

# ANOVA (ANALYSIS OF VARIANCE)

- İki ya da daha çok evrene ait ortalama puanları karşılaştırmada kullanılan güçlü bir istatistiktir.
- Diğer yöntemlerde olduğu gibi ANOVA' da örneklem verisini kullanarak evren hakkında çıkarımlar yapmayı amaçlar
- ANOVA ve t testi aynı amaca yönelik iki farklı yöntemdir
- t testi iki grup ile sınırlıdır
- ANOVA grup sayısı 2'den fazla olduğunda da kullanılabilir.



# ANOVA (ANALYSIS OF VARIANCE)

- t testinde olduğu gibi ANOVA hem bağımsız gruplarda hem de bağımlı gruplarda kullanılabilir.
- ANOVA' da birden fazla bağımsız değişken (gruplama değişkeni) olabilir.
- Bağımlı grup ve bağımsız grup desenleri birleştirilebilir.
- Birden fazla bağımsız değişken kullanmak ve desenleri karıştırabilmek araştırmacılara esneklik sağlar.

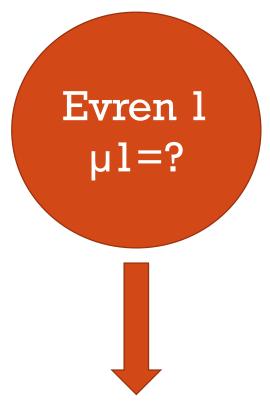


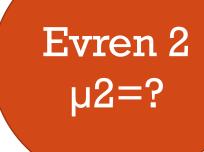
- •İki grup karşılaştırılırken t testi kullanıyoruz.
- •Araştırmamızda 3 grup olduğunda (A,B ve C grupları) bu grupları karşılaştırırken 3 tane t testi yapsak nasıl olur?
- •Önce A ile B'yi sonra A ile C'yi daha sonra da B ile C'yi t-testi alfa=0.05 kriterini kullanarak karşılaştırabilir miyiz?

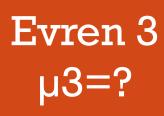
Bir t –testi yaparken alfa kriterimiz 0.05 yani %95 olasılığımız ve
% 5 hata (I.Tür hata) yapma ihtimalimiz bulunmaktadır.

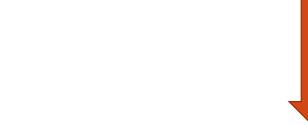
- Eğer 3 tane t-testi uygularsak ve alfa=0.05 alırsak
  - I. Tür hata **yapmama** ihtimalimiz: 0.95x0.95x0.95=.857 olur.
  - I. Tür hata yapma ihtimalimiz de 1-.857=0.143 yani %14.3 olur.
  - Bu olay familywise veya experimentwise hata oranı olarak bilinir.
  - Birden fazla t testi yaptığımızda I. Tip hata yapma olasılığımız artar.













Örneklem 1 n=15  $\bar{X}=23.1$ KT=114 Örneklem 2 n=15  $\bar{X} = 20.5$ KT=130 Örneklem 3 n=15  $\bar{X}=20.8$ KT=114



- İki olası sonuç:
  - (1)Evren ortalamaları arasında bir fark yoktur. Örneklem ortalamaları arasında gözlenen fark örnekleme hatasından kaynaklanır.
  - (2)Evren ortalamaları birbirinden farklıdır ve bu farklılıklar örneklem ortalamalarına yansımıştır.
- Bu iki sonuç aslında bizim hipotezlerimizi göstermektedir
  - H0:  $\mu 1 = \mu 2 = \mu 3$
  - H1: Evren ortalamalarından en az biri farklıdır.
- Çoklu örneklem ortalamaları arasındaki değişkenliği örneklem içindeki değişimle karşılatıran F testi kullanılır.



### TEMEL KAVRAVIAR

 ANOVA'da gruplama değişkeni (Bağımsız değişkenler) faktör olarak isimlendirilir.

 Her bir faktörün (gruplama değişkeninin) değerleri düzey olarak isimlendirilir.

 ANOVA' da birden fazla faktör olduğunda faktöriyel ANOVA olarak adlandırılır.



### VARSAYIMLARI

- Bağımlı değişkenin ölçüldüğü ölçek en az eşit aralık düzeyindedir.
- K örneklem bağımsızdır ve evrenden yansız olarak seçilmiştir.
- Örneklemlerin seçildiği evrenlere ait puanların dağılımı normaldir.
- K evrene ait varyanslar homojendir.
- Bağımsız değişken kategoriktir.



# VARSAYIMLAR-VARYANSLARIN HOMOJENLIĞİ

- t-testinde olduğu gibi ANOVA'da da grupların varyanslarının homojen olduğu yani birbirlerine yakın olduğu varsayılır.
- t-testinde olduğu gibi burada da varyans homojenliğini Levene's test ile kontrol edebiliriz.
- Varyans homojenliği varsayımının ihlali durumunda Brown– Forsythe F (1974), ve Welch's F (1951) istatistikleri kullanılabilir.



# VARSAYIMLAR-VARYANSLARIN HOMOJENLIĞİ

- Varyans homojenliği varsayımının ihlali durumunda Brown–Forsythe F (1974), ve Welch's F (1951) istatistikleri kullanılabilir.
- Brown ve Forsythe grup varyanslarını ağırlıklandırarak (n/N) varyansların homojen olmama durumunu düzeltmeye çalışır ve büyük varyansa sahip olan büyük örneklemlerin etkisini azaltır.
- Brown-Forsythe ile Welch test I.Tür hatayı kontrol etmeyi sağlar.
- Bu iki test arasında Welch testi daha güçlü bulunmaktadır



# VARSAYIMLAR-NORMALLIK

- Burada dikkat edilmesi gereken bağımlı değişken değerinin her grup için normallik koşulunu sağlamasıdır. (Her bir düzeyde)
- Normallik sağlanmadığında:
  - örneklem büyüklüğünü artırmak
  - veriyi dönüştürmek ve dönüştürülmüş veri üzerinde ANOVA yapmak
  - ya da parametrik olmayan testleri (Kruskal Wallis)
- Grup büyüklükleri eşit olduğunda (n1=n2=n3) ANOVA normallik ve grup varyans homojenliği varsayımı ihlallerine dirençlidir.

# ANOVA'NIN MANTIĞI

- F istatistiğini kullanılır
- Varyansları analiz eder
- grup içi değişkenlikleri/varyansı (örnekleme hatası nedeniyle grup içerinde oluşan fark) gruplar arası değişkenlikler (örnekleme hatası ve olası deneysel etkilerden dolayı oluşan değişkenlikler) ile karşılaştırır.



# GRUP İÇİ VARYANS

- Aynı grup içerisindeki bireylerin hepsine aynı deneysel işlem uygulandığında grup içindeki farklılıklar hata olarak kabul edilir.
- Bu farklar sistematik olmayan varyansı temsil eder.
- Bu hatalar ölçüm hatasından veya çalışmanın yürütülmesi sırasında ortaya çıkan tesadüfi etkilerden oluşabileceği gibi bu iki etkiden de kaynaklanıyor olabilir.
- Grup içi değişkenlik her bir grup için evren ortalaması kestirimin elde edilmesindeki karıştırıcı değişkenliği ifade eder.



### GRUPLAR ARASI VARYANS

- Deneysel işlem ile hata varyansı etkisinin birleşimidir.
- Deneysel işlem etkisi, bağımsız değişkene bağlı oldugu düşünülen sistematik değişkenliktir.
- Bu etki, araştırmacının hatalardan izole etmek istediği değişkenliktir.
- Gruplar arası varyans açıklanan varyanstır.





Gruplar arası varyans

-Deneysel işlem

-Hata varyansı

Grup içi varyans

-Hata varyansı



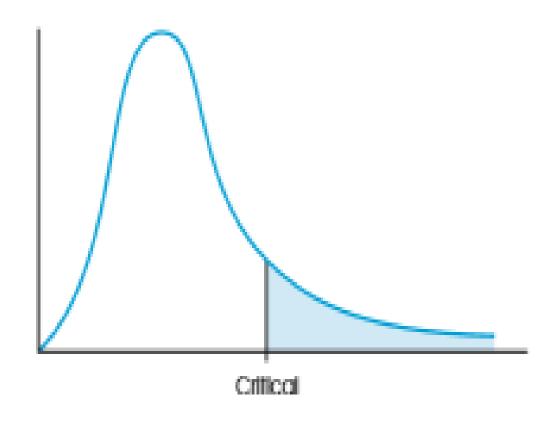
- İlk olarak gruplar arası varyans ve gruplar içi varyans hesaplanır.
- Daha sonra gruplar arası varyans grup içi varyanasa bölünerek Fgözlenen değeri elde edilir.
- F gözlenen F kritik değer ile karşılaştırılıp karar verilir.
- Gözlenen değer kritik değeri aşarsa farkın şanstan kaynaklanmadığı sonucuna varılır.

$$F = \frac{Gruplar\ arası\ varyans}{Grup\ içi\ varyans} = \frac{Sistematik\ etki + rastgele, sistematik\ olmayan\ etki}{rastgele, sistematik\ olmayan\ etki}$$



# F DAĞILIMI

- Örneklem büyüklüğü arttıkça ortalaması l'e yaklaşan pozitif çarpık bir dağılımdır.
- F dağılımı 0 alt sınırına sahiptir.
- Gruplar arası serbestlik derecesi grup sayısı (k)-l
- Gruplar içi serbestlik derecesi N-k





# GÖSTERIMLER

- K= grup sayısı
- N= örneklem büyüklüğü –tüm grup
- n= her bir grubun örneklem büyüklüğü (n1, n2, n3,...)
- T= her bir grup için puanların toplamı (T1, T2, T3, ...)
- KT= kareler toplamı (KT1, KT2 Her bir grup için)
- KT GA= Kareler toplamı gruplar arası
- KT Gi= Kareler toplamı grup içi
- KT Top= Toplam kareler toplamı
- KOGA= Kareler ortalaması gruplar arası
- KOGi= Kareler ortalaması grup ici için



# TER YONLU ANOVA



# TEK FAKTÖRLÜ(TEK YÖNLÜ) ANOVA

Düzeyleri iki veya daha fazla olan tek bir gruplama değişkeni/ faktör

### Örnek Araştırma (Gravetter, 2013):

Bir araştırmacı cep telefonu kullanımının sürücülerin dikkatleri üzerindeki etkisini araştırıyor. Bu kapsamda sürücülerin dikkatleri bir sürüş simülasyonu içerisinde ölçülüyor.

Sürücüler üç gruba ayrılıyor.

- Birinci grup: telefon kullanmadan araba kullananlar
- İkinci grup: telefon kullanarak araba kullananlar
- Üçüncü grup: kulaklık kullanarak telefonda koşanlar.



Bağımlı Değişken: Dikkat

Faktör Değişken (Bağımsız): telefon kullanımı

 Düzeyler: kullanmayanlar, elinde kullananlar, kulaklıkla kullananlar



# HIPOTEZ TESTININ AŞAWALARI

- 1. Evren dağılım özelliklerinin belirlenmesi
- 2. Hipotezlerin kurulması
- 3. Karar kuralının belirlenmesi
- 4. Orneklem istatistiğinin hesaplanması
- 5. Karar verme
- 6. Sonuçların yorumlanması



### 1. Evren Dağılım özelliklerinin belirlenmesi

Bağımlı değişken,

- 1. Sürekli,
- 2. En az eşit aralıklı ölçekte,
- 3. Veri sayısı yeterli (n>=30)
- 4. Dağılımı normaldir.

Gruplar bağımsız (X1 ve X2'nin her kategorisindeki kişiler farklı kişilerdir.)

Bağımlı değişkenin değerleri, her bir kategoride homejendir. Alt gruplarda varyansları eşit kabul edilir.

Bağımsız değişken düzey sayısı 2'den fazla olan kategorik bir değişkendir.



#### 2. Adım hipotezlerin oluşturulması

- H0:  $\mu 1 = \mu 2 = \mu 3$ Sürüş sırasında telefon kullanmak sürücülerin dikkatlerini etkilemez
- H1: Evren ortalamalarından en az biri farklıdır. Sürüş sırasında telefon kullanmak sürücülerin dikkatlerini etkiler

#### 3. Adım karar kriteri

- Manidarlık düzeyi olarak alfa= 0.05 kullanalım
- 0.05 düzeyi için F tablosunda yer alan kritik değerin okunması (Bu örnek için: 2. sütun 12. satır)
- Kritik değer 3.88
- F gozlenen F kritikten buyukse H0 reddedilir.



	Grup l Telefon kullanmayanl ar	Grup 2 Telefon kullananlar (ellerinde tutarak)	Grup 3 Kulakla telefon kullananlar	
	4	0	1	
	3	1	2	
	6	3	2	
	3	1	0	
	4	0	0	
n	n1=5	n2=5	n3=5	N=15
Toplam	T1=20	T2=5	T3=5	T=30
Ortalama	M1=4	M2=1	M3=1	M=1
Kareler toplami	KT1=	KT2=	KT3=	

KTGi	KT1 + KT2 + K3+ + KTk			
KTGA	KTtop-KTGi veya			
	n(KT ortamalar)			
	KT ortalamalar: grup ortalamalarının genel			
	ortalamadan farklarının kareleri toplamı			
Sd gi	N-k			
Sd ga	k-l			



### KARELER TOPLAMI HESAPLAMA FORMÜLLERİ

### 1. Toplam Kareler Toplamı (KT\_t)

$$KT_t = \sum_{i=1}^n X_{ij}^2 - \frac{(T)^2}{N}$$

2. Gruplar Arası Kareler Toplamı (KT\_Ga)

$$\sum_{i=1}^{n} \frac{T_j^2}{n_j} - \frac{(T)^2}{N}$$

3. Grup İçi Kareler Toplamı (KT\_Gi)

$$KT_{Gi} = \sum_{i=1}^{n} X_{ij}^2 - \frac{T_j^2}{n_i}$$



F oranı:	Gruplar arası varyans Grup içi varyans	
Gruplar arası varyans	Gruplar arası kareler toplamı sd gruplar arası	
Grup içi varyans	Gruplar içi kareler toplamı sd grup içi	
Toplam kareler toplamı (KTTOP)	KTGA+ KTGi	
Toplam serbestlik derecesi	Sd gruplar arası + sd grup içi	

#### ANOVA TABLOSU

Değişkenlik Kaynağı	Kareler Toplamı	Serbestlik Derecesi	Kareler Ortalaması (Varyans)	F Değeri
Gruplar Arası	30	2 (k-1)	30/2	Gavaryans /Givaryans
Gruplar İçi	16	15-3=12 N-k	16/12	
Toplam	46	14 N-1	46/14	



#### 3. Adım: Veri toplama ve örneklem istatistiği hesaplama

$$F = \frac{\frac{KTGA}{k-1}}{\frac{KTGi}{N-k}} = \frac{30/2}{16/12} = 11.28$$

#### 4. adım: Karar verme

- Karar: Hesapladığımız F değeri kritik değerden büyük olduğundan  $H_0$  reddedilir.
- Veya p değeri tanımladığımız alfa değerinden küçükse H<sub>0</sub> reddedilir.
- Sonuç: Gruplardan en az birinin evren ortalaması diğerlerinden farklıdır.

### ANOVA'da etki büyüklüğü

- ANOVA analizlerinde yaygın olarak kullanılan en temel etki büyüklüğü ölçümü eta karedir.
- Eta kare bize varyansın yüzde kaçının deneysel etki ile açıklanabileceğini söyler
- Eta kare= Ktga/Kttop
- •0 ile 1 arasında değer alır
- .01, .06 ve .14 'e göre yorumlanır.



### ANOVA'da açıklanan varyans

 Bağımlı Değişkendeki varyansın ne kadarının bağımsız değişken tarafından açıklanabileceğini omega-kare gösterir.

$$w^2 = \frac{KT_{GA} - sd_{GA}KO_{G\dot{1}}}{KT_{TOP} + KT_{G\dot{1}}}$$



### ANOVA

- ANOVA sonucundan elde edilen F değeri bize gruplar arasında deneysel etkiden dolayı oluşan bir fark var mı yok mu onu söyler.
- Eğer F değeri anlamlı çıkarsa ilk düşünmemiz gereken üç grup arasında anlamlı bir fark olduğudur.
- Bu anlamlı fark birçok değişik durumda karşımıza çıkabilir. Örn:
  - 3 grubun ortalamasının birbirinden farklı olabilir.
  - 1 . ve 2. grubunun ortalamalarının eşit olması ama 3. grubun bunlardan farklı bir ortalamaya sahip olabilir.
- Bunun gibi birçok durumdan dolayı F değeri anlamlı çıkmış olabilir. F değeri sadece gruplar arasında anlamlı bir fark olduğunu söyler ve bundan başka ek bir bilgi sunmaz.



## POST HOC

- Varyans analizi sonrasında F değerinin anlamlı bulunması sonucunda hangi ortalamaların birbirinden farklı olduğunu görmek amacıyla post-hoc karşılaştırmaları yapılır.
- Post hoc çoklu karşılaştırma testleri aslında her bir grup çiftini kullanarak yapılan t-testlerinden başka bir şey değildir.
  - Örn: Diyelim ki elimizde 3 grup olsun (A, B ve C). Bu 3 grubu kullanarak yaptığımız ANOVA sonucu anlamlı çıktı ve farkın nereden kaynaklandığını anlamaya çalışıyoruz. Bunu anlamak için 3 ayrı t-testi yaparak A ile B'yi, A ile C'yi, B ile C'yi karşılaştırmamız gerekmektedir.
  - Daha önce uyarmıştık ANOVA yerine 3 ayrı t-testi yaparsak familywise hata oranını artırırız ve I.Tür hata yapma olasılığımız artar.
- Literatürde familywise hata oranını kontrol ederek I.tür hata riskimizi 0.05'te tutmaya çalışan birçok yöntem önerilmiştir.



## POST -HOC

- Önerilen Post Hoc Çoklu Karşılaştırma Testleri arasında en popüler olanı alfa değerini yaptığınız karşılaştırma testleri adedine bölmektir. Örneğin 10tane t-testi uyguladıysanız 0.05 olan alfa değerini 10 a bölerek 0.005 elde ederiz ve alfa değeri olarak 0.05 yerine 0.005 kullanırız. Bu yönteme Bonferroni düzeltmesi denir.
- Post Hoc Çoklu Karşılaştırma Testlerini kullanarak I.Tür hata oranını düşürürken II.tür hata oranını artırıyoruz. II.Tür hata da 1-Power (güç) eşit olduğu için I.Tür hatayı düşürmek analizlerimizin gücünü de olumsuz yönde etkilemektedir. Bu sebepten dolayı Post Hoc Çoklu Karşılaştırma Testleri arasında karar verirken
  - I.Tür hata oranı kontrolü
  - II.Tür hata kontrolü
  - varsayımların ihlali ve testin güvenirliği konularına dikkat etmeliyiz.



- Post Hoc Çoklu Karşılaştırma Testleri üzerine yapılan araştırmalar genelde normallik varsayımı ve varyans homojenliğinin ihmalleri durumu üzerine çalışmışlardır.
- Ortaya atılan Post Hoc Çoklu Karşılaştırma Testleri normallikten küçük sapmalar olması durumundan çok fazla etkilenmemektedir.
- Fakat grupların popülasyon varyanslarının ve eleman sayılarının eşit olmaması durumu Post Hoc Çoklu Karşılaştırma Testlerini etkilemektedir.



- Hochberg's GT2 ve Gabriel's pairwise testleri grup eleman sayılarının eşit olmadığı durumlar için geliştirilmiştir.
- Popülasyon varyanslarının farklı olduğu durumlar için birçok Post Hoc Çoklu Karşılaştırma Testi geliştirilmiştir:
- Tamhane's T2, Dunnett's T3, Games-Howell ve Dunnett's C.
- Tamhane's T2 is daha katı (conservative) and Dunnett's T3 ve C çok küçük I.Tür hata kontrolüne sahiptir.
- The Games-Howell prosedürü örneklemler küçük olduğunda daha esnek (liberal) fakat grup eleman sayıları eşit olmadığında da doğru sonuçlar vermektedir.



- Eşit sayıda grup elemanlarına ve varyans homojenliğine sahipseniz REGWQ or Tukey's HSD testlerini kullanabilirsiniz.
- Eğer grup eleman sayıları arasında çok az fark varsa Gabriel's procedure kullanılabilir.
- Eğer grup eleman sayıları çok farklı ise Hochberg's GT2.
- Eğer varyans homojenliği hakkında şüpheler varsa Games–Howell procedure kullanılabilir.
- Eğer çok tutucu iseniz Bonferroni tercih edilebilir.
- Bunlara ek olarak birçok test bulunmaktadır (Dunnett test )



• Scheffe Testi : post hoc karşılaştırmalarında en yaygın olarak kullanılan testlerden biridir. Scheffe testti hem ortalama çiftlerini hem de ortalamaların karışık kombinasyonlarının karşılaştırılmasına izin veriri.

 Tukey testi: ortalamaların ikili karşılaştırmasında Scheffe testine göre daha güçlüdür.



Çizelge 1. Varyansların eşit olması durumunda seçilebilecek post-hoc test istatistikleri

Test Turu	Post-hoc	Varyans Eşit	Örneklem Eşit	Örneklem Eşit Değil
laştırma stikleri	LSD	X		X
	Sidak	X		X
	Bonferroni	X		X
ary atte	Tukey HSD	X	X	
Çoklu Karşılaştırma Test istatistikleri	Hochberg's GT2	X	X	21000
	Gabriel	X		X
	Scheffe	X		X
Çoklu Aralık Test istatsitikleri	SNK	X	X	
	Tukey's B	X	X	
	Duncan	X		X
	R-E-G-W-F	X	X	
	R-E-G-W-Q	X	X	
	Waller Duncan	X		X
	Dunnet	X	X	80

Çizelge 2. Varyansların Eşit Olmaması Durumunda Seçilebilecek POST-HOC Test istatistikleri

Post-hoc	Varyans Eşit Değil	Örneklem Eşit Değil	
Games-Howell	X	X	
Tamhane's T2	X	X	
Tamhane's T3	X	X	
Dunnet's C	X	X	
Dunnet's T3	X	X	



# RAPORLAMA

ANOVA sonuçlarını rapor ederken mutlaka

- F değeri,
- serbestlik dereceleri ve
- p değerini rapor etmeliyiz.
- Eta-squared (eta-kare) etki bütüklüğü



# ÖRNEK

- Bir mühendislik fakültesinde birinci sınıfların temel matematik dersine giren bir öğretim üyesi, 4 ayrı fen lisesinden gelen öğrencilerin problem çözme becerilerine bakarak, bu liselerdeki matematik öğreti minin niteliği hakkında fikir sahibi olmak istemektedir.
- Öğrencilerin geldikleri Fen lisesinin matematik başarıları üzerinde anlamlı bir etkisi var mıdır?



# ÖRNEK

Varyansın Kaynağı	Kareler Toplamı	sd	Kareler ortalaması	F
Gruplar arası	KT GA	Sdga	KOGA	KOGA
Grup içi	KT Gi	Sdgi	KOGi,	KOGi,
Toplam	KT top	sdtop		



Varyansın Kaynağı	Kareler Toplamı	sd	Kareler Ortalamsı	F	p	Analmlı fark
Gruplar arası	1036,314	3	345,438			
Gruplar içi	17344,174	154	112,625	3,067	0,030	A-C
Toplam	18380,487	157				

Dört farklı fen lisesinden gelen öğrencilerin matermatik başarıları arasında fark olup olmadığını incelemek üzere yapılan test sonucunda A fen lisesinden gelen öğrencilerin Ortalamalarının (Xa=71,95), B fen lisesinden gelen öğrencilerin ortalaması (Xb=66,21), C Fen lisesinden gelen öğrencilerin (Xc=65,11) ve D fen lisesinden gelen öğrencilerin (Xd=68,64) en az ikisi arasında istatistiksel olarak anlamlı fark gözlenmiştir (F(3,154)=3,07,p<0.05).

Test sonucu hesaplanan etki büyüklüğü bu farkın orta düzeyde olduğunu göstermektedir.



# KRUSKAL-WALIS H TESTI

- Veriler aralık ya da oran ölçeğinde olmakla birlikte normallik varsayımını karşılamıyorsa
- Aralık ya da oran ölçeği düzeyindeki ölçümlerin sıra değerlerine dönüştürülmesi ile ikiden fazla gruba ilişkin sıra ortalamalar arasında fark olup olmadığının araştırılmasında kullanılır.
- Bu test, incelenen değişken altında yatan sürekli dağılımı varsayar.
- Kruskal wallis testinde k örneklemin aynı evrenden ya da ortalamalar açısından eşit evrenlerden geldiği şeklindeki  $H_0$  hipotezini test eder.



- Kruskal Wallis Testi çoklu karşılaştırma seçeneği içermemektedir.
- Eğer ikiden fazla grup arasında Kruskal Wallis test sonucu anlamlı bir fark bulunursa tüm grupların olası ikililerinin Mann-Whitney U testi ile kıyaslanması yapılır.



 Bir öğretim üyesi 3 farklı mühendislik bölümünde öğrenim gören toplam 14 öğrencinin bilişsel esneklik ölçeğinde aldıkları puanları karşılaştırmak istemektedir.

Makine	Gıda	bilgisayar
120	140	140
130	130	135
125	145	160
135	150	175
130	135	



Gruplar	N	$X^2$	р	Anlamlı fark
Makine	5	7.51	0.023	
Gıda	5			
İnsaat	4			

