Dyah Pitaloka A.P.P, 2007. **Estimation of Multipredictor Nonparametric Regression Model with Lognormal Error Based of Local Polynomial Estimator**. This S*kripsi* in guided by Nur Chamidah, S.Si., M.Si and Toha Saifudin, S.Si., M.Si. Mathematics Departement, Faculty of Mathematic and Natural Sciences, Airlangga University.

ABSTRACT

The regression analysis is a method in statistica which used to explain relationship between two or more variables. Suppose Y is response variable and $X_1,X_2,\ldots,X_p$ is a predictor variables for n observation, the relationship of that variables can be expressed as:

$$y_i = m(x_{1i},...,x_{di})\varepsilon_i, \qquad ,i = 1,2,...,n \qquad \varepsilon_i \sim LN(0,\sigma^2)$$

where $m(x_{1i},...,x_{di})$ is unknown regression function. The model would be transformed by adding natural logarithm and the resulting is

$$y_i^* = m^*(x_{1i},...,x_{di}) + \varepsilon_i^* \qquad ,i = 1,2,...,n \qquad \varepsilon_i^* \sim N(0,\sigma^2)$$

m^* functions is assumed as additive function and the resulting is

$$y_i^* = \sum_{j=1}^{d} m_j^*(x_{ji}) + \varepsilon_i^* \qquad , i = 1,2,...,n$$

The purpose of this *skripsi* is to estimate the functions in the additive nonparametric regression models with the local polynomial estimator approach and make program in S-PLUS based on the backfitting algorithm and then applied at "Physical Fitness Course" data. From the result of applying estimation model based of local polynomial estimator, we got

$$\hat{m}(x_{1i},...,x_{di}) = \exp(\hat{m}^*(x_{1i},...,x_{di})) = \exp(\sum_{i=1}^{d} \hat{m}_j^*(x_{ji}))$$

with $\hat{m}_j^*(x_{ji}) = X_j^* \hat{\beta}_j$, $\hat{\beta}_j = \left(X_j^T(\lambda) W_j X_j(\lambda)\right)^{-1} X_j^T W_j \left(Y^* - \sum_{k \neq j} X_k(\lambda)\beta_k \right)$

Result of applying the model in "Physical Fitness Course" data by D.R.A.C Courtesy Linnerud we got the estimation model with value of *Mean Square Error* 0.0022

$$\hat{y} = \exp(\sum_{i=1}^{d} \hat{m}_j^*(x_{ji})) = \exp(3.78187691845 + 0.001017563\, X_1 + 0.032510548 X_2^2 + -0.032510548 X_2 + 0.0009851808 X_3 - 0.07207139 X_4)$$

From the model, increasing weight and average of heart beat when take a rest will increasing average of oxygen taken by body. In other side increasing average of heart beat when after running and time to run for 1.5 mil will decreasing average of oxygen taken by body.

Key Words : Nonparametric Regression, Local Polynomial Estimator, Backfitting Alghorithm, Lognormal Distribution

Dyah Pitaloka A.P.P, 2007. **Estimasi Model Regresi Nonparametrik Multiprediktor Dengan Error Lognormal Berdasarkan Estimator Polinomial Lokal.** Skripsi ini dibawah bimbingan Nur Chamidah, S.Si, M.Si dan Toha Saifudin, S.Si, M.Si. Jurusan Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Airlangga.

ABSTRAK

Analisis regresi merupakan suatu metode statistika yang digunakan untuk menjelaskan hubungan antara dua variabel atau lebih. Misalkan Y adalah variabel respon dan $X_1, X_2, ..., X_d$ adalah variabel-variabel prediktor untuk n pengamatan, maka hubungan antar variabel-variabel tersebut dapat dinyatakan dalam model multiplikatif sebagai berikut :

$$y_i = m(x_{1i},...,x_{di})\varepsilon_i, \qquad ,i = 1,2,...,n \qquad \varepsilon_i \sim LN(0,\sigma^2)$$

$m(x_{1i},...,x_{di})$ merupakan fungsi regresi yang tidak diketahui. Model multiplikatif tersebut ditransformasi dengan melogaritmakan basis e sehingga diperoleh model:

$$y_i^* = m^*(x_{1i},...,x_{di}) + \varepsilon_i^* \qquad ,i = 1,2,...,n \qquad \varepsilon_i^* \sim N(0,\sigma^2)$$

Fungsi m^* diasumsikan sebagai fungsi aditif sehingga diperoleh model :

$$y_i^* = \sum_{j=1}^{d} m_j^*(x_{ji}) + \varepsilon_i^* \qquad , i = 1,2,...,n$$

Tujuan dari skripsi ini adalah untuk mengestimasi fungsi-fungsi dalam model regresi nonparametrik dengan pendekatan estimator polinomial lokal dan membuat program pada *software* S-Plus berdasarkan algoritma *backfitting*, kemudian menerapkannya pada data "*Physical Fitness Course*". Dari hasil estimasi model berdasarkan estimator polinomial lokal diperoleh

$$\hat{m}(x_{1i},...,x_{di}) = \exp(\hat{m}^*(x_{1i},...,x_{di})) = \exp(\sum_{i=1}^{d} \hat{m}_j^*(x_{ji}))$$

dengan $\hat{m}_j^*(x_{ji}) = X_j^* \hat{\beta}_j$, $\hat{\beta}_j = \left(X_j^T(\lambda) W_j X_j(\lambda)\right)^{-1} X_j^T W_j \left(Y^* - \sum_{k \neq j} X_k(\lambda)\beta_k \right)$

Selanjutnya dari hasil penerapan pada data "*Physical Fitness Course*" yang dilakukan oleh D.R.A.C. Courtesy Linnerud diperoleh nilai *Mean Square Error* 0.0022 dan estimasi model

$$\hat{y} = \exp(\sum_{i=1}^{d} \hat{m}_j^*(x_{ji})) = \exp(3.78187691845 + 0.001017563\,X_1 + 0.032510548 X_2^2 + - 0.032510548 X_2 + 0.0009851808 X_3 - 0.07207139 X_4)$$

Berdasarkan model diatas, semakin bertambahnya berat badan dan rata-rata detak jantung saat istirahat akan meningkatkan rata-rata oksigen yang diambil oleh tubuh. Sedangkan semakin bertambahnya rata-rata detak jantung sesaat setelah berlari dan waktu berlari sejauh 1.5 mil akan menurunkan rata-rata oksigen yang diambil oleh tubuh.

Kata kunci : Regresi Nonparametrik, Estimator Polinomial Lokal, Algoritma *Backfitting*, Distribusi Lognormal.