



Quantifying the within-group contribution to the variability of count traits

Izabela R. C. Oliveira

ESALQ, University of São Paulo, Piracicaba, Brazil - izabela.rco@gmail.com

Geert Molenberghs

I-BioStat, Universiteit Hasselt, Hasselt, Belgium - geert.molenberghs@uhasselt.be

Clarice G.B. Demétrio*

ESALQ, University of São Paulo, Piracicaba, Brazil - clarice.demetrio@usp.br

Carlos T. S. Dias

ESALQ, University of São Paulo, Piracicaba, Brazil - ctsdias@usp.br

Cláudio L. Souza

ESALQ, University of São Paulo, Piracicaba, Brazil - clsouza@usp.br

Abstract: Heritability and repeatability are important concepts in animal and plant breeding and are quantified based on fitting a model to hierarchical data. When linear models can be used to fit to the data, these attributes are defined as ratios of variance components. Matters are less simple for non-Gaussian outcomes. The focus here is on count outcomes where extensions of the Poisson model are used to describe the data. Expressions for heritability of count traits are derived using the Poisson combined model, which combines a Poisson outcome distribution with normal as well as gamma random effects, to capture both correlation among repeated observations as well as overdispersion, and admits closed-form expressions for the mean, variances and, hence, ratio of variances. The proposed methodology is illustrated using data from plant breeding programs. In the potato study, we have used our methodology to calculate the heritability for the count trait number of large tubers per plot. We considered two scenarios, with and without covariates. In the latter one, we showed how the combined model plays an important role in accommodating extra variability besides the intra-cluster correlation. In the tomato study, we have extended our approach to calculate the repeatability on trichomes count. We want to reiterate that, in these models, heritability is a function rather than a constant. At first sight, this is a drawback. However, it is a consequence from the mean-variance relationship in the models considered. If the model fits the data well, it can also be claimed to be a feature of the data. Practically, heritability and repeatability change with the effects present in the predictor functions. Evidently, one can summarize the functions in a variety of ways, using averages, medians, quartiles, ranges, etc.

Keywords: Combined model; Gamma distribution; Generalized linear mixed model; Overdispersion; Poisson distribution; Random effect.