

Bölüm 4

MADDE TEPKİ KURAMINDA YENİ YAKLAŞIMLAR: PARAMETRİK OLMAYAN MADDE TEPKİ KURAMI

Ezgi MOR DİRLİK¹

Giriş

Psikolojik yapıların ölçülmesinde evrensel tek bir yaklaşım yoktur, bunun yerine birden çok yöntem ve her yöntemin kendine özgü avantaj ve dezavantajları vardır (C-rocker ve Algina, 1986). Yaygın olarak benimsenmiş ve 1900'li yılların başında ortaya çıkmış Klasik Test Kuramı (KTK) psikometri alanına egemen olmuş, halen de bir çok ölçme değerlendirme uygulamasında kullanılmaktadır. KTK'nın uzantısı olarak nitelendirilebilecek varyans analizine dayanan Genellenebilirlik Kuramı da psikolojik yapıların ölçülmesinde kullanılmaktadır, özellikle hata kaynaklarının incelenmesi konusunda alana katkı sağlamaktadır. Madde Tepki Kuramı (MTK) ise KTK'nın barındırdığı bir çok sınırlılığa çözüm üretme amacıyla geliştirilmiştir. Psikometri alanına bir çok yenilik getiren MTK maddeden bağımsız yetenek kestirimi, değişmez madde parametreleri kestirimleri gibi konularda öncüsü KTK'nın dezavantajlarını ortadan kaldırmıştır (Hambleton ve Swaminathan, 1985).

Madde tepki kuramı modelleri ve uygulamaları psikolojik ölçmelerde birçok yeniliğe olanak sağlamış ve karmaşık psikometri problemlerine çözümler üretmiştir. Bireylerden bağımsız olarak kestirilen madde parametreleri ve maddelerden bağımsız olarak kestirilen yetenek parametreleri ile test eşitleme, bireye uyarlanmış test uygulamaları ve madde yanlılığı gibi önemli psikometrik konular bu kuram çerçevesinde araştırılmakta ve uygulamaya dönüştürülmektedir. Bu avantajlarına rağmen Parametrik Madde Tepki Kuramı (PMTK) modellerinin doğasından kaynaklanan doğrulayıcı özelliği modellerin esnekliğini sınırlandırmakta ve gerçek uygulamalarda çeşitli nedenlerden dolayı model veri uyumunun sağlanmasını zorlaştırmaktadır. Yeterli düzeyde sağlanamayan model veri uyumu ise PMTK'nın söz konusu avantajlarını ortadan kaldırmakta, modele uyum sağlamayan maddelerin ölçekten çıkarılmasına neden olmakta ve bu nedenle aynı zamanda geçerliğe tehdit oluşturmaktadır. PMTK'nın sınırlayıcı özelliğine yönelik daha esnek modelleme yaklaşımlarının aranması sonucu ortaya çıkan Parametrik Olmayan Madde Tepki Kuramı (POMTK) ise daha zayıf varsayımlara sahip olması sayesinde daha fazla veri setine uyum sağlayabilmektedir. Doğrulayıcı bir özellikten çok açıklayıcı özellik taşıyan POMTK modelleri yanıt fonksiyonlarının belli bir forma sahip olmasını gerektirmediğinden, verilerin doğal yapısının değişmesine yol açmadan ölçülen gizil değişken hakkında daha gerçekçi bilgiler sağlamaktadır.

¹Dr. Öğretim Üyesi, Kastamonu Üniversitesi, emor@kastamonu.edu.tr

sahip olduğu çalışma kapsamında belirlenmiş ancak Sijtsma ve Molenaar (1987), söz konusu farkların küçük ve pratikte manidar olma olasılığının düşük olduğunu belirtmişlerdir.

Cronbach Alfa katsayısı da POMTK kapsamında incelenen veri setleri için hesaplanabilecek bir diğer güvenilirlik belirleme katsayısıdır. Cronbach Alfa katsayısı, Kuder Richardson katsayısının genel hali olarak tanımlanmakta ve testin olası tüm iki yarılarından kestirilen iki yarı korelasyonlarının ortalaması olarak hesaplanmaktadır (Cronbach, 1951). Guttman (1945) bu katsayının güvenilirliğin sınır değeri olarak kullanılabilirliğini belirtmekte ancak testlerin tek boyutluluk varsayımını karşıladığı durumlarda gerçek güvenilirliğe eş değerler üretilebileceği belirtilmiştir (Feldt ve Qualls, 1996). Söz konusu katsayının POMTK uygulamaları için de güvenilirliğin alt sınır değerine ilişkin bilgi verdiği, dolayısıyla kullanılabilirliği belirtilmektedir (Molenaar ve Sijtsma, 1984). Son olarak Van der Ark, Van der Palm ve Sijtsma (2011) tarafından gizil sınıf analizinde kullanılmak üzere geliştirilen LCRC (Latent Class Reliability Coefficient) katsayısı da POMTK yaklaşımında kullanılabilirliği belirtilmiştir.

Sonuç

Bölüm boyunca tartışılan ve tanıtılan Parametrik Olmayan Madde Tepki Kuramı yaklaşımı gelişmeye devam eden bir alandır. Halen yeni yöntemler önerilmekte, PMTK ile kıyaslanmakta ve ülke çapında da çeşitli çalışmalar yapılarak kuramın gelişmesine ve tanıtılmasına katkı sağlanmaktadır (Avşar Şengül, 2015; Koğar, 2014; Mor Dirlik, 2017). Yapılan uygulamalı ve kuramsal çalışmalardan elde edilen sonuçlar ve yaklaşımın özellikleri incelendiğinde, MTK'nın daha az sayıda kişiden oluşan örneklemelere ve daha az maddeden oluşan veri setlerine uygulanmasını sağlayan bir yaklaşım olduğu sonucuna ulaşılmaktadır. Ancak bu durum parametrik yaklaşımın tamamen parametrik olmayana tercih edilmesi sonucunu doğurmamaktadır. Parametrik yaklaşım doğası gereği daha güçlü kestirimler yapmakta, bireyselleştirilmiş test uygulamaları, test eşitleme gibi birçok psikometri uygulamasına çözüm üretmektedir. Bir diğer ifadeyle, parametrik MTK modelleri parametrik olmayan modellere göre ölçme uygulamalarında daha yaygın olarak kullanılmakta ve daha etkili kestirimler yapabilmektedir. Parametrik olmayan yaklaşım ise PMTK modellerinin sınırlılığı olan zayıf model veri uyumunun ortaya çıkması durumunda, ya da geniş ölçekli uygulamalara gidilemediği durumlarda özellikle de tıp bilimlerindeki gibi hiyerarşik yapıların ölçüldüğü durumlarda oldukça etkili sonuçlar üretmektedir. Ayrıca günden güne gelişen bir yaklaşım olan POMTK'nın ilerleyen yıllarda PMTK'nın kullanıldığı alanların birçoğunda kullanılabilirliği ön görülmektedir.

Kaynakça

Avşar, Şengül. A. (2015). Çok kategorili puanlanan maddelerin psikometrik özelliklerinin farklı test koşullarında parametrik olmayan Madde Tepki Kuramı modellerine göre incelenmesi. Yayınlanmamış doktora tezi. Ankara Üniversitesi Eğitim Bilimleri Fakültesi.

Baker, F. B. and Kim. S.H. (2004). Item response theory: Parameter estimation techniques (2nd Edition). New York: Marcel Dekker.

Christensen, K. and Kreiner, S. (2010). Monte Carlo tests of the Rasch model based on scalability coef-

ficients. *British Journal of Mathematical and Statistical Psychology*, 63, 101-111.

Crocker, L., and Algina, J. (1986). *Introduction classical and modern test theory*. New York: Harcourt Brace Javonovich College Publishers.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 6, 297-334.

Douglas, J. and Cohen, A. (2001). Nonparametric item response function estimation for assessing parametric model fit. *Applied Psychological Measurement*. 25 (3). 234-243.

Embretson, S. E. and Reise, S. P. (2000). *Item response theory for psychologists*. Mahwah, NJ: Erlbaum.

Feldt, L. S. & Qualls, A. L. (1996). Bias in coefficient alpha arising from heterogeneity of test content. *Applied Measurement in Education*, 9 (3), 277-286.

Guttman, L. (1945). A basis for analyzing test-retest reliability. *Psychometrika*, 10, 255-282.

Hambleton, R. K. and Swaminathan, H. (1985). *Item response theory. Principles and applications*. Boston: Kluwer.

Junker, B. (2000). Some topics in nonparametric and parametric IRT. with some thoughts about the future. Carnegie Mellon University: Department of Statistics.

Junker, B. W. and Sijtsma, K. (2001). Nonparametric item response theory in action: an overview of the special issue. *Applied Psychological Measurement*, 25, 211-220.

Koğar, H. (2014). Madde tepki kuramının farklı uygulamalarından elde edilen parametrelerin ve model uyumlarının örneklem büyüklüğü ve test uzunluğu açısından karşılaştırılması. Yayımlanmamış Doktora Tezi. Hacettepe Üniversitesi Eğitim Bilimleri Enstitüsü.

Kujipers, R. E., Van der Ark, L. A. , Croon, M. A. and Sijtsma, K. (2016). Bias in point estimates and standard errors of Mokken's scalability coefficients. *Applied Psychological Measurement*. 40 (5), 331-345.

Lee, Y. S. (2007). A comparison of methods for nonparametric estimation of item characteristic curves for binary items. *Applied Psychological Measurement*. 31. 121-134.

Loevinger, J. (1947). A systematic approach to the construction and evaluation of tests of ability. *Psychological Monographs*. 61 (A).

Loevinger, J. (1948). The technic of homogeneous tests compared with some aspects of "scale analysis" and factor analysis. *Psychological Bulletin*. 45 (6). 507-529.

Meijer, R.R., and Baneke, J.J. (2004). Analyzing psychopathology items: a case for nonparametric item response theory modeling. *Psychological Methods*, 9 (3), 354-368.

Meijer, R. R., and Egberink, I. J. L. (2012). *Investigating Invariant Item Ordering in Personality and Clinical Scales: Some Empirical Findings and a Discussion. Educational and Psychological Measurement*, 72 (4), 589-607. DOI: [10.1177/0013164411429344](https://doi.org/10.1177/0013164411429344)

Meijer, R., Tendeiro, J., and Wanders, R. (2014). [The use of nonparametric item response theory to explore data quality](#). S. P. Reise & D. Revicki (Ed.). *Handbook of item response theory modeling: Applications to typical performance assessment*. In S. P. Reise & D. Revicki (Eds.). *Handbook of item response theory modeling: Applications to typical performance assessment*. (s. 85-110). New York: Routledge.

Mokken, R. J. (1971). A theory and procedure of scale analysis with applications in political research. New York. Berlin: Walter de Gruyter. Mouton.

Mokken, R. J. and Lewis, C. (1982). A nonparametric approach to the analysis of dichotomous item responses. *Applied Psychological Measurement*, 6 (A), 417-430

Molenaar, I. W. (1982). Mokken scaling revised. *Kwantitatieve Methoden*. 3. 145-164.

Molenaar, I.W. ve Sijtsma, K. (1984). Internal consistency and reliability in Mokken's nonparametric item response model. *Tijdschrift voor onderwijsresearch*. 9 (5). 257-268.

Molenaar, I. W. (1991). A weighted Loevinger H coefficient extending Mokken scaling to multicategory items. *Kwantitatieve Methoden*. 37. 97- 117.

Molenaar, I. W. (1997). Nonparametric models for polytomous responses. In W.J. Man der Linden and R.K. Hambleton (Ed.). *Handbook of modern item response theory* (s. 369-380). New York: Springer.

Molenaar, I. W. (2001). Thirty years of nonparametric item response theory. *Applied Psychological Measurement*, 25 (3), 295-299.

Molenaar, I. W., & Sijtsma, K. (2000). *User's manual MSP5 for Windows*. Groningen: iecProGAMMA.

Mor Dirlik, E. (2017). Parametrik ve Parametrik Olmayan Madde Tepki Kuramı Modellerinden Çeşitli Faktörlere Göre Elde Edilen Madde ve yetenek Kestirimlerinin Karşılaştırılması. Yayımlanmamış doktora tezi. Ankara Üniversitesi Eğitim Bilimleri Fakültesi.

Ramsay, J. O. and Abrahamowicz M. (1989). Binomial regression with monotone splines: A psychometric application. *Journal of the American Statistical Association*. 54 (408). 906-915

Ramsay, J. O. (1991). Kernel smoothing approaches to nonparametric item characteristic curve esti-

mation. *Psychometrika*, 56 (4), 611-630.

Ramsay, J. O. (1997). A functional approach to modeling test data. In W. J. Man der Linden and R. K. Hambleton (Ed.). *Handbook of Modern Item Response Theory* (s. 381-394). New York: Springer.

Ramsay, J. O. (2000). *TestGraf* A computer program for nonparametric analysis of testing data. Unpublished manuscript. McGill University.

Roussos, L. A., & Ozbek, O. Y. (2006). Formulation of the DETECT population parameter and evaluation of DETECT estimator bias. *Journal of Educational Measurement*, 43(3), 215-243.

Santor, D. and Ramsay, J. O. (1998). Progress in the technology of measurement: Applications of item response models. *Psychological Assessment*, 10,345-359.

Sijtsma, K., & Molenaar, I. W. (1987). Reliability of test scores in nonparametric item response theory. *Psychometrika*, 52(1), 79-97.

Sijtsma, K., and Molenaar, I. W. (2002). *Introduction to nonparametric item response theory*. Newbury Park, CA: Sage.

Sijtsma, K. and Junker, B.W. (2006). *Item Response Theory: Past Performance. Present Developments and Future Expectations*. *Behaviormetrika*, 1, 75-102.

Sijtsma, K. and Meijer, R. R. (2007). *Nonparametric item response theory*. C. R. Rao. & S. Sinharay (Ed.) *Handbook of statistics26: psychometrics*. (s.719-746). Amsterdam: Elsevier.

Sijtsma, K., Meijer, R.,R. and Van der Ark, L.A. (2011). Mokken scales analysis astime goes by:an update for scalig practitioners. *Personality and Individual Differences*, 50, (31-37).

Stout, W. F. (1987). A nonparametric approach for assessing latent trait unidimensionality. *Psychometrika*. 52. 589-617.

Stout, W. F. (1990). A new item response theory modeling approach with applications to unidimensionality assessment and ability estimation. *Psychometrika*. 55, 293-325.

Stout, W., Habing, B., Douglas, J., Kim, H. R., Roussos, L., & Zhang, J. (1996). Conditional covariance-based nonparametric multidimensionality assessment. *Applied Psychological Measurement*, 20 (4), 331-354.

Straat, J. H., Van der Ark, L.A. and Sijtsma, K. (2014). Minimum sample size requirements for Mokken scale analysis. *Educational and Psychological Measurement*, 74, 809-822.

van der Ark, L. A. , Croon, M.A. and Sijtsma, K. (2008). Mokken scale analysis for dichotomus items using marginal models.*Psychometrica*, 73, 183-208.

van der Ark, L. A., van der Palm, D. W., & Sijtsma, K. (2011). A latent class approach to estimating test-score reliability. *Applied Psychological Measurement*, 35 (5), 380-392.

Van Schuur, W. H. (2003). Mokken scale analysis between the Guttman scale and parametric item response theory. *Political Analysis*. 11, 139-163.

Van Schuur, W. H. (2011). *Ordinal item response theory: Mokken scale analysis*. Los Angeles: Sage Publications.

Watson, R., Dreary, I. And Shipley, b. (2008). A hierarchy of distress:Mokken scaling of the GHQ-30. *Psychological Medicine*, 38 (4), 575-579.