

ANALISIS SURVIVAL UNTUK PARAMETER SKALA DARI DISTRIBUSI WEIBULL MENGGUNAKAN MLE DAN METODE BAYESIAN

Survival Analysis for Scale Parameter of Weibull Distribution with MLE and Bayesian Method

Ferra Yanuar^{1*}, Sisca Wulandari², Izzati Rahmi HG³.

^{1,2,3} Prodi Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Andalas
Kampus Limau Manis, Padang, 25163, Indonesia

Correspondence author e-mail: ^{2*} ferrayanuar@sci.unand.ac.id

Abstrak

Pemodelan data tahan hidup (analisis *survival*) perlu dan penting untuk dilakukan. Data tahan hidup pada umumnya diasumsikan berdistribusi Weibull. Pendekatan Bayesian telah diimplementasikan untuk mengestimasi parameter dalam analisis *survival* tersebut. Penelitian ini bertujuan untuk membandingkan metode *Maximum Likelihood* dan Bayesian dengan menggunakan Invers Gamma sebagai prior konjugat untuk mengestimasi fungsi *survival* parameter skala distribusi Weibull. Perbandingan dilakukan melalui studi simulasi. Performa terbaik dari kedua estimator tersebut dipilih berdasarkan nilai bias absolut terendah dan mean square error. Dua sampel ukuran yang berbeda dihasilkan untuk menggambarkan data waktu hidup yang digunakan dalam penelitian ini. Hasil penelitian menunjukkan bahwa *Maximum Likelihood* adalah penduga terbaik dibandingkan Bayes dengan distribusi Invers Gamma sebagai konjugat sebelumnya.

Kata Kunci : *Distribusi Weibull, Analysis Survival, Maximum Likelihood Estimation, Bayesian, Absolute bias, Mean Square Error.*

Abstract

Modeling of survival data is necessary and important to do. Survival data is generally assumed to have a Weibull distribution. Bayesian approach has been implemented to estimate the parameter in such this survival analysis. This study purposes to compare the performance of the Maximum Likelihood and Bayesian using Invers Gamma as prior conjugate for estimating the survival function of scale parameter of Weibull distribution. The comparisons are made through simulation study. The best performance of both estimators is chosen based on the lowest value of absolute bias and the mean square error. Two different size samples are generated to illustrate the life time data which are used in this study. This study results that maximum likelihood is the best estimator compared to Bayes with Invers Gamma distribution as conjugate prior.

Keywords: *Weibull distribution, Survival analysis, Maximum Likelihood, Bayesian, Absolute bias, Mean Square Error.*

Article info:

Submitted: 01st January 2021

Accepted: 26th February 2021

How to cite this article:

F. Yanuar, S. Wulandari, I. Rahmi HG., "ANALISIS SURVIVAL UNTUK PARAMETER SKALA DARI DISTRIBUSI WEIBULL MENGGUNAKAN MLE DAN METODE BAYESIAN", *BAREKENG: J. Il. Mat. & Ter.*, vol. 15, no. 1, pp. 147-156, Mar. 2021.



This work is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International License](https://creativecommons.org/licenses/by-sa/4.0/).
Copyright © 2021 Ferra Yanuar, Sisca Wulandari, Izzati Rahmi HG.

1. PENDAHULUAN

Analisis *survival* adalah suatu metode untuk menganalisis data yang berhubungan dengan waktu, mulai dari *start-point* sampai dengan terjadinya suatu kejadian khusus atau *end-point*. Dalam analisis *survival* terdapat tiga istilah yang perlu dipahami. Pertama, *survival time* atau waktu suatu objek untuk tetap bertahan dalam periode pengamatan. Kedua, kejadian (*event*) yang menjadi fokus dalam penelitian. Ketiga, suatu kejadian terjadi apabila peneliti mempunyai waktu ketahanan yang menjadi objek penelitian, walaupun sesungguhnya peneliti tidak mengetahui waktu ketahanan yang pasti, istilah ini disebut dengan sensor [1]–[3].

Pada analisis *survival* sering terjadi data tersensor (*censored data*) yaitu adanya informasi mengenai waktu ketahanan hidup tetapi tidak diketahui secara pasti berapa lama waktu ketahanannya [4]. Penyebabnya adalah hingga pengamatan berakhir belum muncul kejadian yang diinginkan. Analisis *survival* dalam penelitian ini menggunakan model *survival* parametrik, yakni model yang waktu *survival*-nya diasumsikan mengikuti suatu distribusi yang diketahui. Asumsi waktu *survival* yang sering digunakan antara lain distribusi Weibull, Eksponensial, Logistik dan lain sebagainya. Pada penelitian ini digunakan analisis *survival* untuk data yang berdistribusi Weibull. Distribusi Weibull adalah salah satu distribusi yang paling banyak digunakan untuk memodelkan dan menganalisis waktu hidup terutama dalam kasus medis. Distribusi Weibull biasanya digunakan untuk menyelesaikan masalah-masalah yang menyangkut lama waktu ketahanan hidup suatu objek mulai berfungsi hingga tidak bisa digunakan lagi atau mati. Distribusi Weibull digunakan secara efektif untuk menganalisis data waktu ketahanan hidup khususnya untuk data tersensor. Distribusi Weibull memiliki dua parameter, yakni parameter bentuk (β) dan parameter skala (θ) [4], [6], [7].

Kajian terkait estimasi parameter model ketahanan hidup (*survival analysis*) dari distribusi Weibull sudah banyak dilakukan oleh peneliti sebelumnya. Diantaranya yaitu kajian oleh Guure dan Ibrahim [1] yang mengkaji perbandingan hasil estimasi parameter dari distribusi Weibull dengan metode MLE dan metode Bayesian (menggunakan prior non informatif yaitu prior Jeffrey dan modifikasinya). Hazhiah, dkk [5] mengkaji proses pengestimasi parameter dari distribusi Weibull terhadap data tak tersensor dan data tersensor tipe II dengan metode Bayes (distribusi prior yang digunakan yaitu distribusi Gamma). Ismed dan Yudi [8] menggunakan metode Bayes dalam membuat model sistem risiko daya tahan yang bersaing (*computing risk systems*) dengan data diasumsikan berdistribusi Weibull. Kemudian Ahmed [6] dan Thamrin, dkk, [9] mengestimasi parameter dari distribusi Weibull dengan metode Bayesian *survival* (menggunakan prior non informatif distribusi Gamma) dan metode maksimum likelihood tetapi kedua hasil kajian mereka saling bertolak belakang.

Untuk itulah penelitian ini dilakukan untuk mengkaji lagi terkait perbandingan hasil pendugaan parameter dari distribusi Weibull dengan menggunakan metode MLE dan metode Bayes. Pada kajian ini akan diestimasi parameter skala dengan mengasumsikan parameter bentuk diketahui. Adapun distribusi prior yang digunakan pada metode Bayes adalah prior konjugat yaitu distribusi Invers Gamma. Data yang digunakan dalam analisis adalah data bangkitan yang diasumsikan memenuhi syarat-syarat sebagai data daya tahan hidup. Penduga terbaik dari kedua analisis metode akan dibandingkan berdasarkan nilai absolut bias dan *Mean Square Error* (MSE) terkecil. Dari metode terbaik yang diperoleh selanjutnya akan diestimasi nilai dari fungsi *survival* dan fungsi *hazard* untuk data bangkitan tersebut.

2. METODE PENELITIAN

Langkah 1: Studi Literatur

Studi literatur yang dilakukan oleh peneliti adalah dengan mempelajari pembahasan mengenai konsep-konsep terkait analisis *survival*, seperti fungsi kepekatan peluang, fungsi *survival*, fungsi *hazard*, data tersensor, metode MLE (*Maximum Likelihood Estimator*), dan metode Bayes. Berikut diuraikan secara ringkas definisi atau teori dasar terkait konsep-konsep tersebut.

a. Fungsi Kepekatan Peluang [10]

Fungsi kepekatan peluang didefinisikan sebagai limit dari peluang suatu objek mati atau gagal dalam interval waktu t sampai $t + \Delta t$, dengan waktu (T) merupakan peubah acak. Jika T merupakan peubah acak non-negatif pada interval $[0; 1)$, maka $F(t)$ merupakan fungsi distribusi kumulatif kontinu dari T yaitu:

$$F(t) = P(T \leq t) = \int_0^t f(x) dx \quad (1)$$

Maka fungsi kepekatan peluang dinyatakan dengan:

$$f(t) = \lim_{\Delta t \rightarrow 0} \left[\frac{P(t < T < (t + \Delta t))}{\Delta t} \right] = \lim_{\Delta t \rightarrow 0} \left[\frac{F(t + \Delta t) - F(t)}{\Delta t} \right] \quad (2)$$

b. Fungsi Survival [11]

Fungsi *survival* merupakan peluang suatu objek yang masih dapat bertahan hidup sampai dengan waktu $t (t > 0)$. Misal T adalah peubah acak dari waktu ketahanan hidup suatu individu dalam interval $[0; 1)$, maka fungsi distribusi kumulatif $F(t)$ untuk distribusi kontinu dengan fungsi kepekatan peluang $f(t)$ dinyatakan sebagai berikut.

$$F(t) = P(T \leq t) \quad (3)$$

Oleh karena itu fungsi *survival* $S(t)$ yang didefinisikan dengan:

$$\begin{aligned} S(t) &= P(T > t) \\ &= 1 - P(T \leq t) \\ &= 1 - F(t) \end{aligned} \quad (4)$$

c. Fungsi Hazard [1]

Fungsi *hazard* merupakan peluang suatu objek tertentu mengalami kejadian kegagalan (mati) dalam selang waktu yang singkat t sampai $t + \Delta t$ jika diketahui objek belum mengalami kejadian sampai waktu t [14].

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t < T \leq t + \Delta t | T > t)}{\Delta t}$$

Berdasarkan teori peluang bersyarat, diperoleh:

$$h(t) = \frac{f(t)}{S(t)} \quad (5)$$

d. Data Tersensor [12]

Data tersensor merupakan data yang diperoleh dari pengamatan yang tidak diamati secara menyeluruh dikarenakan objek penelitian tak dapat diambil datanya, atau sampai akhir penelitian objek tersebut belum mengalami kejadian tertentu. Terdapat tiga alasan terjadinya suatu penyensoran, yaitu :

- objek pengamatan yang diamati tidak mengalami suatu kejadian sampai penelitian berakhir (loss to follow-up),
- objek pengamatan hilang selama penelitian,
- objek mengalami kegagalan lain yang menyebabkan pengamatan tidak dapat diteruskan.

Penyensoran merupakan suatu hal yang membedakan analisis *survival* dengan analisis statistik lainnya. Penyensoran dilakukan untuk mengatasi beberapa permasalahan dalam suatu analisis, misalnya dalam suatu penelitian dibutuhkan waktu yang lama untuk mendapatkan data yang lengkap sampai objek pengamatan mengalami suatu kejadian yang diinginkan. Terdapat tiga macam penyensoran pada analisis *survival*, yaitu:

- Data Tersensor Kanan
Data *survival* biasanya merupakan data yang tersensor kanan. Sensor kanan dapat terjadi karena beberapa alasan, yaitu:
 - objek pengamatan belum mengalami suatu kejadian sampai masa penelitian berakhir,
 - objek pengamatan keluar pada saat masa penelitian berlangsung,
 - objek pengamatan meninggal pada saat penelitian, akan tetapi penyebab meninggal tidak berhubungan dengan kejadian yang diperhatikan.

b. Data Tersensor Kiri

Data tersensor kiri terjadi ketika objek pengamatan tidak teramati pada awal waktu pengamatan, akan tetapi sebelum penelitian berakhir semua kejadian sudah dapat diamati secara penuh. Dapat pula dikatakan bahwa kejadian yang ingin diperhatikan pada objek pengamatan tersebut sudah terjadi saat objek pengamatan tersebut sudah masuk ke dalam penelitian.

c. Data Tersensor Interval

Sensor interval terjadi ketika suatu kejadian yang diamati pada objek pengamatan terjadi pada selang waktu tertentu.

e. Metode *Maximum Likelihood Estimation* (MLE) [13], [14]

Salah satu metode dalam penaksiran parameter adalah *Maximum Likelihood Estimation* (MLE). Prinsip dari MLE adalah menemukan estimasi dari suatu parameter. Misalkan akan ditentukan estimasi dari $\hat{\theta}$ dengan cara memaksimalkan fungsi likelihood dan kemudian disamakan dengan nol. Fungsi *likelihood* dinyatakan dengan [6, 7]:

$$L(\theta) = f(x_1; \theta) \dots f(x_n; \theta) = \prod_{i=1}^n f(x_i; \theta) \quad (6)$$

Dalam analisis *survival* dengan data tersensor, konstruksi fungsi *likelihood* berbeda dari fungsi *likelihood* yang biasa. Asumsi utama yang dibutuhkan untuk mengkonstruksi fungsi *likelihood* untuk data *survival* yang tersensor adalah waktu hidup dan waktu sensor saling bebas. Jika asumsi tersebut tidak terpenuhi, maka dibutuhkan teknik lainnya. Pada data *survival* yang mengandung pengamatan tersensor kanan fungsi *likelihood*nya adalah:

$$L = \prod_{i=1}^n [f(x_i)]^{\delta_i} [S(x_i)]^{1-\delta_i} \quad (7)$$

dengan δ_i adalah indikator sensor, $\delta_i = 1$ untuk tersensor dan $\delta_i = 0$ untuk tidak tersensor, $S(x_i)$ merupakan fungsi *survival* untuk x_i .

Nilai parameter $\hat{\theta}$ dapat diperoleh dengan memaksimalkan fungsi kepekatan peluang bersama atau disebut *Maximum Likelihood Estimation* (MLE). Hal tersebut dilakukan dengan metode turunan pertama dari fungsi *likelihood*-nya terhadap parameternya sama dengan nol. Namun biasanya sulit untuk mencari turunan fungsi *likelihood* sehingga yang dilakukan adalah menentukan nilai maksimum dari logaritma natural fungsi *likelihood* tersebut atau disebut dengan fungsi *log-likelihood*. Fungsi *log-likelihood* merupakan fungsi kepekatan peluang bersama yang diubah menjadi bentuk logaritma, tujuannya untuk mempermudah dalam menaksir parameter. Fungsi *log-likelihood* yang mengandung parameter yang akan diestimasi θ , ditulis dalam bentuk:

$$\ln L(\theta) = \prod_{i=1}^n \ln f(x_i, \theta) \quad (8)$$

f. Metode Bayes [10], [15]

Metode Bayes merupakan suatu metode yang digunakan untuk menduga parameter dengan menggabungkan informasi yang terkandung dalam objek dengan informasi yang telah tersedia sebelumnya. Pada metode Bayes, jika suatu populasi mengikuti distribusi tertentu dengan suatu parameter didalamnya, maka kemungkinan parameter mengikuti suatu distribusi peluang tertentu disebut sebagai prior. Distribusi prior merupakan distribusi awal yang memberi informasi mengenai parameter. Pada kajian ini menggunakan distribusi prior konjugat yaitu distribusi prior yang diperoleh dari hasil identifikasi pola fungsi *likelihood*-nya. Metode Bayes menggabungkan fungsi *likelihood* dengan distribusi prior dari parameter yang akan diduga sehingga didapatkan distribusi posterior yang akan menjadi dasar pendugaan parameter.

Distribusi posterior merupakan fungsi kepekatan peluang bersyarat dari suatu parameter yang akan diestimasi θ jika diberikan pengamatan $\mathbf{x} = (x_1, x_2, \dots, x_n)$ yang diformulasikan dengan [16]:

$$f(\theta|\mathbf{x}) = \frac{f(\mathbf{x}|\theta)f(\theta)}{\int f(\mathbf{x}|\theta)f(\theta)d\theta} = \frac{f(\mathbf{x}, \theta)}{f(\mathbf{x})} \quad (9)$$

Mean dari distribusi posterior $f(\theta|x_1, x_2, \dots, x_n)$ dinyatakan dengan T, disebut penduga Bayes untuk $\tau(\theta)$.

Langkah 2 : Menentukan formula fungsi *hazard* dan fungsi *survival*, dan MSE untuk parameter skala θ yang berdistribusi Weibull (dengan parameter bentuk β diasumsikan diketahui).

Dengan berdasarkan konsep teori sebagaimana yang telah diuraikan pada Langkah 1 di atas, selanjutnya akan ditentukan rumusan konsep-konsep tersebut untuk distribusi Weibull. Sebagaimana diketahui distribusi Weibull memiliki dua parameter yaitu parameter skala θ dan parameter bentuk β . Misalkan $\mathbf{x} = (x_1, x_2, \dots, x_n)$ merupakan nilai dari variabel acak berdistribusi Weibull, maka fungsi kepekatan peluangnya adalah [9], [10]:

$$f(x; \theta) = \frac{\beta}{\theta} x^{\beta-1} e^{-\frac{x^\beta}{\theta}} \text{ dengan } \theta, \beta > 0 \text{ dan } x \geq 0$$

Fungsi likelihood dari distribusi Weibull untuk data tersensor kanan memiliki bentuk sebagai berikut:

$$\begin{aligned} L(\mathbf{x}; \theta) &= \prod_{i=1}^n [f(x_i)]^{\delta_i} [S(x_i)]^{1-\delta_i} \\ &= \left(\frac{\beta}{\theta}\right)^{\sum_{i=1}^n \delta_i} \prod_{i=1}^n (x_i^{(\beta-1)\delta_i}) e^{-\frac{\sum_{i=1}^n x_i^\beta}{\theta}} \end{aligned} \quad (10)$$

Selanjutnya akan ditentukan formula untuk fungsi hazard, fungsi survival serta MSE untuk parameter skala θ yang tersebut, yang merupakan parameter dari distribusi Weibull. Adapun pada kajian ini, parameter bentuk β diasumsikan diketahui.

Langkah 3 : Membangkitkan Data.

Setelah diperoleh formula untuk mengestimasi parameter skala θ pada Langkah 2 di atas, kemudian dibangkitkan sekelompok data untuk mengilustrasikan penerapan dari formula tersebut. Data yang digunakan pada penelitian ini adalah data bangkitan dari program R berjumlah 25 dan 50 data sampel yang diasumsikan sebagai data daya tahan hidup yang berdistribusi Weibull dengan menetapkan parameter bentuk $\beta = 1$ dan parameter skala $\theta = 1$. Nilai kedua parameter ditetapkan sama dengan 1 adalah untuk kemudahan dalam estimasi dan mengilustrasikan metode. Data jumlah 25 untuk mewakili data berukuran kecil dan data jumlah 50 untuk mewakili data berukuran besar. Untuk merekayasa munculnya data tersensor kanan pada data bangkitan ini, ditetapkan bahwa untuk data bernilai kurang atau sama dengan satu maka data dikatakan tersensor, dan data yang bernilai besar dari satu dikatakan data tidak tersensor. Pada studi simulasi ini juga dipilih nilai parameter skala $\theta = \{0,5, 0,8, 1, 1,5\}$ untuk kemudian nilai-nilai ini nantinya akan dibandingkan dengan nilai estimasi θ yang diperoleh dari formula hasil Langkah 4 berikut. Parameter skala θ ini akan diestimasi dengan menggunakan metode MLE dan metode Bayes dengan memilih distribusi Invers Gamma sebagai prior konjugat. Pada studi simulasi dengan menggunakan program R ini digunakan data dimana banyaknya data tersensor kanan pada setiap ukuran data sekitar 40%.

Langkah 4: Pengolahan Data.

Pengolahan data dilakukan dengan berdasarkan formula yang diperoleh pada Langkah 2 di atas. Adapun perhitungan untuk pengolahan data adalah menggunakan aplikasi program R.

3. HASIL DAN PEMBAHASAN

3.1. Analisis *Survival* Distribusi Weibull Dengan MLE

Misalkan $\mathbf{x} = (x_1, x_2, \dots, x_n)$ merupakan nilai dari variabel acak berdistribusi Weibull, dengan fungsi likelihoodnya sesuai persamaan (10). Dengan menggunakan fungsi likelihood tersebut, dapat diestimasi nilai harapan untuk parameter skala θ dimana parameter bentuk β diketahui, yaitu [6]:

$$\hat{\theta}_{ML} = \frac{\sum_{i=1}^n x_i^\beta}{\sum_{i=1}^n \delta_i} \quad (11)$$

Kemudian dirumuskan nilai harapan dan ragam untuk estimator $\hat{\theta}_{ML}$ dengan metode MLE masing-masing adalah:

$$\begin{aligned} E(\hat{\theta}_{ML}) &= E\left(\frac{\sum_{i=1}^n x_i^\beta}{\sum_{i=1}^n \delta_i}\right) \\ &= \frac{1}{\sum_{i=1}^n \delta_i} (n\theta) \end{aligned} \quad (12)$$

dan

$$\begin{aligned} Var(\hat{\theta}_{ML}) &= Var\left(\frac{\sum_{i=1}^n x_i^\beta}{\sum_{i=1}^n \delta_i}\right) \\ &= \frac{1}{(\sum_{i=1}^n \delta_i)^2} (n\theta^2) \end{aligned} \quad (13)$$

Selanjutnya dapat dirumuskan juga fungsi *survival* dan fungsi *hazard* untuk parameter $\hat{\theta}$ dari metode MLE ini masing-masing sebagai berikut :

$$\begin{aligned} \hat{S}_{ML}(x_i; \hat{\theta}_{ML}) &= e^{-\frac{x_i^\beta}{\hat{\theta}_{ML}}} \\ &= e^{-\frac{x_i^\beta \sum_{i=1}^n \delta_i}{\sum_{i=1}^n x_i^\beta}} \end{aligned} \quad (14)$$

dan

$$\begin{aligned} \hat{h}_{ML}(x_i; \hat{\theta}_{ML}) &= \beta \frac{1}{\hat{\theta}_{ML}} x_i^{\beta-1} \\ &= \beta \frac{\sum_{i=1}^n \delta_i}{\sum_{i=1}^n x_i^\beta} x_i^{\beta-1} \end{aligned} \quad (15)$$

3.2 Analisis Survival Distribusi Weibull Dengan Metode Bayesian

Pada metode Bayesian, prior yang digunakan adalah prior konjugat distribusi Invers Gamma dengan memilih parameter $\theta \sim IG(1,1)$. Fungsi kepekatan peluang dari $\theta \sim IG(1,1)$ dapat dinyatakan dalam bentuk [10]:

$$\begin{aligned} f(\theta) &= \frac{1^1}{\Gamma(1)} \theta^{-(1+1)} \exp\left(-\frac{1}{\theta}\right) \\ &= \theta^{-2} \exp\left(-\frac{1}{\theta}\right), \theta > 0 \end{aligned}$$

Fungsi kepekatan peluang bersama adalah sebagai berikut:

$$\begin{aligned} f(\mathbf{x}; \theta) &= \left(\left(\frac{\beta}{\theta}\right)^{\sum_{i=1}^n \delta_i} \prod_{i=1}^n (x_i^{(\beta-1) \delta_i}) e^{-\frac{\sum_{i=1}^n x_i^\beta}{\theta}} \right) (\theta^{-2} e^{-\frac{1}{\theta}}) \\ &= \beta^{\sum_{i=1}^n \delta_i} \prod_{i=1}^n (x_i^{(\beta-1) \delta_i}) e^{-\frac{1}{\theta}(\sum_{i=1}^n x_i^\beta + 1)} \theta^{-(\sum_{i=1}^n \delta_i + 2)} \end{aligned} \quad (16)$$

Fungsi kepekatan peluang marjinal dari persamaan (16) adalah:

$$\begin{aligned} f(x) &= \int_0^\infty \beta^{\sum_{i=1}^n \delta_i} \prod_{i=1}^n x_i^{(\beta-1) \delta_i} e^{-\frac{1}{\theta}(\sum_{i=1}^n x_i^\beta + 1)} \theta^{-(\sum_{i=1}^n \delta_i + 2)} d\theta \\ &= \beta^{\sum_{i=1}^n \delta_i} \prod_{i=1}^n (x_i^{(\beta-1) \delta_i}) \left(\sum_{i=1}^n x_i^\beta + 1\right)^{-(\sum_{i=1}^n \delta_i + 2)} \Gamma(\sum_{i=1}^n \delta_i + 1) \end{aligned} \quad (17)$$

Diperoleh distribusi posterior:

$$f(\theta|x) = \frac{f(\mathbf{x};\theta)}{f(\mathbf{x})}$$

$$= \frac{(\sum_{i=1}^n x_i^\beta + 1)^{(\sum_{i=1}^n \delta_i + 1)}}{\Gamma(\sum_{i=1}^n \delta_i + 1)} \theta^{-(\sum_{i=1}^n \delta_i + 2)} e^{-\frac{1}{\theta}(\sum_{i=1}^n x_i^\beta + 1)} \quad (18)$$

Dengan menggunakan prior $IG(1,1)$ diperoleh distribusi posterior $\theta_{IG} | X \sim IG(\sum_{i=1}^n \delta_i + 1, \sum_{i=1}^n x_i^\beta + 1)$. Sehingga untuk $\theta \sim IG(\sum_{i=1}^n \delta_i + 1, \sum_{i=1}^n x_i^\beta + 1)$ diperoleh nilai harapan dan ragam posterior untuk masing-masing adalah sebagai berikut:

$$\hat{\theta}_{IG} = E(\theta_{IG}) = \frac{\sum_{i=1}^n x_i^\beta + 1}{\sum_{i=1}^n \delta_i}, \quad \sum_{i=1}^n \delta_i > 0 \quad (19)$$

$$Var(\theta_{IG}) = \frac{(\sum_{i=1}^n x_i^\beta + 1)^2}{(\sum_{i=1}^n \delta_i)^2 (\sum_{i=1}^n \delta_i - 1)}, \quad \sum_{i=1}^n \delta_i > 1 \quad (20)$$

Selanjutnya akan ditentukan nilai harapan dan ragam bagi penduga Bayesian. Nilai harapan untuk $\hat{\theta}_{IG}$ dengan metode Bayesian adalah :

$$E(\hat{\theta}_{IG}) = E\left(\frac{\sum_{i=1}^n x_i^\beta + 1}{\sum_{i=1}^n \delta_i}\right) = \frac{1}{\sum_{i=1}^n \delta_i} (n\theta + 1) \quad (21)$$

Ragam untuk $\hat{\theta}_{IG}$ dengan metode Bayesian adalah:

$$Var(\hat{\theta}_{IG}) = Var\left(\frac{\sum_{i=1}^n x_i^\beta + 1}{\sum_{i=1}^n \delta_i}\right) = \frac{1}{(\sum_{i=1}^n \delta_i)^2} (n\theta^2) \quad (22)$$

Estimasi parameter $\hat{\theta}$ fungsi *survival* metode Bayesian sebagai berikut:

$$\begin{aligned} \hat{S}_B(x_i; \hat{\theta}_{IG}) &= e^{-\frac{x_i^\beta}{\hat{\theta}_{IG}}} \\ &= \exp\left(-\frac{x_i^\beta \sum_{i=1}^n \delta_i}{\sum_{i=1}^n x_i^\beta + 1}\right) \end{aligned} \quad (23)$$

Estimasi parameter $\hat{\theta}$ fungsi *hazard* metode Bayesian sebagai berikut:

$$\begin{aligned} \hat{h}_B(x_i; \hat{\theta}_{IG}) &= \beta \frac{1}{\hat{\theta}_{IG}} x_i^{\beta-1} \\ &= \beta \frac{\sum_{i=1}^n \delta_i}{\sum_{i=1}^n x_i^\beta + 1} x_i^{\beta-1} \end{aligned} \quad (24)$$

3.3 Indikator Keباikan Estimator

Indikator kebaikan estimator yang digunakan pada penelitian ini adalah nilai absolut bias dan nilai MSE (*Mean Square Error*) yang kecil [1]. Bias diberikan oleh $b(\hat{\theta}) = E(\hat{\theta}) - \theta$ dengan $\hat{\theta}$ merupakan penduga bagi θ . Formula nilai absolut bias untuk metode MLE adalah :

$$bias(\hat{\theta}_{ML}) = E(\hat{\theta}_{ML}) - \theta = \frac{1}{\sum_{i=1}^n \delta_i} (n\theta) - \theta \quad (25)$$

Dan nilai absolut bias untuk metode Bayesian adalah :

$$bias(\hat{\theta}_{IG}) = E(\hat{\theta}_{IG}) - \theta = \frac{1}{\sum_{i=1}^n \delta_i} (n\theta + 1) - \theta \quad (26)$$

Diketahui ragam penduga MLE adalah seperti yang dirumuskan pada persamaan (13), sehingga MSE untuk metode MLE adalah :

$$\begin{aligned} MSE(\hat{\theta}_{ML}) &= Var(\hat{\theta}_{ML}) + (bias(\hat{\theta}_{ML}))^2 \\ &= \frac{1}{(\sum_{i=1}^n \delta_i)^2} (n\theta^2) + \left(\frac{1}{\sum_{i=1}^n \delta_i} (n\theta) - \theta\right)^2 \end{aligned} \quad (27)$$

Adapun ragam penduga Bayesian adalah seperti yang dirumuskan pada persamaan (22), sehingga formula MSE untuk penduga Bayes adalah :

$$MSE(\hat{\theta}_{IG}) = Var(\hat{\theta}_{IG}) + (bias(\hat{\theta}_{IG}))^2$$

$$= \frac{1}{(\sum_{i=1}^n \delta_i)^2} (n\theta^2) + \left(\frac{1}{\sum_{i=1}^n \delta_i} (n\theta + 1) - \theta \right)^2 \quad (28)$$

3.4 Studi Simulasi Penerapan Analisis Survival Distribusi Weibull

Pada bagian ini akan diilustrasikan penerapan metode yang telah diuraikan pada bagian sebelumnya pada data bangkitan dengan mengambil beberapa ukuran data yaitu 25 dan 50. Pada proses pembangkitan data, ditetapkan dan dipilih nilai untuk parameter skala (θ) yaitu $\theta = \{0,5, 0,8, 1, 1,5\}$. Hasil estimasi untuk $\hat{\theta}$ kemudian dibandingkan dengan nilai θ yang ditetapkan tersebut, perbedaannya dinyatakan dengan nilai Absolut Bias. Adapun nilai MSE diperoleh dengan menggunakan persamaan (28). Tabel 1 menyajikan hasil perbandingan absolut bias dan MSE hasil penerapan metode MLE dan metode Bayesian.

Tabel 1. Hasil Perbandingan Absolut Bias dan MSE Menggunakan Metode MLE dan Metode Bayesian

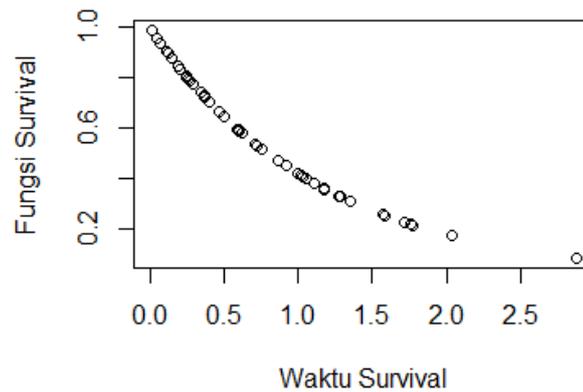
Banyak data (n)	θ	Absolut Bias		MSE	
		MLE	Bayesian	MLE	Bayesian
25	0,5	0,02083	0,06250	0,0113	0,01480
	0,8	0,20000	0,25000	0,08000	0,10250
	1,0	0,47058	0,52941	0,30795	0,36678
	1,5	1,62500	1,70833	3,03125	3,30902
	3,0	15,7500	16,000	262,125	270,063
50	0,5	0,02083	0,04167	0,00590	0,00720
	0,8	0,20000	0,22500	0,06000	0,07062
	1,0	0,47058	0,50000	0,26470	0,29325
	1,5	1,27778	1,31481	1,78704	1,88305
	3,0	7,00000	7,06670	51,0000	51,9382

Dari Tabel 1, diatas diperoleh bahwa nilai absolut bias hasil estimasi dari MLE cenderung menghasilkan nilai yang lebih kecil daripada hasil metode Bayes untuk kedua ukuran data dan untuk semua nilai θ terpilih. Dengan demikian dapat disimpulkan bahwa estimasi parameter untuk kasus analisis waktu ketahanan hidup (*survival analysis*) pada distribusi Weibull dengan metode MLE akan menghasilkan nilai estimasi yang lebih baik daripada metode Bayesian yang menggunakan prior konjugat dari distribusi Invers Gamma. Hasil kajian ini linier dengan hasil kajian yang dilakukan oleh Ahmed [6] tetapi bertentangan dengan kajian yang dihasilkan oleh Thamrin dkk [9].

Selanjutnya karena MLE mampu menghasilkan nilai estimasi yang lebih baik, maka analisis *survival* lanjutan dilakukan berdasarkan metode MLE ini, yaitu mengestimasi fungsi *survival* dan fungsi *hazard* untuk parameter skala $\hat{\theta}$ tersebut. Rumusan untuk mengestimasi fungsi *survival* adalah menggunakan persamaan (14) dan untuk mengestimasi fungsi *hazard* digunakan persamaan (15), masing-masing dinyatakan sebagai berikut :

$$\hat{S}_{ML}(x_i; \hat{\theta}_{ML}) = e^{-\frac{x_i^\beta \sum_{i=1}^n \delta_i}{\sum_{i=1}^n x_i^\beta}} = e^{-\frac{x^{(1)} \times 27}{288.7744^{(1)}}$$

Dengan mengambil beberapa nilai untuk x , fungsi *survival* ini dapat diilustrasikan dengan grafik fungsi *survival*, berikut:

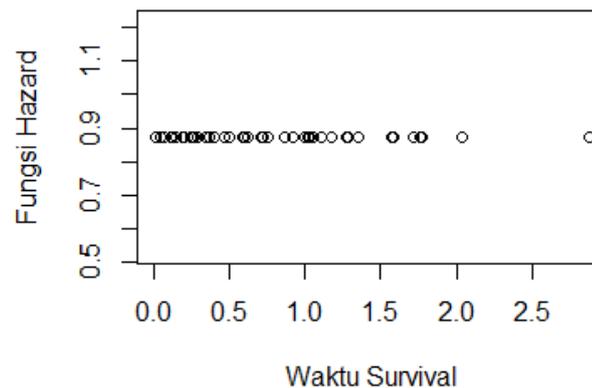


Gambar 1 Grafik Fungsi *Survival* Metode MLE

Dan fungsi *Hazard* dirumuskan dengan :

$$\hat{h}_{ML}(x_i; \hat{\theta}_{ML}) = \beta \frac{\sum_{i=1}^n \delta_i}{\sum_{i=1}^n x_i^\beta} x^{\beta-1} = 1 \times \frac{27}{288.7744^{(1)} + 1} \times x^{(1-1)}$$

Dengan mengambil beberapa nilai untuk x , $x > 0$, fungsi *hazard* ini dapat diilustrasikan dengan grafik fungsi *hazard* berikut.



Gambar 2 Grafik Fungsi *Hazard* Metode MLE

4. KESIMPULAN

Studi simulasi pada kajian ini yang mengambil dua ukuran data berbeda membuktikan bahwa penduga parameter dengan metode MLE memberikan estimasi nilai absolut bias dan MSE yg lebih kecil dibandingkan dengan metode Bayesian untuk kedua ukuran data dan untuk semua nilai θ terpilih. Distribusi prior yang digunakan pada analisis Bayes adalah prior konjugat yaitu distribusi Invers Gamma. Dapat disimpulkan dalam penelitian ini bahwa MLE merupakan penduga yang lebih baik dari metode Bayes. Kesimpulan ini belum tentu berlaku untuk prior jenis lain dan dari distribusi yang lain. Oleh karena itu untuk pembuktiannya perlu dilakukan studi simulasi lanjutan untuk kasus khusus seperti demikian.

DAFTAR PUSTAKA

- [1] C. B. Guure and N. A. Ibrahim, "Bayesian Analysis of the Survival Function and Failure Rate of Weibull Distribution with Censored Data," *Mathematical Problems in Engineering*, vol. 2012, pp. 1–18, 2012, doi: 10.1155/2012/329489.
- [2] D. Lukitasari, A. Setiawan, and L. R. Sasangko, "Bayesian Survival Analysis Untuk Mengestimasi Parameter Model Weibull-Regression Pada Kasus Ketahanan Hidup Pasien Penderita Jantung Koroner," *dC*, vol. 4, no. 1, p. 26, Feb. 2015, doi: 10.35799/dc.4.1.2015.7531.

- [3] S. Dey, "Bayesian Estimation and Prediction on Inverse Rayleigh Distribution," *Malaysian Journal of Mathematical Sciences*, vol. 6, no. 1, pp. 113–124, 2012.
- [4] T. G. Ieren and P. E. Oguntunde, "A Comparison between Maximum Likelihood and Bayesian Estimation Methods for a Shape Parameter of the Weibull-Exponential Distribution," *AJPAS*, pp. 1–12, May 2018, doi: 10.9734/ajpas/2018/v1i124504.
- [5] I. T. Hazhiah, S. Sugito, and R. Rahmawati, "Estimasi Parameter Distribusi Weibull Dua Parameter Menggunakan Metode Bayes," *Medstat*, vol. 5, no. 1, pp. 27–35, Jun. 2012, doi: 10.14710/medstat.5.1.27-35.
- [6] Ahmed, "Comparison of the Bayesian and Maximum Likelihood Estimation for Weibull Distribution," *Journal of Mathematics and Statistics*, vol. 6, no. 2, pp. 100–104, Apr. 2010, doi: 10.3844/jmssp.2010.100.104.
- [7] D. Kundu and H. Howlader, "Bayesian inference and prediction of the inverse Weibull distribution for Type-II censored data," *Computational Statistics & Data Analysis*, vol. 54, no. 6, pp. 1547–1558, Jun. 2010, doi: 10.1016/j.csda.2010.01.003.
- [8] I. Iskandar and Y. S. Gondokaryono, "Competing risk models in reliability systems, a weibull distribution model with bayesian analysis approach," *IOP Conf. Ser.: Mater. Sci. Eng.*, vol. 114, p. 012064, Feb. 2016, doi: 10.1088/1757-899X/114/1/012064.
- [9] S. A. Thamrin, "Penaksiran Parameter Distribusi Weibull dengan Metode Bayesian Survival dan Maksimum Likelihood," *Jurnal Keteknik dan Sains*, vol. 1, no. 2, pp. 22–27, 2018.
- [10] F. Yanuar, H. Yozza, and R. V. Rescha, "Comparison of Two Priors in Bayesian Estimation for Parameter of Weibull Distribution," *sci. technol. indones.*, vol. 4, no. 3, p. 82, Jul. 2019, doi: 10.26554/sti.2019.4.3.82-87.
- [11] E. Avci, "Bayesian survival analysis: comparison of survival probability of hormone receptor status for breast cancer data," *IJDATS*, vol. 9, no. 1, p. 63, 2017, doi: 10.1504/IJDATS.2017.10003994.
- [12] H. Rinne, *The Weibull distribution: a handbook*. Boca Raton: CRC Press, 2009.
- [13] H. M. Badjana, B. Renard, J. Helmschrot, K. S. Edjamé, A. Afouda, and K. Wala, "Bayesian trend analysis in annual rainfall total, duration and maximum in the Kara River basin (West Africa)," *Journal of Hydrology: Regional Studies*, vol. 13, pp. 255–273, Oct. 2017, doi: 10.1016/j.ejrh.2017.08.009.
- [14] A. A. Al-Shomrani, A. I. Shawky, O. H. Arif, and M. Aslam, "Log-logistic distribution for survival data analysis using MCMC," *SpringerPlus*, vol. 5, no. 1, p. 1774, Dec. 2016, doi: 10.1186/s40064-016-3476-7.
- [15] F. Yanuar, "The Use of Uniformative and Informative Prior Distribution in Bayesian SEM," *Global Journal of Pure and Applied Mathematics*, vol. 11, no. 5, pp. 3259–64, 2015.
- [16] W. M. Bolstad and J. M. Curran, *Introduction to Bayesian Statistics*, Third Edition. John Wiley & Sons, Inc., 2017.