Criteria for testing equality of multivariate scatters.

Ellipse area covered by the recorded COP traces is a common index of individual overall postural performance [1]. After calculating the area of each participant, it can be used as dependent variable within appropriate inferential procedures testing the significance of the difference among groups (e.g. ANOVA). In this study, individuals COP (and COM) positions were gathered together as a group specific three-dimensional scatter, resumable by a sample covariance matrix. In this connection, it will be necessary to define a scalar measure of group's data spread for testing equality of several covariance matrices. We considered the sample Generalized Variance (GV) recurring in many Likelihood Ratio criteria for hypothesis testing [2].

Given an unbiased sample covariance matrix of COP (and COM) three-dimensional positions $S_i = \begin{bmatrix} s_x^2 & s_{xy} & s_{xz} \\ s_{xy} & s_y^2 & s_{yz} \\ s_{xz} & s_{yz} & s_z^2 \end{bmatrix}$, with eigenvalues $\lambda_1 > \lambda_2 > \lambda_3$, GV_i is defined as the determinant of S_i , formally $|S_i| = \lambda_1\lambda_2\lambda_3$. The volume of 95% *prediction ellipsoid* is given by 4/3πabc, where a = $\sqrt[2]{\chi^2_{3,0.95}\lambda_1}$, b = $\sqrt[2]{\chi^2_{3,0.95}\lambda_2}$, c = $\sqrt[2]{\chi^2_{3,0.95}\lambda_3}$ are semi-axes, $\chi^2_{3,0.95} = 7{,}815$ is the 95th percentile of a Chi-square distribution with 3 degrees of freedom, and can be conveniently expressed as

$$vpe^{95\%} = \frac{4}{3}\pi\left(\chi^2_{3,0.95}\right)^{\frac{3}{2}}(\lambda_1\lambda_2\lambda_3)^{\frac{1}{2}}. \qquad (1)$$

For intergroup (and inter-studies) comparison purposes, if the relative volume of the ellipsoid is considered essential, it is possible to remove the inferential part of the equation, setting $\chi^2 = 1$, and leading to a so called *standard prediction ellipsoid*

$$vpe^{st} = \frac{4}{3}\pi GV_i^{\frac{1}{2}}, \qquad (2)$$

that is completely specified by group's specific GV_i and that unlike the 95% prediction ellipsoid, does not largely depend on the sample size N_i, and on specific distributional assumptions [3,4]. In this connection, GV is a standard measure of multivariate scatter of the data, and can be used to conduct between groups inferential tests directly with a Chi-square log-Likelihood Ratio Test [2,5]

$$\log\_LRT = -2\rho\left[log\left(GV_1^{\frac{n_1}{2}}\right) + log\left(GV_2^{\frac{n_2}{2}}\right) - log\left(\left|\frac{n_1S_1 + n_2S_2}{n_1 + n_2}\right|^{\frac{n_1+n_2}{2}}\right)\right], \quad (3)$$

where $n_i = N_i - 1$ are degrees of freedom of sample variances.

Asymptotically and under strong normality assumption, when the null hypothesis of equality of scatters parameters between groups is true, Eq. (3) is distributed as a complex Chi-square distribution

$$P(\log\_LRT \leq z) = P(\chi_6^2 \leq z) + \omega_2[P(\chi_{6+4}^2 \leq z) - P(\chi_6^2 \leq z)], \quad (4)$$

with technical details on how to calculate ω_2 and ρ that can be found elsewhere [2,6]. The important point to stress here is that this testing procedure is very sensitive to outliers besides its sensitivity to the distribution, and its robustification - replacing group's covariance matrix S_i by robust estimators of multivariate scatter - is challenging [5,7]. Multivariate outliers can be hard to detect because we can no longer rely on visual inspection, and using a weighted Euclidean (Mahalanobis) distance suffers from the well-known masking effect [8]. We approached the problem using a classical Minimum Volume Ellipsoid (MVE) algorithm for detection of multidimensional outlier [9], defining the smallest ellipsoid that encloses all the "good" data, leaving all the bad observations outside the ellipsoid itself, and conducting groups comparison with Eq. (3,4) on selected robust subsets ($N_{85\%}$ and $N_{95\%}$) besides raw data ($N_{100\%}$).

References

[1] Paillard T, Noé F. Techniques and Methods for Testing the Postural Function in Healthy and Pathological Subjects. BioMed Research International 2015;2015.

[2] Anderson TW. An Introduction to Multivariate Statistical Analysis. John Wiley & Sons; 2003.

[3] Schubert P, Kirchner M. Ellipse area calculations and their applicability in posturography. Gait & Posture 2014;39:518–22. doi:10.1016/j.gaitpost.2013.09.001.

[4] Rocchi M, Sisti D, Ditroilo M, Calavalle A. The misuse of the confidence ellipse in evaluating statokinesigram. Ital J Sport Sci 2005.

[5] Aslam S, Rocke DM. A robust testing procedure for the equality of covariance matrices. Computational Statistics & Data Analysis 2005;49:863–74. doi:10.1016/j.csda.2004.06.009.

[6] Muirhead RJ. On the Distribution of the Likelihood Ratio Test of Equality of Normal Populations. The Canadian Journal of Statistics / La Revue Canadienne De Statistique 1982;10:59–62. doi:10.2307/3315076?ref=search-gateway:ada83a9d39d0ef35868d4b363604ec72.

[7] Nordhausen K, Tyler DE. A cautionary note on robust covariance plug-in methods. Biometrika 2015;102:573–88. doi:10.1093/biomet/asv022.

[8] Rousseeuw PJ, van Zomeren BC. Unmasking Multivariate Outliers and Leverage Points. Journal of the American Statistical Association 1990;85:633–9. doi:10.2307/2289995?ref=search-gateway:de7ff480d0a8890ff902f4b643ab29ec.

[9] Van Aelst S, Rousseeuw P. Minimum volume ellipsoid. Wiley Interdisciplinary Reviews: Computational Statistics 2009;1:71–82. doi:10.1002/wics.19.