Hacettepe Üniversitesi Eğitim Fakültesi Dergisi (H. U. Journal of Education) 35:397-405 [2008]

MINIMUM SAMPLE SIZE FOR CRONBACH'S COEFFICIENT ALPHA: A MONTE-CARLO STUDY

CRONBACH ALFA KATSAYISI İÇİN MİNİMUM ÖRNEKLEM GENİŞLİĞİ: MONTE-CARLO ÇALIŞMASI

Halil YURDUGÜL*

ABSTRACT: The coefficient alpha is the most widely used measure of internal consistency for composite scores in the educational and psychological studies. However, due to the difficulties of data gathering in psychometric studies, the minimum sample size for the sample coefficient alpha has been frequently debated. There are various suggested minimum sample sizes for the robust estimate of the population coefficient alpha. This research indicates that the performance of an estimator of the coefficient alpha depends not only on the sample size but also on the largest eigenvalue of the sample data set. Thus, when the largest eigenvalue increases, unbiased estimation of the population coefficient alpha is possible, even though the sample size is small. The simulations in this study were based on Monte-Carlo method with bootstrap technique.
Keywords: Reliability, sample size, coefficient alpha, Monte-Carlo, Simulation

ÖZET: Eğitimsel ve psikolojik çalışmalarda birleşik ölçmelerin güvenirliğinin hesaplamasında yaygın olarak Cronbach (1951) tarafından geliştirilen alfa katsayısı kullanılmaktadır. Ancak bu tür çalışmalarda veri toplamanın zorluğu nedeniyle alfa katsayısı için gerekli olan minimum örneklem genişliği tartışma konusudur. Evren alfa katsayısının sağlam kestirimi için gerekli olan minimum örneklem genişliği için farklı öneriler vardır. Bu çalışmanın sonuçlarına göre; yansız ve tutarlı bir alfa kestirimi örneklem genişliğinin büyüklüğü kadar aynı zamanda ölçmelerin birinci özdeğerinin büyüklüğüne bağlıdır. Buna göre, birinci özdeğer büyüdükçe, düşük örneklem genişliklerinde alfa katsayısının yansız bir kestirimi olanaklıdır. Bu çalışmada kullanılan simulasyonlar bootstrap tekniği ile birlikte Monte-Carlo yöntemi üzerine kurulmuştur.
Anahtar sözcükler: Güvenirlik, örneklem genişliği, alfa katsayısı, Monte-Carlo, simülasyon

1. INTRODUCTION

A major issue in psychometric studies is the precision of measurement, which is frequently quantified by the reliability coefficients. The coefficient alpha, developed by Cronbach (1951), is the most commonly used index for estimating the reliability of measurement instruments such as scales, multiple item tests, questionnaires, or inventories (Raykov 1997) in the fields of psychology, education, statistics, sociology, medicine, counselling, nursing, political science, and economics (Cortina 1993). Cronbach (2004) reported that his 1951 article had been cited no less than 5,590 times and, in recent years, had been cited approximately 325 times per year in the Social Sciences Citation Index (Liu & Zumbo; 2007). In some of those studies, the sample coefficient alpha in Peterson's meta-analysis study (1994) on coefficient alpha is obtained from a small sample size ($n \leq 200$). The general view on this subject is that the sample coefficient alpha obtained from larger samples tends to produce a more accurate estimate of the population coefficient alpha. In reliability literature, Kline (1986) suggested a minimum sample size of 300, as did Nunnally and Bernstein (1994). Segall (1994) called a sample size of 300 "small". Charter (1999) stated that a minimum sample size of 400 was needed for a sufficiently precise estimate of the population coefficient alpha. Charter (2003) has noted that with low sample sizes alpha coefficients can be unstable. However, minimum sample size for the sample coefficient alpha has been frequently debated due to the difficulty of data collection in psychometric research.

Factor analysis prompts a similar debate. In factor analysis, there are several suggestions about necessary sample size. However, MacCallum *et al*. (1999) declare that necessary minimum sample size for factor analysis is dependent on the level of communalities of item variables. The approach by MacCallum *et al*. (1999) can also be applied to the reliability coefficient in terms of principal

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

component analysis or PCA. This current study investigates the effects of sample size and of the magnitude of the largest eigenvalues obtained from PCA on the performance of an estimator of the population coefficient alpha.

1.1. The Coefficient Alpha

The coefficient alpha, α, is used as a measure of internal consistency for a scale with k items. Suppose $\mathbf{X} = \left[X_1 \cdots X_k\right]'$ denotes a vector of k responses. The population coefficient α of the test is given by

$$\alpha = \frac{k}{k-1}\left[1 - \frac{tr(\mathbf{\Sigma})}{\mathbf{1}'\mathbf{\Sigma}\mathbf{1}}\right]$$

where $\mathbf{\Sigma}$ is the variance-covariance matrix of the population data set, and $\mathbf{1}$ is a kx1 vector of 1s. The sample counterpart of the population coefficient alpha is given by

$$\hat{\alpha} = \frac{k}{k-1}\left[1 - \frac{tr(\mathbf{S})}{\mathbf{1}'\mathbf{S}\mathbf{1}}\right]$$

where $\mathbf{S}$ is the unbiased sample covariance matrix. The coefficient $\hat{\alpha}$ is also the maximum likelihood estimate of the coefficient alpha when data are multivariate normal. Under the assumption of a multivariate normal distribution, van Zyl et al. (2000) obtained the asymptotic distribution of the $\hat{\alpha}$ as

$$\sqrt{n}(\hat{\alpha} - \alpha) \sim N(0, \phi).$$

Here

$$\phi = \left[\frac{2k^2}{(k-1)^2(\mathbf{1}'\mathbf{\Sigma}\mathbf{1})^3}\right]\left[(\mathbf{1}'\mathbf{\Sigma}\mathbf{1})(tr\mathbf{\Sigma}^2 + tr^2\mathbf{\Sigma}) - 2(tr\mathbf{\Sigma})(\mathbf{1}'\mathbf{\Sigma}\mathbf{\Sigma}\mathbf{1})\right]$$

where n denotes sample size and $\hat{\alpha}$ is the MLE of α (van Zyl *et al*. 2000; Yuan and Bentler, 2002).

The reliability coefficients (e.g. coefficient alpha) for composite tests are also a measure of first-factor saturation of the data set (Crano and Brewer; 1973). Thus, standardized alpha based on a correlation matrix of item scores is directly related to the eigenvalue of the first unrotated principal component (Cortina 1993). Depending on this relation, Armor (1974) developed the coefficient theta, θ, as a version of the coefficient α:

$$\theta = \frac{k}{k-1}\left[1 - \frac{1}{\lambda_1}\right]$$

where λ_1 is the largest eigenvalue obtained from principal component analysis (PCA). Essentially, the coefficient alpha is also calculated from the correlation mean ($\bar{\rho}$) of $\binom{k}{2}$:

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

The magnitude of $\bar{\rho}$ is related to λ_1. In this study, the performance of the coefficient $\hat{\alpha}$ was investigated for different sample sizes and different values of $\hat{\lambda}_1$.

2. METHOD

2.1. The Design of Simulation

For sample size studies, Yuan et al. (2003) recommended that bootstrap estimates be used to compute the $\hat{\alpha}$ rather than asymptotic estimates. The bootstrap method uses raw data, whereas the asymptotic method uses a covariance or correlation matrix that does not include any information regarding sample size. Therefore, the bootstrap method was used and Likert-type normal data was generated by Box-Muller (1958) transformation in the simulation phase of this study. In the first stage of simulation, D=10000 population data were generated (each of which included N = 5000 observations and different numbers of randomly determined variables) and the coefficient α and λ_1 obtained from PCA was calculated from each population data set. In the second stage of simulation, for each population data set, 100 samples were drawn by simple random sampling with replacement for each n=30, 100, 300, and 500. For each sample data set, bootstrap estimators of the coefficient α and of the λ_1 were computed. In the last stage of simulation, the relative bias (R-Bias) and relative root mean square error (R-RMSE) were computed for every 100 samples (replication) in each sample size level. The relative bias and the relative RMSE were computed for each estimator.

$$\text{R - Bias}(\hat{\alpha}_{pi}) = \frac{1}{M}\sum_{j=1}^{M}\left[\frac{(\hat{\alpha}_{pij} - \alpha_p)}{\alpha_p}\right]$$

and

$$\text{R - RMSE}(\hat{\alpha}_{pi}) = \sqrt{\frac{1}{M}\sum_{j=1}^{M}\frac{(\hat{\alpha}_{pij} - \alpha_p)^2}{\alpha_p}}$$

where $\hat{\alpha}_{pij}$ is an estimator of the α_p for the jth replication and ith sample size (i = 30, 100, 300, 500) in the pth population data set (p = 1, 2,…, 10000), and M is the number of replications. In this study, when the values of relative bias were less than 0.01 and the values of relative RMSE were less than 0.05, the estimator for the parameter was deemed acceptable. In this way, the performance of $\hat{\alpha}$ for each sample size (n=30, 100, 300, 500) may be observed. Another issue is the item numbers regarding population data sets. Even though the item numbers were randomly determined in the simulation, lower and upper limits for item numbers were determined in advance. Accordingly, they are partially controlled as $5 \leq k \leq 20$. Table 1 presents the frequency of item numbers in the D = 10000 population data set generated.

Table 1: Item Numbers of Generated Data in Simulation

k	frequency	*k*	frequency	*k*	frequency	*k*	frequency
5	621	*9*	637	*13*	627	*17*	594
6	623	*10*	661	*14*	634	*18*	652
7	627	*11*	614	*15*	610	*19*	621
8	635	*12*	620	*16*	630	*20*	594

The simulations in this study were performed the SIMREL software (Yurdugul, in press). A modified form of this software application was developed for the simulation phases of this study. The algorithm of this application is constructed as Figure 1. This software application, developed for Microsoft operating systems, is coded in Visual Basic programming language. The more details of SIMREL can be seen in Yurdugul (in press) by researchers.

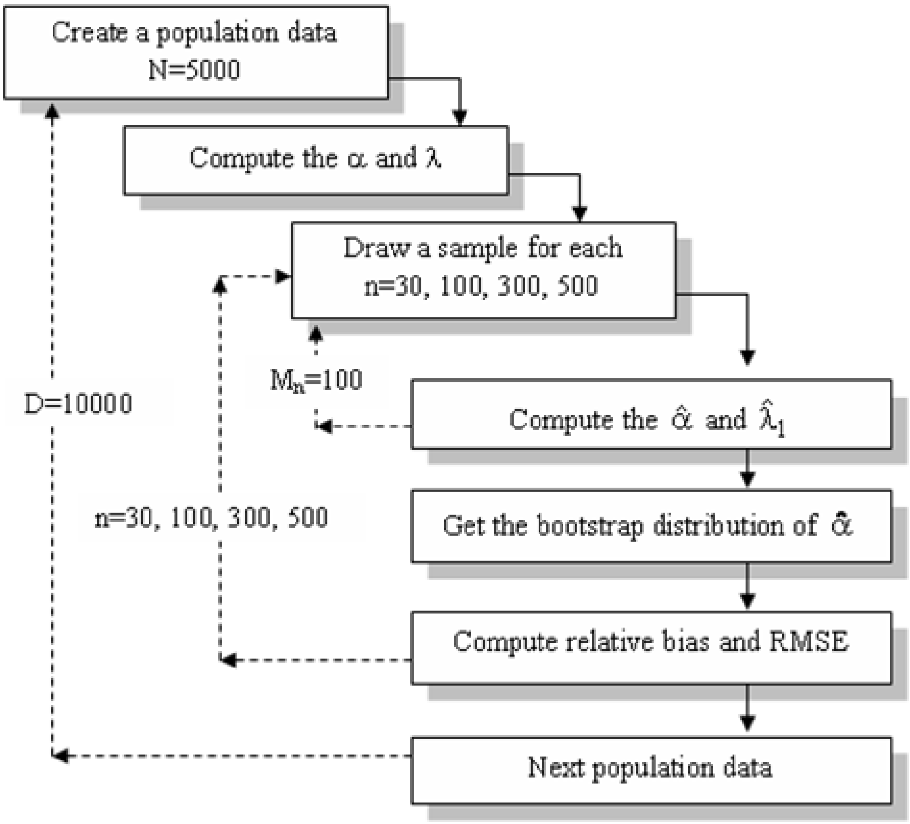

Figure 1: The simulation algorithm of SIMREL

3. FINDINGS

The results of the simulation were examined in two different ways. In the first approach, the $\hat{\lambda}_1$ was considered as a continuous variable, $1 \leq \hat{\lambda}_1 < \infty$. Thus, the behaviors of the R-RMSE and the R-Bias values of coefficient alpha estimates were observed for each $\hat{\lambda}_1$ value in different sample sizes. In the second approach, the $\hat{\lambda}_1$ values were categorized as discrete variables in order to investigate the effects of both sample size and levels of the λ_1 on the performance of the estimator.

In Figure 2, the relative bias increases with a decrease in sample size and in the largest eigenvalue. The accuracy of the coefficient α improved with larger sample sizes and higher values of the $\hat{\lambda}_1$. According to the values of R-Bias (Figure 2) and R-RMSE (Figure 3), when n=30, the magnitude of $\hat{\lambda}_1$ had a critical role due to the robust estimator of coefficient alpha. Although the suggested sample size for coefficient alpha is 300 in psychometric literature, it is observed in this study that when n=300, the estimator coefficient alpha is biased for the small values of $\hat{\lambda}_1$ (Figure 2 and 3). However, when n = 500, whatever the value of $\hat{\lambda}_1$, the estimated coefficient alpha is more precise.

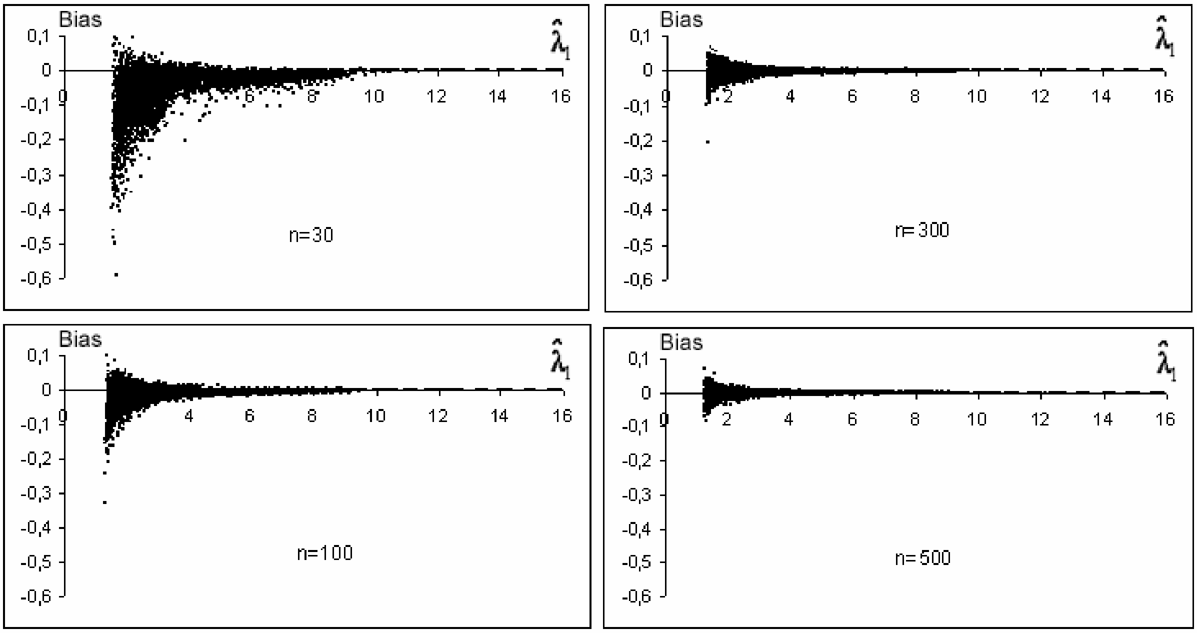

Figure 2: The performance of sample coefficient alpha with relative bias.

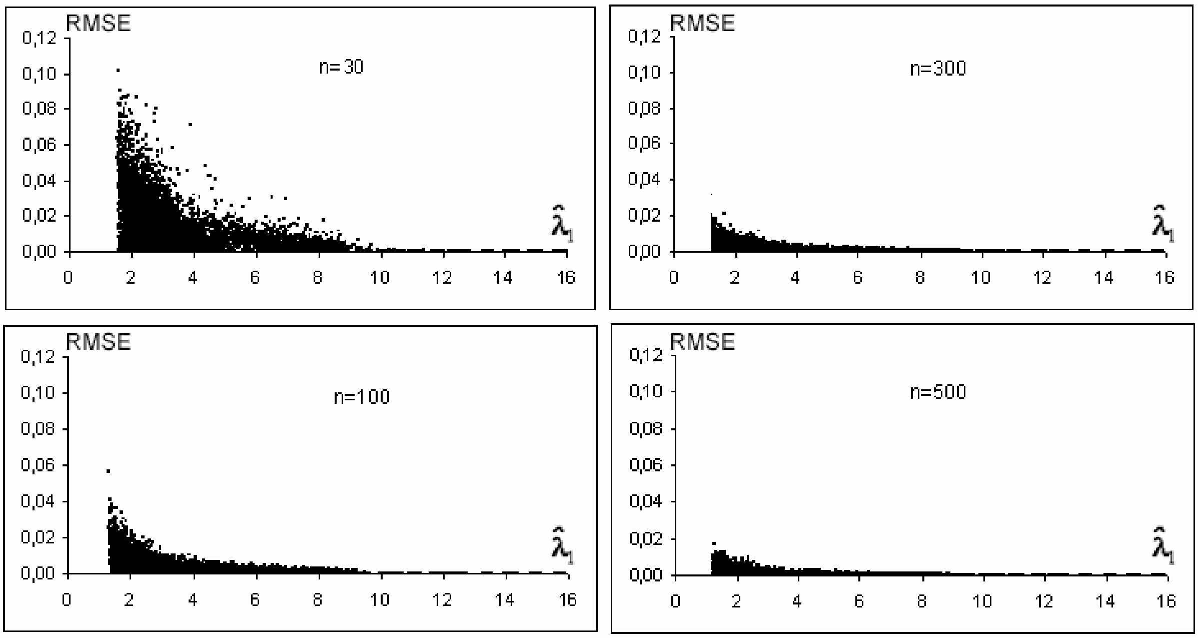

Figure 3: The performance of sample coefficient alpha with relative RMSE.

In the second approach, the $\hat{\lambda}_1$ values were arbitrarily categorized as "small" when $\hat{\lambda} < 3$, "moderate" when $3 \leq \hat{\lambda} < 6$, and "large" when $\hat{\lambda} \geq 6$. The R-Bias and R-RMSE of the coefficient α estimates were treated separately as dependent variables in the design of the ANOVA. In each analysis, the independent variables were a) sample size, b) number of items (Table 1), and c) level of the $\hat{\lambda}_1$ (*l*). Table 2 presents the results of the three-way ANOVA for R-Bias and R-RMSE, including effect-size estimate (ω) values.

Table 2: Analysis of variance results for the R-Bias and R-RMSE in Monte Carlo study

		Bias as dependent variables				RMSE as dependent variables			
Source	df	SS	F	P	ω	SS	F	P	ω
n	3	3.45	2197.65	0.000	0.142	0.243	3042.95	0.000	0.186
k	15	0.04	5.55	0.000	0.002	0.002	4.84	0.000	0.002
l	2	2.19	2093.86	0.000	0.095	0.179	3359.96	0.000	0.144
*n*x*k*	45	0.08	3.32	0.000	0.004	0.004	3.15	0.000	0.004
*n*x*l*	6	1.69	539.45	0.000	0.075	0.112	697.68	0.000	0.095
*k*x*l*	27	0.31	22.21	0.000	0.015	0.028	39.10	0.000	0.026
*n*x*k*x*l*	81	0.42	9.81	0.000	0.020	0.030	14.07	0.000	0.028
Error	39820	20.81				1.061			

SS: Sum of squares; P: probability; ω: effect size estimations, *n*: Sample size; *k*: Numbers of item; *l*: Level of first eigenvalue

As seen in Table 2, all main effects and interactions were significant. The largest effect-size was of sample size (ω_n = 0.142 for R-Bias and ω_n = 0.186 for R-RMSE), followed by the level of eigenvalue (ω_l = 0.095 for R-Bias and ω_l = 0.144 for R-RMSE).

As for interactions, the largest effect-sizes were observed in the *sample size* x *level of first eigenvalue* interaction (ω_{nxl} = 0.075 for R-Bias and ω_{nxl} = 0.095 for R-RMSE). The other effect-sizes for the interactions and main effects were small, but they were statistically significant. Instead of the post-hoc analysis for interactive effects, the graph of the interaction between sample size and the level of the first eigenvalue for relative bias is given in Figure 4.

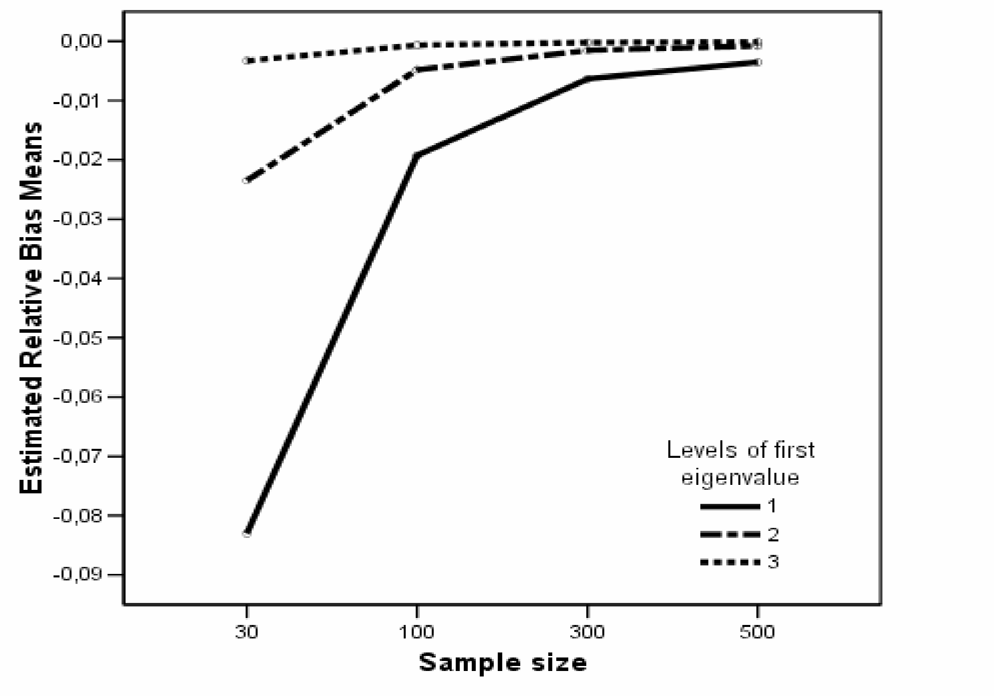

Figure 4: The performance of sample coefficient alpha in levels of first eigenvalue.

According to the simulation results of this study, when $\lambda_1 \geq 6$ (Level 3), the estimates of coefficient alpha are unbiased for each sample size. However, the estimates of coefficient alpha are only unbiased for data sets whose eigenvalue is $3 \leq \lambda_1 < 6$ (Level 2) when $n > 100$. Similarly, the estimates of coefficient alpha are unbiased if sample size is larger than 300 when $\lambda_1 < 3$ (Level 1). On the other hand, when sample size is 500 or larger, the performance of the coefficient alpha estimator behaves independently from the largest eigenvalue of the sample data set. When the sample size is smaller than 500, the performance of the sample coefficient alpha also depends on the first eigenvalue of the data set.

4. CONCLUSION

In literature regarding reliability, necessary minimum sample size for coefficient alpha is commonly suggested as 200, 300, or 500. In this study, minimum sample size for estimating coefficient alpha is dependent on the level of the first (largest) eigenvalue obtained from PCA. If the value of that first eigenvalue of the sample data set is higher than 6.00, the sample coefficient alpha, even when n=30, is an especially robust estimator of the population coefficient alpha. Similarly, if the first eigenvalue is between 3.00 and 6.00, the required minimum n=100 will be adequate for an unbiased estimator of coefficient alpha.

One of the limitations of this study is that the simulation is based on normal data. Yuan and Bentler (2002) reported that the asymptotic distribution for the coefficient alpha estimates, obtained based on a normal sampling distribution, is valid within a large class of non-normal distributions. Given this, the findings of this study can further be extended for non-normal sampling distributions.

REFERENCES

Armor, D. J. (1974). *Theta reliability and factor scaling.* In Costner, H. L. (ed.), Sociological Methodology. Jossey-Bass, San Francisco. 17−50.

Box, G. E. P., & Muller, M. E. (1958). A note on the generation of random normal deviates. *Annals of Mathematical Statistics* 29: 610–611.

Charter, R. A. (1999). Sample Size Requirements for Precise Estimates of Reliability, Generalizability, and Validity Coefficients. *Journal of Clinical and Experimental Neuropsychology*, 21, 559-566.

Charter, R.A. (2003). Study Samples Are Too Small to Produce Sufficiently Precise Reliability Coefficients. *The Journal of General Psychology*, Vol. 130, 117-129.

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology, 78* (1), 98-104.

Crano, W D., & Brewer, M. B. (1973). *Principles of research in social psychology.* New York: McGraw-Hill.

Cronbach, L.J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika,16,* 297-334.

Cronbach, L. J. (2004). My current thoughts on coefficient alpha and successor procedures. *Educational and Psychological Measurement, 64*, 391-418.

Kline, P. (1986). *A handbook of test construction: Introduction to psychometric design*. New York: Methune & Company.

Liu, Y., & Zumbo, B. D. (2007). The Impact of Outliers on Cronbach's Coefficient Alpha Estimate of Reliability: Visual Analogue Scales. *Educational and Psychological Measurement*, *67,* 620-634.

MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods*, *4*, 84-99.

Nunnally, J.C.&Bernstein, I.H. (1994). *Psychometric theory* (3rd ed.). Neew York: McGraw-Hill.

Peterson, R.A. (1994) A meta-analysis of Cronbach's coefficient alpha. *Journal of Consumer Research* 21, 381-391.

Raykov, T. (1997). Scale reliability, Cronbach's coefficient alpha, and violations of essential tau-equivalence with fixed congeneric components. *Multivariate Behavioral Research*, 32, 329-353.

Segall, D. O. (1994). The reliability of linearly equated tests. *Psychometrika, 59,* 361-375.

van Zyl, J. M., Neudecker, H., & Nel, D. G. (2000). On the distribution of the maximum likelihood estimator of Cronbach's alpha. *Psychometrika*, *65*, 271-280.

Yuan, K., & Bentler, P. M. (2002). On robustness of the normal-theory based asymptotic distributions of three reliability coefficient estimates. *Psychometrika, 67(2)*, 251-259.

Yuan, K., Guarnaccia, C. A., & Hayslip, B. (2003). A study of the distribution of the sample coefficient alpha with the Hopkins Symptom Checklist: Bootstrap versus asymptotics. *Educational and Psychological Measurement, 63*, 5-23.

Yurdugul, H. (In press). SIMREL: A software for the coefficient alpha and its confidence intervals with Monte Carlo studies, *Applied Psychological Measurement.* DOI: 10.1177/0146621608322657

GENİŞLETİLMİŞ ÖZET

Eğitimsel ve psikolojik araştırmalarda ölçme araçlarının/sonuçlarının güvenirliği önemli konulardan birisidir. Özellikle birleşik ölçmelerde tek bir teste ilişkin içtutarlılık anlamındaki güvenirliğin elde edilmesinde Cronbach (1951) tarafından geliştirilen alfa katsayısının yaygın bir şekilde kullanıldığı görülmektedir. Öyle ki; Cronbach (1951)'ın orijinal makalesine 2004 yılı itibari ile 5590 kere referans verilmiş (Cronbach, 2004) ve Social Sciences Citation Index kapsamındaki makalelerden ise yılda ortalama 325 kere referanslar verilmektedir (Liu ve Zumbo; 2007). Ancak Peterson (1994)'ın Cronbach alfa katsayısına ilişkin yaptığı meta-analiz çalışmasında bu katsayısının küçük örneklem genişliklerinden elde edildiğini rapor etmiştir. Oysa ki örneklem alfa katsayısının evren alfa katsayısını yansız bir şekilde kestirebilmek için büyük örneklem genişliklerine ihtiyaç duyulduğu ifade edilmektedir (Charter; 2003). Alfa katsayısına ilişkin ideal minimum örneklem genişliği konusunda da ilgili literatürde yapılan öneriler ise; Kline (1986) bu örneklem genişliğini (*n*=) 300 olarak açıklamış ancak Nunnally ve Bernstein (1994) ile Segall (1994) ise *n*=300 genişliğini "küçük" olarak nitelenmdirmişlerdir. Diğer taraftan Charter (1999) ise bu örneklem genişliğinin 400 olması gerektiğini ifade etmiştir. Bu tartışmaların temelinde ise eğitimsel ve psikolojik alanda veri toplama güçlüğü yatmaktadır.

Diğer taraftan; benzer bir tartışma uzun süredir faktör analizi için yapılmıştır. Faktör analizi için yapılan farklı önerilerden sonra MacCallum *vd.* (1999) bu tür mutlak ölçü ile verilen örneklem genişliklerinin bir kavram yanılgısı olduğunu, örneklem genişliklerinin aslında ölçmelerin ortak varyanslarının (communality) bir fonksiyonu olduğunu ifade etmiştir. Öyle ki; kaliteli ölçmelerde (yüksek ortak varyansa sahip ölçmeler) düşük örneklem genişliği yeterli olabilmektedir (MacCallum *vd.*, 1999). Bu çalışma da benzer şekilde aynı tartışmayı güvenirlik hesaplamasında alfa katsayısı için yapmayı amaçlanmıştır. Ancak bu çalışmada MacCallum vd. (1999)'in aksine ortak faktör analizine dayalı olarak değil, alfa katsayısına ilişkin örneklem genişliğinin temel bileşenler analizinden elde edilen en büyük özdeğerin bir fonksiyonu olması üzerine kurulmuştur. Bunun en önemli nedeni ise; aslında güvenirlik katsayılarının özelikle ölçme kümesinin ilk faktörünü ölçmesi (Crano and Brewer 1973) ve diğer taraftan alfa katsayısının ilk döndürülmemiş özdeğerin fonksiyonu olmasından (Cortina 1993) kaynaklanmaktadır.

Çalışmanın amacına uygun olarak, bootstrap yöntemine dayalı Monte-Carlo simülasyonu kullanılmıştır. Yuan vd. (2003) bu tür örneklem genişliği çalışmalarında bootstrap yaklaşımının asimtotik yaklaşıma göre daha geçerli olduğunu ifade etmiştir. Simülasyonlarda her biri N=5000 gözleme sahip D=10000 adet normal dağılıma uygun likert türü veriler elde edilerek bu veri kümeleri evren (population) olarak adlandırılmış ve her biri için evren alfa katsayısı elde edilmiştir. Daha sonra her bir evrenden her bir örneklem genişliğinde *n*=30, 100, 300, 500 basit rastgele örnekleme yöntemi ile 500 örneklem çekilerek *örneklem alfa katsayısı* elde edilmiş ve *evren alfa katsayısı* ile *örneklem alfa katsayısı* arasındaki tutarlılık incelenmiştir. Bu incelemeler göreli yan (relative bias) ve göreli hata (relative RMSE) değerleri ile gözlenmiştir. Burada kestirimlere ilişkin yanlılık ve hata değerleri 0 değerine ne kadar yakın ise örneklem alfa katsayısı, evren alfa katsayısının o denli sağlam kestiricisi (robust estimation) olduğu ifade edilir. Bu çalışmada yanlılık ölçütü olarak 0,01 ve RMSE ölç.ütü ise 0,05 olarak alınmıştır.

Simülasyon sonuçları iki aşamalı olarak incelenmiştir. Bunlardan ilkinde farklı örneklem genişliklerinde alfa kestiricisinin yanlılık ve hata değerlerinin veri kümesinin birinci özdeğerine göre davranışı incelenmiştir. Elde edilen sonuçlara göre; Çizim 1 ve Çizim 2'de görüldüğü gibi; örneklem genişliği en düşük *n*=30 olduğunda eğer birinci özdeğer düşük değerleri ise bu durumda alfa kestiricisi sağlam bir kestirici (robust estimator) ve/veya tutarlı bir kestirici olmadığı görülmektedir. Ancak *n*=30 olmasına rağmen birinci özdeğer artıkça alfa kestiricisi daha sağlam kestirici özelliğine yönelmekte ve daha tutarlı davranmaktadır. Özellikle birinci özdeğer ≥ 10 olduğunda örneklem genişliği *n*=30 olsa bile örneklem alfa katsayısının evren alfa katsayısı ile aynı değeri almakta olduğu gözlenmiştir. Benzer şekilde; birinci özdeğer 1'e doğru yaklaştıkça *n*=30 örneklem enişliğinde çok tutarsız sonuçlar elde edilmiştir. Çizim 1 ve Çizim 2'nin diğer grafiklerinde de görüldüğü gibi; örneklem genişliği arttıkça, özellikle *n*=500 olduğunda tüm alfa kestirimleri yaklaşık *hatasız* olarak elde edilmiştir.

Çalışmanın ilk aşamasında birinci özdeğerin sürekli bir değişken olarak kabul edildiği durumlarda örneklem alfa katsayısının tutarlığı incelenmişti. Çalışmanın ikinci aşamasında ise; birinci özdeğerin değerleri (yorumlamayı kolaylaştırmak ve etkileri/etkileşimleri inceleyebilmek amacıyla) yapay bir şekilde kategorik hale getirilmiş ve örneklem alfa katsayısının tutarlığı üzerindeki çeşitli faktörlerin etkileri incelenmiştir. Bunun için birinci özdeğer a) düşük ($\hat{\lambda}_1 < 3$), b= orta ($3 \le \hat{\lambda}_1 < 6$), ve c) geniş ($\hat{\lambda}_1 \ge 6$) olarak sınıflanmıştır. Buna göre; birinci özdeğer 3 farklı düzeyde (*l*=düşük, orta, geniş), örneklem genişliği (*n*=30, 100, 300,500) 4 farklı düzeyde ve madde sayıları ise 16 (Tablo 1) farklı düzeyde ele alınarak örneklem alfa kestirimlerindeki yanlılık ve RMSE değerlerine göre etkileri varyans analiz (ANOVA) yöntemi ile incelenmiştir (Tablo 2). Elde edilen sonuçlara göre; örneklem alfa katsayısının tutarlı kestirimleri üzerinde test edilen tüm değişkenler istatistiksel olarak anlamlı iken en önemli etki ve etkileşimlerin sırasıyla örneklem genişliği (*n*) ve birinci özdeğer düzeyi (*l*) ve her iki değişken arasındaki etkileşim (*n*x*l*) olduğu görülmüştür. Bu sonuç bir önceki aşamada elde edilen bulguları tamamen desteklemektedir.

Sonuç olarak; örneklem alfa katsayısının evren alfa katsayısının tutarlı bir kestirimi için gerekli olan minimum örneklem genişliği için mutlak bir değer önerilemeyeceği, çünkü sağlam ve tutarlı bir kestirimin örneklem genişliğinin yanı sıra ölçme kümesinin birinci özdeğerine bağlı olduğu ortaya konmuştur. Bununla birlikte yapılan simülasyon sonucunda eğer birinci özdeğer > 6,00 ise *n*=30, birinci özdeğer 3,00 ile 6,00 arasında ise *n*=100, birinci özdeğer<3,00 ise n≥300 örneklem genişlikleri yansız bir alfa katsayısı kestirimi için yeterli olarak gözlenmiştir.

Bu çalışmadaki veri kümesi normal dağılıma uygun olarak üretilmiştir. Yuan ve Bentler (2002) asimptotik yaklaşıma dayalı ve normal veriler üzerinde yaptıkları simülasyon çalışmasında elde ettikleri sonuçların normal olmayan dağılımlar içinde geçerli olduğunu ifade etmişlerdir. Benzer olarak; bu çalışmada elde edilen sonuçlarda normal olmayan veri kümeleri için genişletilebilir.