Examination of the Effects of Different Missing Data Techniques on Item Parameters Obtained by CTT and IRT

Examination of the Effects of Different Missing Data Techniques on Item Parameters Obtained by CTT and IRT

,Ufuk Akbaş

Özet

The purpose of this study is to examine the effect of different missing data techniques on the item parameters estimated for Classical Test Theory (CTT) and Item Response Theory (IRT) comparatively through simulated and real data sets. For this purpose, data sets with missing completely at random pattern with different sample sizes and missing data rates have been generated. Item parameters based on CTT and IRT are estimated after applying different missing data techniques (listwise deletion, regression imputation and expectation – maximization). Estimated parameters were compared with the parameters of complete data sets. The same procedure was performed on the data of the PM3 subtest in the PISA 2012 application. It is found that higher difficulty and lower discrimination values are obtained with listwise deletion while accurate item difficulties can be obtained with regression imputation and expectation - maximization algorithm, but the same situation is not valid for item discriminations. It is also seen that missing data at a ratio of %2 can lead to serious problems even if they have missing completely at random pattern.

Anahtar Kelimeler

null

Abstract

The purpose of this study is to examine the effect of different missing data techniques on the item parameters estimated for Classical Test Theory (CTT) and Item Response Theory (IRT) comparatively through simulated and real data sets. For this purpose, data sets with missing completely at random pattern with different sample sizes and missing data rates have been generated. Item parameters based on CTT and IRT are estimated after applying different missing data techniques (listwise deletion, regression imputation and expectation – maximization). Estimated parameters were compared with the parameters of complete data sets. The same procedure was performed on the data of the PM3 subtest in the PISA 2012 application. It is found that higher difficulty and lower discrimination values are obtained with listwise deletion while accurate item difficulties can be obtained with regression imputation and expectation - maximization algorithm, but the same situation is not valid for item discriminations. It is also seen that missing data at a ratio of %2 can lead to serious problems even if they have missing completely at random pattern.

Keywords

null

References

Akbaş, U. (2014). Farklı örneklem büyüklüklerinde ve kayıp veri örüntülerinde ölçeklerin psikometrik özelliklerinin kayıp veri baş etme teknikleri ile incelenmesi. Yayınlanmamış doktora tezi, Ankara Üniversitesi, Eğitim Bilimleri Enstitüsü, Ankara.

Allison, P. D. (2002). Missing data. Newbury Park, CA: Sage.

Alpar, R. (2011). Uygulamalı çok değişkenli istatistiksel yöntemler. Ankara: Detay Yayıncılık.

Bal, C. (2003). Çok gruplu veri setlerinde eksik gözlem sorununun çözümlenmesi ve sağlık alanında bir uygulama. Yayınlanmamış doktora tezi, Osmangazi Üniversitesi, Sağlık Bilimleri Enstitüsü, Eskişehir.

Bernaards, C. A. and Sijtsma, K. (1999). Factor analysis of multidimensional polytomous item response data suffering from ignorable item nonresponse, Multivariate Behavioral Research. 34(3), 277 – 313.
doi:10.1207/S15327906MBR3403_1

Tam metne ulaşmak ve tüm referansları görmek için tıklayın.