

Different estimation methods for discrete survival time data with frailty

*Barroeta, Carolina*¹; *Julià, Olga*²; *Espinal, Anna*³

¹cbarroeta@yahoo.es, Department of Social Sciences, University C.L.A. Venezuela

²olgajulia@ub.edu, Departament de Probabilitat, Lògica i Estadística, Universitat de Barcelona,

³anna.espinal@uab.cat, Servei d'Estadística Aplicada, Universitat Autònoma de Barcelona

Different estimation methods for discrete survival time data with frailty

Our interest is to compare different methods as well as available software for estimating models for discrete time to event data. The most known models for dealing with discrete survival data are the logit model and the clog-log model. Moreover, specific methods for continuous survival, as Cox model with ties treatment, can be used. To this end we analyse the first diagnosis of mastitis in cows measured as number of lactations (discrete time).

Unobserved heterogeneity is the variability among individuals due to unmeasured characteristics. Due to cows are nested in herd models with hierarchical structure are required. Usually in survival analysis a way of introducing this unobserved heterogeneity is using a frailty term (or random effect).

The frailty term may be added either in discrete data clog-log and logit models as in Cox models with different treatments for ties. These frailty models have three parts: the random effect, the baseline function and the covariates with the corresponding regression coefficients β . Our goal is to compare the estimators obtained for both the coefficients of the covariates as for the frailty term using different models and several estimation methods.

In this study a sample of 437 cows that come from different geographic zones of Venezuela has been analyzed. The median follow-up time was 4.5 lactations (range, 1 – 8). Of the 437 cows, 338 cows (77.3%) were diagnosed with mastitis infection and 99 cows (22.7%) had a censored time. The observed covariates were: prior disease, season, food supplement, type of milking and the categorized lactation period (1 = from 7 to 100 days; 2 = from 101 to 244 days and 3 => 244 days). With the goal of controlling for unobserved heterogeneity the herd was introduced as a frailty term.

Keywords: discrete-time models, ties, frailty model, survival analysis