

ANALISIS TAHAN HIDUP DATA TERSENSOR TIPE II MENGUNAKAN MODEL DISTRIBUSI WEIBULL PADA PENDERITA HEPATITIS C

Budi Santoso, Respatiwiulan, dan Tri Atmojo Kusmayadi
Program Studi Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam
Universitas Sebelas Maret

Abstrak. Setiap individu mempunyai kemampuan yang berbeda untuk bertahan dalam menghadapi hepatitis C. Waktu yang diperlukan setiap individu dalam bertahan hidup dapat dijadikan waktu tahan hidup. Waktu tahan hidup dapat berbentuk data lengkap, data tersensor tipe I, dan data tersensor tipe II. Data tersensor tipe II lebih sering digunakan karena lebih menghemat biaya dan waktu dibandingkan dua bentuk data yang lain. Tujuan penelitian ini adalah menurunkan ulang model tahan hidup berdasarkan distribusi Weibull, mengestimasi parameter, melakukan simulasi pembangkitan data tersensor tipe II penderita hepatitis C dan menginterpretasikannya. Estimasi parameter menggunakan metode *maximum likelihood estimation* (MLE) didapatkan persamaan nonlinier. Untuk menyelesaikannya, digunakan metode Newton-Raphson. Nilai estimasi parameter pada data simulasi diperoleh nilai $\hat{\alpha} = 0.002927443$ dan $\hat{\beta} = 3.282639$. Selanjutnya, dengan melakukan substitusi nilai estimasi ke fungsi waktu tahan hidup dapat disimpulkan bahwa sebesar 0.37% penderita hepatitis C akan mengalami kematian pada $t = 307$ bulan. Peluang penderita hepatitis C untuk bertahan hidup semakin lama semakin kecil. Semakin lama terinfeksi, semakin besar peluang penderita hepatitis C mengalami kematian sesaat.

Kata Kunci : analisis tahan hidup, data tersensor tipe II, distribusi Weibull, MLE, Newton-Raphson, hepatitis C.

1. PENDAHULUAN

Faktor cuaca, lingkungan dan sistem kekebalan tubuh yang lemah menyebabkan terjadinya penularan penyakit. Penularan penyakit merupakan kendala manusia sebagai individu dalam bertahan hidup. Penyakit menular disebabkan oleh virus atau bakteri yang menginfeksi individu. Beberapa penyakit menular antara lain *HIV/AIDS*, *SARS* dan hepatitis. Penularan hepatitis disebabkan oleh virus melalui cairan tubuh seperti darah, sperma, dan air susu ibu. Menurut Kementerian Kesehatan RI [1], hepatitis terdiri dari hepatitis A, B, C, D, dan E. Hepatitis A dan E sering muncul sebagai kejadian luar biasa yang berhubungan dengan perilaku hidup sehat dan bersifat akut. Namun, individu yang terinfeksi hepatitis ini dapat sembuh. Berbeda dengan hepatitis A dan E, hepatitis B, C, dan D dapat menjadi kronis dan dapat menimbulkan sirosis lalu menjadi kanker hati. Menurut data Organisasi Kesehatan Dunia (*WHO*) sekitar 3% dari 170 juta orang di dunia terinfeksi hepatitis C (*HCV*). Penderita hepatitis C terus bertambah dengan adanya infeksi baru yang setiap tahun mencapai empat juta penderita. Di Indonesia, penderita hepatitis C mencapai tujuh juta orang dan setiap tahun terus meningkat. (Kementerian Kesehatan RI [1])

Menurut Kementerian Kesehatan RI [1], infeksi hepatitis C dini menunjukkan gejala ringan dan tidak khas sehingga disebut juga *silent infection* (infeksi terselubung). Banyak dari penderita hepatitis C yang tidak sadar dirinya telah terinfeksi. Virus hepatitis C dibandingkan dengan virus hepatitis B lebih ganas dan lebih sering menimbulkan penyakit hepatitis yang kronis. Replikasi virus yang sangat cepat dan mampu berubah varian tipe dan sub tipe mengakibatkan sangat sulit untuk membuat vaksin anti-hepatitis C.

Setiap individu mempunyai kemampuan yang berbeda-beda untuk bertahan dalam menghadapi hepatitis C. Hal ini dipengaruhi oleh sistem kekebalan masing-masing individu. Waktu yang diperlukan setiap individu dalam bertahan hidup sampai mengalami kematian dapat dijadikan waktu tahan hidup. Menurut Lawless [2], waktu tahan hidup dapat berbentuk data lengkap, data tersensor tipe I dan data tersensor tipe II. Data tersensor tipe I merupakan data yang diperoleh setelah pengamatan berjalan selama waktu yang ditentukan, sedangkan data tersensor tipe II merupakan data yang diperoleh setelah terjadi kegagalan tertentu. Penyensoran tipe II lebih sering digunakan karena lebih menghemat waktu dan biaya dibandingkan data lengkap dan data tersensor tipe I.

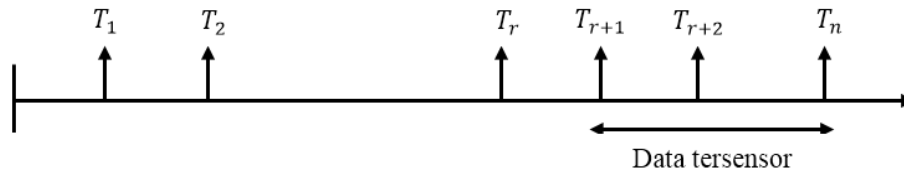
Analisis tahan hidup dengan data tersensor memerlukan asumsi tertentu tentang distribusi populasinya. Beberapa distribusi parametrik yang populer dan dapat digunakan untuk menganalisis model tahan hidup adalah distribusi gamma, distribusi eksponensial, distribusi log-normal, distribusi Weibull dan distribusi log-logistik. Menurut Lawless [2], distribusi Weibull merupakan model distribusi tahan hidup yang paling banyak digunakan dalam analisis tahan hidup. Aplikasi distribusi Weibull meliputi waktu tahan hidup atau daya tahan produk dalam suatu industri seperti komponen mobil dan peralatan listrik. Selain itu, distribusi Weibull juga digunakan dalam bidang biologi dan medis seperti studi tentang waktu sampai terjadinya tumor pada populasi manusia. Penelitian dalam bidang medis yang menggunakan distribusi Weibull, yaitu Simwa dan Mungisha [4] yang memodelkan jumlah cell CD4 pada penderita *HIV/AIDS*.

Penelitian ini menurunkan ulang model waktu tahan hidup berdasarkan distribusi Weibull, mengestimasi parameter dengan *maximum likelihood estimation (MLE)* yang mengacu pada Panahi dan Asadi [3]. Selanjutnya, melakukan simulasi pembangkitan data tersensor tipe II penderita hepatitis C dan menginterpretasikannya.

2. HASIL DAN PEMBAHASAN

commit to user
2.1. Fungsi Waktu Tahan Hidup Data Tersensor Tipe II. Data tersensor tipe II mempunyai pengamatan sebanyak r dari n sampel yang diamati dan percobaan akan dihentikan jika telah terjadi kegagalan ke- r yang terjadi sebelum

waktu t_i . Data terdiri dari r tahan hidup terkecil $T_1 \leq T_2 \leq T_3 \leq \dots \leq T_r$ dari sejumlah sampel random yang terdiri dari n tahan hidup $T_1, T_2, T_3, \dots, T_n$ seperti yang disajikan pada Gambar 1.



Gambar 1. Ilustrasi model tersensor tipe II

Berdasarkan Gambar 1, misalkan T merupakan variabel random dari n individu yang diamati, maka $f(t_1)$ merupakan fungsi densitas peluang dari variabel random individu ke-1, $f(t_2)$ merupakan fungsi densitas peluang dari variabel random individu ke-2, dan seterusnya sampai $f(t_r)$ untuk variabel random individu ke- r . Setiap individu yang mengalami kegagalan dari individu ke-1 sampai ke- r masing-masing mempunyai satu komponen waktu. Individu yang masih dapat bertahan sampai melebihi waktu kegagalan dari individu ke- r dapat dituliskan dengan $T_{r+1}, T_{r+2}, T_{r+3}, \dots, T_n$ dengan banyaknya individu $n - r$. Sampel random sebanyak n dengan jumlah kegagalan r mengikuti distribusi multinomial sehingga terdapat $\frac{n!}{(n-r)!}$ urutan yang mungkin terjadi dari n pengamatan.

Fungsi densitas peluang bersama dari $T_1, T_2, T_3, \dots, T_r$ dari data yang diamati dapat dituliskan sebagai

$$\begin{aligned} f(t_1, t_2, \dots, t_r) &= \frac{n!}{(n-r)!} f(t_1) f(t_2) \dots f(t_r) [P(T_{r+1} \geq t_r) \dots P(T_n \geq t_r)] \\ &= \frac{n!}{(n-r)!} \left[\prod_{i=1}^r f(t_i) \right] [S(t_r)]^{n-r}. \end{aligned} \quad (2.1)$$

Diasumsikan data waktu tahan hidup penderita hepatitis C mengikuti distribusi Weibull sehingga didapatkan tiga fungsi waktu tahan hidup.

(1) Fungsi densitas peluang

$$f(t_i) = \alpha \beta (\alpha t_i)^{\beta-1} (e^{-(\alpha t_i)^\beta}). \quad (2.2)$$

(2) Fungsi tahan hidup

$$S(t_i) = e^{-(\alpha t_i)^\beta}. \quad (2.3)$$

(3) Fungsi *hazard*

$$h(t_i) = \alpha \beta (\alpha t_i)^{\beta-1}. \quad (2.4)$$

Dari persamaan (2.2) dan (2.3) yang disubstitusikan ke dalam persamaan (2.1), diperoleh fungsi densitas peluang bersama data tarsensor tipe II berdistribusi Weibull sebagai

$$f(t_1, t_2, \dots, t_r) = \frac{n!}{(n-r)!} (\alpha^\beta \beta)^r \left(\prod_{i=1}^r t_i^{(\beta-1)} \right) \left(\prod_{i=1}^r e^{-(\alpha t_i)^\beta} \right) \left(e^{-(\alpha t_i)^\beta} \right)^{n-r} \quad (2.5)$$

2.2. Estimasi Parameter. Metode yang digunakan untuk mengestimasi parameter data tarsensor tipe II berdistribusi Weibull, yaitu metode *maximum likelihood estimation*. Dari persamaan (2.5) dapat dibentuk fungsi *likelihood*

$$L(\alpha, \beta) = \frac{n!}{(n-r)!} (\alpha^\beta \beta)^r \left(\prod_{i=1}^r t_i^{(\beta-1)} \right) \left(\prod_{i=1}^r e^{-(\alpha t_i)^\beta} \right) \left(e^{-(\alpha t_i)^\beta} \right)^{n-r}. \quad (2.6)$$

Kemudian fungsi *likelihood* pada persamaan (2.6) ditransformasikan ke bentuk logaritma natural (ln), didapatkan fungsi *ln-likelihood*

$$\ln L(\alpha, \beta) = \ln n! - \ln(n-r)! + r \ln \alpha^\beta + r \ln \beta + (\beta-1) \sum_{i=1}^r \ln t_i - \alpha^\beta (u_1) \quad (2.7)$$

dengan

$$u_1 = \sum_{i=1}^r t_i^\beta + (n-r)t_r^\beta.$$

Persamaan (2.7) diturunkan parsial terhadap α maupun β dan disamadengankan nol untuk memaksimumkan fungsi *ln-likelihood*.

$$\frac{r\beta}{\alpha} - \alpha^{(\beta-1)} \beta (u_1) = 0 \quad (2.8)$$

$$\frac{r}{\beta} + r \ln \alpha + \sum_{i=1}^r \ln t_i - \alpha^\beta \ln \alpha (u_1) - \alpha^\beta (u_2) = 0 \quad (2.9)$$

dengan

$$u_2 = \sum_{i=1}^r t_i^\beta (\ln t_i) + (n-r)t_r^\beta (\ln t_r).$$

Dari persamaan (2.8) diperoleh

$$\alpha = \left(\frac{r}{u_1} \right)^{\frac{1}{\beta}}. \quad (2.10)$$

Dengan melakukan substitusi pada persamaan (2.10) ke persamaan (2.9) diperoleh

$$\frac{r}{\beta} + \sum_{i=1}^r \ln t_i - \left(\frac{r}{u_1} \right) (u_2) = 0 \quad (2.11)$$

Persamaan (2.11) merupakan persamaan nonlinier yang sulit ditentukan penyelesaian eksaknya sehingga diperlukan penyelesaian secara numerik dengan metode Newton - Raphson. Terdapat empat langkah yang dilakukan untuk mengestimasi

parameter β . Langkah pertama menentukan nilai awal dari β . Langkah kedua menentukan fungsi $f(\beta)$ dan turunan pertamanya.

$$f(\beta) = \frac{r}{\beta} + \sum_{i=1}^r \ln t_i - \left(\frac{r}{u_1} \right) (u_2) \quad (2.12)$$

$$f'(\beta) = -\frac{r}{\beta^2} + \frac{r}{(u_1)^2} (u_2)^2 - \frac{r}{(u_1)} (u_3) \quad (2.13)$$

dengan

$$u_3 = \sum_{i=1}^r t_i \beta (\ln t_i)^2 + (n-r) t_r \beta (\ln t_r)^2.$$

Langkah ketiga melakukan substitusi pada kedua persamaan (2.12) dan (2.13) ke persamaan iterasi Newton-Raphson

$$\hat{\beta}_{k+1} = \hat{\beta}_k - \frac{f(\hat{\beta}_k)}{f'(\hat{\beta}_k)}$$

sehingga diperoleh persamaan iterasi Newton-Raphson untuk mengestimasi parameter β yaitu

$$\hat{\beta}_{k+1} = \hat{\beta}_k + \frac{\frac{r}{\hat{\beta}_k} + S_{1,f} - \frac{r}{S_{2,f} + S_{2,s}} (S_{3,f} + S_{3,s})}{\frac{r}{\hat{\beta}_k^2} - \frac{r}{S_{2,f} + S_{2,s}} (S_{3,f} + S_{3,s})^2 + \frac{r}{S_{2,f} + S_{2,s}} (S_{4,f} + S_{4,s})} \quad (2.14)$$

dengan

$$S_{1,f} = \sum_{i=1}^r \ln t_i$$

$$S_{2,f} = \sum_{i=1}^r t_i \hat{\beta} \quad ; \quad S_{2,s} = (n-r) t_r \hat{\beta}$$

$$S_{3,f} = \sum_{i=1}^r t_i \hat{\beta} (\ln t_i) \quad ; \quad S_{3,s} = (n-r) t_r \hat{\beta} (\ln t_r)$$

$$S_{4,f} = \sum_{i=1}^r t_i \hat{\beta} (\ln t_i)^2 \quad ; \quad S_{4,s} = (n-r) t_r \hat{\beta} (\ln t_r)^2$$

Langkah keempat adalah menentukan nilai $d = \hat{\beta}_n - \hat{\beta}_{n-1}$. Nilai dari d digunakan untuk menghentikan iterasi dengan kriteria $|d| < \text{toleransi eror}$. Toleransi eror ditentukan sebesar 0.0000001. Jika limit toleransi yang ditentukan terpenuhi, iterasi berhenti. Sebaliknya, jika $|d| \geq \text{toleransi eror}$ maka iterasi akan kembali pada langkah pertama dengan nilai β yang diperoleh digunakan untuk iterasi selanjutnya. Jika sudah diperoleh nilai estimasi β yang memenuhi kriteria yang ditentukan, nilai estimasi parameter α ditentukan sebagai

$$\hat{\alpha} = \left(\frac{r}{S_{2,f} + S_{2,s}} \right)^{\frac{1}{\hat{\beta}}} \quad (2.15)$$

Berdasarkan metode Newton-Raphson diperoleh estimasi parameternya, yaitu $\hat{\beta} = \hat{\beta}_{k+1}$ yang dihitung berdasarkan persamaan (2.14) dan $\hat{\alpha}$ yang dihitung berdasarkan persamaan (2.15)

2.3. Simulasi. Hepatitis C merupakan penyakit dengan waktu infeksi yang lama. Waktu pertama terinfeksi sampai kronis (sirosis) dapat bertahun - tahun. Menurut *Centers for Disease Control and Prevention (cdc.gov)* [5], lama waktu infeksi hepatitis C sampai mengalami sirosis dan kematian paling sering terjadi antara 240 bulan sampai 360 bulan dan dapat mencapai 780 bulan. Mengacu pada *Centers for Disease Control and Prevention (cdc.gov)* [5], dilakukan simulasi pembangkitan data waktu tahan hidup penderita hepatitis C.

Misalkan terdapat 50 penderita yang terinfeksi hepatitis C. Dari 50 penderita diambil sebanyak 30 pengamatan pertama yang telah ditentukan dahulu sebelum pengamatan dilakukan. Diperoleh 30 data penderita yang meninggal karena terinfeksi hepatitis C dan 20 data tersensor. Berdasarkan ilustrasi Gambar 1, T_1 merupakan waktu tahan hidup penderita hepatitis C yang pertama meninggal dalam pengamatan sampai T_{30} yang merupakan waktu tahan hidup penderita hepatitis C yang ke-30 meninggal. Waktu individu yang masih dapat bertahan hidup setelah T_{30} , yaitu T_{31} sampai T_{50} merupakan data tersensor dari penderita hepatitis C. Waktu tahan hidup 30 penderita disajikan pada Tabel 1.

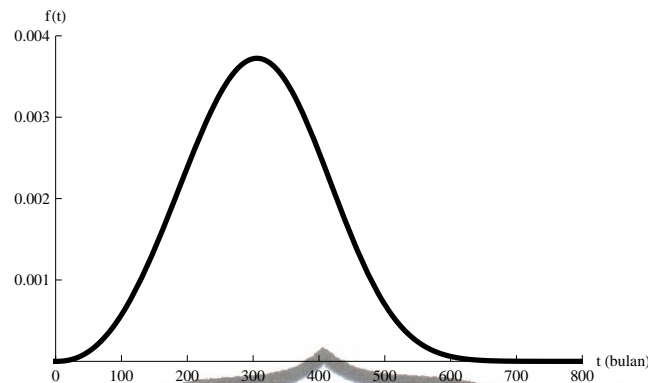
Tabel 1. Waktu tahan hidup penderita hepatitis C

No	$t(\text{bulan})$	No	$t(\text{bulan})$	No	$t(\text{bulan})$	No	$t(\text{bulan})$	No	$t(\text{bulan})$
1	113.931	7	176.043	13	230.758	19	257.510	25	308.177
2	126.944	8	202.794	14	234.784	20	265.204	26	313.805
3	133.892	9	209.582	15	240.199	21	290.310	27	313.841
4	148.848	10	210.504	16	243.139	22	297.448	28	316.217
5	151.443	11	212.354	17	246.193	23	302.131	29	318.145
6	168.973	12	226.583	18	250.369	24	305.658	30	333.250

Tabel 1 merupakan data hasil simulasi pembangkitan data dengan nilai rata - rata adalah 310.5. Menurut *Centers for Disease Control and Prevention (cdc.gov)* [5], peluang individu yang terinfeksi hepatitis C mengalami kematian semakin besar pada interval waktu tertentu lalu semakin kecil dengan bertambahnya waktu. Karakteristik hepatitis C tersebut dapat dijelaskan oleh fungsi densitas peluang dengan nilai β antara 2 sampai 3. Diambil nilai $\beta = 2.5$ serta $\alpha = 0.00286$ untuk membatasi waktu infeksi hepatitis C dengan lama waktu infeksi maksimal 780 bulan. Selanjutnya, untuk mencari estimasi parameter α dan β pada data tersensor tipe II digunakan metode *maximum likelihood estimation* dengan metode Newton-Raphson.

Estimasi parameter distribusi Weibull pada data tersensor II menggunakan bantuan software dengan nilai awal parameter $\beta = 2$ diperoleh nilai $|d| = 4.973799 \times 10^{-14}$ pada iterasi kelima dengan nilai estimasi $\hat{\alpha} = 0.002927443$ dan $\hat{\beta} = 3.282639$. Nilai estimasi kemudian disubstitusi ke dalam fungsi densitas peluang, fungsi tahan hidup, dan fungsi hazard.

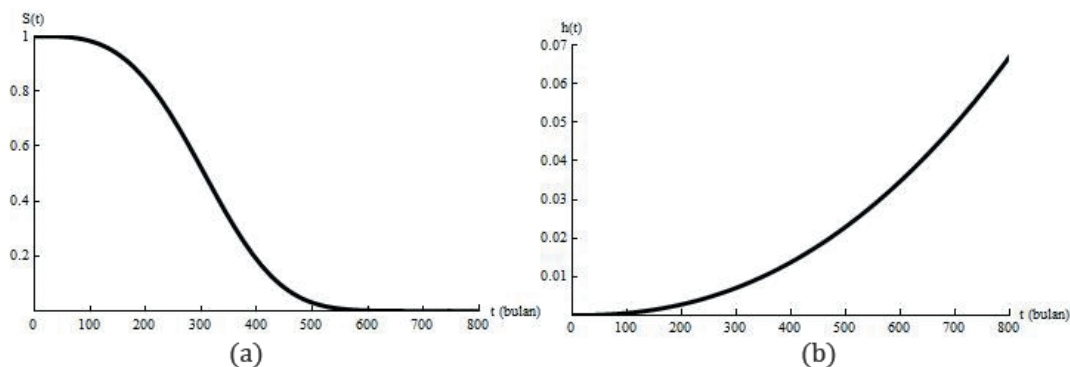
Fungsi densitas peluang yang menunjukkan peluang terjadinya kematian pada penderita hepatitis C dapat disajikan pada Gambar 2.



Gambar 2. Fungsi densitas peluang

Dari Gambar 2, terlihat peluang terjadinya kematian pada penderita hepatitis C semakin lama semakin besar sampai $t = 307$ bulan. Setelah $t = 307$ bulan, peluang terjadinya kematian pada penderita hepatitis C semakin kecil mendekati 0. Kematian paling banyak terjadi pada saat $t = 307$ bulan dengan peluang sebesar 0.00372369 yang artinya 0.37% penderita hepatitis C akan mengalami kematian pada $t = 307$ bulan.

Fungsi tahan hidup yang menunjukkan peluang penderita hepatitis C bertahan hidup dapat disajikan pada Gambar 3(a). Fungsi *hazard* yang menunjukkan peluang penderita hepatitis C mengalami kematian dalam interval waktu yang sangat kecil (sesaat) dapat disajikan pada Gambar 3(b).



Gambar 3. (a) Fungsi tahan hidup dan (b) Fungsi *hazard*

Dari Gambar 3(a) terlihat bahwa saat $t = 0$ bulan, peluang penderita hepatitis C untuk bertahan hidup adalah 1. Saat $t = 800$ bulan, peluang penderita hepatitis C untuk bertahan hidup mendekati 0 yang artinya peluang penderita hepatitis C untuk bertahan hidup selama 800 bulan sangat kecil. Hal ini menunjukkan bahwa semakin lama terinfeksi, semakin kecil peluang penderita hepatitis C untuk bertahan hidup. Dari Gambar 3(b), terlihat bahwa peluang penderita hepatitis C mengalami kematian sesaat berbanding lurus terhadap waktu. Saat

$t = 0$ bulan, peluang penderita hepatitis C mengalami kematian sesaat adalah 0 lalu semakin besar sampai $t = 800$ bulan dengan peluang sebesar 0.0670387 yang artinya penderita hepatitis C memiliki peluang kematian sebesar 6.7% jika penderita dapat bertahan hidup selama $t = 800$ bulan. Hal ini menunjukkan bahwa semakin lama terinfeksi, semakin besar peluang penderita hepatitis C mengalami kematian sesaat.

3. KESIMPULAN

Dari pembahasan yang telah dilakukan, dapat dituliskan dua kesimpulan.

- (1) Berdasarkan distribusi Weibull, diperoleh fungsi waktu tahan hidup dan estimasi parameter data tersensor tipe II.

- (a) Fungsi waktu tahan hidup

$$f(t_i) = \alpha\beta(\alpha t_i)^{\beta-1}(e^{-(\alpha t_i)^\beta})$$

$$S(t_i) = (e^{-(\alpha t_i)^\beta})$$

$$h(t_i) = \alpha\beta(\alpha t_i)^{\beta-1}$$

- (b) Estimasi parameter α dan β

$$\hat{\alpha} = \frac{r}{S_{2,f} + S_{2,s}}$$

$$\hat{\beta}_{k+1} = \hat{\beta}_k + \frac{\frac{r}{\hat{\beta}_n} + S_{1,f} - \frac{r}{S_{2,f} + S_{2,s}}(S_{3,f} + S_{3,s})}{\frac{r}{\hat{\beta}_n^2} - \frac{r}{S_{2,f} + S_{2,s}}(S_{3,f} + S_{3,s})^2 + \frac{r}{S_{2,f} + S_{2,s}}(S_{4,f} + S_{4,s})}$$

- (2) Dari simulasi data tersensor tipe II penderita hepatitis C diperoleh dua kesimpulan.

- (a) Nilai estimasi $\hat{\alpha} = 0.00372369$ dan $\hat{\beta} = 3.282639$.
 - (b) Sebesar 0.37% penderita hepatitis C akan mengalami kematian pada $t = 307$ bulan. Peluang penderita hepatitis C untuk bertahan hidup semakin lama semakin kecil. Semakin lama terinfeksi, semakin besar peluang penderita hepatitis C mengalami kematian sesaat.

DAFTAR PUSTAKA

- [1] Kementerian Kesehatan Republik Indonesia, *Situasi dan Analisis Hepatitis*, Pusat Data dan Informasi Kementerian Kesehatan RI, Jakarta Selatan, 2014.
- [2] Lawless, J. F., *Statistical Model and Methods for Life Time Data*, John Wiley and Sons, Inc., New York, 1982.
- [3] Panahi, H. and S. Asadi, *Estimation of the Weibull Distribution Based on Type-II Censored Samples*, Journal of Applied Mathematical Sciences, **5** (2011), no. 52, 2549-2558.
- [4] Simwa, O. R and T. Y. J. Mungisha, *A Model for the CD4 Cell Counts in an HIV/AIDS Patient and its Application in Treatment Interventions*, American Journal of Infectious Diseases, **1** (2005), no. 1, 61-65.
- [5] www.cdc.gov, *Viral Hepatitis-Hepatitis C Information*, <http://www.cdc.gov/hepatitis/hcv/hcvfaq.htm>, (diakses tanggal 07 Januari 2016).