ORTALAMALARIN KARŞILAŞTIRILMASI

Varyans analizi sonucunda yapılan F testi ile muamele ortalamaları arasındaki farkın istatistiksel olarak önemli olup olmadığı test edilir. Eğer Ho ile ortaya atılan "muamele ortalamaları arasında fark yoktur" hipotezi red edilmişse bu, "en az bir muamele ortalaması diğerlerinden farklıdır" şeklinde kurulan alternatif hipotezin kabul edilmesi demektir. Varyans analizi ile ortaya konan bu farklılığı hangi muamele ortalamasından kaynaklandığının belirlenmesi gerekir. Ortalamaları karşılaştırmak amacıyla genel olarak üç yöntem kullanılır.

- Regresyon tekniği
- Kontrast tanımlama tekniği
- Çoklu karşılaştırma testleri

Bu bölümde bu üç yöntemden çoklu karşılaştırma testleri verilmiştir.

Çoklu karşılaştırma testleri, F testi ile genel olarak ortalamalar arasında fark vardır şeklinde ortaya konan genel sonuçta ortalamalar arasındaki farklılığın hangi muameleden kaynaklandığını tespit etmeye çalışılır. Bu açıdan çoklu karşılaştırma testleri varyans analizini tamamlayıcı bir nitelik taşımaktadır ve denemeye alınan bütün muamele ortalamalarının karşılıklı olarak mukayesesine imkan sağlarlar.

Çoklu karşılaştırma testlerinde birden fazla hipotez birlikte test edildiğinden I. tip hata seviyeleri değişik olmaktadır. Bazı testlerde deneme başına hata (α_d), I. tip hataya(α_c) eşit olurken(($\alpha_d=\alpha_c$) , bazı testlerde bu değer sabit olmayıp karşılaştırılmak istenen ortalama sayısı ile $\alpha_d = 1-(1-\alpha)^{t-1}$ ilişkisi vardır. Ortalamaları karşılaştırmak amacıyla birçok test önerilmiştir. Bu testler arasındaki en önemli farklılık ise işlenecek deneme hatası oranını sınırlamada aldıkları tedbirlerin farklı olmasıdır. Burada en yaygın kullanılan çoklu karşılaştırma testlerinden birkaçı üzerinde durulmuştur. Bunlar;

- LSD (En küçük önemli fark)(Korunmuş)
- TUKEY testi
- DUNCAN testi
- SNK testi

LSD testi

LSD testi, F testinin önemli çıkması halinde kullanılır. Deneme hatasındaki büyümeyi önlemek için hiçbir tedbir öngörmez. Bu nedenle karşılaştırılacak ortalama sayısı arttıkça denem hatası da büyür. Test yaparken I. tip hata seviyesi 0.05 alınırsa bu testte deneme başına hata ile ortalama sayısı arasında $\alpha_d = 1-(1-\alpha)^{t-1}$ ilişkisi olduğundan örneğin ortalama sayısı 5 olduğunda deneme başına hata oranı 0.05 yerine hızla artarak $\alpha_d = 1-(1-0.05)^{5-1} = 0.185$ olmaktadır. Bu nedenle bu testin kullanılması 3'ten fazla ortalama olması halinde önerilmemektedir.

LSD istatistiği

$$LSD = t_{\alpha,v} * \sqrt{\frac{2S^2}{n}}$$

Burada,

$t_{\alpha,v}$: v serbestlik dereceli, α önem seviyesindeki t-cetvel değerini,

S^2 = Hata kareler ortalamasını (varyansı),

n = Her bir muameledeki tekerrür sayısını göstermektedir.

Gruplardaki gözlem sayıları farklı olduğu zaman bir ortak n değeri (n_0) hesaplanır.

$$n_o = \frac{1}{t-1}[N - \frac{\Sigma n_i^2}{N}]$$

Fosfor örneğine LSD testini uygulayalım…

Çoklu karşılaştırma testlerini uygulamada ilk olarak grup ortalamaları büyüklük sırasına dizilir. Buna göre ortalamalar;

p0	p1	p2	p3
16.2	26	27.5	52.2

VK	SD	KT	KO	F
Fosfor	4-1=3	4006.64	1335.55	17.36
Hata	22-4=18	1384.63	76.94	
Genel	22-1=21	5391.27	-	

Test istatistiği;

$$LSD = t_{\alpha,v} * \sqrt{\frac{2S^2}{n}} = 2.101 * \sqrt{\frac{2*76.94}{6}} = 10.64$$

$t_{\alpha,v}=t_{0.025,18}=2.101$

Gruplardaki gözlem sayıları eşit olmadığı için no hesaplanmalıdır.

$$n_o = \frac{1}{t-1}[N - \frac{\Sigma n_i^2}{N}] = \frac{1}{4-1}[22 - \frac{5^2+7^2+4^2+6^2}{22}] = 5.48 \cong 6$$

Tüm grup ortalamaları karşılıklı olarak birbirleriyle karşılaştırılacağına göre, tüm grup ortalamaları arasındaki farklar alınarak, bu farklar LSD değeri ile karşılaştırılacaktır. Eğer Fark>LSD olursa, iki grup arasındaki fark istatistik olarak önemlidir anlamına gelecektir.

Pratik bir karşılaştırma için grup ortalamaları fark tablosu oluşturulup karşılaştırma tablo üzerinde yapılabilir.

	-	p1 26	p2 27.5	p3 52.2
p0	16.2	9.8-	11.3*	36*
p1	26	-	1.5-	26.2*
p2	27.5	-	-	24.7*

9.8<10.64 olduğu için p0 ile p1 arasındaki fark önemsizdir. Benzer şekilde diğer karşılaştırmalarda yapılırsa…

Tüm bu karşılaştırmalar aşağıdaki gibi aynı olan gruplar aynı harflerle ifade edilerek özetlenebilir.

p0	p1	p2	p3
16.2	26	27.5	52.2
a	a		
	b	b	c

Tukey Testi

Tukey'in HSD testi olarak da bilinir. Tukey testi de LSD'de olduğu gibi bir kritik değer kullanır

$$W = q_{t,\alpha,v} * \sqrt{\frac{S^2}{n}}$$

Burada,

$q_{\alpha,t,v}$: t tane grup için v serbestlik dereceli, α önem seviyesindeki Tukey cetvel değerini,

S^2 = Hata kareler ortalamasını (varyansı),

n = Her bir muameledeki tekerrür sayısını göstermektedir.

$$W = q_{t,\alpha,v} * \sqrt{\frac{S^2}{n}} = q_{4,0.05,18} * \sqrt{\frac{S^2}{n}} = 4 * \sqrt{\frac{76.94}{6}} = 14.32$$

Fosfor örneğine LSD testini uygulayalım…

Grup ortalamaları fark tablosundaki fark değerleri ile bu değer karşılaştırılırsa…

	-	p1 26	p2 27.5	p3 52.2
p0	16.2	9.8-	11.3-	36*
p1	26	-	1.5-	26.2*
p2	27.5	-	-	24.7*

9.8<14.32 olduğu için p0 ile p1 arasındaki fark önemsizdir. Benzer şekilde diğer karşılaştırmalarda yapılırsa…

Tüm bu karşılaştırmalar aşağıdaki gibi aynı olan gruplar aynı harflerle ifade edilerek özetlenir.

p0	p1	p2	p3
16.2	26	27.5	52.2
a	a	a	b

Duncan Testi

Kademeli bir testtir. Büyüklük sırasına dizilmiş ortalamalar önce yanındaki ile daha sonra bir atlayarak diğeri ile karşılaştırılırken farklı düzeltmeler yapıldığından farklı Duncan değerleri ile karşılaştırılır. Daha küçük farkları bile önemli bulan bir test olduğundan araştırıcılar tarafından sıklıkla tercih edilir.

$$Dp = Q_{\alpha,p,v} * \sqrt{\frac{S^2}{n}}$$

Burada,

$Q_{\alpha,p,v}$: t tane grup için v serbestlik dereceli, α önem seviyesindeki Duncan cetvel değerini,

S^2 = Hata kareler ortalamasını (varyansı),

n = Her bir muameledeki tekerrür sayısını göstermektedir.

Fosfor örneğine LSD testini uygulayalım…

Grup ortalamaları fark tablosundaki fark değerleri ile bu değer karşılaştırılırsa…

p	2	3	4
$Q_{0.05,p,18}$	2.97	3.12	3.21
$S\bar{x}$	3.58	3.58	3.58
Dp	10.6	11.16	11.49

		p1	p2	p3	
	-	26	27.5	52.2	
p0	16.2	9.8-	11.3*	36*	→ D4=11.49
p1	26	-	1.5-	26.2*	→ D3=11.16
p2	27.5	-	-	24.7*	→ D2=10.6

Tüm bu karşılaştırmalar aşağıdaki gibi aynı olan gruplar aynı harflerle ifade edilerek özetlenir.

p0	p1	p2	p3
16.2	26	27.5	52.2
a	a		
	b	b	c

SNK Testi

Duncan istatistiği gibi kademeli bir testtir. Uygulanışı aynı Duncan testi gibidir. SNK değerleri kademe sayısına göre SNK cetvel değerleri ile standart hatanın çarpılması ile hesaplanır ve kademe sayısına göre alınan farklar ile karşılaştırılır. SNK'nın en son hesaplanan kritik değeri Tukey testindeki kritik değerle aynıdır.

$$SNK = Wp = q_{\alpha,p,v} * \sqrt{\frac{S^2}{n}}$$

Burada,

$q_{\alpha,p,v}$: t tane grup için v serbestlik dereceli, α önem seviyesindeki SNK cetvel değerini,

S^2 = Hata kareler ortalamasını (varyansı),

n = Her bir muameledeki tekerrür sayısını göstermektedir.

Fosfor örneğine LSD testini uygulayalım…

p	2	3	4
$Q_{0.05,p,18}$	2.97	3.61	4.0
Sx	3.58	3.58	3.58
Wp	10.6	12.92	14.32

Grup ortalamaları fark tablosundaki fark değerleri ile bu değer karşılaştırılırsa…

	-	p1 26	p2 27.5	p3 52.2	
p0	16.2	9.8-	11.3-	36*	→ W4=14.32
p1	26	-	1.5-	26.2*	→ W3=12.92
p2	27.5	-	-	24.7*	→ W2=10.6

Tüm bu karşılaştırmalar aşağıdaki gibi aynı olan gruplar aynı harflerle ifade edilerek özetlenir.

p0	p1	p2	p3
16.2	26	27.5	52.2
a	a	a	b

Tukey, Duncan ve SNK testleri F testine bağlı değildir. Yani F testi önemli çıkmasa dahi bu testler yapılır ve ortalamalar arasında farklılık olabilir.