_0	V	U	m_0
Yanlış H_0	S	T	$M-m_0$
	R	M-R	M

$\mu_1, \mu_2, \dots, \mu_J$, J tane ortalama olduğunu varsayalım ve istenilen $M=J(J-1)/2$ ikili hipotezler ailesinin testidir. S, R adet ret kümesinden doğru biçimde reddedilen hipotezlerin sayısıdır. V yanlış biçimde reddedilen çift sayısıdır.

Benjamini ve Hochberg (38), yanlış reddedilen sıfır hipotezlerin Q raslantı değişkeni ile ifade edilebileceğini belirtmişlerdir. $Q=V/(V+S)$ eşittir. $R=0$ iken Q sıfır olarak tanımlanmaktadır. Bunun anlamı reddedilme olmadığında hata oranı sıfırdır. Benjamini ve Hochberg (38) tarafından YBO, Q'nun ortalaması olarak tanımlanmıştır.

Buna bağlı olarak;

$$E\left(\frac{V}{M}\right) = \text{Karşılaştırma Başına Hata Oranı}$$

$P(V \geq 1) = \text{Deneysel Hata Oranı}$

$$E(Q) = E\left(\frac{V}{V+S}\right) = E\left(\frac{V}{R}\right) = E\left(\frac{\text{Yanlış Ret Sayısı}}{\text{Toplam Ret Sayısı}}\right)$$

Benjamini ve Hochberg (38), bu hata oranı ile ilgili önemli özellikler sunmuşlardır:

- $\mu_1 = \mu_2 = \dots = \mu_J$ ise YBO deneysel hata oranına eşit olur. Diğer bir deyişle sıfır hipotez doğruysa YBO'nun kontrolü deneysel hata oranının kontrolünü beraberinde getirir. Sıfır hipotez doğruysa $S=0$ ve böylece $V=R$ dir. Eğer $V=0$ ise $Q=0$ olacaktır. Eğer $V>0$ ise $Q=1$ olacaktır ve buna bağlı olarak $P(V \geq 1) = E(Q)$ olur.

- b. $m_0 < M$ olduğu zaman YBO, deneysel hata oranından küçük veya bu orana eşittir. Eğer $V \geq 1$ ise $V/R < 1$ ve eğer $V=0$ ise $V/R=0$ olur. Sonuç olarak $P(V \geq 1) \geq E(Q)$ olur. Bu sonuç bir süreç için deneysel hata oranı kontrol edildiğinde YBO'da kontrol edilmiş olur.
- c. Az sayıda ortalama çifti birbirine eşit ise V/R oranı daha küçük olur. Eşit olmayan çiftler birbirinden çok farklı ise YBO'da ve deneysel hata oranında daha büyük farklılık oluşur.

2.5.4. İstatistiksel Güç Tanımları

Tip II hata olasılığı β 'yi elde etmek için hipotezin yanlış olduğu ve belirli bir hipotezin doğru olduğu varsayılır. Böylece belirlenen H_1 alternatif hipotezin gerçekten doğru olduğunda H_0 'ı kabul etme olasılığı için bir değer belirlenir. İstatistiksel güç ise yanlış H_0 hipotezinin reddedilmesi olasılığıdır ($1 - \beta$). Çoklu karşılaştırma testlerinde güç farklı durumlar altında incelenebilmektedir. Einot ve Gabriel (7), 1975 yılında yayınladıkları makalede çoklu karşılaştırma testlerinin gücünü her bir hipotez kümesi için incelemişlerdir. Buna göre ikili, üçlü, dörtlü ve beşli hipotezlerin güç hesaplamalarını özetlemişlerdir. Daha sonra Ramsey tarafından yapılan çalışmalarda ise güç tanımları daha farklı şekilde belirtilmiştir (19, 39-41). Bu çalışmalardan sonra çoklu karşılaştırma testlerinin güç incelemesinde yaygın olarak Ramsey'in tanımları kullanılmıştır. Ramsey'in güç tanımlarını farklı isimlendiren literatürler mevcuttur (11, 42).

Ramsey'in çoklu karşılaştırma testlerinde kullanılan güç tanımları aşağıdaki gibidir:

i) Herhangi bir çift için güç: Farklı olan ortalama çiftleri arasında en az bir gerçek farkı doğru olarak yakalayabilme olasılığıdır. Farklı olan ortalama çiftlerinden hiçbirini yakalamadığı durumda istenen durum gerçekleşmez.

ii) Bütün çiftler için güç: Farklı olan ortalama çiftleri arasında tüm gerçek farkları doğru olarak yakalayabilme olasılığıdır. Eğer farklı olan ortalama çiftlerinden en az bir tanesi bile yakalanmadıysa istenen durum gerçekleşmez.

iii) Her bir çift için güç: Farklı verilen ortalama çiftinin reddedilme olasılığıdır. Klasik güç tanımıyla eş değerdir. Ortalama çiftlerinin testinde her bir hipotezi birbirinden bağımsız olarak düşünür ve buna bağlı olarak güç hesaplanır.

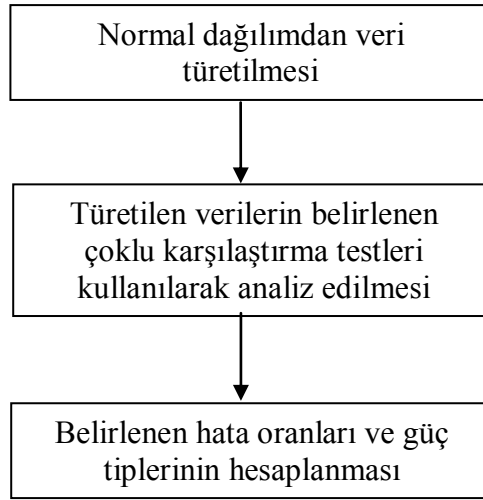
3. GEREÇ VE YÖNTEM

Belirlenen çoklu karşılaştırma testlerinin hata oranları ve güç bakımından birbiriyle karşılaştırılmasında izlenen algoritma Şekil-1’de verilmiştir. Normal dağılıma uygun veri türetilirken ortalaması 40 ve varyansı 2, 4 ve 8 olarak belirlenmiştir.

Testlerin birbiriyle karşılaştırılmasında grup sayısı, varyans, örneklem büyüklükleri dikkate alınmıştır. k , grup sayısı olmak üzere 3, 5 ve 7 grup için karşılaştırmalar gerçekleştirilmiştir. Grupların varyansının homojen olması dikkate alınmıştır. Grupların örneklem büyüklüklerinin eşit olduğu duruma ilişkin simülasyon bilgileri Tablo-5’de belirtilmiştir. Örneklem büyüklüklerinin farklı olduğu duruma ilişkin simülasyon bilgileri Tablo-6’da belirtilmiştir.

Testlerin istatistiksel güç açısından karşılaştırılmasında kullanılan simülasyon tasarımları örneklem büyüklüklerinin eşit olduğu ve farklı olduğu durumlar için Tablo-7 ve Tablo-8’de belirtilmiştir.

Tablo-5’ten Tablo-8’e kadar yer alan simülasyon tasarımlarının her biri için 100 tekrar gerçekleştirilmiştir.



Şekil 1:Veri türetme algoritması

Tablo-5: Örneklem genişliklerinin eşit olduğu durumda Tip I hatanın incelendiği simülasyon senaryoları ($\mu=40$)

Senaryo	Grup Sayısı (k)	Örneklem Büyüklüğü	σ^2
1	3	10/10/10	2
2	3	10/10/10	4
3	3	10/10/10	8
4	3	30/30/30	2
5	3	30/30/30	4
6	3	30/30/30	8
7	3	100/100/100	2
8	3	100/100/100	4
9	3	100/100/100	8
10	5	10/10/10/10/10	2
11	5	10/10/10/10/10	4
12	5	10/10/10/10/10	8
13	5	30/30/30/30/30	2
14	5	30/30/30/30/30	4
15	5	30/30/30/30/30	8
16	5	100/100/100/100/100	2
17	5	100/100/100/100/100	4
18	5	100/100/100/100/100	8
19	7	10/10/10/10/10/10/10	2
20	7	10/10/10/10/10/10/10	4
21	7	10/10/10/10/10/10/10	8
22	7	30/30/30/30/30/30/30	2
23	7	30/30/30/30/30/30/30	4
24	7	30/30/30/30/30/30/30	8
25	7	100/100/100/100/100/100/100	2
26	7	100/100/100/100/100/100/100	4
27	7	100/100/100/100/100/100/100	8

Tablo-6: Örneklem genişliklerinin eşit olmadığı durumda Tip I hatanın incelendiği simülasyon senaryoları ($\mu=40$)

Senaryo	Grup Sayısı (k)	Örneklem Büyüklüğü (n)	σ^2
28	3	10/12/14	2
29	3	10/12/14	4
30	3	10/12/14	8
31	3	30/35/40	2
32	3	30/35/40	4
33	3	30/35/40	8
34	3	100/110/120	2
35	3	100/110/120	4
36	3	100/110/120	8
37	5	10/12/14/16/18	2
38	5	10/12/14/16/18	4
39	5	10/12/14/16/18	8
40	5	30/35/40/45/50	2
41	5	30/35/40/45/50	4
42	5	30/35/40/45/50	8
43	5	100/110/120/130/140	2
44	5	100/110/120/130/140	4
45	5	100/110/120/130/140	8
46	7	10/12/14/16/18/20/22	2
47	7	10/12/14/16/18/20/22	4
48	7	10/12/14/16/18/20/22	8
49	7	30/35/40/45/50/55/60	2
50	7	30/35/40/45/50/55/60	4
51	7	30/35/40/45/50/55/60	8
52	7	100/110/120/130/140/150/160	2
53	7	100/110/120/130/140/150/160	4
54	7	100/110/120/130/140/150/160	8

Tablo-7: Örneklem genişliklerinin eşit olduğu durumda testlerin gücünün karşılaştırılmasında kullanılan simülasyon senaryoları

Senaryo	Grup Sayısı (k)	Örneklem Büyüklüğü	Ortalamalar	σ^2
55	3	10/10/10	40/40/42	2
56	3	10/10/10	40/40/42	4
57	3	10/10/10	40/40/42	8
58	3	30/30/30	40/40/42	2
59	3	30/30/30	40/40/42	4
60	3	30/30/30	40/40/42	8
61	3	100/100/100	40/40/42	2
62	3	100/100/100	40/40/42	4
63	3	100/100/100	40/40/42	8
64	5	10/10/10/10/10	40/40/42/44/46	2
65	5	10/10/10/10/10	40/40/42/44/46	4
66	5	10/10/10/10/10	40/40/42/44/46	8
67	5	30/30/30/30/30	40/40/42/44/46	2
68	5	30/30/30/30/30	40/40/42/44/46	4
69	5	30/30/30/30/30	40/40/42/44/46	8
70	5	100/100/100/100/100	40/40/42/44/46	2
71	5	100/100/100/100/100	40/40/42/44/46	4
72	5	100/100/100/100/100	40/40/42/44/46	8
73	7	10/10/10/10/10/10/10	40/40/42/44/46/48/50	2
74	7	10/10/10/10/10/10/10	40/40/42/44/46/48/50	4
75	7	10/10/10/10/10/10/10	40/40/42/44/46/48/50	8
76	7	30/30/30/30/30/30/30	40/40/42/44/46/48/50	2
77	7	30/30/30/30/30/30/30	40/40/42/44/46/48/50	4
78	7	30/30/30/30/30/30/30	40/40/42/44/46/48/50	8
79	7	100/100/100/100/100/100/100	40/40/42/44/46/48/50	2
80	7	100/100/100/100/100/100/100	40/40/42/44/46/48/50	4
81	7	100/100/100/100/100/100/100	40/40/42/44/46/48/50	8

Tablo-8: Örneklem genişliklerinin eşit olmadığı durumda testlerin gücünün karşılaştırılmasında kullanılan simülasyon senaryoları

Senaryo	Grup Sayısı (k)	Örneklem Büyüklüğü	Ortalamalar	σ^2
82	3	10/12/14	40/40/42	2
83	3	10/12/14	40/40/42	4
84	3	10/12/14	40/40/42	8
85	3	30/35/40	40/40/42	2
86	3	30/35/40	40/40/42	4
87	3	30/35/40	40/40/42	8
88	3	100/110/120	40/40/42	2
89	3	100/110/120	40/40/42	4
90	3	100/110/120	40/40/42	8
91	5	10/12/14/16/18	40/40/42/44/46	2
92	5	10/12/14/16/18	40/40/42/44/46	4
93	5	10/12/14/16/18	40/40/42/44/46	8
94	5	30/35/40/45/50	40/40/42/44/46	2
95	5	30/35/40/45/50	40/40/42/44/46	4
96	5	30/35/40/45/50	40/40/42/44/46	8
97	5	100/110/120/130/140	40/40/42/44/46	2
98	5	100/110/120/130/140	40/40/42/44/46	4
99	5	100/110/120/130/140	40/40/42/44/46	8
100	7	10/12/14/16/18/20/22	40/40/42/44/46/48/50	2
101	7	10/12/14/16/18/20/22	40/40/42/44/46/48/50	4
102	7	10/12/14/16/18/20/22	40/40/42/44/46/48/50	8
103	7	30/35/40/45/50/55/60	40/40/42/44/46/48/50	2
104	7	30/35/40/45/50/55/60	40/40/42/44/46/48/50	4
105	7	30/35/40/45/50/55/60	40/40/42/44/46/48/50	8
106	7	100/110/120/130/140/150/160	40/40/42/44/46/48/50	2
107	7	100/110/120/130/140/150/160	40/40/42/44/46/48/50	4
108	7	100/110/120/130/140/150/160	40/40/42/44/46/48/50	8

4. BULGULAR

4.1. Deneysel Hata Oranı ve Karşılaştırma Başına Hata Oranı Sonuçları

Belirlenen 12 adet çoklu karşılaştırma testlerinin deneysel hata oranı ve karşılaştırma başına hata oranı simülasyon tasarımları Tablo-5 ve Tablo-6'da belirtilmiştir. Bu iki tabloda yer alan 54 senaryoya ilişkin sonuçlar Tablo-9 ile Tablo-20 arasında yer almaktadır. Tablo-9 ve Tablo-14 arası grupların deneysel hata oranı sonuçlarını gösterirken Tablo-15 ve Tablo-20 arası grupların karşılaştırma başına hata oranı sonuçları göstermektedir.

Tablo-9: Grup sayısının 3 ve örneklem büyüklüklerinin eşit olduğu durum için deneysel hata oranı

Çoklu Karşılaştırma Testleri	$n_1=n_2=n_3=10$			$n_1=n_2=n_3=30$			$n_1=n_2=n_3=100$		
	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$
LSD	0.130	0.120	0.120	0.110	0.080	0.100	0.080	0.130	0.100
Bonferroni	0.030	0.040	0.030	0.020	0.040	0.050	0.040	0.030	0.030
Dunn-Sidak	0.030	0.040	0.030	0.020	0.040	0.050	0.040	0.030	0.030
Scheffe	0.030	0.040	0.030	0.020	0.040	0.040	0.030	0.030	0.020
REGW - F	0.040	0.040	0.030	0.020	0.040	0.060	0.040	0.050	0.020
REGW - Q	0.030	0.050	0.040	0.020	0.040	0.060	0.050	0.040	0.050
SNK	0.030	0.050	0.040	0.020	0.040	0.060	0.050	0.040	0.050
Tukey'in a	0.030	0.050	0.040	0.020	0.040	0.060	0.050	0.040	0.050
Tukey'in b	0.030	0.050	0.040	0.020	0.040	0.060	0.050	0.040	0.050
Duncan	0.080	0.090	0.110	0.070	0.080	0.100	0.060	0.120	0.090
Hochberg'in GT2	0.030	0.040	0.040	0.020	0.040	0.050	0.040	0.040	0.030
Gabriel	0.030	0.040	0.040	0.020	0.040	0.050	0.040	0.040	0.030

Tablo-10: Grup sayısının 5 ve örneklem büyüklüklerinin eşit olduğu durum için deneysel hata oranı

Çoklu Karşılaştırma Testleri	$n_1=n_2=n_3=n_4=n_5=10$			$n_1=n_2=n_3=n_4=n_5=30$			$n_1=n_2=n_3=n_4=n_5=100$		
	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$
LSD	0.300	0.280	0.160	0.270	0.310	0.260	0.250	0.240	0.280
Bonferroni	0.020	0.040	0.030	0.060	0.070	0.050	0.060	0.010	0.030
Dunn-Sidak	0.020	0.040	0.030	0.060	0.070	0.050	0.060	0.030	0.030
Scheffe	0.020	0.020	0.020	0.010	0.040	0.010	0.010	0.010	0.010
REGW - F	0.040	0.050	0.030	0.100	0.080	0.070	0.070	0.030	0.040
REGW - Q	0.030	0.060	0.030	0.070	0.080	0.090	0.070	0.030	0.030
SNK	0.030	0.060	0.030	0.070	0.080	0.090	0.070	0.030	0.030
Tukey'in a	0.030	0.060	0.030	0.070	0.080	0.090	0.070	0.030	0.030
Tukey'in b	0.030	0.060	0.030	0.070	0.080	0.090	0.070	0.030	0.030
Duncan	0.180	0.190	0.120	0.220	0.210	0.170	0.190	0.150	0.190
Hochberg'in GT2	0.020	0.040	0.030	0.060	0.070	0.050	0.060	0.030	0.030
Gabriel	0.020	0.040	0.030	0.060	0.070	0.050	0.060	0.030	0.030

Tablo-11: Grup sayısının 7 ve örneklem büyüklüklerinin eşit olduğu durum için deneysel hata oranı

Çoklu Karşılaştırma Testleri	$n_1=n_2=n_3=n_4=n_5=n_6=n_7=10$			$n_1=n_2=n_3=n_4=n_5=n_6=n_7=30$			$n_1=n_2=n_3=n_4=n_5=n_6=n_7=100$		
	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$
LSD	0.420	0.420	0.500	0.440	0.410	0.380	0.440	0.470	0.380
Bonferroni	0.060	0.040	0.020	0.050	0.060	0.080	0.050	0.020	0.050
Dunn-Sidak	0.060	0.040	0.020	0.050	0.060	0.080	0.050	0.020	0.050
Scheffe	0.020	0.010	0.010	0.020	0.020	0.030	0.000	0.000	0.020
REGW - F	0.080	0.040	0.060	0.080	0.050	0.060	0.070	0.050	0.070
REGW - Q	0.070	0.040	0.040	0.060	0.070	0.080	0.050	0.030	0.060
SNK	0.070	0.040	0.040	0.060	0.070	0.080	0.050	0.030	0.060
Tukey'in a	0.070	0.040	0.040	0.060	0.090	0.080	0.050	0.030	0.060
Tukey'in b	0.070	0.040	0.040	0.060	0.080	0.080	0.050	0.030	0.060
Duncan	0.290	0.230	0.330	0.290	0.270	0.220	0.290	0.240	0.250
Hochberg'in GT2	0.060	0.040	0.020	0.050	0.060	0.080	0.050	0.020	0.050
Gabriel	0.060	0.040	0.020	0.050	0.070	0.080	0.050	0.020	0.050

Tablo-12: Grup sayısının 3 ve örneklem büyüklüklerinin farklı olduğu durum için deneysel hata oranı

Çoklu Karşılaştırma Testleri	$n_1=10; n_2=12; n_3=14$			$n_1=30; n_2=35; n_3=40$			$n_1=100; n_2=110; n_3=120$		
	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$
LSD	0.180	0.100	0.090	0.110	0.200	0.150	0.110	0.130	0.140
Bonferroni	0.050	0.040	0.040	0.060	0.060	0.060	0.060	0.060	0.020
Dunn-Sidak	0.050	0.040	0.040	0.060	0.060	0.060	0.060	0.060	0.020
Scheffe	0.040	0.030	0.020	0.060	0.070	0.050	0.050	0.050	0.020
REGW - F	0.050	0.050	0.040	0.060	0.090	0.060	0.060	0.080	0.030
REGW - Q	0.050	0.040	0.030	0.060	0.060	0.060	0.060	0.050	0.020
SNK	0.070	0.040	0.040	0.060	0.100	0.060	0.080	0.070	0.020
Tukey'in a	0.080	0.040	0.040	0.060	0.100	0.060	0.080	0.070	0.020
Tukey'in b	0.080	0.040	0.040	0.060	0.100	0.060	0.080	0.070	0.020
Duncan	0.130	0.080	0.070	0.090	0.170	0.130	0.100	0.100	0.090
Hochberg'in GT2	0.060	0.030	0.040	0.060	0.080	0.050	0.050	0.060	0.020
Gabriel	0.060	0.030	0.040	0.060	0.080	0.050	0.050	0.060	0.020

Tablo-13: Grup sayısının 5 ve örneklem büyüklüklerinin farklı olduğu durum için deneysel hata oranı

Çoklu Karşılaştırma Testleri	$n_1=10; n_2=12; n_3=14; n_4=16; n_5=18$			$n_1=30; n_2=35; n_3=40; n_4=45; n_5=50$			$n_1=100; n_2=110; n_3=120; n_4=130; n_5=140$		
	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$
LSD	0.380	0.260	0.220	0.290	0.210	0.350	0.370	0.270	0.250
Bonferroni	0.030	0.050	0.030	0.050	0.030	0.040	0.040	0.040	0.030
Dunn-Sidak	0.030	0.050	0.030	0.050	0.030	0.050	0.040	0.040	0.030
Scheffe	0.020	0.020	0.020	0.030	0.000	0.030	0.010	0.030	0.010
REGW - F	0.040	0.030	0.020	0.050	0.030	0.060	0.050	0.040	0.030
REGW - Q	0.020	0.040	0.020	0.050	0.020	0.030	0.030	0.040	0.020
SNK	0.030	0.040	0.030	0.060	0.020	0.060	0.060	0.050	0.030
Tukey'in a	0.030	0.040	0.030	0.060	0.020	0.060	0.060	0.050	0.030
Tukey'in b	0.030	0.040	0.030	0.060	0.020	0.060	0.060	0.050	0.030
Duncan	0.200	0.190	0.150	0.200	0.170	0.220	0.240	0.180	0.150
Hochberg'in GT2	0.030	0.020	0.030	0.050	0.020	0.050	0.050	0.040	0.030
Gabriel	0.030	0.020	0.030	0.050	0.020	0.050	0.050	0.040	0.030

Tablo-14: Grup sayısının 7 ve örneklem büyüklüklerinin farklı olduğu durum için deneysel hata oranı

Çoklu Karşılaştırma Testleri	$n_1=10; n_2=12; n_3=14; n_4=16;$ $n_5=18; n_6=20; n_7=22$			$n_1=30; n_2=35; n_3=40; n_4=45;$ $n_5=50; n_6=55; n_7=60$			$n_1=100; n_2=110; n_3=120;$ $n_4=130; n_5=140; n_6=150;$ $n_7=160$		
	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$
LSD	0.410	0.450	0.420	0.490	0.400	0.430	0.460	0.400	0.480
Bonferroni	0.040	0.060	0.010	0.020	0.010	0.050	0.010	0.030	0.010
Dunn-Sidak	0.040	0.060	0.010	0.020	0.020	0.050	0.010	0.030	0.010
Scheffe	0.030	0.020	0.010	0.010	0.000	0.000	0.000	0.020	0.000
REGW - F	0.070	0.080	0.030	0.020	0.040	0.060	0.050	0.040	0.020
REGW - Q	0.040	0.050	0.010	0.020	0.020	0.050	0.010	0.030	0.010
SNK	0.070	0.070	0.090	0.040	0.020	0.050	0.020	0.050	0.030
Tukey'in a	0.070	0.070	0.090	0.040	0.020	0.050	0.020	0.050	0.030
Tukey'in b	0.070	0.070	0.090	0.040	0.020	0.050	0.020	0.050	0.030
Duncan	0.290	0.300	0.250	0.310	0.240	0.250	0.280	0.240	0.240
Hochberg'in GT2	0.060	0.060	0.080	0.030	0.020	0.050	0.010	0.030	0.020
Gabriel	0.060	0.060	0.080	0.030	0.020	0.050	0.010	0.030	0.020

Tablo-15: Grup sayısının 3 ve örneklem büyüklüklerinin eşit olduğu durum için karşılaştırma başına hata oranı

Çoklu Karşılaştırma Testleri	$n_1=n_2=n_3=10$			$n_1=n_2=n_3=30$			$n_1=n_2=n_3=100$		
	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$
LSD	0.043	0.050	0.047	0.047	0.037	0.040	0.033	0.063	0.040
Bonferroni	0.010	0.017	0.010	0.007	0.020	0.020	0.013	0.013	0.010
Dunn-Sidak	0.010	0.017	0.010	0.007	0.020	0.020	0.013	0.013	0.010
Scheffe	0.010	0.017	0.010	0.007	0.017	0.017	0.010	0.010	0.007
REGW - F	0.020	0.020	0.017	0.013	0.023	0.027	0.020	0.030	0.010
REGW - Q	0.013	0.023	0.020	0.013	0.023	0.023	0.023	0.023	0.020
SNK	0.013	0.023	0.020	0.013	0.023	0.023	0.023	0.023	0.020
Tukey'in a	0.010	0.023	0.013	0.010	0.020	0.023	0.017	0.013	0.017
Tukey'in b	0.010	0.023	0.017	0.010	0.020	0.023	0.020	0.020	0.017
Duncan	0.033	0.037	0.043	0.033	0.037	0.040	0.027	0.057	0.033
Hochberg'in GT2	0.010	0.017	0.013	0.007	0.020	0.020	0.013	0.013	0.010
Gabriel	0.010	0.017	0.013	0.007	0.020	0.020	0.013	0.013	0.010

Tablo-16: Grup sayısının 5 ve örneklem büyüklüklerinin eşit olduğu durum için karşılaştırma başına hata oranı

Çoklu Karşılaştırma Testleri	$n_1=n_2=n_3=n_4=n_5=10$			$n_1=n_2=n_3=n_4=n_5=30$			$n_1=n_2=n_3=n_4=n_5=100$		
	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$
LSD	0.054	0.058	0.029	0.062	0.058	0.048	0.048	0.041	0.054
Bonferroni	0.002	0.004	0.004	0.010	0.010	0.006	0.006	0.001	0.004
Dunn-Sidak	0.002	0.004	0.004	0.010	0.011	0.006	0.006	0.003	0.004
Scheffe	0.007	0.007	0.007	0.003	0.017	0.003	0.003	0.003	0.003
REGW - F	0.023	0.017	0.013	0.050	0.067	0.037	0.033	0.010	0.020
REGW - Q	0.013	0.020	0.013	0.040	0.060	0.043	0.027	0.010	0.017
SNK	0.020	0.020	0.013	0.040	0.060	0.047	0.030	0.010	0.017
Tukey'in a	0.013	0.020	0.013	0.037	0.043	0.037	0.023	0.010	0.013
Tukey'in b	0.013	0.020	0.013	0.037	0.060	0.040	0.027	0.010	0.017
Duncan	0.103	0.110	0.070	0.143	0.143	0.113	0.120	0.077	0.107
Hochberg'in GT2	0.007	0.013	0.013	0.033	0.037	0.020	0.020	0.010	0.013
Gabriel	0.007	0.013	0.013	0.033	0.037	0.020	0.020	0.010	0.013

Tablo-17: Grup sayısının 7 ve örneklem büyüklüklerinin eşit olduğu durum için karşılaştırma başına hata oranı

Çoklu Karşılaştırma Testleri	$n_1=n_2=n_3=n_4=n_5=n_6=n_7=10$			$n_1=n_2=n_3=n_4=n_5=n_6=n_7=30$			$n_1=n_2=n_3=n_4=n_5=n_6=n_7=100$		
	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$
LSD	0.057	0.043	0.058	0.056	0.052	0.046	0.051	0.057	0.053
Bonferroni	0.004	0.002	0.002	0.003	0.004	0.005	0.004	0.001	0.004
Dunn-Sidak	0.004	0.002	0.002	0.003	0.004	0.006	0.004	0.001	0.005
Scheffe	0.007	0.007	0.003	0.007	0.010	0.010	0.000	0.000	0.013
REGW - F	0.033	0.023	0.030	0.040	0.027	0.043	0.050	0.027	0.053
REGW - Q	0.030	0.017	0.020	0.027	0.033	0.050	0.043	0.010	0.053
SNK	0.037	0.023	0.020	0.027	0.037	0.057	0.043	0.010	0.060
Tukey'in a	0.030	0.017	0.020	0.027	0.073	0.040	0.027	0.010	0.047
Tukey'in b	0.030	0.017	0.020	0.027	0.043	0.050	0.040	0.010	0.053
Duncan	0.240	0.160	0.220	0.250	0.213	0.177	0.207	0.187	0.223
Hochberg'in GT2	0.027	0.017	0.013	0.023	0.027	0.040	0.027	0.007	0.033
Gabriel	0.027	0.017	0.013	0.023	0.040	0.040	0.027	0.007	0.033

Tablo-18: Grup sayısının 3 ve örneklem büyüklüklerinin farklı olduğu durum için karşılaştırma başına hata oranı

Çoklu Karşılaştırma Testleri	$n_1=10; n_2=12; n_3=14$			$n_1=30; n_2=35; n_3=40$			$n_1=100; n_2=110; n_3=120$		
	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$
LSD	0.063	0.033	0.033	0.043	0.083	0.067	0.047	0.060	0.050
Bonferroni	0.020	0.013	0.013	0.020	0.023	0.027	0.020	0.020	0.007
Dunn-Sidak	0.020	0.013	0.013	0.020	0.023	0.027	0.020	0.020	0.007
Scheffe	0.017	0.010	0.007	0.020	0.027	0.023	0.017	0.017	0.007
REGW - F	0.020	0.017	0.017	0.027	0.047	0.037	0.030	0.043	0.013
REGW - Q	0.020	0.013	0.013	0.023	0.023	0.037	0.030	0.023	0.007
SNK	0.027	0.013	0.017	0.030	0.043	0.037	0.037	0.037	0.007
Tukey'in a	0.030	0.013	0.013	0.020	0.037	0.027	0.027	0.023	0.007
Tukey'in b	0.030	0.013	0.013	0.023	0.037	0.033	0.030	0.023	0.007
Duncan	0.047	0.027	0.027	0.040	0.070	0.060	0.043	0.050	0.033
Hochberg'in GT2	0.023	0.010	0.013	0.020	0.030	0.023	0.017	0.020	0.007
Gabriel	0.023	0.010	0.013	0.020	0.030	0.023	0.017	0.020	0.007

Tablo-19: Grup sayısının 5 ve örneklem büyüklüklerinin farklı olduğu durum için karşılaştırma başına hata oranı

Çoklu Karşılaştırma Testleri	$n_1=10; n_2=12; n_3=14; n_4=16; n_5=18$			$n_1=30; n_2=35; n_3=40; n_4=45; n_5=50$			$n_1=100; n_2=110; n_3=120; n_4=130; n_5=140$		
	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$
LSD	0.060	0.053	0.037	0.054	0.036	0.053	0.072	0.043	0.044
Bonferroni	0.004	0.007	0.005	0.006	0.003	0.006	0.005	0.004	0.003
Dunn-Sidak	0.004	0.007	0.005	0.006	0.003	0.007	0.005	0.004	0.003
Scheffe	0.007	0.007	0.010	0.013	0.000	0.010	0.007	0.010	0.003
REGW - F	0.027	0.027	0.017	0.023	0.013	0.037	0.023	0.020	0.020
REGW - Q	0.010	0.020	0.013	0.023	0.007	0.017	0.017	0.013	0.017
SNK	0.023	0.030	0.027	0.033	0.013	0.033	0.027	0.017	0.020
Tukey'in a	0.023	0.017	0.020	0.027	0.007	0.027	0.023	0.017	0.010
Tukey'in b	0.023	0.023	0.020	0.027	0.013	0.033	0.027	0.017	0.020
Duncan	0.110	0.113	0.080	0.130	0.093	0.117	0.143	0.093	0.083
Hochberg'in GT2	0.020	0.010	0.017	0.023	0.007	0.023	0.020	0.013	0.010
Gabriel	0.020	0.010	0.017	0.023	0.007	0.023	0.020	0.013	0.010

Tablo-20: Grup sayısının 7 ve örneklem büyüklüklerinin farklı olduğu durum için karşılaştırma başına hata oranı

Çoklu Karşılaştırma Testleri	$n_1=10; n_2=12; n_3=14; n_4=16;$ $n_5=18; n_6=20; n_7=22$			$n_1=30; n_2=35; n_3=40; n_4=45;$ $n_5=50; n_6=55; n_7=60$			$n_1=100; n_2=110; n_3=120;$ $n_4=130; n_5=140; n_6=150;$ $n_7=160$		
	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$
LSD	0.060	0.055	0.048	0.050	0.039	0.044	0.054	0.045	0.052
Bonferroni	0.003	0.003	0.000	0.001	0.001	0.003	0.001	0.003	0.001
Dunn-Sidak	0.003	0.003	0.000	0.001	0.001	0.003	0.001	0.003	0.001
Scheffe	0.013	0.007	0.003	0.003	0.000	0.000	0.000	0.013	0.000
REGW - F	0.040	0.030	0.010	0.007	0.013	0.027	0.030	0.027	0.023
REGW - Q	0.020	0.017	0.003	0.007	0.010	0.027	0.003	0.023	0.007
SNK	0.057	0.047	0.037	0.020	0.010	0.030	0.010	0.043	0.013
Tukey'in a	0.040	0.030	0.030	0.017	0.007	0.027	0.010	0.033	0.010
Tukey'in b	0.050	0.037	0.033	0.017	0.007	0.030	0.010	0.033	0.013
Duncan	0.263	0.213	0.170	0.177	0.150	0.177	0.213	0.170	0.183
Hochberg'in GT2	0.030	0.027	0.027	0.013	0.007	0.023	0.003	0.023	0.007
Gabriel	0.030	0.027	0.027	0.013	0.007	0.023	0.003	0.023	0.007

4.2. İstatistiksel Güç ve Yanlış Bulgu Oranı Bakımından Karşılaştırılması

Çoklu karşılaştırma testlerinin istatistiksel güç bakımından karşılaştırılmasında farklı güç tanımları üzerinden hesaplamalar yapılmıştır. Buna göre testlerin “her bir çift” için, “herhangi bir çift” için, “bütün çiftler” için güç hesaplamaları gerçekleştirilmiştir. Çoklu karşılaştırma testlerinin birbiriyle kıyaslanmasında kullanılan “yanlış bulgu oranı” hesaplanmıştır. Güç ve yanlış bulgu oranı nın kıyaslanmasında kullanılan simülasyon senaryoları örneklem büyüklüklerinin eşit olduğu durum için Tablo-7’de, örneklem büyüklüklerinin farklı olduğu durum için Tablo-8’de verilmiştir. Tablo-21 ile Tablo-26 arasında her bir çift için, Tablo-27 ile Tablo-32 arası her hangi bir çift için güç, Tablo-33 ile Tablo-38 arası bütün çiftler için güç, Tablo-39 ile Tablo-44 arası yanlış bulgu oranı sonuçlarını göstermektedir.

Tablo-21: Grup sayısının 3 ve örneklem büyüklüklerinin eşit olduğu durum için her bir çift için güç değerleri

Çoklu Karşılaştırma Testleri	$n_1=n_2=n_3=10$			$n_1=n_2=n_3=30$			$n_1=n_2=n_3=100$		
	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$
LSD	0.855	0.550	0.355	1.000	0.975	0.755	1.000	1.000	1.000
Bonferroni	0.845	0.540	0.355	0.990	0.965	0.745	0.990	0.990	0.990
Dunn-Sidak	0.835	0.530	0.355	0.980	0.955	0.745	0.980	0.980	0.980
Scheffe	0.825	0.520	0.350	0.970	0.945	0.735	0.970	0.970	0.970
REGW - F	0.830	0.450	0.260	1.000	0.975	0.735	1.000	1.000	1.000
REGW - Q	0.830	0.440	0.255	1.000	0.975	0.715	1.000	1.000	1.000
SNK	0.830	0.440	0.255	1.000	0.975	0.715	1.000	1.000	1.000
Tukey'in a	0.735	0.385	0.220	1.000	0.930	0.665	1.000	1.000	0.990
Tukey'in b	0.790	0.410	0.235	1.000	0.955	0.675	1.000	1.000	0.995
Duncan	0.850	0.535	0.320	1.000	0.975	0.745	1.000	1.000	1.000
Hochberg'in GT2	0.720	0.370	0.215	1.000	0.955	0.650	1.000	1.000	0.985
Gabriel	0.720	0.370	0.215	1.000	0.930	0.650	1.000	1.000	0.985

Tablo-22: Grup sayısının 5 ve örneklem büyüklüklerinin eşit olduğu durum için her bir çift için güç değerleri

Çoklu Karşılaştırma Testleri	$n_1=n_2=n_3=n_4=n_5=10$			$n_1=n_2=n_3=n_4=n_5=30$			$n_1=n_2=n_3=n_4=n_5=100$		
	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$
LSD	0.943	0.816	0.656	1.000	0.989	0.903	1.000	1.000	0.999
Bonferroni	0.798	0.643	0.451	0.998	0.916	0.747	1.000	1.000	0.992
Dunn-Sidak	0.800	0.648	0.453	0.998	0.916	0.748	1.000	1.000	0.992
Scheffe	0.760	0.598	0.390	0.996	0.874	0.704	1.000	1.000	0.992
REGW - F	0.890	0.739	0.540	0.999	0.957	0.841	1.000	1.000	0.999
REGW - Q	0.888	0.736	0.530	0.999	0.953	0.841	1.000	1.000	0.999
SNK	0.932	0.793	0.599	1.000	0.983	0.886	1.000	1.000	0.999
Tukey'in a	0.824	0.668	0.469	0.998	0.928	0.763	1.000	1.000	0.993
Tukey'in b	0.884	0.729	0.519	0.999	0.949	0.838	1.000	1.000	0.999
Duncan	0.941	0.810	0.642	1.000	0.989	0.901	1.000	1.000	0.999
Hochberg'in GT2	0.800	0.649	0.454	0.998	0.916	0.748	1.000	1.000	0.992
Gabriel	0.792	0.642	0.448	0.988	0.907	0.740	0.990	0.990	0.982

Tablo-23: Grup sayısının 7 ve örneklem büyüklüklerinin eşit olduğu durum için her bir çift için güç değerleri

Çoklu Karşılaştırma Testleri	$n_1=n_2=n_3=n_4=n_5=n_6=n_7=10$			$n_1=n_2=n_3=n_4=n_5=n_6=n_7=30$			$n_1=n_2=n_3=n_4=n_5=n_6=n_7=100$		
	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$
LSD	0.959	0.881	0.778	1.000	0.996	0.931	1.000	1.000	1.000
Bonferroni	0.850	0.735	0.580	0.998	0.934	0.814	1.000	1.000	0.994
Dunn-Sidak	0.851	0.736	0.582	0.998	0.935	0.814	1.000	1.000	0.995
Scheffe	0.797	0.674	0.512	0.991	0.889	0.756	1.000	1.000	0.975
REGW - F	0.921	0.802	0.661	1.000	0.975	0.879	1.000	1.000	0.999
REGW - Q	0.918	0.804	0.665	1.000	0.974	0.877	1.000	1.000	0.999
SNK	0.956	0.869	0.729	1.000	0.995	0.924	1.000	1.000	1.000
Tukey'in a	0.861	0.744	0.597	0.998	0.944	0.821	1.000	1.000	0.996
Tukey'in b	0.919	0.800	0.660	1.000	0.974	0.875	1.000	1.000	0.999
Duncan	0.958	0.879	0.768	1.000	0.996	0.929	1.000	1.000	1.000
Hochberg'in GT2	0.853	0.736	0.583	0.998	0.935	0.814	1.000	1.000	0.995
Gabriel	0.844	0.729	0.576	0.988	0.926	0.806	0.990	0.990	0.985

Tablo-24: Grup sayısının 3 ve örneklem büyüklüklerinin farklı olduğu durum için her bir çift için güç değerleri

Çoklu Karşılaştırma Testleri	$n_1=10, n_2=12, n_3=14$			$n_1=30, n_2=35, n_3=40$			$n_1=100, n_2=110, n_3=120$		
	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$
LSD	0.940	0.620	0.445	1.000	0.990	0.860	1.000	1.000	1.000
Bonferroni	0.930	0.615	0.445	0.990	0.980	0.850	0.990	0.990	0.990
Dunn-Sidak	0.920	0.605	0.435	0.980	0.970	0.840	0.980	0.980	0.980
Scheffe	0.910	0.605	0.430	0.970	0.960	0.835	0.970	0.970	0.970
REGW - F	0.935	0.555	0.385	1.000	0.990	0.810	1.000	1.000	1.000
REGW - Q	0.905	0.475	0.280	1.000	0.985	0.770	1.000	1.000	1.000
SNK	0.915	0.510	0.300	1.000	0.990	0.780	1.000	1.000	1.000
Tukey'in a	0.905	0.460	0.285	1.000	0.980	0.740	1.000	1.000	1.000
Tukey'in b	0.915	0.475	0.270	1.000	0.985	0.765	1.000	1.000	1.000
Duncan	0.925	0.605	0.415	1.000	0.990	0.830	1.000	1.000	1.000
Hochberg'in GT2	0.885	0.430	0.270	1.000	0.980	0.720	1.000	1.000	1.000
Gabriel	0.885	0.435	0.270	1.000	0.980	0.720	1.000	1.000	1.000

Tablo-25: Grup sayısının 5 ve örneklem büyüklüklerinin farklı olduğu durum için her bir çift için güç değerleri

Çoklu Karşılaştırma Testleri	$n_1=10, n_2=12, n_3=14, n_4=16, n_5=18$			$n_1=30, n_2=35, n_3=40, n_4=45, n_5=50$			$n_1=100, n_2=110, n_3=120, n_4=130, n_5=140$		
	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$
LSD	0.977	0.894	0.744	1.000	0.993	0.949	1.000	1.000	1.000
Bonferroni	0.903	0.723	0.559	1.000	0.967	0.840	1.000	1.000	1.000
Dunn-Sidak	0.903	0.723	0.561	1.000	0.967	0.841	1.000	1.000	1.000
Scheffe	0.871	0.668	0.527	0.999	0.948	0.792	1.000	1.000	0.997
REGW - F	0.953	0.811	0.640	1.000	0.984	0.909	1.000	1.000	1.000
REGW - Q	0.941	0.776	0.604	1.000	0.980	0.898	1.000	1.000	1.000
SNK	0.977	0.877	0.688	1.000	0.993	0.943	1.000	1.000	1.000
Tukey'in a	0.923	0.740	0.578	1.000	0.969	0.859	1.000	1.000	1.000
Tukey'in b	0.954	0.801	0.628	1.000	0.986	0.906	1.000	1.000	1.000
Duncan	0.978	0.888	0.731	1.000	0.993	0.948	1.000	1.000	1.000
Hochberg'in GT2	0.906	0.723	0.561	1.000	0.967	0.841	1.000	1.000	1.000
Gabriel	0.898	0.718	0.557	0.990	0.957	0.832	0.990	0.990	0.990

Tablo-26: Grup sayısının 7 ve örneklem büyüklüklerinin farklı olduğu durum için her bir çift için güç değerleri

Çoklu Karşılaştırma Testleri	$n_1=10, n_2=12, n_3=14, n_4=16, n_5=18, n_6=20, n_7=22$			$n_1=30, n_2=35, n_3=40, n_4=45, n_5=50, n_6=55, n_7=60$			$n_1=100, n_2=110, n_3=120, n_4=130, n_5=140, n_6=150, n_7=160$		
	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$
LSD	0.992	0.939	0.846	1.000	0.999	0.968	1.000	1.000	1.000
Bonferroni	0.939	0.816	0.692	1.000	0.979	0.887	1.000	1.000	0.999
Dunn-Sidak	0.939	0.817	0.692	1.000	0.979	0.887	1.000	1.000	0.999
Scheffe	0.893	0.762	0.628	0.999	0.949	0.824	1.000	1.000	0.994
REGW - F	0.980	0.884	0.765	1.000	0.995	0.935	1.000	1.000	1.000
REGW - Q	0.973	0.869	0.743	1.000	0.993	0.926	1.000	1.000	1.000
SNK	0.992	0.932	0.820	1.000	0.999	0.963	1.000	1.000	1.000
Tukey'in a	0.946	0.826	0.709	1.000	0.983	0.896	1.000	1.000	0.999
Tukey'in b	0.980	0.873	0.757	1.000	0.995	0.932	1.000	1.000	1.000
Duncan	0.992	0.936	0.835	1.000	0.999	0.966	1.000	1.000	1.000
Hochberg'in GT2	0.939	0.817	0.693	1.000	0.979	0.887	1.000	1.000	0.999
Gabriel	0.931	0.809	0.688	0.990	0.970	0.878	0.990	0.990	0.989

Tablo-27: Grup sayısının 3 ve örneklem büyüklüklerinin eşit olduğu durum için herhangi bir çift için güç değerleri

Çoklu Karşılaştırma Testleri	$n_1=n_2=n_3=10$			$n_1=n_2=n_3=30$			$n_1=n_2=n_3=100$		
	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$
LSD	0.940	0.680	0.510	1.000	1.000	0.870	1.000	1.000	1.000
Bonferroni	0.880	0.490	0.310	1.000	1.000	0.820	1.000	1.000	1.000
Dunn-Sidak	0.880	0.490	0.320	1.000	1.000	0.820	1.000	1.000	1.000
Scheffe	0.870	0.470	0.290	1.000	1.000	0.790	1.000	1.000	1.000
REGW - F	0.890	0.510	0.320	1.000	1.000	0.840	1.000	1.000	1.000
REGW - Q	0.900	0.500	0.330	1.000	1.000	0.820	1.000	1.000	1.000
SNK	0.900	0.500	0.330	1.000	1.000	0.820	1.000	1.000	1.000
Tukey'in a	0.900	0.500	0.330	1.000	1.000	0.820	1.000	1.000	1.000
Tukey'in b	0.900	0.500	0.330	1.000	1.000	0.820	1.000	1.000	1.000
Duncan	0.930	0.650	0.440	1.000	1.000	0.860	1.000	1.000	1.000
Hochberg'in GT2	0.890	0.490	0.320	1.000	1.000	0.820	1.000	1.000	1.000
Gabriel	0.890	0.490	0.320	1.000	1.000	0.820	1.000	1.000	1.000

Tablo-28: Grup sayısının 5 ve örneklem büyüklüklerinin eşit olduğu durum için herhangi bir çift için güç değerleri

Çoklu Karşılaştırma Testleri	$n_1=n_2=n_3=n_4=n_5=10$			$n_1=n_2=n_3=n_4=n_5=30$			$n_1=n_2=n_3=n_4=n_5=100$		
	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$
LSD	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Bonferroni	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Dunn-Sidak	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Scheffe	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
REGW - F	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
REGW - Q	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
SNK	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Tukey'in a	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Tukey'in b	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Duncan	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Hochberg'in GT2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Gabriel	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Tablo-29: Grup sayısının 7 ve örneklem büyüklüklerinin eşit olduğu durum için herhangi bir çift için güç değerleri

Çoklu Karşılaştırma Testleri	$n_1=n_2=n_3=n_4=n_5=n_6=n_7=10$			$n_1=n_2=n_3=n_4=n_5=n_6=n_7=30$			$n_1=n_2=n_3=n_4=n_5=n_6=n_7=100$		
	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$
LSD	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Bonferroni	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Dunn-Sidak	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Scheffe	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
REGW - F	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
REGW - Q	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
SNK	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Tukey'in a	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Tukey'in b	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Duncan	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Hochberg'in GT2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Gabriel	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Tablo-30: Grup sayısının 3 ve örneklem büyüklüklerinin farklı olduğu durum için herhangi bir çift için güç değerleri

Çoklu Karşılaştırma Testleri	$n_1=10, n_2=12, n_3=14$			$n_1=30, n_2=35, n_3=40$			$n_1=100, n_2=110, n_3=120$		
	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$
LSD	0.990	0.770	0.630	1.000	1.000	0.920	1.000	1.000	1.000
Bonferroni	0.960	0.560	0.430	1.000	1.000	0.830	1.000	1.000	1.000
Dunn-Sidak	0.960	0.560	0.430	1.000	1.000	0.830	1.000	1.000	1.000
Scheffe	0.950	0.560	0.400	1.000	1.000	0.820	1.000	1.000	1.000
REGW - F	0.980	0.640	0.510	1.000	1.000	0.860	1.000	1.000	1.000
REGW - Q	0.950	0.560	0.400	1.000	1.000	0.830	1.000	1.000	1.000
SNK	0.970	0.600	0.410	1.000	1.000	0.830	1.000	1.000	1.000
Tukey'in a	0.970	0.620	0.450	1.000	1.000	0.840	1.000	1.000	1.000
Tukey'in b	0.970	0.600	0.410	1.000	1.000	0.830	1.000	1.000	1.000
Duncan	0.990	0.750	0.580	1.000	1.000	0.890	1.000	1.000	1.000
Hochberg'in GT2	0.960	0.560	0.430	1.000	1.000	0.830	1.000	1.000	1.000
Gabriel	0.960	0.570	0.430	1.000	1.000	0.830	1.000	1.000	1.000

Tablo-31: Grup sayısının 5 ve örneklem büyüklüklerinin farklı olduğu durum için herhangi bir çift için güç değerleri

Çoklu Karşılaştırma Testleri	$n_1=10, n_2=12, n_3=14, n_4=16, n_5=18$			$n_1=30, n_2=35, n_3=40, n_4=45, n_5=50$			$n_1=100, n_2=110, n_3=120, n_4=130, n_5=140$		
	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$
LSD	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Bonferroni	1.000	1.000	0.990	1.000	1.000	1.000	1.000	1.000	1.000
Dunn-Sidak	1.000	1.000	0.990	1.000	1.000	1.000	1.000	1.000	1.000
Scheffe	1.000	1.000	0.990	1.000	1.000	1.000	1.000	1.000	1.000
REGW - F	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
REGW - Q	1.000	1.000	0.990	1.000	1.000	1.000	1.000	1.000	1.000
SNK	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Tukey'in a	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Tukey'in b	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Duncan	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Hochberg'in GT2	1.000	1.000	0.990	1.000	1.000	1.000	1.000	1.000	1.000
Gabriel	1.000	1.000	0.990	1.000	1.000	1.000	1.000	1.000	1.000

Tablo-32: Grup sayısının 7 ve örneklem büyüklüklerinin farklı olduğu durum için herhangi bir çift için güç değerleri

Çoklu Karşılaştırma Testleri	$n_1=10, n_2=12, n_3=14, n_4=16, n_5=18, n_6=20, n_7=22$			$n_1=30, n_2=35, n_3=40, n_4=45, n_5=50, n_6=55, n_7=60$			$n_1=100, n_2=110, n_3=120, n_4=130, n_5=140, n_6=150, n_7=160$		
	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$
LSD	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Bonferroni	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Dunn-Sidak	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Scheffe	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
REGW - F	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
REGW - Q	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
SNK	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Tukey'in a	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Tukey'in b	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Duncan	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Hochberg'in GT2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Gabriel	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Tablo-33: Grup sayısının 3 ve örneklem büyüklüklerinin eşit olduğu durum için bütün çift için güç değerleri

Çoklu Karşılaştırma Testleri	$n_1=n_2=n_3=10$			$n_1=n_2=n_3=30$			$n_1=n_2=n_3=100$		
	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$
LSD	0.770	0.420	0.200	1.000	0.950	0.640	1.000	1.000	1.000
Bonferroni	0.540	0.250	0.110	1.000	0.860	0.480	1.000	1.000	0.970
Dunn-Sidak	0.550	0.250	0.110	1.000	0.860	0.480	1.000	1.000	0.970
Scheffe	0.530	0.240	0.110	1.000	0.830	0.470	1.000	1.000	0.970
REGW - F	0.770	0.390	0.200	1.000	0.950	0.630	1.000	1.000	1.000
REGW - Q	0.760	0.380	0.180	1.000	0.950	0.610	1.000	1.000	1.000
SNK	0.760	0.380	0.180	1.000	0.950	0.610	1.000	1.000	1.000
Tukey'in a	0.570	0.270	0.110	1.000	0.860	0.510	1.000	1.000	0.980
Tukey'in b	0.680	0.320	0.140	1.000	0.910	0.530	1.000	1.000	0.990
Duncan	0.770	0.420	0.200	1.000	0.950	0.630	1.000	1.000	1.000
Hochberg'in GT2	0.550	0.250	0.110	1.000	0.910	0.480	1.000	1.000	0.970
Gabriel	0.550	0.250	0.110	1.000	0.860	0.480	1.000	1.000	0.970

Tablo-34: Grup sayısının 5 ve örneklem büyüklüklerinin eşit olduğu durum için bütün çift için güç değerleri

Çoklu Karşılaştırma Testleri	$n_1=n_2=n_3=n_4=n_5=10$			$n_1=n_2=n_3=n_4=n_5=30$			$n_1=n_2=n_3=n_4=n_5=100$		
	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$
LSD	0.560	0.100	0.010	1.000	0.910	0.300	1.000	1.000	0.990
Bonferroni	0.060	0.000	0.000	0.980	0.440	0.000	1.000	1.000	0.930
Dunn-Sidak	0.080	0.000	0.000	0.980	0.440	0.000	1.000	1.000	0.930
Scheffe	0.030	0.000	0.000	0.960	0.220	0.000	1.000	1.000	0.930
REGW - F	0.340	0.030	0.000	0.990	0.680	0.110	1.000	1.000	0.990
REGW - Q	0.330	0.020	0.000	0.990	0.670	0.110	1.000	1.000	0.990
SNK	0.530	0.100	0.010	1.000	0.890	0.270	1.000	1.000	0.990
Tukey'in a	0.120	0.000	0.000	0.980	0.500	0.010	1.000	1.000	0.940
Tukey'in b	0.320	0.020	0.000	0.990	0.630	0.100	1.000	1.000	0.990
Duncan	0.560	0.100	0.010	1.000	0.910	0.300	1.000	1.000	0.990
Hochberg'in GT2	0.080	0.000	0.000	0.980	0.440	0.000	1.000	1.000	0.930
Gabriel	0.080	0.000	0.000	0.980	0.440	0.000	1.000	1.000	0.930

Tablo-35: Grup sayısının 7 ve örneklem büyüklüklerinin eşit olduğu durum için bütün çift için güç değerleri

Çoklu Karşılaştırma Testleri	$n_1=n_2=n_3=n_4=n_5=n_6=n_7=10$			$n_1=n_2=n_3=n_4=n_5=n_6=n_7=30$			$n_1=n_2=n_3=n_4=n_5=n_6=n_7=100$		
	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$
LSD	0.380	0.020	0.000	1.000	0.920	0.160	1.000	1.000	1.000
Bonferroni	0.000	0.000	0.000	0.950	0.200	0.000	1.000	1.000	0.880
Dunn-Sidak	0.000	0.000	0.000	0.950	0.200	0.000	1.000	1.000	0.890
Scheffe	0.000	0.000	0.000	0.820	0.010	0.000	1.000	0.990	0.580
REGW - F	0.090	0.000	0.000	1.000	0.600	0.010	1.000	1.000	0.970
REGW - Q	0.090	0.000	0.000	1.000	0.590	0.010	1.000	1.000	0.970
SNK	0.380	0.020	0.000	1.000	0.910	0.150	1.000	1.000	1.000
Tukey'in a	0.010	0.000	0.000	0.950	0.240	0.000	1.000	1.000	0.920
Tukey'in b	0.090	0.000	0.000	1.000	0.580	0.000	1.000	1.000	0.970
Duncan	0.380	0.020	0.000	1.000	0.920	0.150	1.000	1.000	1.000
Hochberg'in GT2	0.000	0.000	0.000	0.950	0.200	0.000	1.000	1.000	0.890
Gabriel	0.000	0.000	0.000	0.950	0.200	0.000	1.000	1.000	0.890

Tablo-36: Grup sayısının 3 ve örneklem büyüklüklerinin farklı olduğu durum için bütün çift için güç değerleri

Çoklu Karşılaştırma Testleri	$n_1=10, n_2=12, n_3=14$			$n_1=30, n_2=35, n_3=40$			$n_1=100, n_2=110, n_3=120$		
	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$
LSD	0.890	0.470	0.260	1.000	0.980	0.800	1.000	1.000	1.000
Bonferroni	0.790	0.290	0.110	1.000	0.960	0.600	1.000	1.000	1.000
Dunn-Sidak	0.800	0.300	0.110	1.000	0.960	0.610	1.000	1.000	1.000
Scheffe	0.770	0.270	0.090	1.000	0.960	0.580	1.000	1.000	0.990
REGW - F	0.890	0.470	0.260	1.000	0.980	0.760	1.000	1.000	1.000
REGW - Q	0.860	0.390	0.160	1.000	0.970	0.710	1.000	1.000	1.000
SNK	0.860	0.420	0.190	1.000	0.980	0.730	1.000	1.000	1.000
Tukey'in a	0.840	0.300	0.120	1.000	0.960	0.640	1.000	1.000	1.000
Tukey'in b	0.860	0.350	0.130	1.000	0.970	0.700	1.000	1.000	1.000
Duncan	0.860	0.460	0.250	1.000	0.980	0.770	1.000	1.000	1.000
Hochberg'in GT2	0.810	0.300	0.110	1.000	0.960	0.610	1.000	1.000	1.000
Gabriel	0.810	0.300	0.110	1.000	0.960	0.610	1.000	1.000	1.000

Tablo-37: Grup sayısının 5 ve örneklem büyüklüklerinin farklı olduğu durum için bütün çift için güç değerleri

Çoklu Karşılaştırma Testleri	$n_1=10, n_2=12, n_3=14, n_4=16, n_5=18$			$n_1=30, n_2=35, n_3=40, n_4=45, n_5=50$			$n_1=100, n_2=110, n_3=120, n_4=130, n_5=140$		
	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$
LSD	0.820	0.300	0.000	1.000	0.940	0.580	1.000	1.000	1.000
Bonferroni	0.370	0.000	0.000	1.000	0.740	0.070	1.000	1.000	1.000
Dunn-Sidak	0.370	0.000	0.000	1.000	0.740	0.070	1.000	1.000	1.000
Scheffe	0.230	0.000	0.000	0.990	0.590	0.030	1.000	1.000	0.970
REGW - F	0.660	0.100	0.000	1.000	0.880	0.380	1.000	1.000	1.000
REGW - Q	0.600	0.050	0.000	1.000	0.840	0.320	1.000	1.000	1.000
SNK	0.820	0.260	0.010	1.000	0.940	0.590	1.000	1.000	1.000
Tukey'in a	0.470	0.020	0.000	1.000	0.750	0.110	1.000	1.000	1.000
Tukey'in b	0.650	0.080	0.000	1.000	0.890	0.360	1.000	1.000	1.000
Duncan	0.820	0.260	0.010	1.000	0.940	0.600	1.000	1.000	1.000
Hochberg'in GT2	0.370	0.000	0.000	1.000	0.740	0.070	1.000	1.000	1.000
Gabriel	0.370	0.000	0.000	1.000	0.740	0.070	1.000	1.000	1.000

Tablo-38: Grup sayısının 7 ve örneklem büyüklüklerinin farklı olduğu durum için bütün çift için güç değerleri

Çoklu Karşılaştırma Testleri	$n_1=10, n_2=12, n_3=14, n_4=16, n_5=18, n_6=20, n_7=22$			$n_1=30, n_2=35, n_3=40, n_4=45, n_5=50, n_6=55, n_7=60$			$n_1=100, n_2=110, n_3=120, n_4=130, n_5=140, n_6=150, n_7=160$		
	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$
LSD	0.850	0.210	0.010	1.000	0.970	0.500	1.000	1.000	1.000
Bonferroni	0.270	0.000	0.000	1.000	0.650	0.030	1.000	1.000	0.970
Dunn-Sidak	0.270	0.000	0.000	1.000	0.650	0.030	1.000	1.000	0.970
Scheffe	0.060	0.000	0.000	0.970	0.310	0.000	1.000	1.000	0.890
REGW - F	0.690	0.030	0.000	1.000	0.900	0.170	1.000	1.000	1.000
REGW - Q	0.630	0.010	0.000	1.000	0.860	0.150	1.000	1.000	1.000
SNK	0.860	0.220	0.010	1.000	0.980	0.470	1.000	1.000	1.000
Tukey'in a	0.340	0.000	0.000	1.000	0.710	0.060	1.000	1.000	0.970
Tukey'in b	0.660	0.010	0.000	1.000	0.890	0.140	1.000	1.000	1.000
Duncan	0.860	0.220	0.010	1.000	0.980	0.470	1.000	1.000	1.000
Hochberg'in GT2	0.270	0.000	0.000	1.000	0.650	0.030	1.000	1.000	0.970
Gabriel	0.280	0.000	0.000	1.000	0.650	0.040	1.000	1.000	0.970

Tablo-39: Grup sayısının 3 ve örneklem büyüklüklerinin eşit olduğu durum için yanlış bulgu oranı

Çoklu Karşılaştırma Testleri	$n_1=n_2=n_3=10$			$n_1=n_2=n_3=30$			$n_1=n_2=n_3=100$		
	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$
LSD	0.034	0.052	0.027	0.038	0.053	0.050	0.038	0.024	0.078
Bonferroni	0.014	0.039	0.023	0.005	0.016	0.023	0.010	0.010	0.025
Dunn-Sidak	0.014	0.039	0.023	0.005	0.016	0.023	0.010	0.010	0.025
Scheffe	0.014	0.041	0.024	0.005	0.016	0.023	0.010	0.005	0.025
REGW - F	0.035	0.063	0.019	0.038	0.053	0.052	0.038	0.024	0.078
REGW - Q	0.035	0.054	0.019	0.038	0.053	0.053	0.038	0.024	0.078
SNK	0.035	0.054	0.019	0.038	0.053	0.053	0.038	0.024	0.078
Tukey'in a	0.013	0.049	0.022	0.010	0.016	0.022	0.015	0.010	0.025
Tukey'in b	0.025	0.057	0.021	0.024	0.026	0.029	0.029	0.015	0.043
Duncan	0.034	0.053	0.030	0.038	0.053	0.051	0.038	0.024	0.078
Hochberg'in GT2	0.014	0.039	0.023	0.005	0.026	0.023	0.010	0.010	0.025
Gabriel	0.014	0.039	0.023	0.005	0.016	0.023	0.010	0.010	0.025

Tablo-40: Grup sayısının 5 ve örneklem büyüklüklerinin eşit olduğu durum için yanlış bulgu oranı

Çoklu Karşılaştırma Testleri	$n_1=n_2=n_3=n_4=n_5=10$			$n_1=n_2=n_3=n_4=n_5=30$			$n_1=n_2=n_3=n_4=n_5=100$		
	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$
LSD	0.004	0.011	0.008	0.003	0.008	0.006	0.007	0.006	0.003
Bonferroni	0.000	0.002	0.000	0.000	0.000	0.000	0.001	0.001	0.000
Dunn-Sidak	0.000	0.002	0.000	0.000	0.000	0.000	0.001	0.001	0.000
Scheffe	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
REGW - F	0.002	0.006	0.002	0.001	0.002	0.003	0.002	0.001	0.001
REGW - Q	0.002	0.006	0.000	0.001	0.002	0.003	0.002	0.001	0.001
SNK	0.004	0.010	0.004	0.003	0.008	0.006	0.007	0.006	0.003
Tukey'in a	0.000	0.002	0.000	0.001	0.001	0.000	0.001	0.001	0.000
Tukey'in b	0.003	0.006	0.000	0.001	0.002	0.003	0.002	0.001	0.001
Duncan	0.004	0.011	0.003	0.003	0.008	0.006	0.007	0.006	0.003
Hochberg'in GT2	0.000	0.002	0.000	0.000	0.000	0.000	0.001	0.001	0.000
Gabriel	0.000	0.002	0.000	0.000	0.000	0.000	0.001	0.001	0.000

Tablo-41: Grup sayısının 7 ve örneklem büyüklüklerinin eşit olduğu durum için yanlış bulgu oranı

Çoklu Karşılaştırma Testleri	$n_1=n_2=n_3=n_4=n_5=n_6=n_7=10$			$n_1=n_2=n_3=n_4=n_5=n_6=n_7=30$			$n_1=n_2=n_3=n_4=n_5=n_6=n_7=100$		
	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$
LSD	0.003	0.002	0.002	0.002	0.002	0.002	0.003	0.002	0.002
Bonferroni	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001
Dunn-Sidak	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001
Scheffe	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
REGW - F	0.001	0.000	0.000	0.001	0.001	0.001	0.001	0.000	0.001
REGW - Q	0.001	0.000	0.000	0.001	0.001	0.001	0.001	0.000	0.002
SNK	0.003	0.002	0.001	0.002	0.002	0.002	0.003	0.002	0.002
Tukey'in a	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001
Tukey'in b	0.001	0.000	0.000	0.001	0.001	0.001	0.001	0.000	0.001
Duncan	0.003	0.002	0.002	0.002	0.002	0.002	0.003	0.002	0.002
Hochberg'in GT2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001
Gabriel	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001

Tablo-42: Grup sayısının 3 ve örneklem büyüklüklerinin farklı olduğu durum için yanlış bulgu oranı

Çoklu Karşılaştırma Testleri	$n_1=10, n_2=12, n_3=14$			$n_1=30, n_2=35, n_3=40$			$n_1=100, n_2=110, n_3=120$		
	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$
LSD	0.026	0.031	0.082	0.005	0.020	0.023	0.029	0.010	0.029
Bonferroni	0.000	0.034	0.036	0.000	0.010	0.014	0.010	0.005	0.010
Dunn-Sidak	0.000	0.034	0.036	0.000	0.010	0.014	0.010	0.005	0.010
Scheffe	0.000	0.035	0.020	0.000	0.010	0.007	0.005	0.005	0.010
REGW - F	0.026	0.035	0.061	0.005	0.020	0.024	0.029	0.010	0.029
REGW - Q	0.016	0.040	0.067	0.005	0.020	0.025	0.029	0.010	0.029
SNK	0.042	0.047	0.091	0.010	0.025	0.025	0.029	0.010	0.029
Tukey'in a	0.000	0.032	0.034	0.000	0.010	0.013	0.010	0.005	0.010
Tukey'in b	0.011	0.040	0.085	0.005	0.015	0.025	0.024	0.010	0.024
Duncan	0.041	0.040	0.078	0.010	0.025	0.024	0.029	0.010	0.029
Hochberg'in GT2	0.000	0.034	0.036	0.000	0.010	0.014	0.010	0.005	0.010
Gabriel	0.000	0.033	0.036	0.000	0.010	0.014	0.010	0.005	0.010

Tablo-43: Grup sayısının 5 ve örneklem büyüklüklerinin farklı olduğu durum için yanlış bulgu oranı

Çoklu Karşılaştırma Testleri	$n_1=10, n_2=12, n_3=14, n_4=16, n_5=18$			$n_1=30, n_2=35, n_3=40, n_4=45, n_5=50$			$n_1=100, n_2=110, n_3=120, n_4=130, n_5=140$		
	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$
LSD	0.009	0.006	0.010	0.002	0.007	0.007	0.007	0.004	0.006
Bonferroni	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.001
Dunn-Sidak	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.001
Scheffe	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.001
REGW - F	0.002	0.004	0.005	0.001	0.004	0.002	0.000	0.001	0.003
REGW - Q	0.001	0.003	0.004	0.000	0.003	0.001	0.000	0.000	0.003
SNK	0.011	0.008	0.013	0.003	0.009	0.009	0.007	0.007	0.009
Tukey'in a	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.001
Tukey'in b	0.007	0.004	0.009	0.002	0.004	0.002	0.002	0.001	0.004
Duncan	0.011	0.009	0.015	0.003	0.009	0.009	0.007	0.007	0.009
Hochberg'in GT2	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.001
Gabriel	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.001

Tablo-44: Grup sayısının 7 ve örneklem büyüklüklerinin farklı olduğu durum için yanlış bulgu oranı

Çoklu Karşılaştırma Testleri	$n_1=10, n_2=12, n_3=14, n_4=16, n_5=18, n_6=20, n_7=22$			$n_1=30, n_2=35, n_3=40, n_4=45, n_5=50, n_6=55, n_7=60$			$n_1=100, n_2=110, n_3=120, n_4=130, n_5=140, n_6=150, n_7=160$		
	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$	$\sigma^2=2$	$\sigma^2=4$	$\sigma^2=8$
LSD	0.001	0.001	0.002	0.000	0.004	0.002	0.002	0.002	0.001
Bonferroni	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Dunn-Sidak	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Scheffe	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
REGW - F	0.000	0.000	0.001	0.000	0.002	0.001	0.001	0.000	0.000
REGW - Q	0.000	0.000	0.000	0.000	0.002	0.001	0.000	0.000	0.000
SNK	0.003	0.004	0.004	0.002	0.006	0.005	0.005	0.003	0.003
Tukey'in a	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Tukey'in b	0.002	0.001	0.001	0.000	0.004	0.002	0.002	0.001	0.000
Duncan	0.003	0.003	0.004	0.002	0.006	0.005	0.005	0.003	0.003
Hochberg'in GT2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Gabriel	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

5. TARTIŞMA VE SONUÇ

Çoklu karşılaştırma testlerinin oluşturulması 1940'lı yılların sonu ve 1950'li yılların başına dayanmaktadır. David B. Duncan, Charles W. Dunnett, HO. Hartley, Henry Scheffe, SN. Roy ve John W. Tukey'in bu konuya önemli katkıları olmuştur. Bazılarının çalışmaları daha önce Fisher, Student ve diğerlerinin yaptığı çalışmalara dayanmaktadır (3). Özellikle Tukey'in 1953 yılında hazırladığı, ancak basılmamış notları bu konu üzerine önemli katkılar sağlamıştır (3). Miller'in, 1966 yılında yayımladığı kitabı sonrasında da çoklu karşılaştırma testleri popüler hale gelmiştir (2, 3). Miller'in kitabı sayesinde 1970'li yıllardan itibaren bu konu üzerinde birçok çalışma gerçekleştirilmiştir (3). Bu çalışmalarla birlikte birçok yeni çoklu karşılaştırma testi önerilmiştir. Bilgisayar teknolojisinin ilerlemesiyle birlikte geliştirilen istatistiksel yazılımlarda bu testlerden bazıları yer almıştır. Çalışmamızda, istatistik yazılımlarında yer alan ve literatürde sıklıkla kullanılan 12 adet çoklu karşılaştırma testi yer almıştır.

Çoklu karşılaştırma testleri, temellendiği istatistik ve dayandığı varsayımlar bakımından birbirinden farklıdır (9). Bu farklılıkların araştırmacılar tarafından bilinmemesi veya anlaşılabilmesi sonucunda testlerin seçilmesine ilişkin hatalar oluşmaktadır. Birçok çoklu karşılaştırma testinin, test istatistiğinin gerektirdiği varsayımlar sağlanmadığında kullanılması uygun değildir (11). Ayrıca araştırmacılar, karşılaştırılacak grupların belli bir yapıda olduğunu göz ardı ederek tüm ortalamaları ikili olarak karşılaştırmak isteyebilir. Bu sıklıkla karşılaşılan hatalardandır (4). Eğer ortalamalar arasında bir yapı söz konusu ise ortogonal bağıntılar kullanılmalıdır (4). Ortalama çiftlerinin karşılaştırılması belli bir yapı olmadığında mümkündür (3, 4). Çalışmamızda herhangi bir yapının var olmadığı ve ortalama çiftlerinin karşılaştırıldığı durumlar göz önünde bulundurulmuştur. Çoklu karşılaştırma testinin seçilmesi grup varyanslarına bağlı olarak da değişmektedir. Homojen veya heterojen varyans durumunda farklı çoklu karşılaştırma testleri kullanılmaktadır.

Normallik ve homojen varyans varsayımı altında çalışmamızda yer alan çoklu karşılaştırma testleri, farklı varyans değerleri, eşit ve farklı örneklem büyüklükleri ve çeşitli grup sayıları için simülasyon yardımıyla incelenmesi amaçlanmıştır. Testlerin birbirleriyle karşılaştırılmasında deneysel hata, karşılaştırma başına hata, güç ve yanlış bulgu oranı gibi istatistiksel kriterler göz önüne alınmıştır.

Ortalamaların farklı olup olmadığını test eden araştırmacılar, genellikle anlamlılık derecesini belirleyerek anlamlı farklılık olup olmadığı kararıyla karşı karşıya kalırlar. Anlamlılık düzeyi, sıfır hipotezinin doğru olduğu durumda en yüksek reddedilme olasılığını

belirler. Tip I hata azaldıkça tip II hatanın artmasına rağmen, araştırmacılar sıfır hipotezinin reddedilmesinin gerçek bir farklılık gösterdiğinden emin olabilirler (11). Fakat bu durum, tip II hata olasılığını artırır. Tip I ve tip II hatalarının göreceli önemini belirleme işiyle karşı karşıya kalan araştırmacılar kabul edilebilir bir anlamlılık düzeyi belirlerler, örneğin $\alpha=0.05$ (11). Fakat tip I hataların nasıl kontrol edileceğine karar vermek, çoklu karşılaştırma testleri yapıldığında daha zordur (43, 44). Çünkü çoklu karşılaştırma testleri kullanıldığında tip I hatanın nasıl kontrol edileceğinin seçimi, etkilerin istatistiksel olarak anlamlı bulunmasıyla sonuçlanabilir. Tip I hatayı kontrol etme biçimi, testin çeşitlilik sorunu ile nasıl başa çıkacağıyla ilişkili olmalıdır (11).

Deneylerdeki çeşitliliğe nasıl yaklaşılması gerektiği uzun yıllardır pek çok literatürde tartışılmıştır. Hatalı pozitiflerin her durumda korunması gerektiğine inananlar çeşitlilik sorunuyla başa çıkmak için hesaplanan tüm deneylerde α 'yı sembol olarak seçmişlerdir (45). Bu çeşit bir kontrol, literatürde deneysel hata kontrolü olarak yer bulur (11).

Buna karşıt grubu oluşturanlar ise, katı tip I hata kontrolünün istatistiksel gücün kaybıyla sonuçlandığını iddia ederler. Bunun sonucu olarak önemli grup etkileri fark edilmez (35, 46). Bu gruptakiler, hata oranının karşılaştırma başına gelecek miktar hesaplanarak oluşturulmasını savunurlar (11, 28). Bu hata oranı, karşılaştırma başına hata oranı olarak adlandırılır. Bu oranın, 0.05 anlamlılık değerine sahip olmasını önerirler, böylelikle genel hata oranının hesaplanan test sayısı ile artmasına olanak sağlanmış olur (11).

Geçmişte, birçok araştırmacı tarafından deneysel hata oranının kontrol edilmesi tavsiye edilmiştir (11, 28, 47). Seaman ve arkadaşlarına göre (48), deneysel hata oranının kontrolünün sağlanması en yaygın ve kabul gören yaklaşımdır. Ayrıca Keselman ve arkadaşları (26), araştırmacıların yaklaşık %85'inin ortalama çiftlerinin karşılaştırma yöntemini kullanırken deneysel hata kontrolünü baz aldıklarını belirtmişlerdir.

Bender ve Lange (29), hata oranlarının kontrolü ile ilgili olarak öncelikle grupların hangi deneye ait olduğunun belirlenmesi gerektiğini belirtmişlerdir. Örneğin, geliştirilen 3 yeni tedavinin standart bir tedavi ile karşılaştırılması isteniyorsa 4 grup için olası karşılaştırma sayısı 6 olacaktır. Çalışmanın temel amacı, yeni tedavi gruplarının standart tedavi ile karşılaştırılması isteniyorsa ve tedavi gruplarının birbiriyle karşılaştırılması ikinci bir analiz olarak düşünülüyorsa bu çalışma birbirine bağlı karşılaştırmaları içeren iki deneyden oluşur. Bu durumda Bender ve Lange (29), her iki deneme için ayrı çeşitlilik düzeltmesi gerektiğini belirtmişlerdir. Yazarlar, çoklu karşılaştırma testlerinin bulgularını tek bir sonuç içerisinde özetlenebildiği durumda deneysel hatanın kontrolünün mantıklı olacağını belirtmişlerdir. Örneğin, her bir tedavi ile standart tedavi arasında istatistiksel farklılık varsa üç yeni tedavinin

standart tedaviden farklı olduğu yorumu deneysel hata kontrolü varlığında yapılmalıdır (29). Aksi durumda, sonuç üzerindeki tip I hatanın kontrolü sağlanamaz.

Literatürlerdeki tartışmaların büyük çoğunluğu deneysel hata oranı ve karşılaştırma başına hata oranı (11, 13, 18, 28) üzerine olmasına rağmen yanlış bulgu oranı gibi farklı hata oranı da önerilmiştir. Benjamini ve Hochberg (38) tarafından önerilen yanlış bulgu oranı, toplam reddedilen hipotezler içerisinde hatalı olarak reddedilen hipotezlerin oranının beklenen değeri olarak tanımlanmaktadır. Benjamini ve Hochberg (38), yanlış bulgu oranının deneysel hata oranı ve karşılaştırma başına hata oranından daha makul olduğunu belirtmişlerdir.

Çoklu karşılaştırma testlerinin uzun geçmişi göz önüne alındığında yanlış bulgu oranı daha çok yeni olmakla birlikte buna rağmen oldukça fazla ilgi görmüştür. Çünkü pratikte karşılaştırma başına hata oranı ile deneysel hata oranı arasında bir uzlaşma ortaya koymaktadır (12). Benjamini ve Hochberg (38)'e göre, yanlış bulgudan kasıt sıfır hipotezinin hatalı reddedilmesidir. Farklı olan nokta, Benjamini ve Hochberg'in bunu toplam bulgular açısından tanımlamalarıdır. Bir başka deyişle, reddedilmiş hipotezlerin; bazıları hatalı olarak reddedilmiş olsa da, doğru olarak reddedilmiş olanlarla hesaba katmalarıdır. Buradaki amaç, toplam bulgularda yanlış bulguların oranını 0.05 seviyesinde tutmaktır (12).

Tip I (α) ve tip II (β) hatalar arasında ilişki vardır. Bu ilişki, tip I hata azalırken tip II hatanın artması şeklindedir. Tip II hatanın artması ise testin gücünün düşmesine neden olur. Bu sebeple tip I hatanın kontrolü sağlanmak istendiğinde araştırmacılar testin gücünden ödün vermek zorunda kalırlar. Güç, yanlış sıfır hipotezinin reddedilme olasılığıdır. Klasik tanımının yanı sıra çoklu karşılaştırma testleri için tek bir güç değeri yerine alt hipotezlerin reddedilme durumlarında göre farklı güçler türleri de hesaplanabilir. Einot ve Gabriel (7), hipotez kümesini ikili, üçlü, dördü ve beşli hipotezlerin güç hesaplamaları şeklinde özetlemişlerdir. Daha sonra Ramsey tarafından yapılan çalışmalarda ise güç tanımları daha farklı şekilde belirtilmiştir (19, 39-41). Bu çalışmalardan sonra çoklu karşılaştırma testlerinin güç incelemesinde yaygın olarak Ramsey'in tanımları kullanılmıştır. Bütün çiftler için güç ve herhangi bir çift için güç tanımları sırasıyla bütün yanlış hipotezlerin reddedildiği ve en az bir yanlış hipotezin reddedildiği güç olarak tanımlanır. Bunların dışında her bir hipotez ile ayrı olarak ilgilenen her bir çift için güç tanımı klasik güç tanımına en yakın olan tanımdır (41). Belirli bir örneklem büyüklüğü için herhangi bir çift için ve bütün çiftler için güç değerleri gerçekten doğru ve yanlış hipotez sayılarına bağlıdır.

Çalışmamızda çoklu karşılaştırma testleri, deneysel ve karşılaştırma başına hata olmak üzere iki hata oranına göre karşılaştırıldı. Ancak literatürde yaygın olarak deneysel hatanın kontrolü önerilmektedir. Bu sebeple testlerin karşılaştırma başına hata oranı bakımından

incelendiği literatüre daha az rastlanmaktadır. Çalışmamızda, çoklu karşılaştırma testlerini deneysel hata oranı bakımından kıyasladığımızda örneklem büyüklükleri eşit olduğu durumda Duncan ve özellikle de LSD testinin hata oranı oldukça yüksek düzeydedir. Bu testler için grup sayısının artması ile birlikte deneysel hata oranında çok yüksek düzeylerde artış görülmektedir. Çalışmamızla uyumlu olarak Thomas (49), LSD ve Duncan testlerinin hata oranlarını yüksek olarak bulmuştur. Dayton (50), Curran-Everett (51), Maxwell ve Delaney (12) ise yaptıkları çalışmalarda, hata oranlarının yüksek olmasından dolayı LSD, Duncan ve SNK testlerini önermemektedir. Toothaker (13), Harter (2), Klockars ve Sax (18)'a göre LSD ve Duncan testlerinin tip I hata oranının diğer testlere göre yüksek olmasından dolayı daha az tutucu olarak belirtmişlerdir.

İncelediğimiz testler içinde Scheffe testi en düşük deneysel hata oranına sahip testtir ve grup sayısının değişiminden etkilenmediği görülmektedir. Bu sonuçla benzer olarak Toothaker (13), Klockars ve Sax (18) Scheffe testi için tip I hatanın korunduğunu belirtmişlerdir. Bernhardson (52) yaptığı çalışmada LSD, Tukey'in a, SNK, Duncan ve Scheffe testlerini deneysel hata oranı bakımından karşılaştırdığında LSD ve Duncan'ın yüksek, Scheffe testinin ise düşük bulmuştur.

Çalışmamızda Bonferroni, Dunn-Sidak, Hochberg'in GT2 ve Gabriel testleri grup sayısı düşük olduğunda düşük deneysel hata oranına sahip iken eşit örneklem büyüklükleri için grup sayısının artması ile birlikte bu oranda artış görülmektedir. REGW-F, REGW-Q, SNK, Tukey'in a ve b testleri ise grup sayısı az olduğunda diğer testlere göre daha yüksek deneysel hata oranına sahip iken grup sayısının artması ile birlikte eşit örneklem büyüklükleri için hata oranlarında da artış görülmektedir. Bu testlerin deneysel hata oranı LSD ve Duncan testlerinden daha düşük düzeyde iken Bonferroni, Dunn-Sidak, Scheffe, Hochberg'in GT2 ve Gabriel testlerinden daha yüksek düzeydedir. SNK testi için Dayton (50), Curran-Everett (51), Maxwell ve Delaney (12) deneysel hata oranının yüksek olmasından dolayı önermezlerken, Bender ve Lange (29), çalışmalarında incelediği Duncan, SNK, REGW-Q ve REGW-F testlerinden Duncan testi hariç tip I hatanın korunduğunu belirtmişlerdir. Seaman ve arkadaşları (48) tarafından gerçekleştirilen bir başka çalışmada Duncan testinin deneysel hata oranının yüksek olduğunu, LSD ve SNK testlerinin ise grup sayısının az olduğu durumlarda bu oranın kontrol altında tutulduğunu ancak grup sayısının artması ile yükseldiğini belirtmişlerdir. Toothaker (13) ve Maxwell (12), çalışmamızın sonuçlarıyla uyumlu olarak Bonferroni, Dunn-Sidak ve Tukey'in a testlerinin deneysel hatanın kontrolünü sağladığını belirtmişlerdir. Ayrıca Maxwell (12), yine çalışmamızın sonucuyla benzer olarak

Hochberg'in GT2 ve Gabriel testlerinin deneysel hata oranının birbirlerine benzer ve düşük düzeyde olduğunu belirtmiştir.

Çalışmamızda örneklem büyüklükleri farklı olduğu durumda da Duncan ve özellikle de LSD testinin deneysel hata oranı oldukça yüksek düzeyde olduğu ve grup sayısının artmasıyla bu oranda çok yüksek düzeylerde artış olduğu görülmektedir. İncelediğimiz testler içinde Scheffe testi en düşük deneysel hata oranına sahip testtir ve grup sayısının artışından etkilenmediği görülmektedir. Elde edilen sonuçlar Thomas (49), Dayton (50), Curran-Everett (51), Maxwell ve Delaney (12), Toothaker (13), Klockars (18) ve Harter'in (2) yaptığı çalışmalarla uyumludur. Diğer testler için grup sayısının artması ile birlikte deneysel hata oranında bir artış veya azalış görülmemektedir. Örneklem büyüklüklerinin artması veya azalmasına bağlı olarak tüm testlere ait deneysel hata oranında herhangi bir değişim gözlenmemiştir. Ayrıca gruplara ait varyanslardaki değişimin, varyans analizinin temel prensiplerinden homojen varyans koşulunun sağlanmasından dolayı hata oranı üzerinde ciddi bir etkisinin olmadığı düşünülmektedir.

Testleri, karşılaştırma başına hata oranı bakımından incelediğimizde, gerek örneklem büyüklükleri eşit olduğu, gerekse farklı olduğu durumlarda LSD ve özellikle de Duncan testinin diğer testlere göre yüksek hata oranına sahip olduğu görülmüştür. LSD testinin, karşılaştırma başına hata oranı grup sayısının artmasından çok fazla etkilenmemektedir. Fakat Duncan testi için ise grup sayısının artışı ile birlikte karşılaştırma başına hata oranında çok ciddi artışlar gözlenmektedir. Bernhardson (52), yaptığı çalışmada LSD, Tukey'in a, SNK, Duncan ve Scheffe testlerini karşılaştırma başına hata oranı bakımından incelediğinde LSD ve Duncan'ın yüksek, Scheffe testinin ise düşük bulmuştur. Örneklem büyüklüklerinin eşit ve farklı olduğu durumlarda Bonferroni, Dunn-Sidak ve Scheffe testleri genel olarak incelendiğinde en düşük karşılaştırma başına hata oranına sahiptir. Bu testlerde grup sayılarının artmasına bağlı olarak karşılaştırma başına hata oranında çok fazla bir değişim gözlenmemektedir. REGW-F, REGW-Q, SNK, Tukey'in a ve b, Hochberg'in GT2 ve Gabriel testleri Bonferroni, Dunn-Sidak ve Scheffe testleri kadar düşük hata oranına sahip olmasa da LSD ve Duncan testlerinden daha düşük düzeydedir. Menéndez ve arkadaşları (53)'nın yaptığı çalışmada, LSD ve REGW-Q testlerinin karşılaştırma başına hata oranı Tukey a, Scheffe ve SNK testlerinden daha yüksek bulunmuştur. Ancak bizim çalışmamızda REGW-Q testinin oranı LSD testinden daha düşük bulunmuştur. Grup sayısının artması bu testlerin hata oranı üzerinde dikkate değer bir değişim göstermemiştir. Örneklem büyüklükleri ve varyansa bağlı değişimin testlerin karşılaştırma başına hata oranı üzerinde bir etkisinin olmadığı görülmektedir.

Çalışmamızda klasik güç tanımıyla eş değer olan, testlerin her bir çift için gücü incelendiğinde örneklem büyüklükleri eşit ve farklı olduğu durumlarda tüm testlerin varyansın artmasıyla gücü azalırken, örneklem büyüklüğünün artmasıyla güçleri artmaktadır. Ayrıca Bonferroni, Dunn-Sidak ve Scheffe testleri dışındaki testler için grup sayısının artmasıyla her bir çift için güç değeri artmaktadır. Grup sayısının ve örneklem büyüklüğünün az olduğu durumda en güçlü test LSD olarak bulunurken, gücü en düşük olan testler Hochberg GT2 ve Gabriel'dir. Grup sayısının ve örneklem büyüklüklerinin artmasıyla LSD, Duncan ve SNK testleri en yüksek güce sahip iken en düşük güç Scheffe testine aittir. Çalışmamızın sonucuyla benzer olarak Toothaker (13) ve Klockars (18), LSD testinin gücünü yüksek Scheffe testinin gücünü ise düşük olarak belirtmişlerdir. Maxwell ve Delaney (12) ise Duncan testini oldukça güçlü bir test, bir başka deyişle gerçekte var olan farklılıkları tespit etmede oldukça etkili bir test olduğunu belirtmişlerdir (12). Menéndez ve arkadaşlarının (53) yaptığı çalışmada, LSD ve REGW-Q testlerinin her bir test için güç değeri Tukey a, Scheffe ve SNK testlerinden daha yüksek bulunmuştur. Ancak çalışmamızda REGW-Q testinin LSD ve Duncan testleri kadar güçlü bir test olmadığı görülmekle birlikte, SNK testinin gücü bu testlere benzerdir. Stoline (54) ise gerçekleştirdiği çalışmada Tukey testinin Bonferroni, Dunn-Sidak ve Hochberg GT2 testlerinden daha güçlü olduğunu belirtmiştir. Ayrıca Dayton da (50) çalışmasında Scheffe ve Bonferroni testlerinin Tukey'in a testi kadar güçlü olmadığı vurgulamıştır. Çalışmamızda grup sayısı az olduğu durum haricinde LSD, Duncan, SNK ve Scheffe testleri dışındaki testler için güç bakımından aralarında önemli bir farklılık görülmemiştir.

Her bir çift için güç değerinde olduğu gibi herhangi bir çift için ve bütün çiftler için güç değerleri incelendiğinde varyansın artmasıyla birlikte azalmakta ve örneklem büyüklüğü artmasıyla artmaktadır. Grup sayısının ve örneklem büyüklüğünün düşük olduğu durumda LSD ve Duncan testleri en yüksek Scheffe testi en düşük güce sahiptir. Herhangi bir çift için güç değerlendirmesinde grup sayısının artmasıyla güç değeri en yüksek düzeye ulaşmıştır.

Bütün çiftler için güç değeri varyansın artmasıyla birlikte güç azalmakta ve örneklem büyüklüğü artmasıyla güç artmaktadır. Grup sayısının artmasıyla güç değerleri azalmaktadır. Örneklem büyüklüğünün ve grup sayısının az olduğu durumda LSD, Duncan, REGW-F, REGW-Q, SNK, Tukey'in a ve b testleri en yüksek güce sahip iken Scheffe testi en düşük güce sahiptir. Grup sayısının artmasıyla LSD, Duncan ve SNK testleri en yüksek güce sahip iken en düşük güce yine Scheffe testi sahiptir. İncelenen testlerin hem her bir çift için güç hem de bütün çiftler için güç değerlerini karşılaştıran bir literatüre rastlanmamıştır.

Yanlış bulgu oranı bakımından testler incelendiğinde örneklem büyüklükleri ve varyansa bağlı olarak bu oranda herhangi bir değişiklik görülmemektedir. Ancak grup sayısının artmasıyla yanlış bulgu oranı azalmaktadır. Düşük grup sayısı olduğunda LSD, REGW-F, REGW-Q, SNK ve Duncan testlerinin yanlış bulgu oranını diğer testlere göre daha yüksek düzeyde olduğu görülmüştür. Yanlış bulgu oranı, Benjamini ve Hochberg tarafından 1995 yılında önerilmiştir (38). O tarihten itibaren kullanımı yaygınlaşmış olsa da çoklu karşılaştırma testlerini homojen varyans, eşit ve farklı örneklem büyüklükleri altında çeşitli grup sayıları için kıyaslayan bir literatüre rastlanmamıştır.

Çoklu karşılaştırma testlerinden hangisini seçileceğine ilişkin yapılan tartışmalarda Keselman ve arkadaşları (11) karşılaştırılmak istenen grup sayısı 4 veya 5 olduğunda ortalama çiftlerinin karşılaştırılmasında REGW-Q testinin kullanılmasını önermişlerdir. Eğer araştırmacılar, ortalama çiftleri arasındaki farkın eş zamanlı güven aralıklarının belirlenmesini istiyorsa Tukey'in a testinin kullanılması önermişlerdir (11). Jaccard ve Guilamo-Ramos (43) ise varyans analizinde anlamlı çıkan F testi sonucunda ikili karşılaştırmalarda Tukey veya Bonferroni testi kullanılması önermişlerdir. Rafter ve arkadaşları (55) ise bütün ortalama çiftlerinin karşılaştırılmasında Tukey testini, bütün ortalama çiftleri olmasada bazılarının karşılaştırılmasında Bonferroni, Dunn-Sidak, Hochberg'in GT2 testlerini önermiştir.

Literatürde çoklu karşılaştırma testlerinin önerildiği birçok çalışma mevcuttur. Ancak bu testlerin birçoğu çok nadir olarak kullanılmaktadır. Çalışmamızda incelenen testler belirlenirken farklı istatistik yazılımlarında ve literatürde yaygın olarak yer alması dikkate alındı. Çalışmamızın diğer çalışmalardan farklılığı, testlerin homojen varyans, çeşitli grup sayıları, eşit ve farklı örneklem büyüklükleri için belirlenen istatistiksel kriterler doğrultusunda birbirleriyle karşılaştırılmasıdır ve benzer çalışmalar incelendiğinde daha az sayıda testin karşılaştırıldığı görülmektedir. Deneysel hata oranı bakımından incelediğimizde çalışmamızın sonucu genellikle literatürle uyumlu çıkmıştır (2, 13, 18, 29, 49-52). Bazı literatürlere göre farklılık gösteren kısmı deneysel hata oranı için SNK testinin LSD ve Duncan testleriyle (12, 50, 51), karşılaştırma başına hata oranı için REGW-Q testinin LSD testiyle (53) birlikte yüksek hata oranına sahip olduğunun belirtilmesidir. Çalışmamızın sonucunda hem SNK testinin hem de REGW-Q testinin hata oranlarının yüksek olduğu, ancak LSD ve Duncan testleri kadar yüksek olmadığı görülmüştür. Benzer olarak karşılaştırma başına hata oranı için incelediğimizde LSD ve Duncan en yüksek hata oranına sahip iken Scheffe testinin yanı sıra Bonferroni ve Dunn-Sidak testleri en düşük hata oranına sahiptir. Klasik güç tanımıyla eş değer olan her bir çift için güç değerlendirmesinde literatürle benzer sonuçlara ulaşılmıştır (12, 13, 18). LSD, Duncan ve SNK testleri en yüksek

güce sahip iken Scheffe testi en düşük güce sahip olarak bulunmuştur. Diğer güç tanımları için ve yanlış bulgu oranı için yapılan karşılaştırmada bütün testlerin birbiriyle kıyaslandığı bir literatüre rastlanmamıştır. Çalışmamızda herhangi bir çift için güç değerleri, grup sayısının ve örneklem büyüklüğünün düşük olduğu durumda LSD ve Duncan testleri için yüksek ve Scheffe testi için düşük olarak bulunurken grup sayısının artmasıyla güç değerleri tüm testler için en yüksek düzeye ulaşmıştır. Bütün çiftler için güç değeri bakımından incelendiğinde grup sayısının ve örneklem büyüklüğünün düşük olduğu durumda LSD, Duncan, REGW–F, REGW–Q, SNK, Tukey’in a ve b testleri en yüksek güce sahip iken Scheffe testi en düşük güce sahiptir. Grup sayısının artmasıyla LSD, Duncan ve SNK testleri en yüksek güce sahip iken en düşük güce yine Scheffe testi sahiptir. Yanlış bulgu oranı bakımından testler incelendiğinde düşük grup sayısı olduğunda LSD, REGW–F, REGW–Q, SNK ve Duncan testlerinin yanlış bulgu oranı diğer testlere göre daha yüksek düzeyde olduğu görülmüştür.

Yapılan çalışmalar sonucunda çoklu karşılaştırma testlerinden hangisinin en iyi olduğu ve kullanılması gerektiği konusunda ortak bir görüş birliğine varılamamıştır. Bu noktada John W. Tukey’in bir sözü göz önünde bulundurulmalıdır (51): “Bütün deneyler için bir tek çoklu karşılaştırma testinin araştırılmaması konusunda cesaretli olmalıyız”. Çalışmamızda LSD ve Duncan testinin yüksek hata oranına sahip olması, Scheffe testinin de düşük güce sahip olmasından dolayı ortalama çiftlerinin karşılaştırılmasında önerilmemektedir. Diğer testler arasında bu kriterler bakımından büyük farklılıklar görülmemektedir.

6. KAYNAKLAR

1. SAHAI H, AGEEL MI. The analysis of variance: Fixed, random, and mixed models, Birkhauser, Boston, 2000.
2. HARTER H. Early history of multiple comparison tests, Editor: KRISHNAIAH P. Handbook of Statistics, Vol. 1, North-Holland, Amsterdam, 617-622, 1980.
3. HOCHBERG Y, TAMHANE AC. Multiple comparison procedures, John Wiley & Sons Inc. , New York, 1987.
4. LOWRY S. Use and misuse of multiple comparisons in animal experiments. Journal of Animal Science, 70(6): 1971-1977, 1992.
5. HSU JC. Multiple comparisons: theory and methods, Chapman & Hall/CRC, London, 1996.
6. RYAN TA. Multiple comparison in psychological research. Psychological Bulletin, 56(1): 26-47, 1959.
7. EINOT I, GABRIEL KR. A study of the powers of several methods of multiple comparisons. Journal of the American Statistical Association, 70(351): 574-583, 1975.
8. CONAGIN A, BARBIN DB. Bonferroni's and Sidak's modified tests. Scientia Agricola, 63: 70-76, 2006.
9. DEMIRHAN H, DOLGUN NA, DEMIRHAN YP, DOLGUN MÖ. Performance of some multiple comparison tests under heteroscedasticity and dependency. Journal of Statistical Computation and Simulation, 80(10): 1083-1100, 2010.
10. MONTGOMERY DC. Design and analysis of experiments, 5th edition, John Wiley & Sons Inc, New York, 2001.
11. KESELMAN H, CRIBBIE RA, HOLLAND B. Pairwise multiple comparison test procedures: an update for clinical child and adolescent psychologists. Journal of Clinical Child & Adolescent Psychology, 33(3): 623-645, 2004.
12. MAXWELL SE, DELANEY HD. Designing experiments and analyzing data: A model comparison perspective, Lawrence Erlbaum Associates, London, 2004.
13. TOOTHAKER LE. Multiple comparison procedures, Sage Publications, Newbury Park, California, 1993.
14. OTT RL, LONGNECKER M. An introduction to statistical methods and data analysis, 5th Edition, Duxbury Press, Pacific Grove, California, 2001.
15. HICKS C. Fundamental concepts in the design of experiments (Deney düzenleme istatistiksel yöntemler). Çeviren: Muluk Z, Toktamış Ö, Kurt S, Karağaoğlu E, Ege Üniversitesi, İzmir, 1994.
16. KEULS M. The use of the "studentized range" in connection with an analysis of variance. Euphytica, 1(2): 112-122, 1952.
17. KLEINBAUM DG, KUPPER LL, NIZAM A, MULLER KE. Applied regression analysis and other multivariable methods. 5th edition, Duxbury Press, Belmont, California, 2008.
18. KLOCKARS AJ, SAX G. Multiple comparisons, Sage Publications, Newbury Park, California, 1986.
19. RAMSEY PH. Power differences between pairwise multiple comparisons. Journal of the American Statistical Association, 73(363): 479-485, 1978.
20. WELSCH RE. A modification of the Newman-Keuls procedure for multiple comparisons. Massachusetts Institute of Technology Working paper: 612-672, 1972.
21. WELSCH RE. Stepwise multiple comparison procedures. Journal of the American Statistical Association, 72(359): 566-575, 1977.
22. DUNCAN DB. Multiple range and multiple F tests. Biometrics, 11(1): 1-42, 1955.
23. SPSS I, SPSS 17.0 Algorithm Guide. 2007. p. 854-861.

24. GABRIEL KR. A simple method of multiple comparisons of means. *Journal of the American Statistical Association*, 73(364): 724-729, 1978.
25. YAZICI BB, PEKER Ö. Bir beyaz eşya işletmesinin kayıp puanlarına ilişkin bir uygulama çalışması. *İstatistik Kongresi Bildiriler Kitabı*. Antalya, 170-172, 1999.
26. KESELMAN H, HUBERTY CJ, LIX LM, OLEJNIK S, CRIBBIE RA, DONAHUE B, KOWALCHUK RK, LOWMAN LL, PETOSKEY MD, KESELMAN JC. Statistical practices of educational researchers: An analysis of their ANOVA, MANOVA, and ANCOVA analyses. *Review of Educational Research*, 68(3): 350, 1998.
27. TATLIDİL H. Uygulamalı çok değişkenli istatistiksel analiz, Ziraat Matbaacılık, Ankara, 2002.
28. KEMP K. Multiple comparisons: Comparisonwise versus experimentwise type I error rates and their relationship to power. *Journal of Dairy Science*, 58(9): 1374-1378, 1975.
29. BENDER R, LANGE S. Adjusting for multiple testing--when and how? *Journal of Clinical Epidemiology*, 54(4): 343-349, 2001.
30. LUDBROOK J. Multiple comparison procedures updated. *Clinical and Experimental Pharmacology and Physiology*, 25(12): 1032-1037, 1998.
31. CABRAL HJ. Multiple comparisons procedures. *Circulation*, 117(5): 698-701, 2008.
32. O'BRIEN P. The appropriateness of analysis of variance and multiple-comparison procedures. *Biometrics*, 39(3): 787-788, 1983.
33. O'NEILL R, WETHERILL G. The present state of multiple comparison methods. *Journal of the Royal Statistical Society. Series B (Methodological)*, 33(2): 218-250, 1971.
34. PERRY J. Multiple-comparison procedures: a dissenting view. *Journal of Economic Entomology*, 79(5): 1149-1155, 1986.
35. ROTHMAN KJ. No adjustments are needed for multiple comparisons. *Epidemiology*, 1(1): 43-46, 1990.
36. SHAFFER JP. Multiple hypothesis testing. *Annual Review of Psychology*, 46: 561-584, 1995.
37. ROBACK PJ, ASKINS RA. Judicious use of multiple hypothesis tests. *Conservation Biology*, 19(1): 261-267, 2005.
38. BENJAMINI Y, HOCHBERG Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 57(1): 289-300, 1995.
39. RAMSEY PH. Comparison of closed testing procedures for pairwise testing of means. *Psychological Methods*, 7(4): 504-523, 2002.
40. RAMSEY PH, RAMSEY PP. Power of pairwise comparisons in the equal variance and unequal sample size case. *British Journal of Mathematical and Statistical Psychology*, 61(1): 115-131, 2008.
41. HORN M, VOLLANDT R. A survey of sample size formulas for pairwise and many one multiple comparisons in the parametric, nonparametric and binomial Case. *Biometrical Journal*, 42(1): 27-44, 2000.
42. EKENSTIERNA M. Multiple comparison procedures based on marginal p-values. Uppsala University Department of Mathematics Project Report No: 2004(12): 1-26, 2004. <http://www2.math.uu.se/research/pub/Ekenstierna.pdf>.
43. JACCARD J, GUILAMO-RAMOS V. Analysis of variance frameworks in clinical child and adolescent psychology: Issues and recommendations. *Journal of Clinical Child and Adolescent Psychology*, 31: 130-146, 2002.

44. JACCARD J, GUILAMO-RAMOS V. Analysis of variance frameworks in clinical child and adolescent psychology: Advanced issues and recommendations. *Journal of Clinical Child and Adolescent Psychology*, 31: 278-294, 2002.
45. GAMES PA. Multiple comparisons of means. *American Educational Research Journal*, 8(3): 531-565, 1971.
46. SAVILLE DJ. Multiple comparison procedures: the practical solution. *American Statistician*, 44(2): 174-180, 1990.
47. HANCOCK GR, KLOCKARS AJ. The quest for : developments in multiple comparison procedures in the quarter century since Games (1971). *Review of Educational Research*, 66(3): 269, 1996.
48. SEAMAN MA, LEVIN JR, SERLIN RC. New developments in pairwise multiple comparisons: Some powerful and practicable procedures. *Psychological Bulletin*, 110(3): 577-586, 1991.
49. THOMAS D. Error rates in multiple comparisons among means-results of a simulation exercise. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 23(3): 284-294, 1974.
50. DAYTON CM. Information criteria for the paired-comparisons problem. *The American Statistician*, 52(2): 144-151, 1998.
51. CURRAN-EVERETT D. Multiple comparisons: philosophies and illustrations. *American Journal of Physiology- Regulatory, Integrative and Comparative Physiology*, 279(1): R1-R8, 2000.
52. BERNHARDSON C. Type I error rates when multiple comparison procedures follow a significant F test of ANOVA. *Biometrics*, 31: 229-232, 1975.
53. MENÉNDEZ DE LA FUENTE I, VALLEJO SECO G, FERNÁNDEZ GARCÍA P. Multiple comparison procedures for simple one-way ANOVA with dependent data. *Spanish journal of psychology*, 2(1): 55-63, 1999.
54. STOLINE MR. The status of multiple comparisons: simultaneous estimation of all pairwise comparisons in one-way ANOVA designs. *American Statistician*, 35(3): 134-141, 1981.
55. RAFTER JA, ABELL ML, BRASELTON JP. Multiple comparison methods for means. *Society for Industrial and Applied Mathematics Review*, 44(2): 259-278, 2002.

7. TEŞEKKÜR

Doktora eğitimim boyunca ve tezimi gerçekleştirmem sırasında benden desteğini hiç esirgemeyen değerli danışmanım ve Anabilim Dalı Başkanımız Doç. Dr. İlker ERCAN'a güveni ve sabrı için çok teşekkür ederim.

Çalışmam boyunca, tezin değerlendirilmesinde değerli katkılarını benimle paylaşan sayın tez izleme jürisi hocalarıma çok teşekkür ederim.

Doktora eğitimim boyunca ve tez çalışmam sırasında göstermiş olduğu destek ve sabrından dolayı eşime ve aileme teşekkür ederim.

8. ÖZGEÇMİŞ

1980 yılında Antalya’da doğdum. İlkokulu Denizli Gazi Mustafa Kemal İlkokulunda bitirdim. Ortaokul öğrenimimi 1995 yılında Denizli Yunus Emre Anadolu Lisesinde tamamladım. Lise öğrenimimi 1998 yılında Denizli TEV Anadolu Lisesi’nde tamamladım. 1998 yılında Hacettepe Üniversitesi Fen-Edebiyat Fakültesi İstatistik bölümünü kazandım. İngilizce hazırlık okuyarak bu bölümden 2003 yılında mezun oldum. 2006 yılında Uludağ Üniversitesi Sağlık Bilimleri Enstitüsü Biyoistatistik Anabilim Dalında yüksek lisansımı tamamladım. 2008 yılında yine kendi bölümümde doktora programına kabul edildim. Uludağ Üniversitesi Tıp Fakültesi Biyoistatistik Anabilim Dalında 2004 yılından itibaren araştırma görevlisi olarak çalışmaktayım.