A bivariate zero-inflated count data regression model with an unrestricted correlation

Shiferaw Gurmu and John Elder

Georgia State University and North Dakota State University

Appendix A: Correlation Coefficient (Not for Publication)

Using iterated expectations, the second order conditional moments for the BIVARZIPL model can be obtained as follows. The conditional variance for the first count variable is

$$\operatorname{Var}(y_{1i} \mid x_i, z_i) = \pi_i \theta_{1i}^2 + [1-\pi_i]\left[\theta_{1i} + \theta_{1i}^2\left(M_a^{(2,0)}(0,0) - 1\right)\right], \quad (1)$$

where

$$M_a^{(2,0)}(0,0) = \frac{(\alpha_1+1)\left[\alpha_1 + \rho_{11}^2(\alpha_1+6)\right]}{\lambda_1^2(1+\rho_{11}^2)}. \quad (2)$$

Likewise

$$\operatorname{Var}(y_{2i} \mid x_i, z_i) = \pi_i \theta_{2i}^2 + [1-\pi_i]\left[\theta_{2i} + \theta_{2i}^2\left(M_a^{(0,2)}(0,0) - 1\right)\right], \quad (3)$$

where $M_a^{(0,2)}(0,0)$ is obtained by replacing α_1 and λ_1 in equation (2) by α_2 and λ_2, respectively. Letting,

$$M_a^{(1,1)}(0,0) = \left[\alpha_1\alpha_2 + 2\rho_{11}\sqrt{\alpha_1\alpha_2} + \rho_{11}^2(\alpha_1+2)(\alpha_2+2)\right]/\lambda_1\lambda_2 \quad (4)$$

the covariance term can be derived as

$$\operatorname{Cov}(y_{1i}, y_{2i} \mid x_i, z_i) = \pi_i\theta_{1i}\theta_{2i} + [1-\pi_i]\left[\theta_{1i}\theta_{2i}\left(M_a^{(1,1)}(0,0) - 1\right)\right]. \quad (5)$$

Hence, the correlation coefficient for zero-inflated model is

$$\operatorname{Corr}(y_{1i}, y_{2i} \mid x_i, z_i) = \frac{\operatorname{Cov}(y_{1i}, y_{2i} \mid x_i, z_i)}{[\operatorname{Var}(y_{1i} \mid x_i, z_i)\operatorname{Var}(y_{2i} \mid x_i, z_i)]^{1/2}}, \quad (6)$$

where the variance and covariance terms are defined in (1), (3) and (5). This correlation for the bivariate zero-inflated Poisson-Laguerre mixture distribution can be negative or positive