SPSS İLE İSTATİSTİKSEL VERİ ANALİZİ

***S**tatistical **P**ackages for the **S**ocial **S**ciences*

PROF.DR.YÜKSEL TERZİ

ONDOKUZ MAYIS ÜNİVERSİTESİ

FEN-EDEBİYAT FAKÜLTESİ

İSTATİSTİK BÖLÜMÜ

SAMSUN

2019

İSTATİSTİKSEL DENEY TASARIMI

Deney (experiment), kontrol altındaki çeşitli durumların/koşulların deney birimlerinin (experiment unit) bilinmeyen karakteristik özellikleri üzerindeki etkisini test etmek amacıyla uygulanan bir işlem veya süreç olarak tanımlanabilir.

Deney tasarımı (experimental design) deney birimlerinin maruz kalacağı kontrol altındaki durumların/koşulların düzenlenmesiyle ilgilidir.

Deney tasarımı birçok farklı alanda kullanılmaktadır. Tarımda gübre türlerinin buğday verimi üzerindeki etkisini araştırmak için, hayvancılıkta barınak türlerinin buğday verimi üzerindeki etkisini araştırmak için, eczacılıkta piyasaya yeni sürülen bir ilacın etkisini var olan ilaçların etkileriyle karşılaştırmak için, mühendislikte bir üretim sürecinin performansını iyileştirmek için kullanılabilir (Şenoğlu ve Acıtaş, 2010).

İSTATİSTİKSEL DENEY TASARIMI

Deney tasarımında ilgilenilen durumlar/koşullar **faktör (factor)** olarak adlandırılır. Faktörler iki ya da daha fazla **düzeye** sahip olabilir. Düzey sayıları yapan kişinin kontrolü altında olabileceği gibi kontrolü dışında da olabilir. Bir serada yapılan bir deneyde bir ürün için farklı sıcaklık değerlerinin 15 ^{o}C, 20 ^{o}C ve 25 ^{o}C olduğunu varsayalım. Bu deneyde '**sıcaklık**' faktörü üç düzeye sahiptir ve bu düzeyler **deneme (treatment)** olarak adlandırılır. Düzey sayısı denemeyi yapan kişi tarafından belirlendiğinden kontrol altındadır.

Eğer cinsiyetin notlar üzerindeki etkisi araştırılıyorsa, cinsiyet kadın-erkek olmak üzere iki düzeye sahiptir ve düzey sayısı deneyi yapan kişinin kontrolünde değildir.

Deney tasarımında faktörler nitel (qualitative) veya nicel (quantitative) değişken olabilir. Deneyde birden fazla faktör varsa faktör düzeylerinin kombinasyonları deneme olarak adlandırılır. (Şenoğlu ve Acıtaş, 2010).

İSTATİSTİKSEL DENEY TASARIMI

Deney birimlerinden elde edilen gözlem değerlerine R.A.Fisher tarafından geliştirilen ANOVA uygulanarak faktör-lerin istatistiksel olarak anlamlı olup olmadıkları belirlenir.

Deney birimleri, denemelerin rasgele olarak atandığı/uygulandığı varlıklar olarak tanımlanır. Deney birim ile gözlem birim (observational unit) birbirinden farklıdır. **Gözlem birimi**, üzerinde ölçümlerin yapıldığı varlıklar olarak tanımlanır. Deney birim ile gözlem birimi aynı veya farklı olabilir.

Üç farklı yemin (Y1,Y2,Y3) buzağıların doğum ağırlığına olan etkisini araştırılalım. Yem faktörü üç düzeye sahip olup, deneyde üç deneme vardır. Yemlerin rasgele uygulandığı varlıklar inekler olduğundan, inekler deney birimidir. Üzerinde ölçümlerin yapıldığı varlıklar ise buzağılar olduğundan, buzağılar gözlem birimidir. Eğer amaç Y1,Y2,Y3 yemlerinin ineklerin ağırlık artışına olan etkisini araştırmak olsaydı, yemlerin uygulandığı ve üzerinde ölçümlerin yapıldığı varlıklar aynı olacağından, inekler hem deney birimi hem de gözlem birimi olacaktı (Şenoğlu ve Acıtaş, 2010).

İSTATİSTİKSEL DENEY TASARIMI

Deney birimlerinin, hakkında bilgi elde etmek istenen karakteristik özelliklerine yanıt-cevap (response) veya bağımlı (dependent) değişken adı verilir. Bağımlı değişkenin alacağı değerleri etkileyen kontrol edilebilir deneysel değişkenlere bağımsız (independent) değişken veya faktör denir.

Benzer yaş ve kiloda (homojen) 20 kadına dört farklı diyet (D1,D2,D3,D4) uygulanmış ve diyet sonunda diyetlerin kilo kaybına olan etkileri araştırılmış olsun. Bu deneyde; bağımlı değişken kilo, bağımsız değişken diyet programları, bağımsız değişkenin düzeyleri (denemeleri) D1,D2,D3,D4 olacaktır.

Benzer çevresel koşullarda aynı denemeye maruz kalan deney birimlerinden alınan ölçümlerin benzer olması beklenir, ancak bu zordur. Deney birimleri arasında kontrol edilemeyen farklılıklara deneysel hata (experimental error) denir. Deney birimleri arasındaki homojenlik, deneysel hatanın küçülerek deneyin hassasiyetinin artmasını sağlar (Şenoğlu ve Acıtaş, 2010).

DENEY TASARIMININ İLKELERİ

Deney tasarımının üç önemli ilkesi bloklama (blocking), rasgeleleştirme (randomization) ve tekrardır (replication). Bu ilkelerdeki temel amaç deneysel hatanın azaltılmasıdır.

1. Bloklama

Deneyin hassaslığını artırmak için aralarında sistematik farklar bulunan deney birimleri, <u>kendi içinde homojen kendi aralarında heterojen</u> olacak biçimde blok adı verilen gruplara bölünür. Bu işleme bloklama denir. Bloklama ile deneysel hatanın azaltılması hedeflenir.

Y1,Y2,Y3 yemlerinin tavukların ağırlık artışına olan etkisi araştırılmak isteniyor. Bu amaçla yapılan bir deneyde, üç farklı ırkın (I1,I2,I3) her birinden 12'şer tavuk olmak üzere toplam 36 tavuk kullanılıyor. Denemeler (Y1,Y2,Y3) arasında anlamlı karşılaştırma yapabilmek için türdeş olan tavuklar aynı blokta yer alacak şekilde, kendi için homojen, kendi aralarında heterojen üç farklı blok oluşturuluyor. Her deneme her blokta 4'er tavuğa uygulanıyor. Böylece türlerin ağırlık artışına olan etkisi deneme etkilerinden arındırılmış oluyor. Bloklama yapmadan (tavukların farklı türlere ait olduğu göz ardı edilseydi) denemeler tavuklara rasgele uygulansaydı, her deneme her türden farklı sayıda tavuğa uygulanacağından, yemler ile türlerin etkisi karışır ve denemeler arasında anlamlı karşılaştırma yapılamazdı (Şenoğlu ve Acıtaş, 2010).

DENEY TASARIMININ İLKELERİ

2. Rasgeleleştirme

Deney tasarımında deney birimlerinin mümkün olduğunca homojen olması istenir. Bloklama bu amaç için kullanılır. Ancak deney birimleri arasında farklılık her zaman olur. Bu farklılıklar rasgeledir. Denemeler deney birimlerine rasgele olarak uygulanmazsa deneme etkileri arasındaki farklar ile hatanın varyansının tahmin değerleri yanlı olur.

Deney birimleri arasındaki farklılıkların, ölçüm değerleri üzerindeki sistematik etkisini kontrol altına almak için rasgeleleştirme yapılır. Rasgeleleştirme deney birimlerinin denemelerre atanma olasılıklarının eşit olmasını sağlar (Şenoğlu ve Acıtaş, 2010).

DENEY TASARIMININ İLKELERİ

3. Tekrar

Denemelerin uygulandıkları deney birimi sayısına tekrar denir. Denemelere maruz kalan deney birimi sayıları her bir deneme için aynı olmak zorunda değildir. Deney birimi sayıları eşit ise bu tasarımlara dengeli (balanced) tasarımlar, eşit değil ise dengeli olmayan (unbalanced) tasarımlar denir. Tekrar sayısı birden fazla olmalıdır, aksi takdirde deneysel hata hesaplanamaz. Tekrar sayısının artması deneysel hatanın küçülmesine yol açar. Ancak faz sayıda tekrar yapmak deneyin maliyetini artırır.

Deney birimleri arasındaki homojenlik ne kadar fazla ise tekrarlanma sayısı da o kadar az olmalıdır. Çünkü deneysel hata, deney birimleri arasındaki farklılıkların bir ölçüsü olarak ifade edilir (Şenoğlu ve Acıtaş, 2010).

DENEY TASARIMININ AŞAMALARI

i. Deneyin planlanması ve tasarımı
ii. Deneyin yapılması ve verilerin elde edilmesi
iii. Veri analiz ve sonuçların yorumlanması

DENEY TASARIMI

DENEY

- Problemin Tanımı
- Bağımlı değişkenin seçimi
- Bağımsız değişkelerin seçimi
  - Nite-Nicel
  - Sabit yada rasgele
- Etken düzeylerinin birleştirilmesi

DÜZEN

- Gözlem sayısı
- Deneyin sırası
- Kullanılacak rasgeleleştirme yöntemi
- Matematiksel model kurulumu

ÇÖZÜMLEME

- Veri toplama ve işleme
- Test istatistiklerinin hesabı
- Sonuçların yorumu

Muamele: Araştırıcının etkisini araştırdığı etmen (faktör, uygulama).

Tekerrür: Aynı muamelenin uygulandığı birden fazla deneme ünitesine verilen isimdir. Tekerrür sayısı arttıkça denemenin güvenilirliği artar.

Deneme Ünitesi: Araştırıcının ele aldığı en küçük deneme birimidir. Balıkların büyümesi denemesinde her bir deneme havuzundaki balıklar deneme ünitesidir.

Denemenin Kurulması: Muameleler ve sayısı belirlenir. Sonra her bir muamele için tekerrür sayısı belirlenir.

VARYANS ANALİZİ (ANOVA)

- İkiden çok grup
- Gruplar bağımsız
- Nicel veri
- Rastgele örneklem
- Hatalar birbirinden bağımsız, normal dağılışlı.

Normal Dağılım	Varyanslar Homojen	Test İst.	Çoklu Karşılaştırma Testi
✓ Evet	Evet	Anova	Tukey, Duncan
✓ Evet	Hayır	Welch	Tamhane's
✓ Hayır	-	Kruskal-Wallis	Dunn

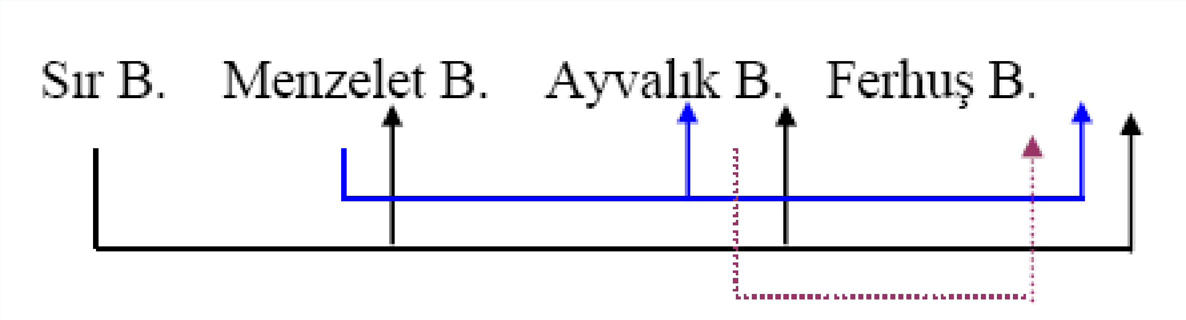
SPSS’TE TESTLERİN YAPILDIĞI YERLER

- ANOVA: Analyze ⇨ Compare Means ⇨ One-way Anova
- WELCH: Analyze ⇨ Compare Means ⇨ One-way Anova ⇨ Options ⇨ Welch
- Kruskal- Wallis: Analyze ⇨ Nonparametric test ⇨ Independent ⇨ Setting Customize tests

VARYANS ANALİZİ (ANOVA)

Muamele sayısının 2'den fazla olması durumunda, t testi ile ikili karşılaştırmaların yapılması gerekir. Örneğin 4 muamele olduğunda 4'ün 2'li kombinasyonu kadar ($_4C_2=6$) karşılaştırma yapmak gerekir. Muamele sayısı arttıkça karşılaştırma sayısı çok daha fazla olacaktır. Bu nedenle 2'den fazla karşılaştırmayı aynı anda yapan varyans analizi tekniği 1925'de Fisher tarafından geliştirilmiştir.

Örnek olarak, 4 farklı barajdaki aynalı sazanların ortalama ağırlıklarının aynı olup olmadığı test edilmek istendiğinde, ikili olarak aşağıdaki gibi 6 karşılaştırma yapmak gerekir.



Dört ortalamanın karşılaştırılması için ikili kombinasyonlar.

VARYANS ANALİZİ (ANOVA)

Varyans analizi yapılabilmesi için bazı varsayımların tutması gerekir. Bunun için bu varsayımların her birinin ayrı ayrı istatistik testleri vardır. Bu testler yapıldıktan sonra tüm varsayımlar tutuyor ise varyans analizine geçilir. Tutmuyor ise verilerin transformasyonları (karekök, logaritmik, açı transformasyonları veya araştırıcının belirlediği bir transformasyon) yapılarak yeniden varsayımlar test edilir. Transformasyonlardan sonra varsayımlar yerine gelmiş ise varyans analizine geçilir. Hala varsayımlar tutmuyorsa, varyans analizi yapılamaz. Bu durumlarda başka yöntemler vardır. Bu yöntemler,

1-Kruskal-Wallis Rank Analizi,

2-Brown ve Forsythe'ın Modifiye F'istatistiği,

3-Box'un Normal F İstatistiği,

4-Tartılı Kareler Ortalaması Yöntemidir.

Varyans analizinin varsayımları

1-Hata terimleri (ε_{ij}) sıfır ortalama ve σ^2 varyans ile normal dağılıma sahiptir.

2- Varyanslar homojendir.

3-Hata terimleri birbirinden bağımsızdır.

Varyanslar heterojen iken varyans analizi yapıp sonuçları yorumlamak hatalı neticeler verebilir. Bu nedenle böyle durumlarda veriler bir transformasyona tabi tutulup varyansların homojenliği sağlandıktan sonra varyans analizi yapılır. Eğer bu şekilde de homojenlik sağlanamaz ise Kruskal-Wallis rank analizi (Siegel, 1956; Gibbons, 1971), Welc analizi (Sokal ve Rholf,1969), Modifiye f istatistiği (Brown ve Forsythe, 1974) veya Box analizi (Box, 1954) gibi analiz şekillerinden biri kullanılabilir.

Sabit Model (Fixed): Etkilerin araştırıcı tarafından tesadüfen değil, bilinçli olarak seçildiği modellerdir.

Şansa Bağlı Model (Random): Etkilerin mümkün tüm seçenekler içerisinden tesadüfen seçildiği ve denemeye alındığı modellerdir.

Karışık Model (Mixed): Hem sabit, hem şansa bağlı etkiler içeren modeldir.

Örneğin denenmesi mümkün biyolojik sınırlar içerisinde 1 mg'dan 100 mg'a kadar hormon dozları (1 mg, 2 mg, 3 mg, ...,100 mg) uygulanabilecekken bilinçli olarak 5,10,15,20 mg düzeyleri denemeye alınıyor ise, hormon dozları muamelesi sabit etkili olur. Eğer 1'den 100'e kadar yazıp kura ile örneğin 3-18-39-55-92 mg düzeylerini denemeye alırsa bu muamele şansa bağlı bir etkiye sahip olur. Model şansa bağlı olduğu zaman deneme sonuçları genelleştirilebilir.

F- Testi

Tüm grupları bir arada ele alarak aralarında anlamlı bir fark bulunup bulunmadığı F testi ile yapılır. Normal dağılan bir anakütleden rasgele çekilen birbirinden farklı n hacimli örneklemlerin istatistiklerinden ki-kare istatistiklerini verir ve bunlarda ki-kare dağılımını meydana getirmektedir.

$$\chi^2 = \frac{(X_1 - \overline{X})^2}{\sigma^2} + \frac{(X_2 - \overline{X})^2}{\sigma^2} + \cdots + \frac{(X_n - \overline{X})^2}{\sigma^2} = Z_1^2 + Z_2^2 + \cdots + Z_n^2 = \sum_{i=1}^{n} Z_i^2$$

Benzer şekilde normal dağılan iki farklı anakütleden n_1 ve n_2 hacimli bütün mümkün rasgele örneklemler seçilip aynı işlemler yapılırsa, $v_1 = n_1 - 1$serbestlik dereceli χ_1^2 dağılımı ve $v_2 = n_2 - 1$ serbestlik dereceli χ_2^2 dağılımı olmak üzere iki dağılım ortaya çıkar. χ_1^2 ve χ_2^2 istatistikleri kendi serbestlik derecelerine bölünüp birbirine oranlanırsa F istatistiği elde edilir.

$$F = \frac{\chi_1^2 / v_1}{\chi_2^2 / v_2} \quad , \qquad \begin{array}{l} H_0 : \mu_1 = \mu_2 = \ldots = \mu_k \\ H_1 : En\, az\, iki\, ortalama\, farklı \end{array}$$

ANOVA TABLOSU

Varyasyon Kaynağı	Kareler Toplamı	Serbestlik Derecesi	Kareler Ortalaması	F Testi
VK	**KT**	**SD**	**KO**	**F**
Gruplar arası	GAKT	k-1	GAKO=GAKT/(k-1)	GAKO/HKO
Gruplar içi	GİKT=HKT	n-k	HKO=HKT/(n-k)	
Genel	GnKT	n-1	GnKO=GnKT/(n-1)	

Burada GAKO gruplar arası varyanstır. Gruplar arası varyans grupların ortalamaları arasındaki farklardan doğan değişkenliği ölçer. Gruplar arası varyans ne kadar büyük ise, grup ortalamalarının birbirinden farklı olma olasılığı o kadar fazladır. HKO ise gruplar içi varyanstır. Gruplar içi varyans her gruptaki değerler arasındaki varyansı ifade eder ve rasgele nedenlere bağlı olan değişkenliği ölçer. Gruplar içi varyans ne kadar büyük ise, grup ortalamalarının birbirinden farklı olma olasılığı o kadar azdır. GnKO ise toplam varyansı gösterir. Toplam varyans gruplar arası varyans ile gruplar içi varyansın toplamına eşit değildir.

$$GnKT = \sum X^2 - \frac{\sum X}{n} \quad , \quad GAKT = \sum_{j=1}^{k}\left[\frac{(\sum X_j)^2}{n_j}\right] - \left[\frac{(\sum X)^2}{n}\right] \quad , \quad HKT = GnKT - GAKT$$

VARYANS HOMOJENLİĞİ TESTLERİ

Varyansların homojenliğinin sağlanması, tip II hatasına karşı araştırmayı korur.

1. F Testi

İki varyans için kullanılan bir testtir. Test istatistiği,

$$F = \frac{S^2 büyük}{S^2 küçük} \sim F_{pay\ S.D.,\ Payda\ S.D.,\ \alpha}$$

şeklindedir ve hipotezler,

$$H_0 : \sigma_1^2 = \sigma_2^2$$

$$H_1 : \sigma_1^2 \neq \sigma_2^2$$

şeklinde oluşturulur.

VARYANS HOMOJENLİĞİ TESTLERİ

2. F_{max} Testi

İki veya daha fazla varyansın homojenlik testi için kullanılır. Test istatistiği,

$$F_{max} = \frac{S^2 en \ büyük}{S^2 en \ küçük} \sim F_{max\,(k,\,v,\alpha)}$$

şeklindedir. Burada,

k = Varyans sayısını,

v = Her bir varyans hesaplanırken eşit sayıda gözlem

var ise herhangi bir varyansın serbestlik derecesini ifade eder.

Bu testte hipotezler,

$$H_0 : \sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \ldots\ldots = \sigma_k^2$$

$$H_1 : \text{en az biri farklı}$$

şeklinde kurulur.

VARYANS HOMOJENLİĞİ TESTLERİ

3. Cochran Testi

İki veya daha fazla varyansın homojenlik testi için kullanılır. Test istatistiği,

$$C = \frac{S^2 \text{en büyük}}{\sum_{i}^{\ell} S_i^2} \sim C_{k,\, v,\, \alpha}$$

şeklindedir. Burada,

k = Varyans sayısını,

v = Varyansların içerdiği gözlemler eşit ise her bir varyansın serbestlik derecesini ifade eder.

VARYANS HOMOJENLİĞİ TESTLERİ

4. Bartlett Testi

H_0: $\sigma_1^2 = \sigma_2^2 = \cdots = \sigma_k^2$

H_a: the variances are not all equal

TS: $$B = \frac{\left[(S_1^2)^{n_1-1}(S_2^2)^{n_2-1}\cdots(S_k^2)^{n_k-1}\right]^{1/(N-k)}}{S_p^2}$$

$$S_p^2 = \frac{\sum_{i=1}^{k}(n_i-1)S_i^2}{N-k}$$

RR: $B \le b_{\alpha,k,n}$ $\quad (n_1 = n_2 = \cdots = n_k = n)$

$B \le b_{\alpha,k,n_1,n_2,\ldots,n_k}$ (when sample sizes are unequal)

where $$b_{\alpha,k,n_1,n_2,\ldots,n_k} \approx \frac{n_1 b_{\alpha,k,n_1} + n_2 b_{\alpha,k,n_2} + \cdots + n_k b_{\alpha,k,n_k}}{N}$$

Brown-Forsythe Testi

Varyanslar heterojen ise ortalamalar arasındaki farklılığı test eden bir istatistiktir.

Test hipotezi,

$$H_0 : \mu_A = \mu_B = \mu_C = \mu_D$$

H_1 : En az biri farklıdır.

Test istatistiği,

$$F' = \frac{MKT}{d} \sim F_{t-1,\hat{v},\alpha}$$

$$d = \sum_{i=1}^{t} (1 - \frac{r_i}{n}) S_i^2$$

$$\hat{V} = \frac{d^2}{\sum_{i=1}^{t} (1 - \frac{r_i}{n})^2 (S_i^2)^2 / (r_i - 1)}$$

Welch Testi (Tartılı Ortalamalar Yöntemi)

Varyanslar heterojen ise ortalamalar arasındaki farklılığı te[illegible] eden bir istatistiktir.

Test Hipotezi,

$$H_0: \mu_A = \mu_B = \mu_C = \mu_D$$

H_1 : En az biri farklıdır.

Test istatistiği,

Tartılı genel ortalama,

$$\overline{Y}^* = \frac{\sum_{i=1}^{t} W_i \overline{Y}_i}{\sum_{i=1}^{t} W_i}$$

Buradaki $\overline{Y}^*$ genellikle, $Y = \left[\sum_{i=1}^{t} r_i \overline{Y}_i / \sum_{i=1}^{t} r_i\right]$ şeklinde verilen basit tartılı ortalamaya eşit değildir. Ağırlıklar ve tartılı genel ortalama hesaplandıktan sonra

$$F^* = \frac{\sum_{i=1}^{t} W_i (\overline{Y}_i - \overline{Y}^*)^2 / (t-1)}{1 + \frac{2(t-2)}{t^2 - 1}\Omega} \sim F_{V_1, V_2, \alpha}$$

t: Grup sayısı,

$$V_1 = t - 1$$

$$V_2 = \frac{t^2 - 1}{3\Omega}$$

$$\Omega = \sum_{i=1}^{t} \frac{1}{f_i}\left(1 - \frac{W_i}{W}\right)^2$$

$$W = \sum W_i$$

$$F^* = \sum_{i=1}^{t} W_i (Y_i - \overline{Y}^*)^2$$

ORTALAMALARIN KARŞILAŞTIRILMASI

Ortalama karşılaştırma yöntemleri üç ana başlık altında toplanır.

1- Regresyon tekniği,

2- Contrast tanımlama,

3- Çoklu karşılaştırma yöntemleri,

3.1. Varyanslar Homojen ise

- a) LSD testi
  - i) Korunmuş LSD testi
  - ii) Korunmamış LSD testi
- b) Tukey testi
- c) Duncan testi
- d) SNK testi
- e) Bonferroni
- f) Scheffe

3.2. Varyanslar Homojen Değilse

- a) Tamhane's testi
- b) Dunnet's T3
- c) Games-Howell

Çoklu Karşılaştırma Testleri
(Multiple Comparison Tests)

ANOVA testi sonucunda H_0 hipotezi red edilirse, ortalamalar arasında istatistiksel olarak anlamlı bir farklılığın olduğu söz edilebilir. Ancak hangi ortalamalar arasında fark olduğu belirlenemez. Bunun için çoklu karşılaştırma testlerine bakmak gerekir.

Örnek büyüklüğüne, varyansın homojenliğine, karşılaştırılacak ortalama sayısına ve önem düzeyine göre hangi çoklu karşılaştırma testin kullanılacağına karar verilir.

Testlerin bazıları ikili karşılaştırmalarda özel hata oranlarını (pairwise error rate), bazıları da eş zamanlı (simultaneous error rate) karşılaştırmalar yapmaktadır.

Çoklu Karşılaştırma Yöntemleri

Varyanslar Homojen.

a.) LSD Testi (En Küçük Önemli Fark Testi)

- **Korunmuş LSD testi:** Korunmuş LSD testi yalnızca F testinin önemli çıkması halinde kullanılır. Üç'ten fazla ortalama olması halinde sakıncalıdır. Çünkü test yaparken I. tip hata düzeyi örneğin %5 seçilirse dahi ortalama sayısı arttıkça deneme başına hata miktarı artmaktadır. Araştırıcının bu hususu unutmaması gerekir.

Örneğin α=%5 iken 10 muamele ortalaması için deneme başına hata miktarı α_d=0,3693'e çıkar

$$\alpha_d=[1-(1-\alpha)^{k-1}]=[1-(1-0,05)^{10-1}]=0,3693.$$

a.) LSD Testi (En Küçük Önemli Fark Testi)

- Korunmuş LSD testi:

s^2 ortak varyans olmak üzere (yani varyanslar homojen) LSD testi aşağıdaki gibi bulunur. s^2=HKO ,$t_{(n-k),\alpha}$ tablo değeri ve r tekerrür sayısı olmak üzere:

$$LSD = t\sqrt{\frac{2s^2}{r}}$$

LSD değeri her grubun ortalama farkları ile karşılaştırılır. LSD değeri küçük ise o ortalamalar birbirinden farklıdır. LSD testi deneme başına hatayı korumaz.

Ortalamalardaki gözlem sayıları farklı olduğu zaman bir ortak n değeri (n_0 veya n_h (harmonik ortalama)) hesaplanır.

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

ve

$$n_h = \frac{t}{\left(\frac{1}{n_1} + \frac{1}{n_2} + \ldots\ldots + \frac{1}{n_t}\right)}$$

şeklindedir.(t= Ortalama sayısı).

ii) Korunmamış LSD testi

F'nin önemli olması gerekmez. Deneme başına hata(αd) en fazla karşılaştırma başına (αc) hata kadar
olur. Bu LSD testi için hazır t- cetveli mevcuttur (Chew, 1976).

b) TUKEY TESTİ

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

Test istatistiği,

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

şeklindedir. Burada,

$q_{\alpha,t,v}$ = Tukey testi için cetvel değerini,

S^2 = Varyans analizindeki hata kareler ortalamasını (varyans),

v = Hata kareler ortalamasının serbestlik derecesini,

α = Önem seviyesini,

n = Her bir ortalamadaki tekerrür sayısını,

ifade eder.

Tukey testi deneme başına hatayı korur. Muamele sayısı arttıkça deneme başına hata artmaz. Ancak **Tukey testi küçük farklara önem vermez.**

Örneğe giren ortalama sayısı C=k(k-1)/2 tüm eşli karşılaştırmalar yapılır.

Tukey testinin güven aralığı LSD testine göre daha kısadır.

Tukey testi k grubun ortalamasını ortak bir hata yaklaşımı ile eşzamanlı ve ikili olarak karşılaştırır. I tip deneme hatası oranındaki korunma değeri sabit ve aynıdır.

c) DUNCAN (DMR) ÇOKLU ARALIK TESTİ

Daha fazla ortalama grubu oluşturan bir testtir. Çünkü daha küçük farklılıkları bile önemli bulan bir testtir. Araştırıcılar ortalamaları arasında bir farklılık olmasını arzu ettiklerinden bu testi sıklıkla kullanırlar. Tukey, Duncan ve SNK testleri F testine bağlı değildir. Yani F testi önemli çıkmasa dahi bu testler yapılır ve ortalamalar arasında farklılık olabilir.

Test istatistiği,

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

şeklindedir. Burada,

$Q_{\alpha,p,v}$ = Kritik cetvel değerini,

p = Büyüklük sırasına dizilmiş ortalamalar arasındaki kademe sayısını,

S^2 = Hata kareler ortalamasını,

v = Hata kareler ortalamasının serbestlik derecesini,

α = Önem seviyesini,

n = Her bir ortalamadaki tekerrür sayısını, ifade eder.

Duncan testinde kullanılan kritik değer sayısı ortalama sayısının bir eksiği kadardır.

Duncan testi sıraya dizilmiş ortalamalar arasındaki farklılıkları ortalamanın sıralamadaki konumunu dikkate alarak değerlendiren bir testtir.

Duncan testi birbiriyle karşılaştırılan ortalamaların konumlarının birbirine göre önemli olduğu karşılaştırmalarda çok sık kullanılır.

ANOVA testi önemsiz çıksa bile Duncan testi kullanılabilir. Duncan testi ortalama sayısı arttıkça koruma seviyesi azaldığından, ortalamalar arasında fark olma olasılığı artar. Böylece ortalamalar arasındaki küçük farklılıkları bile önemli bulma ihtimali artar.

Duncan α =0,05 Önem Düzeyinde, Koruma ve H0 Reddetme Olasılığı (Ott, 1988).

Karşılaştırılacak Ortalama Sayısı p	Koruma Seviyesi $(1-0.05)^{p-1}$	H_0 Reddetme Olasılığı $1-(1-0.05)^{p-1}$
2	0.950	0.050
3	0.903	0.098
4	0.857	0.143
5	0.815	0.185
6	0.774	0.226
7	0.735	0.265
8	0.698	0.302
9	0.663	0.337
10	0.630	0.370

Karşılaştırılacak ortalama sayısı arttıkça, Duncan testinin sıfır hipotezini reddetme olasılığı artar.

d) SNK (Student – Newman – Keuls) TESTİ

Duncan testine benzeyip çoklu aralık testidir. Tukey ile aynı cetveli kullanır. SNK'nın en son hesaplanan kritik değeri Tukey testindeki kritik değerle aynıdır.

Test istatistiği,

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

şeklindedir. Burada,

$q_{\alpha,p,v}$ = Kritik cetvel değeri,

S^2 = Varyans analizindeki hata kareler ortalamasını (varyans),

v = Hata kareler ortalamasının serbestlik derecesini,

α = Deneme başına düşen birinci tip hata oranını (önem seviyesi),

n = Her bir ortalamadaki gözlem sayısını,

ifade eder.

SNK testinde de Duncan testinde olduğu gibi kullanılan kritik değer sayısı ortalama sayısının bir eksiği kadardır.

- Bonferroni Testi

F testi önemli çıkmasa bile uygulanabilir. Tukey testine göre daha güçlüdür.

- Dunnett Testi

Uygulama gruplarının her biri kontrol grubu ile karşılaştırılır. Uygulama gruplarını birbirleriyle karşılaştırmaz. F testi önemsiz olsa bile uygulanabilir.

-Scheffe Testi

Güven sınırlarına dayalı güçlü bir testtir. Grupların tekerrür sayıları eşit olmadığı durumlarda da uygulanabilen bir testtir. Bu testte hata oranı deneme başınadır. Bu nedenle α=0,10 tercih edilir.

Varyanslar Heterojen Olduğunda Kullanılan Çoklu Karşılaştırma Testleri

Dunnett T3:

Varyanslar heterojen olduğunda ve örnek büyüklüğü n<50 olduğunda kullanılır.

Games-Howell:

Varyanslar heterojen olduğunda ve örnek büyüklüğü n>50 olduğunda kullanılır. Dunnett T3 testine göre güven aralığı daha ısadır yani bu teste göre daha güçlü bir testtir (Wilcox, 1987).

Tamhane's Testi:

Varyanslar heterojen ise bu test kullanılabilir.

CONTRAST (ÖZEL KARŞILAŞTIRMALAR)

CONTRAST deyimi muamele ortalamaları arasında karşılaştırma yapmak için kullanılır. Bununla birlikte POLYNOMIAL deyimi kullanılırsa ortogonal polinom katsayıları kullanılarak karşılaştırma yapılır.

İki muamele ortalaması arasındaki bir karşılaştırma veya contrast işaret dikkate alınmadan, iki ortalama arasında alınan farktır. Yani $\bar{Y}_1 - \bar{Y}_2$ farkı bir kontrastır. Üç ortalama arasındaki karşılaştırma çeşitli yollarla yapılabilir.

$$a_1\overline{Y}_1 + a_2\overline{Y}_2 + a_3\overline{Y}_3$$

Burada $\bar{Y}_1 - \bar{Y}_2$ 'dır. Buna göre de ai=(0 1 -1) olabilir. Bu ifade ai'lerin özel bir linear fonksiyonudur ve ortalamalarının tartılı bir toplamıdır.

Örneğin a1=1, a2=0 ve a3=-1 alınırsa karşılaştırma $\sum ai = 0$ şeklinde , a1=0, a2=1, a3=-1 alınırsa karşılaştırma $\bar{Y}_1 - \bar{Y}_3$ şeklindedir. Bu karşılaştırmalardan herbiri bir kontrasttır. Üç veya daha fazla ortalama arasında yapılan karşılaştırmalarda tarif edilen kontrastlar birbiri ile ortogonal olabileceği gibi ortogonal olmayabilirde.

Ortagonal Contrast

Ortagonal Contrast

iki yada daha fazla kontrast söz konusu olduğunda ortagonallıktan söz edilir.

Dengeli deneme planı $\sum_{i=1}^{a} ci\,di = 0$

Dengesiz deneme planı $\sum_{i=1}^{a} r\,ci\,di = 0$

$ci = 1.$kontrastın katsayısı

$di = 2.$kontrastın katsayısı

iki kontrastın çarpımlarının toplamı 0 ise ortagonaldır şeklinde yorumlanır.

Örneğin μ_0 (Kontrol) μ_1 μ_2 $a = 3 =$ muamele

	μ_0 (Kontrol)	μ_1	μ_2
K_1	-2	1	1
K_2	0	-1	1

$c_1=-2$	$C_2=1$	$c_3=1$	
$d_1=0$	$d_2=-1$	$d_3=1$	
0	-1	1	$\sum=0$

$$\sum_{i=1}^{a} ci\,di = 0$$

(2*0)+(1*-1)+(1*1)=0-1+1=0

O halde Ortogonal olduğu görülür.

Örnek. 4 ortalama arasında aşağıdaki şekilde üç contrast tarif edilsin:

C1= (-3) $\bar{Y}_1$+(-1) $\bar{Y}_2$+(+1) $\bar{Y}_3$+(+3) $\bar{Y}_4$ Ho: $\mu_3 + 3\mu_4 = \mu_2 + 3\mu_1$

C2=(+1) $\bar{Y}_1$ +(-1) $\bar{Y}_2$+(-1) $\bar{Y}_3$+(+1) $\bar{Y}_4$ Ho: $\mu_1 + \mu_4 = \mu_2 + \mu_3$

C3=(+1) $\bar{Y}_1$ +(+1) $\bar{Y}_2$+(-1) $\bar{Y}_3$+(-1) $\bar{Y}_4$ Ho: $\mu_1 + \mu_2 = \mu_3 + \mu_4$

Contrastların birbirine tekabül eden terimlerinin katsayılarının çarpımı toplamı sıfır ise, bu karşılaştırmalar veya contrastlar ortogonaldır denir. Buna göre C_1 ve C_2 karşılaştırmalarına bu kuralı uygularsak;

(-3) (+1) + (-1) (-1)+ (+1)(-1)+(+3)(+1)=0

Buradan görülüyorki C_1 ve C_2 kontrastları ortogonaldır.

Buna göre C_1 ve C_3 karşılaştırmalarına bu kuralı uygularsak;

(-3) (+1) +(-1)(+1)+(+1)(-1)+(+3)(-1)=-8 o halde C_1 ve C_3 karşılaştırmalarının ortogonal olmadığını anlarız.

Eğer bir faktörün seviyeleri (muameleler) eşit aralıklarla artan biçimde bir seri oluşturuyorsa ve **kantitatif (nicel)** ise muamele varyasyonu, linear, kuadratik, kübik,..vs şeklinde parçalanabilir. **Bu parçalama yöntemi muamele ölçeği boyunca ele alınan kademelerle ölçüm kriteri arasındaki ilişkinin şekli hakkında bilgi verir.** Ortagonal parçalanma sonucunda tahmin edilen varyasyon kaynaklarına ait (tek serbestlik dereceli) kareler toplamının toplamı muameleler arası kareler toplamına eşit olmalıdır. Ortagonal parçalanma sonucunda elde edilen yeni ilişkilerin denklemleri lineer, kuadratik ve kübik için aşağıdaki gibidir:

$$\text{Linear} \quad \hat{Y} = b_o + b_1 X$$

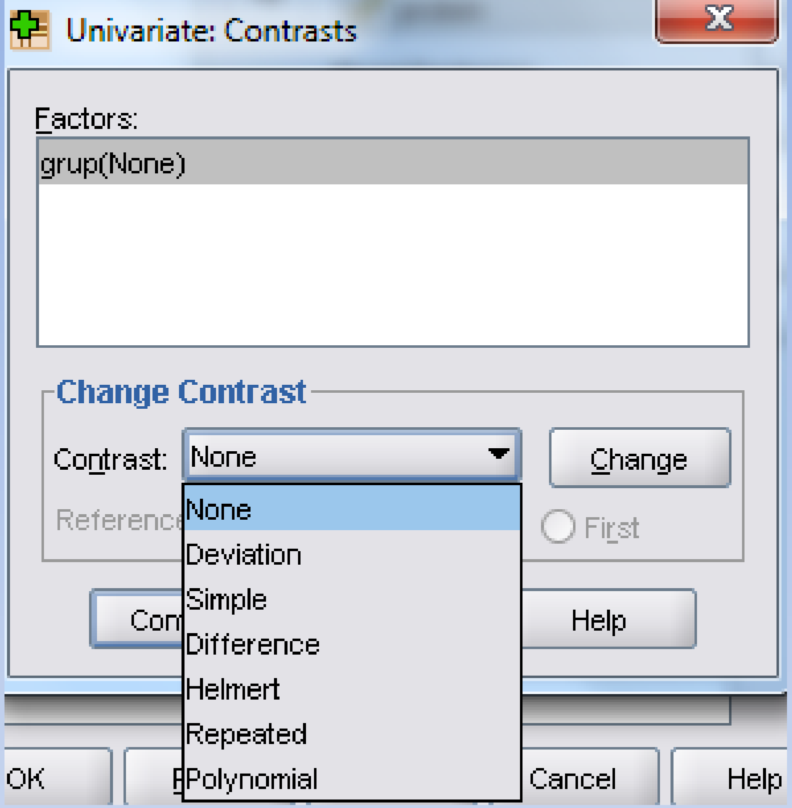
$$\text{Kuadratik} \quad \hat{Y} = b_o + b_1 X + b_2 X^2$$

$$\text{Kübik} \quad \hat{Y} = b_o + b_1 X + b_2 X^2 + b_3 X^3 \quad \text{şeklinde yazıazılabr}$$

Ortogonal Polinom Katsayıları

Grup Sayısı	Polinom Derecesi	1	2	3	4
3	Lineer	-1	0	+1	
	Kuadratik	+1	-2	+1	
	Lineer	-3	-1	+1	+3
4	Kuadratik	+1	-1	-1	+1
	Kübik	-1	+3	-3	+1

Contrast Türleri



Deviation Contrast: Genel Ortalama ile herbir ortalamanın karşılaştırılmasıdır.

Simple Contrast: Her bir ortalamayı tüm ortalamalar ile karşılaştırmak istediğimizde kullanılır.

1	0	-1	0	0
0	1	-1	0	0
0	0	-1	1	0
0	0	-1	0	1

$\mu1=\mu3$ $\mu2=\mu3$ $\mu4=\mu3$ $\mu5=\mu3$

Helmert Contrast : Faktörün herbir düzeyinin ortalamasını kendisini takip eden düzeylerle karşılaştırır. En son faktör düzeyi hariç karşılaştırmalar yapılır.

-4	1	1	1	1
0	-3	1	1	1
0	0	-2	1	1
0	0	0	-1	1

$4\mu 1=(\mu 2+\mu 3+\mu 4+\mu 5)$

$3\mu 2=(\mu 3+\mu 4+\mu 5)$

$2\mu 3=(\mu 4+\mu 5)$

$\mu 4=\mu 5$

Difference Contrast: Herbir ortalamayı kendinden önceki ortalamalar ile karşılaştırır. (Helmerrt Contrast'ın karşıtıdır)

$\mu 2 = \mu 1$

$2\mu 3 = \mu 2 + \mu 1$

$3\mu 4 = \mu 3 + \mu 2 + \mu 1$

$4\mu 5 = \mu 1 + \mu 2 + \mu 3 + \mu 4$

1	-1	0	0	0
1	1	-2	0	0
1	1	1	-3	0
1	1	1	1	-4

Repeated Contrast : Herbir düzeyin ortalamasını sonraki düzeylerin ortalaması ile karşılaştırır. (Helmert'den farkı en son faktör düzeyi hariç karşılaştırma yapılmasıdır.)

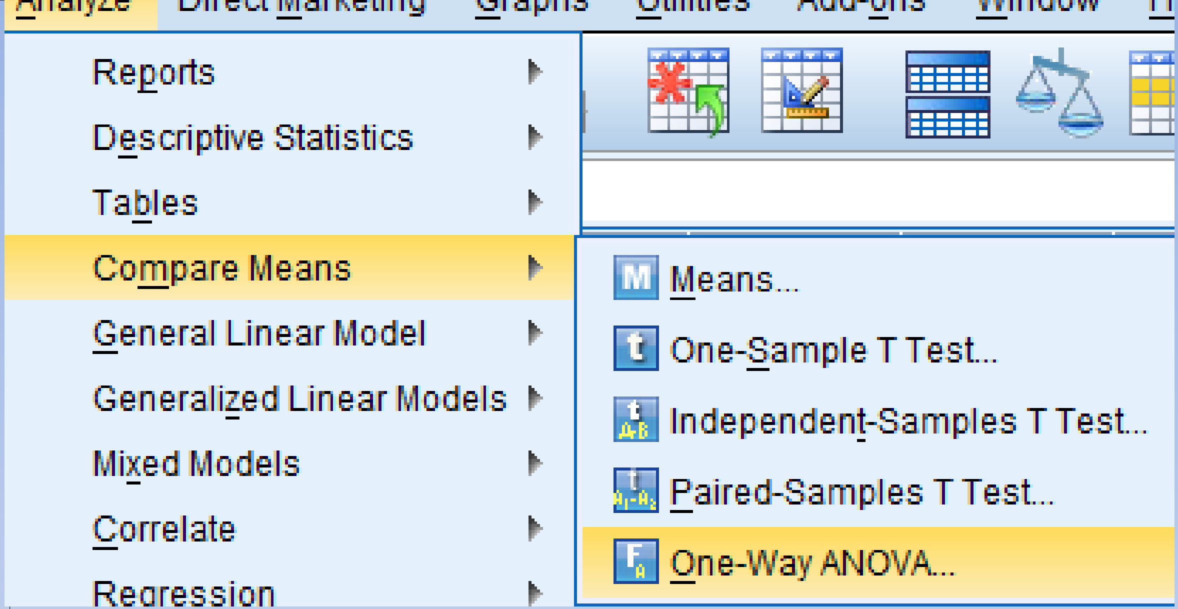
-1	1	0	0	0
0	-1	1	0	0
0	0	-1	1	0
0	0	0	-1	1

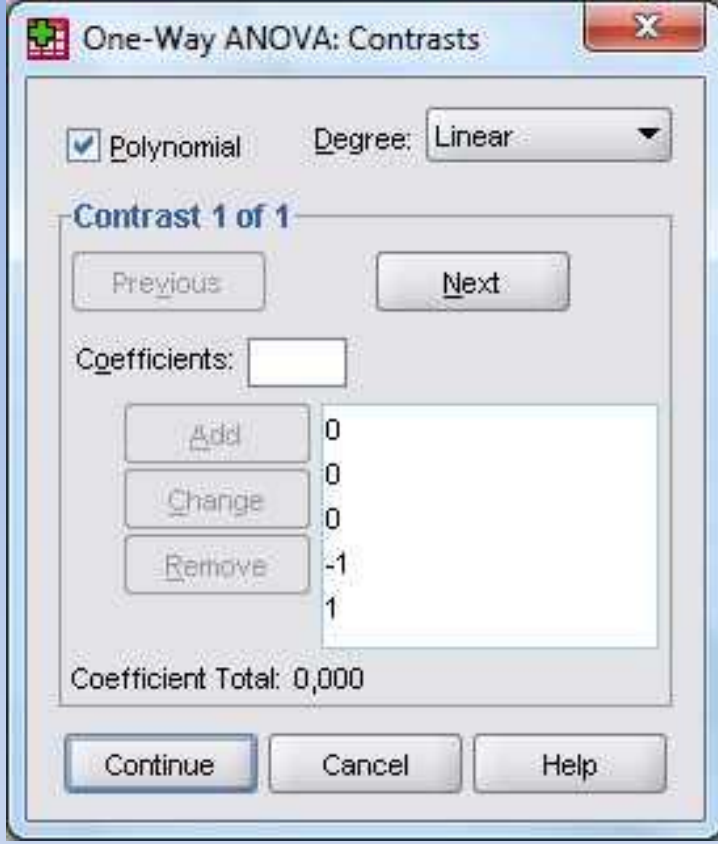
$\mu 1=\mu 2$

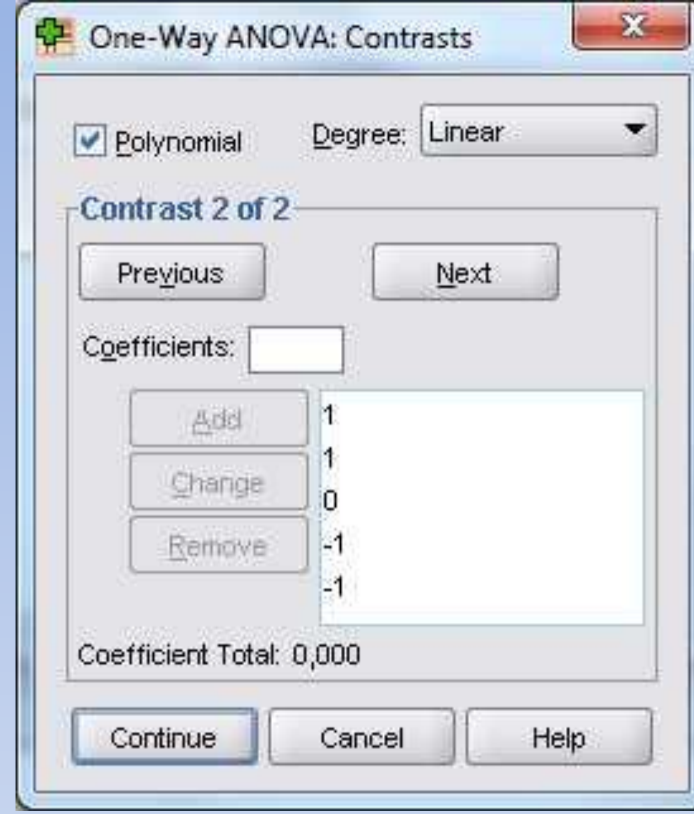
$\mu 2=\mu 3$

$\mu 3=\mu 4$

$\mu 4=\mu 5$


Reports
Descriptive Statistics
Tables
Compare Means
General Linear Model
Generalized Linear Models
Mixed Models
Correlate
Regression
Means...
One-Sample T Test...
Independent-Samples T Test...
Paired-Samples T Test...
One-Way ANOVA...





SPSS Syntax ile Çözüm

```
ONEWAY
  veri BY grup
  /POLYNOMIAL= 1
  /CONTRAST= 0 0 0 -1 1
  /CONTRAST= 1 1 0 -1 -1
  /STATISTICS DESCRIPTIVES.
```

$$H_0 : \mu_4 = \mu_3$$

$$H_0 : \mu_1 + \mu_2 = \mu_3 + \mu_4$$

ANOVA

veri

			Sum of Squares	df	Mean Square	F	Sig.
Between Groups	(Combined)		18,160	4	4,540	1,182	,349
	Linear Term	Contrast	,980	1	,980	,255	,619
		Deviation	17,180	3	5,727	1,491	,247
Within Groups			76,800	20	3,840		
Total			94,960	24			

Contrast Coefficients

Contrast	grup				
	1	2	3	4	5
1	0	0	0	-1	1
2	1	1	0	-1	-1

Contrast Tests

		Contrast	Value of Contrast	Std. Error	t	df	Sig. (2-tailed)
veri	Assume equal variances	1	-2,60	1,239	-2,098	20	,049
		2	,20	1,753	,114	20	,910
	Does not assume equal variances	1	-2,60	1,095	-2,373	7,587	,047
		2	,20	1,726	,116	14,873	,909

ÖRNEK: Bir işletmede bulunan üç eşdeğer makine üretimi aşağıdaki gibidir. Bu üç makine arasında fark var mıdır?

	A	B	C	
	4	6	3	
	5	7	4	
	5	6	5	
	4	8	5	
	6	6	4	
	6	7	4	
	4	9	3	
	5	8	3	
	4	6	4	
	4	5	3	Toplam
Σx	47	68	38	153 (Σx)
Σx^2	227	476	150	853 (Σx^2)
n_j	10	10	10	30 (Σn)

I. Kareler toplamlarının bulunması:

GnKT:Genel Kareler Toplamı

$$GnKT = \sum X^2 - \frac{\sum X}{n} \quad , \quad GAKT = \sum_{j=1}^{k}\left[\frac{(\sum X_j)^2}{n_j}\right] - \left[\frac{(\sum X)^2}{n}\right] \quad , \quad GİİK = GnKT - GAKT$$

$$\text{GnKT} = \sum \text{X}^2 - \frac{(\sum \text{X})^2}{\text{n}} = 853 - \frac{153^2}{30} = 72{,}7$$

GAKT: Gruplar arası kareler toplamı

$$\text{GAKT} = \sum_{\text{j}=1}^{\text{k}}\left[\frac{(\sum \text{X}_\text{j})^2}{\text{n}_\text{j}}\right] - \left[\frac{(\sum \text{X})^2}{\text{n}}\right] = \left[\left(\frac{47^2}{10}\right) + \left(\frac{68^2}{10}\right) + \left(\frac{38^2}{10}\right)\right] - \left(\frac{153^2}{30}\right) = 47{,}4$$

GiKT: Grup içi kareler toplamı

$$\text{GİKT} = \text{GnKT} - \text{GAKT} = 72{,}7 - 47{,}4 = 25{,}3$$

Serbestlik Derecelerinin Bulunması:

Genel serbestlik derecesi: GnSD= n-1 =30-1=29

Gruplar arası serbestlik derecesi: GASD=Grup sayısı-1=3-1=2

Grup içi serbestlik derecesi: GiSD= n-Grup sayısı=30-3=27

Kareler Ortalamasının Bulunması:

Gruplar arası kareler ortalaması:

$$\text{GAKO} = \frac{\text{GAKT}}{\text{k}-1} = \frac{47{,}4}{2} = 23{,}7$$

Grup içi kareler ortalaması:

$$\text{GİKO} = \frac{\text{GİKT}}{\text{n}-\text{k}} = \frac{25{,}3}{27} = 0{,}937$$

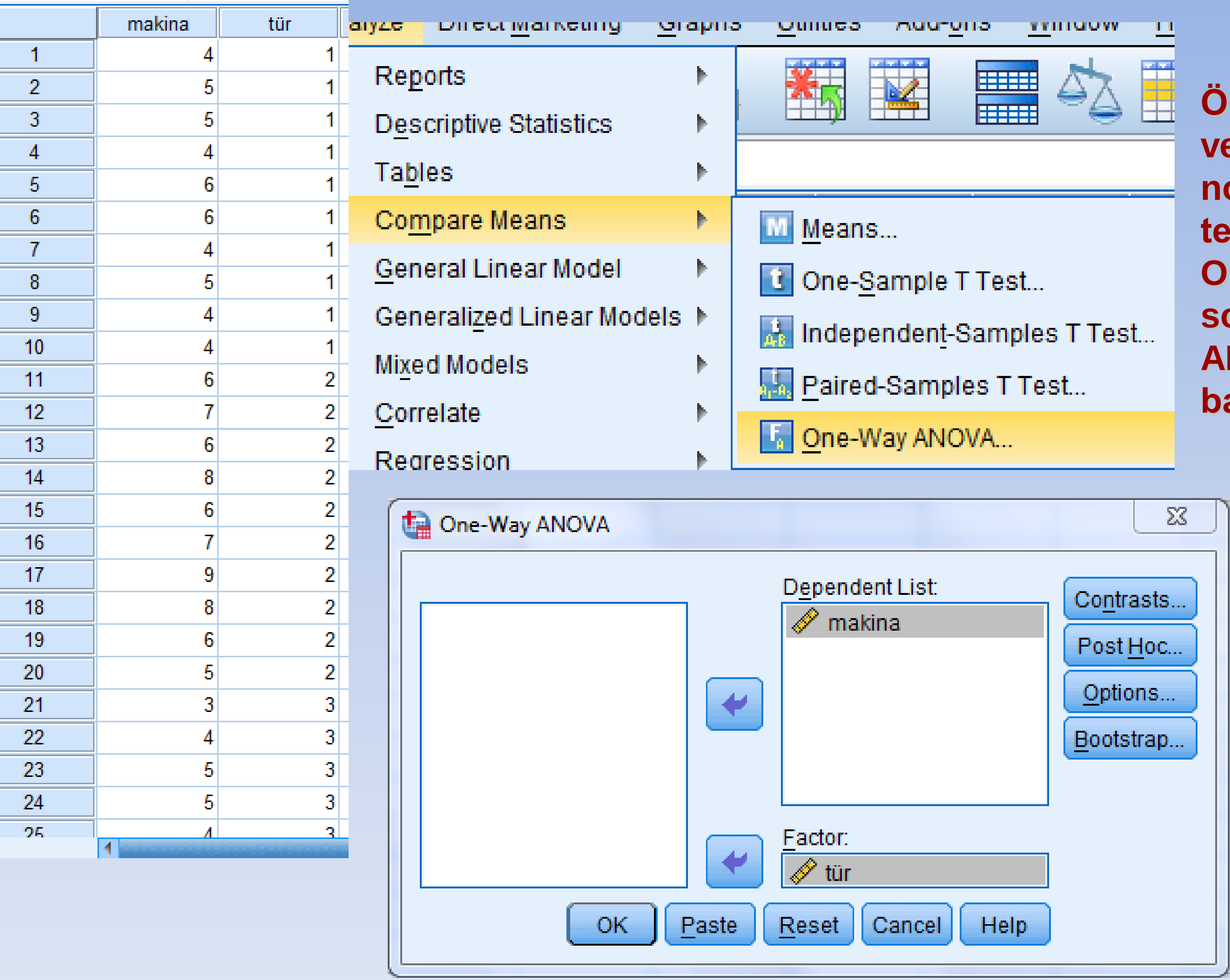
Varyasyon Kaynağı	Kareler Toplamı	Serbestlik Derecesi	Kareler Ortalaması
VK	KT	SD	KO
GA	47.4	2	23.7
Gi	25.3	27	0.937
Gn	72.7	29	----

Hipotezler: H_0: Gruplar arası fark yoktur. H_1: Gruplar arasında fark vardır.

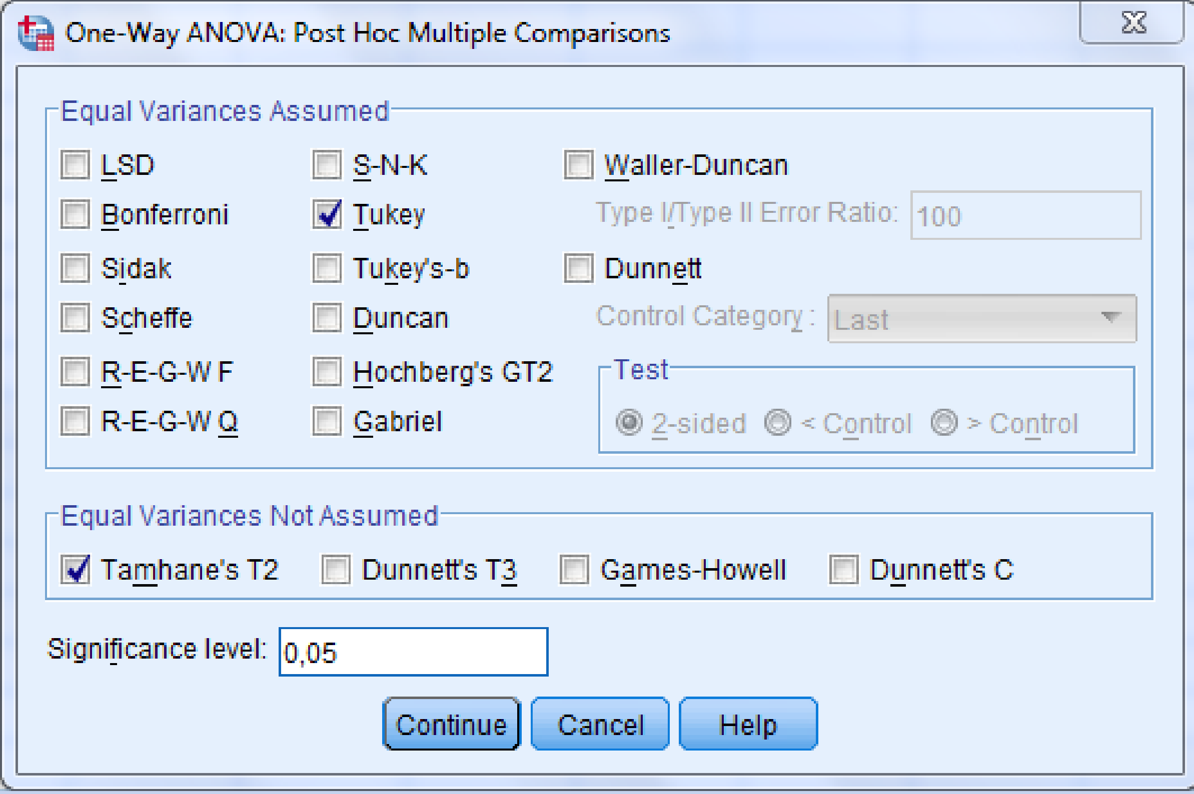
$$F = \frac{\text{GAKO}}{\text{GİKO}} = \frac{23{,}7}{0{,}937} = 25{,}3$$

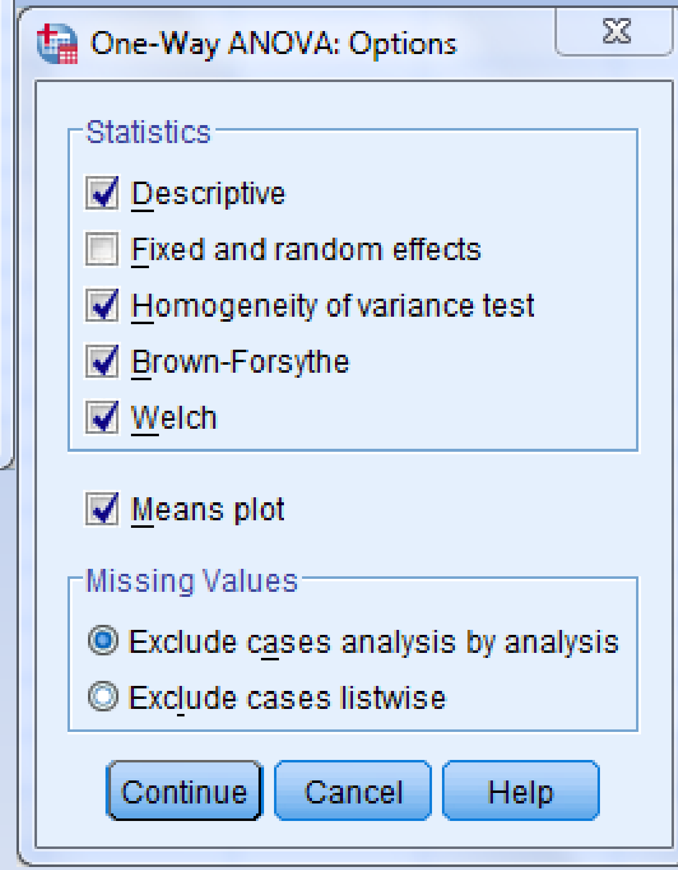
Karşılaştırma: F_{Hesap}=25.3 F_{Tablo} = 2.35 ; 25.3 > 2.35 olduğundan H_0 ret edilir.

Sonuç: Gruplar arasında fark vardır. Üç makinenin üretimi arasında anlamlı bir fark bulunmuştur. Bundan sonra gruplar ikişer ikişer karşılaştırılır. Bu karşılaştırmada t testi kullanılır. Bu şekilde karşılaştırılan ortalamalar sıralanır ve önem denetimi yapılır.



Önce verilerin normallik testi yapılır. Ondan sonra ANOVA ya bakılır.


One-Way ANOVA: Post Hoc Multiple Comparisons
Equal Variances Assumed
LSD
S-N-K
Waller-Duncan
Bonferroni
Tukey
Type I/Type II Error Ratio:
100
Sidak
Tukey's-b
Dunnett
Scheffe
Duncan
Control Category :
Last
R-E-G-W F
Hochberg's GT2
Test
R-E-G-W Q
Gabriel
2-sided
< Control
> Control
Equal Variances Not Assumed
Tamhane's T2
Dunnett's T3
Games-Howell
Dunnett's C
Significance level:
0,05
Continue
Cancel
Help


One-Way ANOVA: Options
Statistics
Descriptive
Fixed and random effects
Homogeneity of variance test
Brown-Forsythe
Welch
Means plot
Missing Values
Exclude cases analysis by analysis
Exclude cases listwise
Continue
Cancel
Help

Test of Homogeneity of Variances

MAKINA

Levene Statistic	df1	df2	Sig.
1,593	2	27	,222

P=0.222>0.05 olduğundan varyanslar homojendir.

ANOVA

MAKINA

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	47,400	2	23,700	25,292	,000
Within Groups	25,300	27	,937		
Total	72,700	29			

P=0.000<0.01 olduğundan en az bir grup ortalaması diğerlerinden farklıdır.

$$H_0 : \mu_A=\mu_B=\mu_C$$

Multiple Comparisons

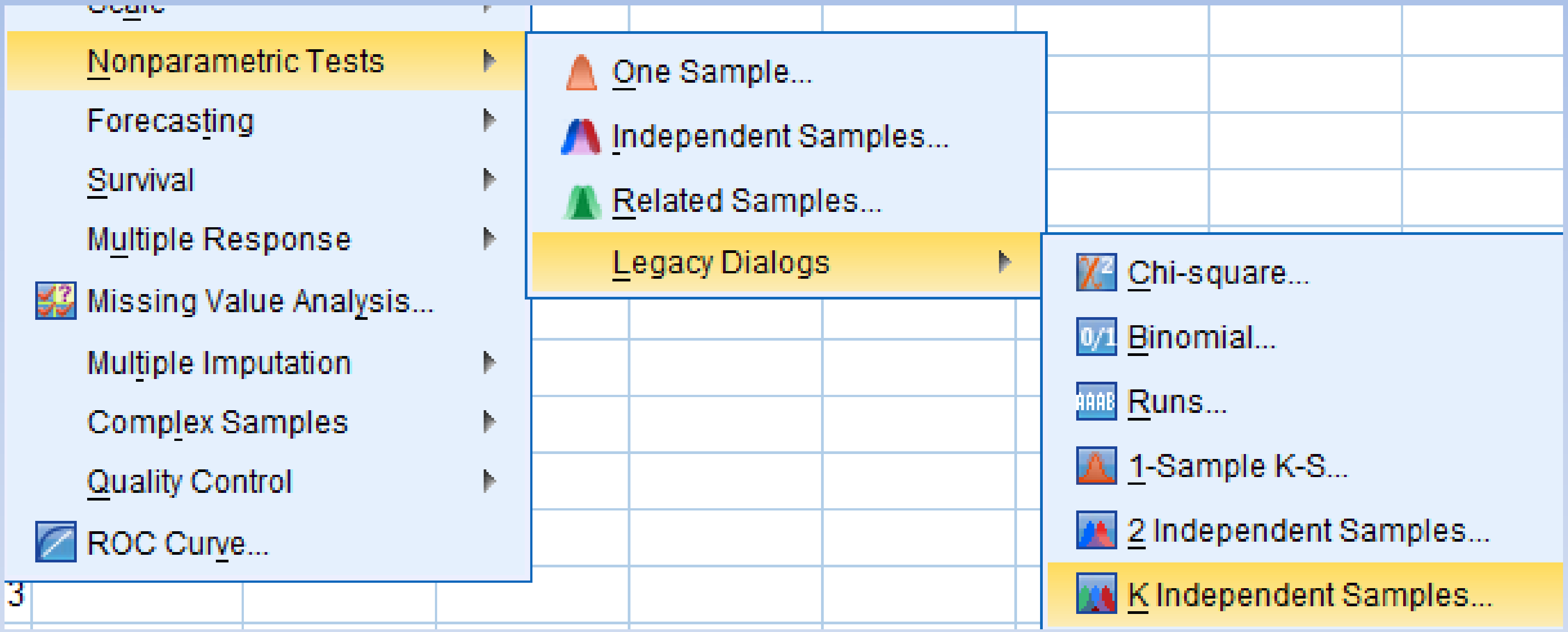
Dependent Variable: MAKINA

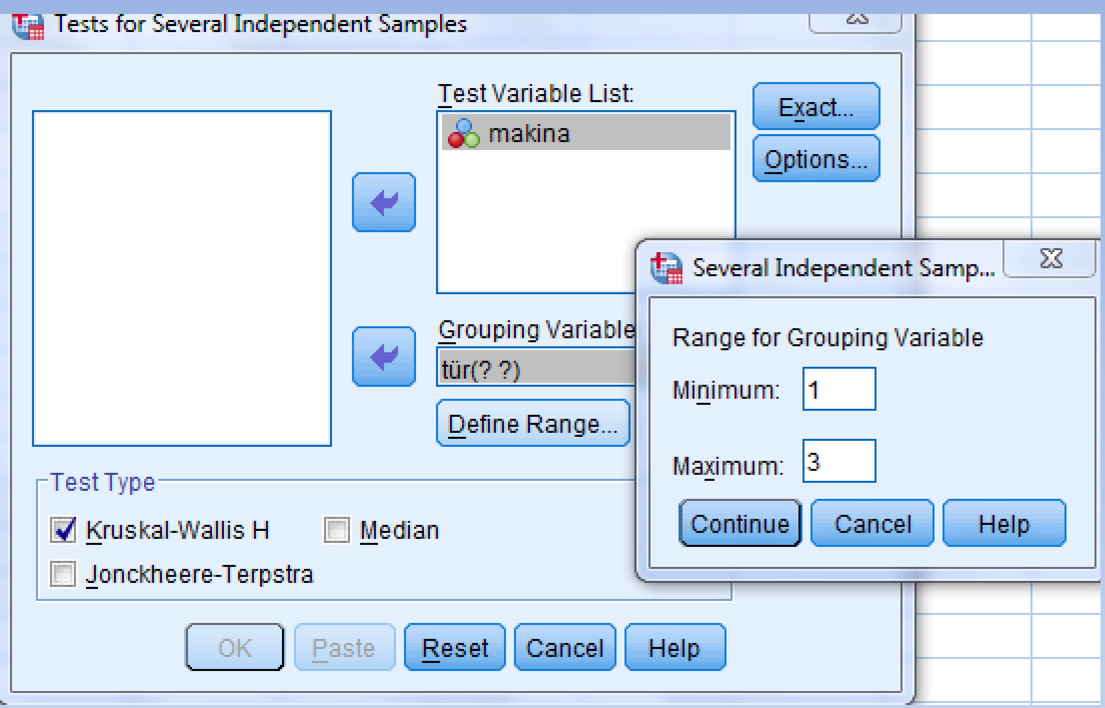
	(I) TUR	(J) TUR	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Tukey HSD	A	A					
		B	-2,10*	,43	,000	-3,17	-1,03
		C	,90	,43	,113	-,17	1,97
	B	A	2,10*	,43	,000	1,03	3,17
		B					
		C	3,00*	,43	,000	1,93	4,07
	C	A	-,90	,43	,113	-1,97	,17
		B	-3,00*	,43	,000	-4,07	-1,93
		C					
Tamhane	A	A					
		B	-2,10*	,43	,001	-3,35	-,85
		C	,90	,43	,066	-4,88E-02	1,85
	B	A	2,10*	,43	,001	,85	3,35
		B					
		C	3,00*	,43	,000	1,76	4,24
	C	A	-,90	,43	,066	-1,85	4,88E-02
		B	-3,00*	,43	,000	-4,24	-1,76
		C					

*. The mean difference is significant at the .05 level.

Mean difference kısmında * işaretli grup ortalamaları birbirinden farklıdır. Buna göre A ile B ve B ile C grup ortalamaları birbirinden farklıdır.

Eğer veriler normal dağılış göstermiyorsa, non-parametric testlerden Kruskal-Wallis testi kullanılır.





Test Statistics[a,b]

	makina
Chi-Square	19,439
df	2
Asymp. Sig.	,000

a. Kruskal Wallis Test

b. Grouping Variable: tur

Eğer ortalamalar farklı çıkarsa, Mann-Whitney U testi ile ikili karşılaştırmalar yapılır. Ancak doğru bir hipotezin red edilmesi anlamına gelen I.Tip hata riskinin azaltılması için önemlilik seviyesi yapılacak olan karşılaştırma sayısına bölünür. Mesela önemlilik seviyesi 0.05 ise ve 3 karşılaştırma yapılacaksa, anlamlılık seviyesi 0.05/3=0.0167 alınır.

P=0,00<0,05 Ho hipotezi kabul edilemez.

KRUSKAL-WALLIS TESTİ

Üç veya daha fazla bağımsız grup nicel verileri normal dağılım göstermiyorsa, parametrik olmayan bir test olan Kruskal-Wallis testi ile gruplar karşılaştırılır. Eğer gruplar arasında anlamlı bir farklılık bulunursa, Mann-Whitney U testi ile ikili karşılaştırmalar yapılır. Ancak doğru bir hipotezin red edilmesi anlamına gelen I.Tip hata riskinin azaltılması için önemlilik seviyesi yapılacak olan karşılaştırma sayısına bölünür. Mesela önemlilik seviyesi 0.05 ise ve 3 karşılaştırma yapılacaksa, anlamlılık seviyesi 0.05/3=0.0167 alınır.

KRUSKAL-WALLIS TESTİ

- Gruplar arası tek yönlü varyans analizinin parametrik olmayan alternatifidir.
- Amaç, her biri n_j, hacimli k sayıda bağımsız örneğin aynı kitleden gelip gelmediğine karar vermektir.
- Sürekli değişkenlere sahip üç yada daha fazla grup için karşılaştırma yapmayı sağlar.
- Değerler sıralı hale çevrilir ve her grup için sıralı ortalamalar karşılaştırılır (Gamgam ve Altunkaynak, 2013).

Varsayımları :

- Veriler n_1, n_2, … , n_k hacimli k tane örnekten oluşur.
- Gözlemler hem örneklem içi hem de örneklemler arası bağımsızdır.
- Değişkenler süreklidir.
- Ölçme düzeyi en az sıralamadır (Gamgam ve Altunkaynak, 2013).

TEST HİPOTEZİ VE TEST ÖLÇÜTÜ

- Hipotezler :

$H_0 : \tau_1 = \tau_2 = \ldots = \tau_k$
H_1 : En az bir τ_j farklıdır (j=1,2,…k).

Ya da ;

H_0 : Örneklerin seçildiği k sayıda yığının hepsi aynı medyana sahiptir.
H_1 : Örneklerin seçildiği k sayıda yığının hepsi aynı medyana sahip değildir.

model aşağıdaki gibi ifade edilebilir.

$X_{ij} = \mu + \tau_j + \varepsilon_{ij}$ i =1,2,…,$n_{i;j}$=1,2,..,k

X_{ij} : j. yığındaki i. birimin değeri

μ : Genel ortalama

τ_j : Faktörün j. düzeyinin bağımlı değişkene etkisi

ε_{ij} : Hata terimi

GÖZLEMLERİN VE RANKLARIN ÖZET TABLOSU

Örneklemeler	Gözlemler	$R_İ = \sum_{j=1}^{n_İ} R(y_{ij})$	n_i
1	$y_{11}\ y_{12} \dots y_{1n1}$	R_1	n_1
2	$y_{21}\ y_{22} \dots y_{2n2}$	R_2	n_2
..	$\dots \dots \dots \dots \dots \dots \dots .$	..	..
k	$y_{k1}\ y_{k2} \dots y_{knk}$	R_k	n_k

Verilerin oluşturduğu tabloda orijinal değerlerin yerine rankları yazılır. Bunun için tüm gözlemler birleştirilir ve birleştirilen gözlemler küçükten büyüğe sıralanır ve en küçük değere 1, en büyük değere n rankı verilir. Burada n, her bir gruptaki gözlem sayılarının toplamıdır. (Kıroğlu, 2001).

Test istatistiği :

- $KW = \frac{12}{n(n+1)} \sum_{i=1}^{k} \frac{{R_i}^2}{n_i} - 3(n+1)$

k = örneklemlerin sayısı,

n_j = j. Örneklemdeki olguların sayısı,

N = $\sum n_j$, birleştirlmiş tüm örneklemlerdeki olgu sayısı,

R_j = j. Örmeklemdeki mertebelerin toplamı

- Eğer veri setinde tekrarlı gözem varsa tekrarlı gözlemler için düzeltme terimi uygulanır.

Düzeltme terimi = 1- $\frac{\sum T}{N^3 - N}$

T = t^3-t

t : tekrarlı gözlem sayısı

$$KW_D = \frac{KW}{1 - \frac{\sum T}{N^3 - N}}$$

- Karar kuralı :

$$\begin{cases} k = 3\,, \quad n_j \leq 5 & Ran\ testı\ tablosu \\ k \geq 3, \quad n_j \geq 5 & x^2{}_{(k-1);a} \end{cases}$$

Hesaplanan KW değeri tablo değerinden büyük ise H_0 hipotezi reddedilir. Örneklerin seçildiği k sayıda yığının hepsi aynı medyana sahip değildir denilebilir .

Bonferroni Düzeltmesi

Eğer *m* test sonunda bir karara varılacaksa, sonuç karar hassasiyetinin *α* olabilmesi için, her bir testin hassasiyetinin *α/m* olması gerekir. α'nın yapılacak karşılaştırma sayısına bölünmesine *Bonferroni düzeltmesi* denir. Örneğin karşılaştırılacak grup sayısı 4 ise karşılaştırma sayısı 6'dır. (1-2, 1-3, 1-4, 2-3, 2-4, 3-4"), α=0,05 olarak belirlenmiş ise 0,05/6=0,0083 olarak bulunur. Bu 6 grup için bulunan test değerine ilişkin p değeri doğrudan 0,0083 ile karşılaştırılır. p<0,0083 ise gruplar arasında fark olduğu söylenir.

Eğer g grubumuz var ve 10 adet test yapmamız gerekirse;

$$(1-\alpha)^k = (1-0,05)^{10} = 0,5987 \quad , \quad \alpha = 0,4013$$

Deneysel hata oranı %40,13 olur. Böylece I. tip hata oranı %5'den %40 lara çıkmış olur.

$$BD = \frac{\alpha}{\binom{k}{2}} = \frac{0,05}{\binom{5}{2}} = \frac{0,05}{10} = 0,005$$

Bonferroni Düzeltmesi

Yeni nominal Alpha seviyemiz 0.005 olduğundan, 10 adet testin herbirinin sonucunun anlamlı olabilmesi için, her bir hesaplanan p değerinin 0.005'ten küçük veya eşit olması gerekir.

Bonferroni düzeltmesinin "konservatif", yani gereğinden fazla tutucu olması nedeniyle, istatistiki gücü (Tip II hatayı arttırmasından ötürü) azalttığını hatırımızda tutmalıyız (bu nedenle, karşılaştırılacak grup sayısı 4-5'ten fazla ise, Bonferroni düzeltmesi tavsiye edilmez.

PARAMETRİK OLMAYAN ÇOKLU KARŞILAŞTIRMA TESTLERİ

İKİDEN ÇOK BAĞIMSIZ GRUPLAR

- Kruskal –Wallis Testi

Gözlem Sayıları Eşit

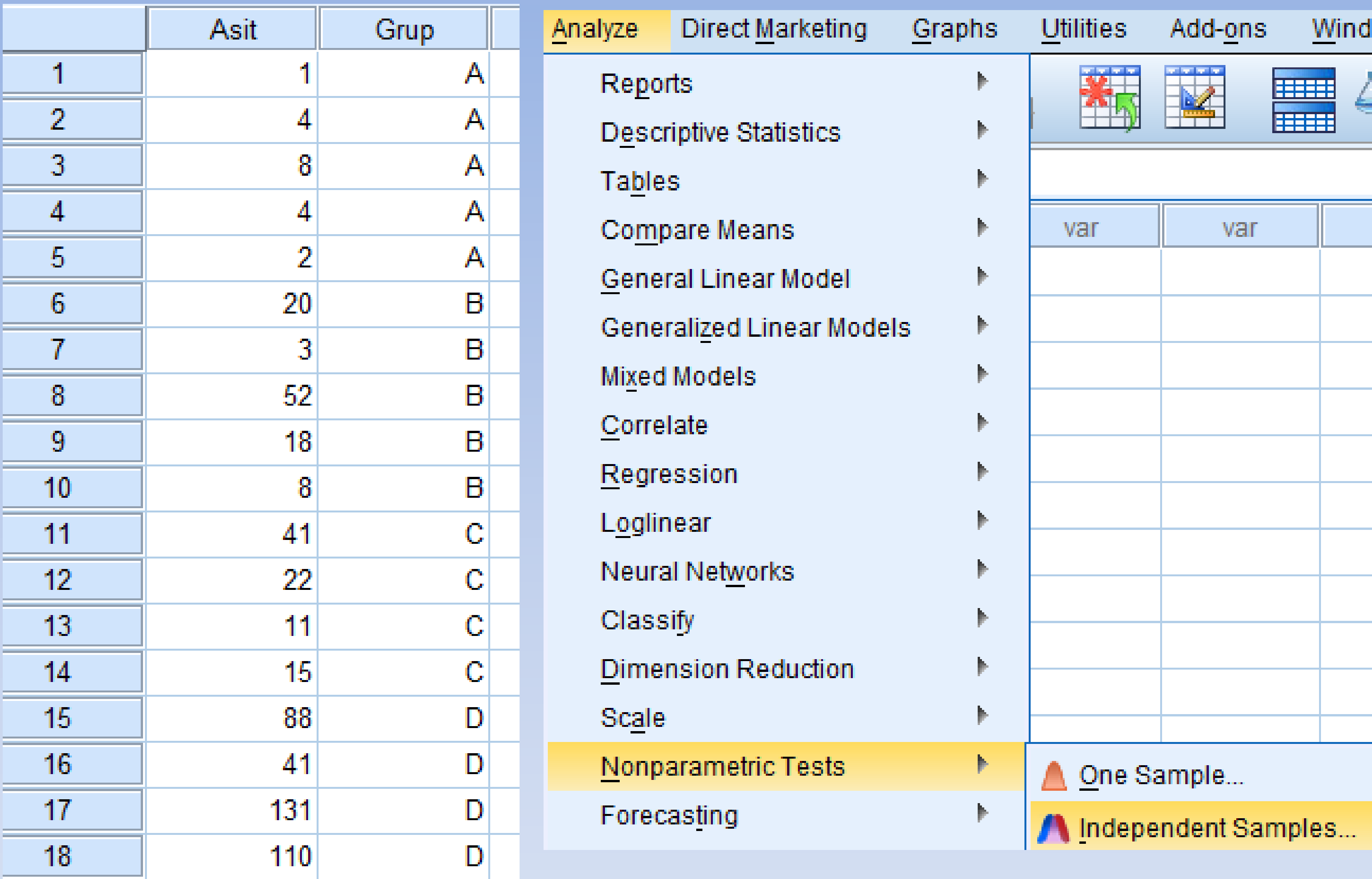
- Tukey Testi
- Student-Newman-Keuls Testi
- Levy Ortanca Testi
- Dwass-Stell Testi
- Nemenyi Testi
- Dunn Testi

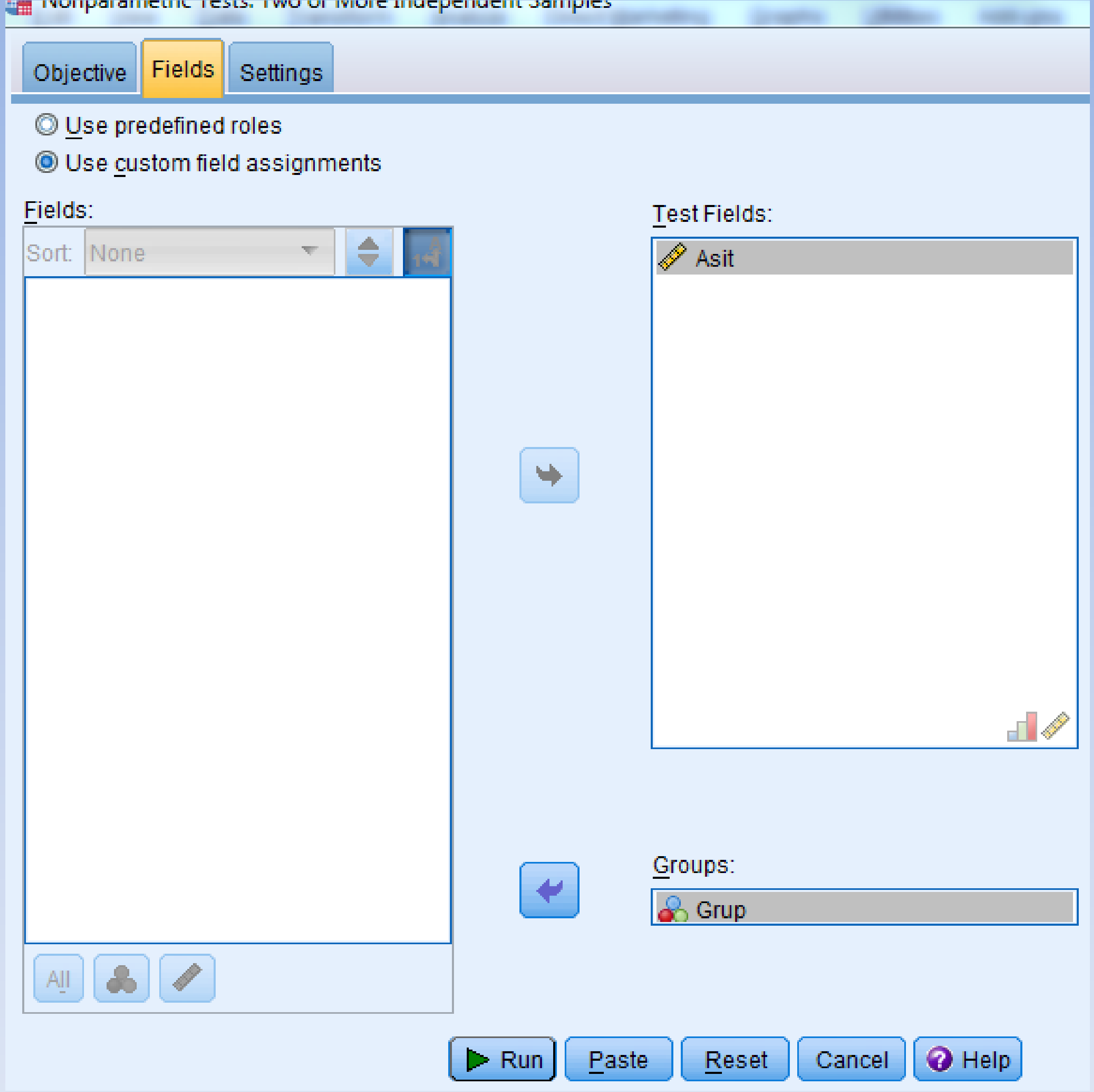
Gözlem Sayıları Eşit Değil

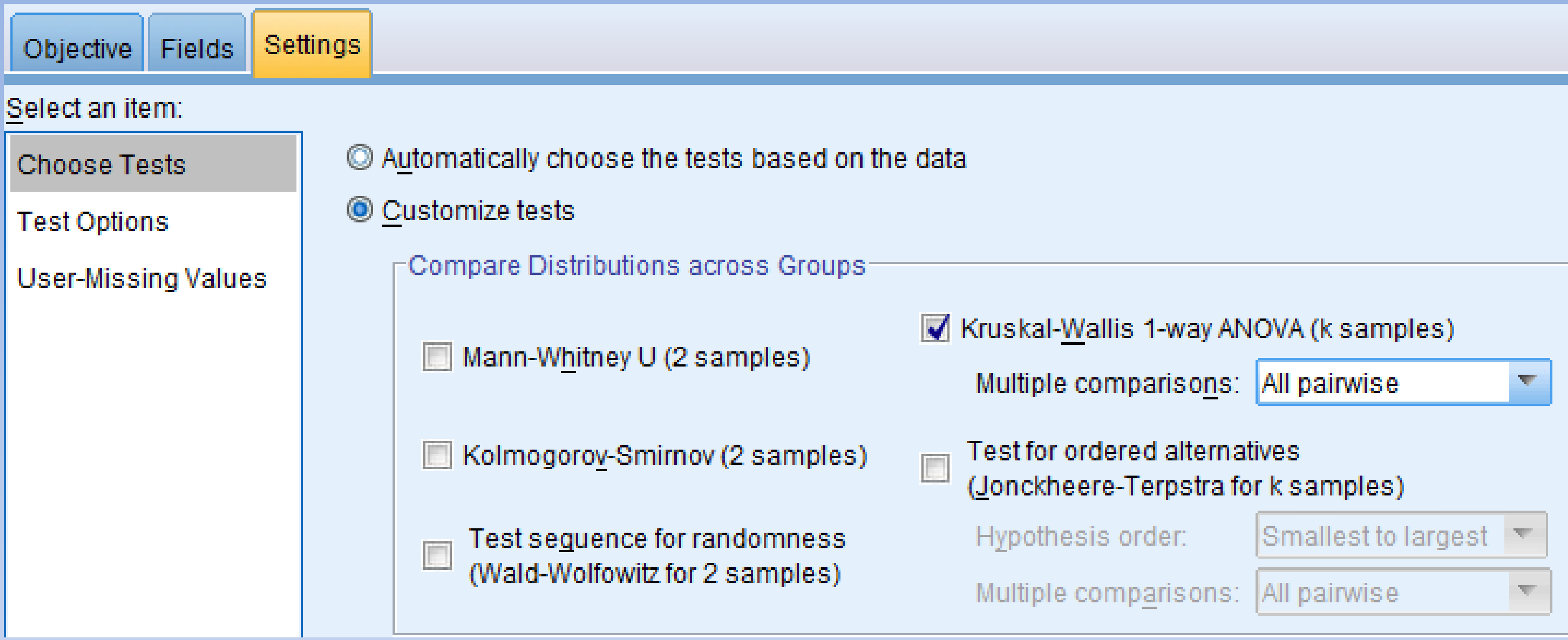
- Conover Testi
- Dunn Testi
- Miller Testi

Örnek : Dört gruptaki hastaların Peak Asit değerleri veriliyor. Bu dört grup ortancaları birbirinden farklı mıdır?

A	B	C	D
1	20	41	88
4	3	22	41
8	52	11	131
4	18	15	110
2	8		


Asit
Grup
1 1 A
2 4 A
3 8 A
4 4 A
5 2 A
6 20 B
7 3 B
8 52 B
9 18 B
10 8 B
11 41 C
12 22 C
13 11 C
14 15 C
15 88 D
16 41 D
17 131 D
18 110 D
Analyze
Direct Marketing
Graphs
Utilities
Add-ons
Reports
Descriptive Statistics
Tables
Compare Means
General Linear Model
Generalized Linear Models
Mixed Models
Correlate
Regression
Loglinear
Neural Networks
Classify
Dimension Reduction
Scale
Nonparametric Tests
Forecasting
One Sample...
Independent Samples...
var
var


Objective
Fields
Settings
Use predefined roles
Use custom field assignments
Fields:
Sort:
None
Test Fields:
Asit
Groups:
Grup
All
Run
Paste
Reset
Cancel
Help

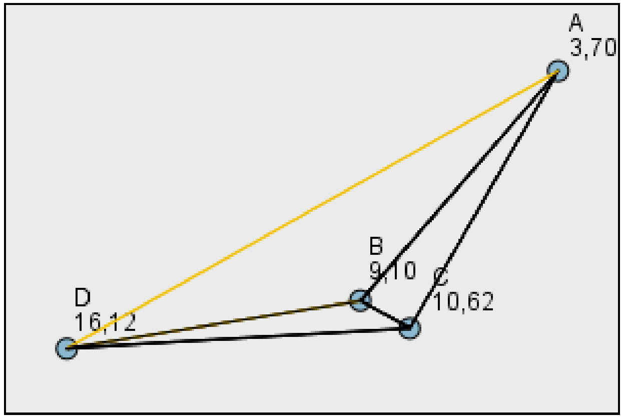


Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of Asit is the same across categories of Grup.	Independent-Samples Kruskal-Wallis Test	,006	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

Pairwise Comparisons of Grup



Each node shows the sample average rank of Grup.

Sample 1-Sam...	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj.Sig.
A-B	-5,400	3,371	-1,602	,109	,655
A-C	-6,925	3,576	-1,937	,053	,317
A-D	-12,425	3,576	-3,475	,001	,003
B-C	-1,525	3,576	-,426	,670	1,000
B-D	-7,025	3,576	-1,965	,049	,297
C-D	-5,500	3,769	-1,459	,144	,867

Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.
Asymptotic significances (2-sided tests) are displayed. The

Pairwise Comparisons | Test: Kruskal-Wallis | Field(s): Asit * Grup(Test 1)

TESADÜFİ PARSELLER DENEME PLANI

Denemenin Kuruluşu: Örneğin arazide, bir ilaç hammaddesi kaynağı olan tıbbi bitki denemesi kurulmak istenirse öncelikle muameleler ve sayısı belirlenir. Sonra her bir muamele (veya bir muamelenin her bir seviyesi) için tekerrür sayısı belirlenir.

Örneğimizde 4 ayrı çeşit muamelemizi oluştursun. Her bir çeşit muamelemizin seviyeleri demektir. Bunların A, B, C, D olarak isimlendirildiğini ve tekerrür sayısının 3 olduğunu varsayalım. Buna

göre,

Muamele sayısı = a = 4

Tekerrür sayısı = b = 3

a x b = 4 x 3 = 12 parsele gerek vardır. Bu durumda arazi 12 eşit parçaya (parsele) ayrılır ve tekerrür sayısı kadar çoğaltılan her bir muamelenin dağıtımı tamamen tesadüfi olarak yapılır.

Tesadüfi parselleri deneme planının özellikleri aşağıdaki gibidir.

1- Homojen bir materyal olmalıdır.

2- Tek yönlü bir heterojenlik söz konusudur. Bu da araştırıcının denemek istediği muameledir.

3- Muamelelerin deneme ünitelerine dağıtımı tamamen tesadüfidir.

Buna göre deneme alanının tamamen homojen olması gerekir. Yani toprak yapısı aynı olmalı, her tarafa eşit miktarda gübre dağıtılmalı, aynı şekilde sulanmalı, yetiştirme ve bakım işleri denemenin tamamında aynı olmalı, denemenin bazı yerleri çukur bazı yerleri tümsek olmamalı, deneme materyalinin tamamına aynı mücadele ilaçları aynı şekilde verilmelidir.

Tesadüfi parselleri deneme planının matematik modeli:

$$Y_{ij} = \mu + \alpha_i + e_{ij}$$

Y_{ij} = i-inci muameleye ait j-inci tekerrürün gözlem değeri

μ = Genel popülasyon ortalaması

α_i = i-inci muamele etkisi

e_{ij} = i-inci muamelenin j-inci tekerrürüne ait tesadüfi hata

$$\text{Genel Kareler Toplamı} = \text{G.K.T.} = \sum_{i=1}^{m} \sum_{j=1}^{n} Y_{ijk}^2 - \frac{Y^2..}{N}$$

$$\text{Gruplar Arası Kareler Toplamı} = \text{GART} = \text{M.K.T.} = \sum_{i=1}^{m} \frac{Y_i^2}{r} - \frac{Y^2..}{N}$$

$$\text{Gruplar içi Kareler Toplamı} = \text{GİİR} = \text{Hata Kareler Toplamı}$$

$$= \text{G.K.T.} - \text{M.K.T.} = \sum_{i=1}^{m} \sum_{j=1}^{r} Y_{ijk}^2 - \sum_{i=1}^{m} Y_{i.}^2 \ / \ r$$

Örnek. Yerleşim sıklığının besi performansına etkisini araştırmak üzere homojen bir sürüden tesadüfi olarak seçilen 5’er buzağı aynı yemle besiye alınıyorlar. Besi sonunda elde edilen canlı ağırlıklar aşağıdaki gibi bulunmuştur.

<u>Yerleşim sıklığı:</u>

A= 21 m^2 , B= 18 m^2 , C= 15 m^2 , D= 12 m^2

Yerleşim sıklığı

Tekerrür	A	B	C	D	
	Y_{11} 123	Y_{21} 127	Y_{31} 127	Y_{41} 129	
	Y_{12} 118	Y_{22} 119	Y_{32} 131	Y_{42} 141	
	Y_{13} 124	Y_{23} 123	Y_{33} 129	Y_{43} 132	
	Y_{14} 125	Y_{24} 117	Y_{34} 143	Y_{44} 140	
	Y_{15} 124	Y_{25} 121	Y_{35} 125	Y_{45} 134	
Σ	**$Y_{1.}$ 614**	**$Y_{2.}$ 607**	**$Y_{3.}$ 655**	**$Y_{4.}$ 676**	**$Y_{..}$ 2552**

Düzeltme katsayısı (D.K.)= $\dfrac{Y^2..}{N} = \dfrac{2552^2}{20} = 325635.2$

Genel Kareler Toplamı (G.K.T.) = $\sum_{İ=1}^{4} \sum_{J=1}^{5} Y_{İJ}^2 - D.K.$

$= (123^2 + ... + 134^2) - D.K.$

$= 326686 - 325635.2$

$= 1050.8$

G.K.T. için serbestlik derecesi= GSD= m(r)-1= 4(5)-1= 19

Yerleşim Sıklığı (Muamele) Kareler Toplamı (M.K.T.)

$= \sum_{i=1}^{4} Y_{i.}^2 / 5 - DK$

$= ((614^2 + ... + 676^2)/5) - DK$

$= 326289.2 - 325635.2$

$= 654$

Muameleye ait serbestlik derecesi=MSD= m-1=4-1= 3

Muamele Kareleri Ortalaması=M.K.T./M.S.D.

$=654/3=218$

H.K.T. = G.K.T. – M.K.T. = 1050.8-654 = 396.8

Hataya ait serbestlik derecesi= HSD= m(r-1)=4(5-1)= 16

Hata Kareleri Ortalaması=H.K.T./H.S.D.

$=396.8/16=24.8$

Varyasyon Kay.	SD	KT	KO	F
Muamele (Yerleşim Sıklığı)	3	654	218	8.79**
Hata	16	396.8	24.8	
Genel	19	1050.8		

Yerleşim sıklığının besi performansına etkisi çok önemli bulunmuştur ($P < 0.01$). Bu test, tesadüf parsellerinde modelin
sabit veya şansa bağlı oluşuna göre değişmez. Modele göre, kareler ortalamalarının beklenen değerleri aşağıdaki gibidir:

Varyasyon Kay.	Model şansa bağlı ise E(K.O.)	Model sabit ise E(K.O.)
Muamele(Yer.Sık.)	$\sigma^2_E + r\ \sigma^2_Y$	$\sigma^2_E + r\,\Sigma\alpha^2_i/(m-1)$
Hata	σ^2_E	σ^2_E

Model şansa bağlı ise yerleşim sıklığına ait varyans unsuru (σ^2_R) ayrıca hesaplanabilir.

$$\sigma^2_E = 24.8$$

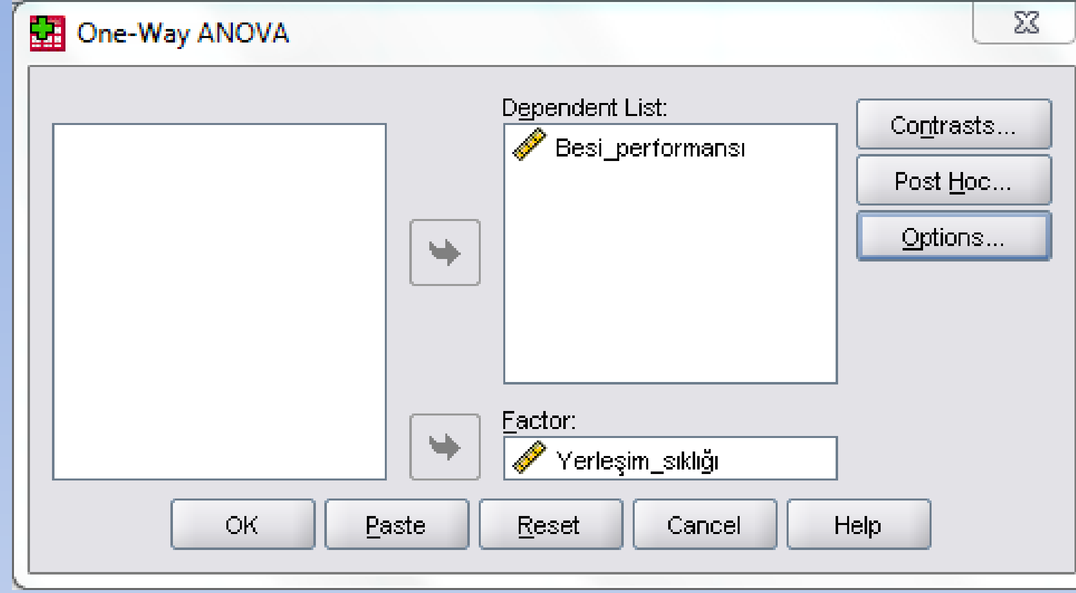
$$\sigma^2_E + r\ \sigma^2_Y = 218$$

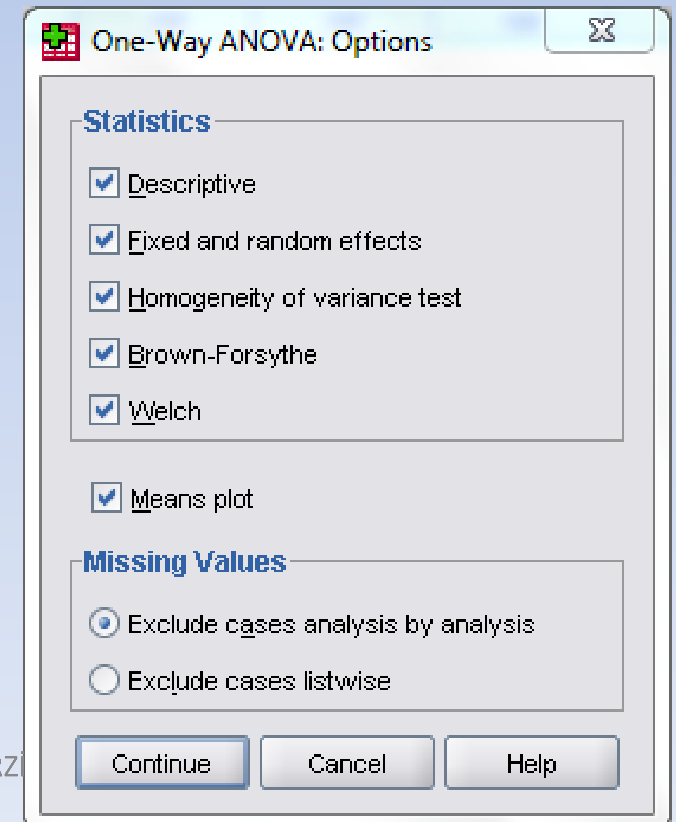
$$\sigma^2_Y = \frac{218 - 24.8}{5} = 38.64$$

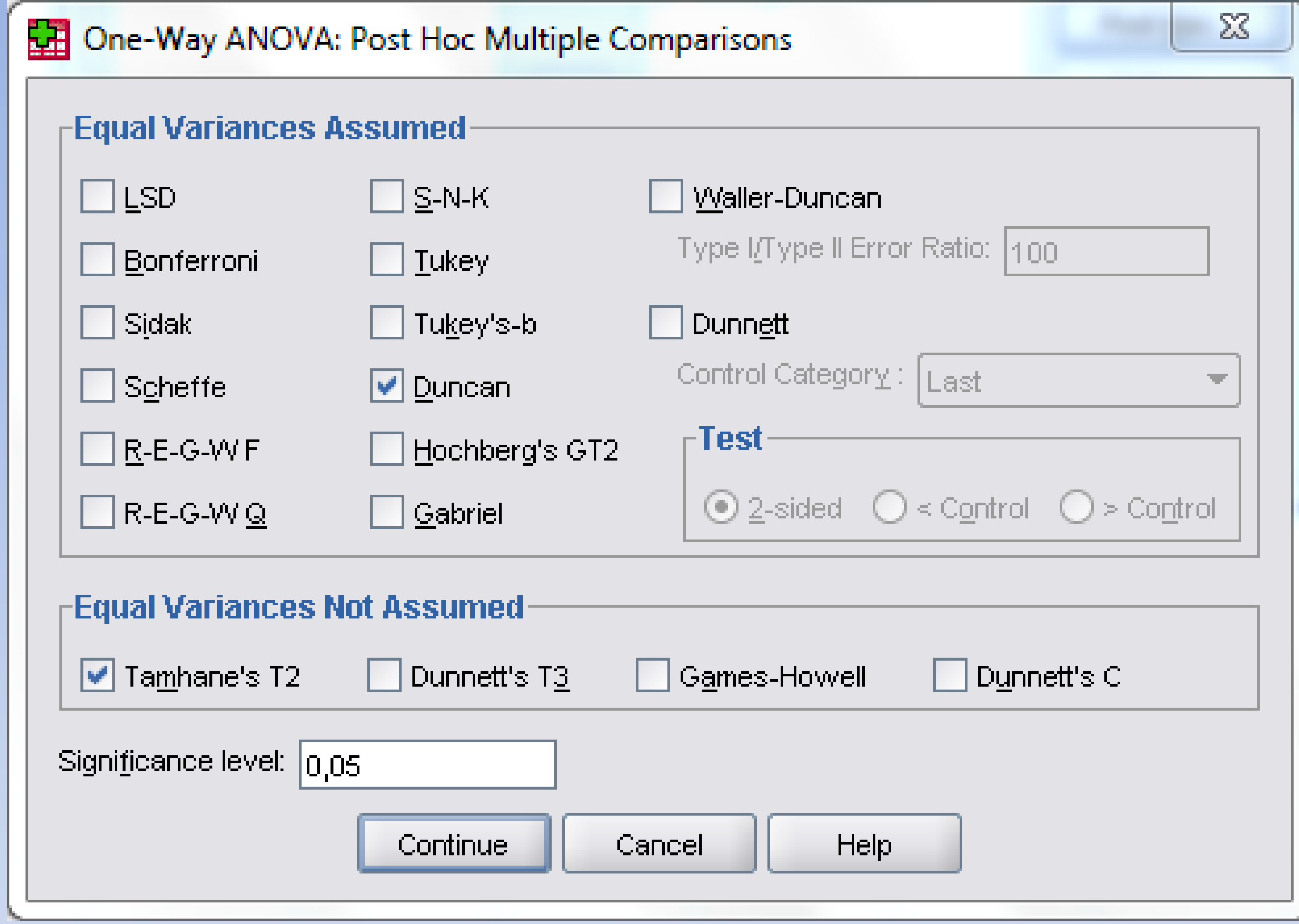
$$\text{Toplam varyasyon} = \sigma^2_E + \sigma^2_Y = 24.8 + 38.64 = 63.44$$

$$\sigma^2_Y \text{'nin toplam varyasyondaki \% payı} = \frac{38.64}{63.44} \text{ x } 100 = \%\ 60.9$$

Besi_performansı	Yerleşim_sıklığı
123	A=21 m.kare
118	A=21 m.kare
124	A=21 m.kare
125	A=21 m.kare
124	A=21 m.kare
127	B=18 m.kare
119	B=18 m.kare
123	B=18 m.kare
117	B=18 m.kare
121	B=18 m.kare
127	C=15 m.kare
131	C=15 m.kare
129	C=15 m.kare
143	C=15 m.kare
125	C=15 m.kare
129	D=12 m.kare
141	D=12 m.kare
132	D=12 m.kare
140	D=12 m.kare
134	D=12 m.kare







Test of Homogeneity of Variances

Besi performansı

Levene Statistic	df1	df2	Sig.
1,034	3	16	,404

Descriptives

Besi performansı

		N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum	Between-Component Variance
						Lower Bound	Upper Bound			
A=21 m.kare		5	122,80	2,775	1,241	119,35	126,25	118	125	
B=18 m.kare		5	121,40	3,847	1,720	116,62	126,18	117	127	
C=15 m.kare		5	131,00	7,071	3,162	122,22	139,78	125	143	
D=12 m.kare		5	135,20	5,167	2,311	128,78	141,62	129	141	
Total		20	127,60	7,437	1,663	124,12	131,08	117	143	
Model	Fixed Effects			4,980	1,114	125,24	129,96			
	Random Effects				3,302	117,09	138,11			38,640

$$\sigma_E^2 = 4{,}98^2 = 24{,}8$$

$$\sigma_E^2 + r\sigma_Y^2 = 218 \Rightarrow \sigma_Y^2 = \frac{218 - 24{,}8}{5} = 38{,}64$$

ANOVA

Besi performansı

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	654,000	3	218,000	8,790	,001
Within Groups	396,800	16	24,800		
Total	1050,800	19			

Besi_performansı

	Yerleşim_sıklığı	N	Subset for alpha = 0.05 1	2
Duncan[a]	B=18 m.kare	5	121,40	
	A=21 m.kare	5	122,80	
	C=15 m.kare	5		131,00
	D=12 m.kare	5		135,20
	Sig.		,663	,201

Means for groups in homogeneous subsets a

a. Uses Harmonic Mean Sample Size = 5,000