

GÜVENİLİRLİK ANALİZİ



DOÇ.DR. YÜKSEL TERZİ

**ONDOKUZ MAYIS ÜNİVERSİTESİ
FEN-EDEBİYAT FAKÜLTESİ
İSTATİSTİK BÖLÜMÜ
SAMSUN**

2017

İÇİDEKİLER

1. GİRİŞ

1.1. Geçerlilik ve Güvenirlik İlişkisi

1.2. Ölçme

1.3. Ölçme Hatası

2. GÜVENİLİRLİK ANALİZİ

2.1. Güvenilirlik Analizi Yöntemleri

2.1.1. Test-yeniden Test (Test-retest)

2.1.2. Paralel Formlar Yöntemi

2.1.3. Gözlemciler Arası Güvenilirlik

2.1.4. İç Tutarlılık Güvenilirliği

2.2. Madde Sayısının Tahmini

2.3. Korelasyon Katsayılarının Fisher Z Puanlarına Dönüştürülmesi

2.3.1. Güvenilirlik Katsayıların Karşılaştırılması

3. GEÇERLİLİK ANALİZİ

3.1. Geçerlilik Analizi Yöntemleri

3.1.1. Yüzey Geçerliliği

3.1.2. İçerik Geçerliliği

3.1.3. Kriter Geçerliliği

3.1.4. Yapısal Geçerlilik

3.2. Faktör Analizi

3.2.1. Açımlayıcı Faktör Analizi

3.2.2. Doğrulayıcı Faktör Analizi

1. GİRİŞ

- Bir oluşumun (fenomen) yapısal ve fonksiyonel özelliklerini ortaya koymak için özgün ölçme araçlarına gereksinim vardır.
- Bir oluşumun gerçek biçimiyle yansıtılması için fiziksel yöntemlere dayalı ölçme araçlarının kullanılması gerekir.



- Fakat bazı davranışsal, yargısal ve bilgi-tutum-davranış türündeki oluşumların ölçülmesinde yararlanılan fiziksel araç ve gereçler bulunmamaktadır.
- Bu tür oluşumların belirlenmesi için bazı ölçme araçları geliştirilmiştir. Bu araçlar k sayıda soru içerirler ve bu sorulardan elde edilen cevaplara göre birimlerin davranışsal, yargısal ve bilgi-tutum-davranışlarına ilişkin bilgi edinilir.



Bir X dersinde uygulanan sınav soru kağıdı bir bilgi ölçme aracıdır. Bu aracın öğrencilerin

- bilgilerini yeterince doğru ölçecek ve
- **yeterli sayıda sorudan oluşacak** biçimde hazırlanmış olması gerekir.



- Sınav kağıdındaki soruların kombinasyonu, sınıfın ya da uygulandığı grubun genel bilgi düzeyini ölçecek sayıda, zorlukta olması ve cevaplanması için yeterli sürede uygulanması gerekir. Ayrıca sorular için seçilen puanlama sistemi de grubun, sınıfın gerçek bilgi düzeyini ölçecek bir yapıda olmalıdır.

- Bir testin puanlama biçimi objektif olmalıdır. Sınav sorularının puanlaması ya da bir testin puanlamasındaki objektiflik, soruların ilgili **uzmanların tümü tarafından aynı ağırlıklı puanlamaya tabi tutulacağı anlamına gelmektedir.**



- Örneğin bir sınavda sorulacak soru içeriği
 - 1.çok zor
 - 2.normal ve
 - 3.kolay olmak üzere üç zorluk düzeyinde olmalıdır.
- Bir test, incelenen olayın (oluşum, davranış, yargı, bilgi vb.) gerçek durumunu ortaya çıkarmak için birbirleriyle tutarlı, ilişkili ve yeterli sayıda soru (madde) içermelidir.

Güvenirlik (güvenilirlik);

Bir ölçme aracında (testte veya ankette) bütün soruların birbirleriyle tutarlılığını, ele alınan oluşumu ölçmede türdeşliğini ortaya koyan bir kavramdır.



- Ölçme araçlarının güvenilirliğini değerlendirmek amacıyla geliştirilmiş yöntemlere Güvenirlik Analizi ve bu araçta yer alan soruların irdelenmesine ise soru analizi (Reliability and Item Analysis) denilmektedir.



- Testlerin gvenirliđini analiz etmek amacıyla gvenirlik katsayıları hesaplanmaktadır. Bugne kadar gvenirlik analizi iin bir ok yntem geliřtirilmiřtir.



Güvenilirlik (reliability) ve geçerlilik (validity) analizleri yapılmadan bir arařtırmanın analiz sonuçlarını yorumlamak doęru deęildir. Güvenilirlik ve geçerlilik herhangi bir řeyin uygun ve saęlam olduęu hakkında bize bilgi verir. Örneęin ayakkabının ayaęımıza hem uygun olması hem ayakkabının saęlam olması gerekir. Burada uygunluk geçerlilięi, saęamlık ise güvenilirlięi temsil etmektedir.

Geçerlilik, bir testin ölçmek istediği özelliği ölçmesi ile ilgilidir. Bir test ölçmek istediği özelliği doğru ve diğer özelliklerle karıştırmadan ölçüyor ise bu testin geçerli olduğu söylenir. Geçerli bir testte bulunması gereken en önemli özellik testin güvenilir olmasıdır.

Güvenilirlik bir ölçüm sürecinde ölçüm işleminin tekrarlanabilirliği ya da tekrarlardaki tutarlılıktır.

- Güvenilirlik bir özelliğin bağımsız ölçümleri arasındaki kararlılıktır.

- Güvenilirlik zamana göre değişmezlik ölçüsüdür.

Güvenilirlik bir testin geçerliliğini etkiler. Geçerli bir test muhakkak güvenilir olmalıdır. **Güvenilir bir test geçerli olmayabilir. Yani güvenilirlik geçerliliği garanti etmemektedir.**

Örneğin kol kuvvetini ölçmek için kullanılan barfix testinin güvenilirliği çok yüksektir. Bu test yanlışlıkla bacak kuvvetini ölçmede kullanılırsa geçerli olmayacaktır. Yine 20-45 yaş arasındaki bayanlar için geliştirilen ve belli bir özelliği ölçen güvenilirliği yüksek bir ölçeğin 55 yaş üzerinelere uygulanırsa geçerli olmayacaktır.

Araştırma çalışmaları yapılırken nesnelerin deęişik özellikleri ile ilgili ölçüm yapmak için deęişik aletler, laboratuvar aygıtları kullanılır veya sosyal bilimler, davranış bilimleri gibi bilim dallarında kullanılan anket çalışmaları veya psikolojik testlerde de farklı ölçme aygıtları kullanılır.

Bir araştırma sonucu elde edilen verilere sadece verinin iyi toplandığı için güvenilmez, yöntem kısmında araştırmada kullanılan aletlerin detay tanımları verilir, kullanım amaçları açıklanır, soru biçimleri ve ölçme ölçeęi (adlandırma sıralama, aralık veya oran) açıklanır. Bunların hepsi geçerlilik ve güvenilirlięin sağlanabilmesi amacıyla yönelik gayretlerdir.

Bir arařtırmada geerlilik ve gvenirliđi artırmak iin ařađıdaki dzenlemelerin yapılmasına dikkat edilir (Thomas RK, Mueller RO.)

Sıra no	Yapılması gerekenler	Hangi Blmde Yapılır
1	alıřmanın hedef ve hipotezleri yalın ve sade bir Őekilde tanımlanmalıdır.	Giriř, Yntem
2	alıřmada kullanılan her aygıt yeterli detayda tanımlanmalıdır. Kullanılan aygıtların alıřmanın amalarına uygunluđu aıklanmalıdır. Uygun geerlilik delilleri toplanmalıdır. alıřmanın amacına uygun mevcut toplanan veri setinin desteklediđi geerliliđin eřidi belirlenmelidir.	Yntem ve sonular
3	Gvenirlik ve geerliliđi artırıcı alınan nlemler aıklanmalıdır. nceki alıřmalardaki gvenirlik indeks deđerleri ile bu alıřmanın amalarına uygun Őekilde hesaplanan indeksler deđerler karřılařtırılmalıdır.	Yntem ve sonular
4	Arařtırmada uygulanan gvenirlik ve geerlilik alıřmaları yorumlanmalı ve rapor edilmelidir. alıřmanın sonuları bu mevcut deliller kapsamında tartiřılmalıdır.	Tartiřma

Geçerlilik ve güvenilirlik ölçme aygıtlarının önemli iki özelliğidir.

Güvenirlik ölçümün kararlılığı ile ilgili bir özelliktir. Alınan ölçümlerin kararlı olması veya değişmemesi istenir, bu kararlılık zamandan zamana, mekandan mekana, sorudan soruya, hakemden hakeme değişmeyen bir özelliktir.

Geçerlilik ise ölçmek istenen özelliğin, ölçülmesi niyet edilen hedef değer gerçekten ölçülebilmesi ile ilgili bir özelliktir. Dolayısıyla geçerlilik çalışmanın amacı ve araştırma sorusu (hipotez) ile yakından ilişkili bir husustur.

Araştırmacılar ölçme aracının farklı kültürler, farklı durumlar, farklı koşullarda aynı hassasiyetle ölçüm yapıp yapmadığı ile yakından ilgilidir. Örneğin Asya ırkından çocuklarda oldukça yüksek güvenilirlik ve geçerliliğe sahip bir psikolojik test, Afrika ırkından çocuklarda olmayabilir.

Güvenirlik ve geçerlilik kavramları ve tanımlamaları zaman içerisinde sürekli gelişmekte bunları ifade edebilmek için farklı ölçütler bulunmaktadır.

Geçerlilik mi, Güvenirlilik mi?

Geçerlilik mi daha önemli güvenirlilik mi daha önemli sorusunun yanıtını vermek biraz zordur.

Geçerli olmayan bir test ne kadar güvenilir olursa olsun bir işe yaramaz. Geçerlilik ön plana çıkmaktadır. En azından yapı ve içerik geçerliliği mutlaka sağlanmalıdır. Bir bakıma güvenirlilikte geçerliliğin bir parçası şeklinde düşünmek kavramların ne kadar iç içe olduğunu vurgulamak için yerinde olur. Burada ayrı ayrı konu edilmesinin amacı aralarındaki farklılıkları vurgulamaktır.

1.1. Geerlilik ve Gvenirlik İliřkisi

Gvenirlik, geerlik iin bir n kořuldur. Bir lme aracının gvenilir olması her zaman geerli olduėu anlamına gelmez. Geerlilik, gvenirliėi kapsayan bir kavramdır.

- Geerlilik ile gvenirlik arasındaki gerek fark bir tanım meselesidir.
- **Gvenirlik yapılan lmn kararlılıėını (consistency) veya daha basit olarak aynı řartlarda, aynı nesnelere, deėiřik zamanlarda veya kiřilerce lldėinde benzer sonuların bulunması hadisesidir.**
- Geerlilik ise lmek istenen, amalanan řeyin gerekten llp llmediėi ile ilgilidir yani kesinlik (accuracy) lsdr.

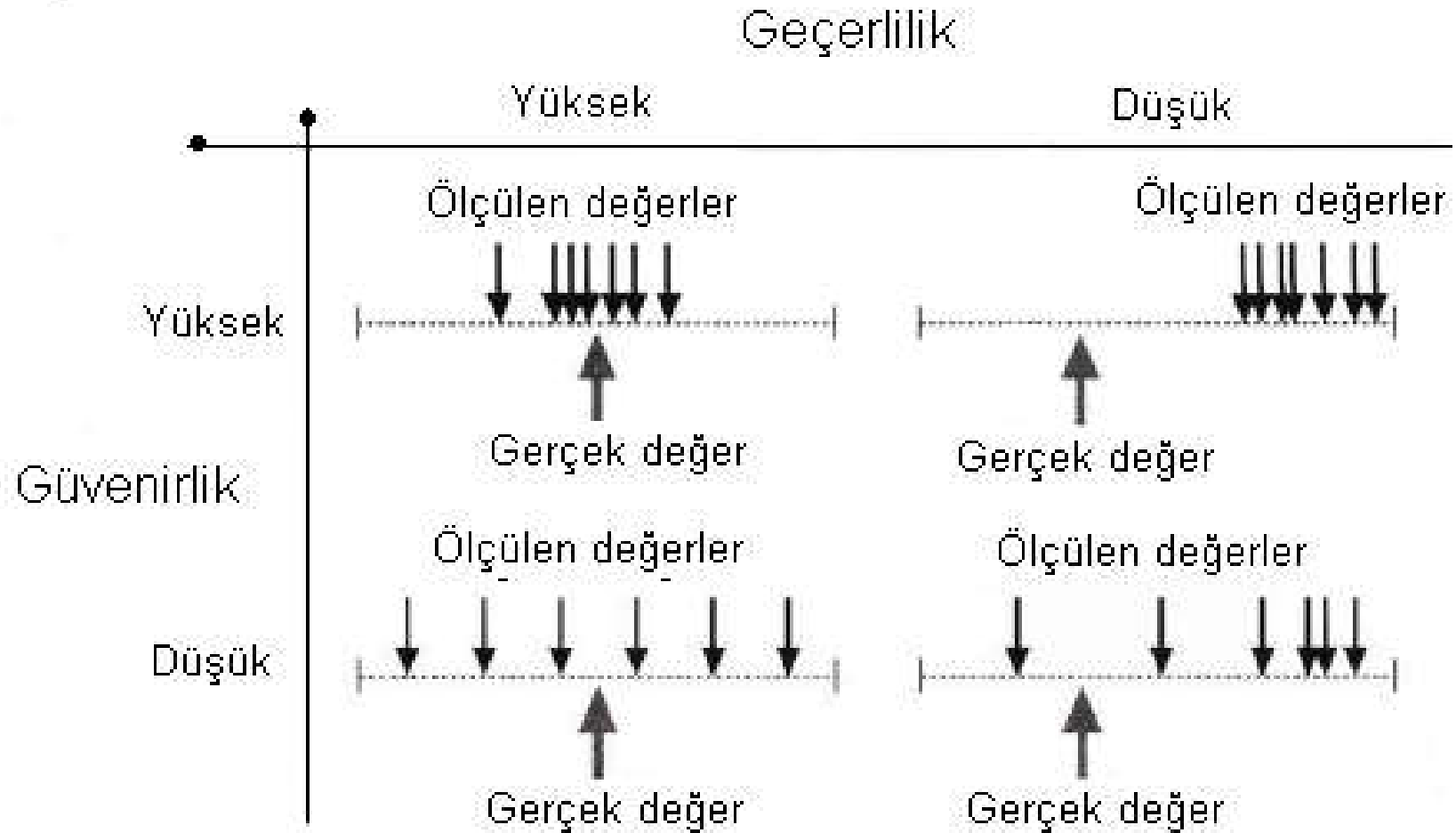
- Amaçlanan hedef için ölçüm yapılmıyorsa ne kadar hassas ölçüm yapılması fazlada bir önem taşımaz, o nedenle geçerlilik daha ön planda düşünülmesi gereken konudur.
- Geçerlilik her türlü hatadan etkilenmesine rağmen, güvenilirlik yalnızca tesadüfî hatadan etkilenir. Bir testin geçerliliği güvenilirlikten etkilenir, bu yüzden testin geçerlilik katsayısı, güvenilirlik katsayısının karekökünden daha büyük olamaz.

Bir testte hata kaynakları kontrol altına alındığı zaman testin güvenilirliği artar. Güvenirliliği artan bir testin geçerliliği de yükselir, ancak bu iki kavram birbirlerinden farklı kavramlardır.

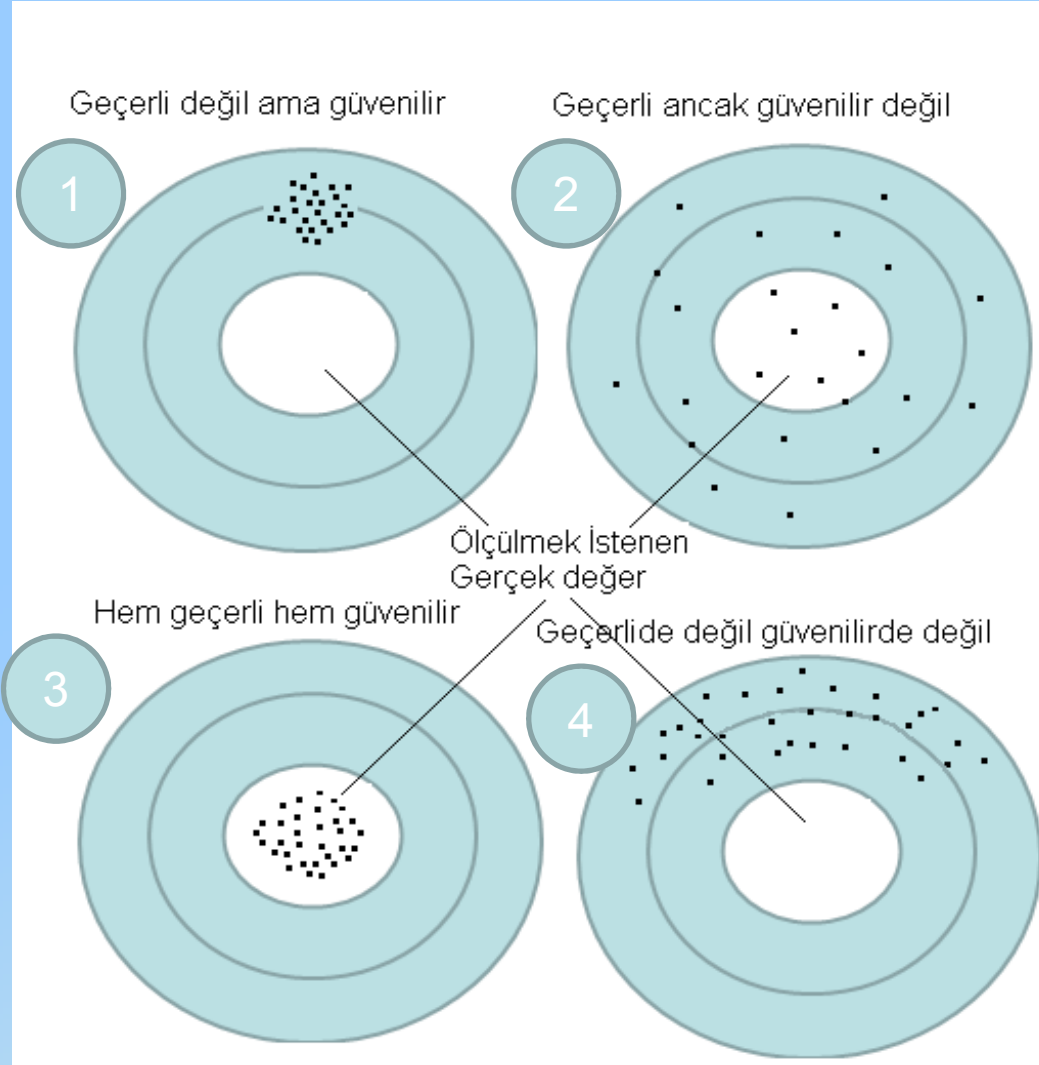
Güvenirlik bir özellik birkaç defa aynı araçla ölçüldüğünde her seferinde benzer sonuçlar elde edilmesi ile ilişkili bir kavramdır.

Geçerlilik ve Güvenirlilik İlişki Şeması

Geçerlilik ve Güvenirlilik



Kavramların şekille ifadesi



1. Her zaman aynı sonucu alıyor, ama istenen sonucu tam kapsamıyor.
2. İstenen sonucun tümünü kapsıyor, ama her zaman farklı sonuç alabiliyor. (Tesadüfî hata yüksek)
3. Her zaman aynı sonucu alıyor ve isteneni tam ve kesin kapsıyor. (Tesadüfî hata düşük)
4. Hem istenen sonucu tam kapsamıyor, hem de her zaman aynı sonucu alamıyor.

1.2. ÖLÇME

Somut veya soyut özelliklerin ölçülmesinde kullanılacak ölçeğin standart bir ölçme aracı olması gerekir.

- Varlıkların ya da olayların özelliklerinin gözlenerek gözlem sonuçlarının sayı veya sembolle ifade edilmesidir.
- Ölçme bir nesneye, bir olguya, tutuma ait özelliği sayısallaştırmak veya sayılabilir simgelerle göstermektir.
- Tanımlanmış kurallara göre belirlenmiş özelliklerin sayısal ifadesidir.
- Ölçme, istatistik birimlerinin ilgilenilen özelliğe sahip olma derecesinin, belirli kurallara uyarak sembolle ve özellikle sayı ile eşleştirilmesidir.

Ölçmenin öğeleri: Varlık, ölçme aracı, ölçme sunucudur

ÖLÇÜM : Ölçme işlemi sonucunda elde edilen sayı veya semboldür.

ÖLÇME KURALI

Ölçmeye konu olan özelliğin hangi miktarına ne değer verileceğinin belirlenmesidir. Örneğin not baremi.

Ölçme belli bir amaç için yapılır. Amaç ölçme konusu olan özellik bakımından bireyler, olaylar yada nesnelere hakkında değerlendirme yapmak ve elde edilen sonuçlara dayanarak belli kararlar vermektir. Sonuçların doğru ve uygun olması için ölçü aracının standardize edilmesi gerekir. Ölçeğin standardize olabilmesi için güvenilirlik ve geçerlilik gibi iki özelliğe sahip olması gerekir.

ÖLÇME VE DEĞERLENDİRME

Ölçme ve değerlendirme aralarında yakın ilişki olan iki kavramdır. Bir test uygulanıp bir skor elde edildiğinde bir ölçüm yapılmış olur. Araştırma konusu ile ilgili sayısal değerleri elde etme işlemine **ölçme**, sayısal değerler elde etmek için kullanılan araç ya da gereçler **ölçek** olarak adlandırılır. Ölçüm sayısal (sürekli-kesikli) ya da nitelik olabilir.

Değerlendirme ise test sonucunda elde edilen bir skorun yorumlanması demektir. Değerlendirme ölçümleri kullanan bir süreç iken ölçmenin amacı bilgi toplamaktır.

ÖLÇMEDE BİRİM

Ölçme araçlarının sayısal değerle ifade edilebilen en küçük parçasıdır. Birim küçüldükçe ölçme aracının duyarlılığı artar, hata miktarı azalır. Ölçmede birim doğal ve tanımlanmış olmak üzere ikiye ayrılır. Bir değişkenin hangi miktarına 1 birim denileceği tanımlanmamış olan birime doğal birim denir.

Örneğin bir sınıftaki öğrenciler sayılırken her öğrenci 1 birim olarak kabul edilir. Tanımlanmış birim ise bilim adamları tarafından sınırları belirlenmiş olan birimlerdir: saat, lira, puan, metre tanımlanmıştır. Ölçme araçlarının özellikleri: Eşitlik, kullanışlılık, genellik

ÖLÇMEDE SIFIR

Doğal ve tanımlanmış olmak üzere iki boyutta ele alınabilir. Bir değişkenin miktar olarak gerçekten yokluğunun ifade eden sifıra **doğal (mutlak)** sıfır denir. Kişinin boyu, ağırlığı doğal sifıra sahiptir.

Bir değişkenin hangi miktarına sıfır denileceği uzmanlarca tanımlanmışsa buna **tanımlanmış (göreceli) sıfır** adı verilir: Sıcaklık, başarı, zeka gibi özellikler tanımlanmış sifıra göre ölçülür. -20 derecede bir değerdir, bir varlığı ifade eder. Sıcaklık **sıfır** denirse bu sıcaklığın olmadığı anlamını taşımaz.

ÖLÇME TÜRLERİ

Doğrudan, dolaylı ve türetilmiş olmak üzere üç tür ölçme vardır.

Doğrudan Ölçme: Ölçülecek özelliğin doğrudan gözlenerek ölçme aracıyla birebir eşleştirilmesi şeklinde yapılan ölçmedir. Ölçülen özellik ölçme aracı üzerindeki gözlenen özellik aynıdır. Ağırlık ağırlık ile, uzunluk da uzunluk ile eşleştirilir.

Dolaylı Ölçme: Ölçülecek özelliğin dolaylı olarak gözlenmesi şeklinde yapılan ölçmedir. Ölçülen özellik ölçme aracı üzerindeki gözlenen özellik aynı değildir. Ölçülecek özellik doğrudan değil de kendisiyle ilgili olduğu kabul edilen başka bir değişken gözlenerek ölçülür. Ağırlık hem doğrudan hem de dolaylı olarak ölçülür. Sıcaklık dolaylı ölçmedir.

Türetilmiş Ölçme: Dolaylı ölçmenin farklı bir türüdür. Ölçmek istediğimiz özellik üzerinde herhangi bir işlem yapmadan bu özellik ilgili başka değişkenler üzerinde matematiksel işlemler yapılarak elde edilen ölçümlerdir. Örneğin yoğunluğun kütle ve hacim bağıntısından, hızın yol ve zaman bağıntısından ölçülmesi, zekânın IQ formülünden ölçülmesi.

1.3. ÖLÇME HATASI

Ölçülen özelliğın gerçek değeriyle ölçme aracından elde edilen değeri arasındaki farka ölçme hatası denir.

Hata artı veya eksi yönde olabilir. Ölçmede sabit (sistematik) ve tesadüfî (şansa bağılı) hata olmak üzere iki tür hata vardır.

Sistematik Hata: Ölçmeden ölçmeye belli oranda artan ya da azalan hatalardır. Sınavda dışarıda gürültü varsa bu durum öğrencilerin puanlarını etkiler ve sistematik olarak puanlarının düşüşüne sebep olabilir. Yani gürültüden etkilenme derecesine göre öğrencilerin puanlarına yansiyabilecek hata, bir öğrenciden başka bir öğrenciye değişiklik gösterebilir. Bu durumda ortaya çıkan hata sistematik hata olur.

Öğrencilerin yazılı sınavlarını yeterli önlem almadan gelişmiş güzel puanlayan bir öğretmenin takdir ettiği puanların hatalı olması yüksektir. Bu durumda puanlara yansıyan hata öğrencilere göre değişiklik göstereceğinden dolayı sistematik hata olacaktır.

Rastgele (Tesadüfi) Hata: Ölçme sonuçlarına belirsiz şekilde karışan hatalardır. Hatanın kaynağı belli olmadığı için bazen eksik bazen fazla ölçüm söz konusu olabilir.

Ölçme işleminin herhangi bir evresinde, ölçme sonuçlarına tesadüfi olarak karışan veya ölçme sonuçlarını tesadüfi olarak etkileyen herhangi bir faktörden kaynaklanan hataya tesadüfi hata denir.

Tesadüfi hata bütün ölçümler üzerinde sürekli bir etkisi yoktur. Bu nedenle tesadüfi hata ölçme sonuçlarının değişkenliğini (varyansını) artırırken, ortalama üzerinde fazla etki yapmaz.

Örneğin öğrencilerin ruhsal durumu başarısını etkileyebilir. Bu durum bazılarının puanını artırırken bazılarını düşürebilir, bazılarını da etkilemeyebilir. Ayrıca öğretmenin sınavları puanlarken toplama hatası yapması, şans başarısı, öğrencinin sınav günü hasta olması tesadüfi hatalar örnek verilebilir.

Not: Sistemik hatalar ölçme araçlarının güvenirliği üzerinde bir etkiye sahip değildir. **Güvenirlik yalnızca tesadüfi hatalardan etkilenir.**

- **Sabit Hatalar:** Her bir ölçme için miktarı değişmeyen hatalara sabit hatalar denir.
- Bir ölçmeden başka bir ölçmeye tekrar eden, ölçme sonuçlarına karışan ve miktar olarak değişmeyen hatalara sabit hatalar denir. Örneğin bir terazinin her ölçümünde 50 gr. eksik tartıyorsa, bu teraziyle yapılacak her ölçümün 50 gr eksik tartılmasına sebep olur. Terazile her ölçülen her ölçüm için hata miktarı tekrar eder, ancak yapılan hata miktarı sabit kalır.
- Sabit hatalar, bireysel ölçümleri ve o ölçümlerin ortalamasını gerçekte olduğundan büyük ya da küçük gösterebilir, fakat ölçümlerin dağılım ölçülerini değiştirici yönde bir etkiye sahip değildir.

- Her ölçüm mutlaka gerçek değeri ile birlikte bir hata terimini içerir. Amaç gerçek değeri elde etmektir, ancak hatadan soyutlanmış gerçek değeri elde etmek mümkün olmamaktadır. Önemli olan bu hatanın nasıl oluştuğunu, nasıl kontrol edilebileceğini bilmektir. Bir anket çalışmasında yanıtlayan kişi maddelerden birine yanıt verdiğinde (X), bu yanıtın gerçek değerini (T) etkileyen birçok faktör vardır.
 - Gerçek değer (T) tahmini (X), gerçek değerden farklı olacaktır. Buna ölçüm hatası (e) denir.
- T : Nesnenin gerçekten o özellikle ilgili sahip olduğu değer,
- X : Araştıracının ölçtüğü değer, (gerçek değer+hata)

$$\text{Gözlem Değeri} = \text{Gerçek Değer} + \text{Hata}$$

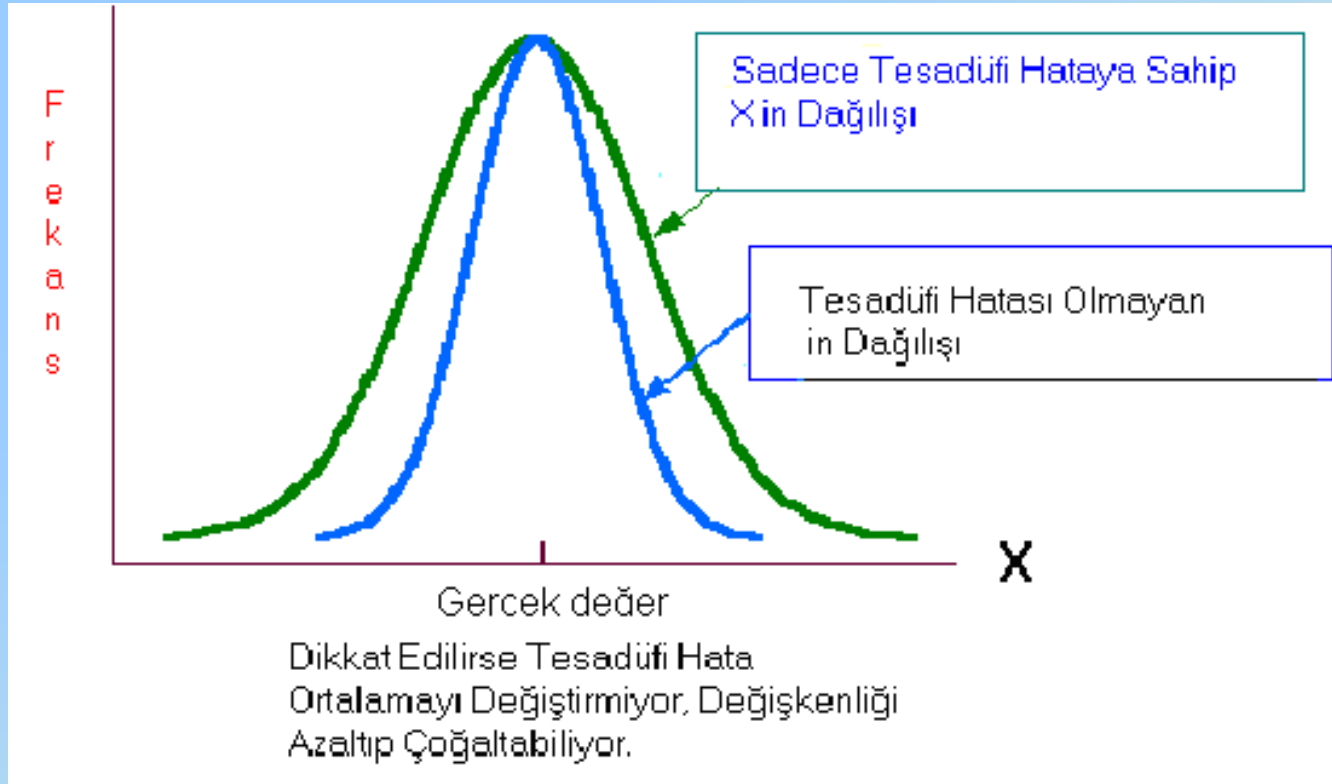
$$X = T + e ; \quad T: \text{Gerçek Değer}, \quad e: \text{Hata}$$

$$\left\{ \begin{array}{l} \text{(Tesadüfi Hata)} \\ e = e_r + e_s \\ \text{(Sistemik Hata)} \end{array} \right\}$$

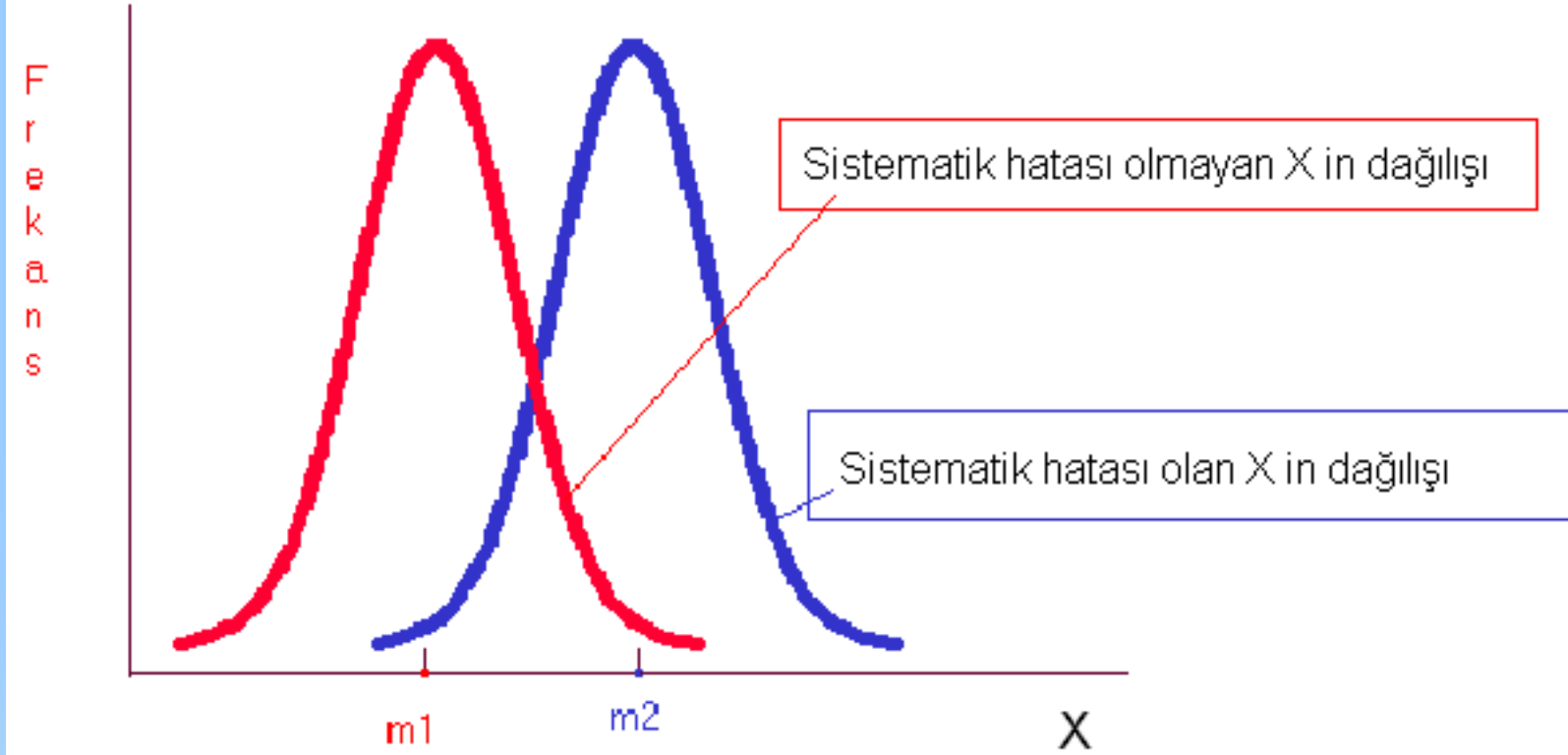
Yani:

$$\text{Gözlem Değeri} = \text{Gerçek Değer} + \left[\begin{array}{l} \text{Tesadüfi} \\ \text{Hata} \end{array} + \begin{array}{l} \text{Sistemik} \\ \text{Hata} \end{array} \right]$$

Tesadüfî Hatanın Ortalamaya Etkisi



Sistemik Hatanın Ortalamaya Etkisi



Dikkat edilirse sistemik hata ortalamanın büyüklüğünü deęiřtirmiyor, ortalamanın bu kaymasına sapma(bias) denir

- Eğer **sistemik hata varsa geçerlilik hiçbir zaman sağlanamayacaktır.** Gerçek ölçmek istenen değer ölçülemeyecektir.
- Tesadüfî hata küçükse güvenilirlik derecesi artacaktır, her seferinde gerçek değere daha yakın değerler ölçülecektir.
- **Güvenirlilik kavramı** birçok araştırmacı tarafından farklı isimlerle adlandırılmaktadır. **Objektiflik** (objectivity), **yeniden elde edilebilirlik** (reproducibility), **kararlılık** (stability), **uyum** (agreement), **ilişki** (association), **duyarlılık** (sensitivity), ve **hassasiyet** (precision) gibi.

Ölçmenin varyansı: $\sigma_x^2 = \sigma_t^2 + \sigma_e^2$

Gözlem değerlerinin varyansı: σ_x^2

Gerçek değerlerin varyansı: σ_t^2

Hatanın varyansı: σ_e^2

Bir ölçme aracının ne derece güvenilir olduğu gerçek değerlerin varyansını gözlenen değerler varyansına oranlayarak elde edilen güvenilirlik katsayısı ile anlaşılabilir.

$$\rho_x = \frac{\sigma_t^2}{\sigma_x^2}$$

Ölçmenin standart hatası: $\sigma_e = \sigma_x \sqrt{1 - \rho_x}$

Ölçmede Hata Kaynakları

Araştırmanın güvenilirliğini tehdit eden değişik hatalar mevcuttur.

Ölçme Aracı Hataları	Uygulama Hataları	Cevaplayıcı Hataları	Kodlama Hataları
<ul style="list-style-type: none">• Alan örnekleme hataları	<ul style="list-style-type: none">• Yetersiz bilgilendirme	<ul style="list-style-type: none">• Bazı maddelere cevap vermeme	<ul style="list-style-type: none">• Birden fazla şık işaretleme
<ul style="list-style-type: none">• Maddelerin belirsiz olması	<ul style="list-style-type: none">• Anketörler arası farklılık	<ul style="list-style-type: none">• Kendini gizleme	<ul style="list-style-type: none">• Eksik şık bırakma
<ul style="list-style-type: none">• Yanlış kelime kullanımı	<ul style="list-style-type: none">• Yetersiz uygulama koşulları	<ul style="list-style-type: none">• Dikkatsiz doldurma	<ul style="list-style-type: none">• Yanlı olarak kodlama
<ul style="list-style-type: none">• Dereceleme hataları	<ul style="list-style-type: none">• Işık, ısı, gürültü	<ul style="list-style-type: none">• İsteksiz doldurma	<ul style="list-style-type: none">• Bilgisayara yanlış kodlama
<ul style="list-style-type: none">• Etiketleme hataları	<ul style="list-style-type: none">• Kopya çekme	<ul style="list-style-type: none">• Yorgun olma, ilaç etkisi	<ul style="list-style-type: none">• Negatif maddeleri ters çevirmeme
<ul style="list-style-type: none">• Çift fikirli cümleler	<ul style="list-style-type: none">• Farklı uygulamalar	<ul style="list-style-type: none">• Duygusal davranma	<ul style="list-style-type: none">• Cevapsız bırakılan maddelere yanlışlıkla puan girme
<ul style="list-style-type: none">• Çift olumsuz cümleler	<ul style="list-style-type: none">• Cevaplayıcının uykusuz oluşu	<ul style="list-style-type: none">• Anlamama	

Anketlerde hatayı azaltmak için genelde bir madde sorulduktan sonra verilen yanıtı unutacak bir süre sonunda aynı maddenin başka bir soru biçiminde yeniden sorulması ile kontrolü şeklinde yapılabilir.

Bunların ortalaması ve deęişkenlięi bulunmaya çalışılır. Bir kişiden elde edilen ölçümlerin **standart hatası ne kadar küçükse** o kadar **güvenilir** bir sonuç elde etmiş oluruz.

Güvenirlikte Hata Kaynakları

Yöntem Hatası: Araştırmacının uyguladığı metodun çok dikkatli şekilde tasarlanmamış olmasından kaynaklanır. Örneğin, Anketin gürültülü ortamda yapılması, anketin çalışanların stresli zamanlarında yapılması gibi.

- **Özellik Hatası:** Anket uygulayan bireylerin kişisel özelliklerinden kaynaklanır. Kendilerine test uygulanan kişilerin yorgun olması gibi.
- **Sistemik Hata:** Testin ölçmeyi amaçladığı alanın dışında başka bir şeyi daha ölçüyor olması durumunda ortaya çıkar .
- **Tesadüfî Hata:** Bilinmeyen kontrol edilmeyen faktörlerden kaynaklanır

Güvenirlikte Hata Kaynakları

- **Denek hatası**, Sözgelimi, çalışanların işlerine bağlılık ve istek derecelerini ölçmeye yönelik bir anket uygulanıyor olsun. Eğer, anket haftanın değişik günlerinde doldurulmuşsa farklı cevapların verileceği açıktır. Pazartesi sabahı ile Cuma öğlen sonrası verilecek cevaplar değişik olur. Böyle bir çalışma için araştırmacı, mümkün olduğunca nötr bir zaman seçmelidir.
- **Denek önyargısı**, kendini güvende hissetmeyen denekler, sorulara genellikle gerçekte düşündükleri şekilde değil de üstünü memnun edecek şekilde cevap verme eğilimindedirler. Araştırmacı, bu riskin farkında olmalı ve bu tür bir önyargıyı en aza indirmek için deneklerin kimliklerini ortaya çıkaracak sorulardan kaçınmalıdır.

- **Gözlemci hatası**, güvenilirliği tehdit eden bir diğer unsurdur.
- Örneğin, araştırmacı veri toplama metodu olarak mülakatı seçmiş olsun. Böyle bir çalışmada ne kadar değişik sayıda mülakatçı varsa o kadar değişik cevap almak mümkündür. Bu tehdidi ortadan kaldırmak için, ya mülakatçı sayısını en aza indirmek veya mülakatları mümkün olduğunca biçimselleştirmek gerekir.
- **Gözlemci önyargısı**, araştırmacının elde ettiği cevapları deneğin kastettiğinden farklı şekilde yorumlamasından kaynaklanır

5'Lİ LİKERT TİPİ ÖLÇEKLERE ÖRNEKLER

- 1.KESİNLİKLE KATILMIYORUM**
- 2.KATILMIYORUM**
- 3.KARARSIZIM**
- 4. KATILYORUM**
- 5. KESİNLİKLE KATILYORUM**

- 1.HİÇ MEMNUN DEĞİLİM**
- 2. MEMNUN DEĞİLİM**
- 3. FİKRİM YOK**
- 4. MEMNUNUM**
- 5. ÇOK MEMNUNUM**

- 1. ÇOK KÖTÜ**
- 2. KÖTÜ**
- 3. ORTA**
- 4. İYİ**
- 5. ÇOK İYİ**

2. GÜVENİLİRLİK (Reliability)

Güvenilirlik, bir testin veya ölçeğin ölçmek istediği şeyi tutarlı ve istikrarlı bir biçimde ölçme derecesidir. Diğer bir deyişle, bireylerin test maddelerine verdikleri cevaplar arasındaki tutarlılıktır.

Güvenilir bir test veya ölçek, benzer şartlarda tekrar uygulandığında aynı benzer sonuçlar verir. Bir test veya ölçek ne derece güvenilir ise ondan elde edilen veriler de o derece güvenilirdir. Güvenilir olmayan bir ölçek yardımıyla elde edilen veriler faydasızdır. Mesela bir zeka testinde bir öğrenci bir gün 100, diğer gün 140 puan alıyorsa yapılan bu testin güvenilirliğinden bahsedilemez.

Güvenirlilik;

- **Aynı şeyin bağımsız ölçümleri arasındaki kararlılıktır;**
- **Ölçülmek istenen belli bir şeyin, sürekli olarak aynı sembolleri almasıdır;**
- **Aynı süreçlerin izlenmesi ve aynı ölçütlerin kullanılması ile aynı sonuçların alınmasıdır;**
- **Ölçmenin, tesadüfî yanılğılardan arınık olmasıdır;**
- **Bir ölçüm sürecinde ölçüm işleminin tekrarlanabilirliği, kararlılığı ya da tekrarlardaki tutarlılıktır.**

GÜVENİLİRLİK

- ✓ Endüstri mühendisine göre güvenilirlik: "üretimde hata oranının veya başarısızlık oranının düşük çıkmasıdır".
- ✓ Bir sosyologa göre güvenilirlik: "ölçüm sonuçlarının farklı anakütlelerde veya aynı ana kütleyle ait farklı örnek kütlelerde benzer sonuçlar vermesidir".
- ✓ Bir insan kaynakları yöneticisine göre güvenilirlik: "psikometrik test sonuçlarının uygulandığı farklı zaman dilimlerinde benzer sonuçlar vermesi ve iş yaşamında personelin göstereceği performansı doğru tahmin etmesidir".

Pilot araştırma verilerinde yapılan güvenilirlik analizleri bir ön yordama niteliğindedir. Güvenilirlik analizleri ise esas araştırma sonuçlarına göre yapılmalıdır.

Güvenilirliği sorgulama :

- **Bir ölçüm aracındaki maddelerin aynı kavramsal yapıyı hatasız bir biçimde ölçmesi**
- **Farklı zamanlarda yapılan ölçüm sonuçlarının benzer çıkması**
- **Bir ölçüm aracına ait sonuçların aynı kavramsal yapıyı ölçen diğer ölçüm araçlarının sonuçlarıyla tutarlı olması**
- **Farklı gözlemciler tarafından yapılan ölçüm sonuçlarının benzer çıkması**

Örnek

Güvenirlilik: (üç senaryo düşünelim)

1. Ağaç masamın boyutları bir zamandan diğerine değişmeyeceğine göre veya bir yerden diğer yere göre değişmeyeceği için ve benim çelik ölçme aletimin boyu değişmeyeceği için ölçüm nerede veya ne zaman yapılırsa yapılsın fark etmez aynı bulunması gerekir. Eğer masa bir yerde 120 cm ölçülmüş ise başka bir yerde de veya başka bir günkü ölçümde de 120 cm gelecektir.

Birbiri peşine 3 ölçüm yapsak her 3 ölçümde de 120 cm bulunacaktır.

2. Diğer bir senaryoya göre ölçüm aletinin esneyebilen bezden yapıldığını ve çekme ile biraz uzayabildiğini varsayalım bu durumda 3 farklı ölçümde değişik ölçümler bulunacaktır, 120 cm, 121 cm ve 120,7 cm gibi.

3. Bir başka senaryoya göre ölçüm aleti sıcaklıktan daha fazla etkilenen daha esnek bir materyalden (plastik) yapılmış olsun. Bu aletle aynı masa oda ısısının farklı olduğu zamanlarda alındığında alet sıcaktan çok etkileneceği için çok daha farklı sonuçlar elde edilecektir. 120 cm, 150 cm ve 80 cm bulunmuş olsun.

Bu kavram ölçümün güvenilirliği ile ilgilidir. **Aynı nesne, aynı işlemle ölçüldüğünde neticenin aynı olması gerekir.** 1. durumda olduğu gibi alınan bir ölçüm **güvenilir** bir ölçümdür. 2. durumda çok az değişiklik vardır, dolayısıyla bu ölçümde **nispeten güvenilirdir**, ama 3. durumda çok büyük farklılıklar olduğu için ölçümler **güvenilirlikten uzaktır**.

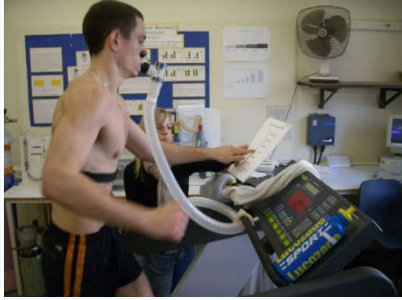



Nesne: Masa



Ölçüm aracı: Çelik metre
Mezura (Kot)
Mezura (plastik)

ÖRNEK 2.1. 1. ve 2. arařtıřıcılar, altın standartla (gerçek deęere en yakın ölçüm) benzer ölçümler yapmıřtır. **Deęerlendiriciler arası güvenirlilik** tamdır.

	Altın standart Ölçüm	1. Arařtırıcı	2. Arařtırıcı
			
1. Kiři	60 ml.kg ⁻¹ .min ⁻¹	60 ml.kg ⁻¹ .min ⁻¹	60 ml.kg ⁻¹ .min ⁻¹
2. Kiři	55 ml.kg ⁻¹ .min ⁻¹	55 ml.kg ⁻¹ .min ⁻¹	55 ml.kg ⁻¹ .min ⁻¹
3. Kiři	70 ml.kg ⁻¹ .min ⁻¹	70 ml.kg ⁻¹ .min ⁻¹	70 ml.kg ⁻¹ .min ⁻¹

ÖRNEK 2.2. 1. ve 2. arařtıřıcılar, altın standartla (gerçek deęere en yakın ölçüm) benzer ölçümler yapmıřtır. **Deęerlendiriciler arası güvenirlilik** tamdır. Ancak ölçümler hatalıdır.

	Altın standart Ölçüm	1. Arařtırıcı	2. Arařtırıcı
			
1. Kiři	60 ml.kg ⁻¹ .min ⁻¹	65 ml.kg ⁻¹ .min ⁻¹	65 ml.kg ⁻¹ .min ⁻¹
2. Kiři	55 ml.kg ⁻¹ .min ⁻¹	60 ml.kg ⁻¹ .min ⁻¹	60 ml.kg ⁻¹ .min ⁻¹
3. Kiři	70 ml.kg ⁻¹ .min ⁻¹	75 ml.kg ⁻¹ .min ⁻¹	75 ml.kg ⁻¹ .min ⁻¹

ÖRNEK 2.3. 1. ve 2. arařtıřıcılar, altın standartla (gerçek deęere en yakın ölçüm) benzer ölçümler yapmıřtır. **Deęerlendiriciler arası güvenirlilik** yoktur, hem de ölçümler hatalıdır.

	Altın standart Ölçüm	1. Arařtırıcı	2. Arařtırıcı
			
1. Kiři	60 ml.kg ⁻¹ .min ⁻¹	72 ml.kg ⁻¹ .min ⁻¹	57 ml.kg ⁻¹ .min ⁻¹
2. Kiři	55 ml.kg ⁻¹ .min ⁻¹	61 ml.kg ⁻¹ .min ⁻¹	52 ml.kg ⁻¹ .min ⁻¹
3. Kiři	70 ml.kg ⁻¹ .min ⁻¹	40 ml.kg ⁻¹ .min ⁻¹	84 ml.kg ⁻¹ .min ⁻¹

2.1. GÜVENİLİRLİK ANALİZİ YÖNTEMLERİ

2.1.1. Test-Yeniden Test (Test-Retest)

2.1.2. Paralel Formlar Yöntemi

2.1.3. Gözlemciler Arası Güvenilirlik

2.1.4. İç Tutarlılık Güvenilirliği

GÜVENİLİRLİK ANALİZİ YÖNTEMLERİ

2.1.1. Test-Yeniden Test (Test-Retest Reliability Analysis)

Buna göre anket formu iki farklı zamanda (2-4 hafta veya 10-20 gün) aynı denekler üzerinde aynı şartlar altında uygulanır. İki farklı zamandaki gözlemler arasındaki korelasyon, güvenilirliğin bir göstergesi olacaktır. Bu ölçeğin bazı sorunları vardır. Önceki sonuçların sonraki sonuçları etkiyebilmesi, tekrar ölçme zorluğu gibi sıkıntılar olabilir.

Bir testin belli bir süre arayla aynı kişilere tekrar uygulanır. Toplam skor puanları ve her bir sorunun tutarlılığı Pearson-Spearman korelasyon ile değerlendirilir. Uyumsuz cevap alınan sorular anketten çıkarılır. Bunun için korelasyon değeri +0,3'ün altında olan sorular uyumsuz soru olarak kabul. Her bir alt ölçek için test re-test için korelasyonlara bakılır. Bulunan korelasyon pozitif ve 1'e yakınsa testin güvenilir olduğuna karar verilir.

2.1.2. Paralel Formlar Yöntemi

Bu yöntemle güvenilirlik tahmini için, bir testin en az iki eşdeğer formunun geliştirilmiş olması gerekir. Paralel iki testin eşdeğer olması için, her iki test, içindeki madde sayısı, niteliği ve ölçtükleri davranış bakımından birbirine denk olmalıdır

Bu yöntem aynı gruba uygulanır ve iki ayrı formdan almış oldukları puanlar arasındaki korelasyona bakılır. Her iki test peş peşe uygulanabileceği gibi belli bir ara ile de uygulanabilir.

Bu yöntemle elde edilen güvenilirlik katsayısı, diğer yöntemlerle elde edilenlerden daha küçük çıkar.

2.1.3. Gözlemciler Arası Güvenilirlik

Gözlemciler önceden belirlenmiş bir puanlama sistemine bağlı olarak belirli bir olguyu, bağımsız bir şekilde değerlendirebilirler. Bu değerlendirmede verilen puanların birbirine benzer olması söz konusu puanların güvenilir olduğunu gösterir.

Birden çok gözlemcinin, birbirinden bağımsız olarak, aynı şeyleri ölçmeye çalıştıkları durumlarda uygulanan bir güvenilirlik ölçütüdür. Bağımsız gözlemciler arası uyumu hesaplamak için Korelasyon - Kendall's Coefficient of Concordance kullanılır.

Ölçüm aracı kullanılarak yapılan değerlendirmede gözlemciler arası uyuşmanın en az 0,70 olması gerekir.

Gözlemcilerin verdikleri puanlar arasındaki uyumu belirlemek için verilerin niteliği göz önünde bulundurulur.

Gözlemci Puanlarında Güvenilirlik Analizleri

Ölçme Düzeyi	İki Gözlemci	İkiden Çok Gözlemci
(Kesikli Değişkenler)		
Nominal	- Uyum İndeksi - Cohen Kappa - Phi Katsayısı	- Uyum İndeksi - Cohen Kappa
Ordinal	-Spearman Sıra Kor. -Kendall Tau (a,b,c)	-Kendall W Uyum Katsayısı
Sürekli Değişken		
Eşit aralıklı /Oran	Pearson Korelasyon An. (Normal Dağılımlı)	Küme İçi Korelasyon Analizi (Intraclass Correlation Coefficient) (Normal Dağılımlı)

Gözlemciler Arası Güvenilirlik

a) Uyum İndeksi

Uyuşma indeksi (Uİ) gözlemcilerin veya değerlendircilerin uyuştukları madde sayısının toplam değerlendirme sayısına oranıdır.

$$\mathbf{Uİ=(Toplam\ uyuşma\ sayısı)/(Toplam\ değerlendirme\ sayısı)*100}$$

Gözlemciler arası değerlendirme sonuçlarının güvenilir sayılabilmesi için Uİ değerinin %75'in üzerinde olması gerekir.

Gözlemciler Arası Güvenilirlik

b) Phi Katsayısı

Phi katsayısı ikili veri (0,1 veya başarılı-başarısız gibi) yapısına sahip iki nitel değişken arasındaki ilişkileri belirlemek için kullanılır. Örneğin iki hakemin kişileri başarılı-başarısız şeklindeki değerlendirmesi arasındaki güvenilirlik Phi katsayısı ile hesaplanır.

Phi katsayısı sınıflayıcı ölçekle elde edilmiş 2x2 lik tablolarda verilerin korelasyon katsayısını bulmak için kullanılan parametrik olmayan bir testtir.

Phi katsayısı varsayımları:

- a) Veriler sınıflayıcı ölçekle ölçülmüş olmalı
- b) 2x2 lik tablolarda kullanılabilir.
- c) Örnekler anakütleden rasgele seçilmiş olmalıdır (Levin ve Fox, 1991).

	X	X^+	Σ
Y	a	b	e
Y^+	c	d	f
Σ	g	h	N

Phi Katsayısı,

$$\phi = \frac{ad - bc}{\sqrt{efgh}} \text{ veya } \phi = \sqrt{\frac{\chi^2}{N}}$$

$$\chi^2 = \frac{N \left(|ad - bc| - \frac{N}{2} \right)^2}{(a + b)(c + d)(a + c)(b + d)} \sim \chi^2_{1,\alpha}$$

Örnek 2.4. İki hakemin 39 adayı değerlendirme neticesinde vermiş oldukları kararlar aşağıdaki gibidir. İki hakemin vermiş olduğu kararlar arasındaki ilişki miktarını bulunuz?

	Başarılı	Başarısız	Toplam
Hakem1	13	7	20
Hakem2	5	14	19
Toplam	18	21	39

$$\phi = \frac{ad - bc}{efgh} = \frac{13 * 14 - 5 * 7}{\sqrt{18 * 21 * 20 * 19}} = 0,388$$

$$\phi = \sqrt{\frac{\chi^2}{N}} = \sqrt{\frac{5,867}{39}} = 0,388$$

$$\chi^2 = \frac{N[|ad - bc| - N/2]^2}{(a + b)(c + d)(a + c)(b + d)} = \frac{39[|13 * 14 - 7 * 5| - 39/2]^2}{20 * 19 * 18 * 21} = 5,867$$

Hakem	Başarı	frekans
1	1	13
1	2	7
2	1	5
2	2	14

*Untitled1 [DataSet0] - SPSS Statistics Data Editor

Crosstabs

Row(s): Hakem

Column(s): Başarı

frekans

Display clustered bar charts
 Suppress tables

Crosstabs: Statistics

Chi-square
 Correlations

Nominal

Contingency coefficient
 Phi and Cramer's V
 Lambda
 Uncertainty coefficient

Ordinal

Gamma
 Somers' d
 Kendall's tau-b
 Kendall's tau-c

Nominal by Interval

Eta
 Kappa
 Risk
 McNemar

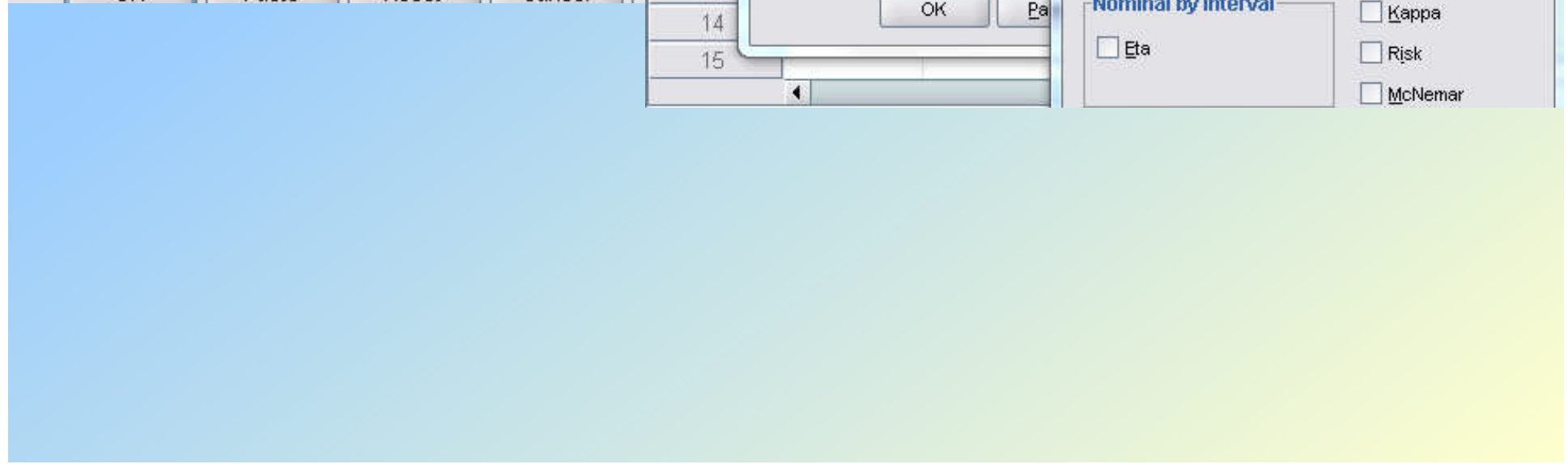
Weight Cases

Do not weight cases
 Weight cases by

Frequency Variable: frekans

Current Status: Weight cases

OK Paste Reset Cancel



Symmetric Measures

		Value	Asymp. Std. Error ^a	Approx. T ^b	Approx. Sig.
Nominal by Nominal	Phi	,388			,015
	Cramer's V	,388			,015
Interval by Interval	Pearson's R	,388	,147	2,560	,015 ^c
Ordinal by Ordinal	Spearman Correlation	,388	,147	2,560	,015 ^c
N of Valid Cases		39			

a. Not assuming the null hypothesis.

b. Using the asymptotic standard error assuming the null hypothesis.

c. Based on normal approximation.

Gözlemciler arası güvenilirlik

c) Cohen Kappa

İki gözlemcinin yaptığı değerlendirmeler arasındaki uyumu belirlemek için kullanılır. Cohen(1960) tarafından geliştirilmiştir.

$$K = \frac{GUO - BUO}{1 - BUO}$$

GUO (Gözlenen uyum oranı): Her iki gözlemcinin benzer puanlar vermesi veya değerlendirmeler yapması

BUO (Beklenen uyum oranı) : Satır toplamları ile sütun toplamlarının çarpımlarının toplam değerlendirme sayısına bölünmesiyle elde edilen toplamların tekrar toplam değerlendirme sayısına oranıdır.

Değerlendirme verilerinde gözlemcilerin uyuştukları puanlar karşılaştırma matrisinin köşegenindeki hücrelere yazılır. Uyuşmadıkları değerler ise köşegen dışındaki harflerin kesiştiği hücrelere yazılır.

Örnek 2.5.

		I. Kişi				
		A	B	C	D	TOP
II. K i ş i	A	9	2	8	15	34
	B	2	18	8	6	34
	C	3	4	8	1	16
	D	2	2	1	11	16
	TOP	16	26	25	33	100

$$GUO=(9+18+8+11)/100=0,46$$

$$BUO=[(16*34)/100+(26*34)/100+(25*16)/100+(33*16)/100]/100=0,23$$

$$Kappa=(0,46-0,23)/(1-0,23)=0,29$$

	hakem2	hakem1	veri	
1	1	1	9	
2	1	2	2	
3	1	3	8	
4	1	4	15	
5	2	1	2	
6	2	2	18	
7	2	3	8	
8	2	4	6	
9	3	1	3	
10	3	2	4	
11	3	3	8	
12	3	4	1	
13	4	1	2	
14	4	2	2	
15	4	3	1	
16	4	4	11	
17				

Weight Cases

Do not weight cases
 Weight cases by
 Frequency Variable:

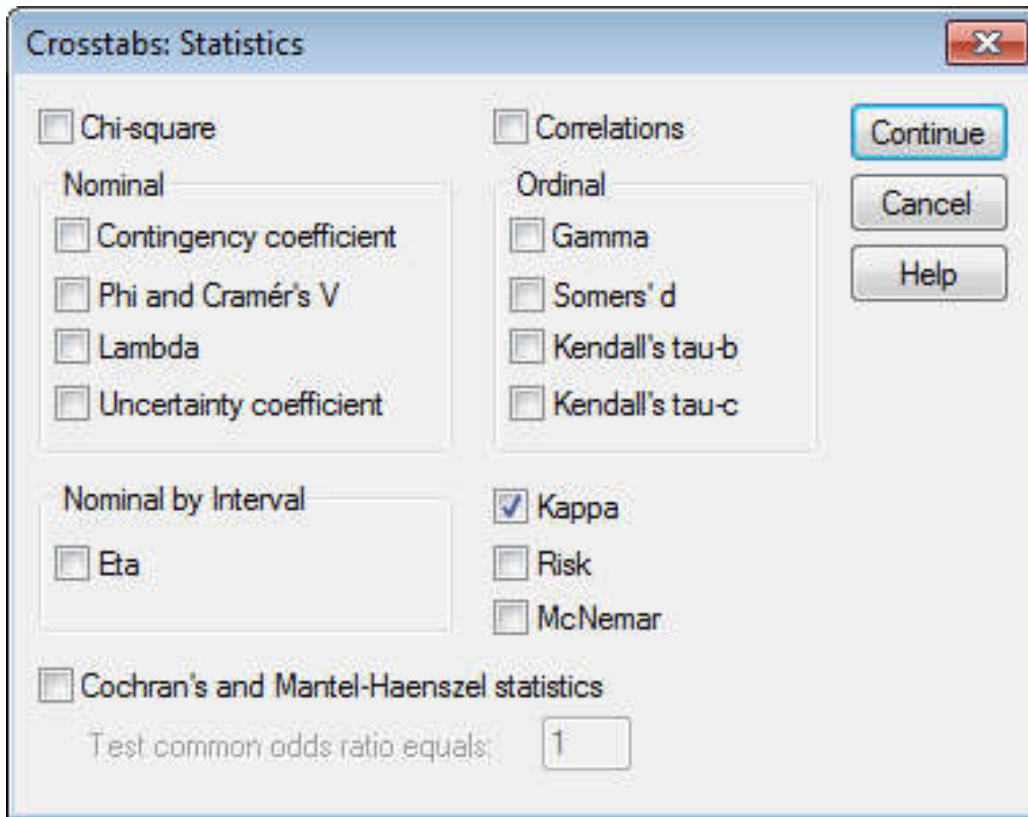
Current Status: Do not weight cases

Crosstabs

Row(s):
 Column(s):

Layer 1 of 1

Display clustered bar charts
 Suppress tables



Symmetric Measures

	Value	Asymp. Std. Error ^a	Approx. T ^b	Approx. Sig.
Measure of Agreement Kappa	,294	,062	5,465	,000
N of Valid Cases	100			

a. Not assuming the null hypothesis.

b. Using the asymptotic standard error assuming the null hypothesis.

Örnek.

I.HAKEM					
II. H A K E M		UYGUN	KALABİLİR	UYGUN DEĞİL	TOPLAM
	UYGUN	11	2	4	17
	KALABİLİR	1	14	2	17
	UYGUN DEĞİL	4	2	10	16
	TOPLAM	16	18	16	50

$$GUO=(11+14+10)/50=0,70$$

$$BUO=[(16*17)/50+(18*17)/50+(16*16)/50]=0,33$$

$$Kappa=(0,70-0,33)/(1-0,33)=0,55$$

Gözlemciler arası güvenilirlik

d) Kendall Uyum Katsayısı W

Çok sayıda hakemin vermiş olduğu puanlar arasında ne ölçüde uyum olduğunu belirler. SPSS'te Non parametric Tests-k Related Samples-Kendall's W kısmından yapılır. Bu testte hakemler veya gözlemciler satırda yer alır. Sütunlarda ise değerlendirilen kişiler, nesnelere yer alır. Kendall W 0-1 arasında değerler alır.

$$W = \frac{12 \sum (R_j)^2}{k^2 N(N^2 - 1)} - \frac{3(N + 1)}{N - 1}$$

N: Sıralanması gereken birimlerin sayısı (değerlendirilenlerin sayısı-sütun sayısı)

K: Sıralamayı yapan birim sayısı (Hakem sayısı-satır sayısı)

R_j: Sıralanan birimlerin aldıkları sıra numaraları puanları toplamı (sütun toplamı)

Örnek 2.6. Bir fakülte'deki staj komisyonununun 6 üyesi yurt dışında staj yapmak isteyen öğrenciler arasından seçilen 5 öğrenciye sıra puanları vermiştir (En iyi öğrenciye 1 sıra puanı veriliyor). Komisyon üyelerinin sıralamaları arasında uygunluk var mıdır?

Üye\Öğr.	A	B	C	D	E	
1	5	1	2	3	4	
2	5	1	2	3	4	
3	5	2	1	4	3	
4	5	2	1	4	3	
5	4	1	2	5	3	
6	4	1	2	5	3	
R_j	28	8	10	24	20	
$(R_j)^2$	784	64	100	576	400	1924

$$W = \frac{12 \sum (R_j)^2}{k^2 N(N^2 - 1)} - \frac{3(N+1)}{N-1} = \frac{12 * 1924}{6^2 * 5(5^2 - 1)} - \frac{3(5+1)}{5-1} = 0,85$$

	A	B	C	D	E
1	5	1	2	3	4
2	5	1	2	3	4
3	5	2	1	4	3
4	5	2	1	4	3
5	4	1	2	5	3
6	4	1	2	5	3
7					

Tests for Several Related Samples

Test Variables:

- A
- B
- C
- D
- E

Test Type

Friedman Kendall's W Cochran's Q

Buttons: OK, Paste, Reset, Cancel, Help, Exact..., Statistics...

Data Reduction

Scale

Nonparametric Tests

- Chi-Square...
- Binomial...
- Runs...
- 1-Sample K-S...
- 2 Independent Samples...
- K Independent Samples...
- 2 Related Samples...
- K Related Samples...**

Missing Value Analysis...

Complex Samples

Quality Control

ROC Curve...

Test Statistics

N	6
Kendall's W ^a	,844
Chi-Square	20,267
df	4
Asymp. Sig.	,000

a. Kendall's Coefficient of Concordance

Örnek 2.7. 10 adaya 3 hakemin vermiş oldukları notların sıra puanları aşağıdaki gibidir. Hakemlerin sıralamaları arasında uygunluk var mıdır?

Criterion	Participant									
	1	2	3	4	5	6	7	8	9	10
X	1	4.5	2	4.5	3	7.5	6	9	7.5	10
Y	2.5	1	2.5	4.5	4.5	8	9	6.5	10	6.5
Z	2	1	4.5	4.5	4.5	4.5	8	8	8	10
Sum	5.5	6.5	9	13.5	12	20	23	23.5	25.5	26.5

Test Statistics

N	3
Kendall's W ^a	,828
Chi-Square	22,349
df	9
Asymp. Sig.	,008

a. Kendall's Coefficient of Concordance

Gözlemciler arası güvenilirlik

e) Kendall Tau a,b,c

Kendall (1938) tarafından geliştirilen bu test, ikili ve sıralı ölçekli veriler arasındaki ilişkileri belirler. İki farklı gözlemcinin yaptığı değerlendirmeler puan büyüklüğü sırası içinde verilmişse veya hakem puanları büyüklük sırasına sokmuşsa Spearman korelasyonunun yanında, Kendall tau b analizi de yapılır. Spearman rho büyüklük sırasına sokulmuş değerler arasındaki Pearson momentler çarpımını temsil ederken, Kendall tau olasılığı temsil eder. Yani iki değişkende aynı sırada yer alan(uyuşan) verilerin gözlenme olasılığı ile farklı sırada yer alan (uyuşmayan) verilerin gözlenme olasılığı arasındaki farktır. İki serideki veriler uyuşuyorsa tau-b +1 değerini alır. Uyuşma yoksa tau-b -1 olur.

Kendall tau-b kare tablolar için, Kendall-tau-c ise dikdörtgen tablolar için kullanılır. Tau-a ise uyuşan ve uyuşmayan çiftler arasındaki farkın toplam çift sayısına bölünmesiyle bulunur.

Kendall tau katsayıları aşağıdaki gibi yorumlanır:

>0,5	Yüksek ilişki
0,36-0,49	Önemli ilişki
0,20-0,35	Orta derecede ilişki
0,10-0,19	Düşük ilişki
<0,1	İlişki yok

Kendall Tau-b

Kendall tau_b baęlı sıralıęa sahip (aynı gözlemleri olan deęerlerin sıra puanlarının ortalaması alınır) ölçek verileri için uygundur. Bu analiz verilerin normal daęılıę göstermedięi ve $n < 20$ olduęu durumlarda daha iyi sonuç vermektedir.

İki serideki veriler uyuşuyorsa tau-b +1 deęerini alır. Uyuşma yoksa tau b -1 olur.

İki sıralı deęişken arasındaki ilişkinin yönünü ve gücünü gösteren parametrik olmayan testlerdir. Kendall $-1 \leq \tau \leq +1$ aralığında deęişim gösterir. Formül için gerekli hesaplamalar Somers'in hesaplanmasında kullanılan yöntem ile aynıdır. Bu testte örnek sayısı 10 geçtiğinde daęılım normal daęılıma yaklaşır (Ergün, 1995).

Kendall Tau-b testi aynı zamanda iki alternatif için Somers'in geometrik ortalaması olarak ifade edilebilir. Bu yüzden, zaman zaman Somers'in simetrik versiyonu olarak ta düşünülebilir. Tau b testi 2x2'lik ve kare şeklindeki tablolar için kullanılması tavsiye edilir. Verilerin sıralayıcı ya da sınıflayıcı özellikte olması gerekir.

	A ₁	A ₂	Σ
B ₁	n ₁₁	n ₁₂	R ₁
B ₂	n ₂₁	n ₂₂	R ₂
Σ	C ₁	C ₂	N

Test hipotezi,

H₀: İlişki yoktur.

H₁: İlişki vardır.

$$\tau_b = \frac{n_{11}n_{22} - n_{12}n_{21}}{\sqrt{R_1 R_2 C_1 C_2}}$$

Tablo 2x2 den büyükse:

m: Satır ya da sütun sayısını (hangisi en küçük ise),
N: Toplam gözlem sayısını,
P: Uyumlu çiftleri,
Q : Uyumsuz çiftleri,
Y₀: Y'deki eşleştirmeleri,
X₀: X'deki eşleştirmeleri ifade etmektedir.

$$S = P - Q$$

$$\tau_b = \frac{S}{\sqrt{(P + Q + Y_0)(P + Q + X_0)}}$$

şeklinde ifade edilir.

Kendal tau b testinin önemliliği,

$$z = \frac{\tau_b}{\sqrt{\frac{2(2N+5)}{9N(N-1)}}} \sim z_\alpha$$

Örnek 2.8. Bir önceki örneğe göre eğitim düzeyi ile davranış değişikliği arasındaki ilişki miktarını bulunuz.

Eğitim Düzeyi	Davranış Değişikliği			Toplam
	Yetersiz	Orta	Çok verimli	
İlköğretim	5	14	21	40
Lise	12	18	13	43
Yüksek	17	7	3	27
Toplam	34	39	37	110

H_0 : İlişki yoktur.

H_1 : İlişki vardır.

Tau değeri,

$$S = 603 - 2155 = -1552$$

$$\tau_b = \frac{-1552}{\sqrt{(603 + 2155 + 1266)(603 + 2155 + 1200)}} = -0.389$$

olarak hesaplanır. Buradan, eğitim düzeyi ile davranış değişikliği arasında negatif orta derecede bir ilişki vardır.

Eğitim	Davranış	frekans
1	1	5
1	2	14
1	3	21
2	1	12
2	2	18
2	3	13
3	1	17
3	2	7
3	3	3

The image shows two dialog boxes from SPSS. The top dialog is 'Crosstabs', where 'frekans' is listed in the dependent variable box, 'Eğitim' is in the Row(s) box, and 'Davranış' is in the Column(s) box. The bottom dialog is 'Crosstabs: Statistics', where 'Kendall's tau-b' is checked under the 'Ordinal' section. Other options like 'Chi-square', 'Correlations', 'Contingency coefficient', 'Phi and Cramer's V', 'Lambda', 'Uncertainty coefficient', 'Gamma', 'Somers' d', and 'Kendall's tau-c' are unchecked.

Symmetric Measures

	Value	Asymp. Std. Error ^a	Approx. T ^b	Approx. Sig.
Ordinal by Ordinal Kendall's tau-b	-,389	,075	-5,144	,000
N of Valid Cases	110			

a. Not assuming the null hypothesis.

b. Using the asymptotic standard error assuming the null hypothesis.

Örnek 2.9.

	A	B
1	2	2
2	3	1
3	2	3
4	1	2
5	3	3
6	1	1
7	2	1
8	3	3
9	1	1
10	2	2
11	1	1
12	3	3
13	1	2
14	2	3
15		

Crosstabs

Row(s): A

Column(s): B

Layer 1 of 1

Previous Next

Display clustered bar charts

Suppress tables

Exact... Statistics... Cells... Format...

OK Paste Reset Cancel Help

Crosstabs: Statistics

Chi-square

Nominal

Contingency coefficient

Phi and Cramér's V

Lambda

Uncertainty coefficient

Correlations

Ordinal

Gamma

Somers' d

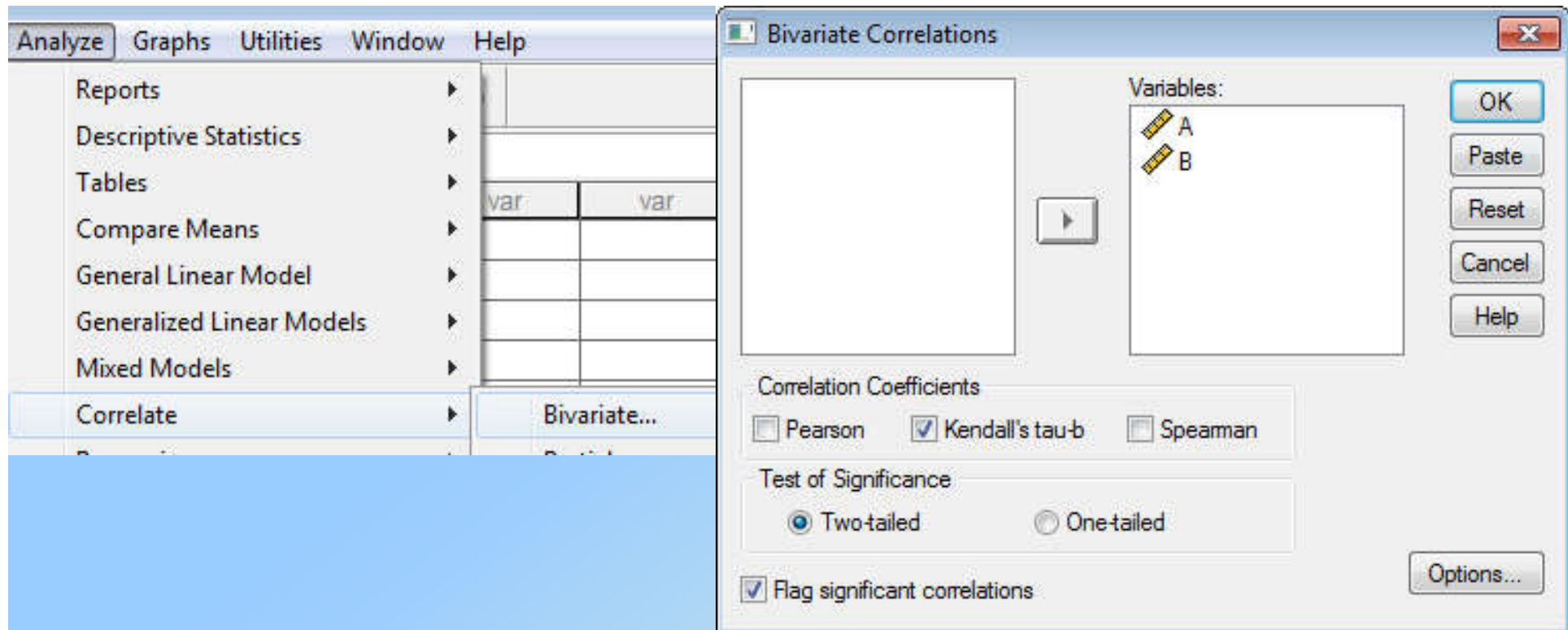
Kendall's tau-b

Kendall's tau-c

Continue Cancel Help

Symmetric Measure

		Value
Ordinal by Ordinal	Kendall's tau-b	,492
N of Valid Cases		14



Correlations

			A	B
Kendall's tau_b	A	Correlation Coefficient	1,000	,492*
		Sig. (2-tailed)	.	,046
		N	14	14

Kendall Tau-c

Kendall'ın tau-c testi kare ya da dikdörtgen şeklinde hazırlanmış daha büyük tablolar için kullanılır. Özellikle, tau-b testinin uygulanamadığı durumlarda da uygulanabilen testlerdir. Kendall'ın tau-c testine farklı kaynaklarda Stuart'ın tau-c ya da Kendall-Stuart tau-c adı da verilmektedir. Tau-c testinde verilerin en az iki tanesinden birinin sıralayıcı özellikte olması istenir (Ergün, 1995).

H_0 : İlişki yoktur.

H_1 : İlişki vardır.

2x2'lik tablolar için,

$$\tau_c = \frac{4(n_{11}n_{22} - n_{12}n_{21})}{N^2}$$

m: Satır ya da sütun sayısını (hangisi en küçük ise),

N: Toplam gözlem sayısını,

S: Uyumlu ve uyumsuz çiftlerin farkını ifade etmektedir.

2x2 den büyük tablolar için:

$$\tau_c = \frac{2m(S)}{N^2(m-1)}$$

şeklinde ifade edilir.

Kendal'ın tau c testinin önemliliği,

$$z = \frac{\tau_c}{\sqrt{\frac{2(2N+5)}{9N(N-1)}}} \sim z_\alpha$$

Örnek 2.10. Bir önceki örneğe göre eğitim düzeyi ile davranış değişikliği arasındaki ilişki miktarını bulunuz.

Eğitim Düzeyi	Davranış Değişikliği			Toplam
	Yetersiz	Orta	Çok verimli	
İlköğretim	5	14	21	40
Lise	12	18	13	43
Yüksek	17	7	3	27
Toplam	34	39	37	110

H_0 : İlişki yoktur.

H_1 : İlişki vardır.

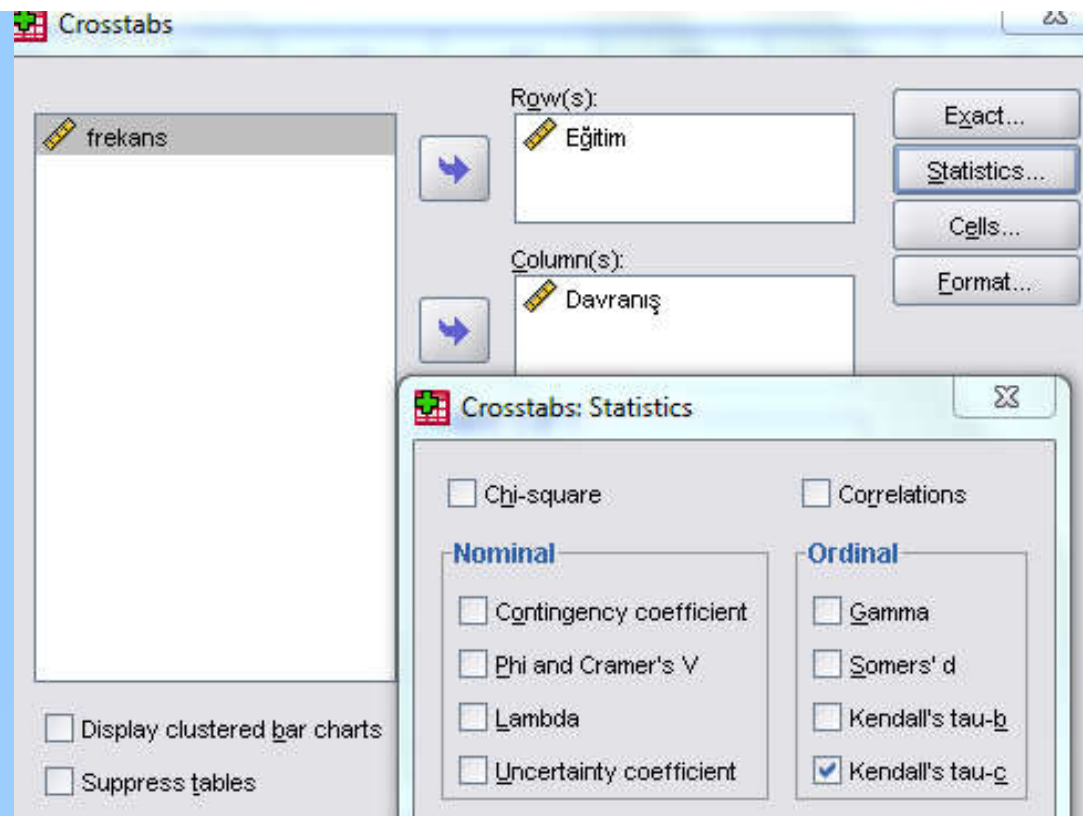
şeklinde kurulabilir ve Kendall'in Tau c katsayısı,

$$\tau_c = \frac{2(3)(-1552)}{110^2(3-1)} = -0.386$$

olarak hesaplanır. Buradan, Eğitim ile davranış değişikliği arasında negatif orta derecede bir ilişki söz konusudur.

Kendall'in Tau c'nin önem testi,

$$z = \frac{0.386}{\sqrt{\frac{2(2(110)+5)}{9(110)(110-1)}}} = |-5.998|$$



Symmetric Measures

	Value	Asymp. Std. Error ^a	Approx. T ^b	Approx. Sig.
Ordinal by Ordinal Kendall's tau-c	-,386	,075	-5,144	,000
N of Valid Cases	110			

a. Not assuming the null hypothesis.

b. Using the asymptotic standard error assuming the null hypothesis.

Gözlemciler arası güvenilirlik

f) Spearman Sıra Korelasyonu

Gözlemcilerin verdikleri puanların kendilerine değil de büyüklük sırasına sokulmuş (rank) değerleri temel alınmışsa, sürekli veriler normal dağılım özelliği göstermiyorsa veya Likert tipi ölçek verileri kullanılıp bu veriler sıralı ölçek niteliğinde değerlendirilmişse Spearman Sıra Korelasyonu kullanılır.

Sıra korelasyon katsayısının hesaplanmasında önce gözlem değerleri büyükten küçüğe doğru sıralanır ve bun sıralamaya göre sıra numarası verilir. Aynı olan gözlemlerin sıra numaraları verilirken sıra no ortalaması alınır.

$$r_s = 1 - \frac{6\sum D_i^2}{n(n^2 - 1)}$$

D_i : X ve Y'nin sıra numaraları arasındaki fark

n : Gözlem sayısı

Örnek 2.11. 7 öğrencinin boy ve ağırlıklarının büyükten küçüğe doğru sıra puanları aşağıda gösterilmiştir. Veriler normal dağılım göstermediğine göre, boy ve ağırlıklar arasındaki ilişkiyi hesaplayınız?

X	Y	D	D ²
2	1	1	1
4	6	-2	4
6	5	1	1
1	2	-1	1
3	3	0	0
7	7	0	0
5	4	1	1

$$r_s = 1 - \frac{6 \sum D_i^2}{n(n^2 - 1)}$$
$$= 1 - \frac{6 * 8}{7(7^2 - 1)} = 0,86$$

Öğrencilerin boy uzunlukları ile ağırlıkları arasında aynı yönde önemli bir ilişki vardır.

	X	Y
1	2,00	1,00
2	4,00	6,00
3	6,00	5,00
4	1,00	2,00
5	3,00	3,00
6	7,00	7,00
7	5,00	4,00

Bivariate Correlations

Variables:

X
Y

Correlation Coefficients

Pearson Kendall's tau-b Spearman

Test of Significance

Two-tailed One-tailed

Flag significant correlations

Options...

Correlations

			X	Y
Spearman's rho	X	Correlation Coefficient	1,000	,857*
		Sig. (2-tailed)	.	,014
		N	7	7
	Y	Correlation Coefficient	,857*	1,000
		Sig. (2-tailed)	,014	.
		N	7	7

*. Correlation is significant at the 0.05 level (2-tailed).

Gözlemciler Arası Güvenilirlik

g) Sınıfıçi Korelasyon Analizi (Intraclass Correlation Coefficient-ICC)

Sınıfıçi korelasyon analizi genelde gözlemci değerlendirmelerinin güvenilirliğini belirlemede kullanılır. Gözlemci değerlendirmeleri değişik şekillerde yapılabilir:

- a) Öğretmenlerin öğrencileri değerlendirmeleri
- b) Öğrencilerin eğitimi değerlendirmeleri
- c) Yöneticilerin astlarını veya astların yöneticileri değerlendirmeleri
- d) Doktorların hastaları değerlendirmeleri
- e) Gözlemcilerin sporcuları değerlendirmeleri

Sınıfıçi korelasyon analizi bir kişiye ait gözlem değerlerinin diğer kişilerin gözlem değerlerine ne ölçüde benzer olduğunu gösterir. Kişiye ait gözlem değerleri “grupıçi-sınıfıçi” değerlendirmeler ve diğer kişilere ait değerlendirmeler ise “gruplar arası-sınıflar arası” olarak tanımlanır. Eğer grup içi ve gruplar arasındaki puanların benzerlikleri yüksekse varyans düşük çıkar ve verilerin güvenilir olduğuna karar verilir.

Sınıfıçı (intraçlass) aynı deęiřkenin tekrarlı ölçümlerini analiz ederken, sınıflararası (interclass) iki farklı deęiřkenin analizinde kullanılır. Farklı ölçüm birimleriyle ölçülmüř iki veri setinin iliřki miktarının ölçümünde sınıflararası analiz uygundur.

Sınıfıçı güvenilirlik katsayısı denemeden denemeye ya da günden güne iki ya da daha fazla ölçümün yapıldığı durumlarda uygulanabilir ve tekrarlı ölçümlerin hem sistematik hem de ortalamalara göre deęiřikliklerine duyarlıdır.

Sınıfıçı güvenilirlik katsayısı (ICC) bulunurken tek yönlü anova ya da tekrarlı ölçümlerde anovadan yararlanılır. Tekrarlı ölçümlerde anova kullanılarak hem sınıfıçı güvenilirlik katsayısı hemde Cronbach's alfa katsayısı bulunabilir.

Fleiss (1981) ve Cicchetti and Sparrow (1981) göre sınıfıçı korelasyon katsayıları aşağıdaki gibi değerlendirilir:

< 0.40	= Zayıf
$0.40 - 0.59$	= Orta
$0.60 - 0.74$	= İyi
> 0.74	= Mükemmel

Sınıfıçı korelasyon katsayıları yapılan ölçümün niteliğine veya araştırmamanın modeline göre değişir. Shrout ve Fleiss sınıfıçı korelasyon analizi için üç farklı model önermişlerdir. Her bir model kendi içinde ikiye ayrılır. Araştırmacı güvenilirliği tek bir gözlemci için (*Single Measure Intraclass Correlation*) veya gözlemcilerin ortalama puanlarına (*Average Measure Intraclass Correlation*) dayalı olarak hesaplar. Seçilen modeller (1,1) veya (1,k) gibi simgelerle gösterilir. Parantez içindeki birinci rakam modeli, ikincisi gözlemci sayısını gösterir. Eğer birden fazla gözlemci veya birden fazla ölçüm yapılmışsa bunların ortalaması alınması gerekir.

Sınıfıçi Korelasyon Katsayısı Modelleri:

One-way random

$$\text{Single score ICC} = \frac{MS_{BC} - MS_{WC}}{MS_{BC} + (k - 1)MS_{WC}}$$

$$\text{Average score ICC} = \frac{MS_{BC} - MS_{WC}}{MS_{BC}}$$

Two-way random

$$\text{Single score ICC} = \frac{MS_{BC} - MS_E}{MS_{BC} + (k - 1)MS_E}$$

Consistency

$$\text{Average score ICC} = \frac{MS_{BC} - MS_E}{MS_{BC}}$$

$$\text{Single score ICC} = \frac{MS_{BC} - MS_E}{MS_{BC} + (k - 1)MS_E + \frac{k}{n}(MS_{BM} - MS_E)}$$

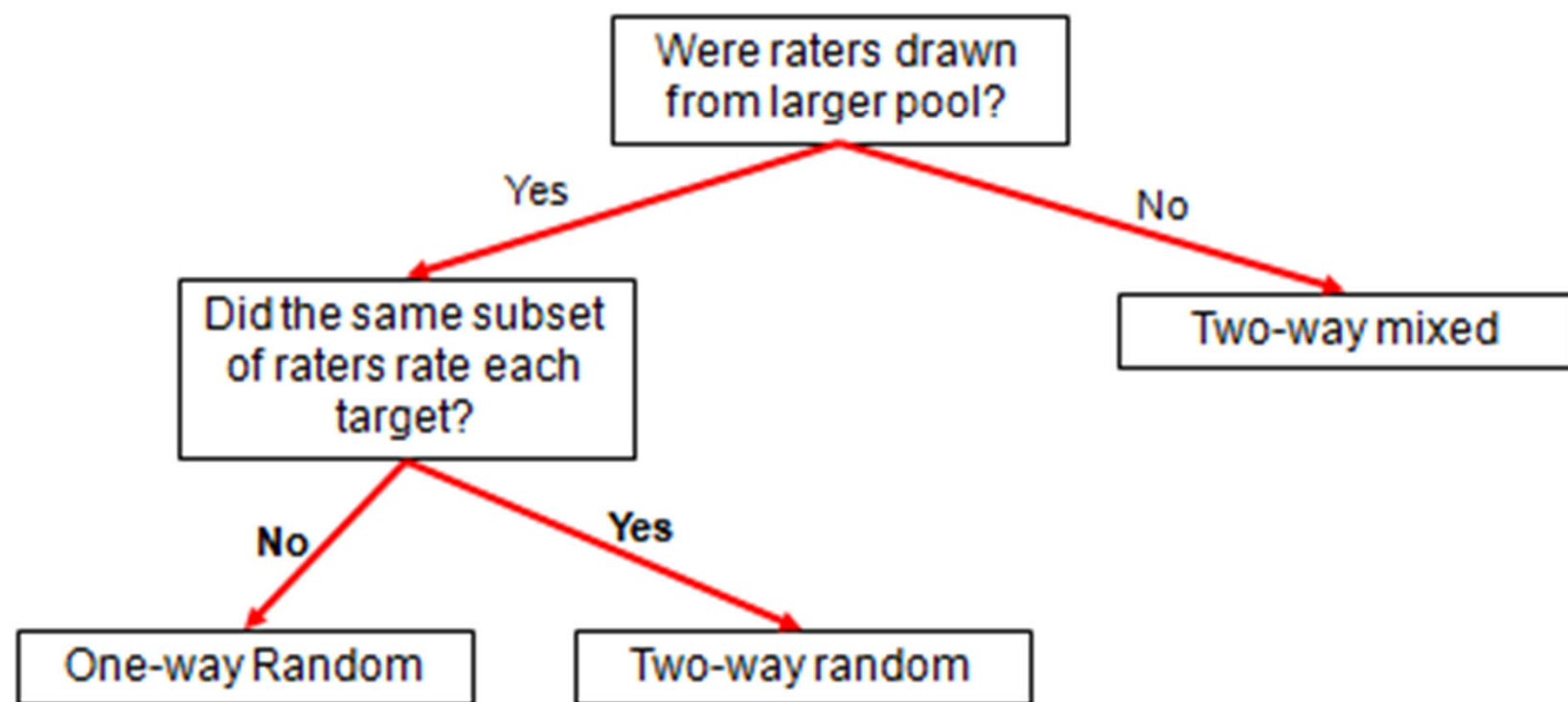
Absolute Agreement

$$\text{Average score ICC} = \frac{MS_{BC} - MS_E}{MS_{BC} + \frac{1}{n}(MS_{BM} - MS_E)}$$

k: Hakem sayısı

n: Gözlenen kişi sayısı

ICC Model Decision Tree (Consideration 1: Raters)



For two-way models you must choose **TYPE** as well
(Consideration 2: Individual rater variability):

- ✓ **consistency** (interested in whether targets ranked the same)
- ✓ **absolute agreement** (interested in whether targets got exact same scores)

i. One-Way Random-(1,1): Bu yöntemde rasgele seçilen kişiler yine rasgele seçilen çok sayıda farklı gözlemciler tarafından değerlendirilir. Gözlemcilerin her biri bütün kişileri değerlendirmez. Kişiler farklı gözlemciler tarafından değerlendirilir. Tek yönlü varyans analizi puan ortalamaları arasında fark bulunup bulunmadığını belirler. Gözlemcilerin rasgele atanmasının zorluğu nedeniyle bu model yaygın olarak kullanılmaz. Örneğin bir araştırmaya rasgele seçilen 6 kişi katılmışsa ve bu 6 kişi rasgele seçilen 4 hakem tarafından farklı puanlar verilmişse bu model ile **puanlar arasındaki tutarlılık araştırılır.** SPSS de gözlemcilerin ortalama güvenilirlik katsayısı için "Single Measure Intraclass Correlation" ile belirlenir.

Reliability Analysis: Statistics

Descriptives for

- Item
- Scale
- Scale if item deleted

Inter-Item

- Correlations
- Covariances

Continue

Cancel

Help

Summaries

- Means
- Variances
- Covariances
- Correlations

ANOVA Table

- None
- F test
- Friedman chi-square
- Cochran chi-square

Hotelling's T-square

Tukey's test of additivity

Intraclass correlation coefficient

Model: One-Way Random

Type: Consistency

Confidence interval: 95 %

Test value: 0

i. One-Way Random-(1,k): Bu yöntemde rasgele seçilen kişiler yine rasgele seçilen çok sayıda farklı gözlemciler tarafından değerlendirilir. Gözlemcilerin her biri bütün kişileri değerlendirir. Gözlemcilerin yaptığı değerlendirmelerin ortalaması alınır. Örneğin personel seçim mülakatlarında bu yöntem kullanılır. SPSS de gözlemcilerin ortalama güvenilirlik katsayısı için “Average Measure Intraclass Correlation” ile belirlenir.

Reliability Analysis: Statistics

Descriptives for
 Item
 Scale
 Scale if item deleted

Inter-Item
 Correlations
 Covariances

Summaries
 Means
 Variances
 Covariances
 Correlations

ANOVA Table
 None
 F test
 Friedman chi-square
 Cochran chi-square

Hotelling's T-square
 Tukey's test of additivity

Intraclass correlation coefficient

Model: One-Way Random Type: Consistency

Confidence interval: 95 % Test value: 0

Buttons: Continue, Cancel, Help

ii. Two-Way Random-(2,1): İki yönlü tesadüfi etki modelinde önceden belirlenmiş belirli sayıda gözlemci/hakem vardır. Bu gözlemciler daha büyük bir gözlemciler grubundan rasgele seçilir. Bu modelde kişilerin tamamı hep aynı gözlemciler tarafından değerlendirilir. Bu modelin tek yönlü anova dan farkı, puanlar üzerinde gözlemcilerin ve kişilerin birlikte etkili olabileceğidir. SPSS de gözlemcilerin ortalama güvenilirlik katsayısı için “Single Measure Intraclass Correlation” ile belirlenir.

Reliability Analysis: Statistics

Descriptives for
 Item
 Scale
 Scale if item deleted

Inter-Item
 Correlations
 Covariances

Summaries
 Means
 Variances
 Covariances
 Correlations

ANOVA Table
 None
 F test
 Friedman chi-square
 Cochran chi-square

Hotelling's T-square
 Tukey's test of additivity

Intraclass correlation coefficient

Model: Two-Way Random Type: Absolute Agreement

Confidence interval: 95 % Test value: 0

Buttons: Continue, Cancel, Help

Absolute agreement:
Gözlemciler arasındaki hatalar sistematik ise tercih edilir.

ii. Two-Way Random-(2,k): Herbir kiři aynı gözlemciler tarafından değerlendirilir. Gözlemciler daha büyük bir ana kütle için temsilcileridir. k sayıdaki gözlemcinin verdiği puanların ortalaması alınarak güvenilirlik katsayısı bulunur. SPSS de gözlemcilerin ortalama güvenilirlik katsayısı için “Average Measure Intraclax Correlation” ile belirlenir.

Reliability Analysis: Statistics

Descriptives for

- Item
- Scale
- Scale if item deleted

Inter-Item

- Correlations
- Covariances

Summaries

- Means
- Variances
- Covariances
- Correlations

ANOVA Table

- None
- F test
- Friedman chi-square
- Cochran chi-square

Hotelling's T-square

Tukey's test of additivity

Intraclass correlation coefficient

Model: Two-Way Random

Type: Consistency

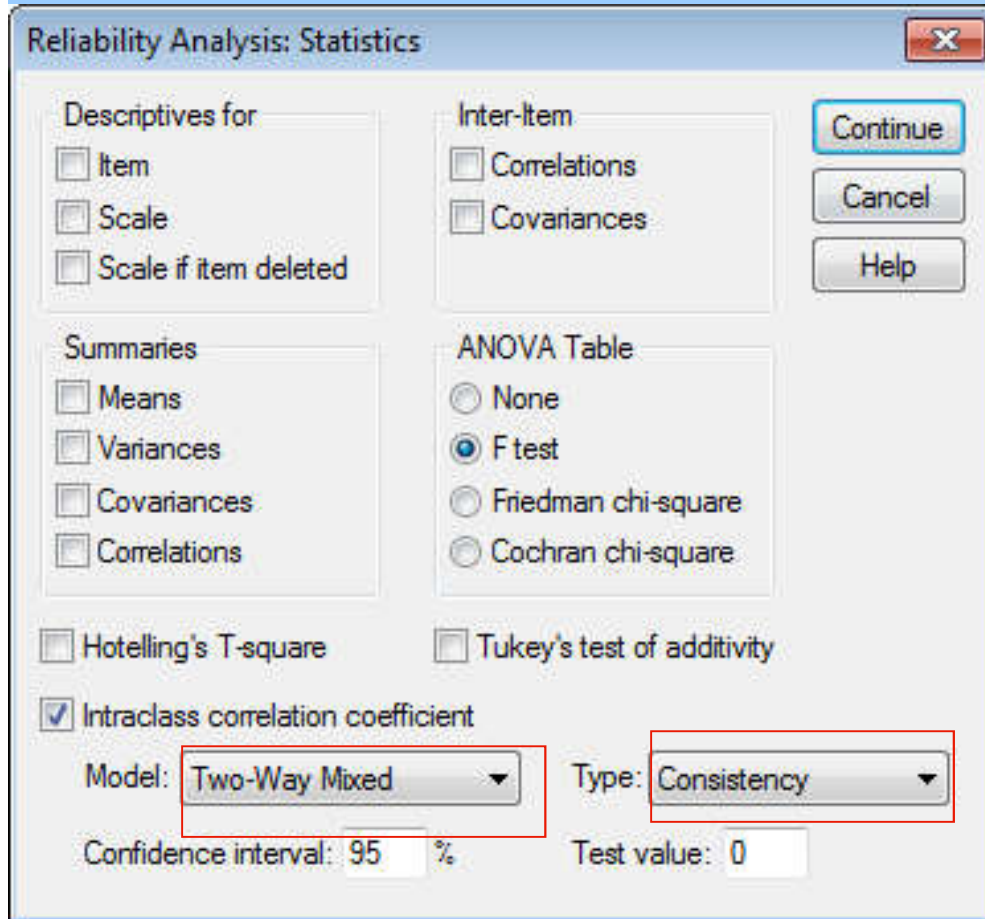
Confidence interval: 95 %

Test value: 0

Buttons: Continue, Cancel, Help

Consistency: Gözlemciler arasındaki hatalar sistematik hata içermiyorsa bu seçenek seçilir.

iii. Two-Way Mixed (3,1): İki yönlü karma etki modelinde kişiler aynı gözlemciler tarafından değerlendirilir. Gözlemciler anakütleyi oluşturan kişilerden rasgele seçilmemiştir. SPSS de gözlemcilerin ortalama güvenilirlik katsayısı için “Single Measure Intraclass Correlation” ile belirlenir.



Reliability Analysis: Statistics

Descriptives for

- Item
- Scale
- Scale if item deleted

Inter-Item

- Correlations
- Covariances

Summaries

- Means
- Variances
- Covariances
- Correlations

ANOVA Table

- None
- F test
- Friedman chi-square
- Cochran chi-square

Hotelling's T-square

Tukey's test of additivity

Intraclass correlation coefficient

Model: Two-Way Mixed

Type: Consistency

Confidence interval: 95 %

Test value: 0

Buttons: Continue, Cancel, Help

Consistency: Gözlemciler arasındaki hatalar sistematik hata içermiyorsa bu seçenek seçilir.

iii. Two-Way Mixed (3,k): Tüm kişiler aynı gözlemciler tarafından değerlendirilir. Gözlemciler anakütleyi oluşturan kişilerden rasgele seçilmemiştir. Güvenilirlik katsayısı k sayıdaki gözlemcinin vermiş olduğu puanların ortalaması alınarak hesaplanır. Bilimsel araştırmalarda hakemlerin rasgele seçilmesi nadirdir, daha çok iradi olarak belirlendiğinden genelde two-way mixed modeli tercih edilir. SPSS de gözlemcilerin ortalama güvenilirlik katsayısı için “Average Measure Intraclass Correlation” ile belirlenir.

Reliability Analysis: Statistics

Descriptives for
 Item
 Scale
 Scale if item deleted

Inter-Item
 Correlations
 Covariances

Summaries
 Means
 Variances
 Covariances
 Correlations

ANOVA Table
 None
 F test
 Friedman chi-square
 Cochran chi-square

Hotelling's T-square
 Tukey's test of additivity

Intraclass correlation coefficient

Model: Two-Way Mixed Type: Consistency

Confidence interval: 95 % Test value: 0

Buttons: Continue, Cancel, Help

Consistency: Gözlemciler arasındaki hatalar sistematik hata içermiyorsa bu seçenek seçilir.

Örnek 2.12. Four Ratings on Six Cases (adapted from Shrout and Fleiss, 1979)

Case	Rater 1	Rater 2	Rater 3	Rater 4
1	9	2	5	8
2	6	1	3	2
3	8	4	6	8
4	7	1	2	6
5	10	5	6	9
6	6	2	4	7
Mean	7.67	2.50	4.33	6.67
SD	1.63	1.64	1.63	2.50

Analysis of Variance

Source	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>
Between cases (BC)	56.21	5	11.24	
Within cases (WC)	112.75	18	6.26	
Between measures (BM)	97.46	3	32.49	31.87
Residuals (E)	15.29	15	1.02	
Total	168.96	23	7.35	

	hakem1	hakem2	hakem3	hakem4
1	9	2	5	8
2	6	1	3	2
3	8	4	6	8
4	7	1	2	6
5	10	5	6	9
6	6	2	4	7
7				

Reliability Analysis

Items:

- hakem1
- hakem2
- hakem3
- hakem4

Model: Alpha

Scale label:

Buttons: OK, Paste, Reset, Cancel, Help, Statistics...

ANOVA

		Sum of Squares	df	Mean Square	F
Between People		56,208	5	11,242	31,866
Within People	Between Items	97,458	3	32,486	
	Residual	15,292	15	1,019	
	Total	112,750	18	6,264	
Total		168,958	23	7,346	

Grand Mean = 5,29

Reliability Analysis: Statistics

Descriptives for:

- Item
- Scale
- Scale if item deleted

Inter-Item:

- Correlations
- Covariances

Summaries:

- Means
- Variances
- Covariances
- Correlations

ANOVA Table:

- None
- F test
- Friedman chi-square
- Cochran chi-square

Hotelling's T-square

Tukey's test of additivity

Intraclass correlation coefficient

Model: One-Way Random

Type: Consistency

Confidence interval: 95 %

Test value: 0

Buttons: Continue, Cancel, Help

Intraclass Correlation Coefficient

	Intraclass Correlation	95% Confidence Interval		Value
		Lower Bound	Upper Bound	
Single Measures	.166	-.133	.723	1,7
Average Measures	.443	-.884	.912	1,7

One-way random effects model where people effects are random.

ANOVA

		Sum of Squares	df	Mean Square	F	Sig
(BC)	Between People	56,208	5	11,242		
	Within People					
	Between Items (BM)	97,458	3	32,486	31,866	,000
	Residual (E)	15,292	15	1,019		
	Total (WC)	112,750	18	6,264		
	Total	168,958	23	7,346		

Grand Mean = 5,29

One-way random

$$\text{Single score ICC} = \frac{MS_{BC} - MS_{WC}}{MS_{BC} + (k-1)MS_{WC}} = .166$$

$$\text{Average score ICC} = \frac{MS_{BC} - MS_{WC}}{MS_{BC}} = .443$$

One-way random:(1,1)One-way random:(1,k)

$$\text{Single ICC} = \frac{11,24 - 6,26}{11,24 + (4-1) * 6,26} = 0,166$$

$$\text{Average ICC} = \frac{11,24 - 6,26}{11,24} = 0,443$$

ANOVA

	Sum of Squares	df	Mean Square	F	Sig
(BC) Between People	56,208	5	11,242		
Within People					
Between Items (BM)	97,458	3	32,486	31,866	,000
Residual (E)	15,292	15	1,019		
Total (WC)	112,750	18	6,264		
Total	168,958	23	7,346		

Grand Mean = 5,29

Two-way random

Consistency

$$\text{Single score ICC} = \frac{MS_{BC} - MS_E}{MS_{BC} + (k - 1)MS_E} = .715$$

$$\text{Average score ICC} = \frac{MS_{BC} - MS_E}{MS_{BC}} = \alpha = .909$$

Two-way random:(2,1)

$$\text{Single ICC} = \frac{11,24 - 1,02}{11,24 + (4 - 1) * 1,02} = 0,715$$

Two-way random:(2,k)

$$\text{Average ICC} = \frac{11,24 - 1,02}{11,24} = 0,909$$

Reliability Analysis: Statistics

Descriptives for

Item

Scale

Scale if item deleted

Inter-Item

Correlations

Covariances

Continue

Cancel

Help

Summaries

Means

Variances

Covariances

Correlations

ANOVA Table

None

F test

Friedman chi-square

Cochran chi-square

Hotelling's T-square

Tukey's test of additivity

Intraclass correlation coefficient

Model: Two-Way Random

Type: Consistency

Confidence interval: 95 %

Test value: 0

Intraclass Correlation Coefficient

	Intraclass Correlation ^a	95% Confidence Interval		Value
		Lower Bound	Upper Bound	
Single Measures	,715 ^b	,342	,946	11,027
Average Measures	,909	,676	,986	11,027

Two-way random effects model where both people effects and measures effects are

ANOVA

(BC)

	Sum of Squares	df	Mean Square	F	Sig
Between People	56,208	5	11,242		
Within People					
Between Items (BM)	97,458	3	32,486	31,866	,000
Residual (E)	15,292	15	1,019		
Total (WC)	112,750	18	6,264		
Total	168,958	23	7,346		

Grand Mean = 5,29

Two-way random

Absolute Agreement

Single score ICC =

$$\frac{MS_{BC} - MS_E}{MS_{BC} + (k-1)MS_E + \frac{k}{n}(MS_{BM} - MS_E)} = .290$$

Average score ICC =

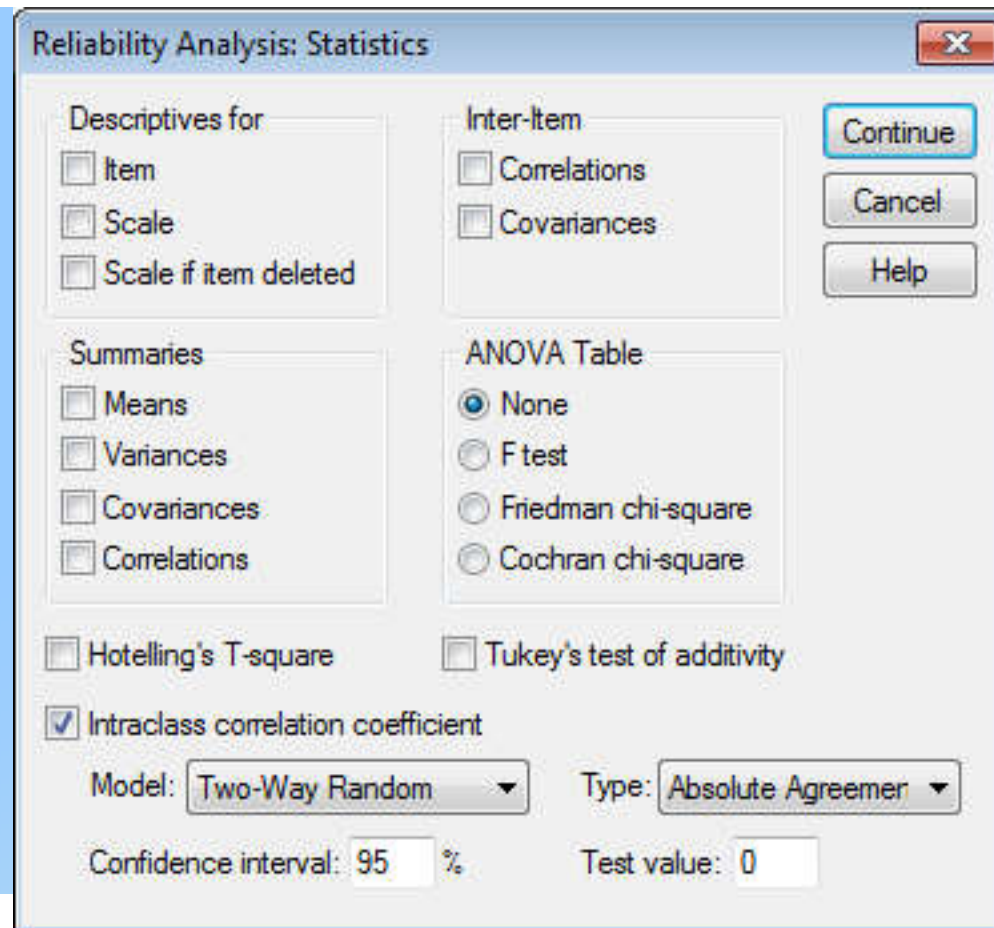
$$\frac{MS_{BC} - MS_E}{MS_{BC} + \frac{1}{n}(MS_{BM} - MS_E)} = .620$$

Two-way random:(2,1)

$$\text{Single ICC} = \frac{11,24 - 1,02}{11,24 + (4-1) * 1,02 + 4 / 6(32,49 - 1,02)} = 0,29$$

Two-way random:(2,k)

$$\text{Average ICC} = \frac{11,24 - 1,02}{11,24 + 1 / 6(32,49 - 1,02)} = 0,620$$



Intraclass Correlation Coefficient

	Intraclass Correlation ^a	95% Confidence Interval		Value
		Lower Bound	Upper Bound	
Single Measures	,290 ^b	,019	,761	11,027
Average Measures	,620	,039	,929	11,027

Two-way random effects model where both people effects and measures effects are

Örnek 2.13.

Kişiler\Nesneler	1.Gözlemci	2.Gözlemci	3.Gözlemci
1	12	18	16
2	18	18	17
3	19	14	16
4	20	15	17
5	12	12	11

Sütunlar gözlemcileri, satırlar kişileri gösterir. Küme içi korelasyon analizi gruplar arasındaki varyansın toplam varyansa oranıdır.

$$ICC = \frac{\sigma_B^2}{\sigma_B^2 + \sigma_W^2}$$

σ_B^2 : Gruplar arası değerlerin varyansı

σ_W^2 : Gruplar içi değerlerin varyansı

ANOVA

(BC)

	Sum of Squares	df	Mean Square	F	Sig
Between People	70,000	4	17,500		
Within People	Between Items(BM)	2,133	2	1,067	,198
	Residual (E)	43,200	8	5,400	,825
	Total (WC)	45,333	10	4,533	
Total	115,333	14	8,238		

Grand Mean = 15,67

One-way random

$$\text{Single score ICC} = \frac{MS_{BC} - MS_{WC}}{MS_{BC} + (k-1)MS_{WC}}$$

$$\text{Average score ICC} = \frac{MS_{BC} - MS_{WC}}{MS_{BC}}$$

	Intraclass Correlation
Single Measures	,488
Average Measures	,741

$$\text{Single ICC} = \frac{17,5 - 4,53}{17,5 + (3-1) * 4,53} = 0,49$$

$$\text{Average ICC} = \frac{17,5 - 4,53}{17,5} = 0,74$$

ANOVA

(BC)

	Sum of Squares	df	Mean Square	F	Sig
Between People	70,000	4	17,500		
Within People	Between Items (BM)	2,133	2	1,067	,198
	Residual (E)	43,200	8	5,400	,825
	Total (WC)	45,333	10	4,533	
Total	115,333	14	8,238		

Grand Mean = 15,67

Two-way random

$$\text{Single score ICC} = \frac{MS_{BC} - MS_E}{MS_{BC} + (k - 1)MS_E}$$

Consistency

$$\text{Average score ICC} = \frac{MS_{BC} - MS_E}{MS_{BC}}$$

	Intraclass Correlation ^a
Single Measures	,428 ^b
Average Measures	,691

Two-way random effects model where

$$\text{Single ICC} = \frac{17,5 - 5,4}{17,5 + (3 - 1) * 5,4} = 0,428$$

$$\text{Average ICC} = \frac{17,5 - 5,4}{17,5} = 0,69$$

ANOVA

(BC)

	Sum of Squares	df	Mean Square	F	Sig
Between People	70,000	4	17,500		
Within People	Between Items (BM)	2,133	2	1,067	,198
	Residual (E)	43,200	8	5,400	,825
	Total (WC)	45,333	10	4,533	
Total	115,333	14	8,238		

Grand Mean = 15,67

Two-way random

Single score ICC =

$$\frac{MS_{BC} - MS_E}{MS_{BC} + (k-1)MS_E + \frac{k}{n}(MS_{BM} - MS_E)}$$

Absolute Agreement

Average score ICC =

$$\frac{MS_{BC} - MS_E}{MS_{BC} + \frac{1}{n}(MS_{BM} - MS_E)}$$

	Intraclass Correlation ^a
Single Measures	,471 ^b
Average Measures	,727

Two-way random effects model where

$$Average ICC = \frac{17,5 - 5,4}{17,5 + 1/5(1,06 - 5,4)} = 0,73$$

2.1.4. İç Tutarlılık (Internal consistency)

Bu yaklaşımda kavramı ölçerken çok sayıda maddeden oluşan (Likert tipi) bir ölçek kullanılarak, ölçekte yer alan maddeler arasındaki korelasyona bakarak içsel tutarlılık analizi yapılır. İç tutarlılık yönteminin amacı bir testin her bir sorusunun aynı niteliği ne kadar ölçtüğünü hesaplamaktır. İç tutarlılık Cronbach alpha, split half, Guttman, Paralel ve Kesin Paralel yöntemler ile hesaplanır. Sonuç 1'e ne kadar yakınsa iç tutarlılık o kadar yüksektir.

Ölçeğin tümü ve alanları için korelasyona dayalı madde analizi yapılır. Hotelling's T^2 analizi ($p < 0,05$ ise ortalama olarak ölçek maddeleri arasında fark var demektir.) ve yarı-test (split half, guttman) güvenilirlik sonuçları elde edilir ve Cronbach Alpha katsayıları hesaplanır. Her bir soru ile total skor puanları arasında korelasyon değeri $+0,3$ 'ün altında olan sorular ölçekten çıkarılır.

Ölçeğin tamamına ait iç tutarlılığa Cronbach Alpha katsayısı ile bakılır. Bazı sorular negatif, bazıları pozitif olur ve bunların Cronbach Alpha güvenilirlik katsayısına ayrı ayrı bakılır.

Güvenilirlik katsayısının hesaplanmasında Spearman ve Guttman değerleri hesaplanır. Bu değerlere “iç tutarlık katsayıları” denir. Güvenilirlik analizi için bu yöntemlerin kullanımında ölçülecek konuya ilişkin ölçekteki soru sayısının $k > 20$ ve denek sayısının $n > 50$ 'den fazla olmasına dikkat edilir. Güvenirlik katsayılarının 0.60 'ın üstünde olması gerekmektedir (Ural ve Kılıç,2005). Bir testin iç tutarlılık katsayısının $0,90$ 'nın üstünde olması, o testin mükemmel güvenilirlikte olduğunu göstermektedir. İçsel tutarlığı ölçmede çeşitli yaklaşımlar vardır.

a) Yarıya Bölme Yöntemi (Split Half)

Bu yöntem, tek bir test formu, tek bir öğrenci grubu ve tek bir test uygulaması gerektirir. Bu yöntemle test güvenilirliğini tahmin etmede, uygulanmış olan test iki eşdeğer yarıya bölünerek öğrencilerin testin iki yarısından aldıkları puanlar arasındaki korelasyon hesaplanır ve daha sonra bu hesaplanan korelasyondan hareketle Sperman-Brown formülünden de yararlanarak testin bütünüünün güvenilirliği kestirilir.

Testi iki eşdeğer yarıya bölmenin bir yolu, test- teki tek numaralı sorularla çift numaralı soruları ayrı puanlamaktır.

Bu yöntemde, ölçekte yer alan sorular ikiye ayrılır ve bu iki parça arasında korelasyon hesaplanır.

Ayrıca bu iki parçadan her biri için ayrı ayrı alfa katsayıları hesaplanır.

Eğer k soru içeren ölçekteki soru sayısı çift ise, her bir alt gruba düşen soru sayısı $k/2$ 'dir.

Yani soruların ilk yarısı birinci alt gruba, ikinci yarısı ikinci alt gruba dahil edilir.

Eğer ölçekteki soru sayısı tek ise, birinci alt gruba dâhil edilecek soru sayısı $k_1=(k+1)/2$ ile bulunur. İkinci alt gruba dâhil edilecek soru sayısı ise $k_2=k-k_1$ şeklinde bulunur.

SPSS'te Split Half yöntemi ile sorular farklı biçimlerde ikiye bölünebilir.

1. Soruların ilk yarısı ve ikinci yarısı
2. Tek numaralı sorular, çift numaralı sorular
3. Rasgele
4. Kolaylık ve zorluk açısından

RELIABILITY

```
/VARIABLES=y1 y2 y3 y4 y5 y6 y7 y8 y9 y10  
/SCALE('ALL VARIABLES') ALL  
/MODEL=SPLIT.
```

RELIABILITY

```
/VARIABLES=y1 y3 y5 y7 y9 y2 y4 y6 y8 y10  
/SCALE('ALL VARIABLES') ALL  
/MODEL=SPLIT.
```

Ölçek ikiye bölümdükten sonra Spearman-Brown formülü ile güvenilirlik katsayısı bulunabilir.

$$r_{SB} = \frac{2r_{xy}}{1 + r_{xy}} \quad r_{xy} = \text{birinci yarı ile ikinci yarı arasındaki korelasyon}$$

Araştırmacı birden fazla yarıya bölme yöntemi uygulamışsa elde edilen farklı güvenilirlik katsayılarının ortalaması alınır. Yarıya bölme güvenilirlik katsayısı en az 0,70 olmalıdır.

İkiye Bölünmüş Yöntem (Split Half) :

Reliability Statistics			
Cronbach's Alpha	Part 1	Value	,836
		N of Items	5 ^a
	Part 2	Value	,796
		N of Items	5 ^b
	Total N of Items		10
Correlation Between Forms			,692
Spearman-Brown Coefficient	Equal Length		,818
	Unequal Length		,818
Guttman Split-Half Coefficient			,816

a. The items are: yontem1, yontem2, yontem3, yontem4, yontem5.

b. The items are: yontem6, yontem7, yontem8, yontem9, yontem10.

Ölçek ilk 5 soru ve son 5 soru diye ikiye bölündüğünde alfa değerleri birbirine yakın ve yüksek çıkmıştır. Bu değerler soruların birbirini izleyen nitelikte düzenlendiğini ifade eder.

Bu modelde ölçeğin güvenilirliği için formlar arası korelasyonlar (Correlation Between Forms) belirler. Aynı zamanda Guttman Split Half, Eşit ve eşit olmayan uzunluk Spearman-Brown katsayılarına güvenilirlik ölçüsü olarak bakılır. Buradaki değerler ölçek güvenilirliğini yüksek olduğunu göstermektedir.

Örnek 2.14.

KATILIMCI	MADDELER						I.YARI (X)	II.YARI (Y)
	1	2	3	4	5	6		
1	4	3	2	2	3	3	9	8
2	5	4	3	4	4	3	12	11
3	2	3	2	3	3	4	7	10
4	5	5	4	2	5	4	14	11
5	2	3	2	2	1	2	7	5

$$r_{SB} = \frac{2r_{xy}}{1 + r_{xy}} = \frac{2 \times 0,69}{1 + 0,69} = 0,81$$

	s1	s2	s3	s4	s5	s6
1	4	3	2	2	3	3
2	5	4	3	4	4	3
3	2	3	2	3	3	4
4	5	5	4	2	5	4
5	2	3	2	2	1	2
6						

Reliability Analysis

Items:

- s1
- s2
- s3
- s4
- s5
- s6

Model: Split-half

Scale label:

OK
Paste
Reset
Cancel
Help

Reliability Statistics

Cronbach's Alpha	Part 1	Value	,897
		N of Items	3 ^a
	Part 2	Value	,646
		N of Items	3 ^b
	Total N of Items		6
Correlation Between Forms			,693
Spearman-Brown Coefficient	Equal Length		,818
	Unequal Length		,818

Yarıya Bölme Yöntemlerinden Rulon Formülü:

Rulon (1939) tarafından geliştirilmiştir. Test ikiye bölünür. Yarı test puanları arasındaki farklılık bulunur. Daha sonra fark puanlarının varyansı ve toplam puanların varyansı bulunarak fark puanlarının varyansı toplam puanların varyansına bölünür.

$$\rho = 1 - \frac{\sigma_{fark}^2}{\sigma_{top}^2}$$

Örnek 2.15. (Rulon Formülü):

KATILIMCI	MADDELER						I.YARI	II.YARI	FARK	TOPLAM
	1	2	3	4	5	6				
1	4	3	2	2	3	3	9	8	1	17
2	5	4	3	4	4	3	12	11	1	23
3	2	3	2	3	3	4	7	10	3	17
4	5	5	4	2	5	4	14	11	3	25
5	2	3	2	2	1	2	7	5	2	12

VARYANS

1,000

27,200

$$\rho = 1 - \frac{1}{27,2} = 0,963$$

Yarıya Bölme Yöntemlerinden Guttman Formülü:

Bu formülde önce maddeler tek ve çift olmak üzere iki gruba ayrılır. İki varsayım söz konusudur:

a) İlk yarının güvenilirliği ile ikinci yarının güvenilirliği eşit değildir.

b) İlk yarının varyansı ile ikinci yarının varyansı eşit değildir.

$$r_G = \frac{2(\sigma_{TOP}^2 - \sigma_1^2 - \sigma_2^2)}{\sigma_{TOP}^2}$$

KATILIMCI	MADDELER						I.YARI	II.YARI	TOPLAM
	1	3	5	2	4	6			
1	4	2	3	3	2	3	9	8	17
2	5	3	4	4	4	3	12	11	23
3	2	2	3	3	3	4	7	10	17
4	5	4	5	5	2	4	14	11	25
5	2	2	1	3	2	2	5	7	12
VARYANS							13,300	3,300	27,200

$$r_G = \frac{2(27,2 - 13,3 - 3,3)}{27,2}$$

$$= 0,779$$

Reliability Statistics

Cronbach's Alpha	Part 1	Value	,902
		N of Items	3 ^a
Cronbach's Alpha	Part 2	Value	,455
		N of Items	3 ^b
Total N of Items			6
Correlation Between Forms			,800
Spearman-Brown Coefficient	Equal Length		,889
	Unequal Length		,889
Guttman Split-Half Coefficient			,779

a. The items are: VAR00001, VAR00003, VAR00005.

b. The items are: VAR00002, VAR00004, VAR00006.

Guttman Katsayıları : Yarı testler arasındaki puan farkları kullanılır, ancak tüm testin güvenilirliğini gösteren bir indeks hesaplanır. Bu yöntemde 6 katsayı hesaplanır. Güvenirliği kovaryans ya da varyans yaklaşımı ile hesaplayan diğer bir yaklaşımdır.

Paralel Model : *Bu model ölçekteki sorular için eşit varyanslılığı ve karşılıklı tekrar soruları içinde eşit hata varyanslılığını varsayar. Bu modelle en büyük benzerlik tahminleri yapılır ve tahminlerin verilere uygunluğu ki-kare ile test edilir.*

Kesin Paralel Model (Strict Parallel): *Bu modelde eşit varyanslılığın yanı sıra sorular arasındaki ortalamalarında eşit olduğu varsayımı söz konusudur. Maddelerin ortalamalarının eşit olmadığı Hotelling T-Square test for equality of means ile test edilir. $P > 0,05$ ise soru ortalamaları arasında fark yoktur denir.*

Ölçeği oluşturan her bir sorunun tanımlayıcı istatistiklerine bakarak Güvenilirlik Analizi için hangi modelin kullanılacağına karar verilir. Sorular arasında eşit varyanslılık (homojen) varsa Alfa ve Paralel model kullanılabilir. Eşit varyanslılığın yanında sorular arasındaki ortalamalarda homojen ise Kesin Paralel model tercih edilmelidir.

Guttman Yöntemi

Reliability Statistics

Lambda	1	,794
	2	,884
	3	,882
	4	,816
	5	,856
	6	,887
N of Items		10

Guttman modeline göre hesaplanan 6 güvenilirlik katsayısının hepsi oldukça güvenilir görülmektedir.

Paralel Yöntemi

Reliability Statistics

Common Variance	1,190
True Variance	,508
Error Variance	,682
Common Inter-Item Correlation	,427
Reliability of Scale	,882
Reliability of Scale (Unbiased)	,882

Bu model ölçekteki sorular için eşit varyanslılığı ve karşılıklı tekrar soruları içinde eşit hata varyanslılığını varsayar. Bu yöntemle göre güvenilirlik katsayısı olarak Reliability of Scale değerine bakılır.

Kesin Paralel Yöntemi

Reliability Statistics

Common Mean	3,141
Common Variance	1,244
True Variance	,504
Error Variance	,740
Common Inter-Item Correlation	,404
Reliability of Scale	,871
Reliability of Scale (Unbiased)	,872

b) Hoy'un varyans analiz yöntemi

Bireyler ve maddeler arasındaki değişimi birlikte hesaba katan bir yöntemdir. Sınıflar arası (Intra class) korelasyon veya sınıflar içi (inter class) korelasyon katsayıları hesaplanabilir.

Bu yöntem maddelerin iç tutarlılığını ölçer ve varyans analizine dayanır.

$$\alpha = \frac{MS_{Between} - MS_{Within(error)}}{MS_{Between}}$$

	df	Sum of Square	Mean Square	F
Between	k-1		$MS_{Between}$	
Within (Error)	n-k		MS_{Within}	
Total	n-1			

İç Tutarlılık

c). Maddeler Arası Korelasyon Katsayılarının Ortalaması

Bu analizde ölçeğin/testin toplam puanları hesaplamaya katılmaz. Burada maddeler arası korelasyon analizi yapılır ve değişkenlerin korelasyon katsayılarının ortalaması alınır. Bu analiz test/ölçek maddelerinin ne ölçüde birbirleriyle ilişkili olduğu hakkında bilgi verir.

Maddeler arasında negatif korelasyon varsa, bu maddeler arasında ters bir ilişki olduğu anlamına gelir. Negatif korelasyon olduğunda bu madde(lerin) diğer maddelerle olan ilişkisine bakılarak ölçekten çıkarılması gerekir.

İç Tutarlılık

d). Madde Toplam Puan Korelasyonu Katsayılarının Ortalaması

Bu yöntem test maddelerinden elde edilen puanlar ile testin toplam puanı arasındaki ilişkiyi açıklar. Madde toplam puan korelasyon katsayılarının ortalaması testin güvenilirliğini verir.

Madde toplam puan korelasyon katsayısı eğer 0,20'nin altında ise (örneklem eksikliğinden olabilir) bu maddelerde ciddi sorun var demektir. 0,2-0,3 çıkan korelasyonlar varsa bu maddeler eğer ciddi öneme sahipse testte kalırlar.

Madde toplam puan korelasyonu negatif çıkmışsa (yanlış kodlama, tersine çevirmeme, kasıtlı cevaplama gibi) bu maddeler ölçekten çıkarılmalıdır.

Yapılan madde toplam korelasyonları için gerçekleştirilen Pearson çarpım moment korelasyon analizi sonucuna tüm maddelerle toplam puan arasındaki ilişki istatistiksel açıdan $p < 0,05$ düzeyinde pozitif yönde anlamlı bulunmalı.

Madde toplam korelasyonu, test maddelerinden alınan puanlar ile testin toplam puanı arasındaki ilişkiyi açıklamaktadır. Madde kalan korelasyonu ise ele alınan maddenin kendisi hariç diğer maddelerden elde edilen toplam puanla olan ilişkisini belirtmektedir.

Ölçmenin Standart Hatası ($SH_{\ddot{o}}$):

Testin güvenilirliği için kullanılan tekniklerden biridir. Ölçmenin standart hatası bireylerin testten aldıkları puanların gerçek puandan olan sapma miktarını açıklamada kullanılır.

$$SH_{\ddot{o}} = s \times \sqrt{1 - r}$$

R : Güvenilirlik katsayısı

s : test puanlarının standart sapması

Güven aralığı: $\bar{X} \pm 1,96 \times SH_{\ddot{o}}$

KR-20 ve KR-21 Formülleri :

Bu formüller bilgi testlerinde sıklıkla kullanılır. Testteki her bir maddeye doğru cevap veren öğrenci yüzdesi hesaplanmışsa, testin güvenilirliği, Kuder-Richardson 20 ve 21 formülleriyle tahmin edilebilir. Bu yolla hesaplanan güvenilirlik katsayısı da testin iç tutarlılığının bir ölçüsüdür.

Test maddelerine verilen cevaplar 2 şıklı ise (evet/hayır, doğru/yanlış gibi) KR-20 ve KR-21 formülleri kullanılabilir. Test maddelerinin aynı ortalama ve varyansa sahip olduğu varsayımı ile hareket eder. Dolayısıyla da maddeler arası korelasyonlara bakılır. Maddeler iki şıklı puanlanmış ise bir madde (soru) nin doğru cevaplanma olasılığı (p) ve yanlış cevaplanma olasılığı (q) hesaplanabilir.

KR-20 formülü, test maddelerinin benzer zorlukta olmadığı varsayımı altında kullanılır. KR-20 formülünde temel varsayım ölçü maddelerinin tek bir yapıyı ölçüyor olmasıdır. Maddelerin içerikleri benzer olmalıdır. Maddeler iki kategorili olduğunda KR-20 ile Cronbach alfa aynı sonucu verir.

KR-20 Formülü

$$KR - 20 = \frac{K}{K - 1} \left[1 - \frac{\sum pq}{s^2} \right]$$

p : Soruların (maddelerin) doğru cevap oranı

q : Soruların (maddelerin) yanlış cevap oranı

K : Testteki soru (madde) sayısı

s²: Toplam puanların (herbir kişinin doğru cevap sayısı) varyansı (varyans bulunurken n-1 yerine n kullanılır.

KR-20>0,70 ise testin iç güvenilirliği yüksek demektir.

Örnek 13.

Öğ\ soru	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	T
1	1	0	0	1	1	1	0	1	1	1	7
2	1	1	0	1	1	0	1	1	1	1	8
3	0	0	1	1	1	0	0	1	0	1	5
4	0	1	0	1	0	1	1	1	1	1	7
5	0	0	0	0	0	0	0	1	0	0	1
6	1	0	0	0	0	1	1	1	0	0	4
7	1	0	0	0	1	0	0	0	1	1	4
8	1	0	1	0	0	1	1	0	1	1	6
9	1	1	1	1	1	0	0	1	1	1	8
10	1	1	1	0	0	1	1	1	1	1	8
p	0,7	0,4	0,4	0,5	0,5	0,5	0,5	0,8	0,7	0,8	
q	0,3	0,6	0,6	0,5	0,5	0,5	0,5	0,2	0,3	0,2	
pq	0,21	0,24	0,24	0,25	0,25	0,25	0,25	0,16	0,21	0,16	2,22

$$KR - 20 = \frac{k}{k-1} \left[1 - \frac{\sum pq}{s^2} \right] = \frac{10}{10-1} \left[1 - \frac{2,22}{4,76} \right] = 0,593$$

Soru.

Öğ\SORU	S1	S2	S3	S4	S5	S6	T
1	1	1	1	1	0	1	5
2	0	0	1	0	1	0	2
3	0	0	0	1	0	1	2
4	0	0	1	1	1	0	3
5	1	1	1	1	0	0	4
6	0	1	1	0	1	0	3
7	0	1	0	0	0	0	1
8	1	1	1	1	1	1	6
9	0	1	1	0	1	0	3
10	1	1	1	0	0	1	4
p	0,4	0,7	0,8	0,5	0,5	0,4	
q	0,6	0,3	0,2	0,5	0,5	0,6	
pq	0,24	0,21	0,16	0,25	0,25	0,24	1,35

$$KR - 20 = \frac{k}{k-1} \left[1 - \frac{\sum pq}{s^2} \right] = \frac{6}{6-1} \left[1 - \frac{1,35}{2,23} \right] = 0,473$$

KR-21 formülü eşit zorluğa sahip maddelerden oluşan (her sorunun güçlük derecesi aynı ise) ve sınıf ortamında uygulanan çoktan seçmeli testler için uygundur. Doğru cevap 1, yanlış cevaplar ise 0 olarak kodlanır. Testteki soruların zorluk dereceleri eşit olmalıdır.

$$KR-21 = \frac{k}{k-1} \left[1 - \frac{\bar{x}(k - \bar{x})}{ks^2} \right]$$

k: Madde (soru) sayısı

KR-20 ve KR-21 formüllerinden elde edilen katsayı 0,7'nin üstünde ise güvenilirlik iyidir. KR-20 ve KR-21 formülleri testin genel olarak güvenilirliği hakkında bilgi verir, tek tek maddelerin güvenilirliği hakkında bilgi vermez.

Soru.

Öğ\ soru	S1	S2	S3	S4	S5	S6	T
1	1	1	1	1	0	1	5
2	0	0	1	0	1	0	2
3	0	0	0	1	0	1	2
4	0	0	1	1	1	0	3
5	1	1	1	1	0	0	4
6	0	1	1	0	1	0	3
7	0	1	0	0	0	0	1
8	1	1	1	1	1	1	6
9	0	1	1	0	1	0	3
10	1	1	1	0	0	1	4
p	0,4	0,7	0,8	0,5	0,5	0,4	
q	0,6	0,3	0,2	0,5	0,5	0,6	
pq	0,24	0,21	0,16	0,25	0,25	0,24	1,35

$$KR - 21 = \frac{k}{k-1} \left[1 - \frac{\bar{x}(k - \bar{x})}{ks^2} \right] = \frac{6}{6-1} \left[1 - \frac{3,3(6-3,3)}{6 \times 2,01} \right] = 0,313$$

f) Cronbach Alfa Katsayısı:

Alfa katsayısı ölçekte yer alan k sorunun varyansları toplamının genel varyansa oranlaması ile bulunan bir ağırlık standart değişim ortalamasıdır. Sorular arasında negatif korelasyon varsa Alfa katsayısı da negatif çıkar. Bu durum güvenilirlik modelin bozulmasına neden olur.

Cronbach Alfa Katsayısı, ölçekte yer alan k sorunun homojen bir yapıyı açıklamak üzere bir bütün oluşturup oluşturmadıklarını araştırır.

Alfa Katsayısı, bireysel puanların k soru içeren bir ölçekte sorulara verilen cevapların toplanması ile bulunduğu durumlarda soruların birbirleri ile benzerliğini, yakınlığını, ortaya koyan bir katsayıdır.

Alfa ne kadar yüksek olursa ölçekteki maddelerin birbiriyle tutarlı ve aynı özelliği yansıtan maddeler olduğu anlaşılır.

- Cronbach alfa katsayısı istatistik temelleri tutarlı ve tüm soruları dikkate alarak hesaplandığından, genel güvenilirlik yapısını diğer katsayılara göre en iyi yansıtan katsayıdır.
- Sorular arası korelasyonu hesaplamanın bir yolu da Cronbach'ın alfasını hesaplamaktır. Eğer formdaki sorular güvenilir ise bu katsayı 1 yakın olur. Cronbach alfa test-tekrar testle hesaplanan güvenilirlik katsayısından daha az konservatiftir.

- Güvenirlilik Analizi, toplam puanlar üzerine kurulu ölçeklere (Likert Ölçeği, Q-Tipi Ölçek) dayalı araçların güvenilirliğini ortaya koymaya yarayan Cronbach Alfa katsayıları hesaplar.

Temel varsayımlar,

- "Her soru toplam skorun bir doğrusal bileşeni olmalıdır." ve
- "Ölçekte toplanabilirlik özelliğinin bulunması gerekir." biçimindedir.

- **Cronbach Alfa katsayısı**, bireysel puanların k soru içeren bir ölçekte sorulara verilen cevapların toplanması ile bulunduğu durumlarda soruların birbirleri ile benzerliğini, yakınlığını ortaya koyan bir katsayıdır.
- Alfa katsayısı, ölçekte yer alan k sorunun türdeş bir yapıyı açıklamak ya da sorgulamak üzere bir bütün oluşturup oluşturmadıklarını sorgulamayı sağlar.

Cronbach Alfa deęişik biçimlerde hesaplanabilir:

1. Maddelerin varyans deęerlerinden hareketle hesaplama

$$\alpha = \frac{k}{k-1} \left[\frac{\sigma_t^2 - \sum \sigma_i^2}{\sigma_t^2} \right]$$

σ_t^2 = Toplam Sütunun Varyansı

σ_i^2 = Deęişkenlerin her birinin varyansı

Tablo 2. Alfa Katsayısının Hesaplanması

Kişiler	Madde 1	Madde 2	Madde 3	Toplam
1	2	3	3	8
2	3	4	3	10
3	3	2	2	7
4	2	4	4	10
5	4	4	1	9
<i>O ortalama</i>	2,8	3,4	2,6	8,8
<i>SS Standart Sapma</i>	0,84	0,89	1,14	1,30
<i>S²Varyans</i>	0,7	0,8	1,3	1,7

$$\alpha = \frac{3}{2} \left[\frac{1,7 - 2,8}{1,7} \right] = -0,97$$

2. Maddelerin kovaryans değerlerinden hareketle hesaplama

$$\alpha = \frac{\frac{k \times Ort(kov)}{Ort(Var)}}{1 + \frac{(k-1) \times Ort(kov)}{Ort(Var)}}$$

k: ölçekte yer alan madde sayısı

Ort(kov): Ortalama kovaryans

Ort(Var): Ortalama varyans

Standardized Cronbach's Alpha

Korelasyon matrisi verilerinden hareketle hesaplama:

$$\alpha = \frac{k\bar{r}}{1 + (k-1)\bar{r}}$$

k: ölçekte yer alan madde sayısı

\bar{r} : maddeler arasındaki ortalama korelasyon katsayısı

Güvenilirlik Analizi Büyüklüğünü Etkileyen Faktörler

- Güvenirlik analizinde k soru ve n birim vardır. Güvenirlik analizleri yapılırken k ve n ye ait iki temel koşulun dikkate alınması gerekir;

i) $K > 20$ olmalıdır. Yani ölçme araçları , bireysel özellikleri (demografik, sosyo ekonomik) sorgulayan sorular dışında, çok sayıda ve birbirleri ile ilişkili soru içermelidir.

ii) $n > 50$ olmalıdır. Yani ölçme araçları çok sayıda rasgele seçilen deneklere uygulanmalıdır.

- Ölçekteki gözlem sayısı, soru (madde) sayısından fazla olmalıdır.
- Madde sayısı ve n arttıkça testin güvenilirliği artabilir.
- Bireyler test edilmeye hazır olmalıdırlar. Bunun için bireyler testten önce test hakkında bilgilendirilmelidirler.
- Bireylerin yaş, cinsiyet, eğitim ve deneyim düzeyleri güvenilirlik katsayısının büyüklüğünün etkileyebilir.
- Testi yapan kişinin özellikleri önemlidir. Testi yapan kişi deneyimli ve objektif ise güvenilirlikte yüksek olacaktır.
- Test özellikleri güvenilirlik katsayısının büyüklüğünü etkiler. Örneğin fiziksel uygunluk testleri spor beceri testlerinden daha güvenilirdir. Çünkü fiziksel uygunluk testleri daha objektif iken, spor beceri testleri daha subjektif testlerdir.

Güvenirlilik analizinin temel varsayımları:

- Her soru toplam skorun bir doğrusal bileşeni olmalıdır.
- Ölçekte toplanabilirlik özelliğinin bulunması gerekir.

Cronbach Alfa Katsayısı

- Sorular arasında negatif korelasyon varsa Cronbach Alfa Katsayısı da negatif çıkar.
- Alfa'nın negatif çıkması, güvenirlik modelinin bozulmasına neden olur.
- Çünkü bu durumda ölçeğin toplanabilirlik varsayımı bozulmuş ve ölçek toplanabilir ölçek olmaktan çıkmış olur.

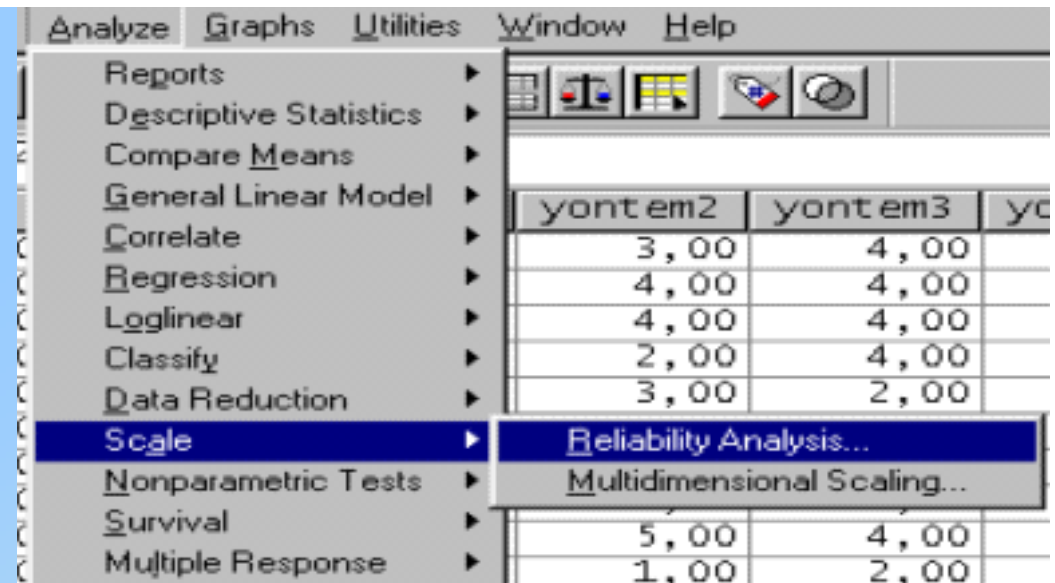
Alfa'nın negatif çıkmasının nedenleri:

- Bireysel maddelere ait varyansların toplamının toplam puanlara ait varyanstan büyük olması
- Araştırmada ciddi ölçüm hatalarının olması
- Ters yönlü maddelerin kodlanmasına dikkat edilmemesi
- Örneklem hacmi küçük veya ölçekteki madde sayısının az olması
- Maddeler arası kovaryans değerlerinin ortalamasının negatif olması
- Testin birden fazla faktör içermesi ve bu faktörlerin birbiriyle negatif yönde ilişkili olması

- **Alfa katsayısının bulunabileceği aralıklar ve buna bağlı olarak da ölçeğin güvenilirlik durumu aşağıda verilmiştir**
- **$0,00 \leq \alpha < 0,40$ ise ölçek güvenilir değildir,**
- **$0,40 \leq \alpha < 0,60$ ise ölçek düşük güvenilirliktedir,**
- **$0,60 \leq \alpha < 0,80$ ise ölçek oldukça güvenilirdir,**
- **$0,80 \leq \alpha < 1,00$ ise ölçek yüksek derecede güvenilir bir ölçektir.**

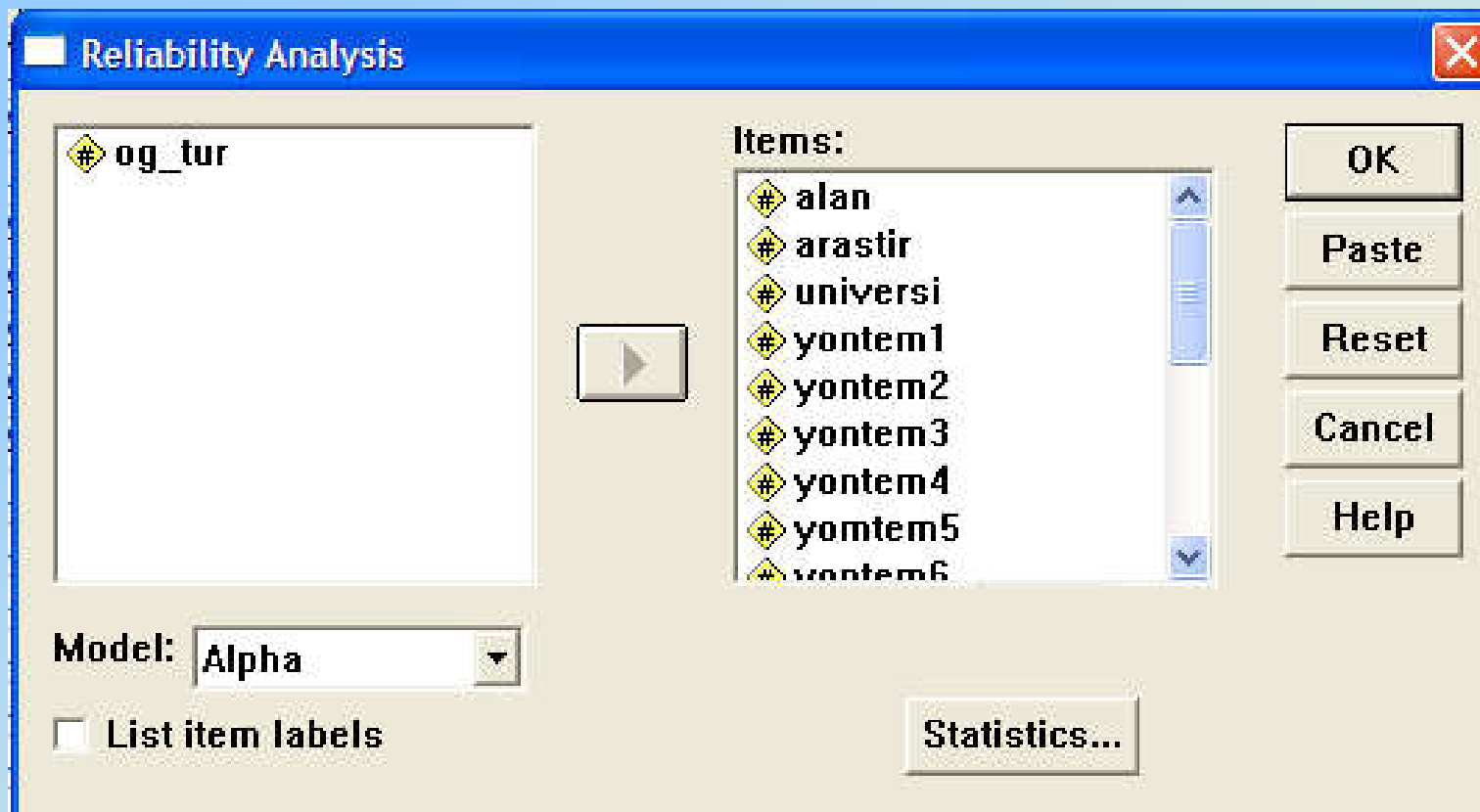
- ❖ Bir ölçekteki sorulara verilen cevapların birey ve sorulara göre önemliliğini belirlemek için iki yönlü varyans analizi yapılır.
- ❖ Sorular arasındaki benzerlikler F testi ile analiz edilir.
- ❖ Sorular sıralama puanları olarak alınmış ise, birey ve soru farklılıklarının analizi için Friedman Ki-Kare testi kullanılır.
- ❖ Sorulara verilen cevaplar 0, 1 biçiminde ikili tipte iseler önemlilik değerlendirilmesi Cochran Ki-Kare testi yardımı ile yapılır.
- ❖ Ölçeğin toplanabilir ölçek tipinde hazırlanıp hazırlanmadığını test etmek için Tukey eklenebilirlik testi (Tukey's Additivity Test) kullanılır.
- ❖ Soru ortalamalarının birbirlerine eşit olup olmadıkları ise Hotelling T² istatistiği ile test edilir.

SPSS'TE CRONBACH ALFA HESABI



The image shows the SPSS 'Analyze' menu with 'Scale' selected. The 'Reliability Analysis...' option is highlighted. A data table is visible in the background.

	yontem2	yontem3	yc
	3,00	4,00	
	4,00	4,00	
	4,00	4,00	
	2,00	4,00	
	3,00	2,00	
	5,00	4,00	
	1,00	2,00	



The 'Reliability Analysis' dialog box is shown. The 'Items:' list contains: alan, arastir, universi, yontem1, yontem2, yontem3, yontem4, yontem5, yontem6. The 'Model:' is set to 'Alpha'. The 'List item labels' checkbox is unchecked. The 'Statistics...' button is visible.

Model: Alpha

List item labels

Statistics...

Reliability Analysis: Statistics

Descriptives for

- Item
- Scale
- Scale if item deleted

Inter-Item

- Correlations
- Covariances

Summaries

- Means
- Variances
- Covariances
- Correlations

ANOVA Table

- None
- F test
- Friedman chi-square
- Cochran chi-square

Hotelling's T-square

Tukey's test of additivity

Intraclass correlation coefficient

Model: Two-Way Random

Type: Absolute Agreement

Confidence: 95 %

Test value: 0

Continue

Cancel

Help

Item : soru

Scale : ölçek

F test: Ölçeği oluşturan sorular arasındaki benzerlik analizini yapar.

Friedman chi-square: Ölçekteki sorulara verilen cevaplar sıralama puanları olarak verilmişse birey ve soru farklılıklarının analizi yapılır.

Cochran chi-square: Ölçekteki sorulara verilen cevaplar 0 yada 1 biçiminde iki değerli ise birey ve sorulara göre önemlilik analizi yapılır.

Tukey's test of additivity : Ölçekte yer alan soruların bir toplamsal ölçek oluşturacak biçimde hazırlanıp hazırlanmadığını, soruların skorlarının toplanabilir olup olmadığını test eder.

Interclass correlation coefficient: Değerlendiriciler arasındaki tutarlılığı ve uyumları test eder ve hesaplama modeli için iki yönlü rasgele (Two-way Random) model seçilir. Birimler içi değerlerin tutarlılığını ve mutlak uyumlarını ölçen katsayıları hesaplar.

Hotelling's T-square : Soru ortalamalarının birbirine eşit olup olmadığını test eder. Soru ortalamalarının birbirine eşit olup olmaması kavramı; soruların denekler tarafından aynı yaklaşım ile algılanıp algılanmadığını yani soruların zorluk derecelerinin birbirine eşit olup olmadığıdır. Ölçek değerlendirmede soruların %75'i orta, %12,5'i kolay ve %12,5'i de zor nitelikte olmalıdır. sıfır olarak seçilir.

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
,882	,886	10

Alfa



Item Statistics

	Mean	Std. Deviation	N
yontem1	3,24	,973	348
yontem2	3,42	,979	348
yontem3	3,35	,944	348
yontem4	3,29	1,128	348
yomtem5	3,04	1,039	348
yontem6	3,11	1,037	348
yontem7	3,20	1,086	348
yontem8	2,69	1,085	348
yontem9	2,76	1,316	348
yontem10	3,31	1,262	348

Inter-Item Correlation Matrix

	yontem1	yontem2	yontem3	yontem4	yontem5	yontem6	yontem7	yontem8	yontem9	yontem10
yontem1	1,000	,559	,521	,445	,426	,447	,455	,345	,203	,294
yontem2	,559	1,000	,568	,554	,407	,449	,508	,402	,246	,305
yontem3	,521	,568	1,000	,560	,517	,456	,511	,416	,322	,340
yontem4	,445	,554	,560	1,000	,518	,560	,502	,396	,292	,447
yontem5	,426	,407	,517	,518	1,000	,573	,529	,367	,400	,405
yontem6	,447	,449	,456	,560	,573	1,000	,609	,393	,351	,327
yontem7	,455	,508	,511	,502	,529	,609	1,000	,571	,436	,436
yontem8	,345	,402	,416	,396	,367	,393	,571	1,000	,348	,500
yontem9	,203	,246	,322	,292	,400	,351	,436	,348	1,000	,487
yontem10	,294	,305	,340	,447	,405	,327	,436	,500	,487	1,000

Summary Item Statistics

	Mean	Minimum	Maximum	Range	Maximum / Minimum	Variance	N of Items
Item Means	3,141	2,690	3,417	,727	1,270	,060	10
Item Variances	1,190	,891	1,731	,840	1,942	,077	10
Inter-Item Covariances	,508	,260	,809	,549	3,108	,012	10
Inter-Item Correlations	,438	,203	,609	,406	2,996	,009	10

Ölçekte yer alan soruların genel ortalaması (item means) 3,141'dir. Ortalanmaların değişim aralığı (range) 0,727 olmaktadır. Soru-Bütün (Inter-item) korelasyonlara bakıldığında 0,203 ile 0,609 arasında değişim gösterdiği görülür. Soru ile bütün arasındaki korelasyonların negatif olmaması gerekir, hatta 0,2'den büyük olması beklenir.

Scale Statistics

Mean	Variance	Std. Deviation	N of Items
31,41	57,661	7,593	10

10 sorudan oluşan ölçek ortalaması (toplam puan) 31,41 ve standart sapma 7,593 bulunmuştur.

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
yontem1	28,18	49,017	,565	,412	,874
yontem2	28,00	48,303	,617	,488	,870
yontem3	28,06	48,152	,658	,495	,868
yontem4	28,13	46,134	,669	,518	,866
yomtem5	28,37	47,254	,653	,469	,867
yontem6	28,30	47,259	,654	,513	,867
yontem7	28,22	45,812	,726	,568	,862
yontem8	28,72	47,659	,589	,420	,872
yontem9	28,65	47,197	,483	,329	,883
yontem10	28,10	46,408	,562	,420	,875

Scale Mean/Variance if Item Deleted: Ölçekten bir soru çıkartıldığında geriye kalan soruların ölçek ortalaması/varyansı

Corrected Item-Total Correlation : Ölçekten çıkarılan ilgili soru ile ölçekteki toplam puan arasındaki korelasyon (Point Biserial Korelasyon). Bu değerler negatif olmamalı ve 0,20'den büyük olmalıdır.

Squared Multiple Correlation : Ölçekten ilgili soru çıkartıldıktan sonra kalan kısma ait çoklu korelasyon katsayısı

Cronbach's Alpha if Item Deleted: Bir sorunun ölçekten çıkarılması için; soru silinirse Alfa katsayısında (Alpha if Item Deleted) değişimine bakılır. Bu sütunda en yüksek değerli soru ölçekten çıkarıldığında, eğer alfa katsayısı yükseliyorsa o soru güvenilirliği azaltan bir sorudur ve ölçekten çıkarılması gerekir.

Madde Silindiğinde Ortalama-Varyanslardaki Değişim (Scale Means-Variance if item deleted)

Eğer madde ölçekten çıkarılırsa bireylerin madde ortalamalar ve varyansda büyük değişiklik olmaması istenir.

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
yontem1	28,18	49,017	,565	,412	,874
yontem2	28,00	48,303	,617	,488	,870
yontem3	28,06	48,152	,658	,495	,868
yontem4	28,13	46,134	,669	,518	,866
yomtem5	28,37	47,254	,653	,469	,867
yontem6	28,30	47,259	,654	,513	,867
yontem7	28,22	45,812	,726	,568	,862
yontem8	28,72	47,659	,589	,420	,872
yontem9	28,65	47,197	,483	,329	,883
yontem10	28,10	46,408	,562	,420	,875

Madde Bütün Korelasyon Katsayısı (Corrected item-total correlation)

Ölçekteki herbir maddenin o madde dışındaki maddelerin toplamıyla-bütünüyle olan korelasyon katsayısıdır. Bu Point Biserial Korelasyondur.

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
yontem1	28,18	49,017	,565	,412	,874
yontem2	28,00	48,303	,617	,488	,870
yontem3	28,06	48,152	,658	,495	,868
yontem4	28,13	46,134	,669	,518	,866
yomtem5	28,37	47,254	,653	,469	,867
yontem6	28,30	47,259	,654	,513	,867
yontem7	28,22	45,812	,726	,568	,862
yontem8	28,72	47,659	,589	,420	,872
yontem9	28,65	47,197	,483	,329	,883
yontem10	28,10	46,408	,562	,420	,875

Çoklu Açıklayıcılık Katsayısı (R^2) (Squared Multiple Coreelation)

Bir madde ile geriye kalan diğer maddeler arasındaki ilişki R^2 ile bulunabilir. Bunun için ilgili madde bağımlı, diğer maddeler bağımsız değişken alınarak çoklu doğrusal regresyon analizi yapılarak R^2 bulunabilir. Bir maddenin R^2 değerinin küçük çıkması, o maddenin diğer maddeler tarafından yeterince açıklanmadığını gösterir.

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
yontem1	28,18	49,017	,565	,412	,874
yontem2	28,00	48,303	,617	,488	,870
yontem3	28,06	48,152	,658	,495	,868
yontem4	28,13	46,134	,669	,518	,866
yontem5	28,37	47,254	,653	,469	,867
yontem6	28,30	47,259	,654	,513	,867
yontem7	28,22	45,812	,726	,568	,862
yontem8	28,72	47,659	,589	,420	,872
yontem9	28,65	47,197	,483	,329	,883
yontem10	28,10	46,408	,562	,420	,875

Madde Silinirse Güvenilirlik Katsayısı (Cronbach's alpha if item deleted)

Ele alınan madde ölçekten çıkarıldığında güvenilirlik katsayısının değişimini incelemek amacıyla yararlanılan bir yaklaşımdır. Bir madde ölçekten çıkartıldığında alfa katsayısı ölçeğin tümü için hesaplanan alfa katsayısına göre artış gösteriyorsa, o madde ölçeğin güvenilirliğini azaltan bir maddedir. Eğer bir madde ölçekten çıkarıldığında alfa katsayısı ölçeğin tümü için hesaplanan alfa katsayısından küçük çıkıyorsa o madde ölçekte yer almalıdır. Böyle bir madde olup olmadığına bakmak için **Cronbach's alpha if item deleted** değerlerinden en büyük olana bakılır.

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
yontem1	28,18	49,017	,565	,412	,874
yontem2	28,00	48,303	,617	,488	,870
yontem3	28,06	48,152	,658	,495	,868
yontem4	28,13	46,134	,669	,518	,866
yontem5	28,37	47,254	,653	,469	,867
yontem6	28,30	47,259	,654	,513	,867
yontem7	28,22	45,812	,726	,568	,862
yontem8	28,72	47,659	,589	,420	,872
yontem9	28,65	47,197	,483	,329	,883
yontem10	28,10	46,408	,562	,420	,875

ANOVA with Tukey's Test for Nonadditivity

			Sum of Squares	df	Mean Square	F	Sig
Between People			2000,841	347	5,766		
Within People	Between Items		187,591	9	20,843	30,560	,000
	Residual	Nonadditivity	,329 ^a	1	,329	,482	,488
		Balance	2129,681	3122	,682		
		Total	2130,009	3123	,682		
Total			2317,600	3132	,740		
Total			4318,441	3479	1,241		

Grand Mean = 3,14

a. Tukey's estimate of power to which observations must be raised to achieve additivity = 1,173.

Yapılan varyans analizi sonucuna göre ölçek toplanabilir özelliktedir (Nonadditivity: $F=0,482$ $P=0,488>0,05$). Ölçümler arası değişime bakıldığında önemli bir farklılık görülmektedir (Between Measures, $F=30,56$ $P=0,00<0,05$).

Yani 10 soruluk bu ölçek toplanabilir özelliktedir ancak ölçümler arasında farklılıklar vardır.

Hotelling's T-Squared Test

Hotelling's T-Squared	F	df1	df2	Sig
282,420	30,657	9	339	,000

Soru ortalamalarının eşitliğini test eden Hotelling T^2 testi sonucunda, soru ortalamaların farklı olduğu görülmektedir (Hotelling's T-Squared=282,42 , $P=0,000<0,05$). Yani en az iki ortalama arasında farklılık vardır. Bu farklılıkların hangi sorulardan kaynaklandığı araştırılmalıdır.

Intraclass Correlation Coefficient

	Intraclass Correlation ^a	95% Confidence Interval		F Test with True Value 0			
		Lower Bound	Upper Bound	Value	df1	df2	Sig
Single Measures	,427 ^b	,385	,472	8,454	347	3123	,000
Average Measures	,882	,862	,899	8,454	347	3123	,000

Two-way random effects model where both people effects and measures effects are random.

- Type C intraclass correlation coefficients using a consistency definition-the between-measure variance is excluded from the denominator variance.
- The estimator is the same, whether the interaction effect is present or not.

Intraclass Correlation (Küme-sınıf içi korelasyon analizi): Gözlemci (hakem) değerlendirmelerinin güvenilirliğini belirlemede kullanılır. Bir kişiye ait gözlem değerlerinin diğer kişilerin gözlem değerlerine ne ölçüde benzer olduğunu gösterir.

Single Measures Intraclass Correlation : Tek ölçüm küme (sınıf) içi korelasyon katsayısı

Average Measures Intraclass Correlation : Rasgele seçilen kişiler çok sayıda farklı gözlemciler(hakemler) tarafından değerlendirilir ve bu değerlendirmelerin ortalaması alınarak korelasyon değeri hesaplanır.

F testi ile ölçeği oluşturan sorular arasında benzerlik olup olmadığı test edilmiş ve benzerlik olmadığı görülmüştür (F=8,454, p=0,000<0,05).

Soru (Madde) ile Bütün Arasındaki Korelasyonlar Yöntemi (Item Total Correlation)

Bir madde ile bu madde dışında diğer maddelerin toplamından oluşan bütün (total) arasındaki korelasyon hesaplanmasına dayanmaktadır. Eğer madde-bütün korelasyon katsayısı düşük ise, o sorunun ölçeğe katkısının düşük olduğu anlamına gelir ve o maddenin ölçekten çıkarılması gerekir.

Madde-bütün korelasyon katsayısının eksi işaretli olmaması ve $>0,25$ olması istenir.

Madde-bütün korelasyon katsayısı maddenin ayırdediciliği ya da maddenin geçerliliği konusunda bilgi verir.

Madde-bütün korelasyon katsayısı hesaplanırken dikkatli olunmalıdır. Özellikle 5'li ve 7'li Likert tipi ölçeklerde seçenekler arasındaki farkın eşit ve bütün puanın sürekli olduğu varsayımı altında, madde-bütün korelasyonları genellikle Pearson korelasyon katsayısı ile hesaplanır. Eğer madde seçenekleri iki kategorili (başarılı-başarısız, evet-hayır, var-yok, katılıyorum-katılmıyorum vb.) ise madde-bütün korelasyonu nokta-iki serili korelasyon (r_{pb}) ile hesaplanmalıdır.

Nokta İki Serili Korelasyon Katsayısı (point biserial correlation coefficient)

İlişki derecesi incelenecek değişkenlerden biri iki kategorili ve doğal şekilde (erkek-bayan, başarılı-başarısız gibi), diğeri sürekli sayısal veri türünde ise Nokta İki Serili Korelasyon katsayısı kullanılır.

$$r_{pb} = \frac{\bar{Y}_1 - \bar{Y}_0}{s_Y} \sqrt{p_0 \times p_1} = \frac{\bar{Y}_1 - \bar{Y}_0}{s_Y} \sqrt{\frac{n_0}{n} \times \frac{n_1}{n}}$$

Örnek. 23 erkek ve 39 bayan sporcunun oluşturduğu bir örnekleme sporcuların boy uzunlukları ve sayıları aşağıdaki gibi bulunmuştur. Cinsiyet ile boy uzunlukları arasında bir ilişki var mıdır.

Boy Uz (cm)	Erkek (p1)	Bayan (p0)	Toplam
160-170	0	7	7
170-180	1	19	20
180-190	4	13	17
190-200	10	0	10
200-210	8	0	8
Toplam	23	39	62

$$p_1 = 23 / 62 = 0.37 \quad p_0 = 39 / 62 = 0.63$$

$$\bar{Y}_1 = \frac{\sum m_i \times f_i}{\sum f_i} = \frac{165 \times 0 + 175 \times 1 + 185 \times 4 + 195 \times 10 + 205 \times 8}{23} = 195.87$$

$$\bar{Y}_0 = \frac{165 \times 7 + 175 \times 19 + 185 \times 13 + 195 \times 0 + 205 \times 0}{39} = 176.54$$

$$\bar{Y} = \frac{165 \times 7 + 175 \times 20 + 185 \times 17 + 195 \times 10 + 205 \times 8}{62} = 183.71$$

$$s_Y = \sqrt{\frac{\sum f_i (m_i - \bar{Y})^2}{\sum f_i}} = 11.98$$

$$r_{pb} = \frac{\bar{Y}_1 - \bar{Y}_0}{s_Y} \sqrt{p_0 \times p_1} = \frac{195.87 - 176.54}{11.98} \sqrt{0.37 \times 0.63} = 0.78$$

Örnek. Bir çalışmada A ve B gibi iki gruba ilişkin ölçümler aşağıdaki gibidir. Nokta iki serili ilişki katsayısını bulunuz.

No	1	2	3	4	5	6	7	8	9	10
Grup	0	1	1	0	1	0	1	0	1	0
Ölçüm	3	6	6	4	4	3	6	2	5	3

No	11	12	13	14	15	16	17	18	19	20
Grup	1	1	0	1	0	1	0	0	1	0
Ölçüm	7	4	3	9	1	5	4	0	5	2

$$p_1 = 10 / 20 = 0.5 \quad p_0 = 10 / 20 = 0.5$$

$$\bar{Y}_0 = 2.5 \quad \bar{Y}_1 = 5.6$$


$$s_Y = \sqrt{\frac{\sum (Y_i - \bar{Y})^2}{n-1}} = 1.96$$

$$r_{pb} = \frac{\bar{Y}_1 - \bar{Y}_0}{s_Y} \sqrt{p_0 \times p_1} = \frac{5.6 - 2.5}{1.96} \sqrt{0.5 \times 0.5} = 0.79$$

Soru. 10 maddeden oluşan bir test çalışmasından aşağıdaki sonuçlar elde edilmiştir. Maddelere verilen cevaplar (0-hayır, 1-evet) şeklindedir. Point Biserial korelasyon katsayılarını bulunuz.

Öğr./Madde	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	Toplam
A	1	1	1	1	1	1	1	1	0	1	9
B	1	1	1	1	1	1	1	0	1	0	8
C	1	1	1	1	1	1	0	1	0	0	7
D	1	1	1	1	1	0	1	0	1	0	7
E	1	1	1	1	1	1	0	1	0	0	7
F	1	1	1	0	1	0	0	0	0	0	4
G	1	1	0	1	0	1	0	0	0	0	4
H	1	0	1	0	1	0	0	0	0	0	3
I	0	1	1	0	0	0	0	0	0	0	2
Toplam	8	8	8	6	7	5	3	3	2	1	

	A	B	C	D	E	F	G	H	I	J	K	L
1	Öğr./Madde	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	Toplam
2	A	1	1	1	1	1	1	1	1	0	1	9
3	B	1	1	1	1	1	1	1	0	1	0	8
4	C	1	1	1	1	1	1	0	1	0	0	7
5	D	1	1	1	1	1	0	1	0	1	0	7
6	E	1	1	1	1	1	1	0	1	0	0	7
7	F	1	1	1	0	1	0	0	0	0	0	4
8	G	1	1	0	1	0	1	0	0	0	0	4
9	H	1	0	1	0	1	0	0	0	0	0	3
10	I	0	1	1	0	0	0	0	0	0	0	2
11	Toplam	8	8	8	6	7	5	3	3	2	1	
12												
13	Öğr./Madde	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	
14	A	8	8	8	8	8	8	8	8	9	8	
15	B	7	7	7	7	7	7	7	8	7	8	
16	C	6	6	6	6	6	6	7	6	7	7	
17	D	6	6	6	6	6	7	6	7	6	7	
18	E	6	6	6	6	6	6	7	6	7	7	
19	F	3	3	3	4	3	4	4	4	4	4	
20	G	3	3	4	3	4	3	4	4	4	4	
21	H	2	3	2	3	2	3	3	3	3	3	
22	I	2	1	1	2	2	2	2	2	2	2	

 =L2-K2

=KORELASYON(B2:B10;B14:B22)



Point Biserial Kor.	0.457	0.286	0.122	0.728	0.486	0.491	0.589	0.459	0.260	0.399
----------------------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------

Point Biserial Korelasyon

Item-Total Statistics



	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
M1	4,7778	5,194	,457	,746
M2	4,7778	5,444	,286	,763
M3	4,7778	5,694	,122	,779
M4	5,0000	4,250	,728	,699
M5	4,8889	4,861	,486	,739
M6	5,1111	4,611	,491	,739
M7	5,3333	4,500	,589	,722
M8	5,3333	4,750	,459	,743
M9	5,4444	5,278	,260	,770
M10	5,5556	5,278	,399	,752

Örnek. 8 maddeden oluşan bir test çalışmasından aşağıdaki sonuçlar elde edilmiştir. Maddelere verilen cevaplar (0-hayır, 1-evet) şeklindedir. Madde1 ile bütün arasındaki korelasyon katsayısını bulunuz.

	M1	M2	M3	M4	M5	M6	M7	M8	Bütün
1	1	0	1	1	0	1	1	1	5
2	1	1	1	1	0	1	0	1	5
3	1	0	0	1	0	0	1	1	3
4	1	0	0	1	1	0	0	1	3
5	0	0	1	1	0	1	1	1	5
6	0	1	0	0	0	0	1	0	2
7	1	1	1	0	0	1	1	1	5
8	0	0	0	1	1	1	0	0	3
9	1	1	1	0	0	1	1	1	5
10	0	0	1	1	0	1	0	0	3

$$r_{pb} = \frac{\bar{Y}_1 - \bar{Y}_0}{S_Y} \sqrt{p_0 \times p_1} = \frac{5.33 - 3.25}{1.5} \sqrt{0.6 \times 0.4} = 0.68$$

İki Serili Korelasyon Katsayısı (rb)

İlişki derecesi incelenecek değişkenlerden biri iki veya daha çok kategorili nitelik, diğeri sürekli sayısal veri türünde ise Eta istatistiği kullanılır. Eta istatistiği ortalama (iki veya daha çok) arasındaki farkın önemlilik testi sonucunda ortalamalardaki değişime ilişkin bir ilişki katsayısıdır. Eğer nitel değişken iki kategorili ise Eta istatistiği **İki Serili veya Nokta İki Serili Korelasyon katsayısı** adını alır. **Nitel değişken doğal 2 kategorili (erkek-bayan, başarılı-başarısız gibi) ise nokta iki serili, doğal olmayan iki kategorili ise (<50yaş, >=50 yaş gibi) iki serili korelasyon katsayısı kullanılır.**

$$p=n_0/n$$

$$, \quad q=n_1/n$$

$$r_b = \frac{\bar{Y}_1 - \bar{Y}_0}{s_Y} \times \frac{p \cdot q}{H}$$

$$P(Z_1 < Z < \infty) = P$$

H yukarıdaki eşitliği sağlayan Z1 değerine karşı gelen değerdir ve Z tablosundan bulunur. P<0.5 ise Z1 artı, p>0.5 ise eksi işaret verilir.

Elde edilen korelasyon katsayısının işareti 0 ya da 1 olarak alınan gruba göre değişir. Bunun için mutlak değere göre yorum yapılır.

Örnek. 10 kişinin yaşları bir testten aldığı skor puanları bir araştırmadan elde edilmiştir. Kişilerin yaşları >70 ve <=69 olmak üzere iki gruba ayrılmış ve aşağıdaki gibi bulunmuştur. İlişki katsayısını bulunuz.

>70	<=69
6	10
7	7
4	8
5	3
2	9

Burada nitel değişken olan yaş grubu doğal olmayan nitel değişkendir. Bu yüzden puan ile görüşme sonucu arasındaki ilişki katsayısı iki serili korelasyon katsayısı ile bulunur.

$$p=5/10=0,5$$

$$q=5/10=0,5$$

$$P(Z_1 < Z < \infty) = 0.5$$

Z tablosundan 0.5'e en yakın değere karşılık gelen Z değeri 0.00 dir. Bu Z değerine karşı gelen Ordinat tablosu değeri 0.3989 bulunur.

$$f(Z) = \frac{1}{\sqrt{2\pi}} \exp\{-Z^2 / 2\} = \frac{1}{\sqrt{2 \times 3.1415}} \exp\{-0.00^2 / 2\} = 0.3989$$

$$r_b = \frac{\bar{Y}_1 - \bar{Y}_0}{s_Y} \times \frac{p.q}{H} = \frac{7.4 - 4.8}{2.6} \times \frac{0.5 \times 0.5}{0.3989} = 0.63$$

Örnek : 10 soruluk bir anket, eğitim teknolojisini kullanma yeterliliğini belirlemeyi amaçlamaktadır. Bu araç 20 kişilik rastgele seçilen deneklere uygulanmıştır. Her bir soru 1 ile 7 arasında kodlanabilen seçenekler içermektedir (örnek varsayımsaldır).

- Anketin güvenilirliği nedir?
- Eğitim teknolojisini kullanma yeterliliğini ölçen ankette yer alan 10 sorunun tümü bu fenomeni değerlendirme yeteneğine sahip midir?
- Gereksiz sorular var mıdır?
- Bu sorular araçtan çıkarılırsa aracın fenomeni ölçme yeteneği ne kadar artar? Test ederek tartışınız.
- Verilerin sembolik olarak analizini göstermek için 10 soru ve 20 birim için bir örnek türetilmiş ve bu örnek veriler aşağıdaki tabloda verilmiştir.

Tablo 1 —10 soruluk ölçek için 20 bireyin cevapları

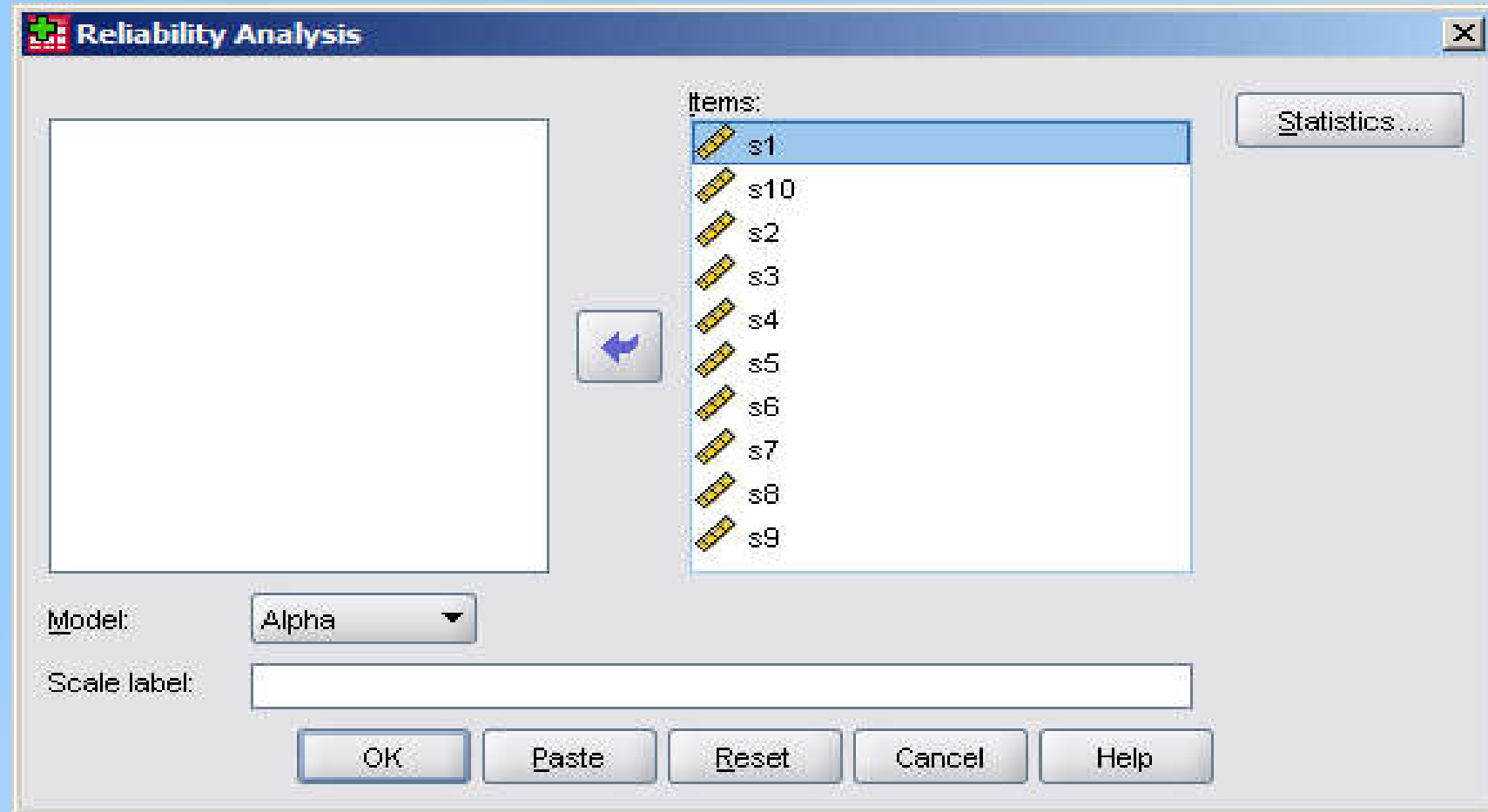
Birim	S O R U L A R									
	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10
1	4	4	3	5	4	3	2	3	5	4
2	3	3	5	3	4	5	4	3	2	4
3	5	5	4	5	4	3	6	5	4	7
4	4	4	3	4	6	6	2	2	3	6
5	2	6	5	3	5	5	4	1	5	4
6	3	7	4	6	4	4	3	4	4	3
7	6	5	3	5	7	3	5	3	6	5
8	5	4	6	7	6	7	4	5	5	6
9	4	3	5	4	3	2	6	4	5	4
10	3	4	7	3	2	4	4	6	5	7
11	4	5	4	5	4	3	3	7	4	5
12	5	4	2	4	5	5	4	1	3	4
13	6	3	3	6	6	4	3	2	2	3
14	5	2	5	3	5	4	3	3	3	5
15	4	1	4	3	5	6	6	4	4	4
16	3	6	1	6	4	5	4	5	6	6
17	4	5	2	5	3	3	7	7	4	5
18	5	4	6	7	5	4	6	5	5	4
19	6	3	5	4	4	5	5	4	6	5
20	4	7	4	2	3	2	4	3	2	3



1 : s1 4 Visible: 10 of 10 Variables

	s1	s10	s2	s3	s4	s5	s6	s7	s8	s9
1	4,00	4,00	4,00	3,00	5,00	4,00	3,00	2,00	3,00	5,00
2	3,00	4,00	3,00	5,00	3,00	4,00	5,00	4,00	3,00	2,00
3	5,00	7,00	5,00	4,00	5,00	4,00	3,00	6,00	5,00	4,00
4	4,00	6,00	4,00	3,00	4,00	6,00	6,00	2,00	2,00	3,00
5	2,00	4,00	6,00	5,00	3,00	5,00	5,00	4,00	1,00	5,00
6	3,00	3,00	7,00	4,00	6,00	4,00	4,00	3,00	4,00	4,00
7	6,00	5,00	5,00	3,00	5,00	7,00	3,00	5,00	3,00	6,00
8	5,00	6,00	4,00	6,00	7,00	6,00	7,00	4,00	5,00	5,00
9	4,00	4,00	3,00	5,00	4,00	3,00	2,00	6,00	4,00	5,00
10	3,00	7,00	4,00	7,00	3,00	2,00	4,00	4,00	6,00	5,00
11	4,00	5,00	5,00	4,00	5,00	4,00	3,00	3,00	7,00	4,00
12	5,00	4,00	4,00	2,00	4,00	5,00	5,00	4,00	1,00	3,00
13	6,00	3,00	3,00	3,00	6,00	6,00	4,00	3,00	2,00	2,00
14	5,00	5,00	2,00	5,00	3,00	5,00	4,00	3,00	3,00	3,00
15	4,00	4,00	1,00	4,00	3,00	5,00	6,00	6,00	4,00	4,00
16	3,00	6,00	6,00	1,00	6,00	4,00	5,00	4,00	5,00	6,00
17	4,00	5,00	5,00	2,00	5,00	3,00	3,00	7,00	7,00	4,00
18	5,00	4,00	4,00	6,00	7,00	5,00	4,00	6,00	5,00	5,00
19	6,00	5,00	3,00	5,00	4,00	4,00	5,00	5,00	4,00	6,00
20	4,00	3,00	7,00	4,00	2,00	3,00	2,00	4,00	3,00	2,00

Alpha Yöntemi





Reliability Analysis: Statistics



Descriptives for

- Item
- Scale
- Scale if item deleted

Inter-Item

- Correlations
- Covariances

Summaries

- Means
- Variances
- Covariances
- Correlations

ANOVA Table

- None
- F test
- Friedman chi-square
- Cochran chi-square

- Hotelling's T-square
- Intraclass correlation coefficient

- Tukey's test of additivity

Model:

Type:

Confidence interval: %

Test value:

Continue

Cancel

Help

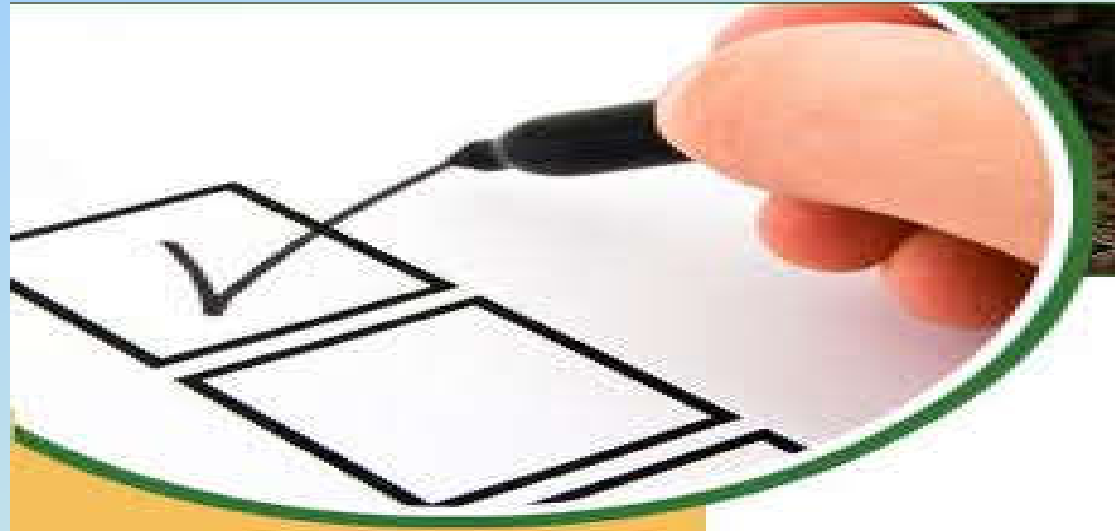
- SPSS hem sorular için hem de ölçek için istatistikler hesaplamaktadır. Inter-Item bölümünde korelasyonları ve kovaryansları görüntülemek için seçenekler işaretlenir. Model uyumu ve soruların uyumluluğu için F testi işaretlenir



Ölçekte yer alan soruların bir toplamsal ölçek oluşturacak biçimde hazırlanıp hazırlanmadığı *Tukey Toplanabilirlik Testi* (***Tukey's test of additivity***) seçeneği ile toplanabilirlik test edilir.



- Soru ortalamalarının birbirlerine eşit olup olmadığı ise Hotelling T^2 seçeneği ile değerlendirilir. Soru ortalamalarının birbirlerine eşit olup olmaması kavramı; soruların denekler tarafından aynı yaklaşım ile algılanıp algılanmadığını, soruların zorluk derecelerinin birbirine eşit olup olmadığını belirtir.



- Diğer bir anlatım ile soruların ölçme yeteneklerinin birbirlerine yakın ve normal dağılım formunda bir yapıda olup olmadıklarını değerlendirmeyi amaçlar. Sorular hedef toplum tarafından benzer sonuçlar verecek biçimde hazırlanmalıdır.
- Bir ölçekte soruların yaklaşık %75'i orta zorlukta %12.5'i kolay ve %12.5'i de zor nitelikte olmalıdır.



Çıktı.1

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
,393	,408	10

Item Statistics

	Mean	Std. Deviation	N
s1	4,2500	1,11803	20
s10	4,7000	1,21828	20
s2	4,2500	1,55174	20
s3	4,0500	1,50350	20
s4	4,5000	1,43270	20
s5	4,4500	1,23438	20
s6	4,1500	1,34849	20
s7	4,2500	1,40955	20
s8	3,8500	1,72520	20
s9	4,1500	1,30888	20

Inter-Item Correlation Matrix

	s1	s10	s2	s3	s4	s5	s6	s7	s8	s9
s1	1,000	,019	-,372	-,070	,312	,486	-,061	,159	-,061	,009
s10	,019	1,000	-,042	,124	,090	-,080	,221	,107	,453	,393
s2	-,372	-,042	1,000	-,254	,178	-,199	-,321	-,126	,093	,110
s3	-,070	,124	-,254	1,000	-,159	-,183	,100	,093	,165	,103
s4	,312	,090	,178	-,159	1,000	,342	,123	,039	,330	,351
s5	,486	-,080	-,199	-,183	,342	1,000	,463	-,219	-,486	-,011
s6	-,061	,221	-,321	,100	,123	,463	1,000	-,187	-,216	,046
s7	,159	,107	-,126	,093	,039	-,219	-,187	1,000	,406	,292
s8	-,061	,453	,093	,165	,330	-,486	-,216	,406	1,000	,337
s9	,009	,393	,110	,103	,351	-,011	,046	,292	,337	1,000

Çıktı 1.

Summary Item Statistics

	Mean	Minimum	Maximum	Range	Maximum / Minimum	Variance	N of Items
Item Means	4,260	3,850	4,700	,850	1,221	,058	10
Item Variances	1,947	1,250	2,976	1,726	2,381	,259	10
Inter-Item Correlations	,064	-,486	,486	,972	-1,001	,056	10

Scale Statistics

Mean	Variance	Std. Deviation	N of Items
42,6000	30,147	5,49066	10

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
s1	38,3500	27,924	,082	,588	,391
s10	37,9000	23,989	,392	,465	,281
s2	38,3500	31,292	-,205	,464	,511
s3	38,5500	28,155	-,017	,477	,437
s4	38,1000	21,884	,463	,621	,229
s5	38,1500	29,082	-,034	,700	,431
s6	38,4500	28,261	,005	,556	,422
s7	38,3500	25,713	,171	,322	,360
s8	38,7500	22,092	,313	,705	,285
s9	38,4500	22,261	,500	,358	,227

Çıktı 1.

ANOVA with Tukey's Test for Nonadditivity

	Sum of Squares	df	Mean Square	F	Sig
Between People	57,280	19	3,015		
Within People					
Between Items	10,480	9	1,164	,637	,765
Residual					
Nonadditivity	,131 ^a	1	,131	,071	,790
Balance	312,589	170	1,839		
Total	312,720	171	1,829		
Total	323,200	180	1,796		
Total	380,480	199	1,912		

Grand Mean = 4,2600

a. Tukey's estimate of power to which observations must be raised to achieve additivity = 1,891.

Hotelling's T-Squared Test

Hotelling's T-Squared	F	df1	df2	Sig
19,351	1,245	9	11	,360

Çıktı 18.1'e göre 10 soruluk ölçekte güvenilirlik katsayısı $\alpha=0.3934$ (Stzd.item $\alpha=0.4077$) olarak bulunmuştur. Bu güvenilirlik oldukça düşük bir güvenirliktir.

Alfa katsayısının değerlendirilmesinde uyulan değerlendirme kriteri;

$0.00 \leq \alpha < 0.40$ ise ölçek **güvenilir değildir.**

$0.40 \leq \alpha < 0.60$ ise ölçek **düşük güvenilirdir.**

$0.60 \leq \alpha < 0.80$ ise ölçek **oldukça güvenilirdir.**

$0.80 \leq \alpha \leq 1.00$ ise ölçek yüksek derecede güvenilir bir ölçektir.

Bu kriterlere göre **ölçeğimizin güvenirligi düşüktür.**

Ölçekte yer alan soruların genel ortalaması 4.26'dır. Ortalamaların değişim aralığı=0.85=4.70-3.85 olarak bulunmuştur. Soru ortalamalarının testi sonucunda ortalamaların farklı olmadığı görülmüştür. (Hotelling $T^2=19.35$, $P=0.3601$ ^{ns}).

H₀ : Sorulara verilen cevapların ortalamaları arasında bir fark yoktur. (birbirine eşittir.)

H₁ : Sorulara verilen cevapların ortalamaları arasında bir fark vardır. (birbirine eşit değildir.)

$p < \alpha$ ise H_0 red edilir.

$p > \alpha$ ise H_0 'ı red edecek yeterli kanıt yoktur.

Soru-Bütün (Item-Total) korelasyonlara bakıldığında -0.2046 ile 0.4999 arasında deęişim görülür.

Soru ile Bütün arasındaki korelasyon katsayılarının negatif olmaması gerekir. **Bu durum ölçeğin toplanabilirlik özelliğini bozar.**

Soru-Bütün korelasyonlarının negatif olmaması ve hatta 0.25 değerinden büyük olması beklenir.

Bu kurala uymayan soruların ölçekten çıkarılması önerilmektedir. Fakat bu kesin kural değildir. Bir sorunun ölçekten çıkarılması için;

Soru silinirse Alfa katsayısında (*Alpha if Item Deleted*) değişime bakmak gerekir.

Soru silinirse ortalamadaki (*Scale mean if item deleted*) değişime bakmak gerekir.

- Örneğimizde genel alfa katsayısı 0.3934 olarak bulunmuştur. Bazı sorular ölçekten çıkarıldığında eğer alfa katsayısı yükseliyorsa o soru güvenilirliği azaltan bir sorudur ve ölçekten çıkarılması gerekir.
- **Örneğin** Soru 2 (S2) ölçekten silinirse alfa değeri 0.5115'e çıkacaktır. Öyleyse S2 homojenliği bozar niteliktedir.

- Yapılan varyans analizi sonuçlarına göre 10 soruluk ölçek toplanabilir özelliktedir (Non-additivity, $F=0.071$, $P=0.79^{ns}$).

H_0 : There is no non-additivity=

Toplanamamazlık yoktur.=Toplanabilirlik vardır.

H_1 : There is non-additivity=

Toplanamamazlık vardır.=Toplanabilirlik yoktur.

$p < \alpha$ ise H_0 red edilir.

$p > \alpha$ ise H_0 'ı red edecek yeterli kanıt yoktur.

- Ölçümler arası (between measures -items) değişkenliğe bakıldığında da önemli bir farklılık görülmemektedir ($F=0.637$, $P=0.765$ ^{ns}).

H_0 : Ölçümler (skorlar-sorular) arası değişkenlik yoktur.

H_1 : Ölçümler (skorlar-sorular) arası değişkenlik vardır.

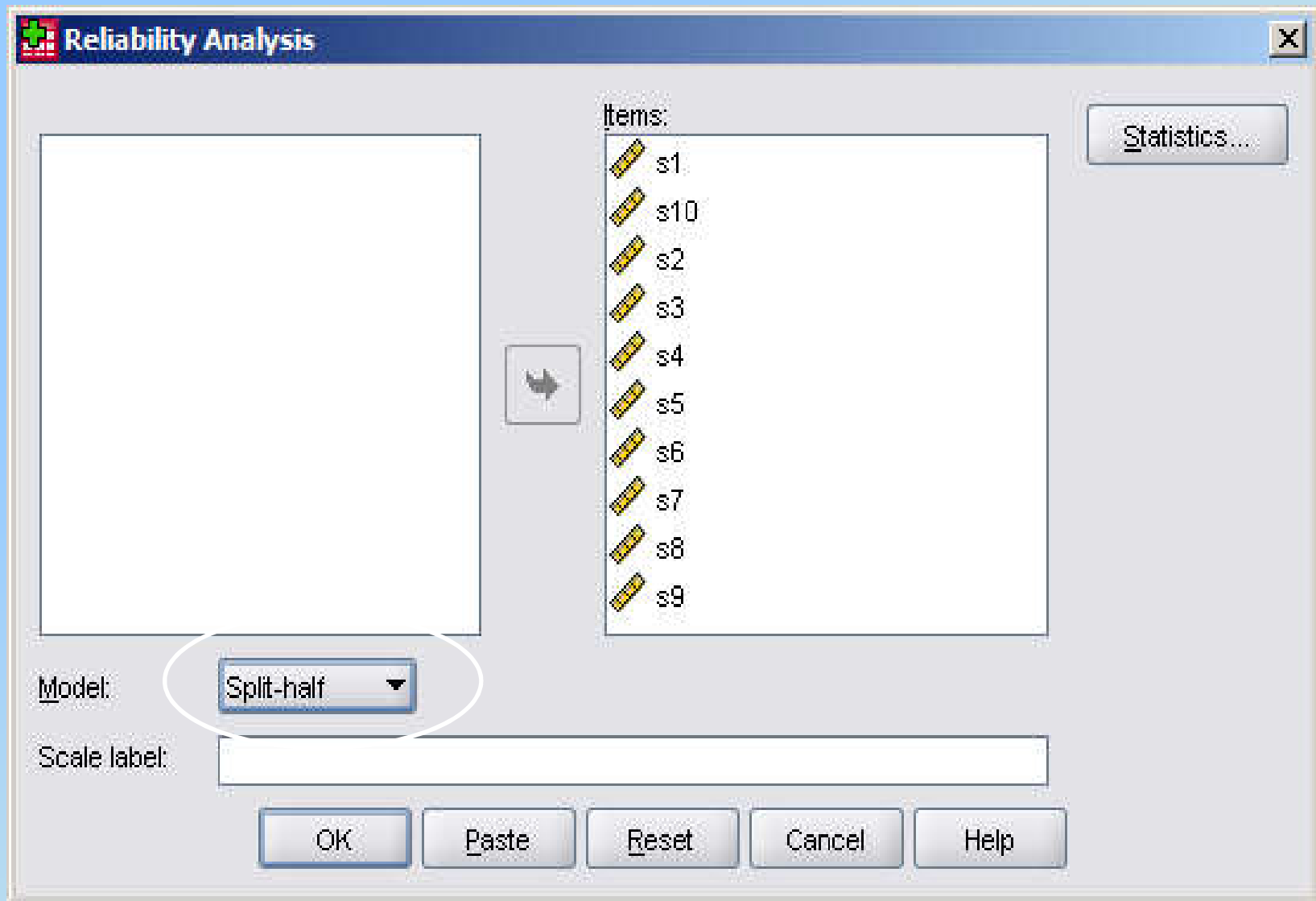
$p < \alpha$ ise H_0 red edilir.

$p > \alpha$ ise H_0 'ı red edecek yeterli delil yoktur.

Split- half Yöntemi

- Örneğimizi, **Split-half yöntemi** aracılığı ile yeniden analiz edersek. Bunun için Ekran 1.'de Alfa yerine **Split-half** seçeneği seçilerek analiz yinelenir. Sonuç Çıktı 2'de yer almaktadır.

Split-half



Çıktı 2.

Reliability Statistics

Cronbach's Alpha	Part 1	Value	-,138 ^a
		N of Items	5,000 ^b
	Part 2	Value	,181
		N of Items	5,000 ^c
	Total N of Items		10,000
Correlation Between Forms			,497
Spearman-Brown Coefficient	Equal Length		,664
	Unequal Length		,664
Guttman Split-Half Coefficient			,658

a. The value is negative due to a negative average covariance among items. This violates reliability model assumptions. You may want to check item codings.

b. The items are: s1, s10, s2, s3, s4.

c. The items are: s5, s6, s7, s8, s9.

Item Statistics

	Mean	Std. Deviation	N
s1	4,2500	1,11803	20
s10	4,7000	1,21828	20
s2	4,2500	1,55174	20
s3	4,0500	1,50350	20
s4	4,5000	1,43270	20
s5	4,4500	1,23438	20
s6	4,1500	1,34849	20
s7	4,2500	1,40955	20
s8	3,8500	1,72520	20
s9	4,1500	1,30888	20

Çıktı 2.

Inter-Item Correlation Matrix

	s1	s10	s2	s3	s4	s5	s6	s7	s8	s9
s1	1,000	,019	-,372	-,070	,312	,486	-,061	,159	-,061	,009
s10	,019	1,000	-,042	,124	,090	-,080	,221	,107	,453	,393
s2	-,372	-,042	1,000	-,254	,178	-,199	-,321	-,126	,093	,110
s3	-,070	,124	-,254	1,000	-,159	-,183	,100	,093	,165	,103
s4	,312	,090	,178	-,159	1,000	,342	,123	,039	,330	,351
s5	,486	-,080	-,199	-,183	,342	1,000	,463	-,219	-,486	-,011
s6	-,061	,221	-,321	,100	,123	,463	1,000	-,187	-,216	,046
s7	,159	,107	-,126	,093	,039	-,219	-,187	1,000	,406	,292
s8	-,061	,453	,093	,165	,330	-,486	-,216	,406	1,000	,337
s9	,009	,393	,110	,103	,351	-,011	,046	,292	,337	1,000

Summary Item Statistics

		Mean	Minimum	Maximum	Range	Maximum / Minimum	Variance	N of Items
Item Means	Part 1	4,350	4,050	4,700	,650	1,160	,064	5 ^a
	Part 2	4,170	3,850	4,450	,600	1,156	,047	5 ^b
	Both Parts	4,260	3,850	4,700	,850	1,221	,058	10
Item Variances	Part 1	1,891	1,250	2,408	1,158	1,926	,252	5 ^a
	Part 2	2,004	1,524	2,976	1,453	1,953	,324	5 ^b
	Both Parts	1,947	1,250	2,976	1,726	2,381	,259	10
Inter-Item Correlations	Part 1	-,017	-,372	,312	,684	-,840	,040	5 ^a
	Part 2	,043	-,486	,463	,949	-,954	,098	5 ^b
	Both Parts	,064	-,486	,486	,972	-1,001	,056	10

a. The items are: s1, s10, s2, s3, s4.

b. The items are: s5, s6, s7, s8, s9.

Çıktı 2.

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
s1	38,3500	27,924	,082	,588	,391
s10	37,9000	23,989	,392	,465	,281
s2	38,3500	31,292	-,205	,464	,511
s3	38,5500	28,155	-,017	,177	,437
s4	38,1000	21,884	,463	,621	,229
s5	38,1500	29,082	-,034	,706	,431
s6	38,4500	28,261	,005	,556	,422
s7	38,3500	25,713	,171	,322	,360
s8	38,7500	22,092	,313	,705	,285
s9	38,4500	22,261	,500	,358	,227

Scale Statistics

	Mean	Variance	Std. Deviation	N of Items
Part 1	21,7500	8,513	2,91773	5 ^a
Part 2	20,8500	11,713	3,42245	5 ^b
Both Parts	42,6000	30,147	5,49066	10

a. The items are: s1, s10, s2, s3, s4.

b. The items are: s5, s6, s7, s8, s9.

Çıktı 2.

ANOVA with Tukey's Test for Nonadditivity

	Sum of Squares	df	Mean Square	F	Sig
Between People	57,280	19	3,015		
Within People					
Between Items	10,480	9	1,164	,637	,765
Residual					
Nonadditivity	,131 ^a	1	,131	,071	,790
Balance	312,589	170	1,839		
Total	312,720	171	1,829		
Total	323,200	180	1,796		
Total	380,480	199	1,912		

Grand Mean = 4,2600

a. Tukey's estimate of power to which observations must be raised to achieve additivity = 1,891.

Hotelling's T-Squared Test

Hotelling's T-Squared	F	df1	df2	Sig
19,351	1,245	9	11	,360

- 10 soru içeren ölçekte

1. kısım güvenirliği $\alpha = -0.138$ ve

2. kısım güvenirliği $\alpha = 0.181$ olarak bulunmuştur.

Bu bilgiler ölçeğin soru sayısının azaltılamayacağını, soruların birbirini izleyen ve soruların fenomeni bir plan içinde sunacak yapıda **düzenlenmediğini** belirtir.

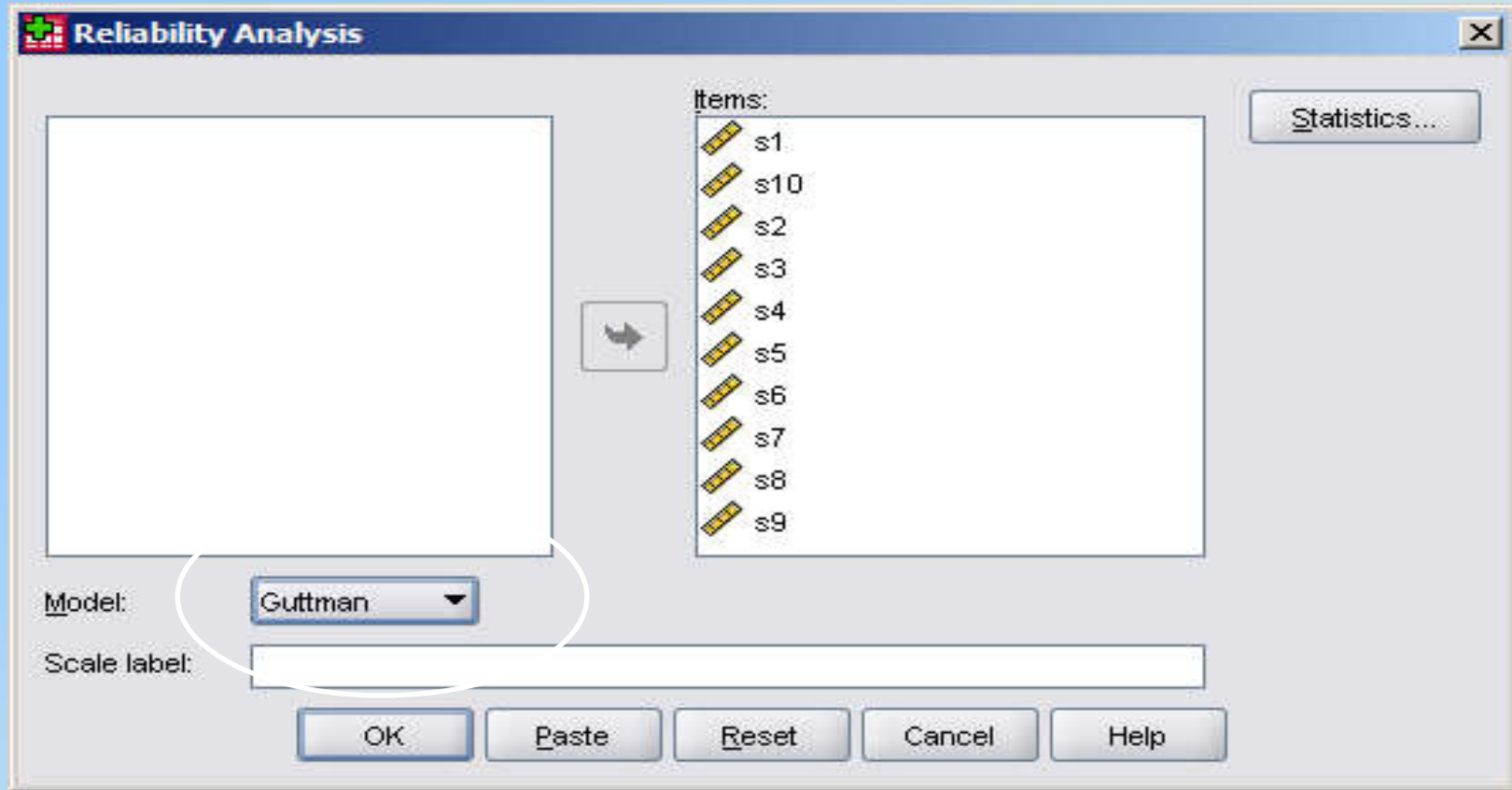
- Anket, split-half yönteminde güvenilirlik korelasyon katsayısı ile belirlenir. Bunun yanında yararlanılan üç güvenilirlik ölçüsü daha vardır. Bunlar
 1. Eşit Spearman-Brown,
 2. Farklı Spearman-Brown katsayıları ve
 3. Guttman Split -Half katsayılarıdır.
- Bu katsayılar güvenirlüğün normal kabul sınırlarında olduğunu belirtirken alfa katsayısı bu bilginin tam tersi sonuçları belirtmektedir.

Guttman Yöntemi

- Örneğimizi Guttman Yöntemine göre yeniden analiz edelim. Sonuçlar Çıktı 3'de verilmiştir.

Çıktı 18.3

- Guttman Yöntemine göre:



Guttman Yöntemine Göre

Reliability Statistics

Lambda	1	,354
	2	,512
	3	,393
	4	,658
	5	,498
	6	,674
N of Items		10,000

Item Statistics

	Mean	Std. Deviation	N
s1	4,2500	1,11803	20
s10	4,7000	1,21828	20
s2	4,2500	1,55174	20
s3	4,0500	1,50350	20
s4	4,5000	1,43270	20
s5	4,4500	1,23438	20
s6	4,1500	1,34849	20
s7	4,2500	1,40955	20
s8	3,8500	1,72520	20
s9	4,1500	1,30888	20

Inter-Item Correlation Matrix

	s1	s10	s2	s3	s4	s5	s6	s7	s8	s9
s1	1,000	,019	-,372	-,070	,312	,486	-,061	,159	-,061	,009
s10	,019	1,000	-,042	,124	,090	-,080	,221	,107	,453	,393
s2	-,372	-,042	1,000	-,254	,178	-,199	-,321	-,126	,093	,110
s3	-,070	,124	-,254	1,000	-,159	-,183	,100	,093	,165	,103
s4	,312	,090	,178	-,159	1,000	,342	,123	,039	,330	,351
s5	,486	-,080	-,199	-,183	,342	1,000	,463	-,219	-,486	-,011
s6	-,061	,221	-,321	,100	,123	,463	1,000	-,187	-,216	,046
s7	,159	,107	-,126	,093	,039	-,219	-,187	1,000	,406	,292
s8	-,061	,453	,093	,165	,330	-,486	-,216	,406	1,000	,337
s9	,009	,393	,110	,103	,351	-,011	,046	,292	,337	1,000

Çıktı 18.3

Summary Item Statistics

		Mean	Minimum	Maximum	Range	Maximum / Minimum	Variance	N of Items
Item Means	Part 1	4,350	4,050	4,700	,650	1,160	,064	5 ^a
	Part 2	4,170	3,850	4,450	,600	1,156	,047	5 ^b
	Both Parts	4,260	3,850	4,700	,850	1,221	,058	10
Item Variances	Part 1	1,891	1,250	2,408	1,158	1,926	,252	5 ^a
	Part 2	2,004	1,524	2,976	1,453	1,953	,324	5 ^b
	Both Parts	1,947	1,250	2,976	1,726	2,381	,259	10
Inter-Item Correlations	Part 1	-,017	-,372	,312	,684	-,840	,040	5 ^a
	Part 2	,043	-,486	,463	,949	-,954	,098	5 ^b
	Both Parts	,064	-,486	,486	,972	-1,001	,056	10

a. The items are: s1, s10, s2, s3, s4.

b. The items are: s5, s6, s7, s8, s9.

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
s1	38,3500	27,924	,082	,588	,391
s10	37,9000	23,989	,392	,465	,281
s2	38,3500	31,292	-,205	,464	,511
s3	38,5500	28,155	-,017	,177	,437
s4	38,1000	21,884	,463	,621	,229
s5	38,1500	29,082	-,034	,706	,431
s6	38,4500	28,261	,005	,556	,422
s7	38,3500	25,713	,171	,322	,360
s8	38,7500	22,092	,313	,705	,285
s9	38,4500	22,261	,500	,358	,227

Çıktı 18.3

Scale Statistics

	Mean	Variance	Std. Deviation	N of Items
Part 1	21,7500	8,513	2,91773	5 ^a
Part 2	20,8500	11,713	3,42245	5 ^b
Both Parts	42,6000	30,147	5,49066	10

a. The items are: s1, s10, s2, s3, s4.

b. The items are: s5, s6, s7, s8, s9.

ANOVA with Tukey's Test for Nonadditivity

		Sum of Squares	df	Mean Square	F	Sig
Between People		57,280	19	3,015		
Within People	Between Items	10,480	9	1,164	,637	,765
	Residual					
	Nonadditivity	,131 ^a	1	,131	,071	,790
	Balance	312,589	170	1,839		
	Total	312,720	171	1,829		
	Total	323,200	180	1,796		
Total		380,480	199	1,912		

Grand Mean = 4,2600

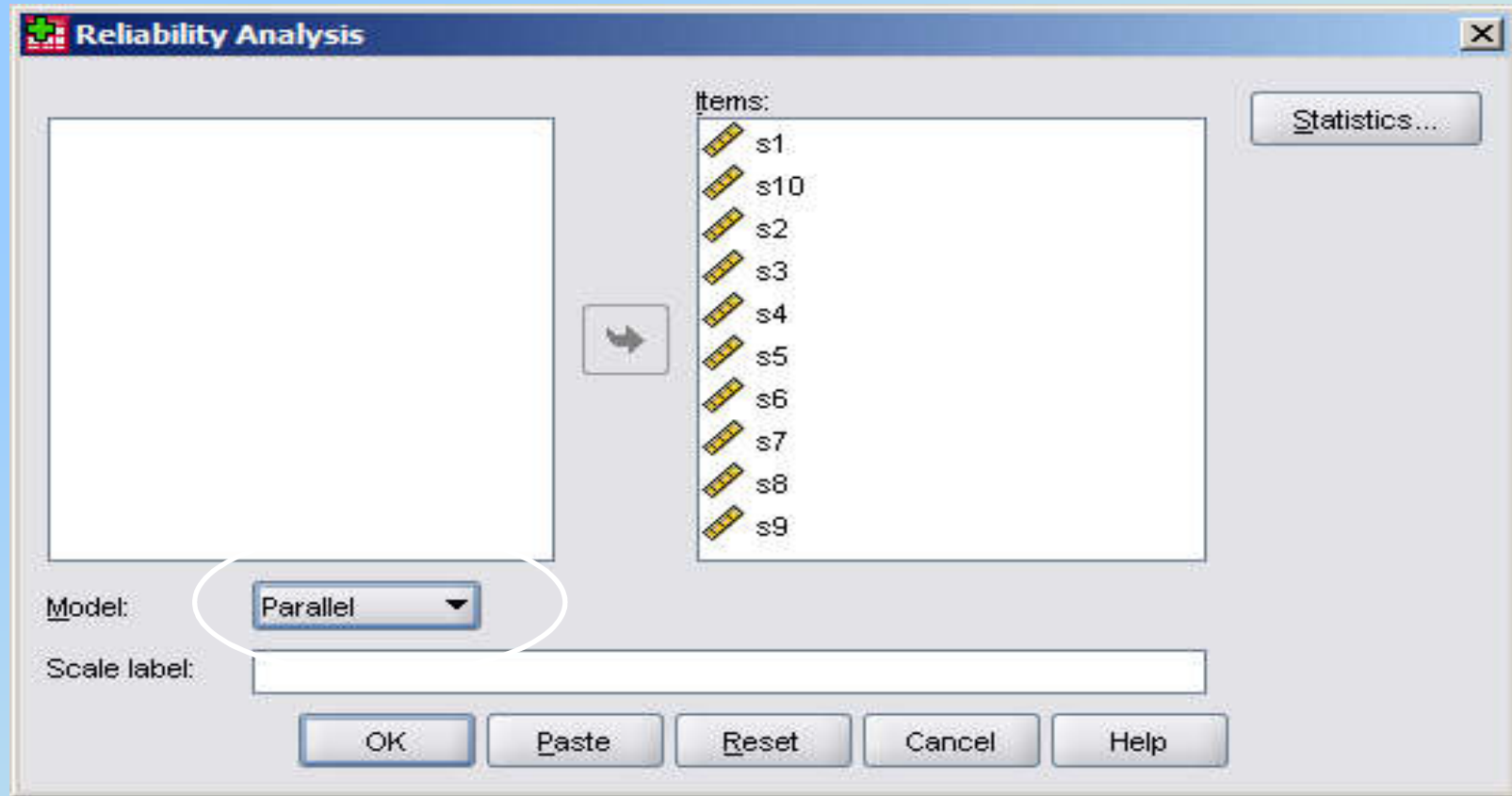
a. Tukey's estimate of power to which observations must be raised to achieve additivity = 1,891.

Hotelling's T-Squared Test

Hotelling's T-Squared	F	df1	df2	Sig
19,351	1,245	9	11	,360

Çıktı 4.

- Parallel Yöntemine göre:



Çıktı 4.

Test for Model Goodness of Fit

Chi-Square	Value	56,407
	df	53,000
	Sig.	,349
Log of Determinant of	Unconstrained Matrix	2,875
	Constrained Matrix	6,536

Reliability Statistics

Common Variance	1,947
True Variance	,119
Error Variance	1,829
Common Inter-Item Correlation	,061
Reliability of Scale	,393
Reliability of Scale (Unbiased)	,457

Item Statistics

	Mean	Std. Deviation	N
s1	4,2500	1,11803	20
s10	4,7000	1,21828	20
s2	4,2500	1,55174	20
s3	4,0500	1,50350	20
s4	4,5000	1,43270	20
s5	4,4500	1,23438	20
s6	4,1500	1,34849	20
s7	4,2500	1,40955	20
s8	3,8500	1,72520	20
s9	4,1500	1,30888	20

Under the parallel model assumption

Inter-Item Correlation Matrix

	s1	s10	s2	s3	s4	s5	s6	s7	s8	s9
s1	1,000	,019	-,372	-,070	,312	,486	-,061	,159	-,061	,009
s10	,019	1,000	-,042	,124	,090	-,080	,221	,107	,453	,393
s2	-,372	-,042	1,000	-,254	,178	-,199	-,321	-,126	,093	,110
s3	-,070	,124	-,254	1,000	-,159	-,183	,100	,093	,165	,103
s4	,312	,090	,178	-,159	1,000	,342	,123	,039	,330	,351
s5	,486	-,080	-,199	-,183	,342	1,000	,463	-,219	-,486	-,011
s6	-,061	,221	-,321	,100	,123	,463	1,000	-,187	-,216	,046
s7	,159	,107	-,126	,093	,039	-,219	-,187	1,000	,406	,292
s8	-,061	,453	,093	,165	,330	-,486	-,216	,406	1,000	,337
s9	,009	,393	,110	,103	,351	-,011	,046	,292	,337	1,000

Çıktı.4.

Summary Item Statistics

	Mean	Minimum	Maximum	Range	Maximum / Minimum	Variance	N of Items
Item Means	4,260	3,850	4,700	,850	1,221	,058	10
Item Variances	1,947	1,250	2,976	1,726	2,381	,259	10
Inter-Item Correlations	,064	-,486	,486	,972	-1,001	,056	10

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
s1	38,3500	27,924	,082	,588	,391
s10	37,9000	23,989	,392	,465	,281
s2	38,3500	31,292	-,205	,464	,511
s3	38,5500	28,155	-,017	,177	,437
s4	38,1000	21,884	,463	,621	,229
s5	38,1500	29,082	-,034	,706	,431
s6	38,4500	28,261	,005	,556	,422
s7	38,3500	25,713	,171	,322	,360
s8	38,7500	22,092	,313	,705	,285
s9	38,4500	22,261	,500	,358	,227

Scale Statistics

Mean	Variance	Std. Deviation	N of Items
42,6000	30,147	5,49066	10

Hotelling's T-Squared Test

Hotelling's T-Squared	F	df1	df2	Sig.
19,351	1,245	9	10	,360

ANOVA with Tukey's Test for Nonadditivity

	Sum of Squares	df	Mean Square	F	Sig.
Between People	57,280	19	3,015		
Within People					
Between Items	10,480	9	1,164	,637	,765
Residual Nonadditivity	,131 ^a	1	,131	,071	,790
Balance	312,589	170	1,839		
Total	312,720	171	1,829		
Total	323,200	180	1,796		
Total	380,480	199	1,912		

Grand Mean = 4,2600

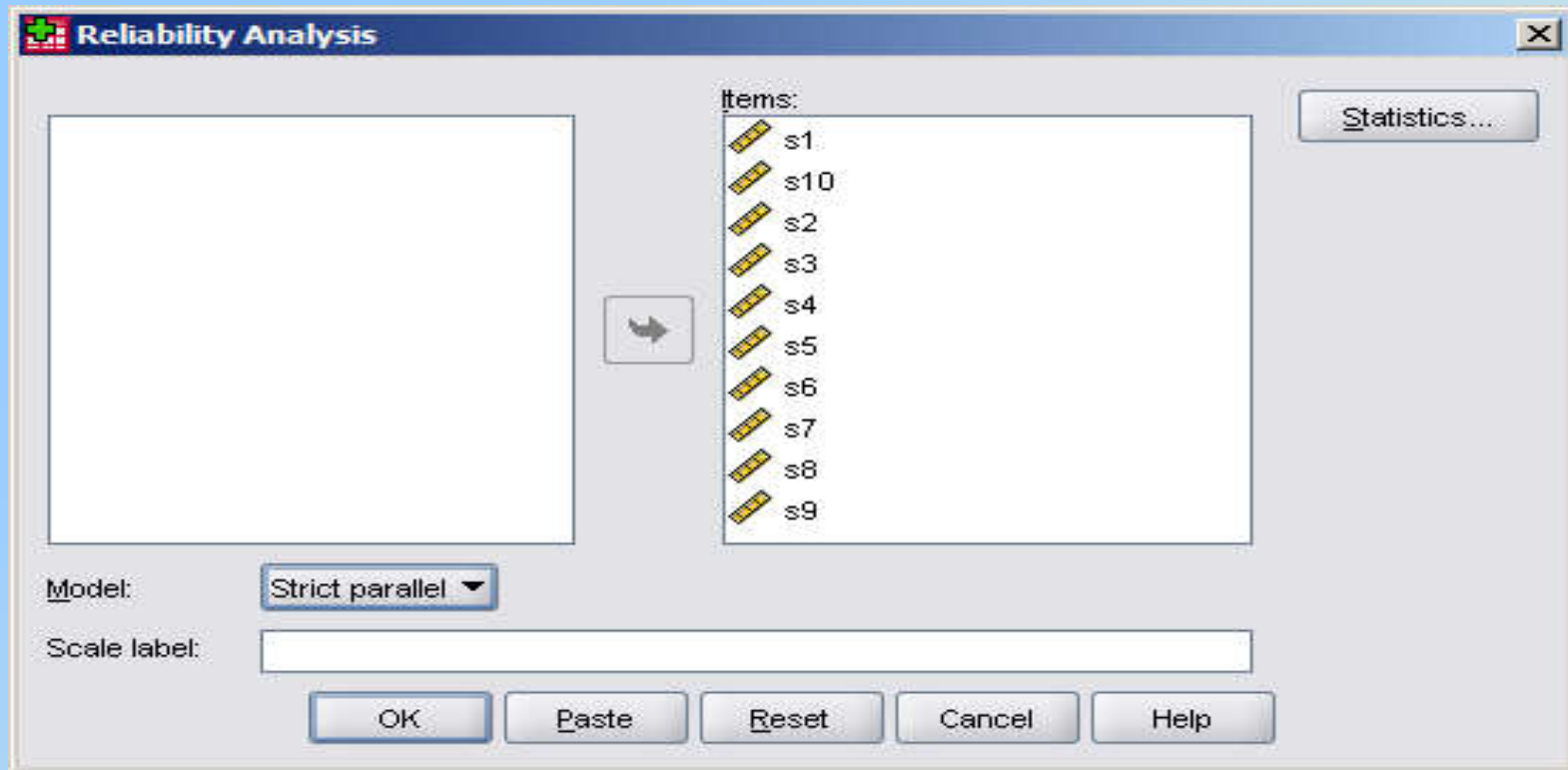
a. Tukey's estimate of power to which observations must be raised to achieve additivity = 1,891.

- Paralel yöntemle göre güvenilirlik katsayısı 0.393 .
- (Sapmasız güvenilirlik katsayısı=0.457) olarak belirlenmiştir.

Kesin Paralel Yöntem

- Örneğimizi Kesin Paralel yöntemle göre analiz ettiğimizde sonuç Çıktı 5'deki gibi elde edilir.

Strict Parallel Yöntemine Göre



Çıktı 5

Test for Model Goodness of Fit

Chi-Square	Value	62,810
	df	62,000
	Sig.	,447
Log of Determinant of	Unconstrained Matrix	2,875
	Constrained Matrix	6,818

1. Model için istatistiksel olarak uygun model varsayımına

Reliability Statistics

Common Mean	4,260
Common Variance	2,000
True Variance	,204
Error Variance	1,796
Common Inter-Item Correlation	,056
Reliability of Scale	,374
Reliability of Scale (Unbiased)	,468

Item Statistics

	Mean	Std. Deviation	N
s1	4,2500	1,11803	20
s10	4,7000	1,21828	20
s2	4,2500	1,55174	20
s3	4,0500	1,50350	20
s4	4,5000	1,43270	20
s5	4,4500	1,23438	20
s6	4,1500	1,34849	20
s7	4,2500	1,40955	20
s8	3,8500	1,72520	20
s9	4,1500	1,30888	20

Inter-Item Correlation Matrix

	s1	s10	s2	s3	s4	s5	s6	s7	s8	s9
s1	1,000	,019	-,372	-,070	,312	,486	-,061	,159	-,061	,009
s10	,019	1,000	-,042	,124	,090	-,080	,221	,107	,453	,393
s2	-,372	-,042	1,000	-,254	,178	-,199	-,321	-,126	,093	,110
s3	-,070	,124	-,254	1,000	-,159	-,183	,100	,093	,165	,103
s4	,312	,090	,178	-,159	1,000	,342	,123	,039	,330	,351
s5	,486	-,080	-,199	-,183	,342	1,000	,463	-,219	-,486	-,011
s6	-,061	,221	-,321	,100	,123	,463	1,000	-,187	-,216	,046
s7	,159	,107	-,126	,093	,039	-,219	-,187	1,000	,406	,292
s8	-,061	,453	,093	,165	,330	-,486	-,216	,406	1,000	,337
s9	,009	,393	,110	,103	,351	-,011	,046	,292	,337	1,000

Summary Item Statistics

	Mean	Minimum	Maximum	Range	Maximum / Minimum	Variance	N of Items
Item Means	4,260	3,850	4,700	,850	1,221	,058	10
Item Variances	1,947	1,250	2,976	1,726	2,381	,259	10
Inter-Item Correlations	,064	-,486	,486	,972	-1,001	,056	10

Çıktı 5

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
s1	38,3500	27,924	,082	,588	,391
s10	37,9000	23,989	,392	,465	,281
s2	38,3500	31,292	-,205	,464	,511
s3	38,5500	28,155	-,017	,177	,437
s4	38,1000	21,884	,463	,621	,229
s5	38,1500	29,082	-,034	,706	,431
s6	38,4500	28,261	,005	,556	,422
s7	38,3500	25,713	,171	,322	,360
s8	38,7500	22,092	,313	,705	,285
s9	38,4500	22,261	,500	,358	,227

Scale Statistics

Mean	Variance	Std. Deviation	N of Items
42,6000	30,147	5,49066	10

Hotelling's T-Squared Test

Hotelling's T-Squared	F	df1	df2	Sig
19,351	1,245	9	11	,360

ANOVA with Tukey's Test for Nonadditivity

	Sum of Squares	df	Mean Square	F	Sig
Between People	57,280	19	3,015		
Within People					
Between Items	10,480	9	1,164	,637	,765
Residual					
Nonadditivity	,131 ^a	1	,131	,071	,790
Balance	312,589	170	1,839		
Total	312,720	171	1,829		
Total	323,200	180	1,796		
Total	380,480	199	1,912		

Grand Mean = 4,2600

a. Tukey's estimate of power to which observations must be raised to achieve additivity = 1,891.

- Kesin paralel yöntemle göre güvenilirlik katsayısı 0.374 (sapmasız tahmin 0.468) bulunmuştur.

- Örnek testin güvenilirliğinde Guttman 3. istatistiği, alfa katsayısı ve kesin paralel ile paralel yöntemin güvenilirlik katsayıları birbirlerine yakın sonuçlar vermektedir. İki parçaya ayırma yönteminde ise parçalar arası korelasyon katsayısı belirtilen katsayılara yakın değerler almaktadır.

- Bu sonuçlara göre soru (madde, item) istatistiklerine bakarak eğer varyanslar birbirine eşit ise (varyanslar homojen ise) alfa katsayısı ve paralel yöntem güvenilirlik katsayılarını ölçeğin güvenilirlik katsayısı olarak kullanmak gerekir. Soruların varyansları homojen ve ortalamaları benzer ise bu durumda Kesin (Tam) Paralel yöntem güvenilirlik katsayılarını kullanmak gerekir.

Alpha Yöntemi (soru 2 çıkartılmış)

- Alpha Yöntemine göre yapılan çözüm sonucunda, 2. sorunun anketten çıkartılması önerilmişti. Soru 2 anketten çıkartılıp Alpha Yöntemini yeni veri setine uyguladığımızda aşağıda verilen sonuçlara ulaşılacaktır.

2.Soru çıkartılırsa

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
s1	38,3500	27,924	,082	,588	,391
s10	37,9000	23,989	,392	,465	,281
s2	38,3500	31,292	-,205	,464	,511
s3	38,5500	28,155	-,017	,177	,437
s4	38,1000	21,884	,463	,621	,229
s5	38,1500	29,082	-,034	,706	,431
s6	38,4500	28,261	,005	,556	,422
s7	38,3500	25,713	,171	,322	,360
s8	38,7500	22,092	,313	,705	,285
s9	38,4500	22,261	,500	,358	,227

*Güvenirlık Analızı Tablo18.1(2 silindi).sav [DataSet2] - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utilities Add-ons Window Help

23: s7 Visible: 10 of 10 Variables

	s1	s10	s3	s4	s5	s6	s7	s8	s9	var
1	4,00	4,00	3,00	5,00	4,00	3,00	2,00	3,00	5,00	
2	3,00	4,00	5,00	3,00	4,00	5,00	4,00	3,00	2,00	
3	5,00	7,00	4,00	5,00	4,00	3,00	6,00	5,00	4,00	
4	4,00	6,00	3,00	4,00	6,00	6,00	2,00	2,00	3,00	
5	2,00	4,00	5,00	3,00	5,00	5,00	4,00	1,00	5,00	
6	3,00	3,00	4,00	6,00	4,00	4,00	3,00	4,00	4,00	
7	6,00	5,00	3,00	5,00	7,00	3,00	5,00	3,00	6,00	
8	5,00	6,00	6,00	7,00	6,00	7,00	4,00	5,00	5,00	
9	4,00	4,00	5,00	4,00	3,00	2,00	6,00	4,00	5,00	
10	3,00	7,00	7,00	3,00	2,00	4,00	4,00	6,00	5,00	
11	4,00	5,00	4,00	5,00	4,00	3,00	3,00	7,00	4,00	
12	5,00	4,00	2,00	4,00	5,00	5,00	4,00	1,00	3,00	
13	6,00	3,00	3,00	6,00	6,00	4,00	3,00	2,00	2,00	
14	5,00	5,00	5,00	3,00	5,00	4,00	3,00	3,00	3,00	
15	4,00	4,00	4,00	3,00	5,00	6,00	6,00	4,00	4,00	
16	3,00	6,00	1,00	6,00	4,00	5,00	4,00	5,00	6,00	
17	4,00	5,00	2,00	5,00	3,00	3,00	7,00	7,00	4,00	
18	5,00	4,00	6,00	7,00	5,00	4,00	6,00	5,00	5,00	
19	6,00	5,00	5,00	4,00	4,00	5,00	5,00	4,00	6,00	
20	4,00	3,00	4,00	2,00	3,00	2,00	4,00	3,00	2,00	
21										

Data View Variable View

SPSS Processor is ready

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
,511	,517	9

Item Statistics

	Mean	Std. Deviation	N
s1	4,2500	1,11803	20
s10	4,7000	1,21828	20
s3	4,0500	1,50350	20
s4	4,5000	1,43270	20
s5	4,4500	1,23438	20
s6	4,1500	1,34849	20
s7	4,2500	1,40955	20
s8	3,8500	1,72520	20
s9	4,1500	1,30888	20

Inter-Item Correlation Matrix

	s1	s10	s3	s4	s5	s6	s7	s8	s9
s1	1,000	,019	-,070	,312	,486	-,061	,159	-,061	,009
s10	,019	1,000	,124	,090	-,080	,221	,107	,453	,393
s3	-,070	,124	1,000	-,159	-,183	,100	,093	,165	,103
s4	,312	,090	-,159	1,000	,342	,123	,039	,330	,351
s5	,486	-,080	-,183	,342	1,000	,463	-,219	-,486	-,011
s6	-,061	,221	,100	,123	,463	1,000	-,187	-,216	,046
s7	,159	,107	,093	,039	-,219	-,187	1,000	,406	,292
s8	-,061	,453	,165	,330	-,486	-,216	,406	1,000	,337
s9	,009	,393	,103	,351	-,011	,046	,292	,337	1,000

Summary Item Statistics

	Mean	Minimum	Maximum	Range	Maximum / Minimum	Variance	N of Items
Item Means	4,261	3,850	4,700	,850	1,221	,065	9
Item Variances	1,896	1,250	2,976	1,726	2,381	,262	9
Inter-Item Correlations	,106	-,486	,486	,972	-1,001	,052	9

Scale Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
s1	34,1000	27,779	,192	,454	,492
s10	33,6500	24,976	,397	,454	,430
s3	34,3000	28,116	,057	,146	,541
s4	33,8500	23,818	,388	,568	,422
s5	33,9000	29,463	,023	,705	,540
s6	34,2000	28,063	,099	,443	,522
s7	34,1000	26,305	,207	,318	,488
s8	34,5000	23,737	,272	,699	,464
s9	34,2000	23,853	,448	,357	,407

Mean	Variance	Std. Deviation	N of Items
38,3500	31,292	5,59393	9

Hotelling's T-Squared Test

Hotelling's T-Squared	F	df1	df2	Sig.
18,640	1,472	8	12	,263

ANOVA with Tukey's Test for Nonadditivity

	Sum of Squares	df	Mean Square	F	Sig.
Between People	66,061	19	3,477		
Within People					
Between Items	10,478	8	1,310	,771	,629
Residual Nonadditivity	,211 ^a	1	,211	,123	,726
Balance	257,978	151	1,708		
Total	258,189	152	1,699		
Total	268,667	160	1,679		
Total	334,728	179	1,870		

Grand Mean = 4,2611

a. Tukey's estimate of power to which observations must be raised to achieve additivity = 1,998.

- Örneğin Soru 2 (S2) ölçekten silinirse alfa değeri 0.5115'e çıkacaktır. Öyleyse S2 homojenliği bozar niteliktedir.
- Ölçeğimiz için mutlaka gerekli soruları soru silindiğinde alfa katsayısının en düşük olduğu sorudan başlayarak sıralayabiliriz: 1-S9, 2-S4, 3-S10, 4-S8, 5-S7, 6-S1 .

Güvenilirliđi deđiřtirmeyen sorular ise ölçeđi destekleyen sorulardır. Ölçekten çıkarılmaması gerekir. Bu sorular ise; S6, S5ve S3 sorularıdır.(Soru 2 ölçekten çıkartıldıđında, item-total correlation'ların hepsi pozitif olduđundan başka bir soruyu çıkartmaya gerek kalmadıđı sonucuna varabiliriz. Bu soruların cronbach's-alpha if item deleted deđerleri 0.511'den büyük olmasına karřın ölçekten çıkartılmaması önerilmektedir.)

Madde Sayısının Artırılmasıyla veya Azaltılmasıyla Alfa Güvenilirlik Katsayısını Tahmin Etme

Alfa güvenilirlik katsayısı ölçekteki/testteki madde sayısına bağlıdır. Madde sayısı artıkça güvenilirlikte artar. Belirli bir güvenilirlik oranına ulaşmak için madde sayısının ne oranda artırılacağına Spearman-Brown formülüyle karar verilir.

Madde Sayısının Tahmini:

Yapılan pilot çalışma sonucunda istenilen güvenilirlik katsayısı elde edilmemişse, madde sayısının artırılmasına gidilir. Bunun için Spearman-Brown formülü kullanılır.

$$m = \frac{r_{SB} (1 - r_m)}{r_m (1 - r_{SB})}$$

m : Mevcut maddelere ilave edilmesi gereken oran

r_{SB} : İstenen güvenilirlik düzeyi

r_m : mevcut değer, hesaplanan güvenilirlik düzeyi

Örnek. Bir arařtırmacı 20 maddeden oluřan bir ölçeęin alfa güvenilirlik katsayısı 0,70 olarak saptanmıřtır. Bu katsayının 0,80 düzeyine çıkarmak için kaç madde ilave edilmesi gerektięini bulunuz?

$$m = \frac{r_{SB} (1 - r_m)}{r_m (1 - r_{SB})} = \frac{0,80(1 - 0,70)}{0,70(1 - 0,80)} = 1,714$$

k: Gerekli toplam madde sayısı

$$k = \text{mevcut madde sayısı} * \text{gerekli oran} = 20 * 1,714 = 34$$

Ölçekteki madde sayısı 20 den 34'e çıkarılırsa güvenilirlik 0,80 e çıkar.

Güvenilirlik Katsayısının Tahmini:

Ölçeğe mevcut sayının belirli bir oranı kadar madde ilave edildiğinde güvenilirlik katsayısının ne olacağı aşağıdaki gibi hesaplanır:

$$r_{SB} = \frac{mr_m}{1 + (m - 1)r_m}$$

r_{SB} : Spearman-Brown güvenilirlik katsayısı
 m : İlave edilmesi gereken oran
 r_m : Mevcut güvenilirlik katsayısı

Örnek. Bir araştırmacı 12 maddeden oluşan bir ölçeğin alfa güvenilirlik katsayısı 0,70 olarak saptamıştır. Madde sayısı 20'ye çıkarıldığında güvenilirlik katsayısı ne olur?

$$m=20/12=1,66$$

ilave edilmesi gereken oran

$$r_{SB} = \frac{1,66 * 0,70}{1 + (1,66 - 1) * 0,70} = 0,79$$

KORELASYON KATSAYILARININ FISHER Z PUANLARINA DÖNÜŞTÜRÜLMESİ

Sıralı ölçek verisi niteliğindeki veriler varsa, normallik varsayımı geçerli olmaz. Fisher (1915) böyle durumlarda elde edilen değerleri Z puanlarına dönüştürerek, verilerin normal dağılım özelliğine sahip olmasını sağlamıştır. Normal dağılım özelliği gösteren Fisher Z değerleri, standart Z puanlarıyla karıştırılmaması gerekir. Her iki yöntemin de hesaplama formülleri farklıdır.

Korelasyon (güvenilirlik) katsayılarının Fisher Z puanlarına dönüştürülmesi işlemi test veya ölçek sonuçlarının daha sağlıklı karşılaştırılmasına imkan sağlar. Bu şekilde güvenilirlik katsayıları yansız biçimde değerlendirilmiş olur.

Korelasyon analizine dayalı güvenilirlik katsayıları Fisher Z puanlarına aşağıdaki amaçla için dönüştürülebilir:

1. Korelasyon katsayısının hipotez testi ile ana kütle için anlamlı olup olmadığını belirlemek.
2. Aynı ölçek veya teste ait farklı güvenilirlik katsayılarını karşılaştırmak.
3. Elde edilen güvenilirlik katsayısının (korelasyon) güven aralığını belirlemek.
4. Farklı güvenilirlik katsayılarını tek bir güvenilirlik katsayısı haline getirmek ve daha sonra bu güvenilirlik katsayısının güven aralığını belirlemek.
5. Anakütle etki büyüklüğünü doğru biçimde ölçmek.

Korelasyon Katsayısının Anlamlılığını Belirleme

Test-Yeniden Test (Re-test) veya Paralel Test uygulamalarından elde edilen korelasyon katsayısının;

a) Ne ölçüde yüksek güvenilirliğe sahip olduğunu

b) Ne ölçüde anakütleye genellenebileceğini

belirlemek için korelasyon (r) değeri Fisher Z puanlarına dönüştürülmelidir.

Örneklem sayısı arttıkça güven aralığının sınırları daralır.

H_0 : $\rho=0$ (Anakütle korelasyon katsayısı sıfırdır, test-yeniden test değerleri birbirinden farklıdır)

H_1 : $\rho \neq 0$ Anakütle korelasyon katsayısı sıfırdan farklıdır, test-yeniden test değerleri arasında önemli bir farklılık yoktur.)

$$Z = 0,5[\ln(1+r) - \ln(1-r)] \quad , \quad Z = 0,5 \ln \left[\frac{1+r}{1-r} \right]$$

$$SS = \sigma_z = \frac{1}{\sqrt{n-3}}$$

$$\text{Güven Aralığı : } Z \pm 1,96 * \frac{1}{\sqrt{n-3}}$$

Güven aralıkları ve bunlara karşı gelen standart Z değerleri:

1- α	: 0,80	0,90	0,95	0,99
Z	: 1,28	1,65	1,96	2,58

Güven aralığı sonucunda alt ve üst sınır korelasyon değerleri bulunur. Bunlar korelasyon katsayısının anakütlerde hangi değerler arasında oynayacağını gösterir. **Eğer güven aralığı sıfırı kapsıyorsa H_0 hipotezinin kabul edildiği yani %95 güven aralığında test-yeniden test uygulamaları arasında istatistiksel olarak anlamlı bir ilişki olmadığı söylenebilir. Sıfır hipotezin kabul edilmesi birinci ölçümdeki verilerle ikinci ölçümdeki veriler arasında önemli farklılık olduğunu gösterir.**

Eğer güven aralığı sıfırı kapsamıyorsa H_0 hipotezi red edilir yani test-yeniden test güvenilirlik katsayısı anlamlıdır, yüksek güvenilirliğe sahiptir denilebilir.

Örnek. 50 soruluk bir ölçek aynı kitle üzerinde iki farklı zamanda uygulanmış ve test-yeniden test yöntemi sonucunda iki ölçümün korelasyon katsayısı 0,45 bulunmuştur. Bulunan güvenilirlik katsayısının %5 önem seviyesinde önemli olup olmadığını test ediniz?

$$H_0 : \rho = 0$$

$$H_1 : \rho \neq 0$$

$$Z = 0,5 \ln \left[\frac{1+r}{1-r} \right] = 0,5 \ln \left[\frac{1+0,45}{1-0,45} \right] = 0,484$$

$$SS = \sigma_z = \frac{1}{\sqrt{n-3}} = \frac{1}{\sqrt{50-3}} = 0,15$$

$$Z \pm 1,96 * \frac{1}{\sqrt{n-3}}$$

$$0,484 \pm 1,96 * 0,15 \Rightarrow (0,19; 0,778)$$

H_0 hipotezi red edilir yani test-yeniden test güvenilirlik katsayısı anlamlıdır.

Güvenilirlik Katsayıların Karşılaştırılması

Eğer birden fazla araştırma yapılmışsa elde edilen güvenilirlik katsayıları arasında önemli bir farklılık bulunup bulunmadığı Fisher Z puanları ile belirlenebilir. **Güvenilirlik katsayısı Cronbach alfa, Küme içi korelasyon katsayısı (ICC) veya korelasyon katsayısı olabilir.**

Alfa değeri ve korelasyon katsayıları eşit aralıklı ölçek verisi değildir. Bu yüzden hesaplama sonucunda elde edilen $r=0,1$ ile $r=0,2$ arasındaki mesafe; bir başka araştırmada elde edilen $r=0,6$ ile $r=0,7$ arasındaki mesafeye eşit değildir. Fisher Z değeri ile güvenilirlik katsayıları arasındaki büyüklüğün konumu daha doğru biçimde belirlenir.

İki bağımsız örneklemden elde edilen korelasyon (güvenilirlik) katsayılarının eşitliğini belirlemek için aşağıdaki işlemler yapılır:

$$H_0 : \rho_1 = \rho_2$$

Birinci çalışmadaki korelasyon katsayısı, ikinci çalışmadaki korelasyon katsayısına eşittir.

$$H_1 : \rho_1 \neq \rho_2$$

$$Z = \frac{Z_1 - Z_2}{\sqrt{\frac{1}{n_1 - 3} + \frac{1}{n_2 - 3}}}$$

Z_1 : Birinci ölçüme ait Fisher Z değeri

Z_2 : İkinci ölçüme ait Fisher Z değeri

n_1 : Birinci ölçümdeki örneklem sayısı

n_2 : İkinci ölçümdeki örneklem sayısı

Örnek. 100 kişi üzerinde uygulanan bir test-yeniden test araştırmasında güvenilirlik katsayısı 0,80 bulunmuştur. 60 kişiden oluşan bir başka grupta ise test-yeniden test uygulaması sonucunda güvenilirlik katsayısı 0,7 çıkmıştır. %5 önem seviyesinde iki güvenilirlik katsayısı birbirinden farklı mıdır?

$$H_0 : \rho_1 = \rho_2$$

$$H_1 : \rho_1 \neq \rho_2$$

$$Z = \frac{0,5 \ln \left[\frac{1+0,8}{1-0,8} \right] - 0,5 \ln \left[\frac{1+0,7}{1-0,7} \right]}{\sqrt{\frac{1}{100-3} + \frac{1}{60-3}}} = 1,43$$

$Z=1,43 < Z_{\text{tablo}}=1,96$ olduğundan H_0 kabul edilir. İki güvenilirlik katsayısı %95 güven seviyesinde birbirinden farklı değildir.

Sorular

1. Ayşe öğretmen sınıfa uyguladığı istatistik testinin aynısını 1 ay sonra tekrar aynı sınıfa uygulamış ise ve testin güvenirliğini bulmaya çalışıyorsa, Ayşe öğretmenin hangi güvenirlik yöntemini kullandığını belirtiniz.
 - a) KR-20
 - b) Paralel form yöntemi
 - c) Yarıya bölme yöntemi
 - d) Cronbach alfa yöntemi
 - e) Test-tekrar test yöntemi

Cevap: E

2. Fatma öğretmen sınav değerlendirmesini yapmış ve kâğıtlar üzerine hiçbir not düşmemiştir, notları farklı bir listeye yazıp kasaya kilitlemiştir. 15 gün sonra tekrar aynı kâğıtları yeniden değerlendirmiş ve notları ayrı bir liste halinde kaydetmiştir. Yani farklı zamanlarda aynı kâğıtları iki kere değerlendirmiştir. Burada iki okuma notları arasındaki korelasyon katsayısı nasıl bir ölçüdür.

- a) Sınavın geçerliliğine dair bir ölçüdür
- b) Cevap anahtarlarının güvenilirliğine dair bir ölçüdür
- c) Soruların zorluk derecesini gösteren bir ölçüdür
- d) Öğrencilerin bilgi güvenilirliğine dair bir ölçüdür
- e) Öğretmenin değerlendirme tutarlılığının bir ölçüsüdür.

Cevap: E

3. Aşağıdaki tanımlanan ölçme araçlarından hangisi diğerlerine göre daha güvenilir ölçüm sonuçları verir?

- a) Kısa yanıtli 5 soru
- b) Doldurma tipli 20 soru
- c) Doğru yanlış tipli 50 soru
- d) Beşer seçenekli 25 soru
- e) Beşer seçenekli 50 soru

Cevap: E

4. Bir test sonucunda KR-20 ve KR-21 indisleri hesaplanmıştır. Buna göre maddelerin zorluk dereceleri aynı olan test hangisi olabilir?
- a) Test-1: KR-20=0.60, KR-21=0.80
 - b) Test-2: KR-20=0.60, KR-21=0.70
 - c) Test-3: KR-20=0.60, KR-21=0.90
 - d) Test-4: KR-20=0.60, KR-21=0.50
 - e) Test-5: KR-20=0.60, KR-21=0.75

Cevap: Her zaman $KR-20 > KR-21$??

5. Sınava girmeniz gerektiğinde okulunuz size 2 testten birini alma olanağı sunsa, siz bunlardan hangisini alırsınız?

a) Birinci testin çok yüksek geçerliliği var ama güvenirliliği ile ilgili hiçbir bilgi verilmiyor,

b) Birinci testin çok yüksek güvenirliği var ama geçerliliği ile ilgili hiçbir bilgi verilmiyor.

Cevap: A

MADDE ANALİZİ

Uygulanan testlerin güvenilir ve geçerli olmasının yanında, bir testte yer alan maddeler(sorular) ile ilgili bir takım analizler yapılarak testte yer alan soruların zorluk derecesi ve bilenle bilmeyeni ayırma gücü yani kalitesi belirlenmeye çalışılır. Bu amaçla madde güçlük indeksi ve madde ayırt edicilik indeksi kullanılır.

Madde Analizi Nedir?

Madde analizi testin bir gruba uygulandıktan sonra, istatistik teknikler yardımıyla test kapsamındaki maddelerin tek tek analiz edilmesidir.

Testin maddeleri ve seçenekleri üzerinde yapılan çalışmaya madde analizi denir.

Bir testte yer alan maddelerin uygulamasından elde edilen sonuçlarının seçilen ölçüte göre işe yarayıp yaramadığını, işe yaramıyorsa bunun muhtemel nedenlerini anlamak ve amaca hizmet etmesini sağlamak için gerekli düzeltmeleri yapmaya “**Madde Analizi**” denir.

Kısaca, hazırlanan bir testin uygulandıktan sonra madde güçlüğüne, soru maddesinin ayırt ediciliğine ve seçeneklerin çeldiriciliğine bakılmasına testin “**madde analizi**” denir.

MADDE ANALİZİ

Madde analizi temelde, belli özelliklere sahip olması gereken bir test içinde yer alacak maddeleri (soruları) irdelemeyi amaçlar. Belli bir amaca yönelik kullanılacak testin son halinin içereceği maddeleri analiz edebilmek için her şeyden önce ön uygulama yapılması ve uygulama sonuçlarının alınması gerekir.

Bir testte yer alan maddelerin uygulamasından elde edilen sonuçlarının belirlenen ölçüte göre işe yarayıp yaramadığını saptamak, işe yaramıyorsa bunun muhtemel nedenlerini anlamak ve testin geliştirilme amacına hizmet etmesini sağlamak için yapılması gerekenleri belirlemek ve gerekli düzeltmeleri yapmaya “madde analizi” denir (Benson, 1977; Ebel ve Frisbie, 1986; McDonald, 1999; Crocker ve Algina, 2008; De Grutijter ve Van der Kamp, 2008; Brennan, 2011).

Test ve Madde Analizlerinin Amacı

- Test ve maddelerin geçerliğini denetlemek, test ve maddelerin amaca uygunluğunu belirlemek
- Test ve maddelerin güvenirliğini denetlemek
- Test ve maddelerin güçlük derecesini belirlemek
- Test ve maddelerin kalitesini belirlemek, test ve maddelerin ayırta edicilik gücünü belirlemek
- Çeldiricilerin iyi çalışıp çalışmadığını belirlemek

Madde Analizi Ne İşe Yarar?

- Madde analizinin amacı, ölçülmek istenilen öğrenme alanı için **güvenirliği ve geçerliği yüksek** verimli bir ölçeğin elde edilmesidir.
- Elimizde geçerliği ve güvenilirliği yüksek bir sınav olsa bile uygulandığı sınıfı daha iyi tanıyabilmesi için **tekrar madde analizi** yapılmasında yarar vardır.
- Madde analizi yoluyla, o sınıftaki **öğrencilerin eksikleri, öğrenme güçlükleri ve yetersiz kaldıkları noktaları** daha iyi görülme fırsatı bulunmuş olur.

- **Madde Analiziyle...**

- Maddenin güçlük düzeyi** (yani ne kadar doğru cevaplandığı)
- Maddenin ayırt ediciliği** (yani bilenle bilmeyeni ne kadar iyi ayırdığı)
- Çeldiricilerin işlerliği** (Çoktan seçmeli maddeler için) hakkında bilgi sahibi olunur.

- **Madde Analizinin Diğer Yararları?**

- Madde analizi ile ilgili veriler, test sonuçlarının sınıfta etkili bir şekilde tartışılması için uygun bir zemin oluşturur.

- Madde analizi verileri, öğrencilerin zayıf yönlerinin iyileştirilmesi çalışmalarına katkı yapar.

- Madde analizi verileri, öğretim sürecinin iyileştirilmesi için bir zemin oluşturur.

- Madde analiz işlemleri, test oluşturma becerisini geliştirebilir.

Testte;

%10 Çok zor güçlükte soru sorulmalıdır

%20 Zor güçlükte soru sorulmalıdır

%40 Orta güçlükte soru sorulmalıdır

%20 Kolay soru sorulmalıdır

%10 Çok kolay soru sorulmalıdır

Madde Analizine başlamadan önce dikkat edilmesi gerekenler:

- ✓ Puanlar sürekli ise verilerin dağılışına bakmak (normal, basık, sivri, çarpık olup olmadığını incelemek).
- ✓ Verilerde aykırı değer durumuna bakmak.
- ✓ Maddeler arası korelasyon matrisine bakmak. Eğer maddeler arasında +1'e yakın korelasyon varsa, o maddelerden birine gerek yok demektir. Bu durum madde analizini etkiler, iç tutarlılığı ve faktörlerin açıklanan varyansını şişirir. Maddeler arası negatif korelasyon birbirinden ayık yapıların işaretini verir.
- ✓ Madde atmadan önce ve madde atıldıktan sonra ölçeğin standart sapmasının incelenmesi gerekir.

Madde Geerliđi

Madde Analizinde maddelerin geerli olması gerekir. Madde geerliđi, bir madenin üretilme amacına hizmet etme derecesini gösterir. Madde hangi kapsam için üretilmiş ise o kapsam içinde ilgili özelliđe sahip olanlarla olmayanları ayırma gücüne sahip olmalıdır.

Geliştirilen ölçek bir sınır yeterliliđi ölçüyorsa o madde bilenle bilmeyeni,; bir tutum ölçüyorsa olumsuz tutum ile olumlu tutumu ayırt ediyorsa o ölçek geerlidir.

Madde geerliđine madde ayırt ediciliđi de denir. Ölçeđin geerliđi madde geerliđinin bir fonksiyonudur.

Madde geerliđi iç ölçüte ve dış ölçüte göre incelenebilir. Eğer ölçüt ölçeđin dışındaki bir ölçüt ise «dış ölçüt»; eđer testin kendi istatistikleriyse «iç ölçüt» olarak adlandırılır.

Madde Geçerliđi

1.) İ Ölüte Gre Madde Geerliđi

Bir leđin i lütü, o lekten elde edilen toplam puanlar ve madde puanları alınır. Toplam puanların maddenin geerliđi iin lüt alınmasının temelinde, denenen maddelerin testle lülmek istenen zelliđi geerli olarak lmesidir (Turgut, 1979).

i) Madde Toplam lek Korelasyon Yöntemiyle Madde Geerliđi

lekteki her bir madde ile leđin tümüne verilen cevapların toplam puanı arasında pozitif ve dođrusal yönde bir iliđki gösteriyorsa o madde ayırt edici (geerli) kabul edilir ve o madde teste dahil edilir. Bu yöntem sadece homojen test veya alt testlere uygulanır. Örneđin bir başarı testi matematik, fizik, kimya bilgisini lüyorsa, maddelerin toplam puanla korelasyonu negatif çıkabilir. Bu gibi durumlarda ayrı ayrı her alt test iin madde-toplam test korelasyonuna bakılması gerekir. Matematik ile maddelerin matematik alt testinden alınan toplam puanla korelasyonuna bakılması gerekir.

Negatif korelasyon gösteren maddeler varsa; ters madde durumu olup olmadığına ve faktör analizine bakılmalıdır.

Madde ve madde toplam korelasyonlarına veri durumuna göre Bi-serial korelasyon, Nokta bi-serial korelasyon, Pearson Korelasyon, Spearman Sıra Korelasyon ile bakılabilir.

ii) Alt-Üst %27'lik Grup Ortalamalarının Karşılaştırılmasıyla Madde Geçerliği

Madde geçerliğinde bir diğer yöntem %27'lik uç grupların ortalamalarının madde bazında karşılaştırılmasıdır (Flanagan, 1952).

Test puanları sürekli biçimde ise iç tutarlılık için testin toplam puanlarına göre oluşturulan alt%27 ve üst%27'lik grupların madde ortalama puanları arasındaki farkları test edilir. Sonucun istatistiksel olarak anlamlı çıkması testin/ölçeğin iç tutarlılığını gösterir.

Analiz sonucu maddelerin bireyleri ölçülen davranış bakımından ne derece ayırt ettiğini gösterir.

Alt-Üst %27'lik Gruplara İlişkin Madde Ortalamalarının Karşılaştırılması

Madde ayırdediciliği yada madde geçerliliği için toplam ölçek puanları en büyükten en küçüğe doğru sıralandıktan sonra üstten ve alttan %27'lik gruplar belirlenerek; ölçekteki maddeler bağımsız iki grup t-testi ile test edilir.

$$t = \frac{\bar{x}_{ust} - \bar{x}_{alt}}{\sqrt{\frac{s_{ust}^2}{n_{ust}} + \frac{s_{alt}^2}{n_{alt}}}}$$

H_0 : Alt ve Üst %27'lik madde puanları arasında fark yoktur.

H_1 : Alt ve Üst %27'lik madde puanları arasında fark vardır.

Madde ayırt ediciliđi aısından bu gruplar arasında istatistiksel olarak anlamlı farklılık ıkması ve t deęerinin (-) iřaretli olmaması istenir.

Üst gruptaki cevaplayıcılar öleđin ölçmek istediđi özelliđe olumlu yönde sahip olanlar, alt gruptaki cevaplayıcılar ise öleđin ölçmek istediđi özelliđe olumsuz yönde sahip olanlardır.

Bir madde için bu iki grup arasında istatistiksel aıdan fark yok ise ($p > 0,05$) bu madde ilgili özelliđe sahip olanlar ile olmayanları ayırt edemeyecektir denilir. Dolayısıyla bu madde öleđe alınmayacaktır.

Test toplam puanlarına göre oluşturulan alt%27 ve üst%27 'lik grupların madde ortalama puanlarının testi

	alan	og_tur	unive rsite	yontem1	yontem2	yontem3	yontem4	yomtem 5	yontem6	yontem7	yontem8	yontem9	yontem1 0	toplam puan	grup
1	sosyal	gundu	omü	1	1	1	1	1	1	1	1	1	3	12	alt%27
2	sosyal	gece	ktu	1	1	1	2	2	1	1	1	1	1	12	alt%27
3	fen	gece	omü	2	1	1	2	2	1	1	1	1	1	13	alt%27
4	sosyal	gece	gazi	2	1	1	1	1	2	1	2	1	2	14	alt%27
5	sosyal	gundu	omü	1	2	2	3	1	1	1	1	1	1	14	alt%27
								2	1	2	1	1	2	15	alt%27
								1	3	1	1	1	3	15	alt%27
								1	1	2	1	1	1	15	alt%27
								1	1	1	1	1	1	15	alt%27
								2	2	1	1	1	2	15	alt%27

Group Statistics

gr	N	Mean	Std. Deviation	Std. Error Mean
yontem1 alt%27	94	2,43	,886	,091
yontem1 üst%27	94	3,94	,827	,085
yontem2 alt%27	94	2,51	,852	,088
yontem2 üst%27	94	4,17	,713	,074
yontem3 alt%27	94	2,48	,852	,088
yontem3 üst%27	94	4,18	,604	,062
yontem4 alt%27	94	2,14	,824	,085
yontem4 üst%27	94	4,23	,739	,076
yomtem5 alt%27	94	2,04	,687	,071
yomtem5 üst%27	94	3,95	,678	,070
yontem6 alt%27	94	2,13	,883	,091
yontem6 üst%27	94	3,95	,694	,072
yontem7 alt%27	94	2,10	,843	,087
yontem7 üst%27	94	4,21	,637	,066
yontem8 alt%27	94	1,79	,760	,078
yontem8 üst%27	94	3,56	,899	,093
yontem9 alt%27	94	1,81	,976	,101
yontem9 üst%27	94	3,87	1,090	,112
yontem10 alt%27	94	2,23	,966	,100
yontem10 üst%27	94	4,35	,851	,088

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means		
		F	Sig.	t	df	Sig. (2-tailed)
yontem1	Equal variances assumed	4,905	,028	-12,085	186	,000
	Equal variances not assumed			-12,085	185,130	,000
yontem2	Equal variances assumed	9,704	,002	-14,484	186	,000
	Equal variances not assumed			-14,484	180,404	,000
yontem3	Equal variances assumed	17,909	,000	-15,806	186	,000
	Equal variances not assumed			-15,806	167,659	,000
yontem4	Equal variances assumed	,000	,985	-18,348	186	,000
	Equal variances not assumed			-18,348	183,844	,000
yontem5	Equal variances assumed	,119	,730	-19,136	186	,000
	Equal variances not assumed			-19,136	185,970	,000
yontem6	Equal variances assumed	8,747	,004	-15,711	186	,000
	Equal variances not assumed			-15,711	176,137	,000
yontem7	Equal variances assumed	4,795	,030	-19,420	186	,000
	Equal variances not assumed			-19,420	173,071	,000
yontem8	Equal variances assumed	3,401	,067	-14,633	186	,000
	Equal variances not assumed			-14,633	181,021	,000
yontem9	Equal variances assumed	,134	,714	-13,678	186	,000
	Equal variances not assumed			-13,678	183,772	,000
yontem10	Equal variances assumed	1,453	,230	-15,937	186	,000
	Equal variances not assumed			-15,937	183,094	,000

Soru. 30 kiřiye uygulanan 10 maddelik ve her madde 8 seenekli (En olumsuzdan-en olumluya doęru sıralanan) bir leęin sonuları ařaęıdaki gibi bulunmuřtur. Maddelerin ayırdedicilik zellięini inceleyiniz. lekten ıkarılması gereken madde olup olmadıęına karar veriniz.

	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	toplam
1	5	4	5	5	4	8	6	5	5	4	51
2	5	3	4	3	5	6	3	4	5	3	41
3	4	5	5	5	3	3	4	4	3	4	40
4	5	4	5	2	6	5	4	6	4	4	45
5	6	6	4	5	4	4	4	6	3	5	47
6	4	4	4	4	4	5	6	5	3	4	43
7	5	7	4	5	5	7	7	6	3	4	53
8	6	4	5	6	3	3	4	3	6	5	45
9	6	6	6	4	8	5	6	6	4	5	56
10	3	5	4	3	3	5	6	4	4	5	42
11	7	4	5	4	7	5	7	5	4	4	52
12	1	2	2	3	3	2	2	2	3	1	21
13	7	6	7	7	4	4	5	5	6	5	56
14	5	5	4	6	6	5	5	6	6	6	54
15	6	7	7	4	6	3	8	7	5	6	59
16	5	6	6	4	4	7	6	6	5	6	55
17	3	5	5	3	5	3	4	4	6	3	41
18	7	6	7	6	6	6	6	7	6	4	61
19	3	4	5	2	3	6	4	4	3	4	38
20	4	4	3	3	4	5	2	1	4	1	31
21	6	6	5	5	3	6	5	4	7	5	52
22	4	5	5	5	6	3	6	4	4	5	47
23	4	5	4	3	2	4	5	5	5	5	42
24	4	5	2	3	3	2	3	4	2	4	32
25	6	7	7	5	5	6	6	5	5	5	57
26	6	6	3	3	5	7	4	6	5	6	51
27	6	6	6	5	5	4	5	5	6	5	53
28	3	4	2	2	4	7	3	3	3	3	34
29	4	4	2	4	5	3	5	2	4	4	37
30	8	7	6	5	3	2	7	5	4	5	52

M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	toplam	grup
7	6	7	6	6	6	6	7	6	4	61	Üst%27
6	7	7	4	6	3	8	7	5	6	59	Üst%27
6	7	7	5	5	6	6	5	5	5	57	Üst%27
6	6	6	4	8	5	6	6	4	5	56	Üst%27
7	6	7	7	4	4	5	5	6	5	56	Üst%27
5	6	6	4	4	7	6	6	5	6	55	Üst%27
5	5	4	6	6	5	5	6	6	6	54	Üst%27
5	7	4	5	5	7	7	6	3	4	53	Üst%27
6	6	6	5	5	4	5	5	6	5	53	..
7	4	5	4	7	5	7	5	4	4	52	..
6	6	5	5	3	6	5	4	7	5	52	..
8	7	6	5	3	2	7	5	4	5	52	..
5	4	5	5	4	8	6	5	5	4	51	..
6	6	3	3	5	7	4	6	5	6	51	..
6	6	4	5	4	4	4	6	3	5	47	..
4	5	5	5	6	3	6	4	4	5	47	..
5	4	5	2	6	5	4	6	4	4	45	..
6	4	5	6	3	3	4	3	6	5	45	..
4	4	4	4	4	5	6	5	3	4	43	..
3	5	4	3	3	5	6	4	4	5	42	..
4	5	4	3	2	4	5	5	5	5	42	..
5	3	4	3	5	6	3	4	5	3	41	..
3	5	5	3	5	3	4	4	6	3	41	Alt%27
4	5	5	5	3	3	4	4	3	4	40	Alt%27
3	4	5	2	3	6	4	4	3	4	38	Alt%27
4	4	2	4	5	3	5	2	4	4	37	Alt%27
3	4	2	2	4	7	3	3	3	3	34	Alt%27
4	5	2	3	3	2	3	4	2	4	32	Alt%27
4	4	3	3	4	5	2	1	4	1	31	Alt%27
1	2	2	3	3	2	2	2	3	1	21	Alt%27

Toplam puanlar büyükten küçüğe sıralanır ve Üst %27'lik dilim ile alt%27'lik dilim seçilir. $30 \cdot 0,27 = 8$ madde üstten ve alttan seçilir.

Group Statistics

	grup	N	Mean	Std. Deviation	Std. Error Mean
M1	Üst%27	8	5,88	,835	,295
	Alt%27	8	3,25	1,035	,366
M2	Üst%27	8	6,25	,707	,250
	Alt%27	8	4,13	,991	,350
M3	Üst%27	8	6,00	1,309	,463
	Alt%27	8	3,25	1,488	,526
M4	Üst%27	8	5,13	1,126	,398
	Alt%27	8	3,13	,991	,350
M5	Üst%27	8	5,50	1,309	,463
	Alt%27	8	3,75	,886	,313
M6	Üst%27	8	5,38	1,408	,498
	Alt%27	8	3,88	1,885	,666
M7	Üst%27	8	6,13	,991	,350
	Alt%27	8	3,38	1,061	,375
M8	Üst%27	8	6,00	,756	,267
	Alt%27	8	3,00	1,195	,423
M9	Üst%27	8	5,00	1,069	,378
	Alt%27	8	3,50	1,195	,423
M10	Üst%27	8	5,13	,835	,295
	Alt%27	8	3,00	1,309	,463
toplam	Üst%27	8	56,38	2,615	,925
	Alt%27	8	34,25	6,453	2,282

Madde	t-test	P
M1	5,584	,000
M2	4,937	,000
M3	3,924	,002
M4	3,771	,002
M5	3,130	,007
M6	1,803	,093
M7	5,358	,000
M8	6,000	,000
M9	2,646	,019
M10	3,871	,002
Toplam	8,987	,000

6.Madde anlamsız bulunmuştur ($p>0,05$). Bu sonuç 6.maddenin işlemediğini gösterir. Yani bu madde ölçeğe ait olumlu ve olumsuz görüşü ayırt edememektedir. Ölçekten çıkarılmalıdır.

Eğer t değeri (-) işaretli bir madde olsaydı bu madde ters bir madde olup olmadığına bakılır. Ters madde değilse bu soru ölçekten çıkarılmalıdır.

iii) Regresyon Yöntemiyle Madde Geçerliđi

Bu yöntemde madde puanları yordayıcı (predictor-açıklayıcı-bağımsız) deđişken, toplam ham puanlar yordanan (bağımlı) deđişken alınır ve tüm maddelere basit doğrusal regresyon analiz yapılır. Anova testinde F deđeri anlamlı olan maddeler geçerli kabul edilir.

Madde Geerliđi

2.) Dış Ölüte Gre Madde Geerliđi

Dış ölüt öleđin zellik ile iliřki gsteren bir bařka deđiřkene iliřkin lmler olabilir. rneđin depresyonu lmek iin Beck depresyon leđi kullanılıyorsa, bireylerin aldıkları puanlar ile yeni geliřtirilen leđin maddelerinden aldıkları puanlar arasında pozitif ve yksek korelasyon varsa o maddeler ayırt edici zelliđe sahip yani geerli maddelerdir denilebilir.

Madde ve dış ltn lme dzeyine gre uygun istatistiksel yntem kullanılır.

Madde analizinde kullanılan iki yöntem vardır.

1. Basit Yöntem
2. Henryson Yöntemi

Bu iki madde analizi yöntemleri arasındaki temel fark hesaplamaya dahil edilen ölçmelerdir.

Henryson yönteminde tüm yanıtlayıcıların puanları hesaplamaya dahil edilirken, basit yöntemde yalnızca üst %27 ve alt %27'lik dilimlerde bulunan yanıtlayıcıların puanları işlemlere dahil edilir. Basit yöntem alt ve üst gruptaki puanların farkına dayalıdır. Basit yöntem gruptaki ölçme sayısının 300'den büyük olduğu durumlarda güvenilir sonuçlar vermektedir.

Henryson yöntemi ile ise ölçmelerin sayısı 60 kişi olsa bile güvenilir sonuçlar alınabilmektedir. Gruplardaki birey sayısı arttıkça her iki yöntemle elde edilen sonuçlar birbirlerine yaklaşmaktadır.

MADDE ANALİZİ

1. BASİT YÖNTEM

Basit yöntemde önce teste verilen cevaplar puanlanır ve doğru cevap sayıları ile test puanları bulunur. Doğru cevaplar «1», yanlış cevaplar «0» olarak puanlanır. Doğru cevap puanları ile test puanları bulunur.

Testteki bireyler en yüksek puandan en düşük puana doğru sıralanır.

En yüksek ve en düşük puanların %27'si seçilerek üst grup ve alt grup oluşturulur. Aradaki puanlar dikkate alınmaz.

Örneğin 378 öğrencinin katıldığı bir sınavda, cevaplar en yüksek puandan en düşük puana sıralandıktan sonra, %27'lik üst ve alt gruplar aşağıdaki gibi belirlenir.

	Sıra	Puanlar
1	100	
2	98	
3	95	
...		
102	80	%27'lik üst üst grup
...	...	%46'lık ortada kalan grup
1	25	
2	24	
3	23	
...		
102	10	%27'lik üst alt grup

Üst ve alt grupta ayrı ayrı ilgili maddeye verilen cevaplardan tüm seçeneklere konulan işaretler, boş bırakılanlar sayılır ve sayım sonuçları aşağıdaki gibi bulunur /(Özgüven, 1994).

Madde Seçenek Analizi

Madde No	A	B	C	D*	E	Boş	Toplam
Üst Grup	n(A _ü)	n(B _ü)	n(C _ü)	n(D _ü)	n(E _ü)	n(Boş- _ü)	N _{üst}
Alt Grup	n(A _a)	n(B _a)	n(C _a)	n(D _a)	n(E _a)	n(Boş-a)	N _{alt}
Toplam	n(A)	n(B)	n(C)	n(D)	n(E)	n(Boş)	N

* Doğru cevap (D seçeneğini işaretleyenler 1, diğerleri 0 puan alacaktır.)

Madde analiz yapıldıktan sonra maddelerin seçimine geçilir. Madde seçiminde madde güçlük indeksi ve madde ayırt edicilik indeksi yeter derecede olanlar seçilir. Madde seçiminde madde güçlük indeksi 0,5 ve civarında olanlar ve madde ayırt edici indeksi $\geq 0,3$ olan maddeler seçilebilir. Madde ayırt edicilik indeksi 0,2-0,29 arasında olanlar düzeltilip geliştirilerek nihai teste dahil edilebilir.

MADDE İSTATİSTİKLERİ

- * Testte yer alan maddeler ile ilgili bir takım analizler yapılarak testte yer alan soruların zorluk derecesi ve bilenle bilmeyeni ayırma gücü yani; kalitesi belirlenmeye çalışılır.
- * Dolayısıyla grubun testi oluşturan maddelere vermiş olduğu yanıtların sayısal göstergelerinden maddelerin işe yarayıp yaramadığı anlaşılır; işe yaramıyorsa bunun nedenini araştırılarak ve ona göre gerekli düzeltmeler yapılabilir.

**MADDE GÜÇLÜK
İNDEKSİ**

**MADDE AYIRT
EDİCİLİK İNDEKSİ**

Madde Güçlük İndeksi (Item Difficult Index-P)

- Madde güçlük indeksi (P), yetenek testleri, başarı testleri gibi bilgi ve becerilerin ölçüldüğü testlerde yer alan maddelerin doğru cevaplanma oranını göstermektedir. Madde güçlük indeksi 0 ile 1 arasında değerler alır.
- **Madde güçlük indeksi bir maddenin kolay ya da zor bir madde olup olmadığı hakkında bilgi verir. Güçlük indeksinin 0'a yaklaşması maddenin zorlaştığını, 1'e yaklaşması maddenin kolaylaştığını, 0.50 olması maddenin orta güçlükte olduğunu gösterir (Tekin, 2000).**

MADDE GÜÇLÜK İNDEKSİ (P, P_j)

- Maddeyi doğru cevaplayanların tüm cevaplayıcı sayısına oranıdır. Oranlama yapıldığında sınıfın yüzde kaçının bu soruyu doğru yanıtladığı görülebilir.

%27'lik Alt ve Üst Gruplar Yöntemi (Basit Yöntem)

Madde güçlük indeksi, sınava katılan herkesi değil sadece en başarılı ve en başarısız olan öğrenci grupları dikkate alınarak hesaplanabilir. Bunun için önce puanlar sıraya dizilir ve testte en başarılı olmuş %27'lik grup üst grup olarak, en başarısız olmuş %27'lik grup da alt grup olarak belirlenir. Daha sonra bu gruplar içinde, ilgili soruya doğru cevap veren öğrencilerin sayısı bulunur ve aşağıdaki formül kullanılarak maddenin güçlük indeksi hesaplanır.

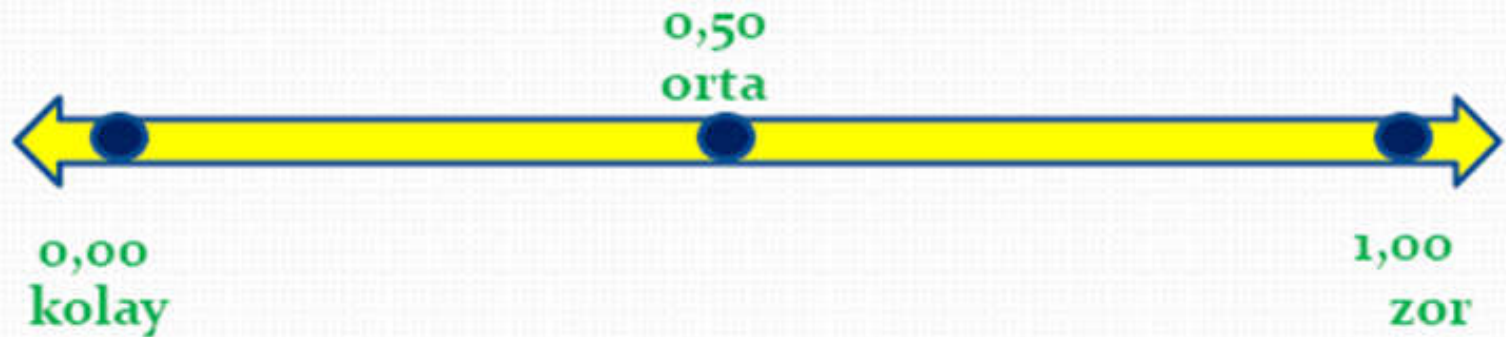
$$P_j = \frac{n(dü) + n(da)}{2n}$$

$n(dü)$: %27'lik üst grupta doğru cevap verenlerin sayısı

$n(da)$: %27'lik alt grupta doğru cevap verenlerin sayısı

n : Üst ya da alt gruptan herhangi birisindeki öğrenci sayısıdır.

- Bir maddeyi doğru cevaplama yüzdesi veya oranı olarak da tanımlanabilir.
- Maddenin doğru yapılma yüzdesidir.
- Maddenin kolaylığını ya da zorluğunu gösterir.
- Maddeyi doğru cevaplayan öğrencilerin oranını
- Madde güçlük indeksi 0,00 ile 1,00 arasında değer
- 0,00'a yaklaştıkça madde zor, 1,00'a yaklaştıkça madde kolay, 0,40 ile 0,60 arasında ise madde güçlüğü orta düzeydedir.



* Bir maddenin güçlük indeksinin 0,50 olmasını isteriz; fakat bir bir testteki tüm maddelerin güçlüklerinin 0,50 olmasını tercih edilmez. Bunun yerine testin bir bölümü zor, bir bölümü kolay, bir bölümü de orta güçlükte hazırlanır. Böylece testin ortalama güçlüğü 0,50 olur. Soruların güçlük dağılımının da normal dağılım olması istenir.

Madde Güçlük İndeksi

Maddenin Değerlendirilmesi

0,29 ve altında bulunan maddeler



ZOR

0,30 ve 0,49 arasında bulunan maddeler



ORTA GÜÇLÜKTE

0,50 ve 0,69 arasında bulunan maddeler



KOLAY

0,70 ve 1,00 arasında bulunan maddeler



ÇOK KOLAY

Madde Güçlüğü Nasıl Yorumlanır?

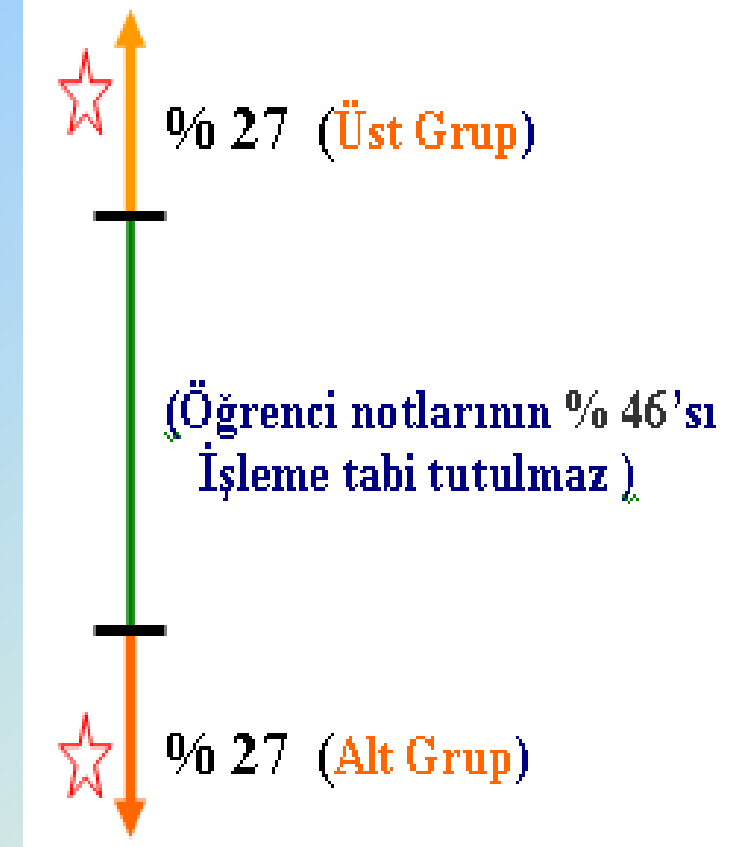
- $0 \leq p_j \leq 1$
- ▶ Madde güçlüğü 0 ile 1 arasında değerler alır!
 - ▶ Bir soruyu 50 kişilik bir sınıfta herkes doğru cevaplamışsa madde güçlüğü $p=50/50=1$ olarak bulunur. Eğer herkes doğru cevap veriyorsa bu soru çok kolay bir soru demektir.
 - ▶ Bir soruya hiç kimse doğru cevap veremiyorsa madde güçlüğü $p=0/50=0$ olarak bulunur. Eğer hiç kimse doğru cevap veremiyorsa bu soru çok zor bir sorudur.
 - ▶ Eğer sınıfın yarısı doğru cevap vermişse o maddenin güçlüğü $p=25/50=0,5$ olarak bulunur ve madde orta derecede güçlüğüne sahip denir.
 - ▶ Madde güçlük değeri 0'a yaklaştıkça madde zorlaşır eğer madde güçlüğü 1'e yaklaşıyorsa madde kolaydır denir.

%27 'lik Alt ve Üst Gruplar Yöntemi ile Madde Güçlük İndeksi Hesaplaması

Öncelikli olarak puanlar sıraya dizilir ve en başarılı %27'lik grup **üst grup**; en başarısız %27'lik grup da **alt grup** olarak belirlenir. Maddeye verilen cevap **doğru** ise **1 yanlış** veya **boş** ise **0** puan verilerek ham puan hesaplanır. Aşağıdaki şekliyle tabloya aktarılır.

I. soru	A	B	C*	D	E	Toplam
Üst grup	10	5	20	10	5	50
Alt grup	14	12	8	7	9	50

* Anahtarlanmış yanıt



Örnek 1. 500 öğrencinin katıldığı 20 soruluk bir test sınavını düşünelim. Bu sınavdaki örneğin **yedinci soruya** üst grupta yer alan öğrencilerin 80 tanesi, alt grupta yer alan öğrencilerin 30 tanesi doğru cevap vermiş olsun. Yedinci sorunun güçlük indeksi kaçtır?

Elimizde 500 tane puan olacaktır. Önce bu puanları sıraya dizilir. Daha sonra, $500 \times (0.27) = 135$ tane en başarılı (üst grup) ve 135 tane en başarısız (alt grup) öğrenciyi tespit ettikten sonra, her grup için yedinci soruya doğru cevap verenlerin sayısı bulunur.

$n = 135$ olur. $n(dü) = 80$ ve $n(da) = 30$ verilmiş.

O halde yedinci sorunun güçlük indeksi,

$$p_j = p_7 = \frac{n(dü) + n(da)}{2n} = \frac{80 + 30}{2(135)} = 0.41$$

bulunur ve yedinci sorunun orta güçlükte bir soru olduğu sonucuna varılır.

Madde güçlük indeksinin aldığı değerlere göre maddenin değerlendirilmesi aşağıda gösterildiği biçimde yapılmaktadır

Madde Güçlük İndeksi

Maddenin Değerlendirilmesi

0.29 ve altında ise

ZOR

0.30 ve 0.49 arasında ise

ORTA GÜÇLÜKTE

0.50 ve 0.69 arasında ise

KOLAY

0.70 ve 1.00 arasında ise

ÇOK KOLAY

ÖRNEK 2.

Madde No 4	A	B	C*	D	E	Toplam
Üst Grup	15	20	25	25	15	100
Alt Grup	20	15	15	22	28	100

$$P_j = (25 + 15) / (2 * 100)$$

= 0,20 Bu madde zor bir maddedir.

Örnek 3. Aşağıda 100 kişilik bir sınıfa uygulanan testin 4. maddesine verilen cevapların seçeneklere göre dağılımı verilmektedir. Üst ve alt gruplarda 27 şer ($100 \cdot 0,27$) öğrenci bulunmaktadır. Buna göre maddenin güçlük indeksi nedir.

Gruplar	Seçenekler				
	A	B	C*	D	E
Üst Grup (27)	3	5	16	1	2
Alt Grup (27)	5	6	9	4	3

* sorunun doğru cevabını göstermektedir.

$$\text{Madde güçlük indeksi (p)} = (16+9)/54 = 0,46$$

4. Madde orta düzeyde bir maddedir.

Madde Ayırt edicilik İndeksi (Item Discrimination Index-rjx)

- **Madde ayırt edicilik indeksi, maddelerin ölçülen özellikle ilgili olarak bireyleri ne derece ayırt ettiğini gösterir.** Diğer bir deyişle, testin ölçmeyi amaçladığı özelliğe yüksek düzeyde sahip olan bireylerle, düşük düzeyde sahip olan bireyleri ayırt etme gücüdür.
- Madde ayırt edicilik indeksi -1 ile +1 arasında değişebilir. Bu değer negatif olması, maddenin ölçülen özellik bakımından bireyleri ters ayırt ettiğini gösterir. Bu nedenle, bu tür maddeler testten çıkarılmalıdır (Büyüköztürk ve diğerleri, 2010).
- Madde ayırt edicilik değeri, testin toplam puanlarına göre oluşturulan alt %27 ve üst %27'lik grupların madde ortalama puanları arasındaki farkların bağımsız t-testi kullanılarak test edilmesiyle bulunabilir. Gruplar arasında gözlenen farkların anlamlı çıkması, testin iç tutarlılığının bir göstergesi olarak değerlendirilebilir (Büyüköztürk, 2011).

MADDE AYIRT EDİCİLİK İNDEKSİ

Madde ayırt edicilik indeksi hesaplanırken de %27 lik üst ve alt gruplar göz önünde bulundurulur ve aşağıdaki formül kullanılır:

$$r_{jx} = \frac{n(dü) - n(da)}{n}$$

$n(dü)$: %27'lik üst grupta doğru cevap verenlerin sayısıdır.

$n(da)$: %27'lik alt grupta doğru cevap verenlerin sayısıdır.

n : Üst ya da alt gruptan herhangi birisindeki öğrenci sayısıdır.

Madde ayırt edicilik indeksinin aldığı değerlere göre karar, aşağıdaki biçimde verilmektedir:

Madde Ayırt Edicilik İndeksi

Maddenin Değerlendirilmesi

0.40 ve daha büyük

Çok iyi madde (Ayırt etme gücü yüksek)

0.30 ve 0.39

Oldukça iyi madde. Yine de geliştirilebilir

0.20 ve 0.29

Düzeltilmesi ve geliştirilmesi gerekir

0.19 ve daha küçük

Çok zayıf madde. Ayırt etme gücü düşük. Mutlaka testten atılmalıdır

Negatif değer

Çok kötü bir maddedir. Düzeltilerek geliştirilemiyorsa kesinlikle testten alınmalıdır.

MADDE AYIRT EDİCİLİK İNDEKSİ

- Bir maddenin bilenle bilmeyeni ne derece ayırdığını gösterir. Bir maddenin kalitesi; yani ne kadar iyi işlediği, bilen (başarılı) öğrenci ile bilmeyen (başarısız) öğrenciyi birbirinden ayırmasıdır.



Pozitif (+) olması; Ortalamanın (%50) altında kalan (başarısız) öğrenciler tarafından bilinmeyip, ortalamanın üstünde kalan (başarılı) öğrenciler tarafından bilinmesi.

Negatif (-) olması; Ortalamanın (%50) üstünde kalan (başarılı) öğrenciler tarafından bilinmeyip, ortalamanın altında kalan (başarısız) öğrenciler tarafından bilinmesi.

Madde Ayırt Ediciliği Nasıl Yorumlanır?

- ▶ Madde ayırt ediciliği -1 ile 1 arasında değerler alır.
- ▶ Diyelim ki bir sınıfta 50 kişi vardır. Eğer üst gruptaki ve alt gruptaki herkes bir maddeyi doğru bilirse madde ayırtediciliği $[(25-25)/25 = 0]$ `0` olur.
- ▶ Eğer üst gruptaki tüm öğrenciler maddeyi doğru cevaplarken alt gruptaki hiçbir öğrenci doğru cevap veremezse madde ayırt ediciliği $[(25-0)/25 = 1]$ `1` olur.> **İSTENİLEN DURUM.**
- ▶ Eğer üst gruptaki tüm öğrenciler maddeyi yanlış cevaplarken alt gruptaki tüm öğrenciler doğru cevap verirse madde ayırt ediciliği $[(0-25)/25 = -1]$ `-1` olur.> **İSTENİLMEYEN DURUM.**
- ▶ ***EĞER BİR MADDENİN AYIRT EDİCİLİĞİ 0,3 VE ÜZERİNDE BİR DEĞER ALIYORSA O MADDE (AYIRTEDİCİLİK BAKIMINDAN) İYİ BİR MADDEDİR DİYEBİLİRİZ!!!***

* Ayırt ediciliği negatif olan maddeler testten hemen çıkarılmaz. Çünkü çeşitli sebeplerle (örneğin; cevap anahtarı yanlışlığı) bilen öğrenciler soruyu yanlış cevaplamış olabilirler. İlgili yanlışlık düzeltilirse negatifken pozitive dönüşebilir. Fakat ister pozitif ister negatif olsun 0,00 – 0,20 aralığındaki tüm maddeler çıkarılmalı çünkü negatifi olanı düzeltseniz bile yine kalitesiz bir madde olarak kalacaktır.

* Madde güçlüğü (P_j) 0,00 ve 1,00 değerlerini aldığı anda madde ayırt ediciliği (r_{jx}) her zaman 0,00 olur. Neden?

Madde Ayırıcılık İndeksi

- Madde ayırt ediciliği **maddenin geçerliği** olarak da yorumlanır.
- Maddenin kalitesi hakkında bilgi verir ve sınavda ölçülecek davranışa **sahip olan ve olmayanları** birbirinden ayırma derecelerini gösterir.
- Madde ayırt ediciliği yüksek maddeler bir **testin güvenilirliğini ve geçerliğini artırır.**
- Alt ve üst grupta maddeyi doğru cevaplayanların sayıları arasındaki fark arttığı oranda (**üst grup lehine**), maddenin ayırıcılığı artar.
- Farkın kapanması ise **ayırt ediciliğini düşürür.**

Madde Ayırıcılık İndeksi

- Bir maddenin ayırt edicilik indeksinin negatif değerde çıkması için alt gruptaki doğru cevaplayıcıların sayısının üst gruptakilerden daha fazla olması gerekir.
- Böyle bir durumda sınıfın üst gruptaki öğrencilerinin alt gruptaki öğrencilerden daha zayıf olduğu, bu maddeyle ilişkili konuyu bilmedikleri gibi bir yorum çıkar ki, bu da mantıksız olur.
- Bu nedenle bu soru çok kötü bir soru olup testten çıkarılması gerekir.
- Ya da bu sorunun cevabı yanlış da verilmiş olabileme ihtimali de göz ardı edilmemelidir.

Çeldiriciler

- Doğru cevap seçeneđi dışındaki seçeneklere **çeldirici** denir.
- Üst grupta doğru cevaba giden öğrenci sayısı alt grupta doğru cevaba giden öğrenci sayısından fazla olmalıdır.
- Üst grupta çeldiricilere giden öğrenci sayısı alt grupta çeldiricilere giden öğrenci sayısından az olmalıdır.
- Çeldiricilere giden öğrenci sayısı dengeli olmalıdır.
- Hiçbir öğrenci tarafından işaretlenmeyen çeldiriciler yeniden düzenlenmelidir.

Çeldirici Analizi

1.madde	A	*B	C	D	E	
Üst grup	7	15	8	20	0	50
Alt grup	22	15	8	0	5	50
TOPLAM	29	30	16	20	5	100

Doğru cevap: B

- ▶ Bir maddeye verilen cevaplar tabloda gösterilmiştir.
- ▶ A, C, D ve E seçenekleri çeldirici olarak düşünüldüğünde çeldiricilerin daha çok alt grup öğrencilerini çeldirmesi gerekir.
- ▶ E seçeneği en iyi çeldiricidir çünkü sadece alt grup öğrencilerini çeldirmiş ve E cevabı vermelerini sağlamıştır.
- ▶ A seçeneği de E seçeneği kadar olmasa da iyi bir çeldiricidir çünkü alt grup öğrencilerini daha çok çeldirmiştir.
- ▶ C seçeneği 2 grubu da eşit oranda çeldirmiş (8'er)
- ▶ D seçeneği çok kötü bir çeldiricidir çünkü sadece üst grup üyelerini çeldirmiştir. Aslında sadece alt grup öğrencilerini yani bilmeyenleri çeldirmesi gerekliydi.

İyi bir test maddesinin sahip olması gereken özellikler

- Kaliteli bir sorunun güçlük indeksi 0,50 civarında, ayırt edicilik indeksi ise 0,40'ın üzerinde olması
- Madde güvenilirlik katsayısının 0,50'ye yakın olması
- Kaliteli bir soruda doğru cevabın üst grupta daha çok öğrenci tarafından cevaplandırılmış olması
- Yanlış cevabın(çeldiricilerin) alt grupta daha çok öğrenci tarafından cevaplandırılmış olması beklenir.
- Yanlış cevabın (çeldiricilerin) tüm seçeneklerde birbirine yakın oranda cevaplandırılmış olması

Bir Madde Ne Zaman Teste Alınır?

- ▶ Madde ayırt ediciliđi 1'e yakın olmalıdır. "İLK BAKILACAK ÖZELLİK"
- ▶ Madde güçlüğü ne çok düşük ne de çok yüksek olmalıdır (0,5 civarı olması gerekir).
- ▶ Çeldiricileri işaretleyen üst gruptaki öğrenci sayısının alt gruptaki öğrenci sayısından az olması gerekir.
- ▶ Çeldiricileri işaretleyen öğrenci sayısının olabildiğince dengeli bir dağılım göstermiş olması istenir.

Örnek 4. 500 öğrencinin katıldığı 20 soruluk bir test sınavını düşünelim. Bu sınavdaki örneğin **yedinci soruya** üst grupta yer alan öğrencilerin 80 tanesi, alt grupta yer alan öğrencilerin 30 tanesi doğru cevap vermiş olsun. Yedinci sorunun ayırt edicilik indeksi kaçtır?

$$n = 135, \quad n(dü) = 80, \quad n(da) = 30$$

$$r_{jx} = \frac{n(dü) - n(da)}{n} = \frac{80 - 30}{135} = 0.37$$

Bu sonuca göre testteki yedinci soru, oldukça iyi bir sorudur. Bilen ve bilmeyen öğrencileri ayırabilmektedir.

Örnek 5. Yapılan bir testte yer alan 5 sorunun madde güçlük indeksleri ve madde ayırt edicilik indeksleri aşağıda verilmiştir. Bu bilgilere göre şıklarda istenenlere cevap veriniz.

- Testten mutlaka çıkarılması gereken soru yada sorular var mı?
- Zor sorular hangileridir?
- Hem kolay hem de ayırt edici soru var mı?
- Hem zor hem de ayırt edici soru var mı?

Sorular	Güçlük indeksi (p_j)	Ayırt edicilik indeksi (r_{jx})
1	0.15	0.13
2	0.60	-0.17
3	0.23	0.89
4	0.85	0.83
5	0.90	0.35

Cevap a) Eđer sorularda d¼zeltme yapılamıyorsa; testten mutlaka ıkarılması gereken(sorulmaması gereken) sorular, madde ayırt edicilięi 0.19 dan d¼ş¼k olan sorulardır. Verilenlere baktıęımızda, **1. ve 2. soruların** ayırt edicilik indeksleri 0.19'dan k¼çük olduęundan bu sorularda d¼zeltme yapılamıyorsa sorulmamalıdır.

Cevap b) Madde g¼çl¼k indeksi 0.29 ve altında olan sorular zor sorulardır.

Verilenlere baktıęımızda, 1. ve 3. soruların g¼çl¼k indeksleri 0.29'dan k¼çük olduęundan **1. ve 3. sorular ZOR** sorulardır.

Cevap c) Madde g¼çl¼k indeksi 0.50 ve 0.69 arasında olan sorular kolay sorulardır.

Madde ayırt edicilik indeksi 0.30'dan b¼y¼k olan sorular da ayırt edici sorulardır. Bu özelliklere sahip soru **YOKTUR.**

Cevap d) Madde g¼çl¼k indeksi 0.29 ve altında olan sorular zor sorulardır.

Madde ayırt edicilik indeksi 0.30'dan b¼y¼k olan sorular da ayırt edici sorulardır. Bu özelliklere sahip olan sadece **3. sorudur.**

Örnek 6. Bir testte elde edilen ilk iki sorunun sonuçları aşağıdaki gibidir. Bu iki sorunun madde güçlük ve madde ayırt edicilik indeksi değerlerini bulunuz.

100 Öğrenciye Uygulanan Test	Soru 1 Doğru Cevap Verenler	Soru 2 Doğru Cevap Verenler
Üst Grup (İlk %27) (27)	25	20
Alt Grup (Son %27) (27)	15	15
Madde Güçlük Endeksi (p)	$p = \frac{25+15}{54} = 0.74$	$p = \frac{20+15}{54} = 0.64$
Madde Ayırt Edicilik Endeksi (r)	$p = \frac{25-15}{27} = 0.36$	$r = \frac{20-15}{27} = 0.19$

Örnek 7. Aşağıda 60 kişilik bir sınıfa uygulanan testin 4. maddesine verilen öğrenci cevaplarının seçeneklere göre dağılımı verilmektedir. Üst ve alt gruplarda yaklaşık 16 ($60 \cdot 0,27$) öğrenci bulunmaktadır. Buna göre maddenin güçlük ve ayırt edicilik düzeylerini yorumlayınız.

Gruplar	Seçenekler				
	A	B	C*	D	E
Üst Grup (16)	2	5	7	1	1
Alt Grup (16)	4	4	2	4	2

$$\text{Madde güçlük indisi (p)} = (7+2)/32 = 0,28$$

$$\text{Madde ayıncılık indisi (d)} = (7-2)/16 = 0,31$$

Yapılan hesaplamalar testin dördüncü maddesinin zor bir madde olduğunu göstermektedir. Çünkü test maddesinin güçlük indisi oldukça küçüktür (0,28). Maddenin ayıncılık indisi ise 0,31 olarak hesaplanmıştır. Madde alt gruptaki öğrencilere oranla üst gruptaki öğrenciler tarafından daha fazla doğru cevaplanmıştır. Buna göre madde yüksek başarılı ve düşük başarılı öğrencileri ayırt etmede uygun ve iyi bir sorudur.

Örnek 8.

Bir testte 129 kişinin 8 soruya verilen doğru cevaplar (1) ve yanlış cevapları (0) ekte verilmiştir.

- Testin iç güvenilirliğini (KR-20),
- Madde güçlük indeksini ,
- Madde ayırt edicilik indekslerini bulunuz ve sonuçları yorumlayınız.

<u>TOPLAM</u>	<u>124</u>	<u>6</u>	<u>30</u>	<u>27</u>	<u>8</u>	<u>73</u>	<u>0</u>	<u>9</u>	<u>277</u>	
P	0,96	0,05	0,23	0,21	0,06	0,57	0,00	0,07	2,15	AO
Q	0,04	0,95	0,77	0,79	0,94	0,43	1,00	0,93		
PQ	0,04	0,04	0,18	0,17	0,06	0,25	0,00	0,06	0,79	TOPLAM

S.SAPMA	0,74
VARYANS	0,55

KR20 -0,508

TESTTE GİREN 129 KİŞİ ORTALAMA 2.1 SORUYA DOĞRU CEVAP VERMİŞTİR.

	S1	S2	S3	S4	S5	S6	S7	S8	Üst %27 toplam	S1	S2	S3	S4	S5	S6	S7	S8	Alt %27 toplam
1	1	1	0	0	1	0	0	1	4	1	0	0	0	0	1	0	0	2
1	1	1	0	0	1	0	0	1	4	1	0	0	0	0	1	0	0	2
1	1	1	0	0	1	0	0	1	4	1	0	0	0	0	1	0	0	2
1	1	1	0	0	1	0	0	1	4	1	0	0	0	0	1	0	0	2
1	1	1	0	0	1	0	0	1	4	1	0	0	0	0	1	0	0	2
1	1	1	0	0	1	0	0	1	4	1	0	0	0	0	1	0	0	2
1	1	0	1	0	1	0	0	1	4	1	0	0	0	0	1	0	0	2
1	1	0	0	0	1	0	0	1	3	1	0	0	0	0	1	0	0	2
1	1	0	1	0	0	0	0	1	3	1	0	0	0	0	1	0	0	2
1	1	0	1	1	0	0	0	0	3	1	0	0	0	0	1	0	0	2
1	1	0	1	1	0	0	0	0	3	1	0	0	0	0	1	0	0	2
1	1	0	1	1	0	0	0	0	3	1	0	0	0	0	1	0	0	2
1	1	0	1	1	0	0	0	0	3	1	0	0	0	0	1	0	0	2
1	1	0	1	0	0	1	0	0	3	1	0	0	0	0	1	0	0	2
1	1	0	1	0	0	1	0	0	3	0	0	0	1	0	0	0	0	1
1	1	0	1	0	0	1	0	0	3	0	0	0	1	0	0	0	0	1
1	1	0	1	0	0	1	0	0	3	0	0	0	1	0	0	0	0	1
1	1	0	1	0	0	1	0	0	3	1	0	0	0	0	0	0	0	1
1	1	0	1	0	0	1	0	0	3	1	0	0	0	0	0	0	0	1
1	1	0	1	0	0	1	0	0	3	1	0	0	0	0	0	0	0	1
1	1	0	1	0	0	1	0	0	3	1	0	0	0	0	0	0	0	1
1	1	0	1	0	0	1	0	0	3	1	0	0	0	0	0	0	0	1
1	1	0	1	0	0	1	0	0	3	1	0	0	0	0	0	0	0	1
1	1	0	1	0	0	1	0	0	3	1	0	0	0	0	0	0	0	1
1	1	0	1	0	0	1	0	0	3	1	0	0	0	0	0	0	0	1
1	1	0	1	0	0	1	0	0	3	1	0	0	0	0	0	0	0	1
1	1	0	1	0	0	1	0	0	3	1	0	0	0	0	0	0	0	1
1	1	0	1	0	0	1	0	0	3	1	0	0	0	0	0	0	0	1
1	1	0	1	0	0	1	0	0	3	1	0	0	0	0	0	0	0	1
1	1	0	0	1	0	0	0	0	2	1	0	0	0	0	0	0	0	1
1	1	0	0	1	0	0	0	0	2	1	0	0	0	0	0	0	0	1
1	1	0	0	1	0	0	0	0	2	0	0	0	0	0	0	0	0	0
TOPLAM	35	6	25	9	8	16	0	9	108	30	0	0	4	0	17	0	0	51

MGi	S1	S2	S3	S4	S5	S6	S7	S8
MADDE	0,93	0,09	0,36	0,19	0,11	0,47	0,00	0,13
GÜÇLÜK İNDEKSİ	ÇOK KOLAY	ZOR	ORTA	ZOR	ZOR	KOLAY	ZOR	ZOR

<u>MGi</u>	<u>Yorum</u>
<0.3	ZOR
0.3-0.49	ORTA
0.5-0.69	KOLAY
0.7-1.0	ÇOK KOLAY

MAI	S1	S2	S3	S4	S5	S6	S7	S8
MADDE	0,14	0,17	0,71	0,14	0,23	0,00	0,00	0,26
AYIRT EDİCİLİK İNDEKSİ	Ç.ZAYIF	ZAYIF	ÇOK İYİ	Ç.ZAYIF	ZAYIF	Ç.ZAYIF	Ç.ZAYIF	İYİ

<u>MAİ</u>	<u>YORUM</u>
<0.2	ÇOK ZAYIF MADDE (TESTTEN ATILMALI)
0.2-0.29	DÜZELTİLESİ-GELİŞTİRİLMESİ GEREKEN MADDE
0.3-0.39	OLDUKÇA İYİ MADDE, FAKAT GELİŞTİRİLEBİLİR
>=0.4	ÇOK İYİ MADDE

Örnek 9. 10 soruya 20 öğrencinin vermiş olduğu cevaplar aşağıdaki gibidir. Madde güçlük ve madde ayırt edicilik indekslerini bulunuz.

	1.soru	2.soru	3.soru	4.soru	5.soru	6.soru	7.soru	8.soru	9.soru	10.soru	Toplam puan
A	1	1	1	1	1	1	1	1	1	1	10
B	1	1	1	1	1	0	1	1	1	1	9
C	1	1	0	1	1	1	1	0	1	1	8
D	1	1	1	1	1	0	0	1	1	1	8
E	1	1	1	1	1	1	1	0	0	0	7
F	1	0	1	0	1	1	0	1	1	1	7
G	1	1	1	1	0	0	0	1	1	1	7
H	1	1	0	1	0	1	0	1	0	1	6
I	1	0	1	1	0	1	1	0	1	0	6
J	0	1	1	1	1	1	0	0	0	1	6
K	1	1	1	0	0	0	1	0	1	1	6
L	1	1	0	1	1	1	0	0	0	0	5
M	1	1	1	1	0	1	0	0	0	0	5
N	1	0	1	1	1	0	0	0	0	0	4
O	0	1	1	1	0	0	0	0	0	0	3
Ö	1	1	0	0	0	0	0	1	0	0	3
P	1	0	1	0	0	0	0	0	0	0	2
R	0	0	0	0	0	0	0	1	0	0	1
S	0	0	0	0	0	0	1	0	0	0	1
T	0	0	0	0	0	0	0	1	0	0	0

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1		M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	TOPLAM		
2	O1	1	1	1	1	1	1	1	1	1	1	10	ÜST %27 GRUP	
3	O2	1	1	1	1	1	0	1	1	1	1	9		
4	O3	1	1	0	1	1	1	1	0	1	1	8		
5	O4	1	1	1	1	1	0	0	1	1	1	8		
6	O5	1	1	1	1	1	1	1	0	0	0	7		
7	O6	1	0	1	0	1	1	0	1	1	1	7		
8	O7	1	1	1	1	0	0	0	1	1	1	7		
9	O8	1	1	0	1	0	1	0	1	0	1	6		
10	O9	1	0	1	1	0	1	1	0	1	0	6		
11	O10	0	1	1	1	1	1	0	0	0	1	6		
12	O11	1	1	1	0	0	0	1	0	1	1	6		
13	O12	1	1	0	1	1	1	0	0	0	0	5		
14	O13	1	1	1	1	0	1	0	0	0	0	5		
15	O14	1	0	1	1	1	0	0	0	0	0	4		
16	O15	0	1	1	1	0	0	0	0	0	0	3		
17	O16	1	1	0	0	0	0	0	1	0	0	3	ALT %27 GRUP	
18	O17	1	0	1	0	0	0	0	0	0	0	2		
19	O18	0	0	0	0	0	0	0	1	0	0	1		
20	O19	0	0	0	0	0	0	1	0	0	0	1		
21	O20	0	0	0	0	0	0	0	1	0	0	1		
22	TOPLAM	15	13	13	13	9	9	7	9	8	9	105	AO	5.25
23	P	0.75	0.65	0.65	0.65	0.45	0.45	0.35	0.45	0.4	0.45	5.25	s	2.643388
24	q	0.25	0.35	0.35	0.35	0.55	0.55	0.65	0.55	0.6	0.55		KR-20	0.741
25	varyans=p*q	0.188	0.228	0.228	0.228	0.248	0.248	0.228	0.248	0.24	0.248	2.3275		
26	ÜST%27	5	5	4	5	5	3	4	3	4	4			
27	ALT%27	2	1	1	0	0	0	1	3	0	0			
28	MGI	0.7	0.6	0.5	0.5	0.5	0.3	0.5	0.6	0.4	0.4			
29	MAI	0.6	0.8	0.6	1	1	0.6	0.6	0	0.8	0.8			
30	rj=s*MAI	0.26	0.382	0.286	0.477	0.497	0.298	0.286	0	0.392	0.398			

Örnek 10. Bir sınavdan aşağıdaki sonuçlar elde edilmiştir.

Madde No	A	B	C*	D	E	Boş	Toplam
Üst Grup	12	12	50	14	11	1	100
Alt Grup	19	22	10	21	26	2	100
Toplam	31	34	60	35	37	3	200

- Verilen maddenin güçlük indeksi kaçtır?
- Verilen maddenin ayırt edicilik indeksi nedir?
- Güçlük indeksi değerine göre, maddenin güçlük düzeyi nasıldır?

Cevaplar:

- 0,3
- 0,4
- Zor

Madde Varyansı ve Standart Sapma

- Bir maddenin varyansı; o maddeyi doğru cevaplayanların oranıyla,dođru cevaplamayanların oranının çarpımıdır.
- Bir maddenin standart sapması; o maddenin varyansının karekökü ile belirtilmesidir.
- Bir maddenin standart sapması ve varyansı, maddenin ve testin güvenilirliğinin hesaplanmasında kullanılır.
- Uygulamada deđişkenlik ölçüsü olarak genellikle standart sapma kullanılır.

Madde Varyansı ve Standart Sapma

- Varyans ve standart sapma, bir dağılımdaki bireyler arasındaki farklılıkları gösteren bir istatistiktir.
- Maddenin varyansı ve standart sapması büyüdükçe farklılıklar artar, varyans küçüldükçe farklılıklar azalır.
- Madde varyansı 0,00 ile 0,25 arasındaki değerleri
- Maddenin standart sapması 0,00 ile 0,50 arasındaki değerleri
- İyi bir test maddenin varyansının 0,25'e, standart sapmasının 0,50'ye yakın olması istenir.
- Varyansı ve standart sapması iyi bir maddenin güçlüğü orta düzeyde yani 0,50 civarında olması

Madde Varyansı ve Standart Sapma

$$\text{Madde varyansı } (S_s)^2 = P_i \cdot q_i$$

P_i Madde güçlüğü, yani maddeyi doğru cevaplama yüzdesi

q_i $(1 - P_i)$, yani maddeyi yanlış cevaplama yüzdesi

$$\text{Standart Sapma} = \sqrt{\text{Madde Varyansı}}$$

DİZİ GENİŞLİĞİ

- Dizi genişliği, bir testten öğrencilerin aldıkları en yüksek ve en düşük puanlar arasındaki farktır. Örnek test puanlarımıza baktığımızda, en yüksek puan **10**, en düşük puan ise **4**'tür. Bu iki puan arasındaki fark **6**'dır. Bu değere dizi genişliği adı verilir. Dizi genişliği küçükse, yapılacak ilk yorum testin güvenilirliğinin ve ayırıcılık gücünün düşük olduğu yönündedir.

Testin Çarpıklık Deęeri

- **Daęılımın çarpıklık ölçüsü:** Bir testten elde edilen puanların daęılımını normal daęılıma yaklaştıkça iyidir diye bilinir. Ancak istenen puan daęılımını, testin kullanılış maksadına baęlıdır.

Çarpıklık deęeri = $3X$ (Aritmetik ortalama - ortanca)/standart sapma
Bu formülün uygulanması sonunda elde edilen deęerlere göre testin güçlüğüne ilişkin olarak şunlar söylenebilir.

Çarpıklık	Grup için testin güçlüğü
Negatif	Kolay
Pozitif	Zor
0,10 dan küçük	Hafif zor
0,10 - 0,25 arası	Orta güçlükte
0,25'den büyük	Çok zor

Örnek 11. Aşağıdaki tablodaki 7.maddenin varyansı kaçtır.

		Maddeler										X_i^*
		1	2	3	4	5	6	7	8	9	10	
Öğrenci	Ahmet	1	1	1	1	1	1	1	1	1	1	10
	Veli	0	1	0	1	0	1	0	0	1	0	4
	Mehmet	1	1	1	1	0	0	0	0	1	0	5
	Ayça	0	1	1	0	0	0	0	0	0	0	2
	Aylin	1	1	1	1	0	0	1	1	1	0	7
	Canan	0	1	0	0	0	0	1	0	1	0	3
I_j		3	6	4	4	1	2	3	2	5	1	31

* öğrencilerin maddelere verdikleri doğru cevapların toplamını, yani her bir öğrencinin testten aldığı ham puanı göstermektedir

$$P_j = \frac{I_j}{N} \Rightarrow P_j = \frac{3}{6} = 0,50$$

$$q_j = 1 - p_j \Rightarrow q_j = 1 - 0,50 = 0,50 \text{ buradan,}$$

$$s_j^2 = p_j \cdot q_j \Rightarrow 0,50 \cdot 0,50 = 0,25$$

Örnek 12. Örnek 6 daki tabloda 4 ve 5. maddelerin standart sapması kaçtır.

$$s_j = \sqrt{0,66 \cdot (1 - 0,66)} \Rightarrow \sqrt{0,66 \cdot 0,44} \Rightarrow 0,54 \quad \text{Madde 4 için.}$$

$$s_j = \sqrt{0,20 \cdot (1 - 0,20)} \Rightarrow \sqrt{0,20 \cdot 0,80} \Rightarrow 0,40 \quad \text{Madde 5 için.}$$

Madde Güvenilirlik Katsayısı

Madde ayırt edicilik indeksi ile maddenin standart sapmasının çarpımıdır. -0,5 ile 0,5 arasında değer alır.

Madde ayırt ediciliği ve madde standart sapması arttıkça, madde güvenilirliği de artar. Güvenilir bir test oluşturabilmek için madde ayırt ediciliği ve standart sapması yüksek yani güvenilir maddelerden testin oluşturulması gerekir.

$$r_j = r_{jx} \times S_j$$

Madde Ayırt Edicilik İndeksi	Maddenin standart sapması
---------------------------------	------------------------------

Testin Ortalaması

Testin ortalaması nihai teste alınan maddelerin güçlük indeksleri toplamı ile kestirilebilir.

$$\bar{X} = \sum_{j=1}^K p_j$$

Testin ortalaması, testi oluşturan maddelerin madde güçlük indekslerinin toplamına eşittir. Ortalaması yüksek bir test oluşturmak isteniyorsa, güçlük indeksi büyük olan maddelerin, ortalaması düşük bir test isteniyorsa güçlük indeksi düşük maddelerin seçilmesi uygun olur (Baykul, 1989).

Testin Ortalama Güçlüğü:

Bir testin güçlük düzeyi hakkında bilgi verir. Bir testin kolay, orta ve zor düzeyde olup olmadığına karar verilir.

$$\bar{P} = \frac{\bar{X}_T}{K}$$

\bar{P} : Testin ortalama güçlüğü

\bar{X}_T : Puanların Aritmetik Ortalaması

K : Testten alınacak en yüksek puan
(madde sayısı)

P değeri **1'e ne kadar yakında** test o kadar **kolay**,

P değeri **0'a ne kadar yakında** o kadar **zor**,

P değeri **0.50** civarında ise **orta güçlüktedir**.

P<0,30 ise **zor** bir test

0.30<=P<=0,70 ise **orta** güçlükte test

P>0.70 ise **kolay** test denir.

Testin Güvenilirlik Katsayısı

Bir testin güvenilirlik katsayısı 0 ile +1 arasında deęer alır. Güvenilirlik katsayısı düşük bir test madde sayısı artırılarak yükseltilebilir (Yılmaz, 1998).

Örnek 13. 10 öğrencinin 10 soruya vermiş oldukları cevapların dağılımı aşağıdaki gibidir. Her bir test maddesi için madde **güçlük** ve **ayırt edicilik indekslerini** hesaplayınız. Maddelerin varyanslarını, standart sapmalarını ve güvenilirlik katsayılarını bulunuz. bulunuz.

	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	Top.
A	1	1	1	1	1	1	1	1	0	1	9
B	1	1	0	0	0	0	0	1	1	0	4
C	0	0	0	0	1	0	1	0	0	0	2
D	1	1	1	1	1	1	1	1	1	1	10
E	0	0	0	0	0	0	1	1	1	0	3
F	1	1	1	1	1	0	1	1	1	0	8
G	0	1	0	0	0	0	0	0	0	0	1
H	1	1	0	1	1	0	1	1	1	0	7
I	1	1	0	0	1	1	0	0	0	0	4
J	1	1	0	0	0	0	1	0	0	0	3
Top.	7	8	3	4	6	3	7	6	5	2	51

	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10
Üst %27	3	3	3	3	3	2	3	3	2	2
Alt %27	1	2	0	0	1	0	2	0	0	0
MGI (pj)	0.667	0.833	0.5	0.5	0.667	0.333	0.833	0.5	0.333	0.333
MAI (rjx)	0.667	0.333	1	1	0.667	0.667	0.333	1	0.667	0.667

Madde Güçlük İndeksi

$$p_j = \frac{n(dü) + n(da)}{2n}$$

Madde Ayırt Edicilik İndeksi

$$r_{jx} = \frac{n(dü) - n(da)}{n}$$

Madde güvenilirlik katsayıları:

$$r_j = r_{jx} \times S_j$$

	<u>M1</u>	<u>M2</u>	<u>M3</u>	<u>M4</u>	<u>M5</u>	<u>M6</u>	<u>M7</u>	<u>M8</u>	<u>M9</u>	<u>M10</u>
MGI (pj)	0.67	0.83	0.50	0.50	0.67	0.33	0.83	0.50	0.33	0.33
MAI (rjx)	0.67	0.33	1.00	1.00	0.67	0.67	0.33	1.00	0.67	0.67
S.Sapma	0.46	0.40	0.46	0.49	0.49	0.46	0.46	0.49	0.50	0.40
Madde Gv. Katsayısı (rj)	0.31	0.13	0.46	0.49	0.33	0.31	0.15	0.49	0.33	0.27

SORULAR

Soru 1. Aşağıdaki tabloda numaralamış beş sorunun doğru cevaplanma yüzdeleri (P_j) ile ayırt edicilik değerleri (r_{jx}) verilmiştir.

Soru	P_j	r_{jx}
1	0,35	0,60
2	0,40	-0,35
3	0,80	0,10
4	0,25	0,15
5	0,85	0,50

Eğer düzeltilemiyorsa yukarıdaki sorulardan hangisi, testten mutlaka çıkarılmalıdır?

A) 1

B) 2

C) 3

D) 4

E) 5

Soru 2. Aşağıdaki tabloda numaralamış beş sorunun doğru cevaplanma yüzdeleri (P_j) ile ayırt edicilik değerleri (r_{jx}) verilmiştir.

Soru	P_j	r_{jx}
1	0,35	0,60
2	0,40	-0,35
3	0,80	0,10
4	0,25	0,15
5	0,85	0,50

Başvuran kişinin çok, seçilecek kişinin az olduğu bir sınavda kullanılacak soruların çoğu, güçlük ve ayırt edicilik bakımından yukarıdaki sorulardan hangisine benzer olmalıdır?

A) 1

B) 2

C) 3

D) 4

E) 5

Soru 3. Aşağıdaki tabloda numaralamış beş sorunun doğru cevaplanma yüzdeleri (P_j) ile ayırt edicilik değerleri (r_{jx}) verilmiştir.

Soru	P_j	r_{jx}
1	0,35	0,60
2	0,40	-0,35
3	0,80	0,10
4	0,25	0,15
5	0,85	0,50

Yukarıdaki sorulardan hangisi kolay ve ayırt edici bir sorudur?

A) 1

B) 2

C) 3

D) 4

E) 5

Soru 4. Tabloda soru sayıları ve Aritmetik ortalamaları verilmiş dersler bulunmaktadır.

<u>Dersler</u>	(Soru Sayısı)	<u>X</u>
Biyoloji	(10)	5,2
Matematik	(45)	38,1
Fizik	(19)	13,4
Kimya	(12)	6,9
Tarih	(14)	11,7

Madde Güçlük indeksi en düşük olan ders aşağıdakilerden hangisidir?

A) Biyoloji

B) Matematik

C) Tarih

D) Fizik

E) Kimya

Soru 5. Aşağıdaki sorulardan hangisi madde analizi ile elde edilecek bilgilerden yararlanılarak cevaplanamaz?

- A)** Testin zorluk derecesi amaca uygun mudur?
- B)** Testteki sorularla yoklanan davranışlar, kazandırılması amaçlanan davranışlarla aynı mıdır?
- C)** Testteki maddelerin çeldiricileri beklenene yakın biçimde çalışıyor mu?
- D)** Test, ölçülmek istenen nitelikleri ne derece bir tutarlılıkla ölçmektedir?
- E)** Test, çok öğrenmiş ve az öğrenmiş olan öğrencileri birbirinden ayırabiliyor mu?

Soru 6. Bir testteki maddelere ilişkin aşağıdaki değerler elde edilmiştir.

Maddeler	I	II	III	IV	V
Madde güçlüğü	0,1	0,15	0,5	0,6	0,9
Madde ayıricılık gücü	0,08	-0,15	0,5	0,28	0,43

Tabloda verilen bilgilere göre hangi maddenin güvenilirliğe katkısı en fazladır?

A) 1

B) 2

C) 3

D) 4

E) 5

Soru 7.

	A	B	C	D	E	Boş	N
Üst Grup	18	30	10	10	10	22	100
Alt Grup	2	26	15	14	10	33	100

Yukarıda analizi verilen maddede doğru cevap hangi seçenekte olursa ayırt edicilik en yüksek olur?

A) A

B) B

C) C

D) D

E) E

Soru 8.

	A	B	C	D	E	Boş	N
Üst Grup	18	30	10	10	10	22	100
Alt Grup	2	26	15	14	10	33	100

Yukarıda analizi verilen maddenin doğru cevabı E ise madde ayırıcılık gücü indeksi kaç olur?

A) 0,00

B) 0,10

C) 0,16

D) 0,20

E) 1,00

Soru 9.

Yetenek Grubu	Seçenekler (%)					
	a.	b.	c.	d.	e.	Boş
Yüksek	30	5	50	15	0	0
Düşük	10	30	10	20	0	30

Yukarıdaki tabloda; yüksek ve düşük yetenek grubundaki öğrencilerin, sayısal yetenek testinde yer alan bir maddenin seçeneklerini işaretleme veya boş bırakma yüzdeleri verilmiştir.

Buna göre, maddeyle ilgili olarak aşağıdakilerden hangisi yanlıştır?

- A) a. seçeneği çeldirici ise kötü işleyen bir çeldiricidir.
- B) Doğru cevap c. seçeneği ise maddenin ayırt ediciliği 0,25'tir.
- C) Doğru cevap c. seçeneği ise maddenin doğru cevaplanma oranı 0,30'dur.
- D) Soru dört seçenekli olarak işlemiştir.
- E) Doğru cevap d. seçeneği ise maddenin ayırt ediciliği negatiftir.

Cevap B

Bir testteki maddelere ilişkin aşağıdaki değerler elde edilmiştir. (KPSS-2008)

Maddeler	I	II	III	IV	V
Madde gücü	0,1	0,15	0,5	0,6	0,9
Madde ayrıcalık gücü	0,08	-0,15	0,5	0,28	0,43

Soru 10. Tabloda verilen bilgilere göre hangi maddenin güvenilirlik katkısı en yüksektir?

- A) I B) II **C) III** D) IV E) V

Soru 11. Yapılandırılacak bir testte bu maddeleri kullanma zorunluluğu olsaydı, en köklü değişikliğin hangi maddede yapılması gerekirdi?

- A) I **B) II** C) III D) IV E) V

Soru 12.

	A	B	C	D	E	Toplam
Üst Grup	1	6	8	12	3	30
Alt Grup	13	6	7	2	2	30

Tabloya göre;

- I. A seçeneği doğru cevap olursa madde ayırt ediciliği en yüksek olur.
- II. B seçeneği doğru cevap olursa kolay bir madde olur.
- III. E seçeneği doğru cevap olursa zor bir madde olur.

İfadelerinden hangileri doğrudur?

- A) Yalnız I B) Yalnız II C) Yalnız III D) I ve II E) II ve III

Soru 13.

Bir öğretmen sınavında yer ala beş maddenin özelliklerini incelemek amacıyla öğrencilerin her bir maddeden aldıkları puanlar ile testin tümünden aldıkları puanlar arasındaki korelasyon katsayısını hesaplamıştır. Bilenle bilmeyeni en iyi ayırt eden maddeyi bu şekilde görebileceğini iddia eden bu öğretmen tablodaki katsayıları elde etmiştir.

Maddeler	Korelasyon Katsayısı
1	0.2
2	-0.1
3	0.5
4	0.6
5	-0.7

Bu öğretmenin en iyi ayırt edici olduğunu iddia edeceği madde hangisidir?

A) 1

B) 2

C) 3

D) 4

D) 5

Soru 14.

Aşağıdaki tabloda aynı kapsamı ölçmek üzere hazırlanmış beş farklı test yer almaktadır. Bu testlerin her biri beşer tane çoktan seçmeli maddeden oluşmaktadır. Tabloda bu maddelerin madde güçlük indeksleri verilmiştir.

	Madde Güçlükleri				
	Madde 1	Madde 2	Madde 3	Madde 4	Madde 5
I	0.4	0.5	0.6	0.5	0.6
II	0.4	0.8	0.9	0.7	0.3
III	0.3	0.4	0.5	0.2	0.9
IV	0.8	0.7	0.6	0.5	0.9
V	0.2	0.4	0.3	0.1	0.5

1. III. Testte yer alan en zor madde hangisidir?
2. V. Testte madde ayırt edicilik indeksi en yüksek olması beklenen madde hangisidir?
3. Tablodaki testler içinde en kolay olan hangisidir?

Soru 15.

		Sorular									
		1	2	3	4	5	6	7	8	9	10
ÖĞRENCİLER	A	1	1	1	0	0	1	1	0	1	0
	B	1	1	0	0	1	1	1	0	0	1
	C	0	1	0	0	1	1	1	1	1	1
	D	1	1	1	0	1	0	1	1	1	0
	E	0	0	1	0	0	0	1	1	1	0
	F	1	1	0	0	1	1	1	1	1	1
	G	1	1	0	0	1	1	1	0	0	0
	H	1	1	0	0	1	1	1	0	0	1
	I	1	1	1	0	0	0	1	0	0	1
	J	0	0	0	0	0	0	1	1	1	0

Ölçülen özellikler açısından öğrencilerin bireysel farklılıklarını ortaya çıkarmada hiç etkisi olmayan maddeler hangileridir?

A) Yalnız 4

B) Yalnız 7

C) Yalnız 8

D) 4 ve 7

E) 4,7 ve 8

Madde Analizinde Kullanılan Yöntemler

2. Henryson Yöntemi

Bu yöntemde testi cevaplayan tüm kişilerin cevapları analiz dahil edilir. Doğru cevaplar “1”, yanlış cevaplar, boş bırakılanlar veya birden fazla şık işaretlenenler “0” olarak puanlanır. Cevaplayıcıların testin tüm maddelerinden aldıkları puanların gösterildiği matrise “madde puanları matrisi” adı verilir.

Madde Puanları Matrisi

	1	2	...	j	...	K	X_i
A	1	0	...	0	...	1	X_1
B	0	0	...	1	...	1	X_2
...
i	0	1	...	a_{ij}	...	0	X_i
...
N	1	0	0	X_N
Madde Puanı	I_1	I_2	...	I_j	...	I_k	

Satırlar cevaplayıcıları, sütunlar testin maddelerini göstermektedir (Baykul, 1989).

Henryson yöntemi ile madde analizi için ilk olarak her bir yanıtlayıcının toplam puanı hesaplanır. Ardından testin aritmetik ortalaması ve standart sapması hesaplanır. Test puanlarının aritmetik ortalaması ve standart sapması satır toplamlarından bulunur.

$$\bar{X} = \frac{\sum X_i}{N}$$

$$s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n - 1}}$$

Madde puanları matrisinde cevaplayıcıların testten aldıkları puanların toplamı X_i gösterilir. Madde puanları matrisinde satır toplamı (X_i) test puanlarını verir.

Henryson Yöntemi: Madde Güçlük İndeksi

Henryson yöntemi ile madde güçlük indeksi, maddeyi doğru yapanların sayısının testi alan tüm öğrencilerin sayısına bölünmesiyle hesaplanır.

$$p_j = \frac{n(D)}{N}$$

p_j : Madde güçlük indeksi

$n(D)$: Maddeyi doğru yapanların sayısı

N : Testi alan kişi sayısı

Madde Güçlük İndeksi	Madde Güçlük Düzeyi
1'e yakın	Kolay
0,5 civarında	Orta
0'a yakın	Zor

Henryson Yöntemi: Madde Ayırt Edicilik İndeksi

Güvenilir bir testin madde ayırt ediciliği yüksek maddelerden oluşması gerekir. -1 ile +1 arasında değerler alır. Negatif ayırt edicilik maddenin önemli bir kusuru olduğunu gösterir. Bu yüzden negatif ayırt edici maddeler teste alınmaz. Bu yöntemde madde ayırtıcılık gücü indeksi aşağıdaki formülle hesaplanır:

$$r_{jx} = \frac{\bar{X}_{jD} - \bar{X}_x}{S_x} \sqrt{\frac{p_j}{q_j}}$$

\bar{X}_x : Test puanlarının aritmetik ortalaması

\bar{X}_{jD} : j maddesini doğru yapan kişilerin test puanları ortalamasıdır.

r_{jx} : j maddesini ayırt edicilik indeksi

S_x : Test puanlarının standart sapması

q_j : $1 - p_j$

Madde Ayırt Edicilik İndeksi	Madde Deęerlendirmesi
$\geq 0,4$	Çok iyi
0,3-0,39	Oldukça iyi
0,2-0,29	Düzeltilmesi veya geliştirilmesi gerekir
$<0,20$	Çok zayıf ve testten çıkarılması gerekir

Soru. Aşağıdaki gibi verileri kullanarak Henryson yöntemi ile madde güçlük indekslerini ve madde ayırıcılık gücü indekslerini hesaplayınız.

	S1	S2	S3	S4	S5
Ö1	1	1	1	1	1
Ö2	1	0	1	1	1
Ö3	1	0	0	1	0
Ö4	1	0	0	1	0
Ö5	0	0	1	0	0
Ö6	0	1	0	0	0
Ö7	0	0	1	1	0
Ö8	1	0	0	0	1
Ö9	0	0	1	1	0
Ö10	0	0	1	1	1

	A	B	C	D	E	F	G
1		S1	S2	S3	S4	S5	TOPLAM
2		1	1	1	1	1	5
3		1	0	1	1	1	4
4		1	0	0	1	0	2
5		1	0	0	1	0	2
6		0	0	1	0	0	1
7		0	1	0	0	0	1
8		0	0	1	1	0	2
9		1	0	0	0	1	2
10		0	0	1	1	0	2
11		0	0	1	1	1	3
12	TOPLAM	5	2	6	7	4	24
13	p_j	0.5	0.2	0.6	0.7	0.4	
14	q_j	0.5	0.8	0.4	0.3	0.6	
15	$\bar{X}_{.j}$	3.00	3.00	2.83	2.86	3.50	
16	$\text{kok}(P_j/q_j)$	1.00	0.50	1.22	1.53	0.82	
17	r_{jk}	0.47	0.24	0.42	0.55	0.71	
18							
19	AO	2.4					
20	S.S.	1.26					

3. GEÇERLİLİK (Validity)

Geçerlilik;

- ***Kullanılan ölçüm aracının ölçülmek istenen özelliğe uygun olması,***
- ***Ölçümün kurallara uygun olarak doğru yapılıp yapılmadığı,***
- ***Ölçüm verilerinin gerçekten ölçülmek istenen özelliği yansıtıp yansıtmasıdır.***

Geçerlilik, belli bir olguya ait ölçüm rakamları olguyu doğru bir şekilde yansıtıyor, tanımlıyor veya kuramsal açıklamalar getiriyorsa geçerlidir (Hammersley, 1987).

Ölçüm verilerinin doğrulamasını yapmak için önce güvenilirlik analizi yapılır, ancak bu tek başına yeterli değildir. Aynı zamanda geçerlilik analizleri de yapılmalıdır.

Yüksek geçerlilik aynı zamanda yüksek güvenilirlik anlamına gelebilir. Ancak yüksek güvenilirlik geçerlilik hakkında hiçbir bilgi vermez.

Güvenilirlik analizlerinde teknik hesaplamalar ön plana çıkarken, geçerlilik analizlerinde yargısal değerlendirmeler ile teknik hesaplamaların birlikte kullanılması gerekir.

Geçerlilik bir ölçme aracının ölçmeyi amaçladığı özelliği, başka herhangi bir özellikle karıştırmadan, tam ve aynı zamanda doğru olarak ölçmesidir. Geçerlik bir gerecin ölçmesi için düzenlendiği olguyu ölçme derecesiyle ilişkilidir. Geçerlik sayesinde ölçülmek istenen değişkenin ölçülebilme derecesi ortaya konulmuş olur (Ergin, 1995 ; Turgut, 1989; Yılmaz, 1996).

Geçerlik, test değerlendirmede göz önüne alınması gereken en önemli konulardan bir tanesidir. Test puanlarına dayalı olarak varılabilecek çıkarımların uygunluğu, anlamlılığı ve kullanılışlılığı hakkında bilgi verir. Yani kavram, bir teste dayalı çıkarsamaların uygunluk veya anlamlılık derecesidir. Testin geçerli hale getirilmesi, bu tür çıkarsamaları desteklemek için veri toplama sürecidir (American Psychological Association, 1992).

İdeal bir geçerlilik çalışması, çeşitli kategorideki bilgilerin bir araya getirilmesini gerektirmektedir (American Psychological Association, 1992). Geçerlik çalışması için veriler çeşitli yollarla toplanabilir. Geleneksel olarak geçerlik türleri kapsamla bağlantılı, kriterle bağlantılı veya yapıyla bağlantılı geçerlik olarak gruplanmaktadır. Bu kategoriler arasında kesin ayırım yapmak mümkün değildir. İdeal olan bir geçerlik saptama bu üç geleneksel kategoriyi de kapsayan bilgileri içermelidir (Hovardalıoğlu, Sezgin, 1998; Turgut, 1989; Yılmaz, 1996).

Geçerlilikte Temel Esaslar

Geçerlilik kavramı içerisinde genelde üzerinde durulan hususlar özetlenecek olursa 3 temel esas dikkati çekmektedir.

1. Kullanılan ölçüm aracı, ölçülmek istenen özelliğe uygun bir araç mıdır?
2. Ölçüm, kurallara uygun olarak yapılıyor mu?
3. Ölçülen veriler gerçekten ölçmek istediğimiz özelliği yansıtıyor mu?

- ✓ İe kapanıklığı len bir test ie kapanıklığı, kaygıyı len bir test kaygıyı lmelidir. rneğın gen sporcuların ileride sprinter olup olmayacağını belirlemek amacıyla bir test geliştirilmiş ise bu test istenen amacı sağlamalıdır. Yanlıřlıkla orta ya da uzun mesafeci olacakları sprinter grubuna katmamalıdır.
- ✓ Geometri dersi genel sınavında ğrencilerin geometri becerileri lülecekse, sınavda geometriden hari matematik sorusu sorulmamalıdır.

Geçerlilik:

Ölçümlerde güvenilirlik tek başına yeterli değildir. Psikolojik değerlendirme testi “prenology” geçmiş dönemde kişilik ve davranış ölçümlerinin değerlendirilmesinde de kullanılıyordu.

Bu testi kaç psikolog uygularsa uygulasin aynı sonucu buluyordu. Bu ölçüm depresyonun bir göstergesi olarak kullanılırsa oldukça güvenilir bir ölçüm olabilir ama bunu kişilik veya davranışın dolaylı bir göstergesi olarak kullanmak psikolog’u yanlış yere götürür.

Yani test geliştirildiği amaç dışında bir kavramın (davranış) değerlendirmesi için bu ölçüm **geçerli** bir ölçüm değildir.

- Bu üniversite sınavlarında da çoğu zaman tartışma konusu olmaktadır. Üniversite sınavlarının güvenilir olması için, aynı sorular sorulduğunda aynı öğrencinin birbirine çok yakın sonuçlar alması beklenir. Amaç bilgiyi ölçmek ise yapılan üniversite sınavının gerçek bilgiyi ölçtüğü söylenebilir mi? Yani bu sınavın bilginin ötesinde bir takım farklı değerlendirmeleri de içeren bir sınav olduğu her zaman tartışılmıştır. İşte burada itiraz konusu olan güvenilirlik değil **geçerlilik**dir.
- Sözlü sınavlarda öğretmenin kişiler hakkındaki ön kanaatleri güvenilirliği değil geçerliliği zedelemektedir. Çünkü o öğretmen kendi sempatisinin yarattığı puanı bir kişiye ekleyecekse o sınavı 3 kere de yapsa sempati nedeniyle daha kolay sorular soracak ve verilen puan her zaman pozitif yönde etkilenecektir.

Belli bir nesneye ait bir özellik ölçüldüğünde bu ölçümden elde edilen veriler nesnenin ölçülen özelliğini doğru yansıtıyor mu, o özelliğın gerçek değerini tanımlıyor veya bu özelliğe kavramsal bir açıklama getiriyor mu sorusu geçerliliğın konusudur.

Burada geçerli olan ölçme aracı değil onunla elde edilen sonuçlarla ilgili bir husustur. Yani sonuçlar geçerlidir veya değildir şeklinde karar verilir. Söz konusu ölçme aracının geliştirildiğı amaç için kullanılması ile ancak geçerli sonuçlar elde edilebilir. Metreyi ağırlık ölçmek için kullanırsak geçerli sonuç elde edemeyeceğımız açıktır.

Geçerlilik bir test veya ölçüğün ölçülmek istenen şeyi ölçme derecesidir.

Örneğın biz cebirsel işlemleri (toplama, çıkarma, çarpma, bölme) öğrettik ve bunların öğrenilip öğrenilmediğini kontrol için bir test uyguluyoruz. Test soruların bu konular kapsamaması gerekir. Bilgisayarlar nasıl toplama yapar? Şeklinde bir soru kapsam dışı bir soru olacaktır. Testin geçerliliğı zedelenecektir.

3.1. GEÇERLİLİK ANALİZ YÖNTEMLERİ

3.1.1. *Yüzey Geçerliliği*

3.1.2. *Kapsam(İçerik) Geçerliliği*

3.1.3. *Kriter Geçerliliği*

3.1.4. *Yapı Geçerliliği*

GEÇERLİLİK ANALİZ YÖNTEMLERİ

3.1.1. Yüzey (Görünüş) Geçerliliği (Face Validity)

Ölçme aracının neyi ölçtüğü değil de neyi ölçer göründüğünü belirtmektedir. Bir ölçeğin görünüş geçerliliği o ölçeğin ölçmek istediği özelliği ölçüyor gözükmesidir.

Ölçek Türkçeye uzman kişilerce çevrilir ve kontrol edilir. Daha sonra Türkçeye çevrilen ölçek deneme amacıyla ön çalışmada belli sayıda kişi üzerinde uygulanır ve ölçeğin anlaşılabilirliği test edilir. Pilot çalışmada deneklerin anlamada güçlük çektikleri sorular yeniden düzenlenir ve ölçek daha kolay anlaşılır hale getirilir. Türkçe form ve İngilizce form aynı kişilere farklı zamanlarda doldurulur. Sonuçlar paired t-test ile test edilir. Tüm sorularda $p > 0,05$ çıkarsa İngilizce ve Türkçe form a verilen cevaplar aynıdır sonucuna ulaşılır.

Bir testin/ölçeğin araştırılan yapıyı ölçüp ölçmediğine ilişkin olarak arařtırmacının

- i.) Kendisinin,**
- ii.) Yakın çevresindeki arkadaşlarının,**
- iii.) Arařtırılan konu hakkındaki uzman olmayan diđer kiřilerin,**
- iv.) Pilot arařtırmaya katılan cevaplayıcıların kanaat ve görüşlerinin toplanmasıyla belirlenir.**

Yüzey geçerliliğinde ölçek maddelerinin düzgün ve anlamlı bir şekilde ifade edilmesi, doğru terimlerin kullanılması, uygun kelimelerin seçilmesi, kelimelerin anlamı açık ve net olması, belirsiz birden fazla anlama gelen kelimelerden kaçınılması önemlidir.

Kullanılan test/ölçek cevaplayıcıların eğitim düzeylerine, bilgisine, kültürel ve yetenek düzeylerine uygun olmalıdır.

Cevaplayıcıların kendilerini mağdur olmuş göstermek istemeleri, mutlu ve huzurlu olduklarını yansıtmaya çalışmaları, arařtırmacının beklediđi cevapları vermemeleri ölçüm hatalarına sebep olmaktadır.

3.1.2. Kapsam (İçerik) Geçerliliği (Content validity)

Kapsam geçerliliği örneklem olarak belirlenen test/ölçek maddelerinin belirli bir amaca yönelik olarak kavramsal ana kütleyi temsil etme derecesidir.

Geliştirilen test incelenen konuların tüm önemli alt konularını içeriyorsa, testin kapsam-içerik geçerliliğinin olduğu söylenir.

Kapsam geçerliliği bir bütün olarak ölçeğin ve ölçekteki her bir maddenin amaca ne derece hizmet ettiğiidir.

Kapsam geçerliliğinde ölçüm aracının ölçmek istediği yapıyı ölçüp ölçmediği ölçeği geliştiren kişilerin kendilerine değil, uzman kararlarına bırakılmıştır.

Kapsam geçerliliği mantıki ve istatistiki olmak üzere iki yolla araştırılabilir.

- i) Mantıki yoldan ölçeğin kapsam geçerliliği:** Ölçek uygulanmadan ölçeğin geçerliliği tahmin edilir. Ölçüm konusu kavramsal olarak tanımlanır. Ölçekteki her maddenin ve bunların dağılımının ölçüm konusunu örnekleyip örneklemeyeceği araştırılır.
- ii) İstatistiki yoldan ölçeğin kapsam geçerliliği:** Daha önceden geliştirilen ve geçerliliği yapılmış bir ölçek ile yeni geliştirilen ölçek aynı anda bireylere uygulanır ve bu iki ölçeğin ilişki katsayısı hesaplanır ve bulunan değer ölçeğin geçerlilik katsayısı olarak nitelendirilir.

Uzmanlardan ölçekteki her bir maddenin ölçme derecesini 100 puan üzerinden değerlendirmeleri istenir. Alınan uzman görüşleri doğrultusunda yapılan ifade değişikliklerinin ardından Kendall Uyuşum Katsayısı (Kendall Coefficient of Concordance W) korelasyon testi uygulanarak ölçeğin kapsam geçerliliği çalışması yapılır. $P > 0,05$ uzmanlar arasında bir uyum olduğunu gösterir. (SPSS:Non parametric test>k related samples)

Kapsam Geçerliliğinin Aşamaları :

- i. Kavramsal yapı ve test evreninin tanımlanması
- ii. Konunun kavramsal boyutlarının ortaya çıkarılması
- iii. Maddelerin belirlenen boyutlara göre oluşturulması
- iv. Geliştirilen ölçeğin uzmanlara değerlendirilmesi
- v. İstatistiksel ve matematiksel analizlerin yapılması

Lawshe'nin İçerik Geçerliliği Oranı:

Bu oran formülü ile hakemlerin her bir ifadeyi (maddeyi) nasıl değerlendirdikleri hesaplanır.

$$İGO_i = \frac{n_e - \frac{N}{2}}{\frac{N}{2}}$$

$İGO_i$: Ölçeğin i'nci maddesinin içerik geçerliliği oranı

n_e : İfadenin uygun olduğunu belirten hakem sayısı

N : Toplam hakem sayısı

İGO -1 ile +1 arasında bulunur. Hakemlerin yarısından daha azı bir madde için “uygun” işaretlemesi yapmışsa sonuç – çıkar. Eksi işaretli maddeler ölçekten çıkarılmalıdır. Hakemlerin %50'si bir madde için uygun yada uygun değil demişse İGO sıfır çıkar. Bunun için bir maddenin ölçekte yer alması için hakemlerin %50'den fazlası uygun demelidir. İGO hakemler arasındaki uyuşmanın tesadüfe bağlı olup olmadığı belli olmaz. İGO sadece değerlendiriciler arasındaki uyuşmayı gösterir.

Örnek 3.1.

Maddeler	Uygun	Kalabilir	Uygun Değil	Toplam Hakem
1	10	2	0	12
2	2	10	0	12
3	0	2	10	12
4	8	2	2	12
5	8	3	1	12
6	7	4	1	12
7	8	2	2	12
8	8	2	2	12
9	6	3	3	12
10	4	4	4	12

Maddeler	n_e	N/2	iGO	Karar
1	12	6	1,00	Kabul
2	12	6	1,00	Kabul
3	2	6	-0,66	Red
4	10	6	0,66	Kabul
5	11	6	0,83	Kabul
6	11	6	0,83	Kabul
7	10	6	0,66	Kabul
8	10	6	0,66	Kabul
9	9	6	0,50	Kabul
10	8	6	0,33	Kabul

3.1.3. Kriter (Referans) Geçerliliđi (Criterion validity):

Bu geçerlilik türü test puanlarının belirlenen bir veya birkaç dış ölçütle ilişkisini inceler.

Ölçek sonuçlarının aynı kavramsal yapıyı ölçen benzer diğer ölçüm sonuçlarıyla (Benzer ölçekler) tutarlı olması istenir. Benzer bir ölçek aynı deneklere uygulanır ve sonuçlar arasındaki korelasyona bakılır.

i) Kestirim (Yordama) Geerlilięi (Predictive Validity)

Bir leęin kestirim geerlilięi, o lekten elde edilen tahmini puan ile llmek istenen zellikleri ltę bilinen kriter arasındaki korelasyonun hesaplanmasıyla elde edilir.

niversiteye ęrenci alırken ęrencinin lisedeki baęarisının veya niversite giriř sınavındaki baęarisının niversite eęitimindeki baęarısı ile paralel olacaęı dřncesi vardır. niversite giriř testinin niversitedeki performansını tahmin ettięini gsterebilmek iin niversite giriř testinin tahmin geerlilięine sahip olması gerekir. Bu nedenle de Yapılan test ile ilerdeki performans arasında iliřki kurulmak istendięi durumlarda testin tahmin geerlilięinin llmesine gereksinim vardır.

Eęer istatistik dersi iin yapılan bir yılsonu sınavının sonuları banka giriř sınavında sorulan genel istatistik soruların sonuları ile uyumlu deęilse kestirim geerlilięi zayıftır denir.

i) Kestirim (Yordama) Geerlilięi (Predictive Validity)

Kestirim geerlilięi standart test kullanılarak belirlenir. Kestirim geerlilięi standart lümü kestiren test ya da deęişkenlerden elde edilen bir skoru kullanarak standardın kestirilme derecesidir. rneęin LYS sınavından elde edilen puanlar adayların gelecekteki başarısının kestiricisi olarak kullanılır.

İyi bir kestirim deęerine ulaşmak için kestirici (baęımsız-x) deęişken (ler) ile referans (kestirilen-y) deęişken arasındaki korelasyon katsayısının yüksek olması istenir. Ayrıca regresyon denklemini yardımıyla belli bir x için y deęeri kestirilebilir. Burada temel sorun bulunan kestirimin ne kadar güvenilir olduğudur. Bunun için rneklem büyüklüğü ve regresyon denkleminin standart hatasından yararlanılabilir.

Bu tür alışmalarda bir dięer sorun elde kestirim denkleminin yeni bir veri grubuna (rnekleme) uygulanması durumunda, geerlilik katsayısının (korelasyon) azalma eğiliminde olup olmadığıdır. Bunun için **apraz geerlilik (cross validation)** yönteminden yararlanılır.

Çapraz Geçerlilik (cross validation)

Çapraz geçerlilik yöntemi yapılan çalışma sonuçlarının doğruluk açısından test edileme sürecidir. Gözlem sayısı az ve bağımsız değişken sayısı çok olduğunda geçerlilik düşük çıkacaktır.

Çapraz geçerlilik için aşağıdaki aşamalara bakılır.

- i) İlgili anakütleden rasgele en az 200 birimlik örneklem seçilir.
- ii) Kestirici değişken-ler(x) ile ölçüt alınan değişken tüm bireylere uygulanır.
- iii) Geçerlilik katsayısı (korelasyon-çoklu korelasyon) hesaplanır. Geçerlilik katsayısı yeterince büyükse işlemlere devam edilir. Aksi durumunda çalışma durdurulur.
- iv) Kestirim denkleminin çapraz geçerliliği bulunur. Bu amaçla 200 kişilik grup 100'erli rasgele iki gruba ayrılır. I. Grubun verileri kullanılarak kestirim denklemi elde edilir. I.gruptan elde edilen kestirim denkleminde, II.grubun kestirici (x) değişken-lerin değerleri yerine konur ve II.gruptaki kestirici değişkenler için skorlar-değerler elde edilir.

Çapraz Geçerlilik (cross validation)

v.) II.gruptaki kestirici değerler kullanılarak elde edilen kestirilmiş skor-değerler ile II.grupta var olan değerler karşılaştırılır. Bu işleme çapraz geçerlilik denir. Bu değerlerin birbirine yakın olması kestirimin doğru olmasını sağlar. Bunun için kestirimin standart hatasından yararlanır. Kestirimin standart hatası, kestirimin denkleminde skorların hesaplanması ile ilgili olarak ortalama hata miktarı hakkında bilgi verecektir. Standart hata ne kadar küçük ise kestirim geçerliliği o kadar iyidir denilir.

$$SH = \sqrt{\frac{\sum (y_i - y'_i)^2}{n - 1}}$$

y_i : II.gruptaki i.birey için bilinen değer

y'_i : II.gruptaki i.birey için kestirilen değer

n : II.gruptaki birey sayısı

Örnek.

No	y	x1	x2
1	11,9	43,1	29,1
2	22,8	49,8	28,2
3	18,7	51,9	37,0
4	20,1	54,3	31,1
5	12,9	42,2	30,9
6	21,7	53,9	23,7
7	27,1	58,5	27,6
8	25,4	52,1	30,6
9	21,3	49,9	23,2
10	19,3	53,5	24,8
11	25,4	56,6	30,0
12	27,2	56,7	28,3
13	11,7	46,5	23,0
14	17,8	44,2	28,6
15	12,8	42,7	21,3
16	23,9	54,4	30,1
17	22,6	55,3	25,7
18	25,4	58,6	24,6
19	14,8	48,2	27,1
20	21,1	51,0	27,5

Veri seti aşağıdaki gibi rasgele 2 gruba ayrılır.

GRUP I				GRUP II			
No	y1	x1	x2	No	y2	x1	x2
1	11,9	43,1	29,1	3	18,7	51,9	37,0
2	22,8	49,8	28,2	4	20,1	54,3	31,1
7	27,1	58,5	27,6	5	12,9	42,2	30,9
8	25,4	52,1	30,6	6	21,7	53,9	23,7
9	21,3	49,9	23,2	10	19,3	53,5	24,8
11	25,4	56,6	30,0	12	27,2	56,7	28,3
13	11,7	46,5	23,0	15	12,8	42,7	21,3
14	17,8	44,2	28,6	16	23,9	54,4	30,1
17	22,6	55,3	25,7	18	25,4	58,6	24,6
19	14,8	48,2	27,1	20	21,1	51,0	27,5

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,881 ^a	,776	,749	2,55653

a. Predictors: (Constant), x12, x11

Ful model için (20 gözlem) R=0,881 olup geçerlilik katsayısı yüksektir. SH=2,55

Grup I için çoklu regresyon analizi sonuçları:

Model Summary

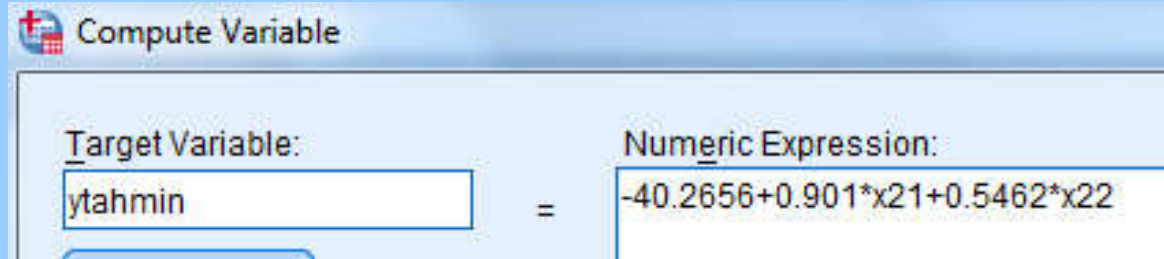
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,889 ^a	,790	,730	2,95424

a. Predictors: (Constant), x12, x11

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-40,266	13,288		-3,030	,019
	x11	,901	,191	,823	4,724	,002
	x12	,546	,378	,252	1,445	,192

a. Dependent Variable: y1



y2	ytahmin	y2_ytahmin	karetoplamlam
18,70	26,71	-8,01	64,09
20,10	25,65	-5,55	30,75
12,90	14,63	-1,73	3,01
21,70	21,24	,46	,21
19,30	21,48	-2,18	4,77
27,20	26,28	,92	,85
12,80	9,84	2,96	8,75
23,90	25,19	-1,29	1,66
25,40	25,97	-,57	,32
21,10	20,71	,39	,16

Toplam=114,65

$$SH = \sqrt{\frac{\sum (y_i - y'_i)^2}{n-1}} = \sqrt{\frac{114,65}{10-1}} = 3,57$$

SH=3,37>2,55 çıkmıştır.

I. Gruba ait $R^2=0,79$ bulundu. II.gruba ait y_2 değerleri ile kestirilen (y_{tahmin}) değerleri arasındaki korelasyon $R=0,82$ bulunur. Bunun karesi ise $R^2=0,67$ çıkar. Burada n $0,79-0,67=0,12$ 'lik bir azalma eğilimi elde edilir. Buda bize $n=20$ birimlik gözlem grubu ile yapılacak kestirimin pek güvenilir olmadığını belirtmektedir.

ii) Uyum (Yakınsama) Geçerliliği (Concurrent Validity)

Aynı yapıyı ölçmek için kullanılan alternatif araçların sonuçlarının birbirlerine yakınsaması veya aynılığının bir ölçüsü olarak kullanılan bir geçerlilik şeklidir.

Ölçmek istediği özelliği kesin bir şekilde ölçen ve yüksek geçerliliğe sahip testlere altın standart ya da referans test denir. Geliştirilen test ile referans test arasındaki korelasyon yüksekse uyum geçerliliği yüksektir denilir.

Yeni geliştirilen bir teste ilişkin uyum geçerliliği için aşağıdaki süreç izlenir.

- Altın Standart-referans test tanımlanır.
- Veriler nicel ise referans test ile yeni geliştirilen test arasındaki korelasyon katsayısı 0,8'den büyük ise yeni testin kullanılabileceğine karar verilir.
- Veriler nitel ise referans teste göre elde edilecek duyarlılık-seçicilik istatistiklerine bakılır.

Nitel Veri İçin Geçerlilik Ölçütleri

Eğer veriler nitel ise (var-yok, başarılı-başarısız vb.) uyum geçerliliği için Duyarlılık (sensitivity)-Özgüllük (Seçicilik-specificity) istatistiklerinden yararlanır.

Geliştirilen Test	Gerçek Durum		Toplam
	Başarılı	Başarısız	
Başarılı	A (GP)	B (YP)	A+B
Başarısız	C (YN)	D (GN)	C+D
Toplam	A+C	B+D	N

A (GP): Gerçek pozitif; gerçekte başarılı olup test sonucunda da başarılı olanların sayısı
B (YP): Yanlış pozitif; gerçekte başarısız olup test sonucunda başarılı olanların sayısı
C (YN): Yanlış negatif; gerçekte başarılı olup test sonucunda başarısız olanların sayısı
D (GN): Gerçek negatif; gerçekte başarısız olup test sonucunda da başarısız olanların sayısı

Duyarlılık (sensitivity) Oranı: Gerçek başarılıları ayırabilme yeteneği

$$D=A / (A+C) = GP / (GP+YN)$$

Duyarlılık; gerçekte başarılı olanların % kaçının geliştirilen test sonucunda başarılı olduğunu gösterir ve gerçek başarılar içinde geliştirilen testin başarılıları ayırtedebilme yeteneği olarak tanımlanır.

Özgüllük (specificity) Oranı: Gerçek sağlamları ayırabilme yeteneği

$$Ö=D / (B + D) = GN / (GN + YP)$$

Gerçekte başarısız olanların % kaçının geliştirilen test sonucunda başarısız olduğunu gösterir ve gerçek başarısızlar içinde, geliştirilen testin başarısızları ayırtedebilme yeteneği olarak tanımlanır.

Doğruluk Oranı (accuracy): Testin toplam doğru tanı koyma oranı

$$DO=(A + D) / N = (GP + GN) / (GP +GN + YP + YN)$$

Doğruluk oranı geçerlilik katsayısı demektir. Bu değer 0,5'in altında ise geliştirilen test ile yapılan sınıflandırmanın şans eseri ortaya çıktığı yorumu yapılır. Bu değer 1'e yakın olması istenir.

Pozitif Tanımlama Oranı (predictive value+): Test pozitifler içerisinde doğruluk oranı

$$PV^+ = A / (A + B) = GP / (GP + YP)$$

Test sonucunda başarılı olanların gerçekten başarılı olma olasılığını verir. Pozitif kestirim değeri diye adlandırılır.

Negatif Tanımlama Oranı (predictive value-): Test negatifler içerisinde doğruluk oranı

$$PV^- = D / (C + D) = GN / (GN + YN)$$

Test sonucunda başarısız olanların gerçekten başarısız olma olasılığını verir ve negatif kestirim değeri adını alır.

3.1.4.Yapı (Kavram) Geçerliliği (Construct validity):

Yapı geçerliliği doğrudan ölçülemeyen bir özelliği ölçen bir testin ölçme derecesi olarak tanımlanabilir. İlgili yapıyı (psikolojik özellikler doğrudan ölçülemeyen özelliklerdir.) ölçeceği düşünülen değişkenlerin oluşturduğu bütünün, öngörülen yapıyı belirleyip belirlemediğinin incelenmesi gerekir (Alpar, 2016).

Ölçme aracını oluşturan maddelerin kuramsal yapısına ilişkin bir geçerlilik kavramıdır. Bir testteki sorular birbiriyle ilişkili olmalı ve aynı yapıyı oluşturmalıdır. Maddelerin neden-sonuç ilişkileri bir bütünlük içerisinde testinin sonucunu gösterir. Bu geçerlilik türü faktör analizi diye bilinen istatistik yöntemle analiz edilerek bulunur. Bu analizde amaç ilişkili maddeleri bir araya toplayarak yeni sanal maddeler (faktörler) oluşturmaktır. Yada önceden geçerliliği saptanmış bir araç ile yeni ölçme aracı karşılaştırılarak veya sonucu bilinen bir gruba uygulanarak yapı geçerliliği belirlenebilir.

3.1.4.Yapı (Kavram) Geçerliliđi (Construct validity):

Ölçme aracı iki gruba uygulanır, bu gruplardan birisi sonuçları bilinen bir gruptur. Ölçme sonunda bilinen grubun yeni ölçme aracı ile beklenen sonuçları uyumlu ise ölçme aracının yapısal geçerliliđi olduđu söylenebilir. Genelde bu şekilde bir grubu bulmak zor olduğundan yapısal geçerlilik faktör analizi ile karara bağlanır.

Aslında yapı geçerliliđi araştırmadaki her kavramın ölçmeye başlamadan önce çok net tanımlarının yapılması ile yakından ilişkilidir.

Testin ölçülmek istenen davranış bağlamında soyut bir kavramı (faktörü) doğru biçimde ölçebilme derecesidir. Bireyin tutum, performans, yetenek gibi özelliklerini ölçmek için sorular sorulabilir. Hazırlanan bu soruların belirtilen özellikleri ne ölçüde doğru ölçtüğü yapı geçerliliği ile araştırılır.

Yapı geçerliği bir testin içerdiği boyutların ve maddelerin belli bir teorik yapıyı (kavramı) ölçmeye yönelik olduklarının ortaya konulması ile ilgilidir (Anastasi, 1982; Öner, 1994). Yapı geçerliği, gerceğin değerlendirmesi için düzenlendiği olguyu değerlendirme derecesinin ölçer (Creswell, 1994). Yapıyla bağlantılı kategoride sınıflanan bilgiler, test puanlarını, ilgilenen psikolojik özelliğin ölçümü olarak ele alır. Bir testin ilgilendiği yapı, kavramsal bir çerçeve içinde ele alınmalıdır. Kavramsal çerçeve yapının anlamını, diğer yapılardan farklarını belirler ve bu yapıya ait ölçümlerin diğer yapılarla nasıl bağlantılı olması gerektiğini gösterir (Hovardalıoğlu, Sezgin, 1998).

3.2.FAKTÖR ANALİZİ (FACTOR ANALYSIS)

Faktör analizi, birbiri ile ilişkili p tane değişkeni bir araya getirerek az sayıda ilişkisiz ve kavramsal olarak anlamlı yeni değişkenler (faktörler, boyutlar) bulmayı amaçlayan çok değişkenli bir istatistiktir.

Faktör analizinin amacı, veri setini küçülterek daha kolay açıklanabilir hale getirmektir. Faktör analizi bir çok değişkenin birkaç başlık altında toplanması tekniğidir. Mesela bir ankette 100 madde olsun. Anket sonucunda deneklerin; sözel, matematiksel ve analitik kabiliyetleri değerlendirilmek istenmiştir. Faktör analizi ile kabiliyetlerin her birisi için “faktör skoru” elde edilebilir. Analiz ile üçten daha az veya daha fazla birbirinden farklı faktörün olup olmadığı ortaya çıkarılır.

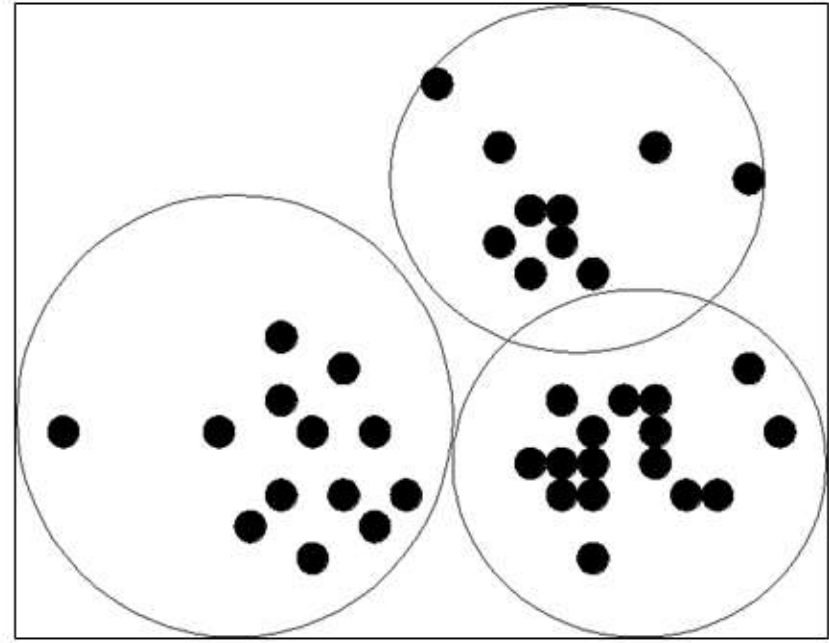
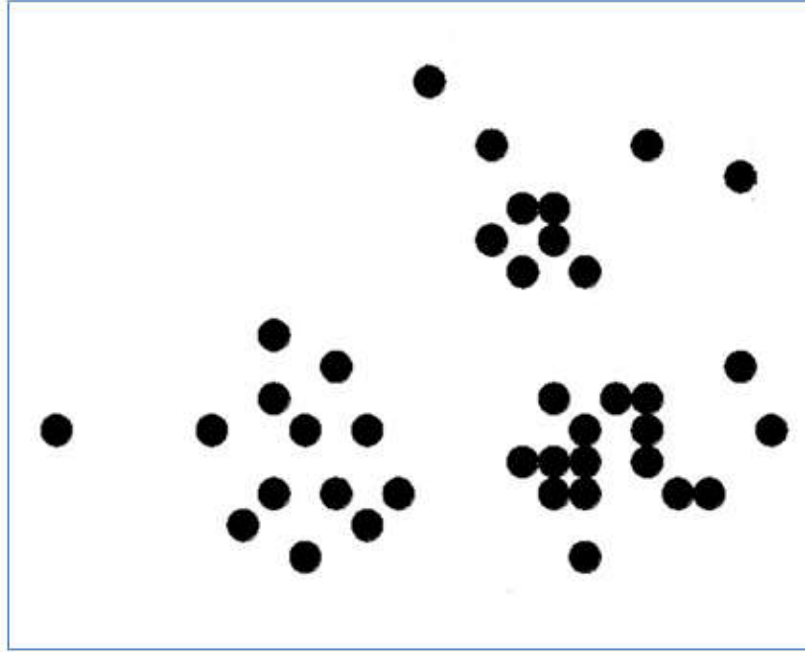
Bir galaksi evren içinde bir faktör gibidir.



Birçok deęişkenin varyansı onu oluřturan kümelenmiř yapılar ve onların arasındaki korelasyonlarca açıklanabilir.

Faktör Analizinin Kavramsal Modeli

FA, ana yapıyı oluşturan kümeleri incelemek için maddeler arasındaki korelasyonları kullanır.



Amaçlar

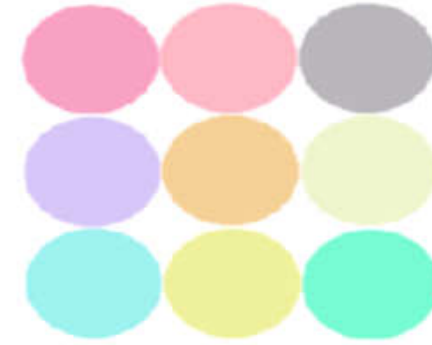
Faktör analizi tekniklerinin iki ana kullanım amacı bulunmaktadır:

- 1. Veri azaltma (Data reduction):** Değişkenlerin sayısının daha az sayıda faktöre indirgenmesi
- 2. Kuram Geliştirme:** Değişkenlerin kendi aralarındaki ilişkileri kullanarak yapıyı (structure) tanımlamak

Faktör analizinin Temel Sorusu: Kaç faktör / bileşen?



Tek Faktör?



Dokuz Faktör?



Üç Faktör?

Basit model

Her soru yalnızca bir faktörden yük alır

Soru 1

Soru 2

Soru 3

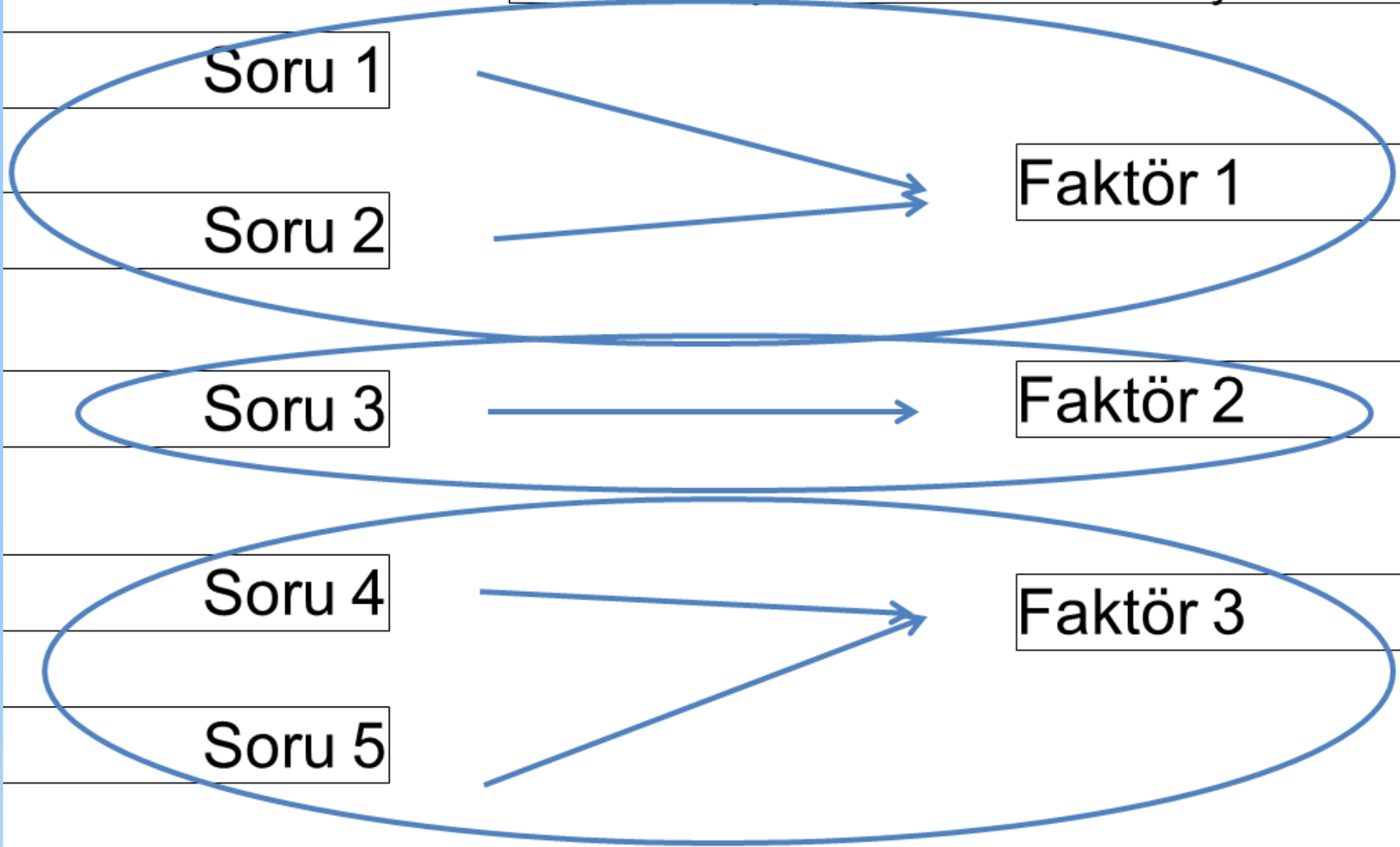
Soru 4

Soru 5

Faktör 1

Faktör 2

Faktör 3



Kompleks Model

Her soru birden fazla faktörden yük alır

Soru 1

Soru 2

Soru 3

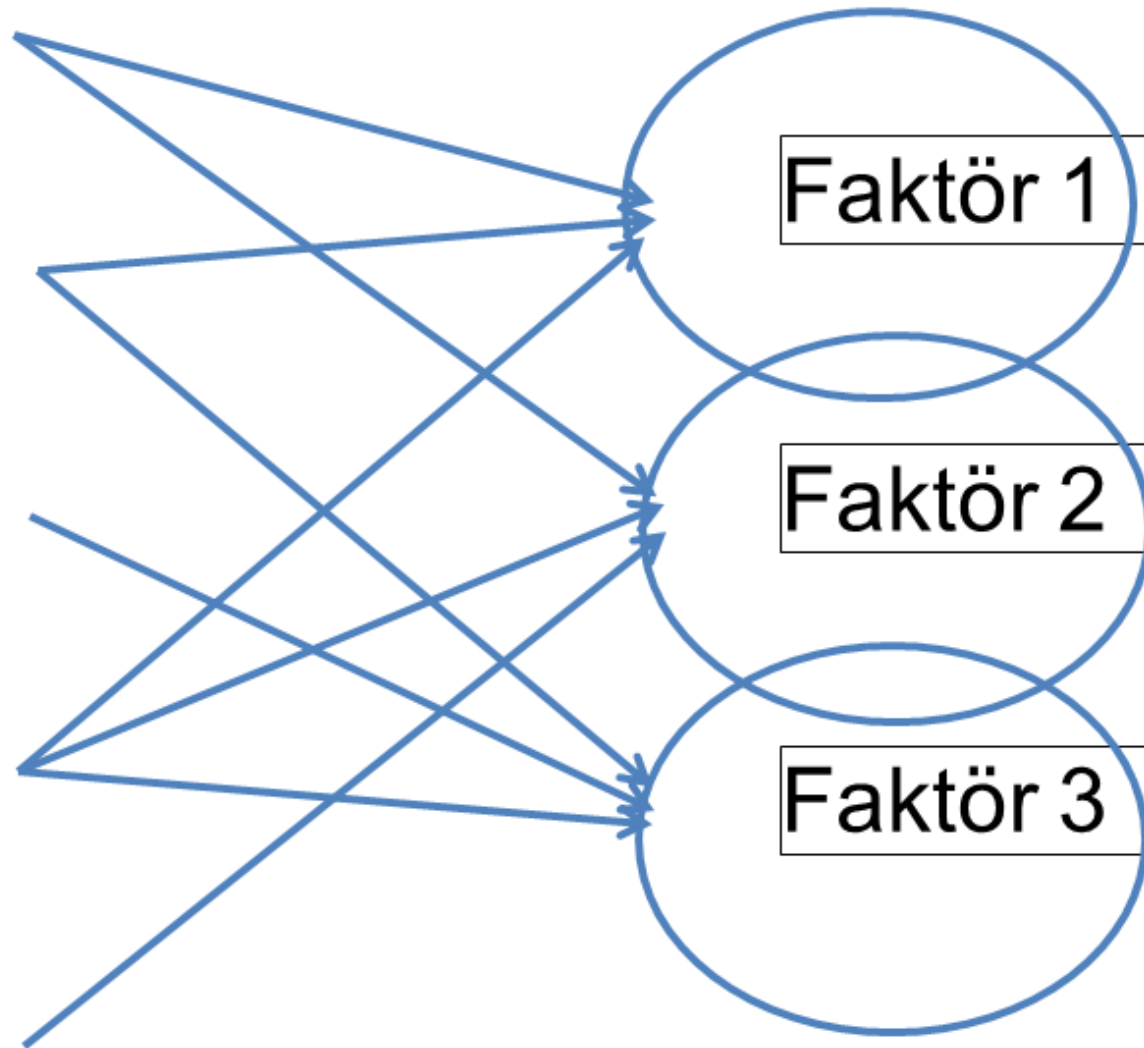
Soru 4

Soru 5

Faktör 1

Faktör 2

Faktör 3



Faktör analizinde arařtırmacı öncelikle arařtırma bağlamında kullanılan deęişkenler setinin temelini oluřturan ana (temel) faktörlerin neler olduęu ve bu faktörlerden her birinin deęişkenlerden her birini açıklama derecesini de görme imkanına sahip olmaktadır. Bu yolla, arařtırmacı elinde bulunan çok sayıda deęişkenden oluřan deęişkenler setini daha az sayıda yeniden oluřturulmuř deęişkenler (faktörler) cinsinden ifade etme ve anlama imkânına sahip olacaktır.

Etkin bir faktör analizinin toplam veri setini en iyi temsil edebilen ancak mümkün olduęunca az sayıda faktörden oluřan bir çözüm olması arzu edilen özelliktir. İyi bir faktör analizi sonucunun hem basit hem de yorumlama becerisinin iyi olması arzu edilir.

Faktör analizi, başlıca amacı aralarında iliřki bulunduęu düşünölen çok sayıdaki deęişken arasındaki iliřkilerin anlaşılmasını ve yorumlanmasını kolaylařtırmak için daha az sayıdaki temel boyuta indirgemek veya özetlemek olan bir grup çok deęişkenli analiz teknięine verilen genel bir isimdir. Dięer bir ifade ile faktör analizi, aralarında iliřki bulunan çok sayıda deęişkenden oluřan bir veri setine ait temel faktörlerin (iliřkinin yapısının) ortaya çıkarılarak arařtırmacı tarafından veri setinde yer alan kavramlar arasındaki iliřkilerin daha kolay anlaşılmasına yardımcı olmaktadır.

3.2.1. Açımlayıcı faktör Analizi

Her bir maddenin hangi diğer maddelerle birliktelik gösterdiğini, bu maddelerin bu gruplara ne kuvvetle bağlandığını görmek amacıyla açımlayıcı faktör analizi; bir kültürde geliştirilen ölçeğin başka bir kültüre uyarlanmasında doğrulayıcı faktör analizi kullanılır.

Yapı geçerliliği açıklayıcı (exploratory) faktör analiziyle incelenir. Faktör analizi tekniği, psikolojik boyutların tanınmasında ve boyutların içeriği hakkında bilgi edinilmesinde kullanılan bir istatistik yaklaşımdır. Faktör analizi test maddelerinin arasındaki ilişkilerden çıkartılan ortak boyutların saptanması işlemidir. Böylece eldeki testle örneklenen davranış, birkaç ortak faktöre indirgenerek, bu davranışın içeriği hakkında bilgi edinilir.

Yapı geçerliliğini incelemek amacıyla **faktör analizi**, iç tutarlılık analizi ve hipotez testlerinden yararlanır.

Faktör analizi ile yapısal geçerlilikte; tek boyutluluğun kanıtlanması, çift boyutluluğun kanıtlanması, boyutlar arasındaki ilişki ve faktöriyel yapıyı doğrulama incelenir.

Açımlayıcı (exploratory) ve doğrulayıcı (confirmatory) olmak üzere iki tür faktör analizi yaklaşımı vardır.

Açımlayıcı faktör analizi (exploratory factor analysis) temel bileşenler analizidir (Kline,1994). Analize alınan maddelerin (değişkenin) öz değeri 1' den büyük olanlar faktör olarak alınır. Faktörlerin ölçeğe ilişkin açıkladıkları toplam varyans bulunur. Önemli faktör sayısını belirlemede Kaiser'in özdeğer \geq 1 kuralı kullanılır. (Kaiser, 1960).

Açımlayıcı faktör analizinde araştırmacı araştırma yaptığı konuyla ilgili olarak değişkenler arasındaki ilişkiye yönelik olarak herhangi bir fikrinin veya öngörüsünün olmaması sebebiyle değişkenler arasındaki muhtemel ilişkiyi ortaya çıkarmaya çalışır.

Doğrulayıcı faktör analizinde ise araştırmacı tarafından daha önceden belirlenen bir ilişkinin doğruluğunu test etmek amaçlanmaktadır.

SPSS'de faktörlerin ortaya çıkarılması değişik metotlar kullanılır. Bunlardan en yaygın kullanılanı Temel Bileşenler (Principal Components) metodudur. Bu metotla bütün değişkenlerdeki maksimum varyansı açıklayacak varyans hesaplanır. Kalan maksimum miktardaki varyansı açıklamak için, ikinci faktör hesaplanır. Bu süreç değişkenlerdeki bütün varyansın açıklanmasına kadar devam eder. Ancak çok faktörün olması iyi değildir. Bunun için öz değer (eigenvalue) kullanılarak, analizde kaç faktörün kullanılacağına karar verilir.

Faktör Analizinden değişkenler arasındaki ilişki doğrusaldır. Faktör analizinde yer alan değişkenler aşağıdaki gibi ifade edilir.

$$X_i = A_{i1}F_1 + A_{i2}F_2 + \dots + A_{im}F_m + U_i$$

X_i : Standartlaştırılmış i.nci değişken

A_{ij} : i.nci değişkenin j.nci ortak faktörü üzerindeki etkisine ilişkin standartlaştırılmış çoklu regresyon katsayısı

F : Ortak faktör

U_i : i.nci değişken için eşsiz faktör (ortak faktörlerce açıklanamayan kısım)

m: Ortak faktör sayısı

Faktör analizi neticesinde elde edilen ortak faktörler değişkenlerin lineer kombinasyonu olarak aşağıdaki gibi gösterilir:

$$F_i = W_{i1}X_1 + W_{i2}X_2 + \dots + W_{ik}X_k$$

F_i : i.nci faktörün tahmini

W_i : Faktör skoru (katsayısı)

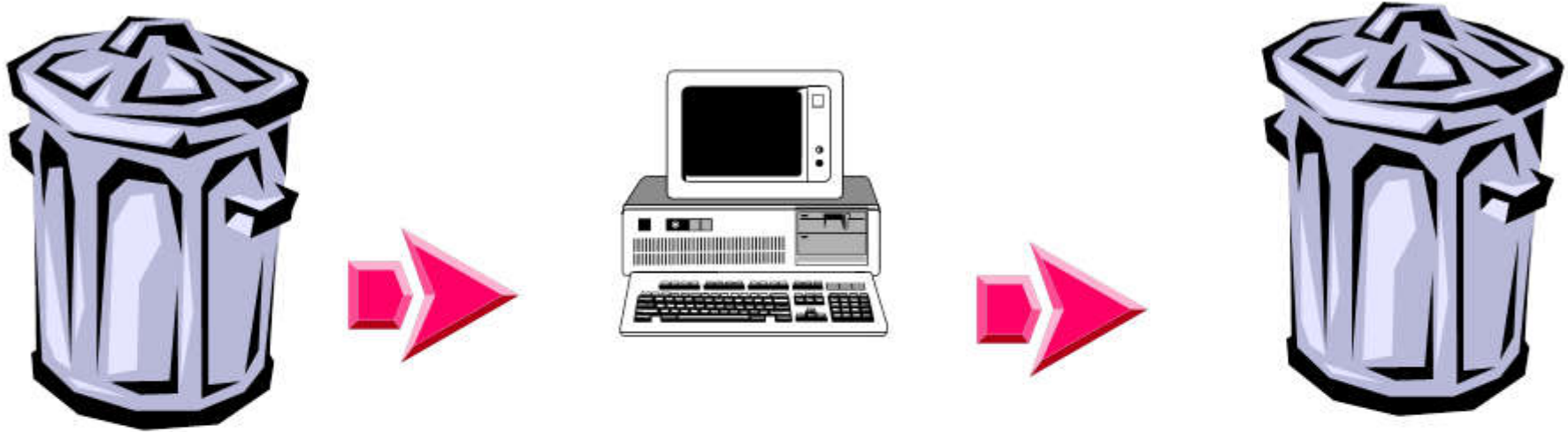
k : Değişken sayısı

Analiz neticesinde kaç faktör elde edilmişse o kadar denklem kurulur. Birinci faktörün ağırlığı (katsayısı) en yüksek olur.

AFA VARSAYIMLARI

- 1 Teorik Uygunluk
- 2 Örneklem büyüklüğü
- 3 Ölçek seviyesi
- 4 Normallik
- 5 Doğrusallık
- 6 Uç Değerler
- 7 Faktörlenebilirlik

Teorik uygunluk



- Literatürü tara, gözden geçir
- Teorik olarak uygun maddeleri kullan.
Uymayanları ele.

Ölçek Düzeyi

Bütün deęişkenler (maddeler) korelasyonel analizler için uygun olmalıdır.

yani eşit aralıklı (interval) ya da eşit oranlı (ratio/metric) ölçek türünde olmalıdır.

Faktörlenebilirlik:
Anti-image Korelasyon matrisi

Anti-image korelasyon matrisi tablosunda madde eliminasyonu için .50 sınır değeri olarak kabul edilir. Düşük değerler, maddenin diğer maddelerle yeterli korelasyonunun olmadığını gösterir.

AFA İKİ YAKLAŞIM İÇERİR

- Tüm varyanslarla yapılan
Temel Bileşenler Analizi
(Principle Components - PC)
- Ortak varyansla yapılan:
Temel Eksen Faktör Analizi
Principle Axis Factoring (PAF)

Temel Bileşenle Analizi (PC)

- Daha yaygındır.
- Daha pratiktir.
- Diğer analizlerde kullanılmak üzere puan hesaplamak ve verileri azaltmak için kullanışlıdır.
- Tüm maddeler için varyansların tamamı analize girer.

Temel Eksen Faktör Analizi (PAF)

- Daha az yaygındır
- Daha kuramsaldır.
- Sadece ortak (shared) varyansları kullanır (Yani özgül varyanslar dışlanır)

Ortak Yüklere (Communalities)

- Her değişkenin (maddenin) bir ortak varyansı bulunmaktadır.
- Bu değer 0 ile 1 arasında değişir.
- PCA ve PAF yaklaşımlarında farklı ortak yük anlayışı tabloya yansır.

Ortak Ykler

- Yksek Ortak Ykler ($>.50$):

ıkan faktrler, analize alınan maddelerin varyansın daha fazlasını aıklamasına neden olur.

- Dşk Ortak Ykler ($<.50$):

Deęerler dşkse yorumu zor daha fazla faktr ıkabileceęini kabul et ya da bu maddeyi elemeyi dşn.

Ortak vükler

Communalities

	Initial	Extraction
mad1	1,000	,759
mad2	1,000	,799
mad3	1,000	,780
mad4	1,000	,804
mad5	1,000	,829
mad6	1,000	,723
mad7	1,000	,699
mad8	1,000	,799
mad9	1,000	,810
mad10	1,000	,770
mad11	1,000	,674
mad12	1,000	,547
mad13	1,000	,546
mad14	1,000	,599
mad15	1,000	,735

Extraction Method: Principal Component Analysis.

Communalities

	Initial	Extraction
mad1	,740	,713
mad2	,815	,762
mad3	,779	,742
mad4	,784	,776
mad5	,811	,812
mad6	,749	,664
mad7	,721	,655
mad8	,800	,792
mad9	,768	,802
mad10	,732	,745
mad11	,382	,512
mad12	,508	,564
mad13	,570	,585
mad14	,464	,552
mad15	,453	,584

Extraction Method: Principal Axis Factoring.

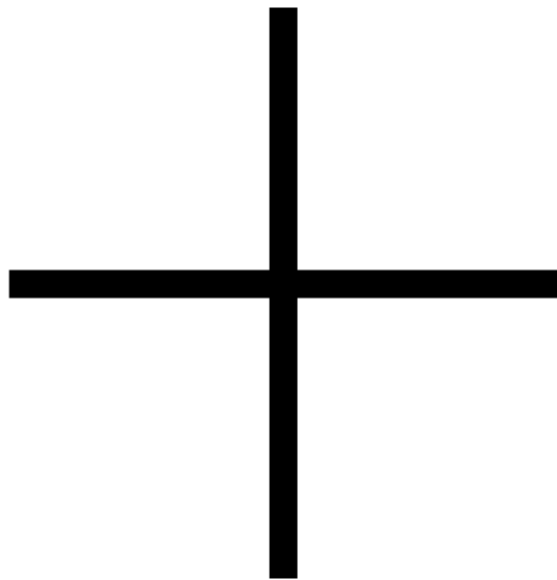
Eigen Deęeri (Özdeęer) (Korelasyonların kareleri toplamı)

- Her faktörün bir eigen deęeri vardır. Eigen deęeri her faktörün açıklama gücünü gösterir
- Ardışık olarak sıralanan faktörler için eigen deęeri giderek düşer.
- Genel olarak: Kaiser kriterlerine göre 1'in üzerindeki eigen deęerleri «kararlı» kabul edilir.
- Eigen deęeri yüzde (%) olarak da ifade edilebilir.
- Tüm eigen deęerlerin toplamı madde sayısını verir.

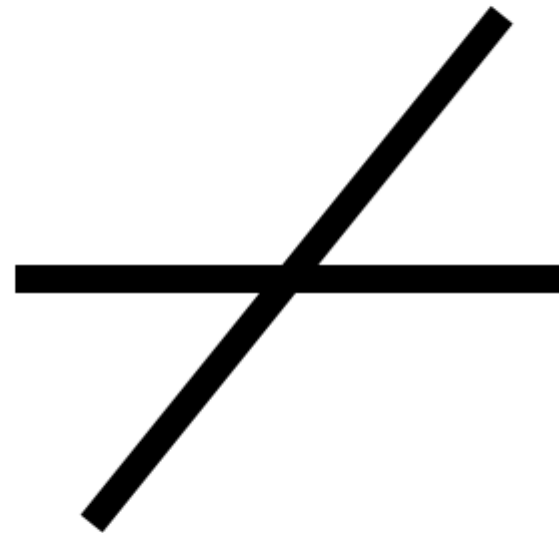
Scree plot

- Eigen değerinin grafik gösterimidir.
- Her faktörün açıkladığı varyans miktarını gösterir.
- Kırılma noktaları arasındaki değişim dikkate alınır.
- İlk faktör en yüksek varyansı açıklar.
- En son faktör en düşük varyansı açıklar.

Faktör Döndürmenin iki Temel Türü

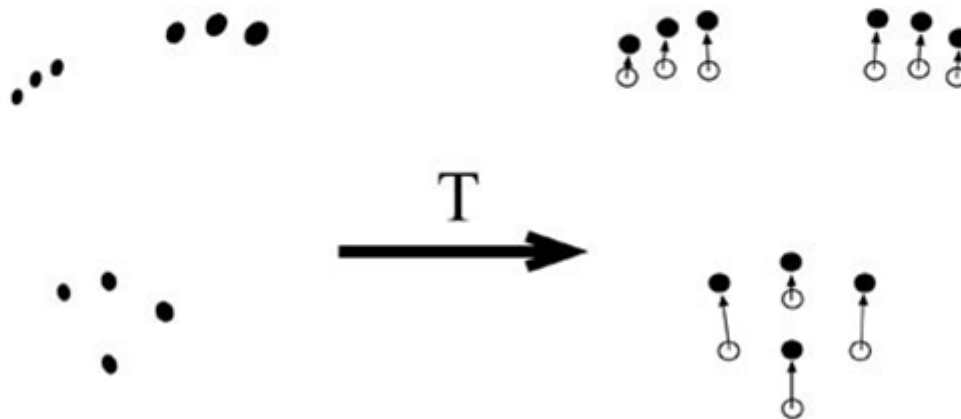
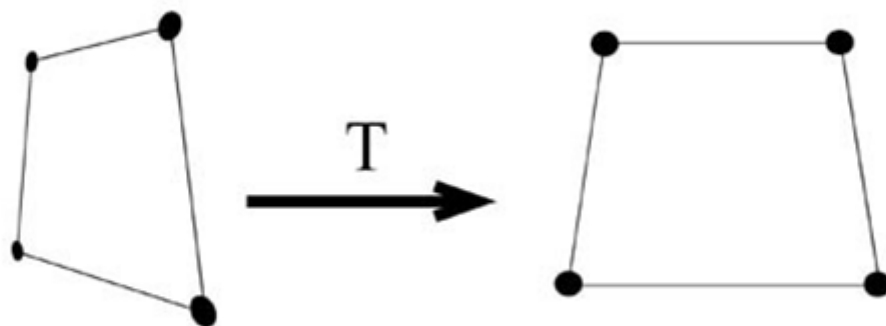


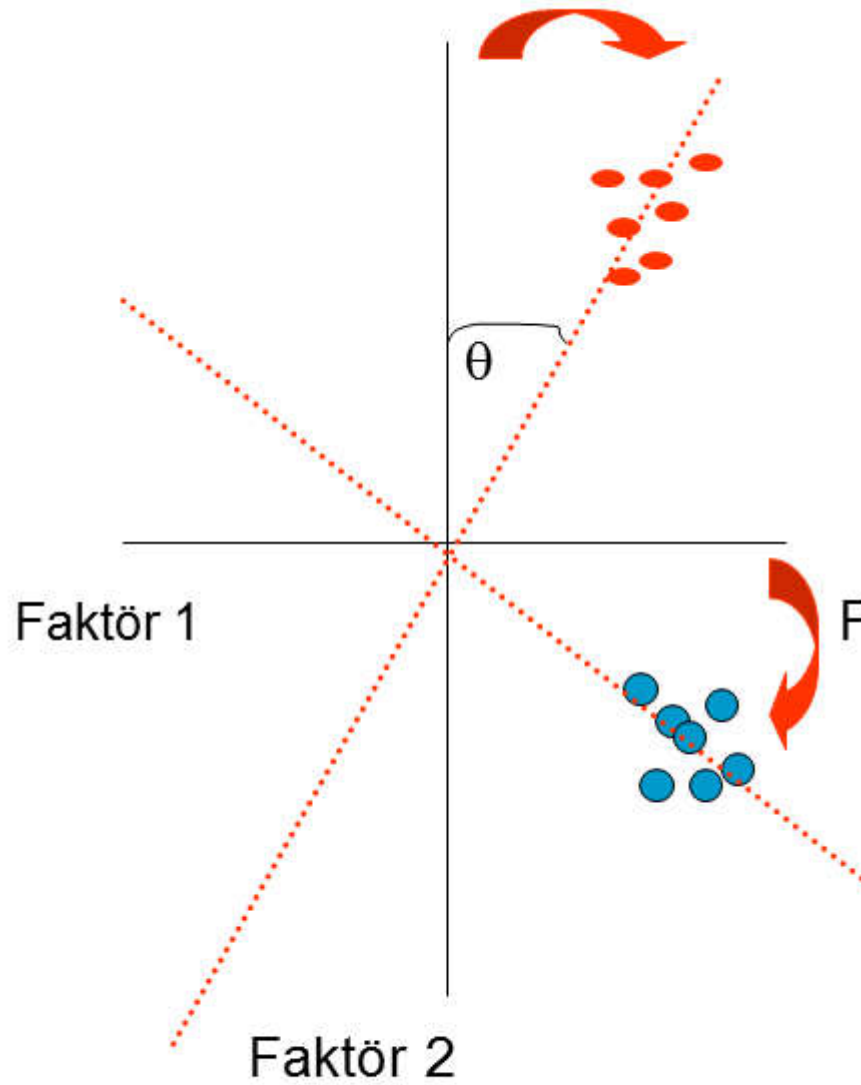
Orthogonal / Dik
(SPSS Varimax)



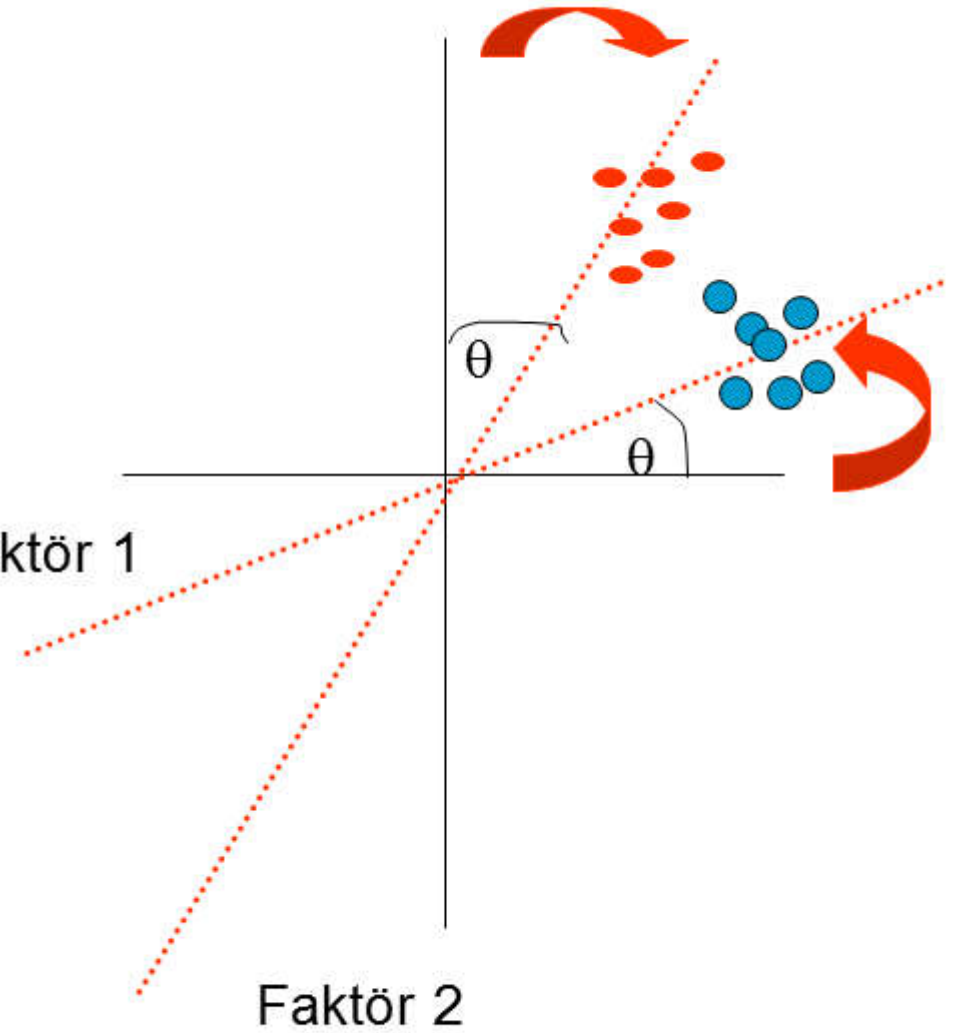
Oblique / Eğik
(SPSS Oblimin)

Döndürmenin Mantığı





İlgisiz (orthogonal) rotasyon



Eđik (oblique) rotasyon

Faktör Analizinin Adımları :

1. Bütün değişkenler için korelasyon matrisi hesaplanır. Birbiri ile ilişkisiz olan değişkenler belirlenir. Böylece faktör modelin uygunluğu değerlendirilir.
2. Faktör sayısı belirlenir. Bu adımda seçilen modelin veriye ne kadar uyumlu olduğu tespit edilir.
3. Rotasyonla faktörler dönüştürülerek, daha iyi yorumlanması sağlanır.
4. Her vaka için faktörün skoru hesaplanır.

m kadar önemli faktör “ bağımsızlık, yorumlamada açıklık ve anlamlılık” için bir eksen döndürmesine (rotation) tabi tutulur. Eksenlerin döndürülmesi sonrasında maddelerin bir faktördeki yükü artarken, diğer faktörlerdeki yükleri azalır. Dik (orthogonal) ve eğik (oblique) olmak üzere iki tür döndürme yaklaşımı vardır. Döndürme sonunda değişkenlerle ilgili toplam varyans değişmezken, faktörlerin açıkladıkları varyans değişir.

Soysal bilimlerde genellikle dik döndürme tercih edilir. Yani faktörler eksenin konumu değiştirilmeksizin (90 derece) döndürülür. Dik döndürme tekniklerinden en sık kullanılan varimax ve quartimax'dır. Bu teknikler maddelerin yüklerini bir faktörde 1'e, diğerlerinde ise sıfıra yaklaştırmayı amaçlar.

Faktör analizi çeşitli aşamalardan oluşan bir analiz tekniğidir. Tipik bir faktör analizinde yer alan aşamalar aşağıdaki şekilde özetlenebilir.

1. Problem tanımı ve veri toplama: Bu aşama faktör analizi için gerekli olan hazırlık çalışmalarını kapsayan ilk aşamadır. Bu aşamada faktör analizinin amacı ve faktör analizinde kullanılacak olan değişkenlerin teori, mevcut araştırmalar ve araştırmacının bilgi ve tecrübeleri veya yaptığı ön çalışmalar (kalitatif veya kantitatif türdeki çalışmalarla) ışığında geliştirilmesi ve uygun ölçüm araçları ile ölçülmesi ve makul yöntemlerle verilerin toplanması işlemleri yapılmaktadır.

2. Korelasyon matrisinin oluşturulması: Faktör analizinin ikinci aşaması analiz sürecinin başladığı aşama olup, bu aşamada korelasyon matrisi oluşturulur. Korelasyon matrisi faktör analizinde yer alan değişkenler arasındaki ilişkiyi gösteren bir matristir.

3. Faktör sayısına karar verme: Üçüncü aşama ise, söz konusu veri seti için faktör analizinin uygun olduğuna karar verdikten sonra, oluşturulan korelasyon matrisini baz alarak, faktör çözümünü ortaya koymak amacıyla uygun bir faktör çıkarma (oluşturma) yönteminin seçilmesi ve başlangıç çözümünün oluşturulmasını kapsamaktadır.

4. Faktör Eksen Döndürme (Factor Rotation) Başlangıç faktör analizi çözümüne ulaşıldıktan sonra (başlangıç faktör matrisi) ortaya çıkan faktörlerin yorumlanması ve isimlendirilmesini kolaylaştırmak için faktörleri temsil eden eksenlerde çeşitli manipülasyonlar veya eksen kaydırmaları yapma yoluna gidilir.

Faktör matrisi faktörler cinsinden standardize edilmiş değişkenleri ifade eden katsayıları içerir. Bu katsayılar faktör yükleri olarak adlandırılıp, değişkenlerle faktörler arasındaki korelasyonları temsil eder. Mutlak değer olarak faktör yükünün büyüklüğü arttıkça değişken ve faktörün birbiriyle yakından ilişkili olduğunu göstermektedir (Altunışık, Coşkun, Bayraktaroğlu, Yıldırım, 2005: 212-226).

Faktör Analizinin Varsayımları

- ✓ Veriler oran ya da aralık ölçeğinde olmalıdır.
- ✓ Veriler çok değişkenli normal dağılım gösteren bir anakütleden çekilmiş olmalıdır. Her değişken çifti için iki değişkenli normal dağılım göstermelidir. Matris grafiklerinde ikili grafiklere bakılarak doğrusal yapı hakkında bilgi edinilebilir.
- ✓ Gözlem sayısı yeterli olmalıdır.

Gözlem Sayısı-Faktör Yüğü Arasındaki İlişki

Faktör yüklerine göre gözlem sayısının ne olacağı testin gücü ve örneklem genişliğine göre hesaplanabilir. Güç=0,80 ve alfa=0,05 ile faktör yükü için örneklem sayısı aşağıdaki gibidir (Hair, 2005).

Faktör Yüğü	Anlamlılık için gereken gözlem sayısı
0,30	350
0,35	250
0,40	200
0,45	150
0,50	120
0,55	100
0,60	85
0,65	70
0,70	60
0,75	50

VERİ SETİNİN FAKTÖR ANALİZİ İÇİN UYGUNLUĞUNUN DEĞERLENDİRİLMESİ

Veri setinin faktör analizi için uygun olup olmadığı Korelasyon Matrisi, Kaiser-Meyer-Olkin (KMO) testi ve Bartlett Küresellik testi ile değerlendirilir.

1. Korelasyon Matrisi:

Değişkenler arasındaki korelasyonlar yüksek olmalıdır. Çünkü değişkenler arasındaki korelasyonlar ne kadar yüksek ise değişkenlerin ortak faktörleri oluşturma olasılıkları o kadar yüksek olacaktır. Korelasyonlar genelde 0,3 den büyük olması istenir.

Aralarında yüksek korelasyon olan değişkenler genel de aynı faktörde yer alırlar.

VERİ SETİNİN FAKTÖR ANALİZİ İÇİN UYGUNLUĞUNUN DEĞERLENDİRİLMESİ

2. Kaiser-Meyer-Olkin (KMO) testi (Örneklem yeterliliği ölçütü):

Bu test örneklem yeterliliğini ölçer ve örneklem büyüklüğü ile ilgilenir. Gözlenen korelasyon katsayıları büyüklüğü ile kısmi korelasyon katsayılarının büyüklüğünü karşılaştıran bir indekstir. Oran ne kadar yüksek ise veri seti faktör analizi yapmak için o kadar iyidir denilebilir. KMO değerinin yüksek çıkması ölçekteki her bir değişkenin ölçekteki diğer değişkenler tarafından mükemmel bir şekilde tahmin edilebileceği anlamına gelir (Kaiser, 1974).

KMO değeri aşağıdaki gibi yorumlanır (Sharma, 1996).

KMO değerinin yorumu:

<0,5	Zayıf
0,5-0,6	Orta
0,7-0,8	İyi
0,8-0,9	Çok iyi
0,9-1,0	Mükemmel

$$KMO = \frac{\sum_{i \neq j} \sum r_{ij}^2}{\sum_{i \neq j} \sum r_{ij}^2 + \sum_{i \neq j} \sum a_{ij}^2}$$

Burada r_{ij}^2 , maddeler arasındaki korelasyonları belirtmektedir. Bu korelasyonlar aynı zamanda korelasyon matrisinin köşegen dışı değerleridir, a_{ij}^2 ise *kısmi* korelasyonları vermektedir. Bu kısmi korelasyonlar özellikle SPSS kullanıcıları için anti-imaj¹ matrisinden elde edilebilmektedir (Pett vd., 2003).

$$ÖUÖ = \frac{\sum_{i \neq j} r_{ij}^2}{\sum_{i \neq j} r_{ij}^2 + \sum_{i \neq j} a_{ij}^2}$$

KMO tüm maddelerin/değişkenlerin oluşturduğu veri kümesi için geçerlidir. KMO'nun özel biçimi olan ve her bir madde/değişken için elde edilen örneklem uygunluk ölçüsü, ÖUÖ, (Measure of Sampling Adequacy-MSA) de söz konusudur.

3. Bartlett Küresellik Testi (Bartlett test of sphericity):

Verilerin öncelikle faktör analizine uygunluğu test edilir, eğer küresellik testi anlamlı bulunursa ($p < 0,05$) faktör analizine geçilir. Bartlett testi ile korelasyon matrisinde değişkenlerin en azından bir kısmı arasında yüksek oranlı korelasyonlar olduğu olasılığını test eder. Analize devam edilebilmesi için aşağıdaki hipotezin reddedilmesi gerekir.

H_0 : Korelasyon matrisi birim matristir.

Eğer $p < 0,05$ ise H_0 hipotezi red edilir yani değişkenler arasında korelasyon olduğu ve verinin faktör analizi için uygun olduğuna karar verilir. Ayrıca $p < 0,05$ ise verilerin çoklu normal dağılımdan geldiği varsayılır. $P > 0,05$ ise faktör analizine devam edilmez (Hair ve ark., 1998).

Küresellik testi özünde değişkenlere ilişkin korelasyon matrisinin, (değişkenler arasında ilişki yoktur varsayımına dayanan) birim matrise karşı test edilme ilkesine dayanır. Bu nedenle Bartlett testi aynı zamanda korelasyon matrisinin anlamlılığının bir testidir.

$$H_0: \mathbf{R} = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix}$$

Bartlett testi bir ki-kare istatistiğidir ve aşağıdaki formül ile elde edilir (Pedhazur ve Schmelkin, 1991).

$$\chi^2 = - \left[(N-1) - \left(\frac{2k+5}{6} \right) \right] \ln |\mathbf{R}|$$

Burada

χ^2 : Bartlett testinden elde edilen değerdir.

N: örneklem genişliği,

k: madde sayısı

$|\mathbf{R}|$: verilerden elde edilen korelasyon matrisinin determinantıdır. Bu determinant değeri aynı zamanda genelleştirilmiş varyansı vermektedir.

Ki-kare test istatistiği için elde edilen serbestlik derecesi ise ikili karşılaştırmanın karşılığı olan

$$s.d. = \binom{k}{2} = k(k-1)/2 \text{ şeklindedir.}$$

Faktör Sayısının Belirlenmesi

Değişkenler arasındaki ilişkileri en yüksek derecede temsil edecek az sayıda faktör belirlenir. Aşağıdaki üç duruma göre faktör sayısı belirlenebilir.

i) Özdeğerlere (Eigenvalues) Göre Belirleme:

Özdeğeri bir ve birden büyük ($\lambda \geq 1$) olan faktörler hesaba katılır. Özdeğeri 1'den küçük olan faktörler hesaba katılmaz. Özdeğer bir faktör tarafından açıklanan toplam varyansı gösterir.

ii) Serpilme Diyagramına (Scree Plot) Göre Belirleme:

Özdeğerlerin grafiği çizilir. Düşey çizgini yataylaştığı yere kadar olan faktörler çözüme dahil edilir. Diyagramda x eksenini faktörleri, y eksenini özdeğerleri gösterir (Lewis-Beck, 1994).

iii) Varyans Oranına Göre Belirleme:

Özdeğerlerin açıkladıkları birikimli varyans miktarının toplam varyansı tahmin edici bir düzeye gelene kadar faktörler modele alınır. Analiz sonucunda elde edilen varyans oranları ne kadar büyükse faktör yapısı da o kadar güçlü olur. Genelde %40' üzerinde olması istenir.

Rotasyonlu Faktör Matrisi

Faktör rotasyonunda amaç isimlendirilebilir ve yorumlanabilir faktörler elde etmektir. Modelin kaç faktörden oluştuğu belirlendikten sonra, her faktörde yer alacak değişken sayısı ve değişkenlerin bu faktörlere göre dağılımı belirlenir.

Elde dileyen olan faktör yükleri faktörlerle ilgili standardize edilmiş bir değişkeni ifade eder. Faktör yükleri değişkenlerin her faktördeki ağırlığını ifade eder. Bu değerler değişkenlerle seçilen faktörler arasındaki ilişki derecesini gösterir. Bir değişkenin hangi faktörle en güçlü korelasyon ilişkisi varsa, o faktörün elemanı demektir (Nakip, 2003).

Rotasyon işlemlerinde ortogonal (dik) ve oblique (eğik) döndürme yöntemleri kullanılır. Ortogonal yöntemde faktörlerin birbiri ile ilişkiye girmemesini sağlar. Oblique (eğik) yöntemde ise faktörler birbirinden bağımsız değildir. Genelde ortogonal (dik) yöntem tercih edilir.

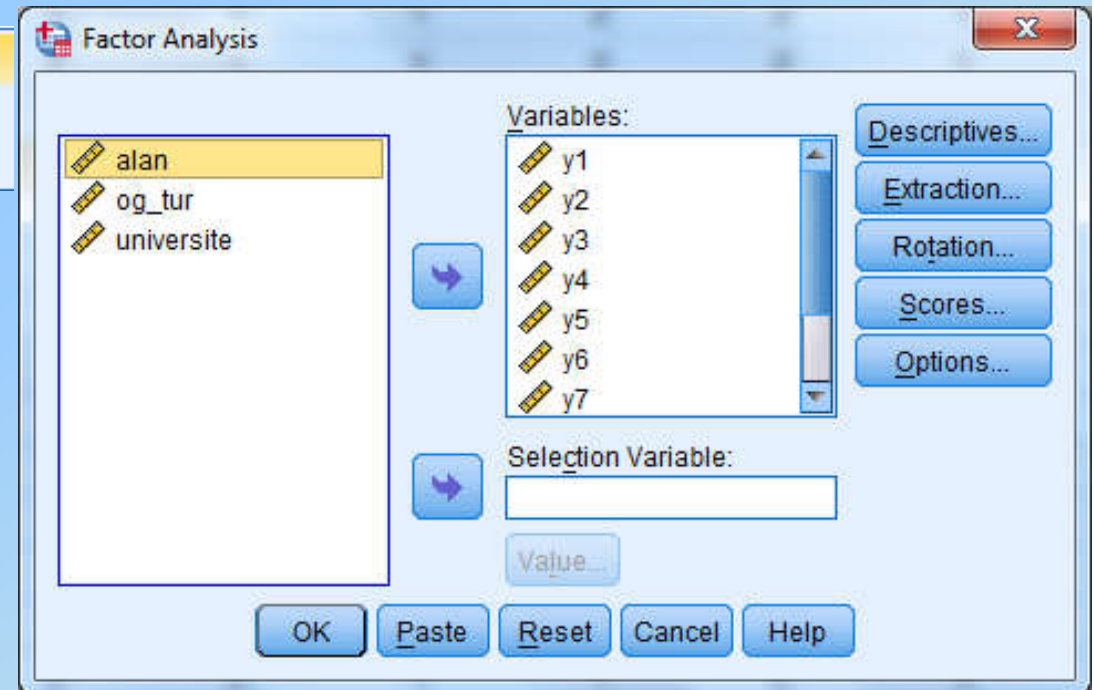
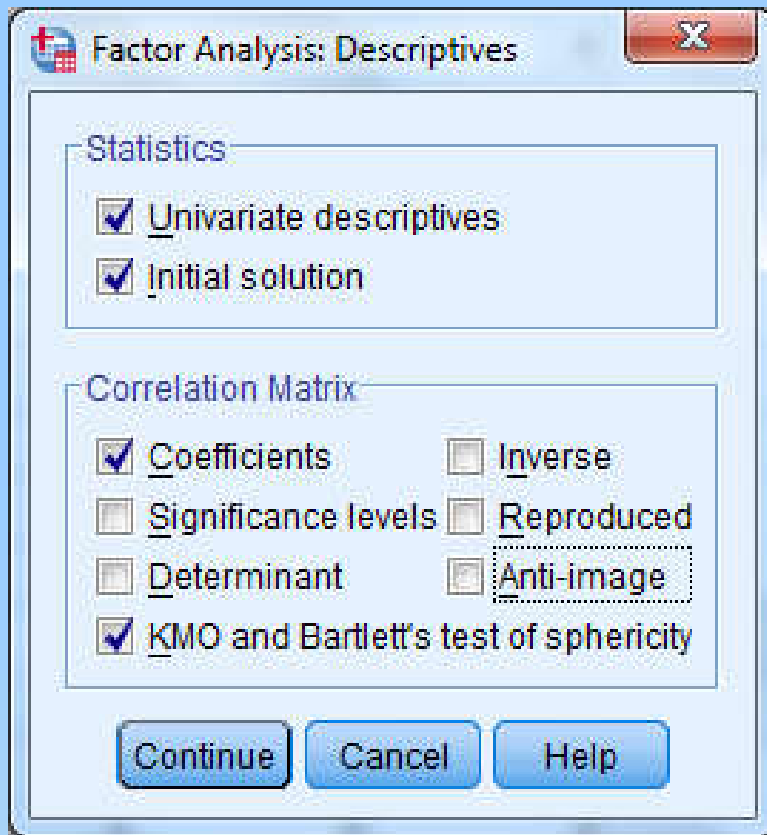
Ortogonal yöntemler içinde en yaygın kullanılan yöntem ise Varimax yöntemidir. Varimax yöntemi basit ve anlamlı faktörlere ulaşmada, faktör yükleri matrisinin sütunlarına öncelik verir. Böylece daha az değişkenle faktör varyanslarının max olması için döndürme yapılır.

Faktör analizinde faktör ağırlıkları orijinal değişken bağımlı değişken ve faktörlerin bağımsız değişkenler olduğu çoklu regresyon eşitliğinde standardize edilmiş regresyon katsayılarıdır.

Faktörlerin Adlandırılması

Faktörde yer alacak deęişken sayısı ve deęişkenlerin bu faktörlere göre dağılımı belirlendikten sonra faktörlere genel bir isim verilir. Eğer ilgisiz deęişkenler bir faktörde toplanmışsa, faktör yükü en fazla olan deęişken esas alınır.

SPSS'DE FAKTÖR ANALİZİ



Factor Analysis: Extraction

Method: Principal components

Analyze

- Correlation matrix
- Covariance matrix

Display

- Unrotated factor solution
- Scree plot

Extract

- Based on Eigenvalue
Eigenvalues greater than:
- Fixed number of factors
Factors to extract:

Maximum iterations for Convergence:

Continue Cancel Help

Factor Analysis: Options

Missing Values

- Exclude cases listwise
- Exclude cases pairwise
- Replace with mean

Coefficient Display Format

- Sorted by size
- Suppress small coefficients

Absolute value below:

Continue Cancel Help

Factor Analysis: Factor Scores

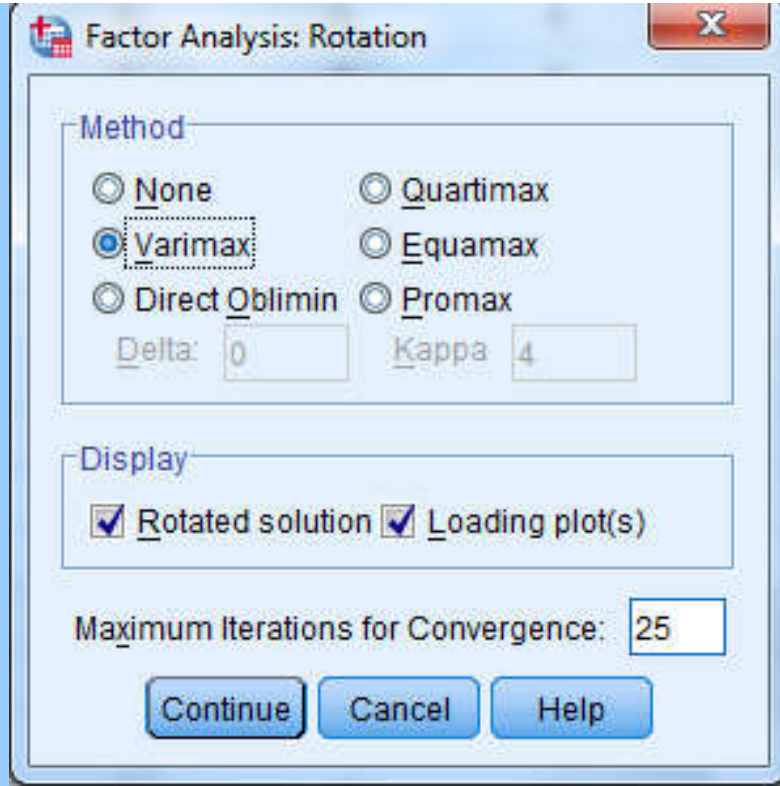
- Save as variables

Method

- Regression
- Bartlett
- Anderson-Rubin

- Display factor score coefficient matrix

Continue Cancel Help



Arařtırmacı, bir faktör analizi tekniğini uygulayarak elde ettiđi p kadar önemli faktörü, “bağımsızlık, yorumlamada açıklık ve anlamlılık” sağlamak için bir **eksen döndürmesine (rotation)** tabi tutar. Eksenlerin döndürülmesi sonucunda maddelerin bir faktördeki yükü artarken, diđer faktörlerdeki yükleri azalır. Böylece faktörler, kendileriyle yüksek ilişki veren maddeleri bulurlar ve faktörler daha kolay yorumlanabilir.

Dik Döndürme (Orthogonal Rotation): Ölçme aracında yer alan alt faktörlerin birbiriyle ilişkisiz olduğunun varsayıldığı ya da belirlendiđi zaman ortogonal döndürme tekniđi kullanılır. Faktörler eksenlerin konumu deđiştirilmeksizin 90 derecelik açıyla döndürülür. Varimax, Quartimax, ve Equamax ortogonal döndürme tekniđidir.

Eđik Döndürme (Oblique Rotation): Ölçme aracında yer alan faktörlerin birbiriyle ilişkili olduğuna varsayıldığında ya da belirlendiđi zaman ortogonal olmayan döndürme teknikleri kullanılır. Eđik döndürmede her faktör birbirinden bağımsız olarak döndürülür. Eđik döndürmede açıklanan toplam varyans oranı deđişmezken, faktörlerin açıkladığı varyans oranlarında deđişme olabilir. Eđik döndürme teknikleri Direct Oblimin ve Promax teknikleridir.

Sosyal bilimlerde genellikle dik döndürme tercih edilir. Dik döndürme tekniklerinden en sık kullanılan varimax ve quartimax'dır. İki teknik de maddelerin yük değerlerini bir faktörde 1,0'a ve diğerlerinde ise 0,0'a yaklaştırmayı amaçlar.

Factor Analysis: Rotation

Method

None Quartimax

Varimax Equamax

Direct Oblimin Promax

Delta: 0 Kappa: 4

Display

Rotated solution Loading plot(s)

Maximum Iterations for Convergence: 25

Continue Cancel Help

Correlation Matrix

	yontem1	yontem2	yontem3	yontem4	yontem5	yontem6	yontem7	yontem8	yontem9	yontem10
Correlation yontem1	1,000	,559	,521	,445	,426	,447	,455	,345	,203	,294
yontem2	,559	1,000	,568	,554	,407	,449	,508	,402	,246	,305
yontem3	,521	,568	1,000	,560	,517	,456	,511	,416	,322	,340
yontem4	,445	,554	,560	1,000	,518	,560	,502	,396	,292	,447
yontem5	,426	,407	,517	,518	1,000	,573	,529	,367	,400	,405
yontem6	,447	,449	,456	,560	,573	1,000	,609	,393	,351	,327
yontem7	,455	,508	,511	,502	,529	,609	1,000	,571	,436	,436
yontem8	,345	,402	,416	,396	,367	,393	,571	1,000	,348	,500
yontem9	,203	,246	,322	,292	,400	,351	,436	,348	1,000	,487
yontem10	,294	,305	,340	,447	,405	,327	,436	,500	,487	1,000

Değişkenler arasındaki korelasyonlar genelde iyi gözükmetedir ($r > 0,20$).

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		,896
Bartlett's Test of Sphericity	Approx. Chi-Square	1489,003
	df	45
	Sig.	,000

KMO örneklem yeterliliği ölçütü 0,896 olup örneklem büyüklüğü açısından faktör analizi için veri yapısının uygun olduğu söylenebilir.

Bartlett testi ile korelasyon matrisinin birim matris olup olmadığı test edilir.

H_0 : Evren korelasyon matrisi birim matristir. $P=0,000 < 0,05$ olduğundan yokluk hipotezi reddedilir. Yani evren korelasyon matrisi birim matris değildir. Ayrıca verilerin çok değişkenli normal dağılımdan geldiği varsayımı sağlanmış olur.

Böylece veri setinin faktör analizi için uygun olduğuna karar verilebilir.

Communalities

	Initial	Extraction
yontem1	1,000	,614
yontem2	1,000	,655
yontem3	1,000	,623
yontem4	1,000	,602
yontem5	1,000	,541
yontem6	1,000	,561
yontem7	1,000	,643
yontem8	1,000	,520
yontem9	1,000	,659
yontem10	1,000	,660

Extraction Method: Principal Component Analysis.

Extraction (Paydaşlık oranı):

Bir değişkenin birden fazla faktöre ait faktör yüklerinin kareleri toplamıdır. Rotated component matrix tablosunda bir değişkenin faktör yüklerinin kareleri toplamı paydaşlık oranını verir.

$$Y2: 0,798^2 + 0,132^2 = 0,655$$

Ortak varyans (Communality) her bir maddenin ortak bir faktördeki varyansı birlikte açıklama oranıdır. Initial (başlangıç) değerleri her bir maddenin başlangıçta varyansı açıklama oranıdır. Extraction (çıkartma değeri) sütununda $Y1=0,614$ şu şekilde yorumlanabilir: $Y1$ varyansın %61,4'ünü açıklamaktadır. Burada maddelerle ilgili olarak verilen ortak varyansların 0,520 ile 0,660 arasında değiştiği görülmektedir.

Eğer Communality değeri 1 den büyük ise ya veri seti çok küçüktür yada araştırmada az veya çok faktör belirlenmiştir.

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	4,980	49,802	49,802	4,980	49,802	49,802	3,577	35,771	35,771
2	1,096	10,956	60,757	1,096	10,956	60,757	2,499	24,986	60,757
3	,735	7,354	68,111						
4	,626	6,258	74,370						
5	,575	5,746	80,116						
6	,502	5,022	85,138						
7	,477	4,773	89,911						
8	,375	3,750	93,661						
9	,333	3,333	96,994						
10	,301	3,006	100,000						

Extraction Method: Principal Component Analysis.

Total variance explained: Açıklanan toplam varyans

Initial Eigenvalues : Başlangıç özdeğerler

Total: Her bir faktörün toplam varyansa olan katkısı açısından toplam özdeğer

% of variance: Her bir faktörün varyansa katkısının yüzdesi

Cumulative: Her bir faktörün varyansa katkısının birikimli yüzdesi

Extraction sums of squared loadings: Faktör sayısı için öneriler

Total: Özdeğeri 1'den büyük olan 2 faktör önerilmiştir.

% of variance: Bu iki faktörün varyansa yaptığı katkılar yüzde olarak verilmiştir.

Cumulative: Her bir faktörün varyansa yaptığı katkının birikimli yüzdesi

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	4,980	49,802	49,802	4,980	49,802	49,802	3,577	35,771	35,771
2	1,096	10,956	60,757	1,096	10,956	60,757	2,499	24,986	60,757
3	,735	7,354	68,111						
4	,626	6,258	74,370						
5	,575	5,746	80,116						
6	,502	5,022	85,138						
7	,477	4,773	89,911						
8	,375	3,750	93,661						
9	,333	3,333	96,994						
10	,301	3,006	100,000						

Extraction Method: Principal Component Analysis.

Rotation sums of squared loadings: Döndürme sonrası faktör sayısı için öneriler

Total: Özdeğeri 1'den büyük olan 2 faktör önerilmiştir.

% of variance: Bu iki faktörün varyansa yaptığı katkılar yüzde olarak verilmiştir.

Cumulative: Her bir faktörün varyansa yaptığı katkının birikimli yüzdesi

Component Matrix^a

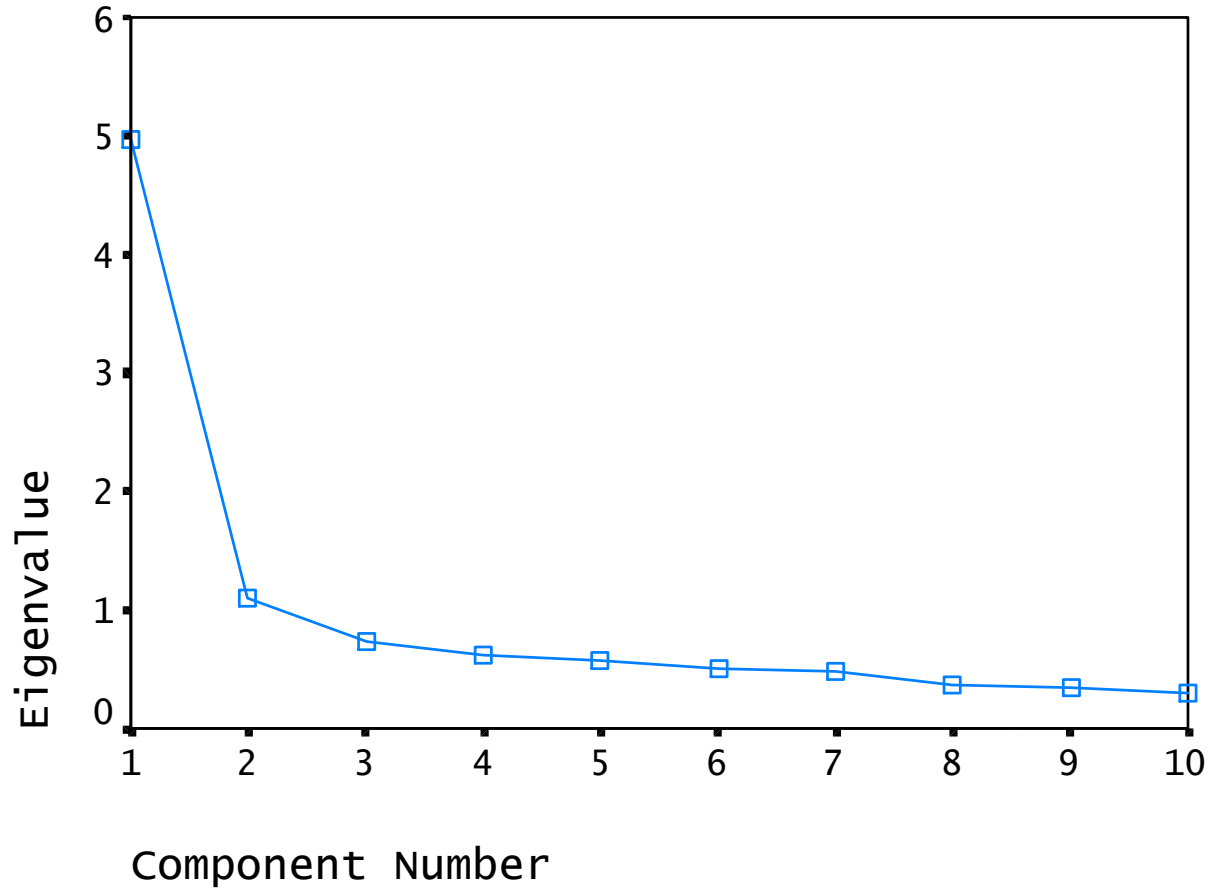
	Component	
	1	2
yontem7	,797	
yontem4	,759	
yontem3	,749	
yontem6	,743	
yomtem5	,736	
yontem2	,717	-,374
yontem1	,670	-,406
yontem8	,668	
yontem10	,629	,514
yontem9	,557	,590

Extraction Method: Principal Component Analysis.

a. 2 components extracted.

Component matrix (Bileşenler matrisi) tablosunda ölçeğe ait maddelerin yük değerinin büyüklüğüne bakılır. Bu yük değerlerinin 0,32 den büyük olması gerekmektedir. Eğer birden fazla maddenin yük değeri 0,32 den küçük ise bunlar sırayla analizden çıkarılırlar. Aynı anda iki madde analizden çıkarılmaz. Tabloda 0,32 den küçük yük değerli madde yoktur.

Scree Plot



Şekilde öz değerlerin nispi değerleri verilmiştir. Burada her faktör ile ilişkili toplam varyans gösterilmiştir. Grafikte 1 ve 2. faktörden sonra belli bir düşme olduğu görülmektedir. Dolayısıyla faktör sayısı 2 olarak tespit edilir. 3 ve diğer faktörlerden sonra önemli bir düşüş eğilimi görülmemektedir. Yani üçüncü ve sonraki faktörlerin varyansa katkıları birbirine yakındır.

Rotated Component Matrix^a

	Component	
	1	2
yontem2	,798	
yontem1	,779	
yontem3	,748	
yontem4	,703	,327
yontem6	,648	,375
yontem7	,584	,549
yontem5	,578	,455
yontem9		,807
yontem10		,789
yontem8	,370	,619

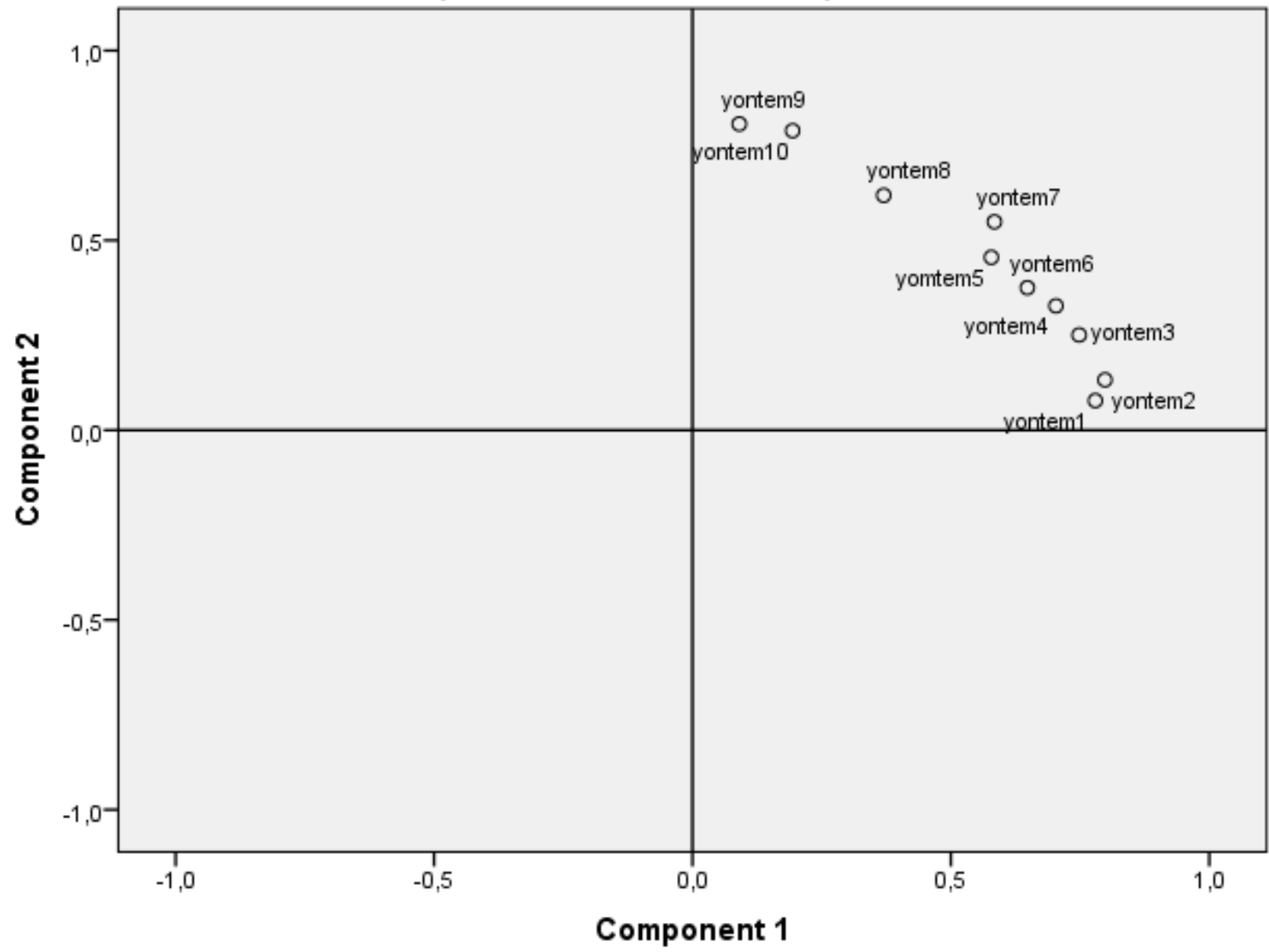
Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.

a. Rotation converged in 3 iterations.

Faktör döndürme sonuçları (rotated component matrix) incelendiğinde maddelerin faktör yük değerlerinin 0,32 den büyük olması gerekir. Ayrıca birden fazla faktörde 0,1'den daha az bir farkla yer alan her hangi bir madde, binişik bir madde olarak değerlendirilir ve bu maddeler ölçekten çıkarılır (Büyüköztürk, 2002).

Tüm maddeler 0,32 den büyüktür. Binişiklik ise Y7 de vardır. Bir maddenin binişik olması için birden fazla faktörde 0,32 den fazla yük değeri olmalı ve maddenin faktörlerdeki yük değeri arasındaki farkın 0,1 den küçük olması gerekir. Y7'nin 2 faktördeki yük değerleri 0,584 ve 0,549 olup 0,32 den büyüktür. Faktör yük değerleri farkı ise $0,549 - 0,549 = 0,039 < 0,1$ olduğundan y7 binişik tir. Bu durum açıklayıcı faktör analizinde arzu edilmeyen bir durumdur. Bu açıdan y7 analizden çıkarılmalıdır.

Component Plot in Rotated Space



y7 çıkarıldıktan sonra yeniden faktör analizi yapılır.

Rotated Component Matrix^a

	Component	
	1	2
y2	,800	
y1	,782	
y3	,753	
y4	,711	
y6	,649	
y5	,585	,456
y9		,812
y10		,806
y8		,601

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.

a. Rotation converged in 3 iterations.

Faktör1 Faktör2

y1

y8

y2

y9

y3

y10

y4

y5

y6

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	4,399	48,879	48,879	4,399	48,879	48,879	3,283	36,478	36,478
2	1,090	12,113	60,992	1,090	12,113	60,992	2,206	24,515	60,992
3	,735	8,171	69,164						
4	,591	6,563	75,727						
5	,533	5,927	81,653						
6	,495	5,501	87,155						
7	,474	5,262	92,417						
8	,375	4,166	96,583						
9	,308	3,417	100,000						

Extraction Method: Principal Component Analysis.

	Faktör1	Faktör2
Y2	0,800²	0,128²
Y1	0,782²	0,000²
Y3	0,753²	0,254²
Y4	0,711²	0,337²
Y6	0,649²	0,355²
Y5	0,585²	0,456²
Y9	0,100²	0,812²
Y10	0,207²	0,806²
Y8	0,374²	0,601²
Toplam	3,283	2,206
Varyans Oranı	(3,283/9)*100 = 36,478	(2,206/9)*100=24,515

Toplam açıklanan varyans oranı =36,478+24,515 =60,992

Çok faktörlü desenlerde açıklanan varyansın %50'den büyük olması gerekir.

	Faktör1	+	Faktör2	=Ortak Faktör Varyansı
Y2	0,800 ²	+	0,128 ²	0,656
Y1	0,782 ²	+	0,000 ²	0,612
Y3	0,753 ²	+	0,254 ²	0,632
Y4	0,711 ²	+	0,337 ²	0,619
Y6	0,649 ²	+	0,355 ²	0,547
Y5	0,585 ²	+	0,456 ²	0,550
Y9	0,100 ²	+	0,812 ²	0,669
Y10	0,207 ²	+	0,806 ²	0,692
Y8	0,374 ²	+	0,601 ²	0,501

Ortak faktör varyansı faktör analizi sonucunda faktörlerin herbir değişken üzerinde yol açtıkları ortak varyanstır. Ortak faktör varyansı bir değişkenin faktör yüklerinin kareleri toplamıdır. Bu değer bir değişkendeki çıkarılmış (extracted) faktörlerce açıklanan varyans oranıdır (Köklü, 2002). Ortak varyans faktörü 0 ile 1 arasında değer alır. 1'e yaklaşması göstergenin varyansa yaptığı katkının yüksek olduğunu, 0'a yaklaşması ise göstergenin varyansa yaptığı katkının düşük olduğunu ifade eder. Örneğin Y2 nin ortak faktör varyansı 0,656 dır yani varyansın %65,6'sının açıklandığı demektir. 0,2 den küçük olan madde heterojenliği bozduğu yani o değişkenin çalışmadan çıkarılması gerektiği anlamına gelir.

3.2.2. DOĐRULAYICI FAKTÖR ANALİZİ-DFA

DFA ölçek geliştirme ve uyarlama sürecinde Açımlayıcı Faktör Analizi ile belirlenmiş olan bir modelin ya da yapının test edilmesi veya doğrulanıp doğrulanmadığının incelenmesine dayanır. DFA önceden belirlenmiş olan yapının sınanmasını veya geliştirilmiş bir ölçme aracının faktör yapısının orijinal formu ile uyarlanmaya çalışıldığı kültürde tutarlılık gösterip göstermediğinin belirlenmesi amaçlanmaktadır.

DFA başka kültürlerde ve örneklerde geliştirilmiş ölçme araçlarının uyarlanmasında kullanılan bir geçerlilik belirleme yöntemidir. Ölçek uyarlama çalışmalarında kullanılan DFA önceden yapı geçerliliği sınanmış olan bir ölçeğin bu yapısını uyarlanmak istenilen dil ve kültürde de koruyup korumadığını test etmenin en iyi yoludur. DFA ölçek uyarlama çalışmalarında başvurulması gereken bir çalışmadır.

DOĐRULAYICI FAKTÖR ANALİZİ-DFA

Modelin uygunluđunun sınanmasında gözlenen deđişkenler arasındaki kovaryans ile modelde önerilen parametreler arasındaki kovaryans matrisi arasındaki farkın, diđer bir deyişle hatanın derecesi temelinde geliştirilmiş olan mutlak uyum indeksleri de kullanılmaktadır.

DFA Lisrel ve AMOS gibi programlarda yapılabilmektedir.

Yapı geçerliliđi için faktör yapılarını belirlenmesinde dođrulayıcı faktör analizi (DFA) kullanılır.

Doğrulayıcı faktör analizi daha önceden belirlenmiş bir yapının doğrulanmasını test etmek amacıyla gerçekleştirilir (Şimşek, 2006). Bu süreçte modelin elde edilen veriyi ne kadar iyi açıkladığı uyum iyiliği indeksleri ile belirlenir. Uyum iyiliği testleri modelin kabul ve reddedilme kararının verilmesini sağlar.

Ki-kare oluşturulan modelin veri tabanına mutlak uygunluğunu değerlendiren önemli bir testtir (Bollen, 1989). Ki-kare testi örneklem büyüklüğüne duyarlıdır ve örneklem sayısı 200'ün üstüne çıktığında genellikle güvenilir sonuçlar vermemektedir (Schumacker ve Lomax, 1996). Bu testte normal ki-kare testinin tersi olarak ki-kare değerinin mümkün olduğunca düşük olması arzulanır. Serbestlik derecesi de ki-kare testinde önemli bir ölçüttür. Serbestlik derecesinin büyük olduğu durumlarda ki-kare anlamlı sonuçlar vermektedir. Bu test ki-kareyi daha az örnek büyüklüğüne bağımlı hale getiren bir yöntem olup ki-karenin serbestlik derecesi bölümünden elde edilir. Bu değer 3'ten küçük olması beklenir (Ayyıldız ve Cengiz, 2006).

Genel Olarak Doğrulamalı Faktör Analizinin Aşamaları



DFA VE AFA ARASINDAKİ TEMEL FARKLAR

AFA ile hızlı bir şekilde maddelerin gireceği alt boyutu ve birden fazla boyuta giren maddeleri görebilir ilgili değişiklikleri yaparak modeli veriye göre oluşturabiliriz. Bu durum AFA'nın en önemli özelliğidir. DFA'nın en önemli özelliği ise bizim kafamızdaki modele verinin uyup uymamasının incelenmesidir (Schumacker and Lomax 2004). Yani AFA 'da uygun modeli kısa sürede oluşturabiliriz fakat bu modelin bilimsel bir açıklaması olmalıdır. Örneğin bir maddenin girdiği alt boyutun teorik olarak da o alt boyuta uygun olması gerekir. AFA ve DFA arasındaki temel farklar aşağıda açıklanmaya çalışılmıştır. ile hızlı bir şekilde maddelerin gireceği alt boyutu ve birden fazla boyuta giren maddeleri görebilir ilgili değişiklikleri yaparak modeli veriye göre oluşturabiliriz. Bu durum AFA'nın en önemli özelliğidir. DFA'nın en önemli özelliği ise bizim kafamızdaki modele verinin uyup uymamasının incelenmesidir (Schumacker and Lomax 2004). Yani AFA 'da uygun modeli kısa sürede oluşturabiliriz fakat bu modelin bilimsel bir açıklaması olmalıdır. Örneğin bir maddenin girdiği alt boyutun teorik olarak da o alt boyuta uygun olması gerekir.

DFA VE AFA ARASINDAKİ TEMEL FARKLAR

1. DFA 'da arařtırmacı ölçekte kaç alt boyut (faktör) olacağını kesin olarak bilmelidir. Hangi maddelerin (gözlenen deęişken) hangi alt boyutta olduğuna yine arařtırmacı kurduğu modelle karar verir. AFA 'da ise arařtırmacı maddelerin hangi alt boyuta girdiğini ve alt boyut sayısını sadece gözlemler (Schumacker ve Lomax 2010). Arařtırmacı eđer isterse alt boyut sayısını sınırlandırabilir.
2. DFA'da, kesin olarak ölçülmek istenen şey için teorik alt yapı gerekir. AFA 'da ölçüğün yapısı için her ne kadar teorik alt yapı gerekse de maddeleri ve alt boyutları AFA belirler.
3. DFA'da birden çok uyum indeksi ve faktör yüklerinin bileşimi modelin uygunluęunu belirler. AFA 'da genelde sadece faktör yüklerine bakarak karar verilir.

ANAHTAR SÖZCÜKLER

Gözlenen Değişken (Observed variable): Bireye ait bir özellik hakkında etkisi incelenen değişkendir. Madde ya da gösterge olarak da adlandırılır.

Gizil Değişken (Latent variable): Birden fazla gözlenen değişkeni etkileyen ve bu gözlenen değişkenler arasındaki ilişkiyi açıklamaya çalışan gözlenemeyen bir değişkendir. Gözlenen değişkenler ortak bir nedeni paylaşmalarından dolayı bir araya gelirler ve bu kümenin geneli gizil değişken olarak adlandırılır (Brown, 2006).

Dışsal Değişken (Exogenous variable): Bağımsız değişken olarak da adlandırılır. Yol şemasında temel başlangıç ya da yordayıcı değişkenlerdir ve modelin şemasal görünümünde sol tarafta yer alırlar.

İçsel Değişken (Endogeneous variable): Bağımlı değişken olarak adlandırılır. Yol şemasında yordanan (aracı değişken-ler) değişkenlerdir. Modelin şemasal görünümünde dışsal değişkenlerin sağında yer alır.

ANAHTAR SÖZCÜKLER

Doğrudan etki (Direct effect): Bir değişkenin bir başka değişkeni tek yönlü bir yolla etkilediği varsayımdır (Kline, 2011).

Dolaylı etki (Indirect effect): Dışsal bir değişkenin içsel bir değişkene olan etkisinin bir ya da daha fazla aracı değişkenle ortaya konulmasıdır.

Karşılıklı etki (Reciprocal effect): İki değişken arasındaki iki yönlü etkidir.

Ölçme modeli (Measurement model): Bir gizil değişken ve göstergelerinden oluşan yapıya denir. Bir gizil değişkenli doğrulayıcı faktör analizi modeli ile aynı anlamda kullanılır.

Yapısal model (Structural model): Gizil değişkenleri birbirine eşzamanlı eşitlik sistemleri bağlayarak oluşturulan modele denir.

ANAHTAR SÖZCÜKLER

Yol şeması-diyagramı (Path Diagram): Faktörlerle ilişkili olduğu varsayılan değişkenleri göstermek amacıyla kullanılabilir. Yol şeması bu nedenle faktörler arasındaki ilişkilerin kurulmasını sağlar ve gözlenen değişkenlerin hangi faktörler altında tanımlanacağını gösterir (Schumacker ve Lomax, 1996).

Hata varyansı (Error variance): Veri setine ilişkin varyansın açıklanamayan kısmını gösterir.

Modifikasyon İndeksi (Modification Index): Gösterge ve gizil değişkenler arasındaki kovaryansa temelinde, modele ilişkin ayrıntılı modifikasyonlar önerir. Bu modifikasyonlar genellikle hata matrisleri temelinde oluşturulur ve modelde orijinal olarak öngörülmemeyen, ancak eklenmesi ya da çıkarılması durumunda modelde kazamılacak ki-kare değerini gösterir.

LISREL

LISREL yazılımı Jöreskog ve Sörbom tarafından geliştirilmiş bir programdır. LISREL yazılımının ismi Linear Structural Models 'in kısaltımıdır. LISREL'in içinde iki program/modül vardır. Bunlardan birisi PRELIS, diğeri LISREL'dir. PRELIS verinin etkili bir şekilde gözden geçirilmesine ve tanımlayıcı bilgilerin incelenmesine olanak verir. Veriyi LISREL'de analize uygun hale getirmek için gerekli olan tüm işlevleri yapar. LISREL ise, PRELIS tarafından oluşturulan veriyi test etmekte kullanılır (Jöreskog and Sörbom 1999).

LISREL'de veri çözümlerken, hiçbir aşamada Türkçe karakter kullanılmamalıdır. Analiz edilen veri tabanını bir klasöre kopyalanıp orada başlanması gerekir. İleride analiz tekrar yapıp sonuçların kontrol edilmesi istenirse, başka klasöre aktarılan LISREL 'e ait dosyalar program tarafından okunmamaktadır. Bu nedenle analiz işlemi tamamen bitinceye kadar dosyalar kopyalanan klasörde kalmalıdır

LISREL

LISREL 'de analize başlarken ilk adım SPSS veya EXEL 'de oluşturulmuş veri tabanını LISREL 'e aktarmaktır. "Import" komutu ile dosya LISREL 'e aktarılır. Dosya import ederken SPSS ile önceden oluşturulmuş veri setinden ölçek maddeleri hariç tüm değişkenler silinmelidir. LISREL sütun hesabına göre çalıştığından yaş, ekonomik durum gibi ölçek maddeleri haricindeki (demografik değişkenler gibi) tüm sütunların çıkarılması gerekmektedir (Çokluk, Şekercioğlu ve Büyüköztürk 2010). Bu aşamadan sonra verinin süreklimi, kategorik mi olduğunun tanımlanması gerekir. Bunun için veri (data) penceresinden "Define Variable" seçilir ve veri türü seçilerek tüm değişkenlere uygula "Apply All Variable" tıklanarak onaylanır. Daha sonra kaydet butonu ile kaydedilerek değişiklikler aktif hale getirilir. Sonraki aşama matrisin oluşturulmasıdır. "Statistics" menüsünden "normal scores" tıklanır ve açılan pencereye ölçek maddeleri atılır, daha sonra "output options"tan matrisin tanımlanması işlemi gerçekleşir. Eğer covaryans matrisi kullanılacaksa dosya adına .COV uzantısı verilir. Korelasyon matrisinde ise .COR uzantısı verilir. Pencere onaylanarak matrisin oluşturulması sağlanır (Çokluk, Şekercioğlu ve Büyüköztürk 2010). Hangi matrisin seçilmesine karar verirken bu makalede "DFA 'da Kullanılan Matris Türleri" başlığında yer alan bilgiler dikkate alınır.

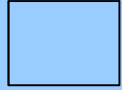
LISREL

Daha sonraki aşama söz dizim (syntax) dosyasının oluşturulmasıdır. SPSS den farklı olarak LISREL’de söz dizim (syntax) dosyası oluşturulur (Çokluk, Şekercioğlu ve Büyüköztürk 2010). New menüsünden “syntax only” sekmesi tıklanır. Açılan sayfaya syntax yazılır. PATH diyagramı çizildikten sonra syntax’ın otomatik yazdırılması yapılabilir. LISREL bu söz dizimini otomatik oluşturur ve analizini oradaki komutlara göre yapar, bazen araştırmacının bu söz dizimi incelemesi ve doğruluğunu teyit etmesi gerekir. Söz dizimi bittikten sonra dosya çalıştırılır ve sonuçlar incelenerek model hakkındaki yargıya varılır. PATH diyagramında görsel olarak maddelerin t değerleri, faktör yükleri, χ^2 , sd, RMSEA gibi gerekli bilgiler okunabilir. Ayrıca output dosyasından daha detaylı bilgiler olan tüm uyum indeksi sonuçlarına ulaşılabilir.

PATH Diyagramı

Yapısal eşitlik modelinde analizler soncunda yol şemaları (path diagrams) elde edilebilmektedir. Uygun matris oluşturulduktan sonra uyum indeksleri ve analizi yapan yazılımın çıktı sayfası haricinde bir PATH diyagramı çizdirilerek modele ait değişkenler, t değerleri, faktör yükleri, açıklanamayan varyans ve bazı uyum iyiliği değerleri bu diyagramda özet olarak görülebilir. Bu şemalar kısaca modele ait çıktıları grafiksel olarak sunar (Gatignon 2011). PATH diyagramında görülmeyen detaylı analiz sonuçları Lisrel programına ait çıktı sayfasında görülebilmektedir.

PATH Diyagramı İçin Şekiller ve Semboller



Gözlenen deęişken, arařtırmacı tarafından ölçölen deęişken



Gizil (latent) deęişken, gözlenemeyen veya varsayımsal yapı



Doęrudan etki (Direct effect), $X \rightarrow Y$ X'in Y'yi tek yönlü bir yolla etkiledięi varsayılır.

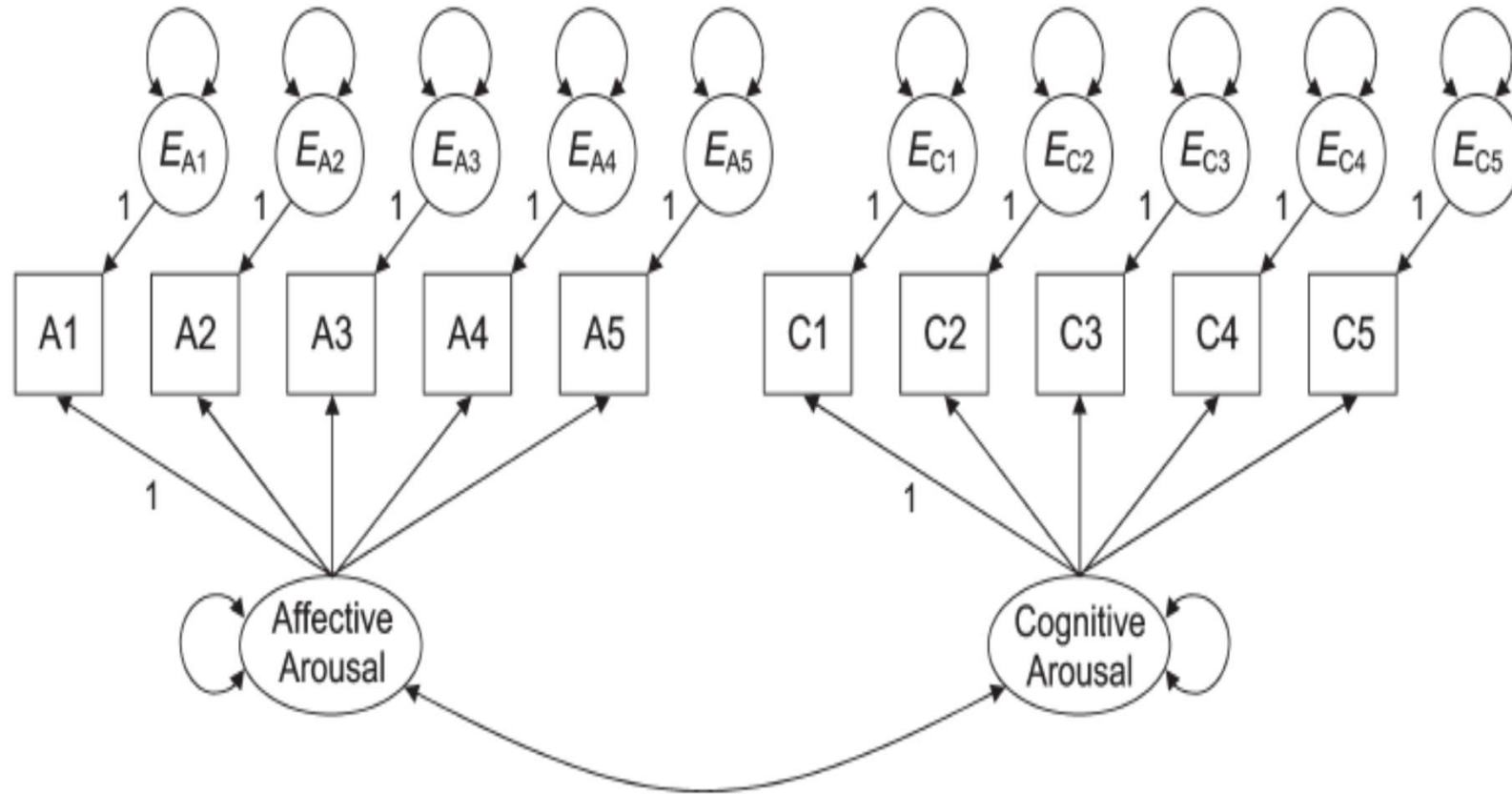


Karşılıklı etki (Reciproral effect),

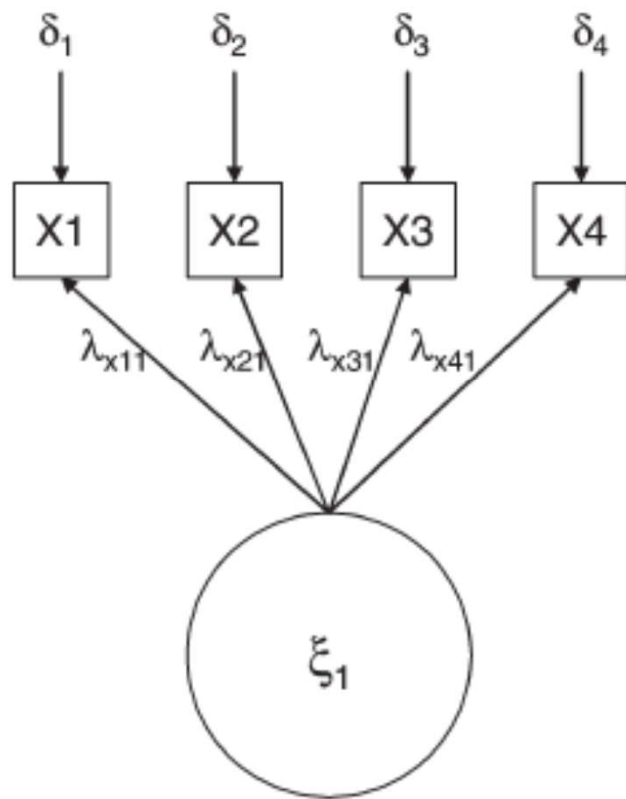


Korelasyon ya da kovaryans, iki deęişken arasında ortak deęişkenlięin olduęu varsayılır.

DOĞRULAYICI FAKTÖR ANALİZİ



Model A: Over-Identified ($df = 2$)

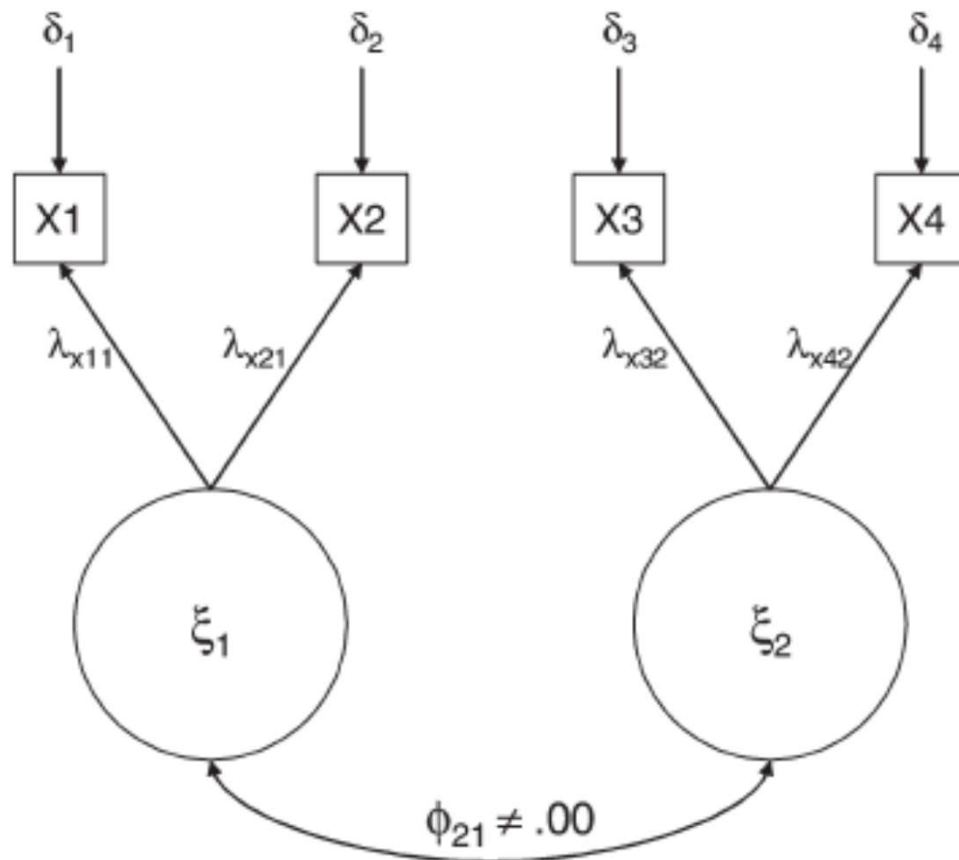


Input Matrix (10 elements)

	X1	X2	X3	X4
X1	σ_{11}			
X2	σ_{21}	σ_{22}		
X3	σ_{31}	σ_{32}	σ_{33}	
X4	σ_{41}	σ_{42}	σ_{43}	σ_{44}

Freely Estimated Model Parameters = 8
(e.g., 4 factor loadings, 4 error variances)

Model B: Over-Identified ($df = 1$)

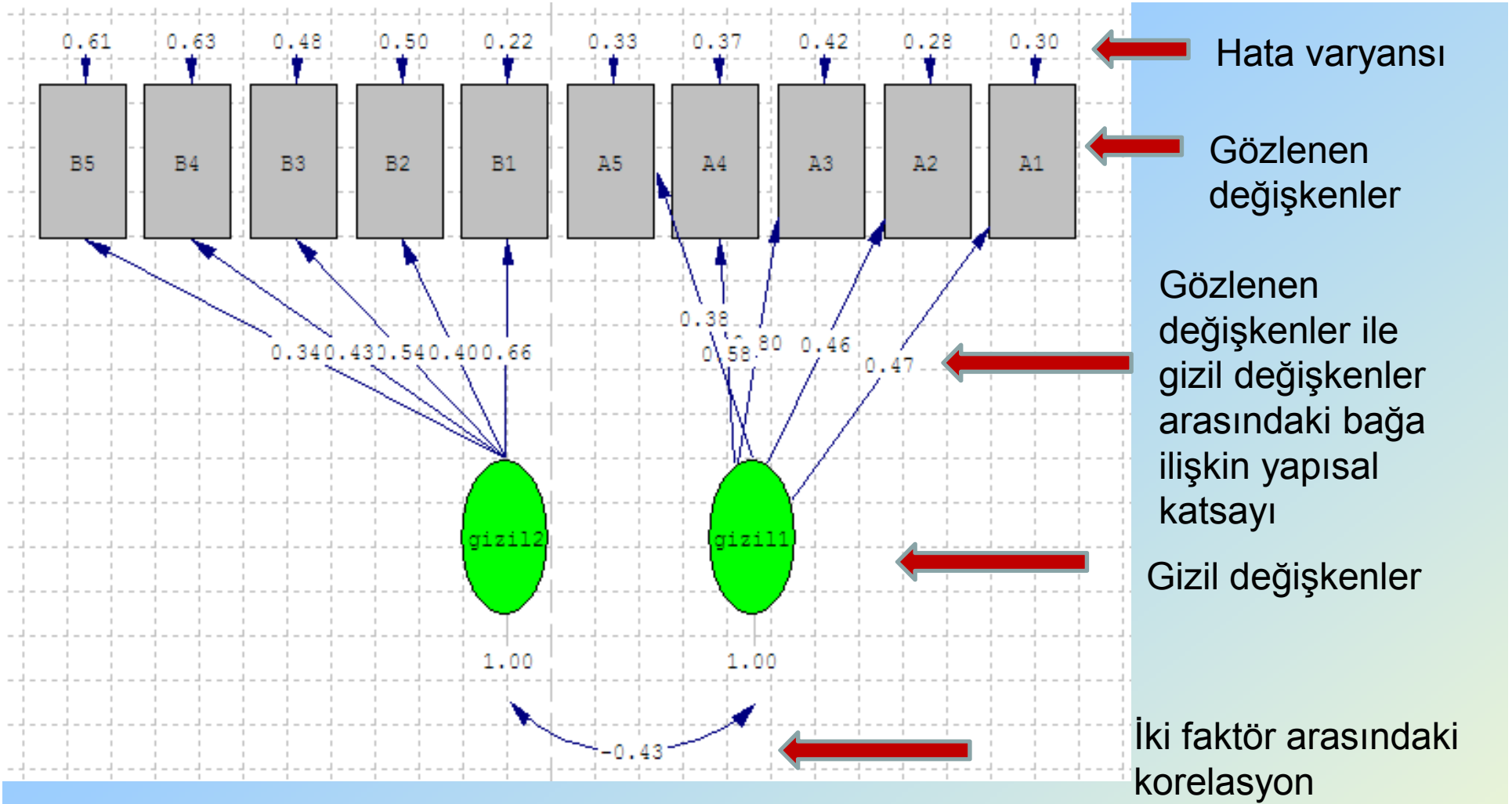


Input Matrix (10 elements)

	X1	X2	X3	X4
X1	σ_{11}			
X2	σ_{21}	σ_{22}		
X3	σ_{31}	σ_{32}	σ_{33}	
X4	σ_{41}	σ_{42}	σ_{43}	σ_{44}

Freely Estimated Model Parameters = 9 (e.g., 4 factor loadings, 4 error variances, 1 factor covariance)

(con.



Uygun bir ölçme modelinde faktör yüklerinin yüksek, hata varyanslarının düşük, faktör korelasyonlarının 0,85'den küçük olması beklenir. Faktör korelasyonlarının 0,85'i geçmesi ise aslında daha az faktörle model veri uyumunun sağlanabileceğini ve varlığı iddia edilen faktörlerin birbirinden ayrı kavramlar olmadıkları düşünülür.

LISREL

PATH diyagramı çizdirildikten sonra ilk iş olarak maddelerin t değerleri kontrol edilir. Tablo t değeri 1.96'yı aşarsa 0.05, 2.56'yı aşarsa 0.01 düzeyinde anlamlıdır. Anlamsız olanların ölçekten çıkarılması gerekir (Şimşek 2007; Schumacker and Lomax 2010). Elimizdeki modelde tüm maddelere ait t değerlerinin anlamlı olması modelin kabul edilebilir olması için gerekli bir koşuldur. Maddelerin hata varyansları da t değerinin yanı sıra incelenmelidir. Hata varyansı çok yüksek olan maddelerin açıklayıcılıkları da düşük çıkacaktır. Eğer birden fazla maddenin t değeri ya da hata varyansı olumsuzsa maddeler teker teker atılıp sonuçlara bakılarak çıkarılmalıdır (Şimşek 2007). Uyum indekslerinin de t değeri ve hata varyansı gibi uygun olması gerekir. Çıktılarda olumsuz sonuçlar var ise, modifikasyonlar değerlendirilir. Birden çok modifikasyon uygulanacak ise teker teker yapılmalıdır (Çokluk, Şekercioğlu ve Büyüköztürk 2010). Harrington'un (2009) aktardığına göre, faktör yüklerinin 0.30'un altında olmaması istenir. 0.71 ve üzeri mükemmel, 0.63 çok iyi, 0.55 iyi, 0.45 güzel/kabul edilebilir ve 0.32 zayıftır. Bu işlemlerden sonra en son sonuç sayfası ve PATH diyagramı incelenerek modelin uyumuna karar verilir. Bu karar verme işleminde makalede anlatılan uyum iyiliği sonuçları, faktör yükleri, t değeri gibi faktörler dikkate alınır (Çapık, 2014).

χ^2 / sd Deęeri:

Doęrulayıcı faktör analizi sonuçlarına göre öncelikle p deęerinin incelenmesi gerekmektedir. Bu deęer, beklenen kovaryans matrisi ile gözlenen kovaryans matrisi arasındaki farkın manidarlığı hakkında bilgi vermektedir. Örneklemin büyük olması sebebiyle p deęeri manidar çıkmaktadır. Bu yüzden genellikle p deęerinin manidar olması pek çok çalışmada tolere edilmektedir.

Ki-kare istatistięi örneklem büyüklüğünden çok çabuk etkilendięi için örneklemden daha az etkilenen χ^2 / sd oranı bunun yerine kullanılabilen bir ölçüttür (Şimşek 2007; Waltz, Strcikland and Lenz 2010). χ^2 deęerinin serbestlik derecesine bölünmesiyle elde edilen bu deęer iki veya altında olmalıdır. Beş ve daha az ise kabul edilebilir bir deęerdir (Munro 2005; Şimşek 2007; Hooper and Mullen 2008).

$0 < \chi^2 / sd \leq 2$ mükemmel uyum vardır.

$2 < \chi^2 / sd \leq 3$ kabul edilebilir düzeyde bir uyum vardır (Kline, 2005).

$3 < \chi^2 / sd < 5$ Orta düzeyde bir uyum vardır (Sümer, 2000).

RMSEA (Root Mean Square Error of Approximation): Tahminin Kök Hata Kareler Ortalaması

Ana kütledeki yaklaşık uyumun bir ölçüsüdür. Yaklaşık ortalamaların karekökü anlamına gelir. Sıfır ve bir arasında değer alır (Munro 2005; Yılmaz ve Çelik 2009; Çokluk, Şekercioğlu ve Büyükoztürk 2010; Schumacker and Lomax 2010).

Modelin anlamlılığı:

$0 < \text{RMSEA} < 0.05$ Normal değer

$0.05 < \text{RMSEA} < 0.08$ Kabul edilebilir uyum

GFI (Goodness of Fit Index): Uyum iyiliđi indeksi

Modelin örneklemdaki kovaryans matrisini ne oranda ölçtüđünü gösterir (Çokluk, Şekerciođlu ve Büyüköztürk 2010; Waltz, Strcikland and Lenz 2010).

GFI, varsayılan modelce hesaplanan gözlenen deđişkenler arasındaki genel kovaryans miktarını gösterir. Regresyon analizindeki R^2 gibi açıklanabilir. Örnek hacminin çok olması GFI deđerini yükselterek dođru sonuç alınmasını önleyebilir. GFI deđeri 0 ile 1 arasında deđişir. GFI'nın 0.90'ı aşması iyi bir model göstergesi olarak alınmaktadır. Bu gözlenen deđişkenler arasında yeterince kovaryansın hesaplandıđı anlamına gelmektedir (Munro 2005; Waltz, Strcikland and Lenz 2010). GFI, iki modelin görelî uyum eksikliđini karşılaştırmaktan çok, toplam varyansa göre açıklanan kovaryansla ilgilenmektedir.

$0.95 \leq \text{GFI} < 1 = .00$ Normal deđer

$0.90 \leq \text{GFI} < 0.95$ Kabul edilebilir deđer

AGFI (Adjusted Goodness of Fit Index):

GFI testinin yüksek örnek hacmindeki eksikliğini gidermek amacıyla kullanılan bir iyi uyum indekstir. Gözlenen değişken sayısına göre modelin serbestlik derecesi için GFI değerini düzeltmektedir. Örneklem sayısının özellikle büyük olduğu durumlarda AGFI daha temsili bir uyum indeksidir. Değeri 0-1 arasında değişir (Munro 2005; Çokluk, Şekercioğlu ve Büyüköztürk 2010).

$0.90 \leq \text{GFI} < 1$

Normal değer

$0.85 \leq \text{GFI} < 0.90$

Kabul edilebilir değer

RMR (Root Mean Square Residual)

SRMR (Standardized Root Mean Square Residual):

SRMR; standartlaştırılmış ortalama hataların kareköküdür. Bu değer sıfıra yaklaştıkça test edilen modelin daha iyi uyum iyiliği gösterdiği anlaşılır. Standardize edilmiş şekline SRMR uyum iyilik indeksi denir (Çokluk, Şekercioğlu ve Büyüköztürk 2010; Wang and Wang 2012).

$0 \leq \text{RMR}, \text{SRMR} < 0.05$

Normal değer

$0.05 < \text{RMR}, \text{SRMR} < 0.10$

Kabul edilebilir değer

CFI (Comparative Fit Index):

Değişkenler arasında hiçbir ilişkinin olmadığını varsayarak kurulan modelin yokluk modelinden (null) farkını verir. Değişkenler arasında ilişkinin olmadığını öngören modeldir. Mevcut modelin uyumu ile gizil değişkenler arası korelasyonu ve kovaryansı yok sayan sıfır hipotez modelinin uyumunu karşılaştırır. Yani model tarafından tahmin edilen kovaryans matrisi ile sıfır hipotezli modelin kovaryans matrisini karşılaştırır. Değeri 0 - 1 arasında değişir (Munro 2005; Çokluk, Şekercioğlu ve Büyüköztürk 2010).

$0.97 \leq \text{CFI} < 1$

Normal değer

$0.90 \leq \text{CFI} < 0.97$

Kabul edilebilir değer

NFI (Normed Fit Index):

NFI; normlaştırılmış uyum indeksi olup, CFI'a alternatif olarak geliştirilmiştir. Örneklem sayısı ile pozitif ilişkilidir. Bu indeks varsayılan modelin temel ya da sıfır hipoteziyle olan uygunluğunu araştırır. 0-1 arasında değişen değerler alır. NFI iç içe model karşılaştırmasına katsı sağlar (Bentler, 1990).

$0.95 \leq \text{NFI} < 1$

Normal değer

$0.90 \leq \text{NFI} < 0.95$

Kabul edilebilir değer

NNFI:

NNFI ya da normlaştırılmamış uyum indeksi; örnek sayısının artmasından etkilenmemektedir. Her ne kadar normalite varsayımından hareket ediyor olsa da NNFI genel olarak 0–1 aralığında olmakla birlikte, bazen bu aralığın dışına çıkabilir (Şehribanoğlu, 2005). Uyum indekslerinin 0.90 değerinden büyük ve hata indekslerinin ise, 0.05 değerinden küçük olmasını önermektedir.

$0.97 \leq \text{NNFI} < 1$	Normal değer
$0.95 \leq \text{NFI} < 0.97$	Kabul edilebilir değer

Modifikasyonlar: Model kurulup test edildikten sonra LISREL arařtırmacıya bazı düzeltmeler önerebilir (Schreiber, Nora, Stage, Barlow and King 2006). Bu düzeltmeler arařtırmacının kurduđu modeli iyileřtirme amacıyla yapılır. Düzeltmeler uyum deđerlerini yakalamaya yeterli deđerse, model teorik yapıya uygun olarak başka bir řekilde yeniden kurulmalıdır (Schreiber, Nora, Stage, Barlow and King 2006).

Modifikasyonlar uygulanacaksa yapılacak deđerikliklerin teorik olarak da mantıklı olması gerekmektedir (Diamantopoulos and Siguaw 2000). Örneđin program bir gözlenen deđeriklenin modelde önerilen örtük deđerikenden başka olan bir örtük deđeriklenin altında olmasını önerirse, bu deđerikliđi yaptığımızda deđeriklenin yeni yerinin teorik olarak da mantıklı olması gerekir. Başka bir anlatımla, program bir maddeyi olduđu alt boyuttan başka bir alt boyuta taşımanın faydalı olacağını önerirse, taşınacak bu maddenin yeni alt boyuta uygun olması gerekir.

LISREL ile elde edilen uyum iyiliği indeksleri

Uyum Ölçüsü	İyi Uyum	Kabul Edilebilir Uyum
χ^2/sd	$0 \leq \chi^2/sd \leq 2$	$2 \leq \chi^2/sd \leq 3$
RMSEA	$0 \leq RMSEA \leq 0,05$	$0,05 \leq RMSEA \leq 0,08$
p değeri (RMSEA<0,05)	$0,10 \leq p \leq 1,00$	$0,05 \leq p \leq 0,10$
CFI	$0,97 \leq CFI \leq 1,00$	$0,95 \leq CFI \leq 0,97$
GFI	$0,95 \leq GFI \leq 1,00$	$0,90 \leq GFI \leq 0,95$
AGFI	$0,90 \leq AGFI \leq 1,00$	$0,85 \leq AGFI \leq 0,90$
AIC	Karşılaştırılan model için AIC'ten daha küçük	
CAIC	Karşılaştırılan model için CAIC'ten daha küçük	

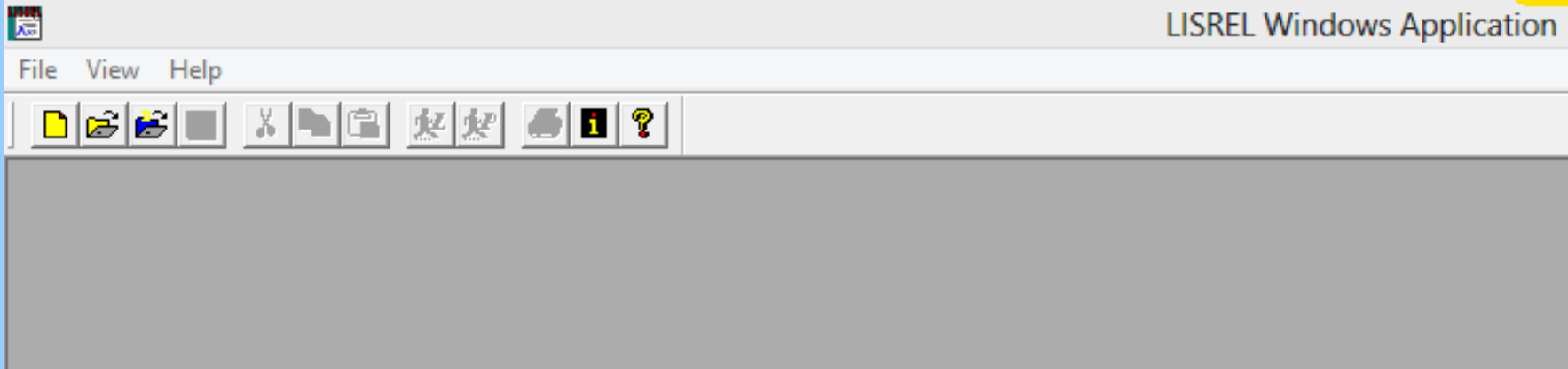
Standart Uyum İyiliği Ölçütleri ile Araştırma Sonuçlarının Karşılaştırılması

Uyum Ölçüleri	İyi Uyum	Kabul Edilebilir Uyum
χ^2	$0 \leq \chi^2 \leq 2df$	$2df \leq \chi^2 \leq 3df$
P değeri	$0.05 \leq p \leq 1$	$0.01 \leq p \leq 0.05$
χ^2/df	$0 \leq \chi^2/df \leq 2$	$2 \leq \chi^2/df \leq 3$
RMSEA	$0 \leq RMSEA \leq 0.05$	$0.05 \leq RMSEA \leq 0.08$
RMR		
SRMR	$0 \leq SRMR \leq 0.05$	$0.05 \leq SRMR \leq 0.10$
NFI	$0.95 \leq NFI \leq 1.00$	$0.90 \leq NFI \leq 0.95$
NNFI	$0.97 \leq NNFI \leq 1.00$	$0.95 \leq NNFI \leq 0.97$
CFI	$0.97 \leq CFI \leq 1.00$	$0.95 \leq CFI \leq 0.97$
GFI	$0.95 \leq GFI \leq 1.00$	$0.90 \leq GFI \leq 0.95$
AGFI	$0.90 \leq AGFI \leq 1.00$	$0.85 \leq AGFI \leq 0.90$
RFI	$0.90 < RFI < 1.00$	$0.85 < RFI < 0.90$

Kaynak: Schermelleh-Engel-Moosbrugger (2003)

Lisrel Programında DFA

i) Programın alıřtırılması ve verilerin yklenmesi:



New: Lisrel'de yeni bir dosya oluřturmak.



Open: Lisrel'de nceden oluřturulmuř bir dosyayı amak



Import: Bařka programlarda (Excel, SPSS gibi) hazırlanan verileri Lisrel programına aktarmak.



Open Data File

File View Help

New Ctrl+N

Open... Ctrl+O

Import Data in Free Format

Import External Data in Other Formats

Print Setup...

1 C:\Users\...\Desktop\YT\yt

2 C:\Users\...\Desktop\YT\YU

3 C:\Users\...\Desktop\YT\YUSUF

4 C:\Users\...\Desktop\YT\PT

Exit

Konum: Masaüstü



Bilgisayar
Sistem Klasörü



Ağ
Sistem Klasörü



YT
Dosya klasörü

Dosya adı:

Aç

Dosya türü:

Free Format Data (*.dat)

Free Format Data (*.dat)

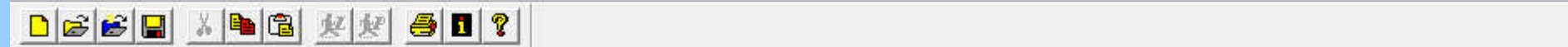
Comma Delimited Data (*.csv)

Tab Delimited Data (*.txt)

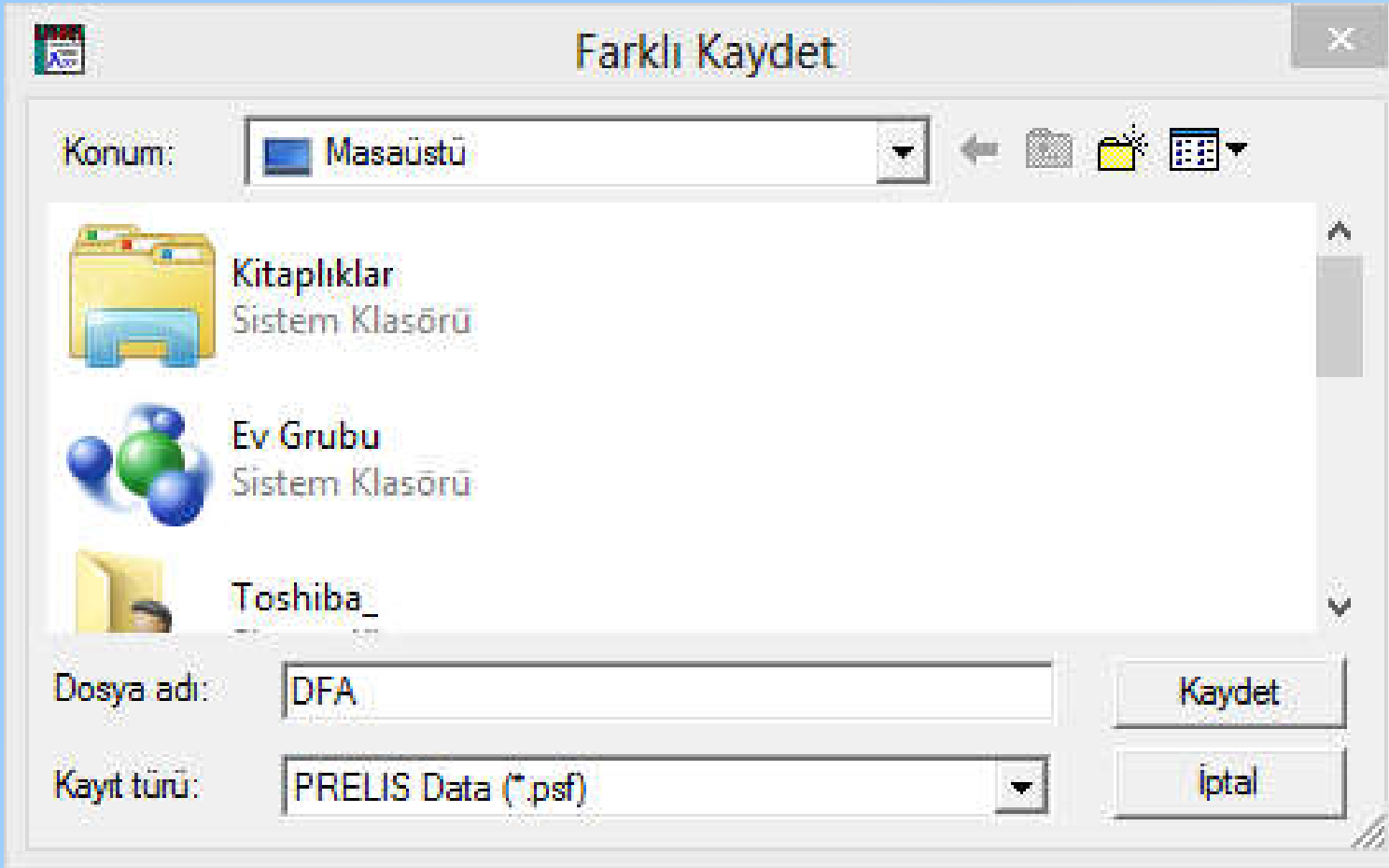
SPSS for Windows (*.sav)

All Files (*.*)

İptal



	Y1	Y2	Y3	Y4	Y5	Y6	Y8	Y9	Y10
1	1,000	1,000	2,000	2,000	2,000	1,000	1,000	1,000	2,000
2	1,000	2,000	2,000	2,000	2,000	1,000	1,000	2,000	3,000
3	1,000	2,000	2,000	2,000	1,000	1,000	1,000	3,000	1,000
4	1,000	2,000	1,000	1,000	1,000	3,000	1,000	1,000	3,000
5	1,000	1,000	3,000	2,000	2,000	2,000	1,000	2,000	1,000
6	1,000	2,000	3,000	1,000	2,000	2,000	1,000	1,000	3,000
7	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	3,000
8	1,000	3,000	4,000	5,000	1,000	1,000	1,000	1,000	1,000
9	1,000	2,000	2,000	3,000	1,000	1,000	1,000	1,000	1,000
10	1,000	1,000	1,000	2,000	2,000	1,000	1,000	1,000	1,000
11	1,000	2,000	1,000	2,000	2,000	2,000	1,000	1,000	2,000
12	2,000	2,000	3,000	3,000	3,000	2,000	1,000	1,000	2,000
13	2,000	2,000	2,000	2,000	3,000	3,000	1,000	2,000	2,000
14	2,000	2,000	2,000	1,000	2,000	1,000	1,000	3,000	2,000
15	2,000	1,000	3,000	2,000	3,000	2,000	1,000	2,000	2,000
16	2,000	2,000	1,000	2,000	2,000	2,000	1,000	1,000	2,000
17	2,000	1,000	3,000	1,000	2,000	1,000	1,000	3,000	2,000
18	2,000	4,000	3,000	5,000	3,000	2,000	1,000	3,000	2,000
19	2,000	2,000	3,000	2,000	3,000	3,000	1,000	2,000	2,000
20	2,000	2,000	2,000	2,000	2,000	2,000	1,000	1,000	2,000
21	2,000	1,000	1,000	2,000	2,000	1,000	1,000	1,000	1,000
22	2,000	3,000	3,000	2,000	3,000	3,000	1,000	1,000	2,000
23	2,000	2,000	2,000	1,000	2,000	3,000	1,000	2,000	2,000
24	2,000	4,000	3,000	3,000	1,000	2,000	1,000	4,000	4,000
25	2,000	2,000	2,000	2,000	1,000	1,000	1,000	1,000	2,000



The image displays three overlapping windows from the SPSS software interface:

- Define Variables Dialog:** Shows a list of variables Y1 through Y10. The 'Insert' button is highlighted. Below the list, a note reads: "To select more than one variable at a time, hold down the CTRL key while clicking on the variables to be selected".
- Variable Types for Y1 ... Dialog:** Shows the 'Continuous' radio button selected. The 'Apply to all' checkbox is checked. Below the dialog is a data table:

3,000	2,000	1,000
3,000	3,000	1,000
2,000	1,000	1,000
3,000	2,000	1,000
2,000	2,000	1,000
2,000	1,000	1,000
3,000	2,000	1,000
3,000	3,000	1,000
2,000	2,000	1,000
2,000	1,000	1,000
- Yeni (New) Dialog:** Shows a list of file types: PRELIS Data, SIMPLIS Project, LISREL Project, and Path Diagram. 'Path Diagram' is selected. Buttons for 'Tamam' (OK) and 'iptal' (Cancel) are visible.

- Data Screening
- Impute Missing Values ...
- Multiple Imputation...
- Equal Thresholds...
- Fix Thresholds...
- Homogeneity Test ...
- Normal Scores...**
- Factor Analysis...
- Censored Regressions...
- Logistic Regressions...
- Probit Regressions...
- Regressions...
- Two-Stage Least-Squares...
- Bootstrapping ...
- Output Options ...

Normal Scores

Variable List:

- Y1
- Y2
- Y3
- Y4
- Y5
- Y6
- Y8
- Y9
- Y10

Cancel

Run

Output Options

Syntax

Add

Remove

- Normal Scores for Y1
- Normal Scores for Y2
- Normal Scores for Y3
- Normal Scores for Y4
- Normal Scores for Y5
- Normal Scores for Y6
- Normal Scores for Y8
- Normal Scores for Y9
- Normal Scores for Y10

To select more than one variable at a time, hold down the CTRL key while clicking on the variables to be selected

Output

Moment Matrix

Covariances

Save to file: LISREL system data

DFA.COV

Means

Save to file:

Standard Deviations

Save to file:

Asymptotic Covariance Matrix

Save to file: Print in output

Asymptotic Variances

Save to file: Print in output

Data

Save the transformed data to file:

Width of fields: 15

Number of decimals: 6

Number of repetitions: 1

Rewind data after each repetition

Print bivariate frequency tables

Print tests of underlying bivariate normality

Perform tests of multivariate normality

Wide print

Random seed

Set seed to 123456

OK

Cancel

Normal Scores

Variable List:

Cancel

Run

Y1
Y2
Y3
Y4
Y5
Y6
Y8
Y9
Y10

Output Options

Syntax

Add

Remove

Normal Scores for Y1
Normal Scores for Y2
Normal Scores for Y3
Normal Scores for Y4
Normal Scores for Y5
Normal Scores for Y6
Normal Scores for Y8

To select more than one variable at a time, hold down the CTRL key while clicking on the variables to be selected



DATE: 03/01/2016
TIME: 11:04

P R E L I S 2.71

BY

Karl G. Jöreskog & Dag Sörbom

This program is published exclusively by
Scientific Software International, Inc.
7383 N. Lincoln Avenue, Suite 100
Lincolnwood, IL 60712, U.S.A.

Phone: (800)247-6113, (847)675-0720, Fax: (847)675-2140

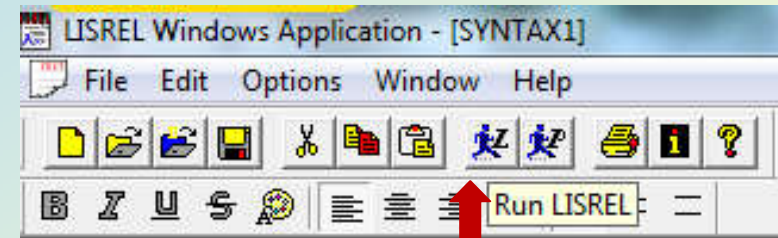
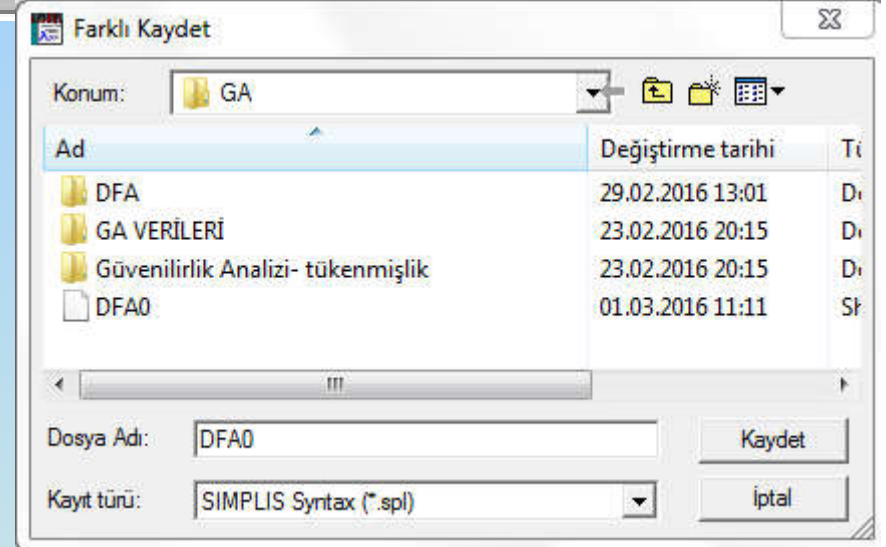
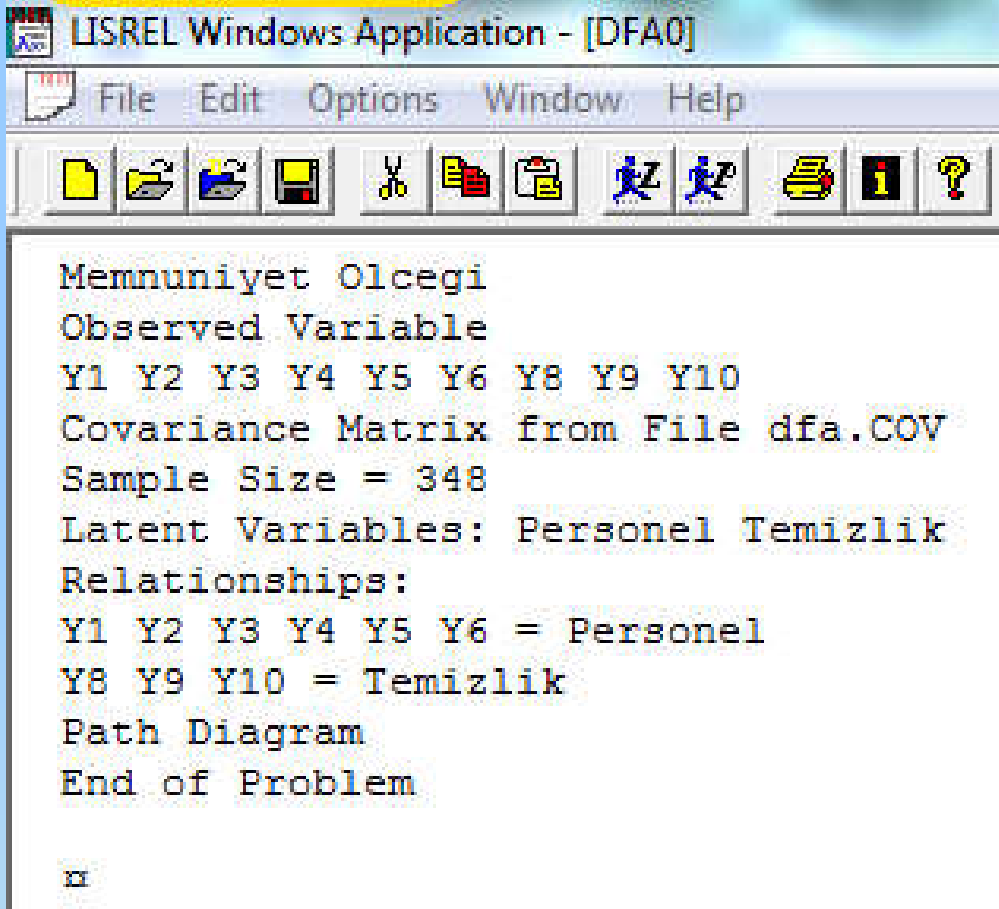
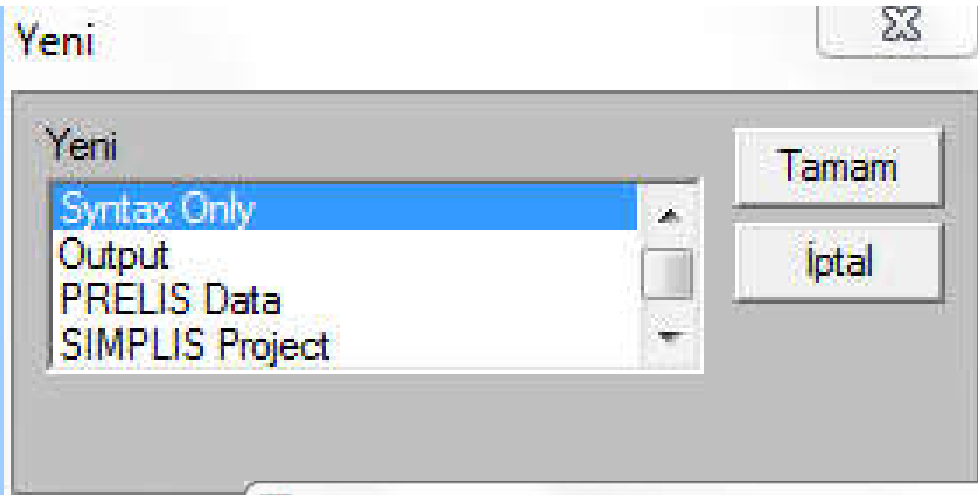
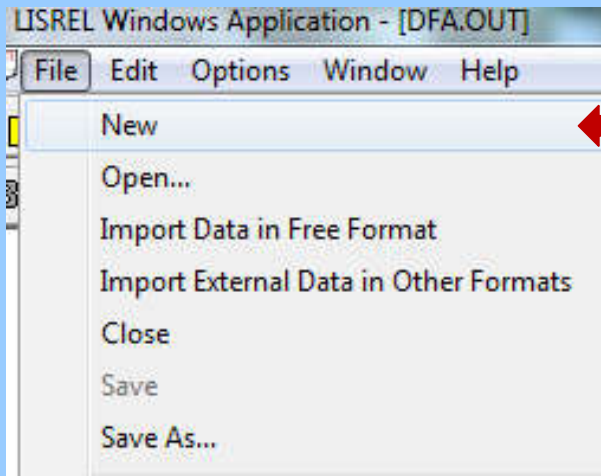
Copyright by Scientific Software International, Inc., 1981-2004

Use of this program is subject to the terms specified in the
Universal Copyright Convention.

Website: www.ssicentral.com

The following lines were read from file C:\Users\terzi\Desktop\;ST NOT\GA\DFA.PR2:

```
!PRELIS SYNTAX: Can be edited  
SY='C:\Users\terzi\Desktop\;ST NOT\GA\DFA.PSF'
```





Groups: Memnuniyet Ölcegi

Models: X-Model

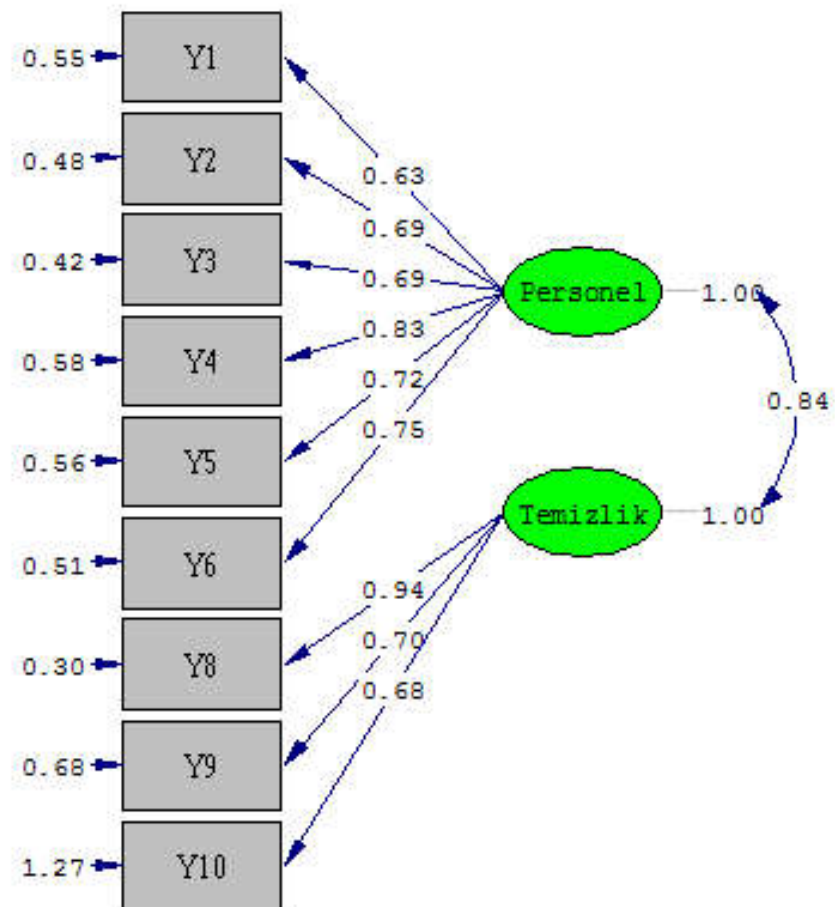
Estimates: Estimates

Observed Y

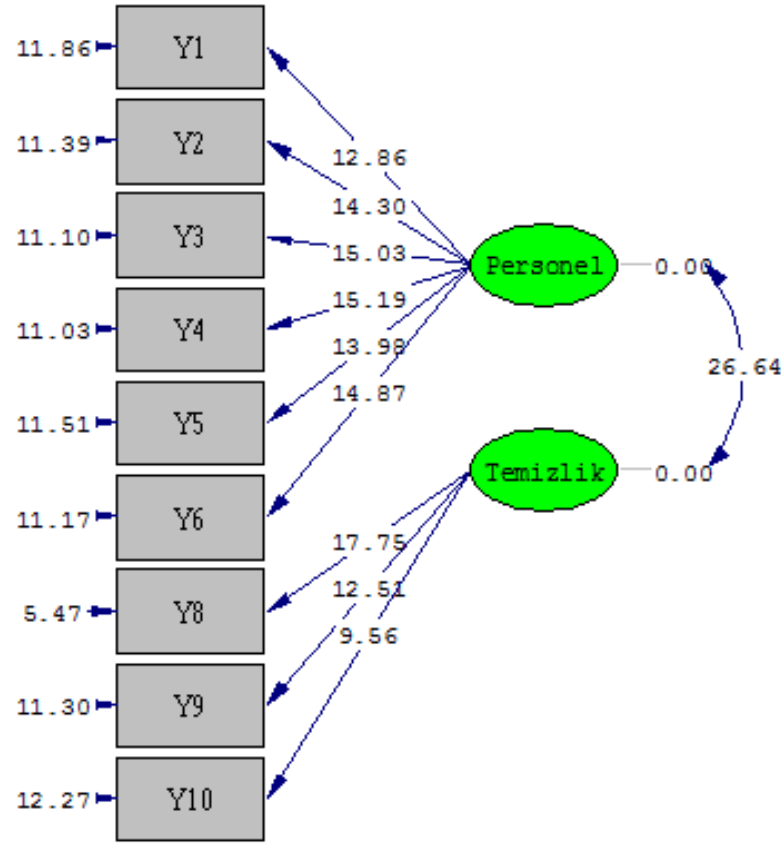
Y1	<input type="checkbox"/>
Y2	<input type="checkbox"/>
Y3	<input type="checkbox"/>
Y4	<input type="checkbox"/>
Y5	<input type="checkbox"/>
Y6	<input type="checkbox"/>
Y8	<input type="checkbox"/>
Y9	<input type="checkbox"/>
Y10	<input type="checkbox"/>

Latent Eta

Personel	<input type="checkbox"/>
Temizlik	<input type="checkbox"/>



Chi-Square=81.06, df=26, P-value=0.00000, RMSEA=0.078

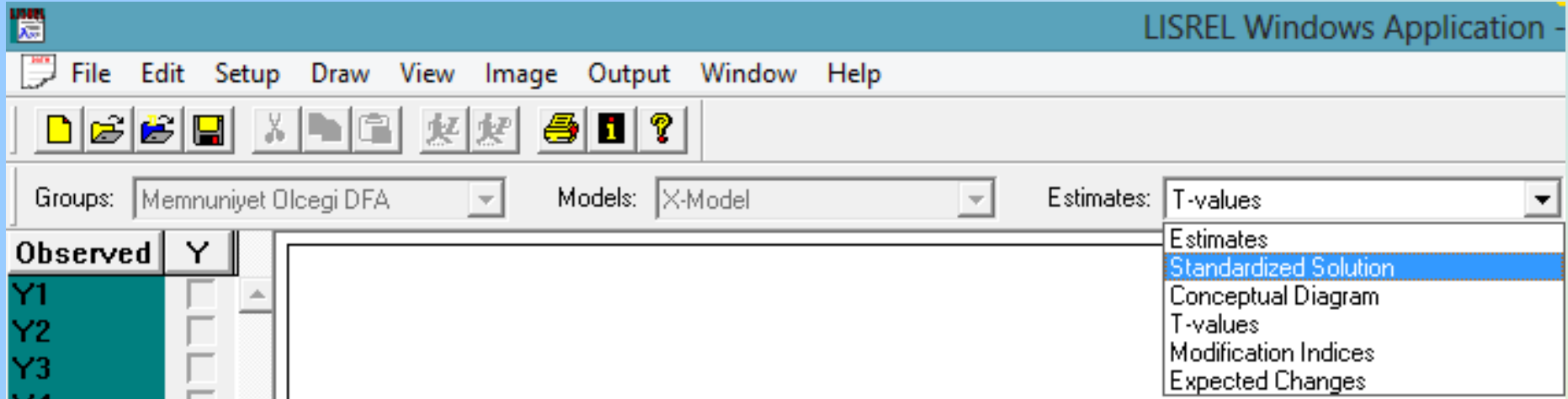


Chi-Square=81.06, df=26, P-value=0.00000, RMSEA=0.078

Gizil deęişkenlerin gözlenen deęişkenleri Açıklama Oranlarının minidarlık düzeyleri

T deęerleri t-tablo=1.96 ile karşılaştırılarak, hangi maddenin önemli olup olmadığı test edilir. Tüm maddeler önemli bulunmuştur. T deęerleri açısından bir problem yoksa analize devam edilir. Önemli bulunmayan maddeler kırmızı okla gösterilir ve o maddede bir sorun olduğu düşünülür. Kırmızı çizgili madde varsa ya modifikasyon önerileri doğrultusunda başka bir faktörle ilişkilendirilerek model uyumu test edilmeli ya da modelden çıkarılarak analize devam edilmelidir.

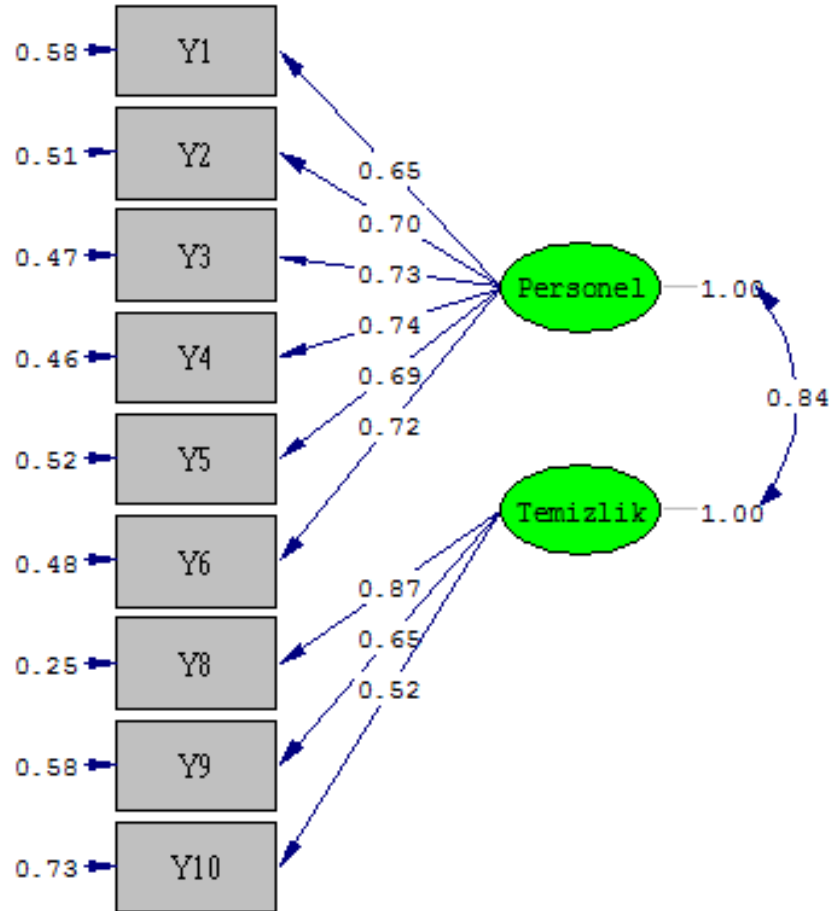
Bir maddeyi işlem dışı bırakmak için göstergelerin hata varyanslarının da kontrol edilmesi gerekir. Her bir maddenin faktör yük değerleri en az 0.30 ve üzeri olmalıdır. Bunun için Standardized Solution (standartlaştırılmış çözüm) yapılır.



Models: X-Model

Estimates: Standardized Solution

Hata varyansları



Chi-Square=81.06, df=26, P-value=0.00000, RMSEA=0.078

En yüksek hata varyansı Y10=0.73 Ancak Y10 önemli bulunduğundan işleminden çıkarılmaz

Window Help

Cascade

Tile

Arrange Icons

Close All

1 DFA.psf

2 DFA.OUT

3 DFA0

4 DFA0.OUT

5 DFA0.PTH

L I S R E L 8.71

BY

Karl G. Jöreskog & Dag Sörbom

This program is published exclusively by
Scientific Software International, Inc.
7383 N. Lincoln Avenue, Suite 100
Lincolnwood, IL 60712, U.S.A.

Phone: (800)247-6113, (847)675-0720, Fax: (847)675-2140
Copyright by Scientific Software International, Inc., 1981-2004
Use of this program is subject to the terms specified in the
Universal Copyright Convention.
Website: www.ssicentral.com

The following lines were read from file C:\Users\terzi\Desktop\;ST NOT\GA\DFA0.spl:

Memnuniyet Olcegi
Observed Variable
Y1 Y2 Y3 Y4 Y5 Y6 Y8 Y9 Y10
Covariance Matrix from File dfa.COV
Sample Size = 348
Latent Variables: Personel Temizlik
Relationships:
Y1 Y2 Y3 Y4 Y5 Y6 = Personel
Y8 Y9 Y10 = Temizlik
Path Diagram
End of Problem

Sample Size = 348

Memnuniyet Olcegi

Covariance Matrix

	Y1	Y2	Y3	Y4	Y5	Y6
Y1	0.95					
Y2	0.52	0.96				
Y3	0.47	0.51	0.89			
Y4	0.48	0.60	0.59	1.27		
Y5	0.43	0.41	0.50	0.60	1.08	
Y6	0.45	0.46	0.44	0.65	0.60	1.08
Y8	0.47	0.54	0.52	0.61	0.58	0.68
Y9	0.36	0.42	0.43	0.47	0.41	0.43
Y10	0.27	0.32	0.41	0.43	0.55	0.48

Covariance Matrix

	Y8	Y9	Y10
Y8	1.18		
Y9	0.66	1.18	
Y10	0.64	0.50	1.73

Y1 = 0.63*Personel, Errorvar.= 0.55 , R² = 0.42
(0.049) (0.046)
12.86 11.86

Y2 = 0.69*Personel, Errorvar.= 0.48 , R² = 0.49
(0.048) (0.042)
14.30 11.39

Y3 = 0.69*Personel, Errorvar.= 0.42 , R² = 0.53
(0.046) (0.038)
15.03 11.10

Y4 = 0.83*Personel, Errorvar.= 0.58 , R² = 0.54
(0.055) (0.053)
15.19 11.03

Y5 = 0.72*Personel, Errorvar.= 0.56 , R² = 0.48
(0.051) (0.049)
13.98 11.51

Y6 = 0.75*Personel, Errorvar.= 0.51 , R² = 0.52
(0.051) (0.046)
14.87 11.17

Y8 = 0.94*Temizlik, Errorvar.= 0.30 , R² = 0.75
(0.053) (0.054)
17.75 5.47

Y9 = 0.70*Temizlik, Errorvar.= 0.68 , R² = 0.42
(0.056) (0.060)
12.51 11.30

Y10 = 0.68*Temizlik, Errorvar.= 1.27 , R² = 0.27
(0.071) (0.10)
9.56 12.27

Correlation Matrix of Independent Variables

	Personel -----	Temizlik -----
Personel	1.00	
Temizlik	0.84 (0.03) 26.64	1.00

Goodness of Fit Statistics

Degrees of Freedom = 26
Minimum Fit Function Chi-Square = 76.17 (P = 0.00)
Normal Theory Weighted Least Squares Chi-Square = 81.06 (P = 0.00)
Estimated Non-centrality Parameter (NCP) = 55.06
90 Percent Confidence Interval for NCP = (31.64 ; 86.10)

Minimum Fit Function Value = 0.22
Population Discrepancy Function Value (F0) = 0.16
90 Percent Confidence Interval for F0 = (0.091 ; 0.25)
Root Mean Square Error of Approximation (RMSEA) = 0.078
90 Percent Confidence Interval for RMSEA = (0.059 ; 0.098)
P-Value for Test of Close Fit (RMSEA < 0.05) = 0.0085

Expected Cross-Validation Index (ECVI) = 0.34
90 Percent Confidence Interval for ECVI = (0.28 ; 0.43)
ECVI for Saturated Model = 0.26
ECVI for Independence Model = 7.50

RMSEA değeri %90 güven aralığıyla birlikte rapor edilmelidir. Üst sınırının 0,10'un altında kalması beklenir. Güven aralığının geniş olması ise daha geniş bir örnekleme ihtiyaç duyulduğuna işaret eder.

```
Chi-Square for Independence Model with 36 Degrees of Freedom = 2582.91
  Independence AIC = 2600.91
    Model AIC = 119.06
    Saturated AIC = 90.00
  Independence CAIC = 2644.58
    Model CAIC = 211.25
    Saturated CAIC = 308.35

  Normed Fit Index (NFI) = 0.97
  Non-Normed Fit Index (NNFI) = 0.97
  Parsimony Normed Fit Index (PNFI) = 0.70
  Comparative Fit Index (CFI) = 0.98
  Incremental Fit Index (IFI) = 0.98
  Relative Fit Index (RFI) = 0.96

  Critical N (CN) = 208.93

  Root Mean Square Residual (RMR) = 0.044
  Standardized RMR = 0.039
  Goodness of Fit Index (GFI) = 0.95
  Adjusted Goodness of Fit Index (AGFI) = 0.91
  Parsimony Goodness of Fit Index (PGFI) = 0.55
```

AIC: Hiyerarşik olmayan modeller kıyaslanırken kullanılır. Daha düşük değere sahip model tercih edilmelidir (Kline, 2005).

CFI: Tüm kovaryansların 0 kabul edildiği bir taban model ile kıyaslama yapar. 0,90 ve üzerinde olması önerilir (Hu & Bentler, 1999).

Stan. SRMR: 0,10'un altında olması önerilir (Klein, 2005).

UYGULAMA SONUÇLARI

H0: Gözlenen ve beklenen varyans-kovaryans matrisleri arasında fark yoktur.

H1: Gözlenen ve beklenen varyans-kovaryans matrisleri arasında fark vardır.

Ki-kare istatistiği: $p=0.000<0.05$

Doğrulayıcı faktör analizi sonuçlarına göre öncelikle p değerinin incelenmesi gerekmektedir. Bu değer, beklenen kovaryans matrisi ile gözlenen kovaryans matrisi arasındaki farkın manidarlığı hakkında bilgi vermektedir.

$\chi^2/sd \leq 2$ mükemmel uyum vardır.

$2 < \chi^2/sd < 5$ kabul edilebilir düzeyde bir uyum vardır (Kline, 2005).

Ki-kare/sd=81,06/26=3,12

RMSEA=0.078 <0.08 iyi bir uyum indeksi görülmektedir.

$0 < RMSEA \leq 0.05$ mükemmel uyum

$0.05 < RMSEA \leq 0.08$ iyi uyum (Jöroskog ve Sörbom, 1993)

$0.08 < RMSEA < 0.10$ zayıf uyum vardır (Tabachnick ve Fidel, 2001).

Uyum İndekslerinin Yorumu

➤ $GFI=0,95$ $AGFI=0,91$

GFI ve $AGFI > 0.95$ Mükemmel uyum

GFI ve $AGFI > 0.90$ İyi uyum (Hooper, Caughlan ve Mullen, 2008)

GFI mükemmel bir uyuma sahipken, $AGFI$ iyi bir uyuma sahiptir.

➤ $RMR=0.044$ $Stan.RMR=0.039$

$RMR-Sta.RMR < 0.05$ Mükemmel uyum

$RMR-Sta.RMR < 0.08$ İyi uyum

$RMR-Sta.RMR < 0.10$ zayıf uyum (Brown, 2006).

RMR ve Standardize RMR mükemmel bir uyuma sahiptir.

➤ $NNFI=0.97$ $CFI=0.98$

$NNFI$, $CFI > 0.95$ Mükemmel uyum

$NNFI$, $CFI > 0.90$ İyi uyum (Sümer, 2000)

$NNFI$ ve CFI indeksleri mükemmel uyuma sahiptir.

Bir modelin kabul edilebilir sayılması için RMSEA değerinin 0.08'in altında, CFI değerinin 0.95'in üzerinde olması koşuluna bakılır.

```
      The Modification Indices Suggest to Add the
Path to from      Decrease in Chi-Square      New Estimate
Y6      Temizlik      11.6      0.42
```

```
      The Modification Indices Suggest to Add an Error Covariance
Between      and      Decrease in Chi-Square      New Estimate
Y2      Y1      13.8      0.12
Y5      Y2      12.0      -0.12
Y6      Y3      14.4      -0.12
Y8      Y6      14.2      0.12
Y10     Y5      9.1      0.15
```

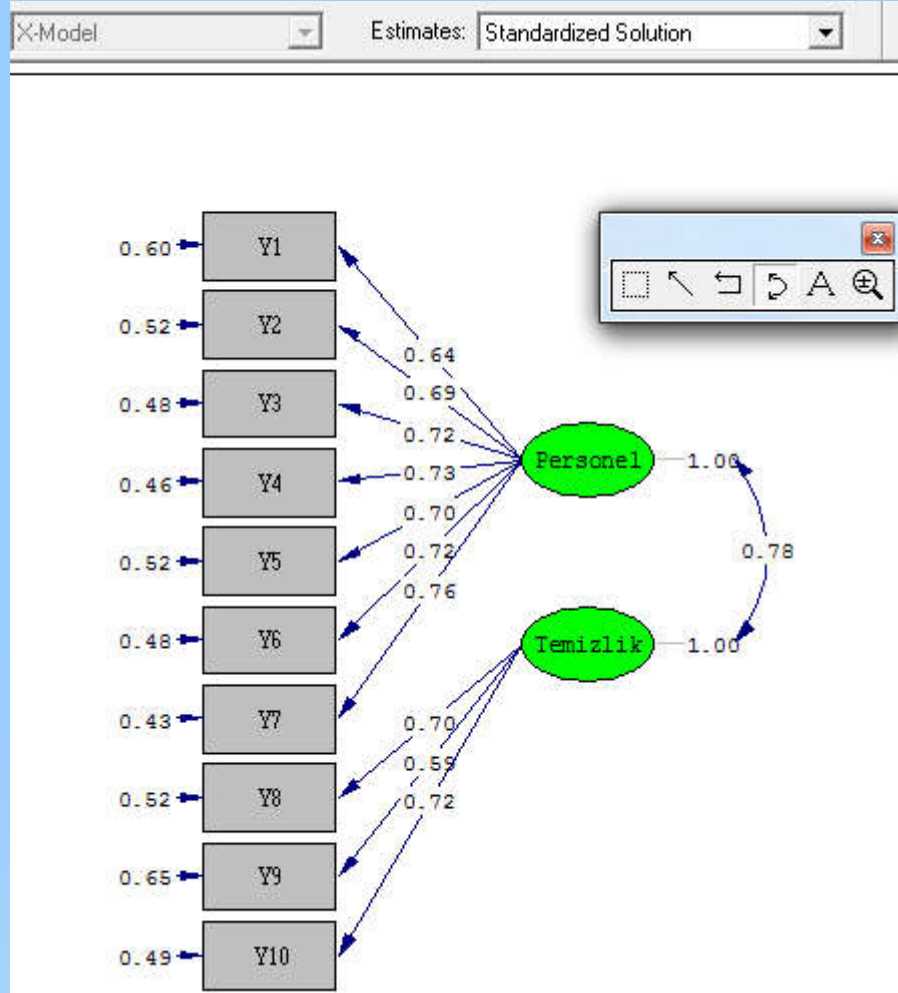
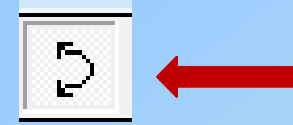
```
Time used:      0.031 Seconds
```



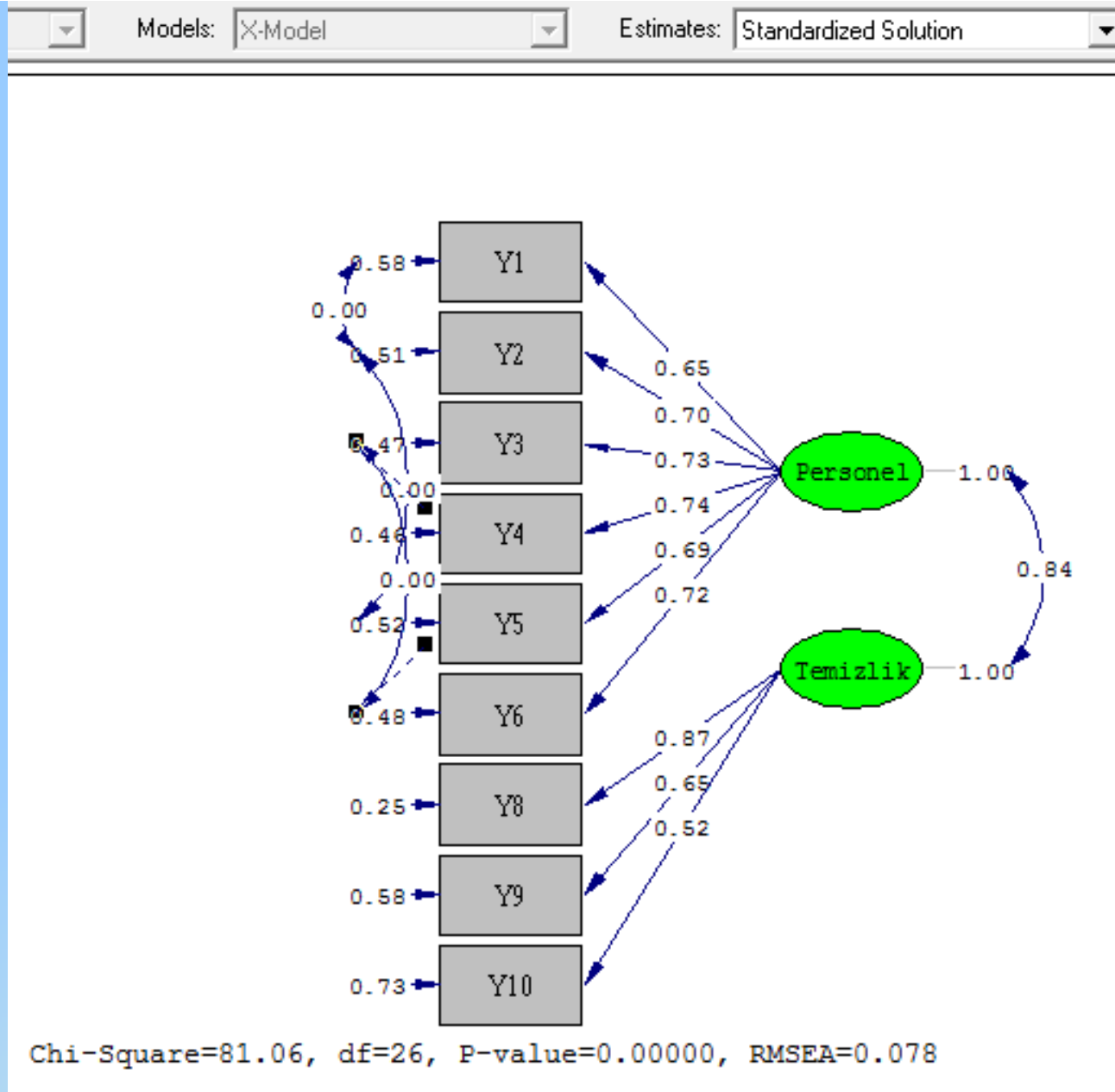
Eğer model uyum indekslerinden bir veya bir kaçı belirtilen kriterler düzeyinde değilse modifikasyon işlemi yapılmalıdır. Modifikasyon önerileri (The modification indices suggest to add an error covariance) incelendiğinde 5 modifikasyon önerisi ortaya çıkmıştır. DFA da sadece aynı faktör içerisinde yer alan maddelerle modifikasyon yapılabilir. Yani farklı iki faktörde yer alan maddeler arasında modifikasyon yapılmaz.

Y5-Y10 VE Y8-Y6 farklı faktörlerde yer aldığından modifikasyon yapılmaz. Y1-Y2, Y5-Y2, Y6-Y3 maddelerine modifikasyon yapılır.

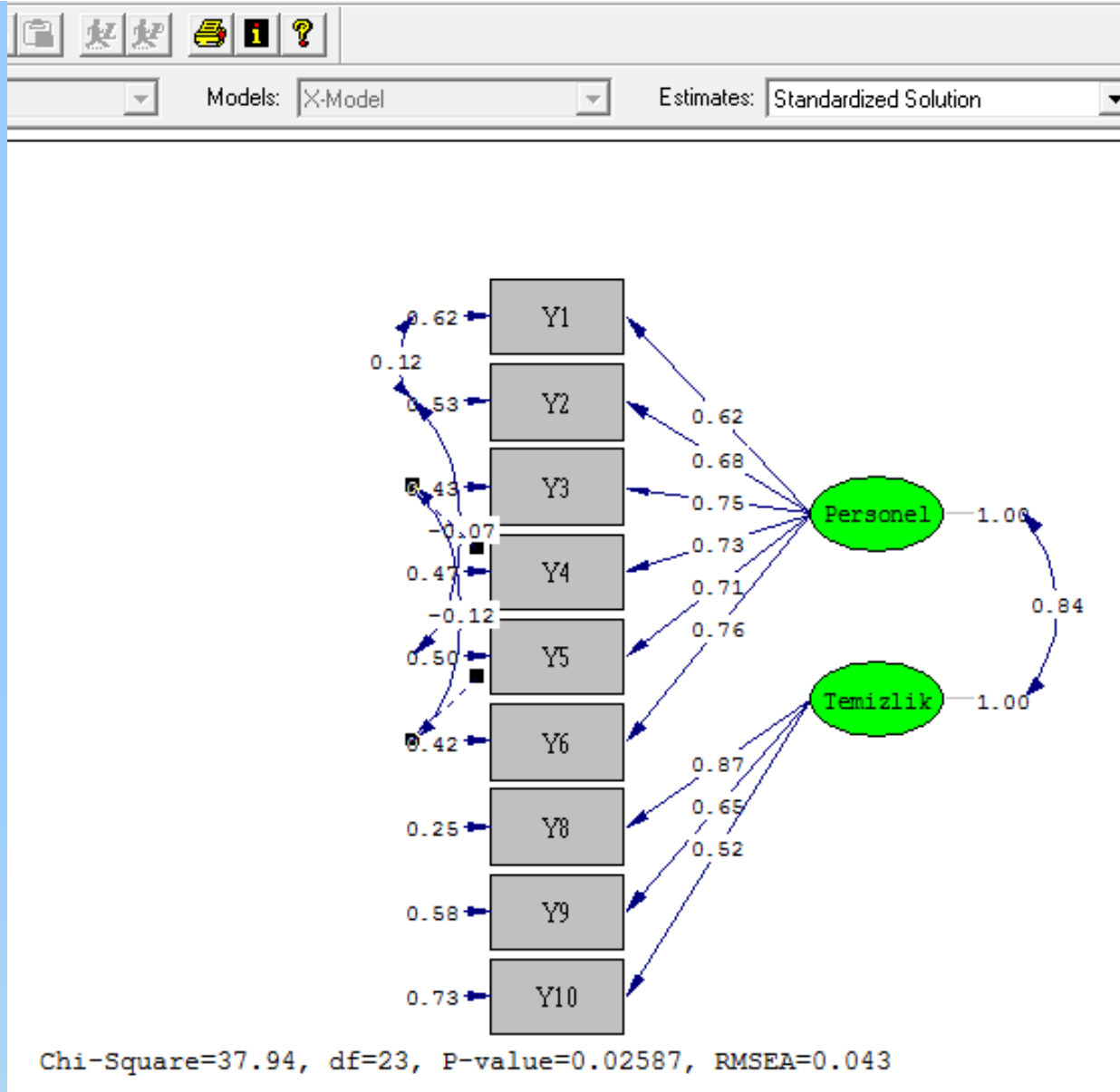
Modifikasyon yapımı:



Yukarıdaki simge seçilir. Hata varyanslarının yanında yer alan çubuğa tıklanır. Farenin sol tuşuna basılı tutulur ve el çekilmeden diğer modifikasyon yapılacak olan göstergenin hata varyansının yanında yer alan modifikasyon yapılacak olan göstergenin hata varyansının yanında yer alan çubuğa yuvarlak işareti çıkıncaya kadar ok çekilir. Modifikasyonlar tamamınca Run Lisrel komutu çalıştırılır.



Tüm modifikasyonlar yapıldıktan sonra F8 ve F5 tuşlarına sırayla basılır.



$Ki-kare/sd=37.94/23=1.65 < 2$ Mükemmel bir uyum vardır.

$RMSEA=0.043 < 0,05$ mükemmel bir uyum indeksi görülmektedir.

Goodness of Fit Statistics

Degrees of Freedom = 23

Minimum Fit Function Chi-Square = 38.62 (P = 0.022)

Normal Theory Weighted Least Squares Chi-Square = 37.94 (P = 0.026)

Chi-Square Difference with 3 Degrees of Freedom = 43.12 (P = 0.00)

Estimated Non-centrality Parameter (NCP) = 14.94

90 Percent Confidence Interval for NCP = (1.84 ; 35.92)

Minimum Fit Function Value = 0.11

Population Discrepancy Function Value (F0) = 0.043

90 Percent Confidence Interval for F0 = (0.0053 ; 0.10)

Root Mean Square Error of Approximation (RMSEA) = 0.043

90 Percent Confidence Interval for RMSEA = (0.015 ; 0.067)

P-Value for Test of Close Fit (RMSEA < 0.05) = 0.65

Expected Cross-Validation Index (ECVI) = 0.24

90 Percent Confidence Interval for ECVI = (0.20 ; 0.30)

ECVI for Saturated Model = 0.26

ECVI for Independence Model = 7.50

Chi-Square for Independence Model with 36 Degrees of Freedom = 2582.91

Independence AIC = 2600.91

Model AIC = 81.94

Saturated AIC = 90.00

Independence CAIC = 2644.58

Model CAIC = 188.69

Saturated CAIC = 308.35

Normed Fit Index (NFI) = 0.99
Non-Normed Fit Index (NNFI) = 0.99
Parsimony Normed Fit Index (PNFI) = 0.63
Comparative Fit Index (CFI) = 0.99
Incremental Fit Index (IFI) = 0.99
Relative Fit Index (RFI) = 0.98

Critical N (CN) = 375.11

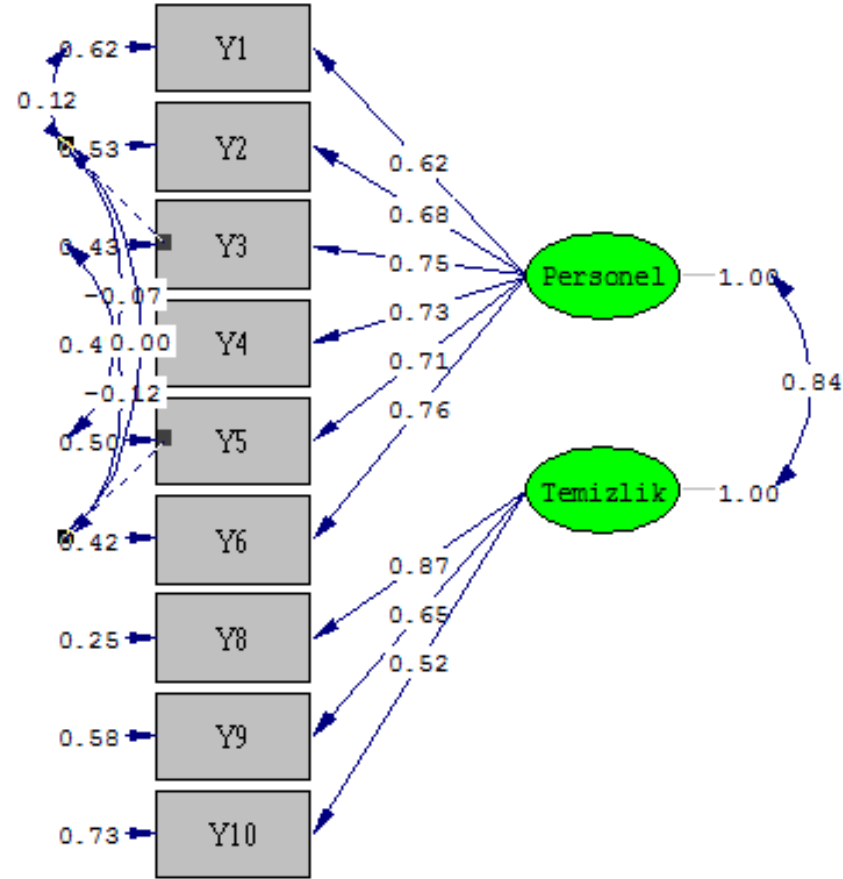
Root Mean Square Residual (RMR) = 0.034
Standardized RMR = 0.028
Goodness of Fit Index (GFI) = 0.98
Adjusted Goodness of Fit Index (AGFI) = 0.95
Parsimony Goodness of Fit Index (PGFI) = 0.50

The Modification Indices Suggest to Add an Error Covariance
Between and Decrease in Chi-Square New Estimate
Y6 Y2 10.3 -0.11
Y8 Y6 8.1 0.09

Time used: 0.016 Seconds

CFI=0.99>0,95
Çok iyi uyum var.

Y2-Y6 modifikasyonu yapılabilir. Y8-y6 farklı faktörlerde olduğu için modifikasyon yapılamaz.



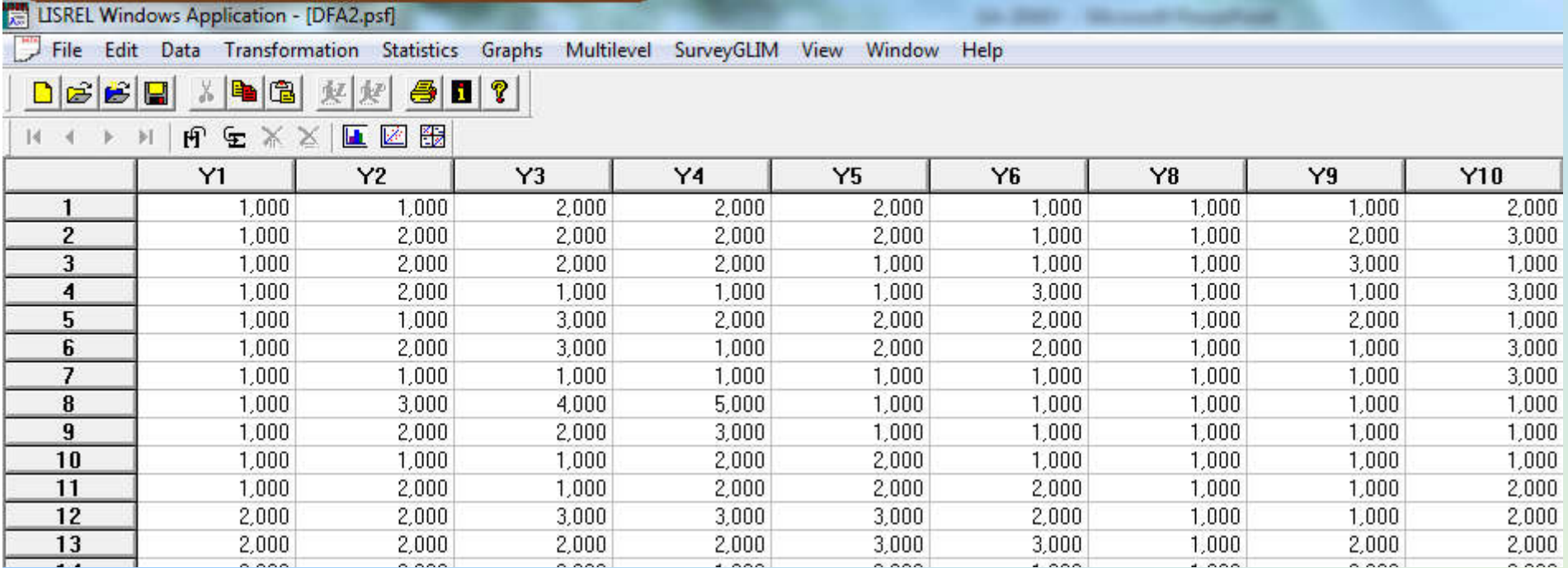
are=151.23, df=22, P-value=0.00000, RMSEA=0.130

Y2-Y6 modifikasyonu ki-kareye önemli bir katkı sağlamadığından (ki-kare=151.23 çok yükselmiştir) dolayı (RMSEA=0,13>0,08 değeri yükseldiğinden) modifikasyon yapılmaz. Ve işlem bitirilir.

Lisrel'de DFA Çözümü-II.Yöntem (Path Diagram)

LISREL Windows Application - [DFA2.psf]

File Edit Data Transformation Statistics Graphs Multilevel SurveyGLIM View Window Help



	Y1	Y2	Y3	Y4	Y5	Y6	Y8	Y9	Y10
1	1,000	1,000	2,000	2,000	2,000	1,000	1,000	1,000	2,000
2	1,000	2,000	2,000	2,000	2,000	1,000	1,000	2,000	3,000
3	1,000	2,000	2,000	2,000	1,000	1,000	1,000	3,000	1,000
4	1,000	2,000	1,000	1,000	1,000	3,000	1,000	1,000	3,000
5	1,000	1,000	3,000	2,000	2,000	2,000	1,000	2,000	1,000
6	1,000	2,000	3,000	1,000	2,000	2,000	1,000	1,000	3,000
7	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	3,000
8	1,000	3,000	4,000	5,000	1,000	1,000	1,000	1,000	1,000
9	1,000	2,000	2,000	3,000	1,000	1,000	1,000	1,000	1,000
10	1,000	1,000	1,000	2,000	2,000	1,000	1,000	1,000	1,000
11	1,000	2,000	1,000	2,000	2,000	2,000	1,000	1,000	2,000
12	2,000	2,000	3,000	3,000	3,000	2,000	1,000	1,000	2,000
13	2,000	2,000	2,000	2,000	3,000	3,000	1,000	2,000	2,000

Define Variables

Y1
Y2
Y3
Y4
Y5
Y6
Y8
Y9
Y10

Insert

Rename

Variable Type

Category Labels

Missing Values

OK

Cancel

To select more than one variable at a time, hold down the CTRL key while clicking on the variables to be selected

Variable Types for Y1 ...

- Ordinal
- Continuous
- Censored above
- Censored below
- Censored above and below

OK

Cancel

Apply to all

3,000	2,000	1,000
2,000	2,000	1,000
2,000	1,000	1,000
3,000	2,000	1,000
3,000	3,000	1,000
2,000	2,000	1,000
2,000	1,000	1,000
3,000	3,000	1,000
2,000	3,000	1,000
1,000	2,000	1,000

Yeni

Yeni

PRELIS Data
SIMPLIS Project
LISREL Project
Path Diagram

Tamam

iptal

LISREL Windows Application - [DFA2.pth]

File Edit Setup Draw View Image Output Window Help

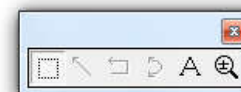


Groups: [dropdown] Models: [dropdown] Estimates: Estimates [dropdown]

Observed Y

VAR 1
VAR 2

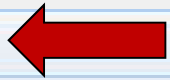
Latent Eta



Setup Draw View Image Output

Title and Comments ...

Groups...

Variables.. 

Data...

Build LISREL Syntax F4

Build SIMPLIS Syntax F8

Labels

Observed Variables	
	Name
1	VAR 1
2	VAR 2

Add/Read Variables

Move Down Move Up

Latent Variables	
	Name

Add Latent Variables

Move Down Move Up

< Previous

Next >

OK

Cancel

Press the Down Arrow to insert one row at a time once a label has been typed in the previous row

Press the Insert key to insert empty rows or the Delete key to delete selected rows

Observed Variables

	Name
1	VAR 1
2	VAR 2
3	Y1
4	Y2
5	Y3
6	Y4
7	Y5
8	Y6
9	Y8
10	Y9

Add/Read Variables

Move Down

Move Up

Press the Down Arrow to insert one row above the previous row

Press the Insert key to insert empty rows or the Delete key to delete selected rows

Add/Read Variables

Read from file: LISREL System File

Add list of variables (e.g., var1-var5):

File Name: Y1-Y6 Browse...

Info
Select one of the two system files. The LISREL data system file has a DSF extension and the PRELIS spreadsheet a PSF extension.

OK Cancel

Observed Variables

	Name
1	VAR 1
2	VAR 2
3	Y1
4	Y2
5	Y3
6	Y4
7	Y5
8	Y6
9	Y8
10	Y9

Add/Read Variables

Move Down

Move Up

Press the Down Arrow to insert one row above the previous row

Press the Insert key to insert empty rows or the Delete key to delete selected rows

Latent Variables

	Name
1	Personel
2	Temizlik

Add Latent Variables

Move Down

Move Up

< Previous

Next >

OK

Cancel

Add Variables

Add one or list of variables here (e.g., var1 - var5):

Personel

OK Cancel


LISREL Windows Application - [DFA2.pth]

File Edit Setup Draw View Image Output Window Help

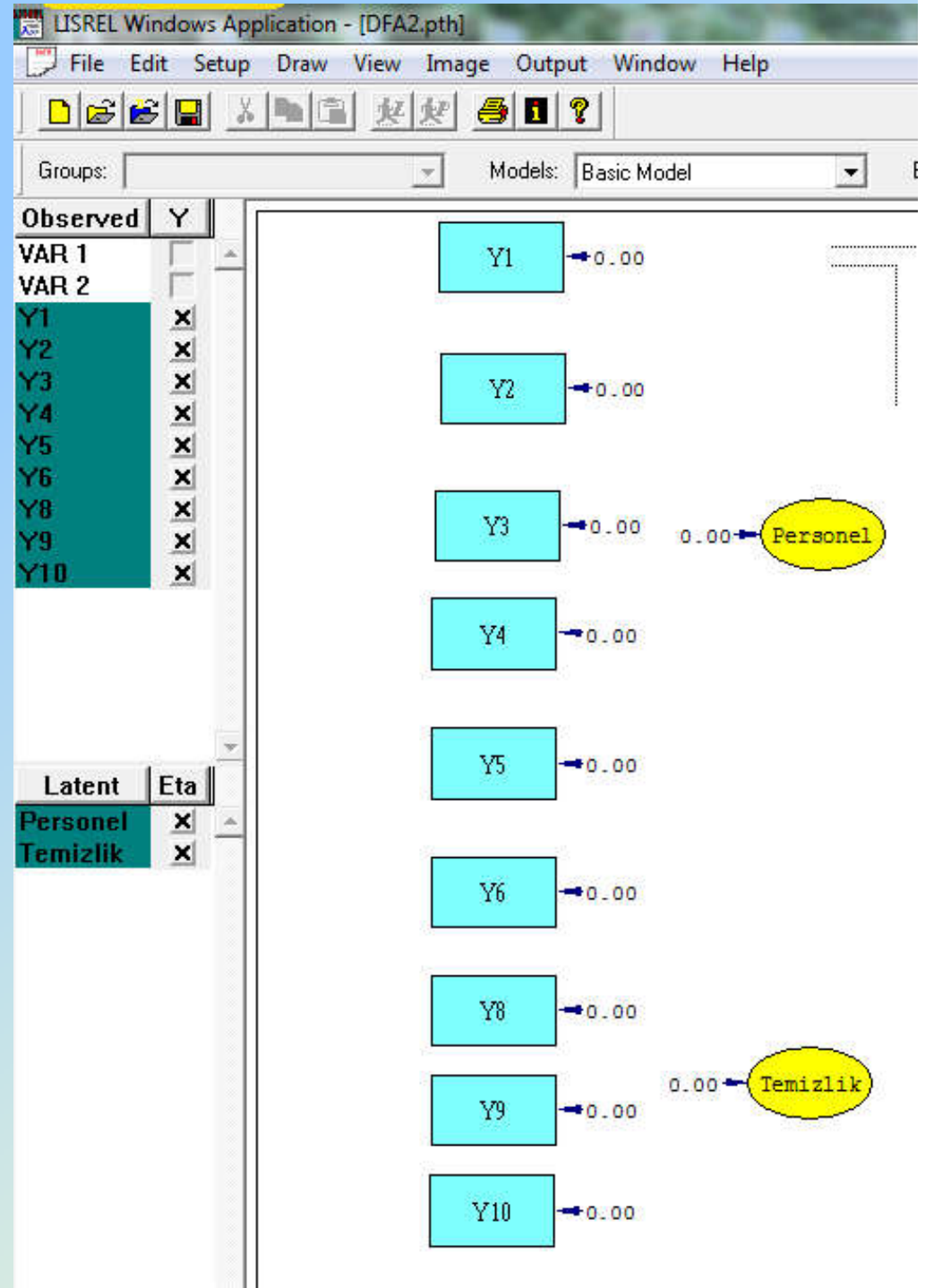
Groups: Models: Estimates: Estimates

Observed	Y
VAR 1	<input type="checkbox"/>
VAR 2	<input type="checkbox"/>
Y1	<input type="checkbox"/>
Y2	<input type="checkbox"/>
Y3	<input type="checkbox"/>
Y4	<input type="checkbox"/>
Y5	<input type="checkbox"/>
Y6	<input type="checkbox"/>
Y8	<input type="checkbox"/>
Y9	<input type="checkbox"/>
Y10	<input type="checkbox"/>

Latent	Eta
Personel	<input type="checkbox"/>
Temizlik	<input type="checkbox"/>



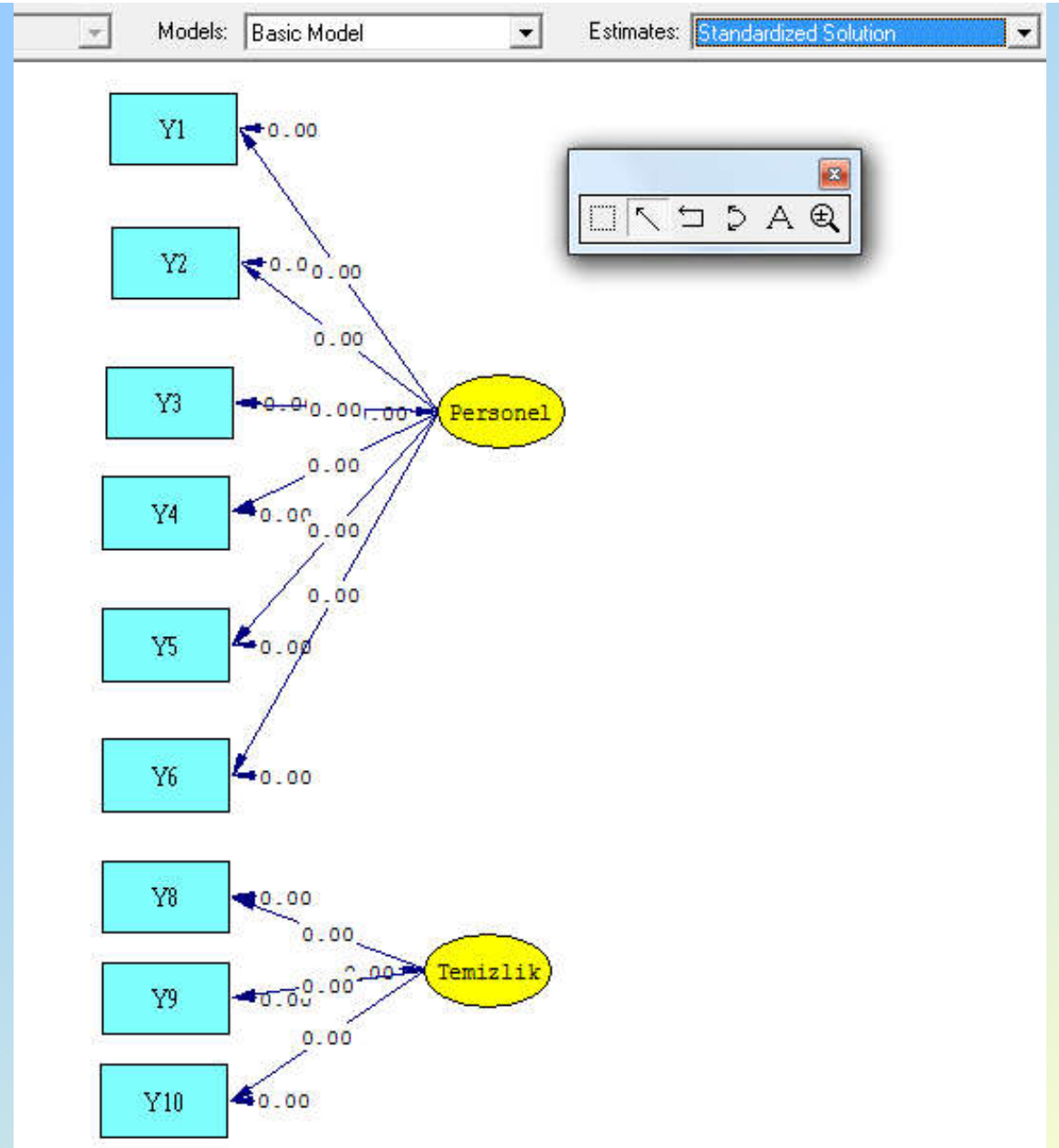
Maddeler seçilip, sağ taraf taşınır .
Ve o maddelere faktör atanır.



Her bir maddenin önceden belirtilmiş faktörlerle ilişkilendirilir. Bunun için araç kutusundan Yararlanılır.



Araç kutusundaki ok işareti tıklanarak, ilgili maddeler ile ilgili faktörler birbirine bağlanır.



Bağlama işlemi bitince F8 ve F5 tuşlarına basılır.

The screenshot shows the LISREL Windows Application interface. The main window displays the following text:

```
SYSTEM FILE from file 'C:\Users\terzi\Desktop\İST NOT\GA\UYGULAMA\Lisrel\DFA2.DSF'  
Latent Variables  Personel Temizlik  
Relationships  
Y1 = Personel  
Y2 = Personel  
Y3 = Personel  
Y4 = Personel  
Y5 = Personel  
Y6 = Personel  
Y8 = Temizlik  
Y9 = Temizlik  
Y10 = Temizlik  
Path Diagram  
End of Problem  
*
```

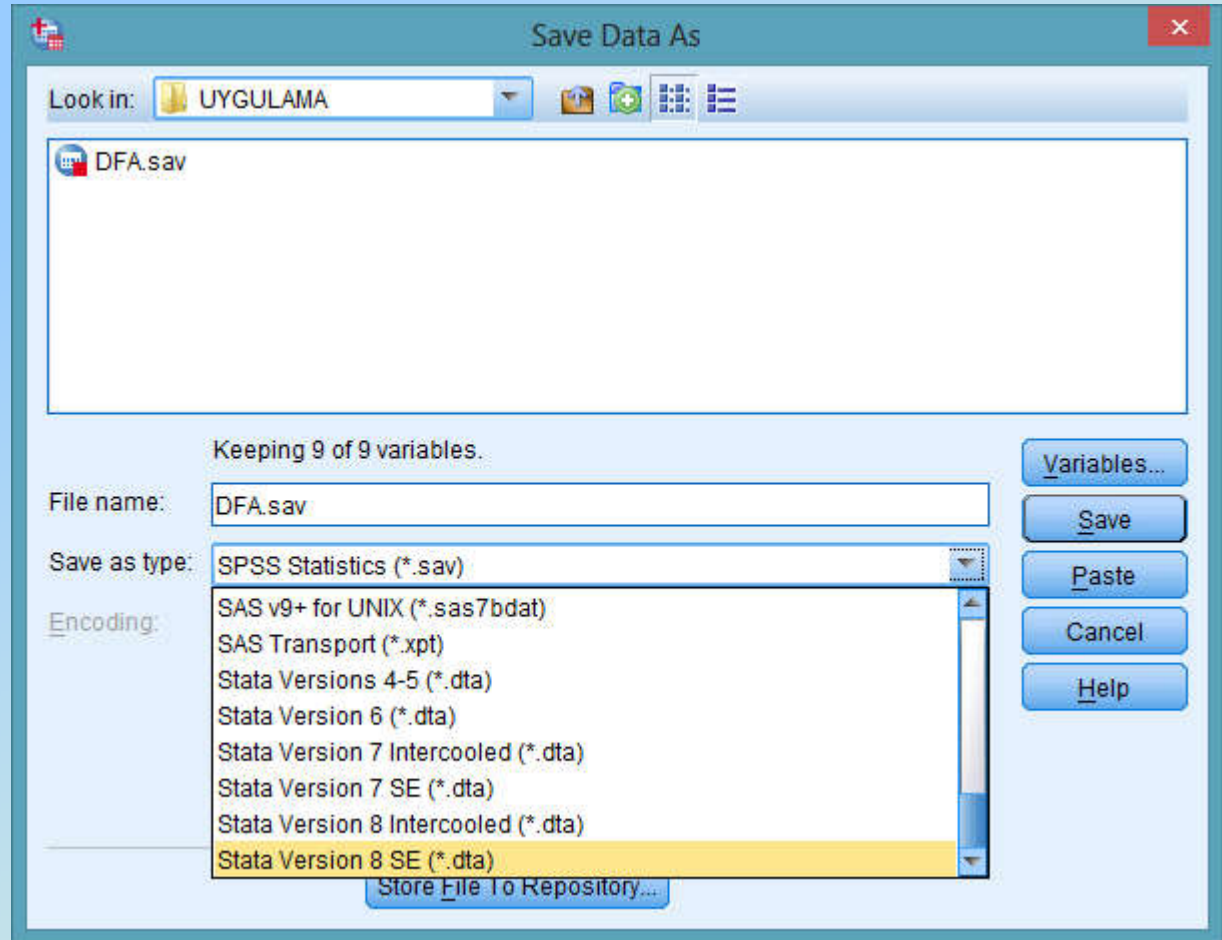
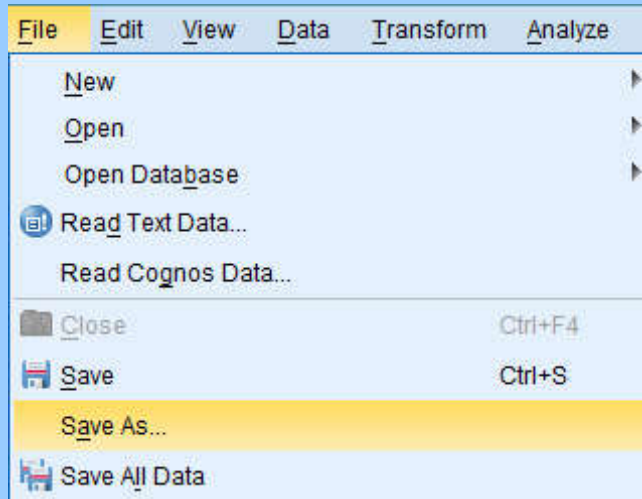
Below the main window is a control panel with the following elements:

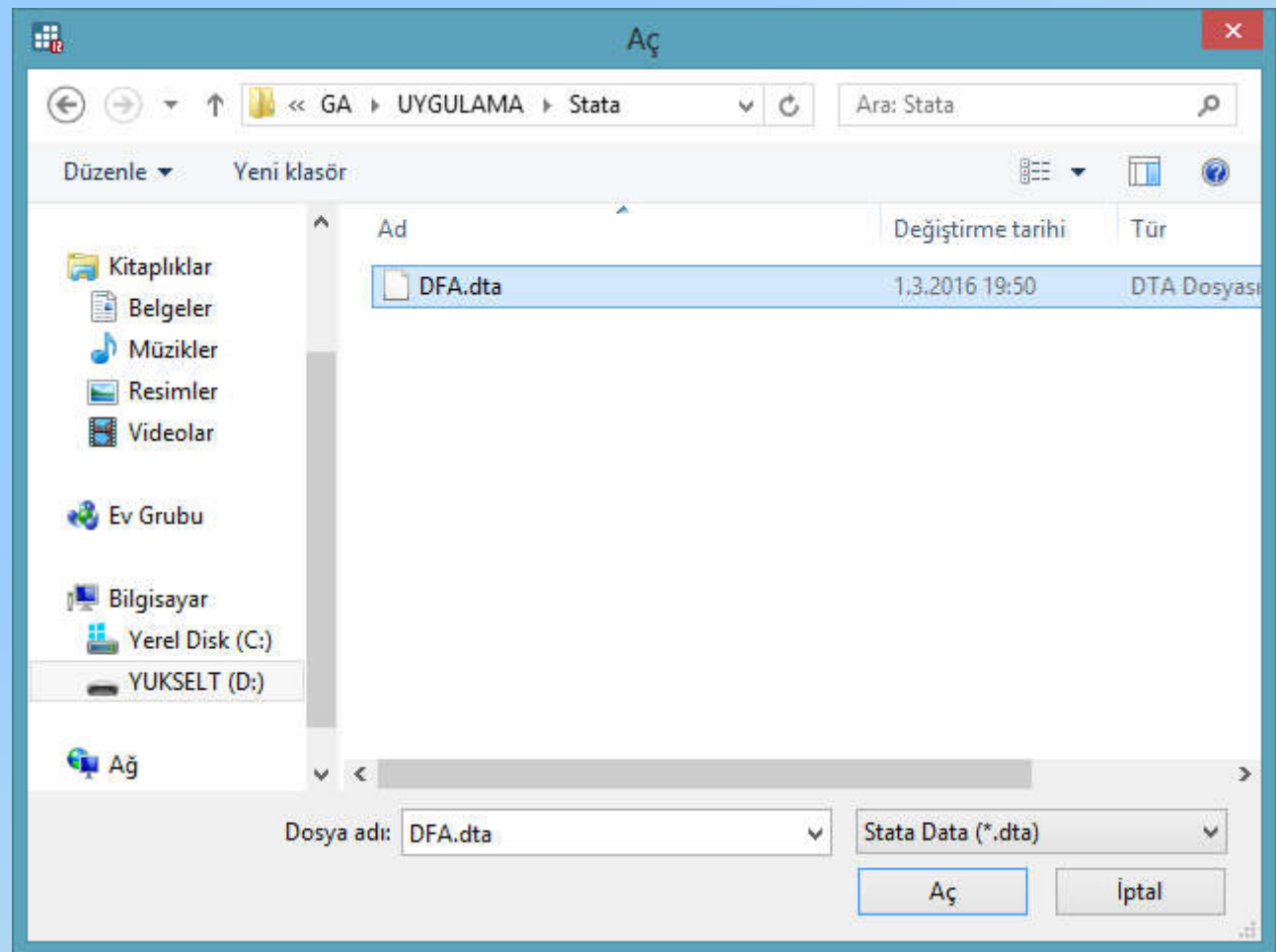
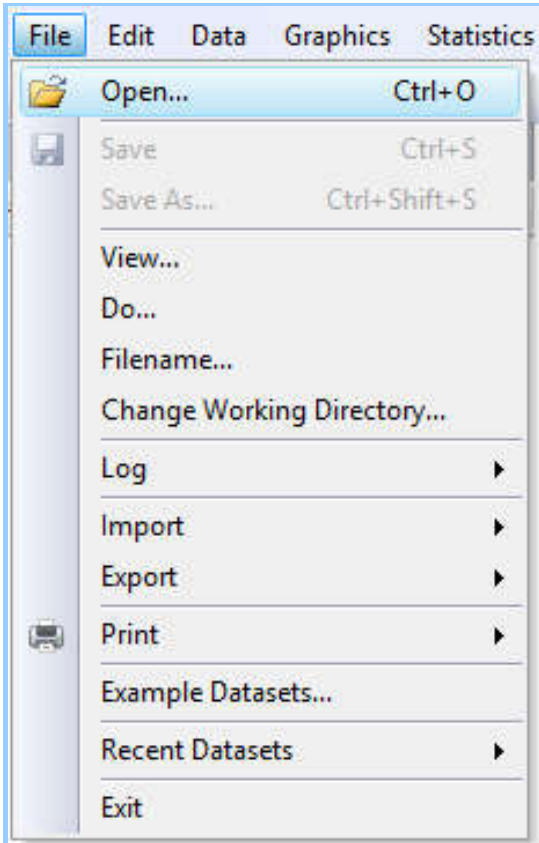
- Observed**: VAR 1, VAR 2, Y1, Y2, Y3, Y4, Y5, Y6, Y8, Y9, Y10
- Latent**: Personel, Temizlik
- Groups**: A dropdown menu.
- From**: Set Path, /, *, ., >
- To**: Set Variance, 7, 8, 9, <==
- Free**: Set Covariance, 4, 5, 6, =
- Fix**: Set Error Variance, 1, 2, 3, {, <-|
- Equal**: Set Error Covariance, 0, ., }



STATA İLE DFA

SPSS'de dosya stata dosyası olarak kayıt edilir.





Statistics User Window Help

- Summaries, tables, and tests ▶
- Linear models and related ▶
- Binary outcomes ▶
- Ordinal outcomes ▶
- Categorical outcomes ▶
- Count outcomes ▶
- Exact statistics ▶
- Endogenous covariates ▶
- Sample-selection models ▶
- Multilevel mixed-effects models ▶
- Generalized linear models ▶
- Nonparametric analysis ▶
- Time series ▶
- Multivariate time series ▶
- State-space models ▶
- Longitudinal/panel data ▶
- Survival analysis ▶
- Epidemiology and related ▶
- SEM (structural equation modeling) ▶**
 - Model building and estimation**
 - Testing and CIs ▶
 - Goodness of fit ▶
 - Group statistics ▶
 - Predictions ▶
 - Other ▶
- Survey data analysis ▶
- Multiple imputation ▶
- Multivariate analysis ▶
- Power and sample size ▶
- Resampling ▶

```

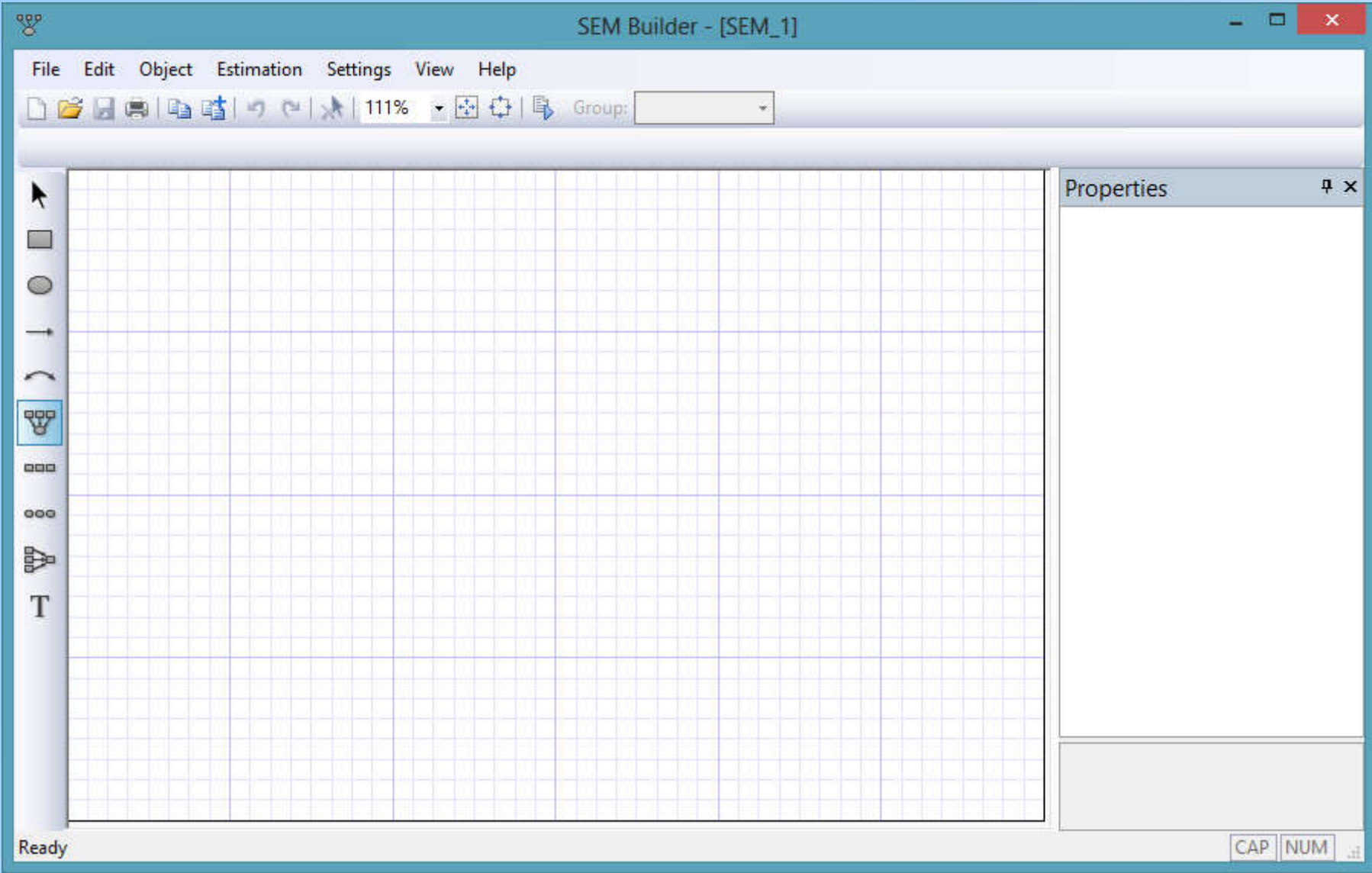
(R)
12.0 Copyright 1985-2011
StataCorp
4905 Lakeway Drive
College Station, Te
800-STATA-PC
979-696-4600
979-696-4601 (fax)

: perpetual license:
611859953
ATAforAll
ATA

-set maxvar-) 5000 maximum varia

ta\DFA.dta", clear

```

Measurement component

Main Distances Connections

Latent variable name:
PERSONEL

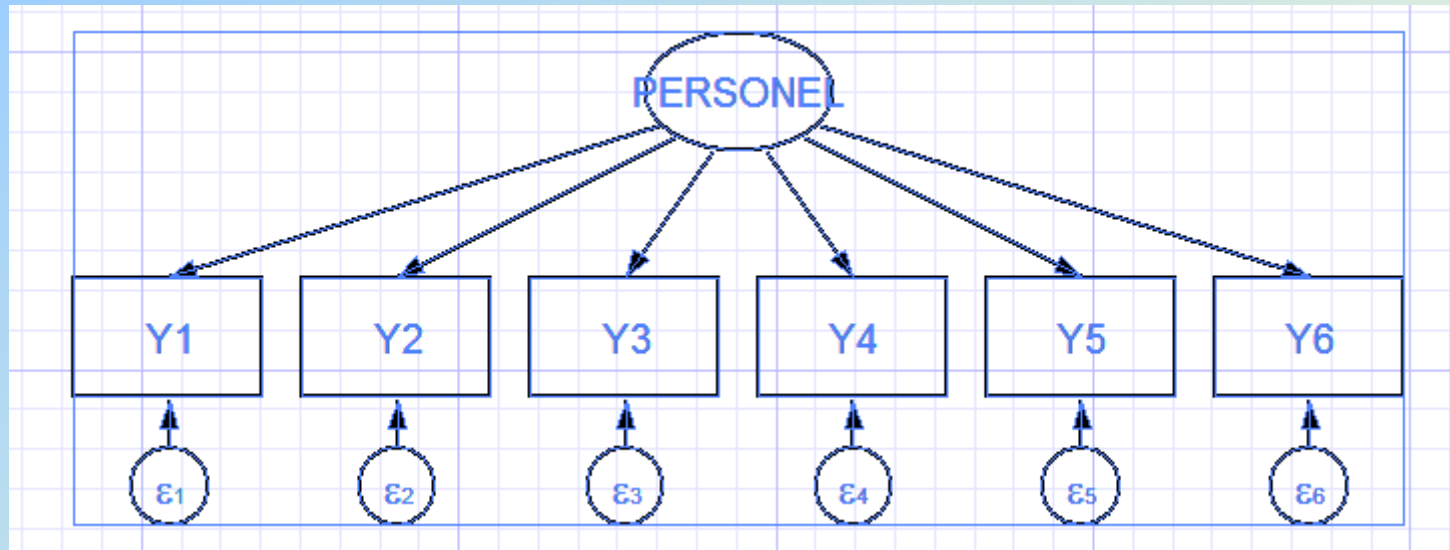
Measurement variables:
 Select variables
 Specify number of variables

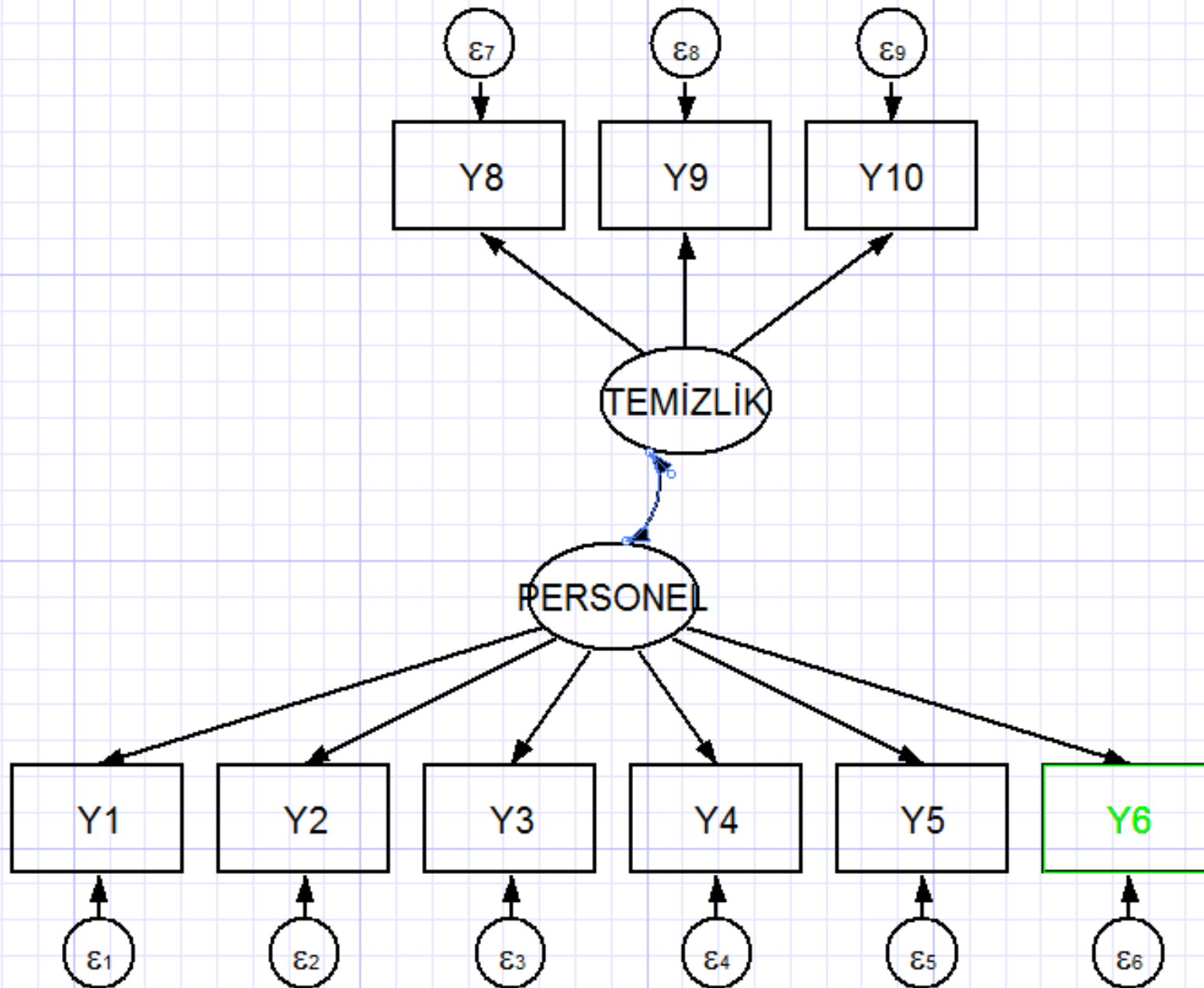
Measurement variables:
Y1 Y2 Y3 Y4 Y5 Y6

Do not estimate constants


Measurement direction:
Down

OK Cancel





Estimation Settings View

Estimate 

Clear Estimates

Testing and CIs ▶

Goodness of fit ▶

Group statistics ▶

Predictions ▶

Other ▶

SEM estimation options

Group Model if/in Weights SE/Robust Reporting Maximization

Standard analysis (no groups)

Group analysis

Group variable:

Parameters that are equal across groups:

Summary statistic data options:

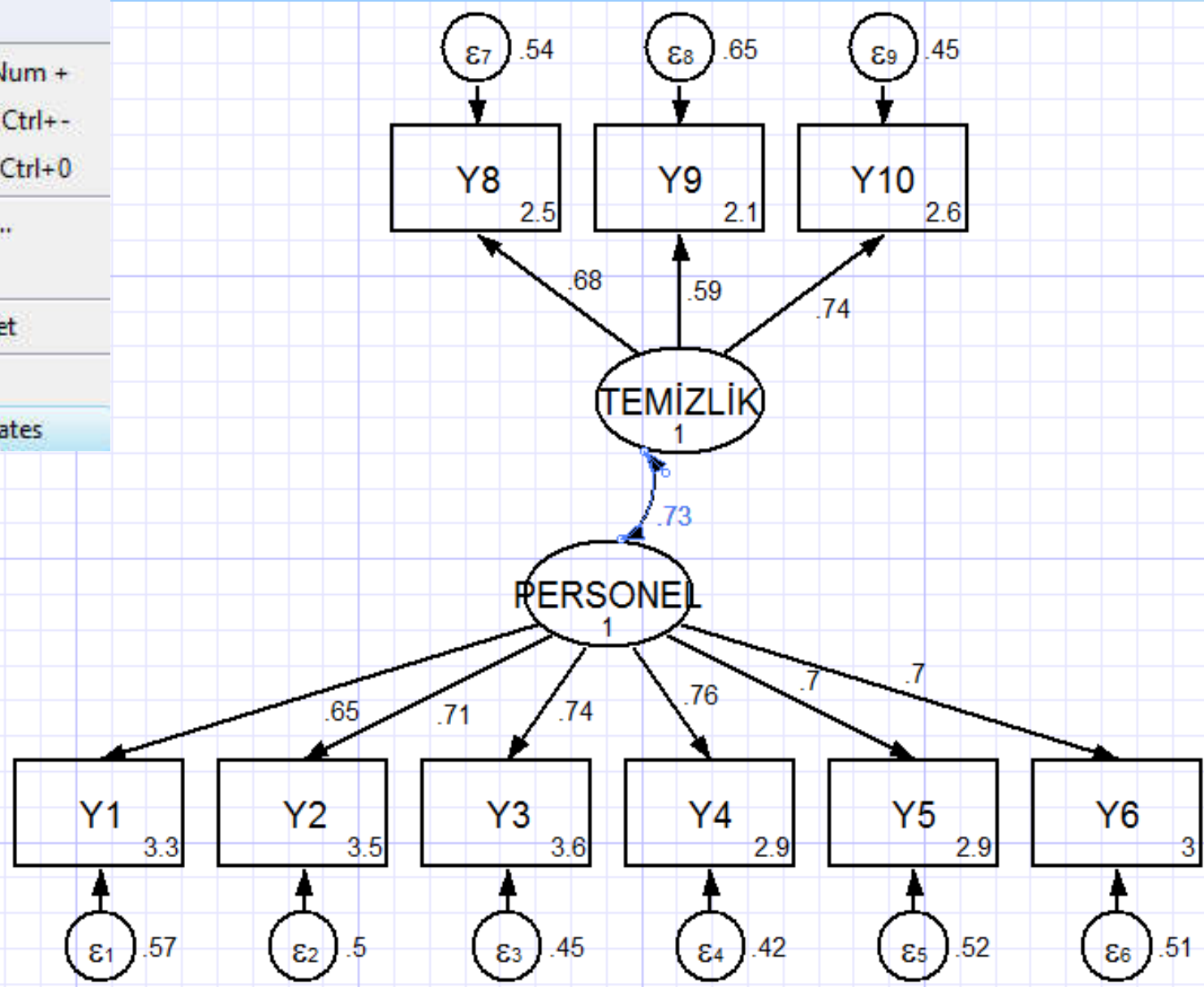
Select summary groups:

Allow groups and pooling of SSD correlations

OK Cancel Submit

View Help

- Zoom In Ctrl+Num +
- Zoom Out Ctrl+-
- Fit in Window Ctrl+0
- Adjust Canvas Size...
- Show Grid
- Show Property Sheet
- Show Estimates
- Standardized Estimates



Estimation Settings View Help

Estimate
Clear Estimates
Testing and CIs
Goodness of fit
Group statistics
Predictions
Other

Group: []

Overall goodness of fit
Equation-level goodness of fit
Matrices of residuals
Information criterion

estat - Postestimation tools for sem

Reports and statistics: (subcommand)

- Equation-level tests that all coefficients are zero (eqtest)
- Display estimation results in modeling framework (framework)
- Group-level goodness-of-fit statistics (ggof)
- Tests for invariance of parameters across groups (ginvariant)
- Goodness-of-fit statistics (gof)**

Statistics to be displayed:

all

Suppress descriptions of statistics

OK Cancel Submit

```
. estat gof, stats(all)
```

Fit statistic	Value	Description
Likelihood ratio		
chi2_ms (26)	95.147	model vs. saturated
p > chi2	0.000	
chi2_bs (36)	1219.754	baseline vs. saturated
p > chi2	0.000	
Population error		
RMSEA	0.087	Root mean squared error of approximation
90% CI, lower bound	0.069	
upper bound	0.107	
pclose	0.001	Probability RMSEA <= 0.05
Information criteria		
AIC	8282.236	Akaike's information criterion
BIC	8390.097	Bayesian information criterion
Baseline comparison		
CFI	0.942	Comparative fit index
TLI	0.919	Tucker-Lewis index
Size of residuals		
SRMR	0.041	Standardized root mean squared residual
CD	0.943	Coefficient of determination



✓ KAYNAKLAR

- ✓ Alpar, R. (2016). Uygulamalı İstatistik ve Geçerlik-Güvenilirlik, Detay Yayıncılık, Ankara.
- ✓ Şencan, H. (2005). Sosyal ve Davranışsal Ölçümlerde Güvenilirlik ve Geçerlilik, Seçkin Yayıncılık, Ankara.
- ✓ Altunışık, R. ve ark. (2005). Sosyal Bilimlerde Araştırma Yöntemleri, Sakarya Kitabevi, Adapazarı.
- ✓ Büyüköztürk, Ş. (2005). Sosyal Bilimler İçin Veri Analizi El Kitabı, Pegema yayıncılık, Ankara.
- ✓ Kalaycı, Ş. Ve ark.(2005). SPSS Uygulamalı Çok Değişkenli İstatistiksel Teknikler, Asil Yayın Dağıtım, Ankara.
- ✓ Özdamar, K. (2002). Paket Programlar ile İstatistiksel veri Analizi 1, Kaan Kitabevi, Eskişehir.
- ✓ Baş, T. (2001). Anket, Seçkin Kitabevi, Ankara.
- ✓ Karagöz, Y. (2014). SPSS 21.1 Uygulamalı Biyoistatistik, Nobel Yayın, Ankara.
- ✓ Çokluk Ö., Şekercioğlu G., Büyüköztürk Ş. Sosyal bilimler için çok değişkenli istatistik SPSS ve Lisrel Uygulamaları, Pegem Akademi, Ankara, 2014.
- ✓ Seçer İ. (2015). SPSS ve Lisrel ile pratik veri analizi- Analiz ve raporlaştırma, Anı yayıncılık, Ankara.
- ✓ Çepni Z. (2010). Yapısal eşitlik modellemesi, Hacettepe Ün., <http://yunus.hacettepe.edu.tr/~cepni/mersinsemsunu.ppt>

- ✓ Atkinson, G. and A. M. (1998). Nevill. Statistical methods for assessing measurement error (Reliability) in variables relevant to sports medicine. *Sports Medicine*. 26:217-238.
- ✓ Baykul, Y. (2000). Eğitimde ve Psikolojide Ölçme: Klasik test Teorisi ve Uygulaması. Ankara: ÖSYM yayınları.
- ✓ Crocker, L. ve Algina, J. (1986). Introduction to Classical and Modern Test Theory. New York: Holt. Rinehart and Winston.
- ✓ Erkuş, A. (2003). Psikometri Üzerine Yazılar. Türk Psikologlar Derneği Yayınları no:24, 1. basım: Ankara.
- ✓ Horst, P.(1966). Psychological measurement and prediction. California: Wadsworth.
- ✓ Ercan, İ., Kan, İ. (2004). Ölçeklerde Güvenilirlik ve Geçerlilik, Uludağ Ün. Tıp Fak. Dergisi, 30(3), 211-216.
- ✓ Karagöz, Y. (2014). SPSS 21.1 Uygulamalı Biyoistatistik, Nobel Yayın, Ankara.
- ✓ Çılan, Ç.A. (2013). Sosyal Bilimlerde Kategorik Verilerle İlişki Analizi, Pegem Akademi, Ankara.
- ✓ Bentler, P.M. (1990). Comparative fit indexes in structural models. *Psychol Bull*, 107(2),238-46.
- ✓ Tekin, H. (2000). *Eğitimde ölçme ve değerlendirme*. Ankara: Yargı Yayınları.Yükseköğretim Kurulu (YÖK) (2012). <http://www.yok.gov.tr/content/view/471/>
- ✓ Şarлак, Y. (2010). Eğitimde Ölçme ve Değerlendirme, İstanbul.
- ✓ Çalık, M. Test Geliştirme ve Madde Analizi, <https://fatihegitim.files.wordpress.com/2011/02/hafta-14-madde-analizi.ppt>
- ✓ Şimşek, Ö. Ve Ceyhun, Ş.Y. Test ve Madde Analizleri, <http://sevimasiroglu.com/wp-content/uploads/2016/12/TEST-VE-MADDE-ANAL%C4%B0ZLER%C4%B0.pdf>
- ✓ Gömleksiz, M. ve Erkan, S. (2016). Eğitimde Ölçme ve Değerlendirme, Nobel Kitabevi, İstanbul.
- ✓ Erkuş, A. (2014). Psikolojide Ölçme ve Ölçek geliştirme-I, Pegem Akademi, Ankara.

- ✓ Ayyıldız, H., Cengiz, ve Ustasüleyman, T. (2006). Üretim ve Pazarlama Bölüm Çalışanları Arası Davranışsal Değişkenlerin Firma Performansı Üzerine Etkisine İlişkin Yapısal Bir Model Önerisi, Muğla Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, Güz 2006, (17,)17.
- ✓ Bollen, K.A. (1989). Structural equations with latent variables. New York: John Wiley & Sons.
- ✓ Schumacker, R.E. & Lomax,R.G.(1996). A beginner's guide to structural equation modeling. Mahwah, New Jersey: Lawrence Erlbaum Associates, Publishers.
- ✓ Schermelleh-Engel, K., Moosbrugger, H., and Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive Goodness-Of-Fit Measures Of Psychological Research Online, (8), 2, 23-74
- ✓ Özabacı, N. (2011). İlişki Niteliği Ölçeği'nin Türkçe Uyarlaması: Geçerlik ve Güvenirlik Çalışması, Eğitim ve Bilim , 36 (162), 159-167.
- ✓ Munro B.H. (2005). Statistical Methods For Health Care Research. Philadelphia: Lippincott Williams & Wilkins, p.351-76.
- ✓ Waltz C.F. (2010). Strcikland OL, Lenz ER. Measurement in Nursing and Health Research. New York: Springer Publishing Company, p.176-8.
- ✓ Çapık, C. (2014). Geçerlik Ve Güvenirlik Çalışmalarında Doğrulayıcı Faktör Analizinin Kullanımı, Anadolu Hemşirelik ve Sağlık Bilimleri Dergisi, 17(3), 196-205.
- ✓ M.Sözbilir (2010). Ölçme ve Değerlendirme.
- ✓ Büyüköztürk, Ş. (2011). *Sosyal bilimler için veri analizi el kitabı*. Ankara: Pegem Yayıncılık.
- ✓ Büyüköztürk, Ş., Çakmak, E.K., Akgün, Ö.E., Karadeniz, Ş. ve Demirel, F. (2010). *Bilimsel araştırma yöntemleri*. Ankara: Pegem Yayıncılık.
- ✓ Güler, N. (2014). Eğitimde Ölçme ve Değerlendirme, Pegem Akademi: Ankara.