CORRESPONDENCE

Age, Period and Cohort Effects

From: Harvey Goldstein
       Department of Mathematics, Statistics & Computing,
       University of London Institute of Education,
       20 Bedford Way,
       London WC1H 0AL, England.

To the Editor of Biometrics:

A quick glance at Equations (1) and (2) in the article by James and Segal (*Biometrics* **38**, 433–443, June 1982) reveals the curious fact that the subscripts on the left-hand side, i, j, do not balance with the subscripts on the right-hand side, i, j, k. Of course, as the authors note, $k = i - j + n$, where i is age, j year and n is a constant, with k known as 'cohort' or year of birth. In particular, (2) can be written

$$\log\{E(Y_{ij}/n_{ij})\} = \alpha_i + \gamma_{i-j+n} + \delta_i\beta_j.$$

While they justify their model partly by reference to earlier work using age, year and cohort models, James and Segal fail to point out that a great deal of the literature on this topic in the social-science area, explicitly or implicitly, has used the third subscript k on the left-hand side of the equations. The epidemiological literature on the other hand has tended to use the same formulation as James and Segal. This incorporation of a strictly redundant subscript has often created confusion and resulted in misleading interpretations; see Goldstein (1979) for a discussion.

REFERENCE

Goldstein, H. (1979). Age, period and cohort effects—a confounded confusion. *Bias* **6**, 19–24.

Dr James replied as follows:

I agree with Professor Goldstein that the literature contains many instances where unwarranted faith has been placed on interpretations of age, period and cohort effects in models containing quite arbitrary restrictions to overcome estimability problems, though I am not sure that it is simply the use of three related subscripts which has led to any misleading interpretations. Our paper was motivated by the contention that too many people find it convenient to assume linear or log-linear models without consideration of the types of underlying processes giving rise to such models. At least in some epidemiological applications we believe we showed that reasonable underlying hazard-function assumptions can lead quite naturally to age–period–cohort models and offer insight into the interpretation of certain combinations of parameters, even though in some cases the combinations of interest may not be estimable.

Ian James
Department of Mathematics,
University of Western Australia,
Nedlands, Western Australia 6009.