💹 Doğrulayıcı Faktör Analizi

📉 Temel Kavramlar

Dr. Kubra Atalay Kabasakal
Bahar 2023

YEM

Yapısal eşitlik modelleri

- yol analizi
- doğrulayıcı faktör analizi
- çoklu grup doğrulayıcı faktör analizi

YEM

- Hipotez edilen **modelin veri ile uyumunun incelendiği YEM analizlerinde**, model ile veri uyumu sağlandığında, veri **hipotez edilen modeli destekler** ancak **alternatif hipotezleri/modelleri reddetmez.**

- Diğer bir ifadeyle **farklı modeller de veriye uyum sağlayabilir.** Dolayısıyla araştırmacı birden fazla modelin veri ile uyumunu inceleyebilir ve
  - **en iyi uyum gösteren modeli belirleyebilir.**

YEM

- YEM analizlerinde Rosseel vd. (2018) tarafından geliştirilen **lavaan** paketi;
- YEM analizleri çıktılarında Tsukahara (2022) tarafından geliştirilen **semoutput** paketi;
- analizlere ilişkin diyagramların elde edilmesinde ise Epskamp vd. (2017) tarafından geliştirilen **semPlot** paketi kullanılacaktır.

```r
# install.packages(c("lavaan","semPlot"))
library(lavaan)
library(semPlot)

# devtools::install_github("dr-JT/semoutput")
library(semoutput)
```

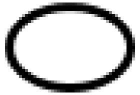
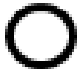
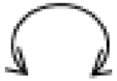
YEM

- Veri modellemenin amacı, **verinin yapısını daha anlaşılabilecek ve kolay yorumlanabilecek şekilde tanımlamaktır.**

- Yapısal eşitlik modelleri (YEM) farklı modellerin test edilebilmesi için **esnek bir çerçeve sağlar.**

- Temelde gözlenenen değişkenlerin **varyanslarına** ve gözlenen değişkenler arasındaki **kovaryanslara** dayalı olan YEM analizlerinin amacı bir grup gözlenen değişken arasındaki kovaryans örüntüsünü anlamak ve **araştırma modeli ile gözlenen değişkenlerin varyanslarını açıklamaktır.**

YEM

Sembol	Tanım	Gösterim
	Gizil değişken	η: içsel gizil ξ: dışsal gizil
	Gözlenen Değişken	Y: içsel gözlenen X: dışsal gözlenen
	Hata/artık	ζ: gizil içsel değişken için hata ε: Y değişkeni için hata δ: X değişkeni için hata
X → Y	X' in Y üzerindeki doğrudan etkisi	
X ↔ Y	X ve Y arasındaki kovaryans	
	Varyans	

Gözlenen Değişken

- **Gözlenen değişkenler** YEM dilinde göstergeler (indicators) olarak ifade edilir ve bunlar araştırmacının **doğrudan ölçtüğü ya da gözlediği değişkenleri ifade eder.**
    - YEM terminolojisinde gözlenen değişkenler, gizil değişkenleri yordamaz; aksine **gizil değişkenler kendi gözlenen değişkenlerini yordar.**

Gizil Değişken

- **Gizil değişkenler** YEM'in en önemli kavramlarından biridir ve araştırmacıların ilgilendikleri **zeka, güdü, duygu, tutum gibi soyut kavramlar ya da psikolojik yapılara** karşılık gelir.
  - Bu yapıları ancak **dolaylı olarak, belirli davranışlar ya da göstergeler temelinde ölçülen değişkenler yardımıyla gözlenebilir .**
  - Diğer bir ifade ile **gizil değişkenler doğrudan gözlenemedikleri** için bağlantılı oldukları **gözlenebilen başka değişkenler aracılığıyla** kestirilmeye çalışılır.

DFA

- **DFA**, daha önceden tanımlanmış ve sınırlandırılmış bir yapının, bir model olarak **doğrulanıp doğrulanmadığının** test edildiği bir analizdir.

- **DFA**, psikoloji alan yazınında daha çok **ölçek geliştirmede ve geçerlik analizlerinde** kullanılmaktadır.

  - Bu analizlerde, önceden belirlenmiş ya da kurgulanmış **bir yapının doğrulanması amaçlanmaktadır.**

DFA

- Araştırmacı, DFA'da kuramsal bilgilere dayalı olarak belirlediği
  - **gözlenen değişkenlerin gizil değişkenlerle** ve
  - **gizil değişkenlerin de kendi aralarında birbiri ile ilişkili** olduğunu kanıtlamaya çalışır.
- Araştırmacı, kurama dayalı olarak
  - **geliştirdiği modelin doğrulanıp doğrulanmadığını** ya da
  - **beklenen modelle gözlenen modelin ne ölçüde uyum gösterdiğini** belirlemeye çalışır.

DFA

- DFA, **gizil değişkenler arasındaki ilişkileri betimleyen (önerilen) model ile elde edilen (gözlenen) verinin ne oranda uyuştuğuna** ilişkin ayrıntılı istatistikler sunar.

- Geleneksel testlerin aksine, **tek bir anlamlılık değeri vermez.** Bu doğrultuda bulgular, verinin uygunluğuna göre ve ölçülen parametrelere ilişkin **çok sayıda istatistiksel ölçüt** kullanılarak değerlendirilir.

- DFA'da süreç, **korelasyon ya da kovaryans matrisi** oluşturarak başlar.

DFA

- Bir ölçme modelinin DFA sonuçlarında
  - **faktörler arasındaki korelasyon** kestirimleri,
- göstergelerin bağlı bulunduğu **faktörler altındaki yükler** ve
- **her bir gösterge için ölçme hatalarının** miktarını verir.

DFA

- Eğer araştırmacının başlangıçtaki ölçme modeli mantıklı bir biçimde doğrulanıyor ise dikkat edilmesi gereken durumlar şunlardır:
  - Ortak bir faktör altında ölçme yapmak için belirlenen göstergelerin tümünün, **o faktörde oldukça yüksek yüklere sahip olması** - **yakınsak geçerlik**
  - **Faktörler arasındaki korelasyon kestirimlerinin aşırı yüksek olmaması** (örneğin >0,85) - **ayırt edici geçerlik**

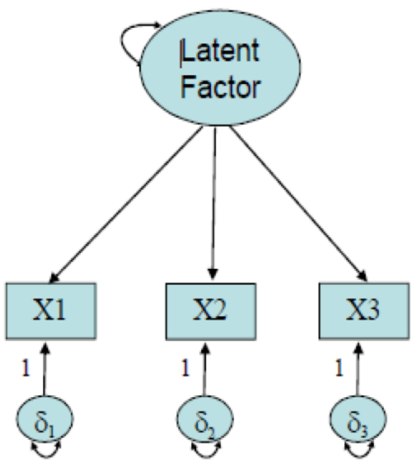
DFA vs AFA

- **Açımlayıcı analizlerde temel amaç**, yapısal bir modele ulaşmak ya da kuram üretmek olmasına karşın, **kurama ilişkin ilk ya da temel bilgilere ulaşılabilir.**

- Diğer taraftan **doğrulayıcı analizlerde**, daha önceki kapsamlı araştırmalardan elde edilen bilgi ya da tecrübeye dayanan durumlardan ve gözlemler çerçevesinde varsayımlar için model oluşturulur.

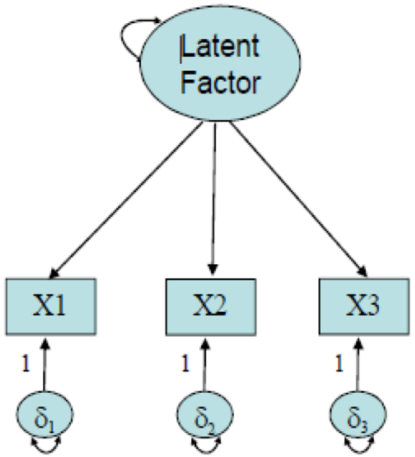
    - Bu varsayımlar temelinde **önceden kurulan modelin, bazı parametreler açısından doğruluğu test edilir.**

Tek Faktörlü DFA Modeli Örneği



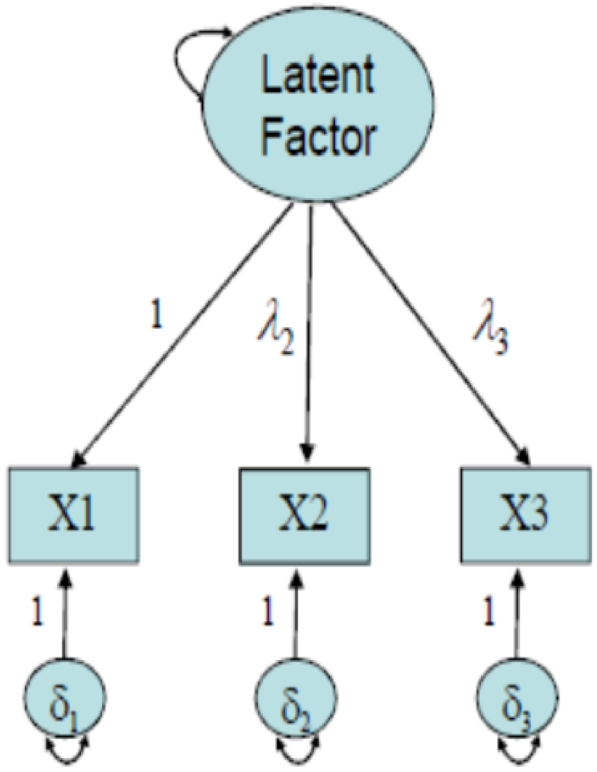
- X1, X2 ve X3 gizil faktörün göstergeleridir. Gizil faktör bu üç göstergenin altında yatan tek ortak yapıdır.
- Diğer bir ifadeyle, **göstergeler arasındaki kovaryansı üretecek gizil faktörden başka kaynaklar yoktur.**
- Her bir gözlenen puan iki bileşene sahiptir: **faktörün neden olduğu puan ve faktörün neden olmadığı puan.**
- Faktörün neden olmadığı kısım bir hata terimi ile gösterilir. Bu hata terimleri şemada δ_1, δ_2 ve δ_3 ile temsil edilmektedir.

Tek Faktörlü DFA Modeli Örneği



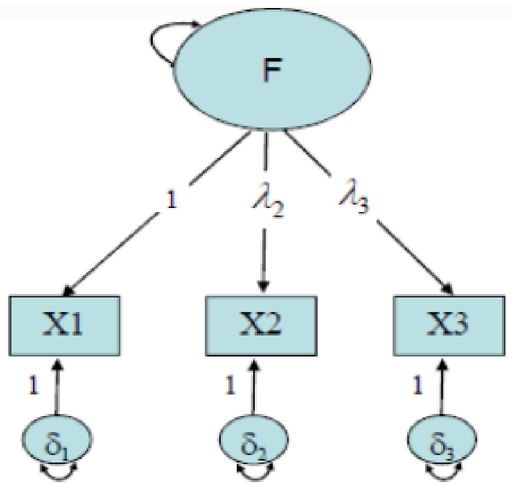
- δ_1 gibi bir hata terimi iki bileşenden oluşur:
  - Her bir göstergeye özel olan güvenilir bir puan
  - Ölçme hatasından dolayı güvenilir olmayan bir hata puanı
- **Kuramsal olarak bu iki bileşen ayırt edilemez.**
- DFA'da δ_1 gibi bir hata terimi ölçme hatası olarak adlandırılır ancak bu terim **iki bileşenin kombinasyonudur.**
- tesadüfi hataları **güvenilir olmayan değerler**, ölçülen özellik dışında ölçmeye karışan diğer özellikler **güvenilir olan değerler**

Tek Faktörlü DFA Modeli Örneği



- Faktör ve göstergeyi bağlayan yola faktör yükü adı verilir. Faktör yükü $\boldsymbol{\lambda}$ simgesi ile gösterilir.
- $\boldsymbol{\lambda_2}$, X2 göstergesinin gizil faktördeki yüküdür.
- $\boldsymbol{\lambda_3}$, X3 göstergesinin gizil faktördeki yüküdür.
- $\boldsymbol{\lambda_1}$, X1 göstergesinin gizil faktördeki yüküdür ancak **1** değerine sabitlenmiştir.

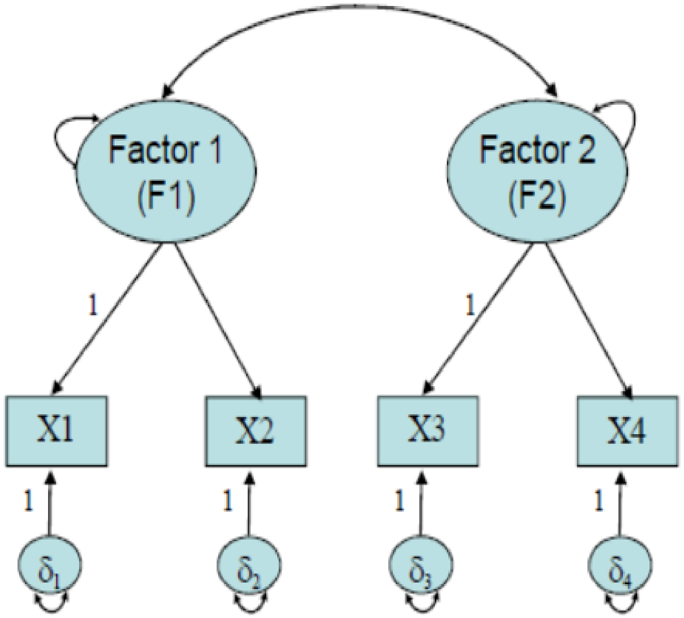
Tek Faktörlü DFA Modeli Örneği



$$x_1 = 1 \quad F + \delta_1$$
$$x_2 = \lambda_2 F + \delta_2$$
$$x_3 = \lambda_3 F + \delta_3$$

- Gizil değişken F, gizil olduğundan sabit bir ölçeğe sahip değildir. Bu nedenle, faktör yüklerinin ölçekleri belirlenemez. İki olası çözüm aşağıdaki gibidir:
  - **Faktör yüklerin ölçeği sabitlenir ve gizil değişkenin ölçeği serbestçe kestirilir.**
  - Gizil değişkenin ölçeği sabitlenir ve faktör yüklerinin ölçekleri serbestçe kestirilir.

İki Faktörlü DFA Modeli Örneği



- X1 ve X2 F1'in göstergeleridir.
- X3 ve X4 F2'nin göstergeleridir.
- F1 ve F2 arasında bir korelasyon vardır.
- X1, X2, X3 ve X4 göstergeleri etki göstergeleri (effect indicators) veya yansıtıcı göstergeler (reflective indicators) olarak adlandırılırlar.

İki Faktörlü DFA Modeli Örneği

$$X = \Lambda \xi + \delta$$

$$\begin{bmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \lambda_{21} & 0 \\ 0 & 1 \\ 0 & \lambda_{42} \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_1 \end{bmatrix} + \begin{bmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \\ \delta_4 \end{bmatrix}$$

İki Faktörlü DFA Modeli Örneği

- DFA modelinde 3 tip kovaryans matrisi tanımlanmalıdır:
    - $\sum$ (sigma): ölçülen gözlenen değişkenleri kovaryans matrisi
    - ϕ (phi): gizil faktörlerin kovaryans matrisi
    - Ψ (psi): ölçme hatalarının kovaryans matrisi

DFA Modeli: Kovaryans Matrisi

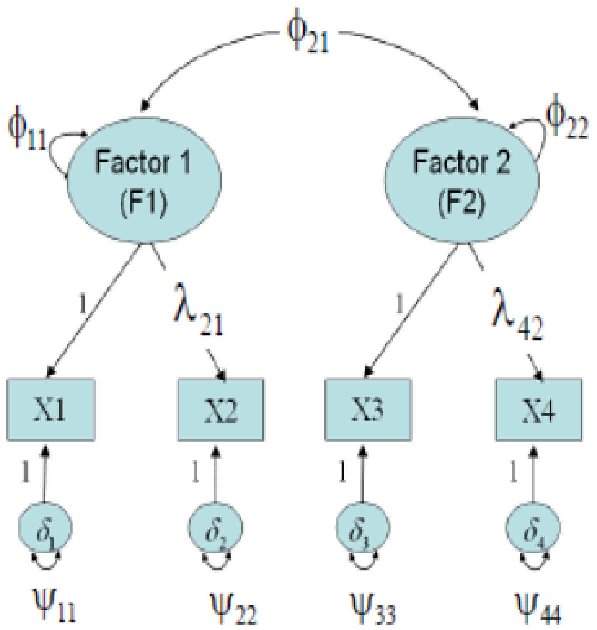
- Model için kovaryans matrisi model parametrelerinin bir fonksiyonu olarak temsil edilebilir:

$$\sum(\theta) = \Lambda\phi\Lambda' + \Psi$$

$$\begin{bmatrix} VAR_{X_1} & & & \\ COV_{X_1,X_2} & VAR_{X_2} & & \\ COV_{X_1,X_3} & COV_{X_2,X_3} & VAR_{X_3} & \\ COV_{X_1,X_4} & COV_{X_2,X_4} & COV_{X_3,X_4} & VAR_{X_4} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \lambda_{21} & 0 \\ 0 & 1 \\ 0 & \lambda_{42} \end{bmatrix} \begin{bmatrix} \phi_{11} & \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} 1 & \lambda_{21} & 0 & 0 \\ 0 & 0 & 1 & \lambda_{42} \end{bmatrix} + \begin{bmatrix} \Psi_{11} & & & \\ 0 & \Psi_{22} & & \\ 0 & 0 & \Psi_{33} & \\ 0 & 0 & 0 & \Psi_{44} \end{bmatrix}$$

$\sum(\theta)$, ϕ, ψ matrisleri simetrik matrislerdir ancak eşitlikte sadece alt üçgen sunulur.

DFA Modeli: Kovaryans Matrisi

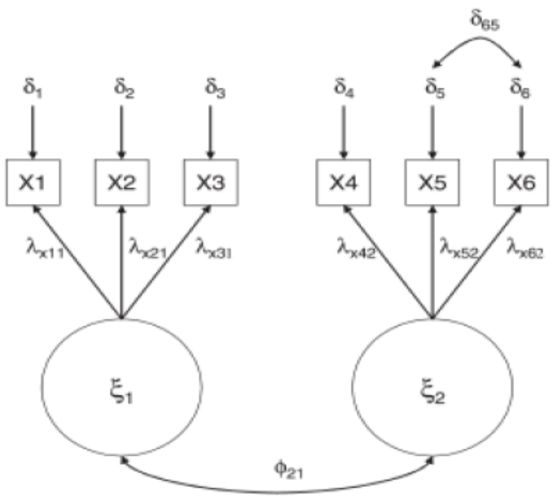


Kovaryans matrislerinden model için serbestçe kestirilecek parametreler sayılabilir:

Faktör varyansları	$\phi_{11}\phi_{22}$
Faktör kovaryansları	ϕ_{21}
Hata varyansları	$\Psi_{11}\Psi_{22}\Psi_{33}\Psi_{44}$
Faktör yükleri	$\lambda_{21}\lambda_{42}$

$$\begin{bmatrix} VAR_{X_1} & & & \\ COV_{X_1,X_2} & VAR_{X_2} & & \\ COV_{X_1,X_3} & COV_{X_2,X_3} & VAR_{X_3} & \\ COV_{X_1,X_4} & COV_{X_2,X_4} & COV_{X_3,X_4} & VAR_{X_4} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \lambda_{21} & 0 \\ 0 & 1 \\ 0 & \lambda_{42} \end{bmatrix} \begin{bmatrix} \phi_{11} \\ \phi_{21} + \phi_{22} \end{bmatrix} \begin{bmatrix} 1 & \lambda_{21} & 0 & 0 \\ 0 & 0 & 1 & \lambda_{42} \end{bmatrix} + \begin{bmatrix} \Psi_{11} & & & \\ 0 & \Psi_{22} & & \\ 0 & 0 & \Psi_{33} & \\ 0 & 0 & 0 & \Psi_{44} \end{bmatrix}$$

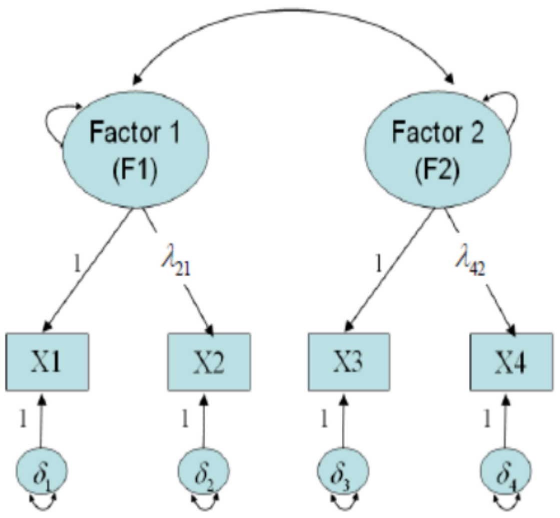
İki Faktörlü DFA Modeli Örneği



Name	Parameter	Matrix	Type	Description
Lambda-X	λ_X	Λ_X	Regression	Factor loadings
Theta delta	δ	Θ_δ	Variance–covariance	Error variances and covariances
Phi	ϕ	Φ	Variance–covariance	Factor variances and covariances
Tau-X	τ_X		Mean vector	Indicator intercepts
Kappa	κ		Mean vector	Latent means
Xi (Ksi)	ξ		Vector	Names of exogenous variables

FIGURE 3.3. Latent X notation for a two-factor CFA model with one error covariance. Factor variances, factor means, and indicator intercepts are not depicted in the path diagram.

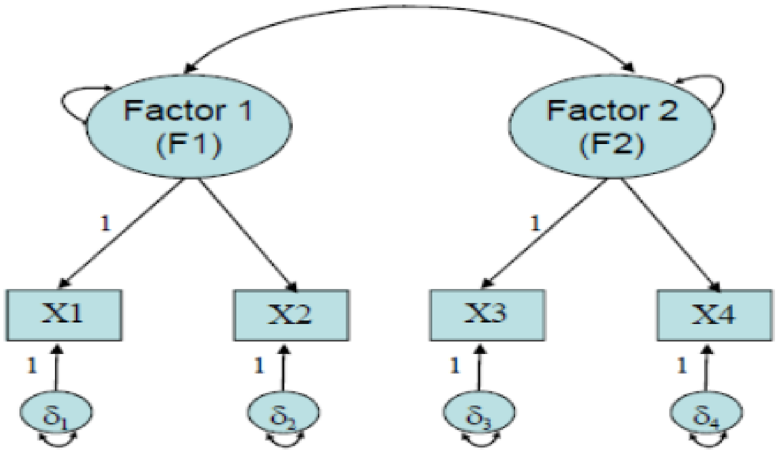
İki Faktörlü DFA Modeli Örneği



Bu nedenle,

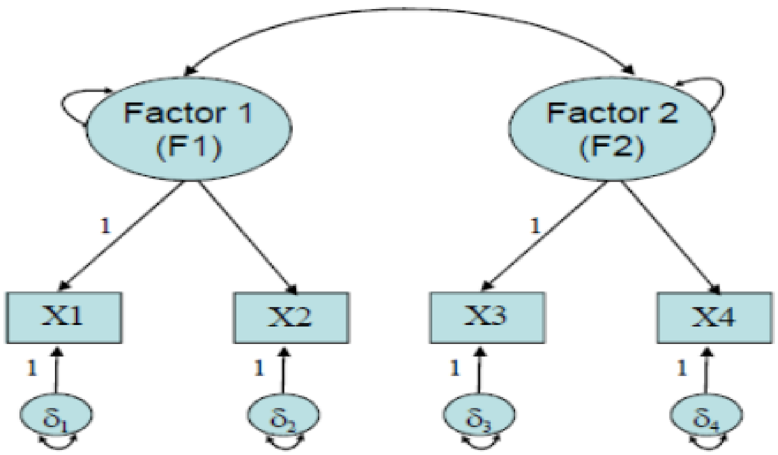
- λ_{21} X2'nin Faktör 1'deki yüküdür.
- λ_{42} X4'ün Faktör 2'deki yüküdür.
- λ_{11} ve λ_{32} şemada değerleri 1'e sabitlenmiş olarak gösterilmektedir

Standart DFA Modelleri



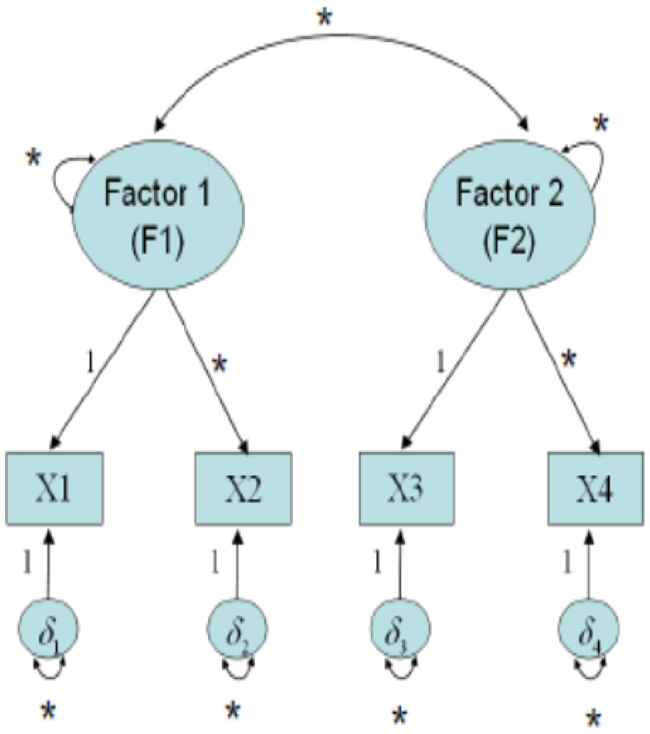
- **Standart DFA** modelleri aşağıdaki özelliklere sahiptir:
  - Her bir gösterge sürekli bir değişkendir ve tek bir faktörün (factor complexity = 1) ve ölçme hatasının bir fonksiyonudur.
  - Ölçme hataları birbirlerinden ve faktörlerden bağımsızdır.
  - Faktörlerin arasında korelasyon vardır.

DFA modellerinin parametreleri:



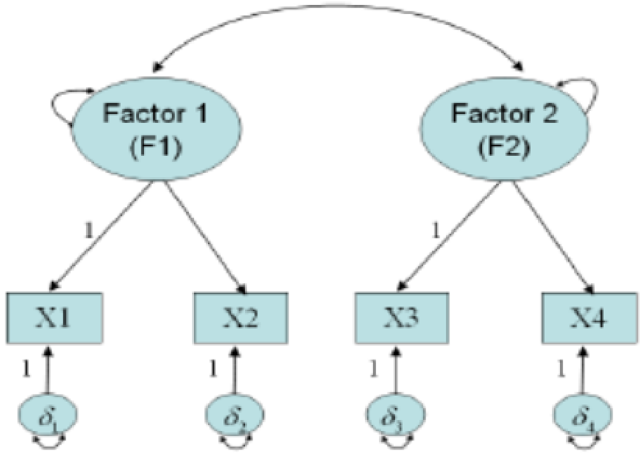
- Standart DFA modellerinde, **faktörler dışsal değişkenlerdir.**
- Dışsal değişkenlerin **varyansları ve kovaryanslarıdır.**
- Faktörlerin göstergeler üzerindeki **doğrudan etkileridir.** (örneğin, faktör yükleri)

DFA modellerinin parametreleri:



- Faktörlerin varyansları: 2
- Faktörlerin kovaryansları: 1
- Hataların varyansları: 4
- Faktör yükleri: 2 (diğer iki faktör yükü 1'e sabitlenmiştir)
- Modelden **serbestçe kestirilecek 9 parametre** vardır.

İki Faktörlü DFA Modeli



$$
\begin{aligned}
x_1 &= 1\,F_1 + 0\,F_2 + \delta_1 \\
x_2 &= \lambda_{21}F_1 + 0\,F_2 + \delta_2 \\
x_3 &= 0\,F_1 + 1\,F_2 + \delta_3 \\
x_4 &= 0\,F_1 + \lambda_{42}F_2 + \delta_4
\end{aligned}
$$

- DFA literatüründe bir gizil faktörü temsil etmek üzere genellikle F yerine ξ kullanılır.

İki Faktörlü DFA Modeli

- Faktör yükleri 1'e

$$\begin{bmatrix} var(x_1) & & & \\ cov(x_2x_1) & var(x_2) & & \\ cov(x_3x_1) & cov(x_3x_2) & var(x_3) & \\ cov(x_4x_1) & cov(x_4x_2) & cov(x_4x_3) & var(x_4) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \lambda_{21} & 0 \\ 0 & 1 \\ 0 & \lambda_{42} \end{bmatrix} \begin{bmatrix} \phi_{11} & \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} 1 & \lambda_{21} & 0 & 0 \\ 0 & 0 & 1 & \lambda_{42} \end{bmatrix} + \begin{bmatrix} \psi_{11} & & & \\ 0 & \psi_{22} & & \\ 0 & 0 & \psi_{33} & \\ 0 & 0 & 0 & \psi_{44} \end{bmatrix}$$

- Faktör varyansları 1'e

$$\begin{bmatrix} var(x_1) & & & \\ cov(x_2x_1) & var(x_2) & & \\ cov(x_3x_1) & cov(x_3x_2) & var(x_3) & \\ cov(x_4x_1) & cov(x_4x_2) & cov(x_4x_3) & var(x_4) \end{bmatrix} = \begin{bmatrix} \lambda_{11} & 0 \\ \lambda_{21} & 0 \\ 0 & \lambda_{32} \\ 0 & \lambda_{42} \end{bmatrix} \begin{bmatrix} 1 & \\ \phi_{21} & 1 \end{bmatrix} \begin{bmatrix} \lambda_{11} & \lambda_{21} & 0 & 0 \\ 0 & 0 & \lambda_{32} & \lambda_{42} \end{bmatrix} + \begin{bmatrix} \psi_{11} & & & \\ 0 & \psi_{22} & & \\ 0 & 0 & \psi_{33} & \\ 0 & 0 & 0 & \psi_{44} \end{bmatrix}$$

- Her ikisinde de 9 tane serbestçe kestirilecek parametre bulunmaktadır.

Modelin Tanımlanması

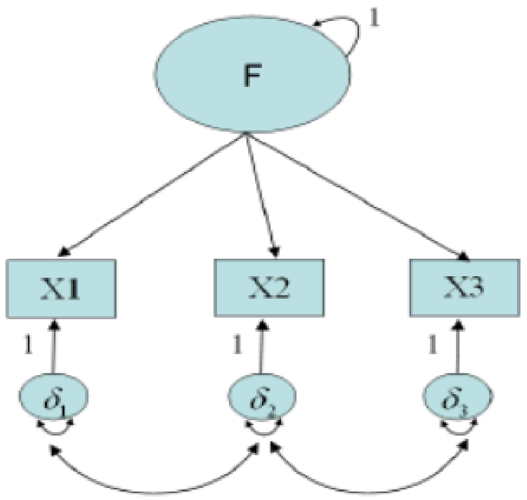
- DFA modelinde gözlemlerin sayısı yol analizi modelindeki gibi hesaplanır:
  - $v(v+1)/2$
- İki faktörlü DFA modeli örneğinde gözlemlerin sayısı:
  - $4(4+1)/2 = 10$ (v = 4, gözlenen değişken sayısı)
- Yönlü ilişkili (recursive) yol modellerinde , eğer gözlemlerin sayısı model parametrelerinin sayısına eşit veya daha büyükse **sd ≥ 0** model tanımlanır. Ancak DFA'da, **sd ≥ 0** koşulunun sağlanması zorunlu fakat yeterli değildir.

Modelin Tanımlanması

- dört gözlenen değişkenli DFA modeli örneğinde gözlemlerin sayısı:

$$\begin{bmatrix} VAR_{X_1} & & & \\ COV_{X_1,X_2} & VAR_{X_2} & & \\ COV_{X_1,X_3} & COV_{X_2,X_3} & VAR_{X_3} & \\ COV_{X_1,X_4} & COV_{X_2,X_4} & COV_{X_3,X_4} & VAR_{X_4} \end{bmatrix}$$

Modifikasyon



- DFA modelleri hatalar arasında korelasyonun tanımlanmasına izin verir.
- Araştırmacılar bir ölçüm için hataların arasında korelasyon olmaması sayıltısının olağan olarak ihlal edildiğini tartışırlar (Schmidt & Hunter, 1996).
- Ancak hatalar arasındaki korelasyonun modele eklenmesi **tanımlama problemlerine neden olabilir.**

Yeterli Gereklilikler

- Zorunlu gerekliliklerin karşılanması DFA modelinin tanımlanmasını garantilemez. Yeterli gereklilikler.

- Eğer

  - **standart bir tek faktörlü DFA modeli en az üç göstergeye** sahipse,

  - **standart bir iki veya daha fazla faktörlü** DFA modelinde her bir faktör için **en az iki göstergeye** varsa,

model tanımlanır.

- **Çapraz yüklü veya hatalar arasında korelasyona** sahip bir DFA modeli için ise kolayca uygulanan yeterli koşul yoktur.

Modelin Tanımlanması

- Model parametresi araştırmacının tanımlamasına bağlı olarak **serbest** (free), **sabit** (fixed) veya **sınırlandırılmış** (constrained) olabilir.
    - **Serbest parametre** (free parameter) örneklem verisinden bilgisayar yazılımı tarafından kestirilen parametredir.
    - **Sabit** parametre (fixed parameter) bir sabite eşit olarak belirlenen parametredir; yazılım bu sabiti veriye bağlı olmaksızın parametrenin kestirimi olarak kabul eder.
    - **Sınırlandırılmış** parametre (constrained parameter) yazılım tarafından belli sınırlılıklar içerisinde kestirilir ancak bir sabite eşit olmak üzere sabitlenmez.

Modelin Tanımlanması

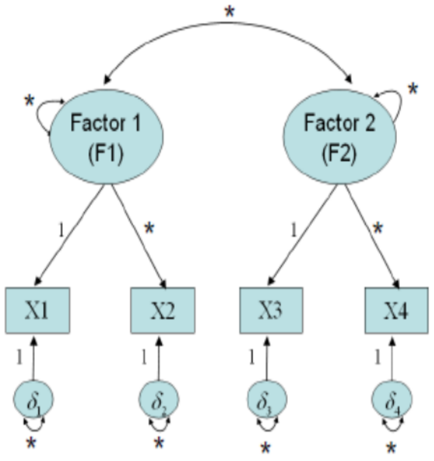
- Bir modelinin tanımlanabilmesi için **kestirilecek olan model parametre sayısının gözlenen parametre sayısına eşit** veya **gözlenen parametre sayısından küçük olması** gerekir.

  - **sd<0** olduğunda model tanımlanamaz.

  - **sd=0**, model ancak tanımlanır (just identification) ve kuramsal olarak her parametrenin tek bir çözümü vardır. Ancak tanımlanan modellerde model veriye mükemmel uyum gösterir.

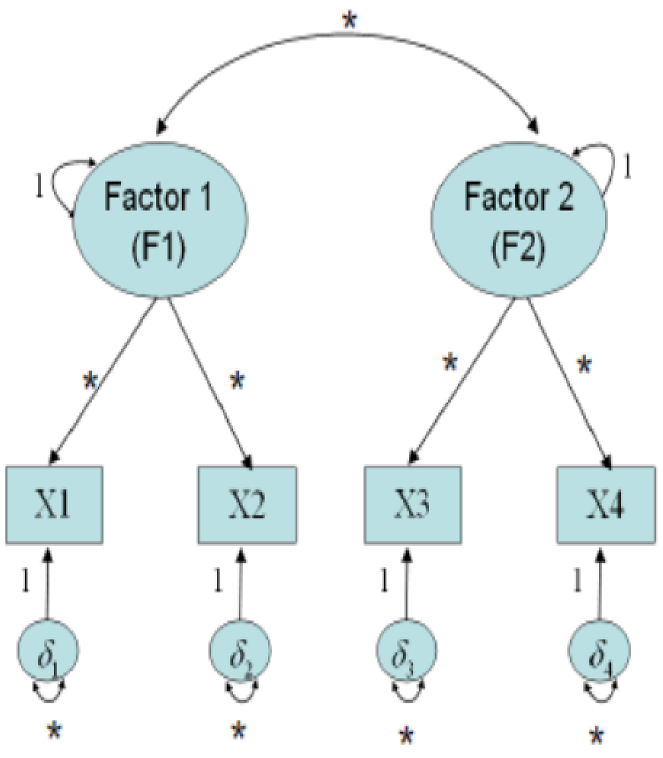
  - **sd>0**, model aşırı tanımlanmış (over identification) olur. Aşırı tanımlanan modellerde kuramsal olarak her bir parametrenin birden fazla çözümü vardır.

Model Tanımlanması



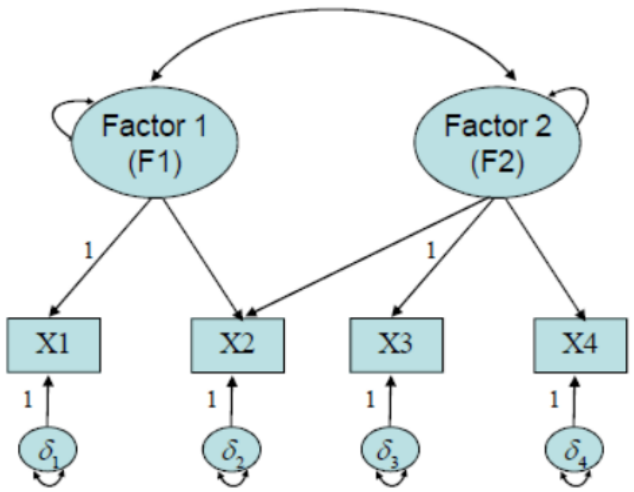
- X1'in faktör yükü **1**'e sabitlendiğinde, F1'in ölçme birimi X1'in ölçme birimiyle aynı olur. Bu durumda X1 **işaret veya referans değişken** adını alır.
- Benzer şekilde, X3 de F2 için **referans değişkendir.**
- Referans değişken birinci gösterge olmak zorunda değildir. Örneğin, X2 de F1 için referans değişken olarak seçilebilir.
  - Referans değişken belirlendiğinde, **değişkenin örneklem varyansının bir kismi gizil değişkene geçer.**

Model Tanımlanması



- hem F1 hem de F2 puanları **ortalaması 0 ve standart sapması 1** olan bir ölçeğe sahip olurlar.
- bütün faktör yükleri serbestçe kestirilir.
- Model parametreleri:
  - Faktörlerin varyansları: 0
    - *1'e sabitlenmişlerdir*
  - Faktörlerin kovaryansları: 1
  - Hataların varyansları: 4
  - Faktör yükleri: 4

Çapraz Yüklü DFA Modelleri



- Verilen örnekte X2 hem F1 hem de F2'nin göstergesidir.

- Bazen bazı göstergeler birden fazla faktörü ölçmek için tasarlanmıştır (factor complexity > 1).

- Böyle bir DFA modeli **artık standart bir DFA modeli** değildir.

Model-Veri Uyumunun Değerlendirilmesi

- Kestirilen parametre sayısından daha fazla sayıda gözleme sahip olan **aşırı tanımlanan** (overidentified) modeller **genellikle veriye mükemmel uyum sağlamaz.**

- Bu durumda böyle **modellerin veriyle ne derece uyumlu** olduğunu ölçmeye ihtiyaç vardır.

- YEM literatüründe tanımlanan **çok sayıda model uyum indeksi vardır** ve sürekli olarak yeni indeksler geliştirilmektedir.

Model-Veri Uyumunun Değerlendirilmesi

- Çok sayıda farklı uyum indeksinin olması bazı problemleri de beraberinde getirir:
  - Farklı makalelerde **farklı uyum indeksleri** rapor edilir.
  - Aynı makale için farklı hakemler kendi bildikleri veya tercih ettikleri farklı indekslerin rapor edilmesini isteyebilirler.
  - Uyum indekslerinin değerlerini rapor ederken seçici davranma olasılığı vardır (örneğin, sadece iyi uyum öneren uyum indekslerinin rapor edilmesi gibi).

Model-Veri Uyumunun Değerlendirilmesi

- YEM uygulamalarına ve simülasyon çalışmalarına göre YEM analizinin sonuçlarını rapor ederken sunulacak ve yorumlanacak uyum indeksleri aşağıdaki gibidir:

- Model **Ki-Kare** Değeri

- Steiger-Lind Root Mean Square Error of Approximation **RMSEA** (Steiger, 1990) (%90 güven aralığı ile birlikte)

- Bentler Comparative Fit Index **CFI** (Bentler, 1990)

- Standardized Root Mean Square Residual **SRMR**

- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107,* 238-246.
- Steiger, J. H. (1990). Structural model evaluation and modification: An interval estimation approach. *Multivariate Behavioral Research, 25,* 173-80.

Model-Veri Uyumunun Değerlendirilmesi

- Uyum indekslerinin değerleri bir modelin sadece ortalama veya genel uyumunu belirtir. Bu nedenle **belli bir indeksin değeri uygun bile görünse, modelin belli kısımları veriye zayıf uyum sağlayabilir.**

- Uyum indeksleri sonuçların **kuramsal olarak anlamlı olup olmadığını belirtmezler.**

- Yeterli uyumu öneren uyum indekslerinin değerleri **yordayıcıların yordama güçlerinin de yüksek olduğunu belirtmezler.**

- Örneğin, **veriye mükemmel uyum sağlayan modellerin bozukluklarının varyansı halen yüksek olabilir.**

Model-Veri Uyumunun Değerlendirilmesi

- **Tek bir indeks** modelin sadece belli bir yönünü yansıttığından, modelin iyi uyum sağladığını belirtmek için **tek başına yeterli olmaz.** Bu nedenle, model uyumu birden fazla indeksin değerine dayanarak değerlendirilir.

- Uyum hem modelin belli kısımlarında **bölgesel** olarak hem de **genel model ve veri uyuşmasının** ne kadar iyi olduğu yönünde global olarak değerlendirilmelidir.

- Genel olarak YEM analizinde model uyumu değerlendirilirken, odak tek bir istatistiksel anlamlılık testinde değildir. **Çeşitli indeksleri incelerken bütüncül bir yaklaşım kullanılmalıdır.**

- **Çoklu indekslerin kullanılması** bir modelin uyumu ile ilgili **en doğru yaklaşımı verecektir.**

Ki-Kare Testi (Chi-Square Test)

- Ki-kare testi gözlenen kovaryans matrisinin tanımlanan modelle tutarlı olup olmadığını değerlendirir.

- Model **ki-kare değeri arttıkça**, modelin veriye uyumu kötüleştiği için model ki-kare aslında bir **kötülük uyum** indeksidir.

Ki-Kare Testi ve Örneklem Büyüklüğü

- **Ki-kare testi örneklem büyüklüğüne** bağlılığından dolayı iyilik uyumunun değerlendirilmesi için **çok ideal değildir.** Ancak geleneksel olarak rapor edilir ve diğer uyum indeksleriyle desteklenir.

- Model ki-kare değerinin örneklem büyüklüğüne hassasiyetini azaltmak için bazı araştırmacılar bu değeri ilgili serbestlik derecesine bölerler. Elde edilen değer normed chi-square (NC) değeri olarak adlandırılır.

- Ancak bu değerin yorumlanması için minimum kabul edilebilirlik düzeyini temsil edecek net bir kesim değeri yoktur.

  - **NC <= 2 ve ya 3 ve ya 5 (Kabul edilir.)**

- Ayrıca **NC örneklem büyüklüğünün etkisini tamamen düzeltmez.**

RMSEA

- Root Mean Square Error of Approximation (RMSEA) **serbestlik derecesinin bir fonksiyonu olarak uyumu değerlendiren bir indekstir:**
- RMSEA indeksi de **kötülük uyum** indeksidir.
- RMSEA indeksinin **daha yüksek değerleri daha kötü uyumu belirtir.**
- RMSEA = 0 değeri **en iyi uyumu belirtir**.
  - Ancak RMSEA = 0 değeri **mükemmel bir uyumu ifade etmez.**

RMSEA

- RMSEA uyumu doğrudan serbestlik derecesinin bir fonksiyonu olarak ele alır; modelin tutumunu hesaba katar (ölçülen değişkenlerin sayısına karşılık kestirilen model parametrelerinin sayısı).

- RMSEA için önerilen kesme noktaları (Hu & Bentler, 1999):

  - RMSEA ≤ 0.05 iyi uyumu belirtir.

  - 0.05 < RMSEA < 0.08 kabul edilebilir uyumu belirtir.

  - RMSEA ≥ 0.08 zayıf uyumu belirtir.

Bentler, P. M. & Hu, L. (1999). Cutoff criteria fpr fit indexes in covariance structure analysis: Conventional criteri versus new alternatives. Structural Equation Modeling, 6(1), 1-55.

RMSEA

- RMSEA tarafından kestirilen evren parametresi için %90 güven aralığı genellikle YEM yazılımlarının outputunda verilir.
- Parametre için güven aralığı kestirilen merkezi olmayan δ parametresine dayanır ve RMSEA örneklem değeri etrafında simetrik olmayabilir.
- Bu güven aralığı nokta kestirimi olarak RMSEA değeri ile ilişkili belirsizlik derecesini yansıtır.
- Eğer Parametre için %90 güven aralığının alt sınırının değeri 0,05'ten küçükse, modelinin evrende tahmini yaklaşık uyuma sahip olduğu hipotezi $H_0 : \epsilon_0 \leq 0.05$ reddedilmeyecektir.

SRMR

- **Standardized Root Mean Square Residual (SRMR)** Bu indeks RMR indeksinin hesaplandığı şekilde hesaplanır ancak standartlaştırılmış artıklar kullanılır.
  - 0.08'den küçük değerler uygun olarak düşünülür (Hu & Bentler, 1999).

Karşılaştırmalı Uyum İndeksleri

(Comparative Fit Indices)

- Bir çok indeks araştırmacının modelinin veriye nasıl uyduğunu, modelin uyumunu daha sınırlandırılmış bir modelle karşılaştırarak değerlendirir.
  - Bu nedenle karşılaştırmalı uyum indeksleri **artımlı uyum indeksleri** (incremental fit indices) olarak da bilinir: daha sınırlandırılmış model (örneğin, sıfır modeli) uyumundan daha esnek model (örneğin, araştırmacının modeli) uyumuna artırım.
- Karşılaştırmalı uyum indekslerinden YEM analizlerinde sık kullanılan iki tanesi **CFI ve NNFI (TLI)** indeksleridir. Ancak iki indeks de örneklem dayanaklı indekslerdir.

Karşılaştırmalı Uyum İndeksleri

(Comparative Fit Indices)

- CFI indeksi Bentler (1990) tarafından geliştirilmiştir
- CFI = 0 değeri araştırmacının modelinin sıfır modeline göre gelişmediğini belirtir.
- CFI değerinin 0.90 veya 0.95'ten daha büyük olması kabul edilebilir uyum için önerilir (Hu & Bentler, 1999).
- CFI = 1 değeri **mükemmel uyumu belirtmez.**

Uyum İndekslerini Raporlarken Öneriler

- Tek bir indeks model uyumunun sadece belli bir yönünü yansıtır. Araştırmacılar aşağıdakilerin rapor edilmesini önerir:
- Model ki-kare değeri: anlamlı olmayan sonuç

Uyum indeksi	İyi uyum	Kabul Edilebilir Uyum
χ^2	p	
χ^2/df	$0 \leq \chi^2/df \leq 2$	$2 < \chi^2/df < 8$
SRMR	0 ≤SRMR ≤ .05	.05 < SRMR < .10
RMSEA	0 ≤RMSEA ≤ .05	.05 < RMSEA < .08
CFI	.95 ≤GFI ≤ 1	.90 < GFI < .95

Not: Bu kesme değerlerin kullanılmasıyla ilgili çok sayıda tartışma vardır.

Marsh, Hau, & Wen, 2004: The Special Issue in Personality and Individual Differences, 2007

Uyum İndekslerini Raporlarken Öneriler

- örneğin, CFI, örneklem büyüklüğüne duyarlıdır.

- Oldukça küçük bir örneklem (örneğin, N = 200) hemen hemen her zaman oldukça büyük bir örneklemden (örneğin, N = 1000) daha küçük **CFI**'ler üretecektir.

- Bu nedenle, .95 gibi CFI için sabit eşikler, küçük örneklem boyutları için daha doğru modelleri uyumlu göstermeyecek ve daha büyük örneklem boyutları için ise muhtemelen daha az uyumlu modelleri kabul edecektir.

- Esnek kesim noktaları belirli bir model ve örneklem özelliklerine uygun alternatif eşik değerleri sağlar. Bu amaçla *FCO* paketi kullanılabilir.

```
library(FCO)
```

Esnek Kesim Noktaları

- Esnek kesim noktaları, gizli değişken sayısı (veya faktörler), gizil değişken başına gösterge sayısı, örneklem boyutları, faktör yükleri ve normal ve normal olmayan veriler için doğru şekilde belirlenmiş Doğrulayıcı Faktör Analizi (DFA) modellerinin simüle edilmiş dağılımlarından türetilir.

- Esnek kesim noktaları, önceden tanımlanmış bir belirsizlik için belirli bir değerin ampirik niceliği olarak anlaşılabilir.

- Öncelikli olarak yüzde 5'lik (veya .05) bir belirsizlik kabul edilirse, verilen model ve örneklem özellikleriyle doğru şekilde belirlenmiş DFA modelleri için simüle edilmiş dağılımın yüzde 5'lik niceliği esnek sınırı belirleyecektir.

Niemand, T., Mai, R. Flexible cutoff values for fit indices in the evaluation of structural equation models. J. of the Acad. Mark. Sci. 46, 1148–1172 (2018)

Model Seçimi

- analizlerinde **birden fazla model veriye uyum sağlayabilir.**
- Bu durumda veriye daha iyi uyum sağlayan modelin belirlenmesi için **model uyumları karşılaştırılır.**
- **Modellerden biri diğerinin alt kümesi** olduğunda modeller yuvalanmıştır. Örneğin, araştırmacılar farklı amaçlarla model içinde bazı parametreleri sınırlandırabilir. Bu durumda sınırlandırılmış olan model sınırlandırılmamış olan serbest model içinde yuvalanmış olur.
- Yuvalanmış modeller karşılaştırılırken, **modellerin ki-kare değerleri arasındaki farkın anlamlılığı incelenir.**

Model Seçimi

- Eğer iki model hiyerarşikse ve her ikisi de veriye kabul edilebilir ölçüde uyum sağlıyorsa, iki modelin veriye uyumunu karşılaştırmak için ki-kare fark test uygulanabilir.

- Yuvalanmamış modellerin karşılaştırılmasında kullanılacak en uygun yaklaşım bilgi kriter değerleridir. Yaygın olarak kullanılan bilgi kriterleri **Akaike bilgi kriteri (AIC) ve Bayes bilgi kriteridir (BIC).**

- Model seçimi yapılırken **bilgi kriteri değeri daha küçük olan model tercih edilir.** Bilgi kriterleri hem yuvalanmış hem de yuvalanmamış modellerin karşılaştırılmasında kullanılabilir.

Varsayımlar

- YEM analizleri genellikle **büyük örneklemler gerektirir.**
- Örneklem büyüklüğü model karmaşıklığı ile oldukça ilişkidir ve daha karmaşık modeller daha büyük örneklemler gerektirir.
- YEM analizlerine ilişkin varsayımlar ise kestirim yöntemine bağlı olarak değişmektedir.
- Kullanılan kestirim yönteminin varsayımlarının sağlanmaması model veri uyumu, parametre ve parametrelere ilişkin hata kestirimlerinde yanlılığa neden olabilir. Dolayısıyla, test edilen kuram hakkında hatalı sonuçlar alınmasına yol açabilir.
- Bu nedenle veri yapısına ve çalışma desenine uygun bir kestirim yöntemi seçilmesi oldukça önemlidir.

Varsayımlar

- YEM analizlerinde yaygın olarak maksimum olabilirlik **ML** kestirim yöntemi kullanılmaktadır.

- Benzer şekilde en küçük kareler **GLS** kestirim yöntemi de normallik varsayımı gerektirir.

- Bu yöntemler normal teori yöntemleri olarak da adlandırılır.

- Normal teori yöntemleri altında yatan beş varsayımdan bahsedilebilir. Bunlar

  - gözlemlerin bağımsızlığı,
  - büyük örneklem, doğru tanımlanmış model,
  - çok değişkenli normallik ve verilerin sürekliliğidir.

Normal teori yöntemleri varsayımları sağlandığında yansız, yeterli ve tutarlı kestirimler üretir.

Varsayımlar

- doğrulayıcı faktör analizi için **cfa()**, yapısal modeller için **sem()** fonksiyonu tanımlanmıştır.

- ML yöntemi paketin varsayın kestirim yöntemidir. Parametre kestirimleri açısından normallik varsayımının ihlaline dayanıklıdır.

- Normal olmayan veriler için alan yazında yaygın olarak kullanılan kestirim yöntemi

  - ağırlıklandırılmış en küçük kareler yöntemidir **WLS**
  - Ayrıca ağırlıklandırılmamış en küçük kareler **ULS**,
  - diyagonal olarak ağırlıklandırılmış en küçük kareler **DWLS**,
  - kestirim yöntemleri de normal dağılmayan verilerde kullanılabilir.

- Kestirim yöntemlerinin kısaltmaları programlar arasında farklılık gösterebilir.

cfa fonksiyonu

Argüman	Açıklama	Değerleri
Model	YEM modeli tanımlanır.	
Data	Gözlenen değişkenlerin yer aldığı veri setidir.	
sampling.weights	Örneklem ağırlıklandırması yapılacağı durumlarda tanımlanır.	Veri çerçevesinde ağırlıklandırma bilgisinin yer aldığı değişkenin adıdır.
group	Çoklu grup analizlerde grup değişkeni tanımlanır.	Veri matrisinde grubu tanımlayan değişkenin adıdır.
cluster	Çok düzeyli analizlerde düzey değişkeni tanımlanır.	Veri matrisinde düzeyi tanımlayan değişkenin adıdır.
constraints	Modele eklenecek diğer sınırlandırmalar tanımlanır.	
estimator	Kestirim yöntemidir.	“ML”, “GLS”, “WLS”, "ULS”, "DWLS” gibi

cfa()

- **cfa()** fonksiyonunun kullanımı aşağıdaki gibidir:

```
cfa(model = NULL,
    data = NULL,
    ordered = NULL, sampling.weights = NULL,
    sample.cov = NULL, sample.mean = NULL, sample.th = NULL,
    sample.nobs = NULL, group = NULL, cluster = NULL,
    constraints = "", WLS.V = NULL, NACOV = NULL, ...)
```

cfa()

- **cfa()** fonksiyonu hem ham veriyle hem de gözlenen değişkenlere ilişkin varyans-kovaryans matrisiyle çalışabilir.

- Girdi olarak **varyans-kovaryans matrisi kullanıldığında,** matriste madde adlarına karşılık gelecek şekilde satır ve sütun adları mutlaka bulunmalıdır. Ayrıca **örneklem ortalamasının (sample.mean) ve gözlem sayısının tanımlanması** (sample.nobs) gerekir.

- Girdi olarak doğrudan **verinin kendisi girildiğindeyse**, örneklem ortalaması ve gözlem sayısının tanımlanmasına gerek yoktur.

- Tüm YEM fonksiyonlarında olduğu gibi **cfa()** fonksiyonu için de **ilk olarak DFA modelinin tanımlanması gerekir.**

cfa()

- DFA modeli tırnak işareti içinde tanımlanır ve faktörler maddeler ile **'=~'** işaretiyle ilişkilendirilir.

- İşaretin sol tarafında faktörler sağ tarafında maddeler yer alır. Her bir faktörde yer alan maddelerin adları sırasıyla **+** operatörüyle eklenir.

```
' F1  =~ m1 + m2 +m3'
```

bitti