

Happy Ramanja Putri, 2010. Estimator Kernel dalam Model Regresi Semiparametrik pada Data Longitudinal. Skripsi ini dibawah bimbingan Nur Chamidah, S.Si, M.Si dan Drs. Suliyanto, M.Si., Departemen Matematika, Fakultas Sains dan Teknologi, Universitas Airlangga, Surabaya

ABSTRAK

Dalam pemodelan data longitudinal, agar dapat mengetahui pengaruh waktu dan faktor-faktor lain terhadap variabel respon, digunakan model regresi semiparametrik. Dalam skripsi ini, teknik *smoothing* yang dipilih untuk mengestimasi model regresi semiparametrik pada data longitudinal adalah estimator Kernel yang dapat diperoleh dengan meminimumkan WLS (*Weighted Least Square*) dan untuk pemilihan *bandwidth* optimal menggunakan metode *Generalized Cross Validation* (GCV). Untuk implementasi pada data dibuat program menggunakan *Software R*. Sumber data diperoleh dari RS.Haji Surabaya 2009 dengan variabel responnya adalah jumlah leukosit dan variable prediktornya adalah waktu dan jumlah trombosit. Dari hasil estimasi diperoleh *Mean Square Error* sebesar 0,00019 dengan nilai $\hat{\alpha}_1$ sebesar 0,8913718. Berdasarkan model yang didapat, setiap penambahan 1.000 sel trombosit penderita, dapat menyebabkan peningkatan jumlah leukosit penderita leukemia sebesar 891 sel.

Kata Kunci : *Regresi Semiparametrik, Data Longitudinal, Estimator Kernel, Generalized Cross Validation , Leukimia*

Happy Ramanja Putri, 2010. Kernel Estimation in Semiparametric Regression Model on Longitudinal Data. This *skripsi* is supervised by Nur Chamidah, S.Si, M.Si and Drs. Suliyanto, M.Si., Mathematics Department, Faculty of Sains and Technology, Airlangga University, Surabaya

ABSTRACT

In longitudinal data modeling, semiparametric regression model were used so the effect of time and the other factors to response variable were known. In this *skripsi*, smoothing technique that were selected to estimate the semiparametric regression model for longitudinal data is Kernel estimator by minimizing WLS (Weighted Least Square) and to optimizing bandwidth is used Generalized Cross Validation (GCV). For data application a program using software R were created. Data source is from RS.Haji Surabaya 2009, with response variable is leukocyte and predictor variables are time and thrombocyte amount. Based on model application, obtained MSE for the effect of time and thrombocyte amount to leukocyte amount is 0,00019, with $\hat{\alpha}_1$ number is 0,8913718 which means for every 1.000 increase of thrombocyte amount, the leukocyte will also increase for 891 cells.

Key Words: *Semiparametric Regression, Longitudinal Data, Estimator Kernel, Generalized Cross Validation , Leukocyte*