DÖRT FARKLI ALFA KATSAYISININ
PSİKOMETRİK AÇIDAN KARŞILAŞTIRILMASI

THE COMPARISON OF FOUR DIFFERENT COEFFICIENT ALPHAS
FROM A PSYCHOMETRIC POINT OF VIEW

Halil YURDUGUL *

ÖZET: Bu araştırmada ölçme araçlarının işaretliliği belirlenmek için kullanılan dört farklı alfa katsayısının psikometrik özellikleri karşılaştırılmıştır. Bu katsayılara sırasıyla; a) ölçmelerin varyans-kovaryans terimlerinden elde edilen standartlaştırılmış alfa (αvK) katsayısı, b) ölçmeler arasındaki korelasyon terimlerinden elde edilen standartlaştırılmış alfa (αSTD) katsayısı, c) temel bileşenler analizinden elde edilen en büyük özdeğer üzerine kurulu olan ve Armor (1974) tarafından önerilen önerilenen alfa (αo) katsayısı ve d) doğruluğu faktör analizinden elde edilen faktör yükleri üzerine kurulu olan ve McDonald (1985) tarafından tanımlanan αm katsayısının özellikleri Monte Carlo simulasyonu ile karşılaştırılmıştır. Karşılaştırmalar 3 aşamada gerçekleştirilmiştir. Bunlar sırasıyla; a) bu katsayılara örneklem karşılaştırılsın hata kareler ortalaması karekökü (RMSE) değerlerine göre sağlam kestirciliği özellikleri, b) tekbeytül ve çokboyutlu veri kümiyendirgeleri değerleri ve c) paralel, özdeğer ve konjenerik veri kümelendeki davranışları karşılaştırılmıştır. Araştırma sonuçunda; αm katsayısı her üç aşamada da diğer katsayılara göre üstün özelliklere sahiptir. Yine her üç aşamada en olumsuz özellik ise αm katsayısında gözlenmiştir.

Anahtar sözcükler: güvenilirlik, işaretlilik, güvenilirlik katsayısı, alfa katsayısı.

ABSTRACT: In this study, the properties of four different coefficient alphas in reliability were compared. Those were a) unstandardized alpha (αvK) based on variance-covariance matrix, b) standardized alpha (αvK) based on correlation matrix, c) maximum alpha (αM) based on first eigenvalues by Armor (1974) and d) αM defined by McDonald (1985) and based on confirmatory factor analysis. The comparisons made to determine the best coefficient conducted on Monte Carlo simulations and received priority consideration were carried out in three stages: a) investigation of robust estimator properties according to root mean square error (RMSE) for four coefficient alphas, b) comparison of coefficient alphas’ values in unidimensional data sets and multidimensional data sets, and c) comparison of coefficient alphas’ behaviors in parallel, tau-equivalent, and congeneric data sets. As a result, the coefficient αM is found to be the optimum coefficient among those three stages. The coefficient αM is found to be non-optimum in all the stages.

Keywords: reliability, internal consistency, reliability coefficient, coefficient alpha.

1. GİRİŞ

Eğitim araştırmalarında ölçme araçlarının güvenilirliği önemli konulardan birisidir. Özellikle çoklu derecelendirilmiş (polytomous) ölçme araçlarının işaretliliğini kestirminde Guttman (1945) ve Cronbach (1951) tarafından geliştirilen alfa katsayısının yaygın bir şekilde kullanıldığı görülmektedir. Alfa katsayısının, tanımlanmış bazı olusuzlukların (ileride açıklanmıştır) karşılıkгресnasındaki kolaylık ve paket programlarda yer almaya nedeniyle halen popülerliğini korumaktadır.

Diğer taraftan ilgili literatürde alfa katsayısının farklı versiyonları tanımlanmıştır. Bunlardan yaygın olarak ele alınanlardan dört tane; a) varyans-kovaryans matrisi üzerinden elde edilen standartlaştırılmış alfa katsayısı (αvK), b) korelasyon matrisi üzerinden elde edilen standartlaştırılmış alfa katsayısı (αSTD), c) Armor (1974) tarafından tanımlanan ve temel bileşenler analizi terimleriley açıklanan maksimum alfa katsayısı (αM) ve d) McDonald (1985) tarafından tanımlanan ve doğrulayıcı faktör analizi terimleriley açıklanan alfa katsayısıdır (αM).

Bu çalışmada yukarıda tanımlanan dört farklı alfa katsayısının karşılaştırılması yapılarak optimallıkları arastırılmıştır. Bu amaçla yönelik olarak üç karşılaştırma ölçütü belirlenmiştir. Bunlardan ilki; alfa katsaylarının örneklem karşılaştırısının sağladığı kestirciliği (robust estimator) özelliğinin artırılması, ikincisi bu katsayılardan farklı boyutlardaki veri kümelendenden elde edilen değerlerinin

* Dr., Hacettepe Üniversitesi, Eğitim Fakültesi, Bilgisayar ve Öğretim Teknolojileri Eğitimi Bölümü, yurdugul@hacettepe.edu.tr

1.1. Güvenilirlik ve Alfa Katsayları

Güvenilirlik ölçümlerin hatasızlığa ilişkin bir kavramdır ve güvenilirliğin nicel değer; ölçülme istenilen gerçek puanlar varyansının [Var(T)], gözlenecek toplam puanlar varyansına [Var(X)] oranını ele eder diilir. "Gercek güvenilirlik" (true reliability) olarak adlandırılan bu güvenilirlik indeksi gerçek puanlar doğrudan elde edilemediğinden dolayı istatistiksel yöntemler ile kestirilir ve kestirim değeri ise güvenilirlik katsayısı olarak adlandırılır. Ölçmelerin içtutarısalılgı anlamında güvenilirliğe yönelik ilk çalışma Spearman (1904) ile başlamıştır. Ancak Spearman (1904) güvenilirliğin kestirimini paralel (ortalama ve varyansları eşit) ölçmeler üzerinde tanımlamıştır. Guttman (1945) ise davranış bilimlerinde paralel ölçmeleri elde etmenin uygulanıla olarak oladelphia ifade etmiş ve Spearman'ın önerdiği katsayıya alternatif olarak 6 adet güvenilirlik katsayısı önermiştir. Guttman (1945), bu katsayılar her zaman gerçek güvenilirliği değil, gerçek güvenilirliğin alt sınırı vereceği ifade ederek bu katsayıları güvenilirlik katsayısı yerine "güvenilirliğin alt sınırı" olarak tanımlamıştır. Guttman'ın sıralaması ve orijinal sembolleri ile bu katsayılar $L_1$, ..., $L_6$ katsayılardır. Bu katsayılardan özellikle $L_3$ katsayısı ölçüme literatüründe diğerlerine göre daha çok ön plana çıkmaktadır.

$$L_3 = \frac{k}{k-1} \left[ \frac{\sum_{i=1}^{k} \text{Var}(X_i)}{\text{Var}(X)} \right]$$  \tag{1}

Burada $k$, ölçmelerin (madde puanlarının) sayısı, $\text{Var}(X_i)$ i. ölçmenin varyansı ve $\text{Var}(X)$ ise tüm ölçüme kümesinin varyansını göstermektedir. Cronbach (1951) ise $L_3$ katsayısinin ölçmelerin paralel olması bile eşdeğer olması durumunda dahi gerçek güvenilirliği vereceğini ifade ederek bu katsayıyı $\alpha$ gösterimyle yeniden önermiştir. Bu nedenle $\alpha$ katsayısı bazı kaynaklarda Guttman-Cronbach katsayısı olarak verilmektedir. Varyans-kovaryans terimleri (matrisi) üzerine kurulu olan $\alpha_{VK}$ katsayısının matrisler ile gösterimi Eşitlik 2'de verilmiştir.

$$\alpha_{VK} = \frac{k}{k-1} \frac{\text{Tr}(\Sigma)}{1 \text{Tr} \Sigma 1}$$ \tag{2}

Burada $\Sigma$, ktk boyutlu varyans-kovaryans matrisi, $1$ ise tüm elemanları 1 olan kx1 boyutlu vektörü ve $\text{Tr}(\Sigma)$ ise $\Sigma$ matrisinin $\text{iz'i}$ (kosegen elemanlarının toplamı) göstermektedir. Evrene ilişkin $\alpha_{VK}$ katsayısının örneklem karşılığı ise;

$$\hat{\alpha}_{VK} = \frac{k}{k-1} \frac{\text{iz}(\text{S})}{1 \text{S} 1},$$ \tag{3}

seçklinde. Burada $S$, örneklemden elde edilen varyans-kovaryans matrisidir. Çokdeğişkenli normal dağılım varsayımlı altında; örneklem alfa katsayısı ($\hat{\alpha}_{VK}$) aynı zamanda evren alfa katsayısının $\alpha_{VK}$ ençoğ olabilirlik kestiricisidir (maximum likelihood estimator) (van Yzel vd., 2000; Yuan & Bentler, 2002).
Bununla birlikte; çoğu uygulamada ve özellikle bilgisayar paket programlarında alfa katsayısı standartlaştırılmış veriler (ya da korelasyon matrisi) üzerinden elde edilmektedir. Bu yaklaşımda Eşitlik 1 ile verilen ifade;

\[
\alpha_{\text{STD}} = \frac{k \bar{p}}{1 + \bar{p}(k-1)}
\]  

(4)

şekline dönüştülmektedir. Burada \(\bar{p}\) ölçmeler arası korelasyonların ortalamanını göstermektedir. Diğer taraftan standartlaştırılmış alfa katsayısının (\(\alpha_{\text{STD}}\)) matris notasyonu ile gösterimi ise;

\[
\alpha_{\text{VK}} = \frac{k}{k-1} \left[ \frac{k}{1^R 1} \right]
\]

(5)

şeklindedir ve \(R\), k x k boyutlu ölçmeler arasındaki korelasyon matrisini göstermektedir.


\[
\alpha_{\text{Max}} = \frac{k}{k-1} \left[ 1 - \frac{1}{\delta_1} \right]
\]

(6)

Burada \(\delta_1\), temel bileşenler analizinden elde edilen en büyük özdegerdir. Bazı kaynaklarda \(\alpha_{\text{Max}}\) katsayısı \(\text{enbüyüklenmiş alfa katsayısı}\) (maximized alfa) olarak da adlandırılmaktadır. Bunun nedenlerinden birisi ise \(\alpha_{\text{VK}}\) ile \(\alpha_{\text{Max}}\) arasındaki bağıntının \(\alpha_{\text{VK}} \leq \alpha_{\text{Max}}\) olmasından kaynaklanmaktadır (Carmines & Zeller, 1979). Diğer taraftan McDonald (1985), \(\alpha_{\text{VK}}\) katsayısının doğrulayıcı faktör analizinden elde edilen faktör yükleri üzerine kurulu versiyonunu (\(\alpha_{\text{M}}\)) tanımlamıştır:

\[
\alpha_{\text{M}} = \frac{k}{k-1} \frac{k(\bar{\lambda})^2 - \bar{\lambda}^2}{k(\bar{\lambda})^2 - \bar{\theta}^2}
\]

(7)

Burada, \(\bar{\lambda}\) doğrulayıcı faktör analizinden elde edilmiş standartlaştırılmış faktör yükleri ortalamasını ve \(\bar{\theta}\) ise ölçmelerin hata terimleri ortalamasını göstermektedir.

Aslında Gutman’ın (1945) çalışması güvenilirlik literatüründe bir lokomotif görevi görmüş ve 1945 yılında sonra, gerçek güvenilirliği en yüksek kestirimi iddia edilen çok sayıda güvenilirlik katsayısı önerilmiş ve/veya en popüler kullanıma sahip olan alfa katsayısını enbüyüklemek için çeşitli katsayılara ya da yöntemler geliştirilmiştir. Ancak günümüzde hesaplama kolaylığı ve paket programlarda yer almış nedeniyle en yaygın kullanıma sahip olan güvenilirlik katsayısının \(\alpha\) ya da \(\alpha_{\text{STD}}\) katsayısı olduğu gözlemlemektir. Öyle ki 2004 yılına kadar 5590’dan fazla araştırma Cronbach’in (1951) makalesine referans verilmiş (Cronbach & Shavelson, 2004) ve yılda ortalama 325 adet Social Sciences Citation Index-SSCI kapsamındaki dergilerden bu makaleye referans verilmiştir (Liu & Zumbo 2007).

Bu çalışmanın kapsamında dört farklı alfa katsayısı ($\alpha_{VK}$, $\alpha_{STD}$, $\alpha_{Max}$ ve $\alpha_{M}$) ele alınmış ve bu katsayılarnın optimallıkları araştırılmıştır. Optimallık araştırması, parametre ile kestirici tutarlılığı, alfa katsaylarının değişik boyut sayısına sahip veri kümeléindeki ve farklı ölçüm yapısına (paralel, eşdeğer ve konjenerik) sahip veri kümeléindeki davranışları üzerine yapılmıştır.

2. YÖNTEM

2.1. Veri Kümesi

Araştırmanın veri kümesi olarak Matthias Jerusalem ve Ralf Schwarzer tarafından geliştirilen “genel özüterlik algoritmosunun” ilişkine ilişkin 25 ülkeden 19120 kişiden toplanan veriler kullanılmış amaçlanmıştır. Ancak, veri kümesinde bazı boş gözlemler (çalışma sonuçlarından kayıp gözlem etkilerinin andırmak için) veri kümesinden çıkarılmış nedeniyle veri kümesi 18178 gözleme inmiştir. Bu veri kümesi araştırmda evren olarak kabul edilmiş ve bootstrap yöntemi ile Monte-Carlo simulasyonuna dayalı olarak farklı genişlikteki örneklemeler elde edilerek araştırımda kullanılmıştır.

2.2. Araştırma ve Simülasyon Tasarımı

2.2.1. Parametre-kestirici tutarlılığı:


Bu aşamada; evren ilişkini veri kümesinden elde edilen parametreler ($\alpha_{VK}$, $\alpha_{STD}$, $\alpha_{Max}$ ve $\alpha_{M}$) ile Monte-Carlo simulasyonu ile elde edilen örneklem değerleri ($\hat{\alpha}_{VK}$, $\hat{\alpha}_{STD}$, $\hat{\alpha}_{Max}$, $\hat{\alpha}_{M}$) karşılaştırılmıştır. Bu karşılaştırma yapılarak belirli örneklem genişliklerinin (N=30, 100, 300 ve 500) her biri için M=1000 adet basit seçkisiz (random) örneklem yöntemi ile örneklem çekilmiş ve bu örneklemlerden alfa katsaylarının örneklem değerleri elde edilerek bu katsayılara ilişkini hata indeksi hesaplanmıştır. Hata indeksi olarak hata kareler ortalaması karekökü (root mean square error, RMSE) kullanılmıştır.

$$RMSE = \sqrt{\frac{1}{M} \sum_{j=1}^{M} (Kestirici_{1} - Parametre)^{2}}$$

Burada $Kestirici_{1}$, j. örneklemenden elde edilen kestirilen alfa değerini göstermektedir. M ise simulasyondaki tekrarlanan örneklem sayısını göstermektedir. RMSE değerleri 0’a yaklaştırıcığa ilgili kestiricinin o derecede sağlam kestirici (robust estimator) ve aynı zamanda tutarlı kestirici olduğu ifade edilir (Poon, Leung, & Lee, 2002). Buna göre; $\alpha_{VK}$, $\alpha_{STD}$, $\alpha_{Max}$ ve $\alpha_{M}$ katsaylarının farklı örneklem genişliklerinde farklı RMSE değerlerine göre tutarlılıkları incelenmiştir.

2.2.2. Veri Kümesinin Boyutluğunu ve Alfa Katsayları:

Alfa katsayısına yönelik eleştirilerin başında; tek boyutlu veri kümlerinde güvenilirlik indeksini daha yansıtır kestiridi, çok boyutlu veri kümlerinde ise güvenilir indeksinden daha düşük değerler ürettiği yönündeki olumsuzluk gelmektedir. Bu araştırmda dört alfa katsayısının bir önceki aşamada elde edilen örneklem değerleri ve kümlerindeki faktör sayılara göre davranışı incelenmiştir. Böylelikle farklı boyutlardaki veri kümleri için $\alpha_{VK}$, $\alpha_{STD}$, $\alpha_{Max}$ ve $\alpha_{M}$ katsaylarının aldığı değerlerin büyükliklerine göre sıralanması amaçlanmıştır. Bu aşamada faktör sayları Kaiser-Guttman kuralı

---

(Gorsuch, 1983) gereği; temel bileşenler analizinden elde edilen özdeğerlerin 1 değerinden bük olma durumuna göre belirlenmiştir.

2.2.2. Ölçmelerin Yapısı ve Alfa Katsayıları:


\[
\phi = \sum_{i=1}^{k-1} (\beta_i - \bar{\beta})^2
\]

Burada \(\beta_i\) i. ölçmeye ilişkin faktör yükünü göstermektedir. Bilindiği gibi; standartlaştırılmış faktör yüklerinin eşit olması ölçmelerin paralel ya da eşdeğer olduğunu göstermektedir. Buna göre tüm yükler eşit olduğunda \(\phi=0\) olacaktır. Diğer taraftan \(\phi\) değeri 0′dan uzaklaşmada (\(\phi \to +\infty\)) ölçmelerde o denli paralel ya da eşdeğer ölçme özelliğinden uzaklaşıp konjenerik ölçmeye yolculuğunu göstermektedir3.

Araştırmının yöntem bölümünde belirlenen karşılaştırma ülkeleri simulasyonlar için SIMREL adi verilen bir program (Yurdugül, baskıda) kullanılmıştır.

3. BULGULAR

3.1. Alfa Katsaylarının Sağlam Kestirici Özellikleri

Araştırmının evrene ilişkin veri kümesinden elde edilen dört adet alfa katsayısının parametre değerleri surasıyla; \(\alpha_{VK}=0.856, \alpha_{STD}=0.857, \alpha_{Max}=0.859\) ve \(\alpha_{Max}=0.856\) bulunmaktadır. Daha sonra evren kümesinden örneklem genişliği surasıyla N=30, 100, 300 ve 500 olacak şekilde (her bir örneklem genişliği için) basit seçkisiz örneklemle yöntemi ile 1000 adet örneklemler çekilerek dört adet alfa katsayısının örneklem değerleri elde edilmiştir. Her bir örneklem değeri için RMSE(\(\alpha_{VK}\)), RMSE(\(\alpha_{STD}\)), RMSE(\(\alpha_{Max}\)) ve RMSE(\(\alpha_{Max}\)) değerleri hesaplanmıştır. Bu simulasyon sonuçlarına ilişkin alfa katsaylarının farklı örneklem genişliklerindeki RMSE değerleri ve grafiği aşağıdaki verilmiştir.

Tablo 1: Simülasyon sonuçlarına göre N=30, 100, 300, 500 örneklem genişliklerindeki alfa katsaylarının RMSE değerleri

<table>
<thead>
<tr>
<th>N</th>
<th>RMSE ((\alpha_{VK}))</th>
<th>RMSE ((\alpha_{STD}))</th>
<th>RMSE ((\alpha_{Max}))</th>
<th>RMSE ((\alpha_{Max}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>0.058</td>
<td>0.055</td>
<td>0.046</td>
<td>0.060</td>
</tr>
<tr>
<td>100</td>
<td>0.030</td>
<td>0.030</td>
<td>0.028</td>
<td>0.030</td>
</tr>
<tr>
<td>300</td>
<td>0.015</td>
<td>0.015</td>
<td>0.015</td>
<td>0.015</td>
</tr>
<tr>
<td>500</td>
<td>0.011</td>
<td>0.011</td>
<td>0.011</td>
<td>0.011</td>
</tr>
</tbody>
</table>

Konjenerik ölçmelerin önemli nedenlerinden birisi (2. aşama karşılaştırmaarda konu edilen) çokboytulu yapılar. Ancak konjenerik ölçmeleri sahip tekboytulu veri kümelere verilen nedeniyle (Lucke, 2005; Zinbarg vd., 2005) ölçmelerin yapılarında göre alfa katsaylarının karşılaştırma alır olarak ele alınmıştır.

Bu çalışmada \(\phi=0\) değeri ölçüme varyanslarının eşit olduğunu göstermektedir. Bu durumda ölçmeler paralel ya da eşdeğer olabilir (Lucke, 2005). Ölçmelerin paralel mi yoksa eşdeğer mi olduğunu anlamak için ortalamalarındaki sapmaları araştırmak gerekmektedir. Ancak bu farklılık ölçüme çalışması ile doğruda ilişkili olmadığı için \(\phi=0\) değeri taşıyan ölçmeler “eşdeğer ölçme” olarak adlandırılmıştır.
Daha önce ifade ettiği gibi; RMSE değerleri incelemeye konu olan istatistiklerin sağlam kestirici özellikleri ortaya koyan bir değerdir ve bu değer 0'a yakın ise inceленen istatistik parametreinin o denli sağlam bir kestiricisi olarak nitelendirilir. RMSE değerleri örneklem genişliklerinden fazlaca etkilenir. Tablo 1’c göre; N=300 ve N=500 olduğunda tüm alfa değerlerinin RMSE değerleri eşi olarak elde edilmiştir. Ancak bu araştırmada ele alınan en düşük örneklem genişliğinde (N=30), en küçük RMSE değeri $\alpha_{\text{Max}}$ ve en büyük RMSE değeri ise $\alpha_{M}$ için gözlenmiştir.

Alfa katsayılarına ilişkin RMSE değerlerinin farklı örneklem genişliklerindeki davranış Çizim 1’de verilmiştir. Buna göre N=30 olduğunda RMSE($\alpha_{\text{Max}}$)<RMSE($\alpha_{\text{STD}}$)<RMSE($\alpha_{V(K)}$)<RMSE($\alpha_{M}$) ilişkisi gözlenmiştir. N=100 için ise RMSE($\alpha_{\text{Max}}$)<RMSE($\alpha_{V(K)}$)=RMSE($\alpha_{\text{STD}}$)= RMSE($\alpha_{M}$) sıralaması söz konusudur. Son olarak N=300 ve N=500 olduğunda ise her dört alfa katsayısının RMSE değerleri eşi olarak elde edilmiştir. Buna göre dört farklı alfa katsaylarının örneklem karşılıklarından küçük örneklemler için (N=30, N=100) en sağlam kestiricinin $\alpha_{\text{Max}}$ olduğu söylenebilir.

Genel olarak RMSE değerleri ele alındığında RMSE($\alpha_{\text{Max}}$) değerinin tüm örneklem genişliklerinde diğer alfa katsayılarının RMSE değerlerinde düşük ya da eşi olduğu Çizim 1’de görülmektedir. Buna göre; sağlam kestirici özelliğinden dolayı en iyi alfa katsayısını Armor (1974) tarafından önerilen $\alpha_{\text{Max}}$ katsayısı olduğu ifade edilebilir.

Çizim 1: Farklı örneklem genişliklerinde alfa katsayılarının RMSE değerlerinin davranışları

3.2. Veri Kümesinin Boyutluğunu ve Alfa Katsayları:

Güvenilirlik konusuna ilişkin çalışmalarla; üzerinde çalışılan veri kümenin boyut sayısı arttıkça alfa katsayısının düşük değerler ürettiği ifade edilir. Bu nedenle; bu araştırma kapsamında alfa katsayılarının temel bileşenler analizinden elde edilen faktör sayımasına göre aldığı değerlerin büyüklikleri incelemiştir. Faktör sayıları, veri kümenin özdeğerlerinin 1’den büyük olma durumlarına göre belirlenmiştir. Diğer taraftan simülasyon sonucu elde edilen veri kümelerini 1, 2, 3 ve 4 faktörü yapılarca sahip olduğu görülmüştür. Tablo 2’de bu faktörlere sahip veri kümelerinden elde edilen dört farklı alfa katsayılarının örneklem ortalamaları verilmiştir.

---

Tablo 2: Simülasyon sonuçlarına göre farklı boyutlardaki örneklemelerden elde edilen alfa katsaylarının ortalama değerleri

<table>
<thead>
<tr>
<th>Faktör</th>
<th>$\alpha_{VK}$</th>
<th>$\alpha_{STD}$</th>
<th>$\alpha_{Max}$</th>
<th>$\alpha_{M}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>0.763</td>
<td>0.769</td>
<td>0.790</td>
<td>0.759</td>
</tr>
<tr>
<td>3</td>
<td>0.824</td>
<td>0.828</td>
<td>0.842</td>
<td>0.825</td>
</tr>
<tr>
<td>2</td>
<td>0.884</td>
<td>0.886</td>
<td>0.892</td>
<td>0.885</td>
</tr>
<tr>
<td>1</td>
<td>0.927</td>
<td>0.928</td>
<td>0.930</td>
<td>0.928</td>
</tr>
</tbody>
</table>

Tablo 2’ye göre; alfa katsaylarının en yüksek ortalama değerleri bir faktörlü (tekboyutlu) veri kümelerinden elde edilmiş ve faktör sayısı arttıkça tüm katsayların değerleri düşme göstermektedir. Alfa katsaylarının ortalama değerleri (faktör sayısına göre) göz önüne alındığında $\alpha_{Max}$ katsayısının diğer katsayılara göre daha yüksek olduğu gözlenmiştir.

Çizim 2’ye göre; tek faktörlü veri kümelerinde tüm katsaylar yaklaşık aynı değerler üretmiştir ($\alpha_{Max} = \alpha_{STD} = \alpha_{VK} = \alpha_{Max} = 0.93$). Ancak faktör sayısı artışta bu eşitlik bozulmaktadır ve özellikle 4 faktörlü veri kümelerinde alfa katsaylarının ortalama değerleri arasında ($\alpha_{Max} > \alpha_{STD} > \alpha_{VK} > \alpha_{M}$) eşitsizlikleri ortaya çıkmaktadır. Burada vurgulanması gereken en önemli bulguun ise Tablo 2’de de görüldüğü gibi her 4 faktörde de $\alpha_{Max}$ değerinin diğer alfa katsayılardan daha yüksek olduğu görülmüştür.

![Alpha Değerleri](image)

**Çizim 2:** Farklı boyutlardaki örneklem veri kümelerinde alfa katsaylarının davranış

3.3. Ölçmelerin Yapısı ve Alfa Katsayları:


$$\delta_{Max-VK} = \hat{\alpha}_{Max} - \hat{\alpha}_{VK}$$  (10)

$$\delta_{Max-STD} = \hat{\alpha}_{Max} - \hat{\alpha}_{STD}$$  (11)

$$\delta_{Max-M} = \hat{\alpha}_{Max} - \hat{\alpha}_{M}$$  (12)

$$\delta_{STD-VK} = \hat{\alpha}_{STD} - \hat{\alpha}_{VK}$$  (13)
\[ \delta_{\text{STD-M}} = \hat{\alpha}_{\text{STD}} - \hat{\alpha}_M \]  
\[ \delta_{M-VK} = \hat{\alpha}_M - \hat{\alpha}_{VK} \]

Bu katsayıların farklıları Eşitlik 9'da verilen doğrulayıcı faktör analizinden elde edilen yüklerin varyanslarına (\(\phi\)) göre ilişkisi incelenmiştir (Bkz: Dipnot 3).

Çizim 3'te \(\alpha_{\text{Max}}\) katsayısının diğer alfa katsayıları ile farklarının, faktör yüklerindeki sapmalara karşın davranışlarını içeren grafikler verilmiştir. Diğer katsayılar arasındaki karşılaştırma grafikleri ise Ek 1'de verilmiştir. Çizim 3'te görüldüğü gibi; \(\phi=0\) olduğu zaman \(\delta_{\text{Max-VK}}=\delta_{\text{Max-STD}}=\delta_{\text{Max-M}}=0\) olmaktadır. Bunun anlamı ise faktör yüklerinin eşit olduğu durumda (paralel veya eşdeğer ölçmeler) \(\alpha_{\text{Max}}\) katsayısı diğer katsayılar ile aynı değerleri almaktadır. Ancak faktör yüklerindeki sapmalar arttıkça, \(\phi>0\), (konjenerik ölçmeler) \(\delta_{\text{Max-VK}}>0\) \(\delta_{\text{Max-STD}}<0\) \(\delta_{\text{Max-M}}>0\) olmaktadır; buna göre ölçme yapısına göre alfa katsayıların karşılaştırmalarında \(\alpha_{\text{Max}}<\alpha_{\text{VK}},\ \alpha_{\text{Max}}<\alpha_{\text{STD}}\) ve \(\alpha_{\text{Max}}<\alpha_{M}\) eşitsizlikleri gözlenmiştir. Diğer taraftan \(\phi\) ile \(\delta\) farkları arasındaki regresif bağıntı incelendiğinde veri kümesindeki konjenerik yapı arttıkça (\(\phi \to +\infty\)) alfa katsayıları arasındaki en hızlı değişim \(\alpha_{\text{Max}}\) ile \(\alpha_{M}\) arasında gerçekleşmektedir.

![Diagram a) \(\phi\) ile \(\delta_{\text{Max-VK}}\) arasındaki bağıntı](image1)

![Diagram b) \(\phi\) ile \(\delta_{\text{Max-STD}}\) arasındaki bağıntı](image2)
Çizim 3: Konjenerik ölçeklerde $\alpha_{\text{Max}}$ katsayısının diğer alfa katsayılara göre davranışı

4. YORUM VE TARTIŞMA

Bu çalışmada klasik güvenilirlik katsayılardan olan alfa katsayısının dört farklı versiyonunu özellikle karşılaştırılmıştır. Eğitimsel ve psikolojik araştırmalara ilişkin güvenilirlik literatüründe yalnızca $\alpha_{VK}$, $\alpha_{STD}$ ve $\alpha_{\text{Max}}$ katsayları büyüklikleri bakımından çeşitli karşılaştırmalar yapılmıştır. Bu araştırmalarla göre; $\alpha_{VK} \leq \alpha_{STD} \leq \alpha_{\text{Max}}$ türünden bir eşitsizlik ortaya konmuştur (Carmines & Zeller, 1979; Iacobucci & Duhachek, 2003; Osburn, 2000; Yurdugül, 2007). Bu araştırmada literatürdeki karşılaştırmalardan farklı olarak bu katsayılarn istatistiksel özellikler (sağlam kestiricilikleri), farklı boyutlardaki ve yapaylardaki (eşdeğer ve konjenerik) davranışları incelenmiştir.

Farklı örneklemelerde elde edilen RMSE değerlerine göre en sağlam kestirici olgununun $\alpha_{\text{Max}}$ katsayısı olduğu görülmüştür. Özellikle örneklem genişliği oldukça olduğunda ($N=30$) en az hata içeren katsayı $\alpha_{\text{Max}}$ katsayısı RMSE($\alpha_{\text{Max}}$)<RMSE($\alpha_{STD}$)<RMSE($\alpha_{VK}$)<RMSE($\alpha_{M}$) iken, örneklem genişliği arttıkça ($N=300$, 500) her dört katsayının RMSE değerleri eşit olmaktadır.

Üzerinde çalışılan veri kümesinin boyutuna göre her dört alfa katsayısının nicel büyüklikleri ele alınırken; tekboyutlu veri kümlerinde dört katsayının da değerlerinin yaklaşıq eşit olduğu ($\alpha_{\text{Max}} = \alpha_{STD} = \alpha_{VK} = \alpha_{M} = 0.93$) gözlenmiş iken, boyut sayısı arttıktça $\alpha_{\text{Max}}$ katsayısının daha büyük değerler ürettiği gözlenmiştir ($\alpha_{\text{Max}} \geq \alpha_{STD} \geq \alpha_{VK} \geq \alpha_{M}$).

Veri kümlerinde yer alan ölçeklerin yapıları göz ömine alınırken ise; paralel ve eşdeğer ölçekler için $\alpha_{\text{Max}}$ katsayısının değerleri ile diğer katsayların değerleri yaklaştır aynı çıkarırken, ölçekler konjenerik yapıya yönelikinde $\alpha_{\text{Max}}$ katsayısı diğer katsayılara göre daha büyük değerler üretmektedir.

Diğer taraftan, karşılaştırmalar standartlaştırılmış ve standartlaştırılmamış alfa katsayılari açısından ele alınırken; her ne kadar Iacobucci, ve Duhachek (2003) olarak standartlaştırılmış alfa katsaylarının rapor edilmesini öneriler olarak ifade etmesine karşın küçük örneklem genişliğinde (Çizim 1) standartlaştırılmamış alfa katsayısının daha sağlam kestirici olduğu görülmektedir. Bununla birlikte veri kümesi tekboyutlu yapıdan çoğaboyutlu bir yapıla yöneländiğinde ise standartlaştırılmış alfa katsayısı güvenilirliği daha büyük değerler ile kestirmektedir (Çizim 2). Ancak ölçekler paralel ölçeklerden konjenerik ölçeklere yönelikinde ise bu iki katsayının belirgin bir üstünüülü ortaya çıkmamaktadır (Ek 1).

Son olarak, tüm karşılaştırmalarda en olumuz özellikse sahip alfa katsayısının McDonald (1985) tarafından tanımlanan $\alpha_{M}$ katsayısı olduğu gözlenmiştir. Bunun olduğu nedenlerinden birisi ise $\alpha_{M}$ katsayısının (Eşitlik 7) faktör yüklerinin ilgili ölçmenin varyansının kestirimine karşılık gelmesi ve böylelikle kestirilen ölçme varyansının doğrulayıcı faktör analizi tarafından gerçek ölçme varyansından yarılı olarak kestirilmeli olabilir.
5. SONUÇLAR VE ÖNERİLER

Bu çalışmada dört farklı alfa katsayısının optimallıkları üzerine üç aşamalı bir karşılaştırma yapılmıştır. Büyük örneklem genişliklerin, tekboylutlu ve paralel/eşdeğer veri kümelerinde tüm katsayılara özdeş olarak elde edilmiştir. Ancak düşük örneklem genişliğinde, çokboylutlu ve konjenerik ölçümlerde $\alpha_{\text{Max}}$ katsayısı en iyi özellikle sahip katsayi olarak belirlenmiştir. Bununla birlikte; en olumsuz katsayı ise $\alpha_{M}$ katsayısı olarak bulunmuştur. Özellikle her üç aşamada ele alınan özellikler genel olarak ele alınırken araçtırmacılar $\alpha_{\text{Max}}$ katsayısının kullanımı önerilir.

KAYNAKLAR


**EXTENDED ABSTRACT**

Reliability in educational and psychological literature can be expressed in terms of stability, equivalence, and consistency. Unlike test-retest for stability and alternate form for equivalence, only a single test is needed for estimating internal consistency. Coefficient alpha suggested by Guttman (1945) and Cronbach (1951) is one of the most commonly used methods in the educational and psychological studies for evaluating internal consistency as an indicator of test quality. Cronbach (2004) reported that his article published in 1951 had been cited no less than 5590 times and, in recent years, had been cited approximately 325 times per year in the Social Sciences Citation Index (Liu, & Zumbo, 2007).

However, in the related literatures, four different coefficient alphas are defined. The coefficient alphas were ranked as: a) *unstandardized alpha* ($\alpha_{UK}$) based on variance-covariance matrix of data set, b) *standardized alpha* $\alpha_{STS}$ based on correlation matrix of data set, c) maximized alpha ($\alpha_{max}$) based on the first eigenvalue obtained principal component analysis (Armor, 1974), and d) *estimated alpha* ($\alpha_{M}$) by McDonald (1985) based on confirmatory factor analysis.

Which coefficient alpha should we use in the calculating reliability of measurement tools? In this study, the optimality of four different coefficient alphas defined above was investigated. For this aim of the study, the research design was conducted within three stages: a) Firstly, the robust estimator properties of coefficient alphas were taken into consideration. The robust estimator properties of those were investigated by root mean square error (RMSE) obtained from Monte-Carlo simulations for different sample sizes (N=30, 100, 300, 500), b) secondly the comparison of coefficient alphas’ values obtained from different dimensional data set, and c) finally comparisons of coefficient alphas’ performance in different data sets which include tau-equivalent and congeneric measures. The test of whether the measures are tau-equivalent (i.e., the factor loadings are equal) or congeneric (i.e., the factor loadings are not equal) measures, was investigated in deviation of unstandardized factor loadings given in Equation 9.

The investigations in the second (b) and third (c) stages were taken into consideration as weakness of coefficient alpha. As known, the coefficient alpha yields a low boundary of real reliability in multidimensional and/or congeneric data set. A necessary and sufficient condition for coefficient alpha to equal the reliability of the test score is that the measures are parallel, tau-equivalent, and
essentially tau-equivalent. An essentially tau-equivalent model includes the essentially tau equivalent measures in which simply the factor loadings are equal for all of measures. The model involves that the off-diagonal elements in the covariance matrix are all equal, but that the variances of measures are not equal for all measures. The congeneric measures may conform to a single-factor model with distinct factor loadings or conform to a multi-factor model. Therefore, this study deals with the coefficient alphas' performance in dimensionality of data set and nature of measures separately. Due to the bias of coefficient alpha under those conditions, the comparisons among coefficient alphas were included in this study. The data set in this study was obtained by 'the general self-efficacy scale' which developed by Schwarzer and Jerusalem (1995). The psychometric properties of data set were reported by Scholz, Gutiérrez Doña, Sud, & Schwarzer (2002). The scale includes 10 items and the data set includes 19120 observations from 25 countries. The data set was used as a population data and several repeated samples were drawn from those population data and those samples were used in Monte-Carlo simulations of this study.

According to results of Monte-Carlo simulation, the smallest RMSE value of coefficient alphas was observed for ($\alpha_{Max}$) in the smallest sample size (N=30), and the ordering of RMSE values in N=30 are RMSE($\alpha_{Max}$)<RMSE($\alpha_{STD}$)<RMSE($\alpha_{VK}$)<RMSE($\alpha_{M}$). In other sample size (N=100, 300, 500) all RMSE values were approximately the same (Table 1 and Figure 1).

According to arithmetical means of coefficient alphas' values in unidimensional data sets, all of the coefficient alpha were approximately equal ($\alpha_{Max} = \alpha_{STD} = \alpha_{VK} = \alpha_{M} = 0.93$). From single-factor data sets to multi-factor data sets, all of the coefficient alphas’ were decreasing. However, the slowest decrease was observed for coefficient $\alpha_{Max}$ and the slowest decrease was observed for coefficient $\alpha_{M}$. Specially in the four factor data sets, the relation of ($\alpha_{Max}$>$\alpha_{STD}$>$\alpha_{VK}$>$\alpha_{M}$) for means of alphas' values was observed (Table 2 and Figure 1).

In stage (c); the coefficient alphas were compared in terms of measures’ nature such as parallel, tau-equivalent, and congeneric. As known, the coefficient alpha underestimates the true reliability in congeneric measures. When those four alpha values were compared in congeneric measures, it is observed that the coefficient $\alpha_{Max}$ yields the highest value. The lowest value was yielded by the coefficient $\alpha_{M}$. According to the results in this study, it is suggested that the coefficient $\alpha_{Max}$ to be used in psychological and educational researches.
EK:

Ölçmelerin Yapılara Göre Alfa Katsaylarının Karşılaştırımları

Yük Sapmalarına (φ) karşın $\alpha_{\text{STD}}$ ve $\alpha_{\text{VK}}$ farklı (δ_{STD-VK}) davranış

Yük Sapmalarına (φ) karşın $\alpha_{\text{STD}}$ ve $\alpha_{\text{M}}$ farklı (δ_{STD-M}) davranış

Yük Sapmalarına (φ) karşın $\alpha_{\text{M}}$ ve $\alpha_{\text{VK}}$ farklı (δ_{M-VK}) davranış