**ESTIMASI PARAMETER DISTRIBUSI LOG-LOGISTIK PADA DATA SURVIVAL TERSENSOR TIPE II MENGGUNAKAN METODE MAKSIMUM LIKELIHOOD**

*1Jurusan Matematika, FMIPA Universitas Negeri Makassar, 90224*
a)e-mail: odiandihasan@gmail.com

**Abstrak.** Analisis *survival* merupakan suatu analisis data mengenai daya tahan hidup atau lamanya waktu hidup suatu individu atau unit pada keadaan tertentu. Tujuan penulisan skripsi ini adalah untuk mendapatkan model *survival* untuk data tersensortipe II, mendapatkan estimasi parameter-parameter, dan contoh penerapannya. Biasanya data *survival* akan mengikuti distribusi tertentu. Dalam skripsi ini akan dibahas mengenai data *survival* yang berdistribusi log-logistik.

 Model *survival* untuk data tersensor tipe II ditentukan dengan mencari estimasi parameter-parameter yaitu γ dan β berdasarkan fungsi maximum likelihood dan menentukan interval konfidensi untuk tiap-tiap parameter dengan mencari matriks informasi dan matriks kovarian terlebih dahulu. Sedangkan contoh data berdistribusi log-logistik didapatkan dengan metode simulasi pembangkitan data dengan *software Minitab* 16. Berdasarkan hasil pembahasan maka dapat disimpulkan bahwa model *survival* untuk data tersensor tipe II yang berdistribusi log-logistik yaitu

Estimasi parameter untuk γ dan β pada model *survival* untuk data tersensor tipe II berdasarkan distribusi log-logistik yaitu menggunakan metode maximum likelihood, dan dari pengolahan data lama perawatan pasien penderita penyakit diabetes mellitus yang berdistribusi log-logistik, didapatkan estimasi parameter untuk dan adalah dan .

**Kata kunci:** *Data tersensor, Distribusi Log-logistik,Maximum Likelihood Estimator*

**Abstract.** Survival analysis is an analysis of data about the survival or length of life time of an individual or unit in certain circumstances. The purpose of writing this thesis is to get survival model for type II censoreddata, get parameter estimation, and examples of its application. Usually the survival data will follow a certain distribution. In this thesis will be discussed about log-logistic-distributed survival data.

The survival model for type II censored data is determined by looking for estimation of the parameters of γ and β based on the maximum likelihood function and determining the confidence interval for each parameter by finding the matrix of information and covariance matrix first. While the example of log-logistic data is obtained by simulation method of data generation with Minitab 16 software.

Based on the results of the discussion it can be concluded that the survival model for type II censored data that log-logistic distributed is

Estimation of the parameters for γ and β in the survival model for type II censored data based on log-logistic distribution using maximum likelihood method, and from the old data processing treatment of patients with diabetes mellitus distributed logistic, obtained parameter estimation for γ and β is 236.22499 and 1.016799.

**Keyword:** *Censored Data, Log-logistic Distribution, Maximum Likelihood Estimator*

**PENDAHULUAN**

Perkembangan ilmu pengetahuan saat ini, khususnya di bidang statistika, menghasilkan penemuan-penemuan alat analisis yang dapat digunakan untuk menganalisis suatu permasalahan. Salah satunya adalah analisis uji hidup yang digunakan untuk meneliti daya tahan hidup suatu unit atau individu pada suatu keadaan tertentu. Analisis yang digunakan untuk menganalisis data waktu hidup tersebut disebut analisis *survival*. Analisis *survival* mencakup *survival time* (waktu tahan hidup) atau *failure time* (Collet, 1997).

Terdapat dua cara yang dapat dilakukan dalam pengambilan sampel pada analisis data tahan hidup yaitu pengamatan tersensor dan pengamatan tidak tersensor (pengamatan lengkap). Pengamatan tersensor dilakukan jika waktu tahan hidup dari individu yang diamati tidak diketahui secara pasti. Pengamatan tersensor diindikasi adanya individu yang tetap hidup sampai jangka waktu yang ditentukan (Lee, 2003).

Ada tiga jenis penyensoran dalam pengamatan terhadap waktu tahan hidup yaitu sensor tipe I, sensor tipe II dan sensor tipe III. Sebagian besar penelitian menggunakan penyensoran tipe II karena waktu pengamatan dapat dihentikan setelah diperoleh sebanyak *r* kegagalan, sehingga peneliti dapat menghemat waktu dan biaya.

Analisis data *survival*data tersensor diperlukan asumsi tertentu tentang distribusi populasinya. Adapun dalam tulisan ini digunakan fungsi *survival* berdistribusi Log-Logistik pada data *survival* tersensor tipe II. Dalam bidang kesehatan distribusi Log-Logistik dapat digunakan untuk meneliti data tahan hidup pasien. Misalnya pada pasien penderita penyakit diabetes mellitus.

**Data *Survival***

Data *survival* adalah data lamanya individu-individu atau unit-unit dari suatu populasi menjalankan fungsinya dengan baik sampai kematian individu- individu tersebut. Dalam mempelajari penerapan data *survival*, terlebih dahulu harus diketahui konsep-konsep statistik pada distribusi *survival*.

Misalkan T merupakan variabel random kontinu non negatif yang menunjukkan tahan hidup individu-individu dari suatu populasi. Pada model kontinu, fungsi-fungsi seperti fungsi densitas peluang, fungsi distribusi kumulatif,fungsi *hazard* dan fungsi *survivor* didefinisikan dalam interval [0, ∞) (Lawless, 1982).

Fungsi densitas peluang pada analisis *survival* adalah peluang suatu individu mati atau gagal dalam interval waktu *t* sampai , dengan waktu T merupakan variabel random. Fungsi densitas peluang dari T dapat dinyatakan sebagai ,

(2.2.)

Menurut Lawless (1982) fungsi *survivor* didefinisikan sebagai peluang suatu individu dapat bertahan hidup sampai waktu *t*. Jika *T*merupakan variabel random dari waktu hidup suatu individu dalam interval [0,∞), maka fungsi*survivor*  S(t) dapat dinyatakan dalam persamaan:

 *S*

Fungsi *hazard* menyatakan peluang kegagalan suatu komponen pada waktu *t*, jika diketahui bahwa komponen tersebut tetap hidup hingga waktu *t*. Menurut Lawless (1982) fungsi *hazard* adalah peluang suatu individu mati dalam interval waktu *t* sampai , jika diketahui individu tersebut masih dapat bertahan hidup sampai dengan waktu *t*, yang dinyatakan sebagai berikut:

**Data Tersensor**

Data tersensor adalah data yang diperoleh sebelum semua data teramati waktu hidupnya, sedangkan waktu pengamatan telah berakhir atau oleh sebab lain. Data tersensor tipe I merupakan data uji hidup yang dihasilkan setelah penelitian berjalan selama waktu yang telah ditentukan. Sedangkan data tersensor tipe II merupakan data hasil penelitian dimana penelitian dihentikan setelah kematian atau kegagalan tertentu telah terjadi (Lawless, 1982).

Data tersensor tipe II merupakan data kematian atau kegagalan yang tidak lengkap (*incomplete mortality data*) yaitu data waktu kematian atau kegagalan dari r observasi terkecil dalam sampel random yang berukuran n dengan 1≤ r ≤ n. Dalam suatu penelitian, penyensoran tipe II lebih sering digunakan, yaitu dalam uji hidup yang terdapat observasi sebanyak n, tetapi penelitian dihentikan ketika observasi mengalami kegagalan ke-r, sehingga dapat menghemat waktu dan biaya. Dalam penyensoran ini, r ditentukan terlebih dahulu sebelum data dikumpulkan.

**Dstribusi log-Logistik**

Variabel random *T* dikatakan mengikuti distribusi log-logistik dengan parameter dan parameter , jika mempunyai fungsi densitas:

Fungsi survivor darididefinisikan sebagai peluang suatu individu dapat bertahan hidup sampai waktu *t*, yaitu

*S*

Fungsi *hazard h(t)* menyatakan peluang suatu komponen mengalami kegagalan pada waktu *t*.

**METODOLOGI PENELITIAN**

Penelitian ini dilakukan pada bulan April sampai Juni 2017 dan menggunakan 35 data yang berdistribusi log-logistik berdasarkan uji distibusi yang diperoleh dari Data Rekam Medis Pasien Rawat Inap Diabetes Mellitus di Rumah Sakit Bhayangkara Makassar.

Mulai

Menentukan model survival untuk data tersensor tipe II berdistribusi log-logistik

Menentukan dan menggunakan metode maksimum *likelihood*

Memberikan contoh penerapan

Memberi kesimpulan hasil penerapan

Kesimpulan

Selesai

Gambar 3.1 Skema Penyelesaian Masalah

**HASIL DAN PEMBAHASAN**

**Estimasi Parameter Log-Logistik pada Data *Survival* Tersensor Tipe II**

Dalam data tersensor tipe II, terdapat r pengamatan dari n sampel yang diamati, dan eksperimen akan dihentikan setelah kegagalan ke-r yang terjadi sebelum waktu. Data terdiri dari *r*tahan hidup terkecil dari sampel random yang terdiri dari *n* tahan hidup .

Fungsi densitas peluang bersama dari dari data yang diamati dapat ditulis sebagai berikut:

Jadi fungsi *likelihood* dari distribusi log-logistik untuk data tersensor tipe II memiliki bentuk:

Dalam mengestimasi parameter distribusi Log-logistik dapat dilakukan dengan menurunkan fungsi likelihoodnya terhadap kemudian disamadengan nol menjadi;

dan

Kedua persamaan tersebut sulit diselesaikan secara manual karena memiliki bentuk yang kompleks, sehingga diperlukan bantuan dengan menggunakan suatu program atau *software* tertentu yang dapat digunakan untuk analisis *survival* dengan distribusi log-logistik.

**Aplikasi Pada Data**

Data yang analisis pada penelitian ini adalah 35 data yang berdistribusi log-logistik berdasarkan uji distibusi . Data berasal dari Data Rekam Medis Pasien Rawat Inap Diabetes Mellitus di Rumah Sakit Bhayangkara Makassar. Dari hasil output *software Minitab* 16, diperoleh nilai parameter lokasi dari data adalah 5,46489 dan nilai parameter skalaadalah 0,983479, sehingga didapatkan;

Berikut ini adalah bentuk fungsi densitas peluang dari data, sedangkan bentuk kurvanya ditunjukkan pada gambar 4.2:

*f*

**

Gambar 4.2. Kurva fungsi densitas peluang dari data lama perawatan

Bentuk fungsi *surviva* untuk data berdistribusi log-logistik setelah mensubtitusi nilai dan  kurva fungsi *survival* dari data pada gambar 4.3.

*S*

**

Gambar 4.3 Kurva fungsi *survival* dari data lama perawatan

Berdasarkan kurva fungsi *survival* (Gambar 4.3.), diketahui bahwa nilai fungsi *survival* semakin turun seiring dengan lama waktu perawatan pasien. Hal ini berarti semakin lama waktu pasien rawat inap, peluang kemampuan untuk bertahan hidup pasien akan semakin rendah.

Bentuk fungsi *hazard* untuk data berdistribusi log-logistik (2.10) setelah mensubtitusi nilai dan ditunjukkan pada persamaan 4.13 dan kurva fungsi *hazard* dari data pada gambar 4.4.

*h*

**

Gambar 4.4 Kurva fungsi *hazard* dari data lama perawatan

Berdasarkan kurva fungsi *hazard* (Gambar 4.4), diketahui bahwa nilai fungsi hazard semakin turun seiring dengan lama waktu perawatan pasien. Hal ini berarti semakin lama waktu pasien rawat inap, laju kesembuhannya semakin rendah.

1. **Kesimpulan dan saran**
	1. **Kesimpulan**

Hasil pengolahan dari data lama perawatan pasien yang menderita penyakit diabetes mellitus yang berdistribusi log-logistik adalah nilai estimasi untuk yaitu dan estimasi untuk yaitu .

* 1. **Saran**

Jurnal ini membahas tentang model *survival* dengan menentukan *maximum likelihood estimator* untuk dan yang merupakan parameter- parameter dari distribusi log-logistik. Dalam penulisan ini hanya membahas model *survival* untuk data tersensor tipe II. Oleh karena itu disarankan adanya penelitian lebih lanjut mengenai model *survival* dengan menggunakan distribusi log-logistik untuk data tersensor tipe yang lain dan juga distribusi-distribusi lain pada data kelompok.

**DAFTAR PUSTAKA**

Abadyo dan Hendro Permadi. 2005. *Metode Statistika Praktis*. Malang: UM Press.

Bain, L.J and Engelhardt. 1992. *Introduction to Probability and Mathematical Statistics*. 2nd ed. California: Duxbury Press.

Collett, David. 2004. *Modelling Survival Data in Medical Research*. 2nd ed. London: Chapman and Hall.

Dixit, Asha. 2008. *Exact Comparison of Hazard Rate Functions of Log-logistic Survival Distribution* [Tesis]. Alabama: Auburn University.

Lawless, J.F. 1982. *Statistical Model and Methods for Lifetime Data*. New York: John Wiley and Sons, Inc.

Lawless, J.F. 2003. *Statistical Model and Methods for Lifetime Data.* 2nd ed. New Jersey: John Wiley and Sons Inc.

Machin, David, Yin Bun C and Mahesh Parmar. 2006. *Survival Analysis Practical Approach.* 2nd ed. Chicester: John Wiley and Sons Ltd.

Rao, G.S, Kantam and K.Rosaih. 2009. “Reliability Estimation in Log-logistic Distribution from Cencored Samples”, *Prob.Stat*.,02,52-67.

Walpole, Ronald E. 1993. *Pengantar Statistika Edisi ke-3*. Jakarta: Gramedia Pustaka Utama.