Applied Psychological Measurement

CopyDetect : An R Package for Computing Statistical Indices to Detect Answer Copying on Multiple-Choice Examinations

Cengiz Zopluoglu

The online version of this article can be found at:
http://apm.sagepub.com/content/37/1/93

Published by:
$SAGE

http://www.sagepublications.com

Additional services and information for ***Applied Psychological Measurement*** can be found at:

Email Alerts: http://apm.sagepub.com/cgi/alerts

Subscriptions: http://apm.sagepub.com/subscriptions

Reprints: http://www.sagepub.com/journalsReprints.nav

Permissions: http://www.sagepub.com/journalsPermissions.nav

Citations: http://apm.sagepub.com/content/37/1/93.refs.html

>> Version of Record - Dec 5, 2012

OnlineFirst Version of Record - Oct 12, 2012

What is This?

CopyDetect: An R Package for Computing Statistical Indices to Detect Answer Copying on Multiple-Choice Examinations

Cengiz Zopluoglu[1]

Keywords

answer copying, answer exchange, cheating, data forensics, test fraud, test security, multiple-choice tests, test integrity, score integrity

Fraud on standardized tests has been an increasing concern (Crouch, 2012; Hildebrand, 2012) because it invalidates the inferences made from test scores. The large-scale, self-report survey results consistently revealed that about 35% of high school students engaged in some type of test fraud two or more times in the previous year (Josephson Institute of Ethics, 2006, 2008, 2010), and answer exchange between two examinees is a type of test fraud commonly observed in multiple-choice examinations (Bopp, Gleason, & Misicka, 2001; Brimble & Clarke, 2005; Hughes & McCabe, 2006; Rakovski & Levy, 2007).

Identifying answer copying is an essential part of maintaining the integrity of test scores, and additional evidence is always necessary when a pair of examinees is suspected of exchanging answers on a multiple-choice test. Many scholars have developed a variety of analytical procedures and addressed the issue from a statistical perspective to provide additional evidence of answer copying between two examinees (Angoff, 1972; Bay, 1995; Bellezza & Bellezza, 1989; Cody, 1985; Frary, Tideman, & Watts, 1977; Hanson, Harris, & Brennan, 1987; Holland, 1996; Saupe, 1960; Sotaridona & Meijer, 2002, 2003; van der Linden & Sotaridona, 2006; Wollack, 1996). However, few of these indices have been shown to be effective and reliable based on the results from simulation studies (Bay, 1995; Hanson et al., 1987; Sotaridona & Meijer, 2002, 2003; Wollack, 1996, 2003, 2006; Wollack & Cohen, 1998; Zopluoglu & Davenport, 2012).

The simulation studies showed that the ω index (Wollack, 1996) and generalized binomial test (GBT; van der Linden & Sotaridona, 2006) provide the highest detection rates, while holding the empirical Type I error rates below the nominal level. In addition, the K index (Holland, 1996), K_1 and K_2 (Sotaridona & Meijer, 2002), and S_1 and S_2 indices (Sotaridona & Meijer, 2003) have provided reasonable detection rates and held Type I error rates below the nominal level in simulation studies. Although much effort has been put into developing statistical indices for detecting answer copying, so far little effort has been made to develop accessible software for practitioners to compute statistical indices that have been found effective in the literature. To my knowledge, none of these useful indices are available for practitioners in any

[1]University of Minnesota, St Paul, MN, USA

Corresponding Author:
Cengiz Zopluoglu, Department of Educational Psychology, University of Minnesota, 193 Education Sciences Building, 56 East River Road, Minneapolis, MN 55455-0364, USA
Email: zoplu001@umn.edu

accessible software, either commercial or otherwise, for detecting answer copying (Integrity, 2005; ScanexII, 2008; Scrutiny!, 2005).

An R package, CopyDetect, is now available to compute the ω, GBT, K, K_1, K_2, S_1, and S_2 indices. The package has two separate functions to compute these indices for dichotomous and nominal response outcomes. The functions require the raw data matrix with either dichotomous or nominal response outcomes, IRT model parameters (1PL, 2PL, 3PL, and nominal response IRT models), and the row numbers of suspected examinees in the data file. If dichotomous responses are provided and users are not able to obtain dichotomous IRT model item parameters from external software, the package internally uses the ltm engine (Rizopoulos, 2006) for estimating 2PL IRT model item parameters and uses them as input. However, if nominal responses are provided, the nominal IRT model item parameters must be obtained from external software and provided to the functions.

The CopyDetect package is available at no charge from the website http://www.r-project.org and works under Windows, Linux, and MacOS platforms. A reference manual is also available from the website. Version 2.14.1 (or later) of the R software (R Development Core Team, 2012) should be installed to be able to use the CopyDetect package.

Acknowledgments

Declaration of Conflicting Interests

Funding

References

Angoff, W. (1972). The *development of statistical indices for detecting cheaters*. Berkeley, CA: Educational Testing Service.

Bay, L. (1995). *Detection of cheating on multiple-choice examinations*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

Bellezza, F. S., & Bellezza, S. F. (1989). Detection of cheating on multiple-choice tests by using error similarity analysis. *Teaching of Psychology, 16*, 151-155.

Bopp, M., Gleason, P., & Misicka, S. (2001). *Reducing incidents of cheating in adolescence* (Masters project). Saint Xavier University, Chicago, IL.

Brimble, M., & Clarke, P. S. (2005). Perceptions of the prevalence and seriousness of academic dishonesty in Australian universities. *Australian Educational Researcher, 32*, 19-44.

Cody, R. P. (1985). Statistical analysis of examinations to detect cheating. *Journal of Medical Education, 60*, 136-137.

Crouch, E. (2012, April 4). School test fraud is target of Missouri bill. *St. Louis Today*. Retrieved from http://www.stltoday.com/news/local/education/school-test-fraud-is-targetof-missouri-bill/article_ab9538 6b-48c2-50ef-b5c6-e110a445019b.html

Frary, R. B., Tideman, T. N., & Watts, T. M. (1977). Indices of cheating on multiple choice tests. *Journal of Educational Statistics, 2*, 235-256.

Hanson, B. A., Harris, D. J., & Brennan, R. L. (1987). *A comparison of several statistical methods for examining allegations of copying*. (ACT Research Report No. 87-15). Iowa City, IA: American College Testing.

Hildebrand, J. (2012, March 15). State creates unit to combat test fraud. *Newsday*. Retrieved from http://www.newsday.com/long-island/education/state-creates-unit-to-combat-test-fraud-1.3605217

Holland, P. W. (1996). *Assessing unusual agreement between the incorrect answers of two examinees using the K-Index: Statistical theory and empirical support* (ETS Research Report No. 96-97). Princeton, NJ: Educational Testing Service.

Hughes, J. M. C., & McCabe, D. L. (2006). Academic misconduct within higher education in Canada. *Canadian Journal of Higher Education*, *36*(2), 1-21.

Integrity. (2005). Castle Rock Research Corporation. Retrieved from http://www.integrity.castlerockresearch.com/integrity.html

Josephson Institute of Ethics. (2006). *Ethics of American youth*. Retrieved from http:// charactercounts.org/pdf/reportcard/2006/reportcard-all.pdf

Josephson Institute of Ethics. (2008). *The ethics of American youth*. Retrieved from http://charactercounts.org/pdf/reportcard/2008/Q_all.pdf

Josephson Institute of Ethics. (2010). *The ethics of American youth*. Retrieved from http://charactercounts.org/pdf/reportcard/2010/ReportCard2010_data-tables.pdf

Rakovski, C. C., & Levy, E. S. (2007). Academic dishonesty: Perceptions of business students. *College Student Journal, 41*, 466-481.

R Development Core Team. (2012). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. ISBN 3-900051-07-0. Available from http://www.R-project.org/

Rizopoulos, D. (2006). ltm: An R package for latent variable modeling and item response theory analyses. *Journal of Statistical Software, 17*(5), 1-25. URL http://www.jstatsoft.org/v17/i05/

Saupe, J. (1960). An empirical model for the corroboration of suspected cheating on multiple choice tests. *Educational and Psychological Measurement, 20*, 475-489.

Sotaridona, L. S., & Meijer, R. R. (2002). Statistical properties of the k-index for detecting answer copying. *Journal of Educational Measurement, 39*, 115-132.

Sotaridona, L. S., & Meijer, R. R. (2003). Two new statistics to detect answer copying. *Journal of Educational Measurement, 40*, 53-69.

ScanexII. (2008). University of Western Ontario. Retrieved from the World Wide Web: http://ssnds.uwo.ca/softwaredownloads.asp

Scrutiny! (2005). Assessment Systems Corporation. Retrieved from the World Wide Web: http://assess.com/xcart/product.php?productid=242

van der Linden, W. J., & Sotaridona, L. (2006). Detecting answer copying when the regular response process follows a known response model. *Journal of Educational and Behavioral Statistics, 31*, 283-304.

Wollack, J. A. (1996). Detection of answer copying using item response theory. *Dissertation Abstracts International, 57/05*, 2015.

Wollack, J. A. (2003). Comparison of answer copying indices with real data. *Journal of Educational Measurement, 40*, 189-205.

Wollack, J. A. (2006). Simultaneous use of multiple answer copying indexes to improve detection rates. *Applied Measurement in Education, 19*, 265-288.

Wollack, J. A., & Cohen, A. S. (1998). Detection of answer copying with unknown item and trait parameters. *Applied Psychological Measurement, 22*, 144-152.

Zopluoglu, C., & Davenport, E.C, Jr. (2012). The empirical power and type I error rates of the GBT and ω indices in detecting answer copying on multiple-choice tests. *Educational and Psychological Measurement.* Advance online publication. doi:10.1177/001316441244294.