LETTER

Statistical errors and asymmetry indices

Savic and Lindström (1) claim to have discovered significant differences in cerebral asymmetry and functional connectivity between homo- and heterosexual subjects. Unfortunately, the statistical analysis they present is strongly undermined by an inappropriate choice of variables. For example, in their analysis of cerebral asymmetry, the authors measure right and left cerebral volumes (R and L, respectively) and use these to construct an "asymmetry index" for each subject defined by $V = (R - L)/(R + L) \equiv X/Y$; they call this index AI. No details about the distribution of R and L are given in ref. 1, but if we assume that they are (correlated) bivariate normal variables then $X = R - L$ and $Y = R + L$ are in fact uncorrelated normal variables for which the exact ratio distribution is known (2) to have very long "tails." This is a general problem with ratio distributions. Standard statistical methods, which assume normal distributions with rapidly decreasing tails, are not generally appropriate for this type of variable. Unfortunately, Savic and Lindström (1) seem unaware of this difficulty and proceed to use a standard t test and ANOVA: the P values emerging from this analysis are meaningless if the variables involved are ill conditioned.

One way to work around this problem is to use a different variable, such as $U = (\mu_x - \mu_y V)/(\sigma_x^2 + \sigma_y^2 V^2)^{1/2}$, which is normally distributed (2) if X and Y are independent normal variables with means μ_x and μ_y and variances σ_x^2 and σ_y^2. Standard methods applied to U would give much more conclusive results than those obtained by using V. An alternative is to use the median instead of the mean (3).

Either way, a much more convincing statistical analysis is needed to establish whether the differences reported in ref. 1 are indeed significant.

Peter Coles[1]
School of Physics and Astronomy, Cardiff University, Queen's Buildings, 5 The Parade, Cardiff CF24 3AA, United Kingdom

1. Savic I, Lindström P (2008) PET and MRI show differences in cerebral asymmetry and functional connectivity between homo- and heterosexual subjects. *Proc Natl Acad Sci USA* 105:9403–9408.
2. Geary RC (1930) The frequency distribution of the quotient of two normal variates. *J R Stat Soc* 93:442–446.
3. Brody JP, Williams BA, Wold BJ, Quake SR (2002) Significance and statistical errors in the analysis of DNA microarray data. *Proc Natl Acad Sci USA* 99:12975–12978.

Author contributions: P.C. wrote the paper.
The author declares no conflict of interest.
[1]To whom correspondence should be addressed. E-mail: peter.coles@astro.cf.ac.uk.
© 2008 by The National Academy of Sciences of the USA