Electronic Journal of Statistics > Vol. 3 (2009) open journal systems

Testing for independence: Saddlepoint approximation to associated permutation distributions

Ehab F. Abd-Elfattah, *Ain Shams University, Cairo, Egypt*

Abstract

One of the most popular class of tests for independence between two random variables is the general class of rank statistics which are invariant under permutations. This class contains Spearman's coefficient of rank correlation statistic, Fisher-Yates statistic, weighted Mann statistic and others. Under the null hypothesis of independence these test statistics have a permutation distribution that is usually approximated by using asymptotic normal theory to determine p-values for these tests. In this note we suggest using a saddlepoint approach that is almost exact and needs no simulations in order to calculate the p-value for tests in this class.

AMS 2000 subject classifications: 62G10, 62G32, 62E17

Keywords: Independence tests, linear rank test, permutation distribution, saddlepoint approximation

Full Text: PDF

Abd-Elfattah, Ehab F., Testing for independence: Saddlepoint approximation to associated permutation distributions, Electronic Journal of Statistics, 3, (2009), 625-632 (electronic). DOI: 10.1214/09-EJS371.

References

Abd-Elfattah, E.F. and Butler, R. (2007). The Weighted Log-Rank Class of Permutation Tests: P-values and Confidence Intervals Using Saddlepoint Methods, Biometrika 94, 3, 543–551. MR2410007

Abd-Elfattah, E.F. and Butler, R. (2009). Log-Rank permutation tests for trend: Saddlepoint p-values and survival rate confidence intervals. Canadian Journal of Statistics 37, 1, 5–16.

Booth, J.G. and Butler, R.W. (1990). Randomization distributions and saddlepoint approximations in generalized linear models. Biometrika 77, 787–796. MR1086689

Broyden, C.G. (1965). A class of methods for solving nonlinear simultaneous equations. Mathematics of Computation 19, 577–593. MR0198670

Burden, R.L. and Faires, J.D. (2003). Numerical analysis. 7th edition, Brooks Cole.

Butler, R.W. (2005). Saddlepoint Approximations with Applications. Cambridge University Press. MR2357347

Cuzick, J. (1982). Rank tests for association with right censored data. Biometrika 69, 2, 351–364. MR0671973

Daniels, H.E. (1958). Discussion of paper by D.R. Cox. Journal of Royal Statistical Society B 20, 236–238.

Davison, A.C. and Hinkley, D.H. (1988). Saddlepoint approximations in resampling method. Biometrika 75, 3, 417–431. MR0967581

Davison, A.C. and Wang, S. (2002). Saddlepoint approximations as smoothers. Biometrika 89, 933–938. MR1946521

Dennis, J.E. and Schnabel, R.B. (1996). Numerical methods for unconstrained optimization and nonlinear equations. SIAM, Philadelphia. MR1376139

Gibbons, J.D. and Chakraborti, S. (2003). Nonparametric statistical inference. 4th edition, Marcel Dekker, New York. MR2064386

Gieser, P.W. and Randles, R.H. (1997). A nonparametric test of independence between two vectors. Journal of the American Statistical Association 92, 438, 561–567. MR1467849

Hajek, J., Sidak, Z. and Sen, P.K. (1999). Theory of Rank Tests. 2nd Ed. Academic Press. MR1680991

Kustra, R., Shi, X., Murdoch, D., Greenwood, C.M. and Rangrej, J. (2008). Efficient p-value estimation in massively parallel testing problems. Biostatistics 9, 4, 601–612.

Lin, D.Y. (2005). An efficient Monte Carlo approach to assessing statistical significance in genomic studies. Bioinformatics 21, 6, 781–787.

Nayak, T.K. (1988). Testing equality of conditionally independent exponential distributions. Communications in Statistics: theory and methods 17, 807–820. MR0939644

Oakes, D. (1982). A concordance test for independence in the present of censoring. Biometrics 38, 451–455.

O'Brien, P. (1978). A nonparametric test for association with censored data. Biometrics 34, 243–250.

Pierce, D.A. and Peters, D. (1992). Practical use of higher order asymptotics for multiparameter exponential families. Journal of Royal Statistical Society B 54, 701–737. MR1185218

Robinson, J. (1982). Saddlepoint approximations for permutation tests and confidence intervals. Journal of Royal Statistical Society B 44, 1, 91–101. MR0655378

Seaman, S.R. and Müller-Myhsok, B. (2005). Rapid simulation of values for product method and multiple-testing adjustment in association studies. American Journal of Human Genetics 76, 399–408.

Skovgaard, I.M. (1987). Saddlepoint expansions for conditional distributions. Journal of Applied Probability 24, 875–87. MR0913828

Wei, L.J. (1980). A generalized Gehan and Gilbert test for paired observations that are subject to arbitrary right censorship. Journal of the American Statistical Association 75, 371, 634–637. MR0590693