Nonparametric Tests for Independence in R

From Herwartz & Maxand (2018), tests for different types of dependence structures:

1. **Bivariate dependence:** Testing for dependence between two random variables x_1 and x_2. The corresponding null hypothesis is $H_0 : F_{x_1,x_2}(x_1, x_2) = F_{x_1}(x_1)F_{x_2}(x_2)$ with joint distribution function F_{x_1,x_2} and marginals F_{x_1}, F_{x_2}.

2. **Groupwise dependence:** Analyzing two sets of variables can be thought of as a generalization of bivariate dependence tests where two disjoint subsets of $\{x_1, \ldots, x_p\}$ are subjected to testing, i.e., $\boldsymbol{x}_1 \in \mathbb{R}^{p_1}$ and $\boldsymbol{x}_2 \in \mathbb{R}^{p_2}$ such that $p_1 + p_2 = p$. The corresponding null hypothesis is $H_0 : F_{\boldsymbol{x}_1,\boldsymbol{x}_2}(\boldsymbol{x}_1, \boldsymbol{x}_2) = F_{\boldsymbol{x}_1}(\boldsymbol{x}_1)F_{\boldsymbol{x}_2}(\boldsymbol{x}_2)$ for multivariate distribution functions $F_{\boldsymbol{x}_1,\boldsymbol{x}_2}$, $F_{\boldsymbol{x}_1}$ and $F_{\boldsymbol{x}_2}$. Furthermore, some tests allow to diagnose the dependence between more than two disjoint subsets, where $p_1 + \ldots + p_c = p$ and $c > 2$.

3. **Mutual dependence:** To test for overall independence within a set of random variables $\{x_1, \ldots, x_p\}$ the null hypothesis is formulated as $H_0 : F_{x_1,\ldots,x_p}(x_1, \ldots, x_p) = F_{x_1}(x_1) \cdots F_{x_p}(x_p)$. The tests exploit the fact that mutual independence is equivalent to independence within all subsets of $\{x_1, \ldots, x_p\}$. This hypothesis is equivalent to stating groupwise independence and choosing subsets of size $p_1 = p_2 = \ldots = p_c = 1$.

Table 1: R packages and functions corresponding to the procedures described in Section 3 of Herwartz & Maxand (2018).

	classical	spatial rank	empirical copula	distance covariance
R package	*Hmisc* (Harrell, 2015)	*SpatialNP* (Sirkia et al., 2013)	*copula* (Hofert et al., 2015)	*energy* (Rizzo & Szekely, 2014) *steadyICA* (Risk et al., 2015)
R function for distinct dependence levels				
bivariate	`hoeffd` (d)	`sr.indep.test` (sr)	`indepTest` (B)	`indep.test` ($dCov$)
mutual	Wilks' Lambda (L)	Fisher comb. of sr	`indepTest` (W, B)	`permTest` ($dCov$)
groupwise	Wilks' Lambda (L)	`sr.indep.test` (sr)	`multIndepTest` (B)	`indep.test` ($dCov$)

References

Harrell, F. E. (2015). *Hmisc*. R package version 3.15-0.

Herwartz, H. & Maxand, S. (2018). Nonparametric tests for independence - a review and comparative simulation study with an application to malnutrition data in india. *Statistical Papers*.

Hofert, M., Kojadinovic, I., Maechler, M., & Yan, J. (2015). *copula: Multivariate Dependence with Copulas*. R package version 0.999-13.

Risk, B. B., James, N. A., & Matteson, D. S. (2015). *steadyICA: ICA and Tests of Independence via Multivariate Distance Covariance*. R package version 1.0.

Rizzo, M. L. & Szekely, G. J. (2014). *energy: E-statistics (energy statistics)*. R package version 1.6.2.

Sirkia, S., Miettinen, J., Nordhausen, K., Oja, H., & Taskinen, S. (2013). *SpatialNP: Multivariate nonparametric methods based on spatial signs and ranks*. R package version 1.1-1.