

INTISARI

ESTIMASI PARAMETER MODEL CAMPURAN DISTRIBUSI PADA DATA SURVIVAL HETEROGEN MENGGUNAKAN ALGORITMA EM

Oleh :

NETI NURHASANAH

13/353882/PPA/04252

Dalam tesis ini dibahas mengenai estimasi parameter model campuran distribusi pada analisis data survival heterogen. Parameter yang tidak diketahui pada model campuran tersebut dapat diestimasi menggunakan algoritma EM. Algoritma EM meliputi dua tahapan, yaitu: tahap ekspektasi (*E-Step*) dan tahap maksimisasi (*M-Step*). Estimasi parameter model campuran distribusi akan diaplikasikan pada untuk menganalisis data survival pasien KNF (*Karsinoma Nasofaring*) yang berobat di RSUP dr. Sardjito Yogyakarta tahun 2008-2010. Hasil estimasi parameter pada data survival pasien KNF menggunakan algoritma EM menunjukkan bahwa model campuran dari distribusi Eksponensial, Gamma, dan Weibul lebih baik dibanding model campuran dari distribusi tunggal dalam merepresentasikan data survival. Dalam hal ini, diperoleh hasil bahwa model campuran dari distribusi berbeda (Eksponensial, Gamma, dan Weibull) menghasilkan nilai AIC dan BIC terkecil.

ABSTRACT

PARAMETER ESTIMATION OF A DISTRIBUTION MIXTURE MODEL ON HETEROGENEOUS SURVIVAL DATA USING EM ALGORITHM

By :

NETI NURHASANAH

13/353882/PPA/04252

In this thesis, we discuss about parameter estimation of a distribution mixture model on analysis of heterogeneous survival data using the EM algorithm. Unknown parameters in the mixture model can be estimated with using the EM algorithm. EM algorithm consists of two steps: the expectation step (E-Step) and the maximization step (M-Step). Parameter estimation of distribution mixture model will be applied to analyze NPC patients survival data who seek treatment at RSUP dr. Sardjito Yogyakarta in 2008-2010. Parameter estimation results in NPC patient survival data using EM algorithm shows that a mixture model of Exponential, Gamma, and Weibull distribution better than a mixture model of the single distribution represents survival data. In this case, the estimation result of a mixture model of different distribution (Exponential, Gamma, and Weibull) has the smallest value of AIC and BIC.

Keywords: heterogeneous survival data, mixture model, the EM algorithm.