



## **PSİKOLOJİDE İSTATİSTİK DERS NOTU:**

### **RAPOR ÖRNEKLERİ VE DERS NOTU NİTELİĞİNDE ÖĞRENCİLERE EK AÇIKLAMALAR**

**DR. ÖĞRETİM ÜYESİ. HANDAN NOYAN**

**BEYKOZ ÜNİVERSİTESİ SBF, PSİKOLOJİ BÖLÜMÜ.**

**E-MAİL: handannoyan@beykoz.edu.tr**

#### **RAPOR 1: PEARSON Kİ KARE TEST SONUCU.**

**Ne zaman kullanırız?** İki nominal ölçüm düzeyindeki (kategorik) değişkenler arasında sıklık dağılımları açısından istatistiksel olarak farklılığın incelenmesi. Bir başka deyişle, hem bağımlı hem de bağımsız değişken nominal/kategorik ölçüm düzeyindedir. Yani, değişkenler alt sınıflara ayrılmıştır. Bu alt sınıfların yüzdelik ve sıklık gibi tanımsal istatistik bilgilerini göstermek amacıyla “çapraz tablo (crosstable)” seçeneği sunmakla beraber, sıklık dağılımları açısından anlamlı bir fark olup olmadığını veren Pearson Ki Kare ( $\chi^2$ ) test istatistiğini de çıkartabiliriz. Bu istatistiksel test sonucu bize neden-sonuç ilişkisi vermez. Ancak, iki değişkenin düzeylerindeki sıklıklarının istatistiksel açıdan farklı olup olmadığı bilgisini verir. Özellikle demografik bilgilerin analiz edileceği betimsel istatistiksel analizlerde sıklıkla başvurulur. Örneğin, kadın ve erkeklerde sigara kullanımına evet ve hayır yanıtı verenlerin nasıl farklılaştığını merak ediyor olabilirsiniz vb. Şimdi varsayımsal OKB datası üzerinden yine varsayımsal bir rapor örneği sunalım:

**Ki Kare Analizi Gerektiren Soru Örneği:** OKB tanısı almış hasta ve sağlıklı kontrol grupları cinsiyet dağılımları açısından istatistiksel açıdan farklılaşmakta mıdır? Gruplara göre cinsiyet dağılımları nasıldır?

#### **Önemli Açıklamalar**

Bağımsız (neden) değişkeni: Tanı Grubu

Bağımlı (sonuç) değişken: Cinsiyet

Her iki değişkenimiz de iki farklı düzeye sahip nominal ölçüm düzeyinde, kategorik değişkenler.

\*Burada 100'e toplananları daha iyi görebilmek adına bağımsız değişkeni satır, bağımlı değişkeni sütunda toplananlar önerilir. Daha az kafa karıştırıcı olacaktır\*.

Ki kare testini en iyi tanımlayan istatistikler, örneklem birim sayısı ve frekans (sıklık) yüzdeliğidir. Çünkü kategorik bir değişkeni ancak sayı ve yüzdelikle tanımlayabiliriz. O yüzden de tanımlayıcı istatistikler bu test için tabloda ve metinde örneklem sayısı (n) ve yüzdelik (%) ile gösterilir. Test istatistik değeri ise  $\chi^2$  ile gösterilir.

### Analiz Basamakları

SPSS>> analyze>> Crosstabs>> Statistics>> Chi-Square>>Cell>> Observation>>Raw, Column, Total. Satır ve sütunlarda yüzdeliklerin gösterimi anlamına gelir.

### Sonuç Raporu ve Çapraz Tablo Gösterimi

Bu basamakların ardından elde ettiğimiz ki kare sonucu aşağıdaki şekilde raporlanmış ve tablo üzerinden gösterilmiştir:

**Tablo 1.** Tanı gruplarının cinsiyet sıklık dağılımları ve istatistiksel farklılıkları

Tanı Grubu	Cinsiyet		Test Değeri $\chi^2$	p (Anlamlılık Düzeyi (2-uçlu)*)
	Kadın n (%)	Erkek n (%)		
Hasta	5 (%50)	5(%50)	0.000	1.00
Kontrol	5 (%50)	5 (%50)		
Toplam	10 (%100)	10 (%100)		

\* Burada anlamlılık düzeyimizin 0.05 önem düzeyinde 2-uçlu belirlediğimizi varsayıyor ve anlamlılığı bu doğrultuda değerlendiriyoruz. Görüldüğü gibi p olasılık değeri > 0.05, yani belirlediğimiz önem düzeyinden büyük, gruplar arasında zaten gözlenen sıklık değerlerinden de (her birinde eşit gözleyeceğimiz gibi, %50 ve %50 şeklinde) anlaşıldığı gibi istatistiksel olarak da anlamlı bir fark yok.

Sütundaki toplam bize cinsiyet içerisinde bir başka deyişle kadın ve erkeklerde hasta-kontrol oranını verirken, satırdaki toplam değer (yazsaydık, gerek yoktur toplamları yazmaya genel olarak, sadece hangisinin içinde baktıysak toplamın 100 olmasına dikkat ediniz, herhangi bir yanlışlık yapmamak için) bize tanı grubu içerisinde cinsiyet oranını verecektir.

**Raporu:** Kadın ve erkeklerde tanı grubuna ilişkin hasta ve kontrol oranları benzer bir şekilde dağılmıştır; cinsiyete göre tanı grubuna ilişkin sıklık dağılımları istatistiksel olarak anlamlı düzeyde farklı değildir, bir başka deyişle cinsiyete göre grupların sıklık dağılımları benzerdir ( $p>0.05$ ) (Tablo 1) (Eğer anlamlı fark bulmazsak test değerlerini ve çıkan anlamlılığı yazmanız şart değildir, sadece p anlamlılık değerinin belirlediğimiz önem düzeyinden bu şekilde büyük olduğunu ifade etmek yeterli

olacaktır. Bu, tüm testler için geçerlidir, sonuna da tablo ve grafiğimiz bu şekilde referans verilir, numarası ile.).

## **NORMAL DAĞILIM KOŞULUNUN SAĞLANDIĞI PARAMETRİK TESTLERDE BAĞIMSIZ DEĞİŞKENİN 2 VE 3 DÜZEY/GRUPLAR ARASINDAKİ ORTALAMA FARKLILIKLARININ DEĞERLENDİRİLDİĞİ TEST SONUÇLARI**

### **Önemli Açıklamalar**

Bağımlı değişkenin interval ya da oransal ölçüm düzeyinde değerlendirilmiş olması ve ayrıca bağımsız değişkenin düzeyleri/gruplarının söz konusu ölçüm düzeyinde yaklaşık bir normal dağılım gösterme koşulunun sağlanmış olması beklenir!

Parametrik testler gerçek büyüklük ve küçüklüğü kıyaslamaya izin veren ortalamaları karşılaştırma üzerinden yapıldığından daha güvenilirdir.

Ortalamaları karşılaştırdığımız düzeneklerde bağımsız değişkenin en az iki düzeyi, bağımlı değişkenin sürekli (interval veya oransal) ölçüm düzeylerinde olması gerekir ki grupları karşılaştırabilelim. İki sürekli veya iki kategorik ölçüm düzeyindeki değişkenler nedensel yani istatistikte etki/gruplar arası farklılık olarak tanımladığımız bir sınamaya izin vermezler!

## **RAPOR ÖRNEĞİ 2: BAĞIMSIZ ÖRNEKLEMLERDE T TESTİ (Independent Samples *t* test)**

### **Önemli Açıklamalar**

1. Her iki grubun da ölçülmek istenen bağımlı değişken ölçümü açısından yaklaşık normal bir dağılımı sağlıyor olması gerekir.
2. Bağımsız değişkeninin iki alt kategori olmalıdır.
3. Bağımlı değişken ortalama karşılaştırmaya izin verecek niceliksel, sürekli değişken türünde yani interval ve/veya oransal ölçüm düzeyinde olmalı.
4. Ölçümler birbirinden bağımsız grupları içermeli, bir başka deyişle deneyin farklı koşullarına farklı bireyler alınmış olmalı.
5. Varyansların homojenliği test edilmeli (Bunu SPSS normal *t* testi sınamasında kendisi yapıyor, Levene'in varyansların homojenliği testi test istatistik tablosunda sunuluyor, bunun test değeri  $F'$ 'dir. Sig. (anlamlılık,  $p$  olasılığı) aynen normal dağılım testinde olduğu gibi  $p > 0.05$  olur ise, varyanslar homojenlikten farklılaşmıyor anlamına geleceği için varyanslar homojen olduğunu gösterir ve bunun karşılığındaki *t* test değeri ve anlamlılığı kullanılacaktır,  $p < 0.05$  olursa varyanslar eşit/homojen değildir (**equal variances not assumed**) karşılığındaki *t* değeri ve sig.(anlamlılık) düzeyi alınmalıdır.

Şimdi OKB datası üzerinden t testi gerektirecek bir soru soruyor ve raporluyoruz.

**SORU:** Hasta ve kontrol grupları Beck depresyon puan ortalamaları açısından farklılık göstermekte midir? (Grupların depresyon ölçüm puanlarının normal dağılım sınamasına göre normal dağılım her iki grup için de doğrulanmıştır, çalışmada anlamlılık düzeyi  $p<0.05$ , 2-uçlu olarak belirlenmiştir).

Şimdi verilen bu bilgiler ışığında raporlayacağız. Burada t test sınamasının yaptığı bağımsız değişkenin iki düzeyini yani hasta ve kontrol gruplarının bu araştırma sorusunda sonuç değişkeni olan bağımlı değişken Beck depresyon puan ortalamaları açısından kıyaslamak ve gruplar arasında gözle gördüğümüz ortalama farklılıkların istatistiksel olarak kayda değer olup olmadığını vermek olacaktır).

Bu tarz veri setlerinin en iyi tanımlayıcı istatistikleri ortalama  $\pm$  standart sapma değerleridir (söz konusu ölçümlerde).

### SPSS<<Analyze<<Compara Means<< Independent Samples *t* test

Bunun sonucunda çıkan değerlerden varsayımsal t test sonucu (tablo ve metni içeren)

**Tablo 2.** Gruplara göre Beck Depresyon ölçek puanları ve istatistiksel farklılıklar veya bağımsız örneklemelerde t test sonuçları.

Bağımlı Değişken Ölçümü	TANI GRUBU				
	Hasta (n=10) (Ort. $\pm$ SS)	Kontrol (n = 10) (Ort. $\pm$ SS)	Test Değeri <i>t</i>	Serbestlik Derecesi ( <i>sd</i> )	<i>p</i> (Anlamlılık Düzeyi) (2-uçlu)*
Beck Depresyon Puanları	18 $\pm$ 3.29	15.3 $\pm$ 2.54	2.05	18	0.05

SS: Standart sapmayı gösterir. Sd (ing: df): Serbestlik derecesini ifade eder. T testi için, toplam örneklem sayısından grup sayısının çıkartılmasıdır, bilinmeyen evren örnekleme dair bir tahmin yürütür. Test değerlerinin yanındaki alt paranteze yazılır.

Bağımsız örneklemelerde *t* test sonucuna göre; hasta ve kontrol gruplarının Beck depresyon skorlarının ortalamaları arasında trend düzeyde bir farklılık vardır ( $t_{(18)} = 2.05$ ,  $p \leq 0.05$ ). Buna göre, hastaların depresyon puanlarının ortalama değerlerinin kontrol grubunkinden trend düzeyde yüksek olduğu gözlenmektedir (Tablo 2). (NOT: Burada anlamlı kabul etsek de eğer başta belirlediğimiz anlamlılık düzeyi eşiği  $p=0.05$  kapsamıyorsa anlamlı diyemiyoruz ama neredeyse anlamlı olacak düzeyde trend bir farklılık,  $0.05 > 0.09$ 'a kadar trend bir farklılık olarak kabul görebilmektedir).

### RAPOR 3: 3 VE 3'TEN DAHA FAZLA GRUP ORTALAMALARININ PARAMETRİK SINAMASI

#### TEK YÖNLÜ VARYANS ANALİZİ (İNG.: ONE-WAY VARYANS ANALİZİ, ANOVA)

Bağımsız gruplar arasındaki ortalama farklılıkları 3 ya da daha fazla düzeyi olan bağımsız değişkenlerde “tek yönlü/faktörlü varyans analizi ile test edilir. Varyans analizleri ANOVA olarak adlandırılır. Tek-

yönlü ANOVA’da hala bir bağımsız değişken, bir bağımlı değişkenimiz söz konusudur. Burada bir bağımsız değişkenin yol açtığı kareler toplamı ve gruplar içi kareler toplamının (hata olarak adlandırılır) hesabı yapılır. F sınaması bize, buna dair bir sonuç verir. Gruplar arası varyans bağımlı değişkendeki değişimin kendi bağımsız değişkenimizden kaynaklandığının ifadesidir. Grup içi varyans ise, bağımlı değişkendeki değişimin/varyansın bizim bağımsız değişkenimiz dışındaki faktörler tarafından ya da katılımcılar arasındaki bireysel farklılıklardan kaynaklandığının gösterebilir. Dolayısıyla ne kadar gruplar arası varyans (gruplar arası kareler toplamı) değeri grup içi varyanstan (kareler toplamı) büyük olursa o kadar büyük F değeri ve buna eş değer anlamlı bir istatistik değeri elde ederiz ki bu araştırma hipotezimizi geçerli kılan bir sonuç verir (gruplar arasında ölçtüğümüz değişkenin ortalamaları açısından bir fark olduğuna dair). Varsayımları bağımsız örneklemdeki t testi ile benzerdir. **Ek olarak birkaç önemli nokta varyans analizi için:**

1. Burada da Levene testi sonucunda  $p > 0.05$  bulursak, F sınamasını veremiyoruz, varyansların homojen olmadığı durumlarda kullanılan iki test değeri olan Robust Brown Forsythe (RBF) ya da Welch test istatistiklerine karşılık gelen test değerleri raporlanmalıdır. F sınaması sadece varyanslar eşit olduğu durumda F test istatistiği ile raporlanır!
2. Diğer önemli bir nokta ise, varyans analizleri bize sadece genel grup karşılaştırmasına dair bağımlı ölçümümüzde bağımsız değişkenimize dair bir fark olup olmadığını verir. Yani, karşılaştırdığımız grup sayısı kaçsa (3,4,5..vb.) o gruplar arasında anlamlı bir fark olup olmadığını gösterir. Ama bu anlamlı farklılığın hangi iki gruptan kaynaklandığına varyans analizleri içerisinde olan ikili karşılaştırma testleri (post-hoc test) sınamaları olarak geçen testlerle bakmamız gerekir. Burada da varyansların homojen olup olmamasına göre bir ayrım vardır: Eğer varyanslarımız homojense en çok kullanılan post-hoc Tukey testidir. Bunun sonucu bize hangi iki grup arasında anlamlı bir fark olup olmadığını gösterecektir. Varyanslarımız homojen değil ise, seçeneklerden Tamhane-2 testi seçilebilir. Yine bu da homojen olmayan veri setlerimiz için anlamlı çıkan varyans analizimizi takiben yapılan bir post-hoc sınamasıdır. Eğer, varyans analizi sonucunda gruplar arasında anlamlı bir fark bulmadıysanız, posthoc analizlere bakmanıza gerek yoktur.
3. Önemli bir nokta daha; ikili karşılaştırmalarda bir Bonferroni düzeltmesi yapmanız beklenebilir, bunun ya varyans analizi altındaki Bonferroni test sınaması ya da elle yapılan bir hesaplama ile uygulayabilirsiniz. Bu elle hesaplamada başta belirlediğiniz anlamlılık düzeyiniz (örn. burada 0.05) grup sayısına (örn. burada 4) bölünür ve çıkan sonuç olan 0.01 sizin ikili karşılaştırmalar için anlamlılık eşliğiniz olur. Yani ikili karşılaştırmalarda 0.05’ten değil,  $p < 0.01$  küçük olmasını beklersiniz. Çünkü grup sayısı arttıkça istatistiksel tip 1 istatistiksel hata riski artar. Bunu önlemenin bir yolu olarak ise böyle bir uyarılama beklenir.

**Tek Faktörlü Varyans Sorusu Örneği:** Mesela OKB datasında, hastalarda birlikte yaşanan kişi durumuna göre depresyon puanlarının nasıl farklılaştığını merak etmiş olalım. Burada birlikte yaşanan

kiři statüsü olan bağımsız deęişkenimiz; yalnız, partner, aile-akraba ve dięer olmak üzere 4 farklı düzeyde deęişimlenmiştir. Yani bağımsız deęişkenlerimizin düzeyleri 3'ten fazla. Bağımlı deęişkenimiz ise sürekli ölçüm düzeyinde ortalama karşılaştırmaya izin veren bir ölçüm düzeyi olan Beck depresyon envanteridir.

**Öncelikle Levene test sonucumuzu kontrol edelim : anlamlılık  $0.23 p > 0.05$  = yani, varyanslarımız homojendir! F test istatistięini kullanabiliriz.**

SPSS çıktısında gruplara göre anlamlı farkı ANOVA tablosunda "Between-groups (gruplar-arası)" seçeneęini takip ederek burada  $F$  test deęeri ve karşılığındaki  $p$  yani sig. deęeri ile gözleriz. Burada  $F = 0.66$ ,  $p = 0.60$  yani  $p > 0.05$  anlamlı bir fark yok. Eęer varyanslarımız homojen çıkmasaydı dięer test istatistiklerimiz hemen alttaki tabloda olacaktı, SPSS outputun içerisinde.  $F$  sınavasını verirken gruplar arası (between- groups) ve grup içi (within-groups)  $df$  (yani serbestlik derecelerini de raporlarız, gruplar arası  $df$  toplam grup sayısının -1'e denk gelir, grup içi ise, toplam kiři sayısının-toplam grup sayımız sonucuna denk gelir, bunlar SPSS outputunda hazırdır, tek yapacağınız  $F$  testinin alt parantezine önce gruplar arası sonra grup içi varyansın serbestlik derecelerini yazarak raporlamak olacaktır). Şimdi raporlayalım:

#### **Tek-yönlü ANOVA raporu:**

Bu çalışmadaki hastaların birlikte yaşanan kiři durumuna göre depresyon düzeyleri istatistiksel olarak anlamlı düzeyde farklılaşmamıştır ( $p > 0.05$ ). Ancak, anlamlı rapor gösterimi için şimdi bu sonucu anlamlı bulduğumuzu varsayalım. O zaman: .....anlamlı fark bulunmuştur ( $F_{(3,6)} = 0.68$ ,  $p < 0.05$ ). Eęer bulduğumuz anlamlılık 0.01 düzeyinde küçük ise 0.05 deęil, 0.01 derdik.

**\*\*Bu anlamlılık bize sadece birlikte yaşanan kişiler arasında depresyon ortalamaları açısından bir fark olup olmadığını verecektir\*\*** ve bizim anlamlı çıkan varyans analizini takiben, post-hoc ikili karşılaştırmalar testini yürütmek (örn burada Tukey) hangi iki grup arasında anlamlı bir fark olduğunu göstermemiz gerekirdi. Bu testler SPSS çıktısında pair-wise comparisons (ikili karşılaştırmalar) başlığı altında gözlenir. İki grup arasında anlamlı fark varsa \* simgesi ile gösterilir ve onun karşılığındaki sig. deęeri raporlarsınız. Buna göre yalnız yaşayanların depresyon puanı ailesi ile yaşayanlardan daha düşüktür ( $p < 0.05$ ) gibi. Dięer ikili karşılaştırmalarda ise anlamlı bir farklılık gözlenmemiştir (hepsi için,  $p > 0.05$ ).

**Not:** İkili karşılaştırmalar ortalama farklarını ve anlamlılık düzeylerini verir, anlamlı ise nasıl  $t$  testinde gidip ortalamaları kontrol ederek hangi grubun hangi gruptan istatistiksel olarak daha iyi olduğunu buluyorsak bunda da ortalamalara göre hangi ikili grubunun dięerinden istatistiksel anlamlılığı doğrultusunda düşük ya da yüksek olduğunu görüyor ve raporluyoruz.

Tablolar t testine bezerdir, her düzeyin benzer şekilde ortalama, standart sapma değerleri ile, serbestlik dereceleri, F test değeri (bu defa  $t$  değil,  $F$ ) ve  $p$  değerleri verilir. Bu doğrultuda örnek tablolardan bu sonucu kendiniz tablo haline getiriniz ve bir egzersiz yapınız.

#### **RAPOR 4 : İKİ FAKTÖRLÜ VARYANS ANALİZ SONUCU**

Burada artık iki bağımsız değişken, bir bağımlı değişkeniniz vardır. Bağımsız değişkenlerinizde 3 düzey şartı aranmaz. 2 düzeyi de olabilir. Ancak, bu analiz mutlaka varyansların eşitliği varsayımını gerektirir. Eşit olmadığı durumlarda varsayımı ihlali vardır ve uygulayamazsınız.

İki faktörlü ANOVA yürütmenizin amacı, iki bağımsız değişkenin birlikte değişiminin bağımlı değişken üzerindeki etkisini merak ediyor olmanızdır. Eğer, birlikte değişime dair bir kanıtınız ya da literatür taramanız yoksa analizleri ayrı ayrı tek yönlü ANOVA ya da  $t$  testi ile yürütebilirsiniz.

Şimdi OKB datası üzerinden iki faktörlü bir ANOVA sorusu soralım ve iki bağımsız değişkenin etkileşiminin veya bir başka deyişle birlikte değişiminin kayda değer olabileceğini düşündüğümüzü varsayalım:

Bu doğrultuda tanı grupları ve cinsiyetin OKB puan ortalamaları üzerindeki etkisini merak ettiğinizi varsayalım. Bu bir 2 faktörlü ANOVA örneği sunacaktır bize.

2-faktörlü ANOVA’da SPSS çıktısı bize üç anlamlılık düzeyi gösterir:

1. **İlk bağımsız değişkenin ana etkisine işaret eden test ve anlamlılık değerleri** (Burada tanı Grup, SPSS çıktısında grup karşısındaki F ve sig. değerine bakmalıyız, burada grup F değeri 12.65,  $p$  (sig.) 0.003  $p < 0.01$ , anlamlı fark VAR, yani bu şu demek oluyor grubun ana etkisi anlamlı. Nedir ana etki? Tanı gruplarının düzeyleri arasında ki burada hasta ve kontrol grupları OKB puan ortalamaları istatistiksel olarak anlamlı bir düzeyde farklılaşmaktadır!).
2. **İkinci bağımsız değişkenin ana etkisine işaret eden test ve anlamlılık değerleri** (Burada cinsiyetin ana etkisi, bu şu demek: kadın ve erkekler arasında OKB puan ortalamaları açısından fark var mı yok mu? Yani hasta ve kontrol ayırt etmeksizin tüm kadın ve erkekler arasında OKB puanları değişiyor mu? Bizim sonucumuzda cinsiyet karşısındaki F değeri = 0.09,  $p$  değeri 0.77,  $p > 0.05$ , yani cinsiyetin ana etkisi anlamsız.
3. **Grup\*Cinsiyet gösterimi ile etkileşim etkisine ilişkin test ve anlamlılık değerlerini verir.** Bu şu demek, grup ve cinsiyetin birlikte değişimi OKB puanları üzerinde etkili olmuş mudur? Bir başka deyişle, tanı gruplarının OKB puanlarındaki değişimi, cinsiyet bağımsız değişkeninin düzeylerine göre farklılaşmakta mıdır? Açıklayalım: Hasta ve kontrollerin OKB puanları kadın ve erkekler arasında her iki grupta farklıdır (**Anlamlı etkileşim etkisinin anlamı budur**). Örneğin, kadın hastalarda erkek hastalara göre OKB puanı düşükken, kontrollerde tam tersi söz konusudur gibi bir sonuç ileri analizle çıkabilir. Bakıyoruz, bizim sonucumuzda etkileşim etkisi

için F değeri 0.51, p değeri 0.49, yani  $p>0.05$ . O halde anlamlı bir etkileşim etkisi bulmadık. Anlamlı bir etkileşim etkisi bize sadece hasta ve kontrollerin OKB puanlarındaki farklılığın cinsiyete bağlı olarak değiştiğini ve gruplarda buna göre farklılaştığını verecekti. Tıpkı post-hoc yapar gibi iki grupta cinsiyetin düzeylerinin nasıl farklılaştığını anlamak için hasta ve kontrollerde ayrı ayrı cinsiyete göre OKB puan ortalamasının nasıl değiştiğini sıyanan burada t testi (çünkü iki düzeyi var) bir ek analiz yaparak sonucu buna göre değerlendirecektik. Hastalarda ve kontrollerde kadın ve erkeklerin OKB puanlarının ne yönde farklılaştığını elde edecektik.

4. Son olarak iki bağımsız değişkenin ana etkileri anlamlı bulunduysa yine farkın hangi ikili gruptan kaynaklandığını test eden ikili post-hoc testlere başvurmamızdır.

#### İki faktörlü ANOVA Sonuç Raporu:

Tanı grubunun ve cinsiyetin birlikte OKB puanı üzerindeki etkisi iki faktörlü varyans analizi ile test edilmiştir. Analiz öncesinde iki faktörlü ANOVA için gereken tüm varsayımlar test edilmiştir. İki faktörlü ANOVA analizinin sonucunda, OKB puan ortalamaları üzerinde tanı grubu ana etkisi istatistiksel olarak anlamlı bulunurken ( $F_{(1,16)}= 12.65, p =0.003$ ) cinsiyet ve cinsiyet ile grup değişkenlerinin etkileşim (ortak) etkisi açısından anlamlı bir farklılık gözlenmemiştir ( $p>0.05$ ). Anlamlı bulunan grup ana etkisi sonucunda yürütülen post-hoc Tukey testi sonucunda farklılığın .../. grup arasında olduğu gözlenmiştir ( $p< 0.05$ ) gibi.

Bu doğrultuda hastaların OKB puan ortalaması kontrollerden anlamlı düzeyde yüksek bulunmuştur (Tablo 3).

**(Ek not:** Ancak, burada zaten iki düzey olduğu için 3 ve daha fazla düzey olmadığı için ikili ikili karşılaştırmaya gerek olacak bir post hoc analize gerek yok, çünkü iki grubun ortalamasından ayırt edebiliriz)

**Tablo 3.** Tanı grubu ve cinsiyete göre OKB puan ortalamalarının karşılaştırılması

	TANI GRUBU		Test Değeri <i>F</i>	<i>Sd</i> <sub>1,2</sub>	<i>p</i> (Anlamlılık Düzeyi (2- uçlu)*)
	Hasta (n=10) Ort. ± SS.	Kontrol (n = 10) Ort. ± SS.			
Beck Depresyon Skorları	28.4 ± 2.39	22.6 ± 2.50			
Grup Ana Etki			12.65	1,16	<b>0.003**</b>
Cinsiyet Ana Etki			0.09	1,16	0.77
Grup*Cinsiyet			0.51	1,16	0.49

\*\* $p<0.01$ .

\*\*Tüm bu parametrik ölçümler bağımsız örneklemeler içindi, tekrarlanan (paired/ repeated measures) için tüm bu ölçümlerin tekrarlı ölçüm karşılığı yapılır. Pair-wise t test, repeated measures ANOVA gibi.\*\*



### Örneğin, bir tekrarlı ölçüm *t* test (Paired-Samples *t* test) sonucu, RAPOR 5:

Tekrarlı ölçümlerde *t* testi sonucunda iki ölçüm puanı açısından yüksek düzeyde anlamlı bir fark gözlenmiştir ( $t_{(15)} = 7.30, p \leq 0.001$ ). Buna göre, katılımcıların fost1 ölçümündeki ortalama değerleri “fost2” ölçümüne (57.19 +/-4.64 vs. 31.69 +/- 9.49) göre anlamlı düzeyde yüksek bulunmuştur.

### NORMAL DAĞILIM GÖSTERMEYEN NON-PARAMETRİK SINAMA İLE İKİ BAĞIMSIZ GRUP KARŞILAŞTIRILMASI

Normal dağılım göstermeyen veri setlerinde medyan ve yüzdelik dilimleri göstermek ya da testin içeriğini karşılayan ortalama sıra farkını vermek (İng: Mean Rank) önemli olacaktır.

Bağımsız örneklemelerde iki düzey karşılaştırılması Mann-Whitney U testi, Z test değeri, p anlamlılık düzeyi.

**OKB datasında derste yürütülen örnek (Mann-Whitney U Testi)** (Bağımsız örneklemelerde *t* test karşılığı)

### RAPOR 6 (Non-parametrik Mann Whitney U)

Hasta ve kontrol grupları arasında istatistiksel açıdan trend düzeyde bir farklılık gözlenmiştir ( $Z = -1.71, p=0.09$ ). Hastaların depresyon puan ortalaması sıralaması kontrollerindekinden (MR: 12.75 vs. 8.25) yüksek seyretmektedir (Tablo 4).

**Tablo 4.** Grupların depresyon puanlarının non-parametrik karşılaştırılması

GRUP	Medyan (IQR)*	Z	p (2- uçlu/tailed)
<b>Hasta</b> (n=16)	17.50 (15-25)		
<b>Kontrol</b> (n=16)	15.50 (13-17.25)	-1.71	0.09

\*IQR %25-%75 yüzdelik dilimler aralığını verir, medyanla birlikte sunulur, nasıl ortalama standart sapma değeri ile birlikte sunuluyorsa medyan da bu çeyrek dilimler ranj aralığı ile sunuma uygundur.

### Tekrarlı ölçümlerin non-parametrik test karşılığı, Wilcoxon Testi

**RAPOR 7 Wilcoxon test sonucu, Z test istatistiği ile gösterilir, yine ortalama sıralarını (mean rank) karşılaştırır, bir non-parametrik ölçüm olduğundan:**

Wilcoxon testi sonucunda iki ölçüm puanı açısından yüksek düzeyde anlamlı bir fark gözlenmiştir ( $Z = -3.54, p \leq 0.001$ ). Buna göre, katılımcıların fost1 ölçümündeki sıralamaları fost2 ölçüm sıralamasından anlamlı düzeyde yüksektir (57.19 vs 42.80 ) göre anlamlı düzeyde yüksektir.

**EK NOT:** Tek faktörlü varyans analizin non-parametrik karşılığı Kruskal Wallis testidir, Z istatistiğini kullanır, içerisinde post-hoc test ayrıca barındırmaz, ikili karşılaştırmalar ayrı ayrı non-parametrik testlerle yürütülür.

**HN.**