

**METODE ESTIMASI ANALISIS *MIXTURE CURE MODEL*  
SEMIPARAMETRIK *ACCELERATED FAILURE TIME*  
(AFT) PADA PASIEN KANKER PAYUDARA  
(ESTIMATION METHOD FOR SEMIPARAMETRIC  
ACCELERATED FAILURE TIME (AFT) MIXTURE CURE  
MODEL FOR PATIENTS BREAST CANCER )**

AGI DEMPITA\*, DANARDONO

**Abstract.** *Analysis of survival is one of the statistical methods used to analyze data related to time, from the time origin or start point to the occurrence of a special end point event. One method for survival analysis is the mixture cure model, which is used to determine the proportion of cured patients and the probability of patients not being cured until a certain time limit with the given treatment. The EM algorithm method is used to determine the value of the estimation parameter. In this study, the data of breast cancer patients and the variables that significantly affect the patient's recovery model are the treatment variables, while in the failure time model there are no significant variables. The chemotherapy treatment given to breast cancer patients is the Taxan Base, Antarsiklin+Taxan and Antarsiklin Base regimens. The highest survival probability of cured patients is patients given the Antarsiklin Base treatment.*

**Keywords:** survival analysis, mixture cure model, EM Algorithm, accelerated failure time, breast cancer.

**Abstrak.** Analisis survival merupakan salah satu metode statistika yang digunakan untuk menganalisis data yang berhubungan dengan waktu, mulai dari *time origin* atau *start point* sampai dengan terjadinya suatu kejadian khusus *end point*. Salah satu metode untuk analisis survival adalah *mixture cure model* yang digunakan untuk mengetahui proporsi pasien sembuh dan probabilitas pasien tidak sembuh sampai pada batas waktu tertentu dengan diberinya *treatment* atau perlakuan. Untuk mengetahui nilai estimasi parameter digunakan metode Algoritma EM. Pada penelitian ini, menggunakan data pasien kanker payudara dan variabel yang berpengaruh signifikan terhadap model pasien sembuh adalah variabel *treatment*, sedangkan pada *failure time model* tidak ada variabel yang signifikan. *Treatment* kemoterapi yang dilakukan pada pasien kanker payudara adalah rejimen Taxan Base, Antarsiklin+Taxan dan Antarsiklin Base. Probabilitas survival pasien sembuh tertinggi yaitu pasien yang diberi *treatment* Antarsiklin Base.

*Kata-kata kunci:* analisis survival, *mixture cure model*, Algoritma EM, *accelerated failure time*, kanker payudara.

## 1. PENDAHULUAN

Analisis survival merupakan salah satu metode statistika yang digunakan untuk menganalisis data yang berhubungan dengan waktu, mulai dari *time origin* atau *start point* sampai dengan terjadinya suatu kejadian khusus *end point*. Analisis survival memerlukan data survival suatu individu untuk mengukur waktu kejadian atau waktu bertahan hidup dari peristiwa penyembuhan, kekambuhan, kematian sampai dengan periode waktu tertentu yang disebut juga *failure event*.

Model *cure rate* dikembangkan untuk mengestimasi proporsi pasien yang sembuh dalam studi klinik dan mengestimasi probabilitas survival pasien yang tidak sembuh sampai dengan batas waktu yang diberikan. Model *mixture* dikatakan model *cure mixture* parametrik jika menggunakan probabilitas standar seperti distribusi eksponensial, Weibull, *Gomperts*, dan *generalized gamma*. Diskusi tentang *mixture* parametrik dapat dilihat pada Boag [1] dan Farwell [4]

Li dan Taylor [7] *cure* model digunakan untuk menganalisis data waktu kegagalan (*failure time*) yang memiliki 2 komponen, yaitu *incidence* menunjukkan terjadinya *event* dan *latency* menunjukkan kapan *event* itu terjadi. Model *cure mixture* AFT merupakan alternatif yang digunakan untuk model *cure mixture Proportional Hazard* (PH). Berbeda dengan model *cure mixture* PH, model *cure mixture* AFT menggunakan AFT sebagai latensi. Pada penelitian ini penulis akan mengaplikasikan model *mixture cure* semiparametrik AFT untuk analisis data pasien kanker payudara yang mendapatkan *treatment* dengan tujuan untuk mengetahui proporsi pasien sembuh dan probabilitas pasien tidak sembuh.

## 2. TINJAUAN PUSTAKA

Penelitian Zhang, Peng [9] yang berjudul "a new estimation method for the semi-parametric accelerated failure time mixture cure model". Dalam penelitian tersebut menggunakan data transplantasi sumsum tulang pasien leukemia limfoblastik akut refraktori dengan tujuan untuk mengetahui proporsi pasien sembuh dan mengestimasi probabilitas survival pasien tak sembuh. Penelitian dianalisis dengan metode estimasi baru yang sebelumnya telah diusulkan yaitu menggunakan mixture cure model semi-parametrik AFT. Metode ini menggunakan Algoritma EM dan estimator *rank* AFT untuk mengestimasi parameter yang diinginkan.

Li dan Taylor [7] sebelumnya pernah melakukan analisis yang berjudul "a semi-parametric accelerated failure time cure model". Penelitian tersebut menggunakan kumpulan data pasien kanker amandel yang diobati atau mendapat *treatment* dengan terapi radiasi. Model penyembuhan (*cure model*) semiparametrik menunjukkan bahwa variabel kovariat dapat mempengaruhi *incidence* dan *latency*. Digunakan model regresi logistik untuk *incidence* dan *latency* ditentukan oleh regresi AFT dengan distribusi *error* yang tidak ditentukan. Algoritma EM digunakan sebagai estimator untuk mengestimasi parameter yang tidak diketahui dalam model *cure mixture* semiparametrik AFT.

**2.1. Estimasi Kaplan-Meier.** Metode Estimasi *Kaplan-Meier* dikenal juga sebagai *Product-Limit estimator* merupakan salah satu metode estimasi non parametrik dalam analisis survival (Harlan) [5]. Karakteristik estimasi *Kaplan Meier* yaitu:

- (1) Interval waktu seringkali tidak sama, perhitungan dilakukan setiap ada kegagalan
- (2) fungsi survival untuk suatu interval waktu adalah proporsi jumlah subjek survive pada awal interval dikurangi jumlah kegagalan dalam interval tersebut

Estimator *Kaplan-Meier* digunakan untuk mengestimasi fungsi survival, diberikan rumus sebagai berikut (Klein dan Moeschberger, 2003):

$$\hat{S}(t) = \begin{cases} 1, & t \leq t_1 \\ \prod_{t_i \leq t} \left(1 - \frac{d_i}{Y_i}\right), & t_i \leq t \end{cases} \quad (2.1)$$

dimana  $d_i$  adalah banyaknya *event* dan  $Y_i$  adalah banyaknya individu yang beresiko (*number at risk*).

### 2.2. Mixture Cure Model Semiparametric.

*tire cure* model digunakan untuk menganalisis data survival. Keuntungan dari metode ini yaitu dapat menghitung proporsi pasien sembuh dan probabilitas pasien yang tidak sembuh dari penyakit kanker payudara sebagai kasus. Diberikan  $T$  menunjukkan sebagai waktu kegagalan dan  $S_{pop}(t|x, z)$  merupakan fungsi survival  $T$ , dimana  $x$  dan  $z$  nilai yang diamati dari dua vektor kovariat. Distribusi survival tergantung pada nilai  $x$  dan  $z$  yang diamati. Oleh karena itu, model *mixture cure* dapat dituliskan sebagai berikut:

$$S_{pop}(t|x, z) = \pi(z) S(t|x) + 1 - \pi(z) \quad (2.2)$$

dengan:

$1 - \pi(z)$  = probabilitas pasien sembuh

$S(t|x)$  = probabilitas hidup waktu  $t$  pasien tidak sembuh

Proporsi pasien sembuh dapat diperoleh dengan persamaan yang bergantung pada fungsi Cloglog. Proporsi tersebut dapat dituliskan sebagai berikut:

$$1 - \pi(z) = 1 - \exp(-\exp(\alpha + \beta x)) \quad (2.3)$$

### 2.3. Estimasi Parameter dengan Algoritma-EM.

gortima EM digunakan untuk mengestimasi parameter yang tidak diketahui  $(\mathbf{b}, \beta)$  dan fungsi survival  $S(\cdot)$  pada model. Misal,  $\mathbf{O} = (t_i, \delta_i, z_i, x_i)$  menyatakan data observasi ke- $i$  dengan  $i = 1, \dots, n$ , dengan  $t_i$  adalah waktu survival yang diobservasi,  $\delta_i$  adalah indikator sensoring dengan  $\delta_i = 1$  untuk waktu tidak tersensor, dan  $\delta_i = 0$  untuk waktu tersensor, dan  $z_i, x_i$  merupakan kovariat kejadian dan laten. Didefinisikan suatu laten variabel random  $y_i$  dengan  $y_i = 1$  jika individu ke- $i$  tidak sembuh dan  $y_i = 0$  sembuh. Dapat dilihat bahwa  $y_i = 1$  jika  $\delta_i = 1$  dan  $y_i$  tidak diketahui jika  $\delta_i = 0$ . Diberikan nilai dari  $y = (y_1, \dots, y_n)$ , melengkapi *log-likelihood*  $l_c(b, \beta, S(\cdot); \mathbf{O}, y) = l_{c1}(b, \mathbf{O}, y) + l_{c2}(\beta, S(\cdot); \mathbf{O}, y)$ , dimana

$$l_{c1}(b, \mathbf{O}, y) = \sum_{i=1}^n y_i \log[\pi(z_i)] + (1 - y_i) \log[1 - \pi(z_i)] \quad (2.4)$$

$$l_{c2}(\beta, S(\cdot); \mathbf{O}, y) = \sum_{i=1}^n y_i \delta_i \log[h(\log(t_i) - \beta x_i)] + y_i \log[S(\log(t_i) - \beta x_i)] \quad (2.5)$$

dengan  $h(\cdot) = \frac{f(\cdot)}{S(\cdot)}$  merupakan fungsi *hazard* dari  $\varepsilon$ .

**E-Step** menghitung ekspektasi bersyarat dari fungsi *log-likelihood* dengan variabel laten  $y_i$ , diberikan estimasi saat ini  $\Theta^{(m)} = \{b^{(m)}, \beta^{(m)}, S^{(m)}(t)\}$  dan  $\mathbf{O}$  data yang diobservasi. Misal  $w_i^{(m)}$  dinotasikan  $E(y_i | \Theta^{(m)}, \mathbf{O})$ , yang diinterpretasikan sebagai probabilitas bersyarat dari individu ke- $i$  yang tidak sembuh pada iterasi ke- $m$  pada algoritma. Dengan rumus sebagai berikut:

$$w_i^{(m)} = \delta_i + (1 - \delta_i) \frac{\pi(z_i) S(\log(t_i) - \beta x_i)}{1 - \pi(z_i) + \pi(z_i) S(\log(t_i) - \beta x_i)} \Big|_{(\Theta^{(m)}, \mathbf{O})} \quad (2.6)$$

Oleh karena itu,  $E[l_c(b, \beta, S(\cdot); \mathbf{O}, y) | \Theta^{(m)}, \mathbf{O}] = l_{c1}(b) + l_{c2}(\beta, S(\cdot))$ , dimana

$$l_{c1}(b) = \sum_{i=1}^n w_i^{(m)} \log[\pi(z_i)] + (1 - w_i^{(m)}) \log[1 - \pi(z_i)] \quad (2.7)$$

$$l_{c2}(\beta, S(\cdot)) = \sum_{i=1}^n w_i^{(m)} \delta_i \log[h(\log(t_i) - \beta x_i)] + w_i^{(m)} \log[S(\log(t_i) - \beta x_i)] \quad (2.8)$$

**M-step** dilakukan untuk memaksimalkan 2.7 dan 2.8 dengan parameter  $b, \beta$  dan  $S(\cdot)$ . Pada penelitian yang dilakukan oleh Zhang, Peng [9] mengusulkan prosedur baru untuk mengestimasi parameter  $\beta$  pada M-step. Dengan menggunakan fakta bahwa

$\delta_i \log w_i^{(m)} \equiv 0$  dan  $\delta_i w_i^{(m)} \equiv \delta_i$ , dapat dituliskan persamaan sebagai berikut:

$$l_{c2}(\beta, S(\cdot)) = \delta_i \log \left[ w_i^{(m)} h(\log(t_i) - \beta x_i) \right] + w_i^{(m)} \log [S(\log(t_i) - \beta x_i)] \quad (2.9)$$

yang dapat ditulis sebagai algoritma dari

$$\prod_{i=1}^n \left[ w_i^{(m)} h(\log(t_i) - \beta x_i) \right]^{\delta_i} \left[ S(\log(t_i) - \beta x_i)^{w_i^{(m)}} \right] \quad (2.10)$$

dengan  $S(\log(t_i) - \beta x_i)^{w_i^{(m)}}$  merupakan fungsi survival dengan fungsi *hazard* yaitu  $w_i^{(m)} h(\log(t_i) - \beta x_i)$ , persamaan 2.10 dapat dikatakan sebagai fungsi *log-likelihood* model AFT dengan

$$\log(T_i) = \beta x_i + \varepsilon_i^* \quad (2.11)$$

dimana fungsi *hazard* dari  $\varepsilon_i^*$  adalah  $w_i^{(m)} h(\varepsilon_i^*)$  dan  $w_i^{(m)}$  adalah konstanta.

Varian dari estimasi parameter yang diusulkan menggunakan algoritma EM pada *mixture cure* model semiparametrik AFT tidak langsung tersedia karena fungsi *log like-lihood* yang tidak lengkap. Oleh karena itu, diusulkan metode *bootstrap* untuk memperoleh varian estimasi dari  $\beta$  dan  $b$ . Dimana metode *bootstrap* digunakan untuk mengestimasi varian dari  $\hat{b}$  dan  $\hat{\beta}$

### 3. METODE PENELITIAN

Data yang digunakan dalam penelitian ini merupakan data sekunder dengan studi kasus pasien kanker payudara yang menerima pengobatan kemoterapi di RSUP Dr. Sardjito Yogyakarta selama periode Januari 2017 – Desember 2021. Penelitian dilakukan terhadap 138 pasien kanker payudara dengan status penyakit pasien stadium awal (stadium I-IIB), dan terdapat sebanyak 125 pasien kanker payudara masih bertahan hidup dan sebanyak 13 pasien meninggal dunia. Variabel yang digunakan dalam penelitian terdiri dari variabel waktu survival, umur, status tersensor, status *treatment*, tingkat stadium.

Analisis data pada penelitian ini menggunakan bantuan *software* RStudio 4.2.0, dengan langkah- langkah sebagai berikut:

- (1) Input data.
- (2) Melakukan analisis dengan estimator Kaplan- Meier untuk mengestimasi fungsi survival dan mengetahui karakteristik fungsi survival.
- (3) Melakukan estimasi parameter dengan beberapa sampel *bootstrap* dan membandingkan *output* dengan melihat nilai *standard error* terkecil dari sampel *bootstrap*.
- (4) Melakukan pengecekan variabel yang berpengaruh signifikan terhadap model waktu survival dari sampel *bootstrap* terpilih
- (5) Memprediksi nilai proporsi dari pasien sembuh
- (6) Memprediksi nilai probabilitas dari pasien tidak sembuh
- (7) Melakukan interpretasi dari hasil analisis

## 4. HASIL DAN PEMBAHASAN

### 4.1. Estimasi Kaplan Meier.

Estimator *Kaplan Meier* digunakan untuk mengestimasi fungsi *survival* dan mengetahui karakteristik kurva *survival*. Pada penelitian ini digunakan tiga variabel dengan analisis *Kaplan Meier* sebagai berikut:

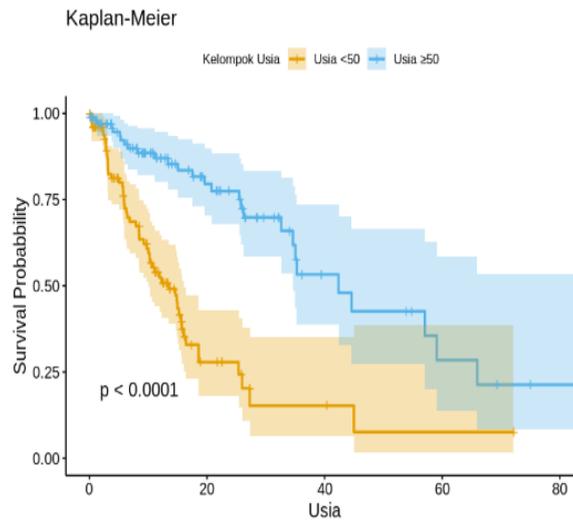


FIGURE 1. Kurva *Kaplan Meier* Penderita Kanker Payudara

Berdasarkan **Gambar 1**, plot variabel umur menunjukkan informasi bahwa kurva pasien yang berumur  $\geq 50$  tahun berada diatas pasien yang berumur  $< 50$  tahun yaitu dengan nilai *survival* 0.9 sampai 1. Sedangkan untuk pasien yang berumur  $< 50$  tahun mempunyai nilai *survival* 0.5 sampai 1. Sehingga dapat diketahui bahwa pasien yang berumur  $\geq 50$  memiliki peluang *survive* yang lebih besar dengan nilai survival 0.9 sampai 1 dibandingkan dengan pasien yang berumur  $< 50$  tahun dengan nilai survival

Plot variabel *treatment* menunjukkan bahwa pasien kanker payudara yang menerima *treatment* kemoterapi rejimen Antarsiklin *Base* memiliki nilai peluang sebesar 1, sedangkan untuk pasien yang menerima kemoterapi rejimen kombinasi Antarsiklin + Taxan memiliki nilai peluang diantara 0.8 sampai 1, dan untuk pasien yang menerima kemoterapi Taxan *Base* memiliki nilai peluang diantara 0.5 sampai 1. Dapat disimpulkan bahwa pasien yang menerima kemoterapi rejimen Antarsiklin *Base* memiliki peluang *survive* yang lebih tinggi dibandingkan dua jenis kemoterapi rejimen lainnya.

Plot variabel stadium menunjukkan bahwa terlihat garis kurva pada pasien dengan stadium IA memiliki nilai peluang sebesar 1. Sedangkan untuk pasien stadium IIA memiliki nilai peluang diantara 0.8 sampai 1. Dan juga untuk pasien stadium IIB memiliki nilai peluang diantara 0.7 sampai 1. Dapat diartikan bahwa pasien kanker

payudara dengan stadium IA memiliki peluang *survive* yang lebih tinggi dibandingkan pasien dengan stadium IIA dan IIB.

#### 4.2. Analisis *Mixture Cure Model* Semiparametrik AFT.

Analisis *Mixture Cure Model* Semiparametrik AFT dilakukan dengan tujuan untuk mengetahui proporsi pasien sembuh dan probabilitas survival pasien yang tidak sembuh pada pasien kanker payudara.

Akan dilakukan estimasi parameter model awal pada fungsi survival dengan algoritma EM. Estimasi menggunakan sampel *bootstrap* dengan  $nboot=15$ ,  $nboot=25$ , dan  $nboot=35$ . Diperoleh dengan estimasi sebagai berikut:

TABEL 1. Estimasi dengan  $nboot = 15$ ,  $nboot = 25$ ,  $nboot = 35$

Variabel	$nboot = 15$	$nboot = 25$	$nboot = 35$
<i>Cure probability model</i>			
(Intercept)	13.5027657	14.3289823	12.2371574
KelUmur	0.4937400	0.4839291	0.5587064
Stadium	4.5755656	4.8696627	0.41672287
Treatment	0.4870494	0.5412209	0.5441484
<i>Failure time distribution model</i>			
(Intercept)	0.4571561	0.10286354	0.4568497
KelUmur	0.2851164	0.3196299	0.2978336
Stadium	0.5644377	0.4919885	0.4781914
Treatment	0.1878518	0.2186291	0.2356550

Pemilihan  $nboot$  terbaik untuk estimasi dapat digunakan dengan memilih nilai *standard error* terkecil. Berdasarkan tabel 1.  $nboot$  terbaik yaitu  $nboot$  sebesar 15 dikarenakan nilai *standard error* terkecil dibandingkan dengan dua  $nboot$  lainnya. Sehingga diperoleh estimasi parameter untuk model *cure* awal adalah sebagai berikut:

TABEL 2. Hasil Estimasi Parameter Model Awal

Variabel	Estimasi	SE	P-Value
<i>Cure probability model</i>			
(Intercept)	-0.99813524	13.5027657	0.9410734
KelUmur	-0.62203567	0.4937400	0.2077254
Stadium	0.09100592	4.5755656	0.9841315
Treatment	-0.78610597	0.4870494	0.1065237
<i>Failure time distribution model</i>			
(Intercept)	0.4868278	0.5422810	0.36932372
KelUmur	-0.5080344	0.2851164	0.07477383
Stadium	0.1751687	0.5644377	0.75630099
Treatment	-0.4632247	0.1878518	0.01366675

Berdasarkan hasil estimasi pada **tabel 2**, dilakukan uji signifikansi parameter model awal. Uji signifikansi digunakan untuk melihat pengaruh dari variabel terhadap model *cure*. Uji hipotesis ( $H_0$ ) adalah variabel tidak signifikan terhadap model dan ( $H_1$ ) adalah variabel signifikan terhadap model. Dengan menggunakan tingkat kepercayaan 95%, diperoleh tolak ( $H_0$ ), karena *p-value* lebih besar dari  $\alpha$ . Sehingga dapat disimpulkan bahwa variabel tidak berpengaruh secara signifikan terhadap model *cure*.

Karena ketiga variabel tidak berpengaruh pada model awal, maka dilakukan eliminasi pada variabel yang mempunyai nilai *p-value* yang paling besar yaitu variabel stadium. Akan tetapi setelah dilakukan estimasi ulang, terdapat variabel yang tidak signifikan, maka dilakukan estimasi untuk mendapatkan model akhir dengan melakukan eliminasi variabel umur. Sehingga diperoleh hasil estimasi model akhir sebagai berikut:

TABEL 3. Estimasi Parameter Model Akhir

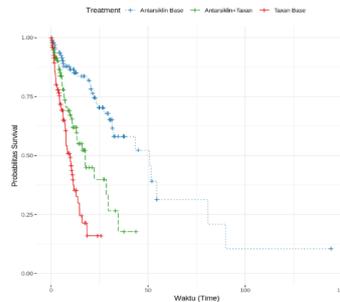
Variabel	Estimasi	SE	P-Value
<i>Cure probability model</i>			
(Intercept)	-0.9576244	0.2686952	0.0003652724
Treatment	-1.0506156	0.4332871	0.0153186699
<i>Failure time distribution model</i>			
(Intercept)	0.5393187	0.2399673	0.01246854
Treatment	-0.6316284	0.3322473	0.05729164

Berdasarkan hasil estimasi pada **tabel 3**, diperoleh bahwa dengan menggunakan tingkat kepercayaan 95% maka *p-value*  $0.015 <$  dari  $\alpha = 0.05$ . Dapat disimpulkan bahwa variabel *treatment* berpengaruh signifikan terhadap model.

Sehingga diperoleh persamaan model akhir dari model *cure* sebagai berikut

$$1 - \pi(z) = \exp(-\exp(-0.9576244 - 1.0506156T_{rt}))$$

dan kurva prediksi probabilitas survival seperti gambar dibawah ini:

FIGURE 2. Kurva Prediksi Survival Model Semiparametrik *Cure* Model

**Gambar 2** menunjukkan bahwa garis putus-putus kecil bagian atas adalah pasien dengan *treatment* Antarsiklin *Base*, garis putus-putus besar bagian tengah adalah pasien

dengan *treatment* Antarsiklin+Taxan, dan garis tebal pada bagian akhir adalah pasien dengan *treatment* Taxan *Base*. Oleh karena itu, diperoleh kesimpulan bahwa pasien penderita kanker payudara yang mendapatkan *treatment* kemoterapi rejimen Antarsiklin *Base* akan memiliki peluang *survival* atau ketahanan hidup yang lebih besar dibandingkan pasien yang mendapatkan *treatment* kemoterapi rejimen kombinasi Antarsiklin+Taxan dan Taxan *Base*.

Contoh perhitungan nilai proporsi pasien sembuh dan probabilitas *survival* pasien tidak sembuh dari pasien kanker payudara dengan menggunakan *mixture cure* model semiparametrik AFT adalah sebagai berikut:

- (1) Jika pasien kanker payudara mendapatkan *treatment* kombinasi Antarsiklin+Taxan dan dirawat inap selama 20 hari maka probabilitas *survival* sembuh dan tidak sembuh diperoleh

$$S_{pop}(20|x, z) = \pi(z) S(20|x) + 1 - \pi(z) = 0.9655$$

Dengan demikian probabilitas *survival* pasien kanker payudara jika diberi *treatment* kombinasi Antarsiklin+Taxan setelah 20 hari dirawat adalah 96.55%. Proporsi pasien sembuh dapat dihitung sebagai berikut

$$\begin{aligned} 1 - \pi(z) &= \exp(-\exp(-0.9576244 - 1.0506156(1))) \\ &= 0.8743889 \end{aligned}$$

sehingga diperoleh proporsi pasien sembuh setelah 20 hari adalah 87.44%. Sedangkan untuk probabilitas pasien tidak sembuh dapat dihitung sebagai berikut:

$$\begin{aligned} S_{pop}(20|x, z) &= \pi(z) S(20|x) + 1 - \pi(z) \\ 0.9655 &= \pi(z) S(20|x) + 0.8744 \\ \pi(z) S(20|x) &= 0.0911 \\ 0.1256 S(20|x) &= 0.0911 \\ S(20|x) &= 0.7253 \end{aligned}$$

Sehingga diperoleh probabilitas pasien tidak sembuh setelah 20 hari adalah 72.53%.

- (2) Jika pasien kanker payudara mendapatkan *treatment* Taxan *Base* dan dirawat inap selama 26 hari maka probabilitas *survival* sembuh dan tidak sembuh diperoleh

$$S_{pop}(26|x, z) = \pi(z) S(26|x) + 1 - \pi(z) = 0.9616$$

Dengan demikian probabilitas *survival* pasien kanker payudara jika diberi *treatment* Taxan *Base* setelah 26 hari dirawat adalah 96.16%. Proporsi pasien sembuh dapat dihitung sebagai berikut

$$\begin{aligned} 1 - \pi(z) &= \exp(-\exp(-0.9576244 - 1.0506156(0))) \\ &= 0.6812589 \end{aligned}$$

sehingga diperoleh proporsi pasien sembuh setelah 26 hari adalah 68.13%. Sedangkan untuk probabilitas pasien tidak sembuh dapat dihitung sebagai berikut:

$$\begin{aligned} S_{pop}(26|x, z) &= \pi(z) S(26|x) + 1 - \pi(z) \\ 0.9616 &= \pi(z) S(26|x) + 0.6813 \\ \pi(z) S(26|x) &= 0.2803 \\ 0.3187 S(26|x) &= 0.2803 \\ S(26|x) &= 0.8795 \end{aligned}$$

Sehingga diperoleh probabilitas pasien tidak sembuh setelah 26 hari adalah 87.95%.

- (3) Jika pasien kanker payudara mendapatkan *treatment* kombinasi Antarsiklin *Base* dan dirawat inap selama 68 hari maka probabilitas survival sembuh dan tidak sembuh diperoleh

$$S_{pop}(20|x, z) = \pi(z) S(20|x) + 1 - \pi(z) = 0.9551$$

Dengan demikian probabilitas survival pasien kanker payudara jika diberi *treatment* kombinasi Antarsiklin *Base* setelah 68 hari dirawat adalah 95.51%. Proporsi pasien sembuh dapat dihitung sebagai berikut

$$\begin{aligned} 1 - \pi(z) &= \exp(-\exp(-0.9576244 - 1.0506156(2))) \\ &= 0.9541408 \end{aligned}$$

sehingga diperoleh proporsi pasien sembuh setelah 68 hari adalah 95.41%. Sedangkan untuk probabilitas pasien tidak sembuh dapat dihitung sebagai berikut:

$$\begin{aligned} S_{pop}(68|x, z) &= \pi(z) S(68|x) + 1 - \pi(z) \\ 0.9551 &= \pi(z) S(68|x) + 0.9541 \\ \pi(z) S(68|x) &= 0.0001 \\ 0.0459 S(68|x) &= 0.0001 \\ S(68|x) &= 0.0022 \end{aligned}$$

Sehingga diperoleh probabilitas pasien tidak sembuh setelah 68 hari adalah 0%.

## 5. KESIMPULAN

Berdasarkan hasil analisis yang telah dilakukan, maka dapat ditarik kesimpulan sebagai berikut:

- (1) Model *mixture cure* dapat ditunjukkan seperti berikut:

$$S_{pop}(t|x, z) = \pi(z) S(t|x) + 1 - \pi(z)$$

dengan  $1 - \pi(z) = 1 - \exp(-\exp(\alpha + \beta x))$  menyatakan proporsi dari pasien sembuh, dan  $S(t|x)$  menyatakan probabilitas pasien tidak sembuh. Estimasi parameter yang digunakan adalah Algoritma EM.

- (2) Estimasi model *cure* pada pasien penyakit kanker payudara diperoleh sebagai berikut:

$$1 - \pi(z) = \exp(-\exp(-0.9576244 - 1.0506156 T_{rt}))$$

- (3) Variabel yang berpengaruh signifikan terhadap pasien sembuh adalah variabel *treatment*, akan tetapi pada *failure time model* variabel *treatment* tidak berpengaruh secara signifikan terhadap pasien tidak sembuh dari penyakit kanker payudara.

### References

- [1] Boag, J. W., Maximum Likelihood Estimates of The Proportion of Patients Cured By Cancer Therapy, *Journal of The Royal Statistical Society* (1949), 11:15-53.
- [2] Cai, C; Peng, Y; Zou, Y and Zhang, J. SMCURE: An R-Package for Estimating Semiparametric Mixture Cure Models, (2012) *Computer Methods and Programs in Biomedicine*, 108:1255-1260
- [3] Danardono, *Analisis Data Survival*, Diktat Kuliah, (2012) UGM, Yogyakarta
- [4] Farewell, V.T. Mixture models in survival analysis. Are they worth the risk ? *The Canadian Journal of Statistics* (1949)14, 257 - 262.
- [5] Harlan, J., *Analisis Survival*, Gunadarma, Depok.
- [6] Kleinbaum, D. G., dan Klein, M., *Survival Analysis: A Self-Learning Text*, 3rd Edition (2012) Springer, New York.
- [7] Li CS, Taylor JMG. A semi-parametric accelerated failure time cure model. *Statistics in Medicine* 2002; 21:3235–3247. DOI: 10.1002/sim.1260
- [8] Mukrinin, W.U , Dampak Penundaan Jadwal Kemoterapi Terhadap Rekurensi dan Ketahanan Hidup Pada Pasien Kanker Payudara di RSUP. Dr. Sardjito Yogyakarta, (2022), Universitas Gadjah Mada, Yogyakarta.
- [9] Zhang, J., dan Peng, Y. A New Estimation Method for The Semiparametric Accelerated Failure Time Mixture Cure Model, *Statistics in Medicine*, 26:3157- 3171

AGI DEMPITA: Universitas Gadjah Mada.  
E-mails: agi.dempita@mail.ugm.ac.id

DANARDONO: Universitas Gadjah Mada.  
E-mails: danardono@ugm.ac.id