# INFERENSI FUNGSI KETAHANAN DENGAN METODE KAPLAN-MEIER

Tatik Widiharih dan Nasichah Siska Andriani
Jurusan Matematika FMIPA UNDIP
Jl. Prof. H. Soedarto, S.H, Semarang 50275

**Abstract.** Let T be a nonnegatif random variable representing the life time of individuals in some population. Life time data of individuals are devided in two kinds, cencored and uncencored data. The probability of an individual surviving till time t is given by the survival function S(t)=P(T≥t). Product Limit estimator (Kaplan-Meier estimator) is a nonparametric method to find the survival function for cencored data.

**Key words:** survival function, cencored data, Kaplan–Meier estimator

## 1. PENDAHULUAN

Analisis waktu hidup adalah metode analisis waktu hidup yang bergantung dari waktu meliputi waktu hidup dan status waktu hidup individu. Data waktu hidup yang diperoleh dapat berupa data tak tersensor dan data tersensor. Data yang dimaksud adalah data hidup individu dalam grup tertentu yang diukur dari periode tertentu pula dan merupakan variabel random yang bernilai nonnegatif sehingga akan membentuk suatu distribusi yang disebut distribusi waktu hidup.

Salah satu permasalahan yang ditemui dalam analisis data ketahanan hidup adalah kemungkinan adanya beberapa individu yang tidak dapat diikuti perkembangannya sampai individu tersebut mati. Waktu tahan hidup adalah ketika individu dapat diikuti perkembangannya sampai individu tersebut mati, sedangkan untuk individu yang tidak dapat lagi diikuti perkembangannya merupakan observasi waktu tahan hidup yang tidak lengkap, biasa disebut observasi tersensor. Fungsi ketahanan suatu individu $S(t)$ didefinisikan sebagai perbandingan antara jumlah observasi yang hidup lebih dari waktu $t$ $(t \geq 0)$ dengan jumlah total observasi [2]. Fungsi ketahanan ini sangat bergantung pada tipe penyensoran yang digunakan. Tipe penyensoran yang sering dipakai adalah sensor tipe I, II dan III.

Dalam tulisan ini diuraikan mengenai inferensi fungsi ketahanan dengan menggunakan metode Kaplan–Meier. Dicari peluang ketahanan hidup individu dengan sensor tipe III dan membandingkan fungsi ketahanan dua sampel yang berbeda dengan salah satu sampel diberi perlakuan tertentu dan sampel yang lain dijadikan observasi terkontrol, lebih lanjut dicari sampel manakah yang mempunyai fungsi ketahanan yang lebih baik.

## 2. KAJIAN TEORI

Analisis ketahanan adalah metode yang sering digunakan untuk mempelajari tentang kegagalan. Untuk analisis ketahanan dimulai observasi suatu kumpulan individu pada beberapa titik waktu yang sudah diketahui secara pasti dan dapat diikuti perkembangannya untuk beberapa periode waktu dan waktu dicatat pada saat kejadian tersebut.

Uji ketahanan hidup berkaitan dengan waktu ketahanan karena dengan diketahui waktu ketahanan individu bisa diketahui berapa besar kemungkinan ketahanan hidup individu tersebut oleh suatu perlakuan tertentu. Waktu ketahanan adalah ukuran data waktu untuk kejadian pasti seperti kegagalan atau kematian. Waktu ketahanan $T$ merupakan variabel random nonnegatif yang mewakili ketahanan hidup dari individu dalam populasi yang merupakan variabel random kontinu dalam interval $[0,\infty)$ atau ketahanan hidup pada waktu t dengan $t > 0$.

## 2.1 Tipe Penyensoran

### 2.1.1 Sensor Tipe I

Ada *n* individu yang diamati dimulai pada waktu yang sama. Eksperimen akan berhenti jika telah dicapai waktu tertentu (waktu penyensoran). Sensor ini membutuhkan waktu yang lama dan biaya yang relatif besar karena peneliti harus menunggu sampai waktu penyensoran berakhir [3].

### 2.1.2 Sensor Tipe II

Ada n individu yang diamati dimulai dari waktu yang sama. Eksperimen dihentikan setelah kegagalan ke–r diperoleh, sehingga sensor tipe II dapat menghemat waktu dan biaya. Jumlah pengamatan harus diputuskan terlebih dahulu sebelum data tersebut dikumpulkan [3].

### 2.1.3 Sensor Tipe III

Individu masuk dalam pengamatan pada waktu yang berbedabeda selama periode waktu tertentu. Beberapa individu mungkin mati/gagal sebelum pengamatan berakhir dimana waktu ketahanannya secara pasti diketahui, kemungkinan yang lain adalah individu keluar sebelum pengamatan berakhir sehingga tidak berpengaruh pada pengamatan berikutnya, dan kemungkinan yang terakhir adalah individu tetap hidup sampai batas akhir waktu pengamatan. Untuk individu yang hilang, waktu ketahanan adalah dari mulai masuk pengamatan sampai dengan waktu terakhir sebelum hilang. Untuk individu yang tetap hidup, waktu ketahanan adalah dari mulai masuk pengamatan sampai dengan waktu pengamatan berakhir. Karena waktu masuk pengamatan tiap individu tidak sama maka waktu sensor yang terjadi juga berbeda–beda [3].

## 3. ESTIMATOR PRODUCT LIMIT (KAPLAN – MEIER )

Jika waktu ketahanan dapat dikelompokkan secara terurut maka waktu ketahanan *T* dapat diperlakukan sebagai variabel random yang diskrit. Misalkan T bernilai $t_1, t_2, \ldots$dengan $0 \le t_1 < t_2 < \ldots$ dan fungsi densitas peluangnya adalah $p(t_i) = P(T = t_i),\ i = 1,2,\ldots,n$ maka fungsi ketahanannya adalah

$$S(t) = P(T \ge t) = \sum_{i=t_i \ge t} p(t_i). \qquad (2.1)$$

Estimasi dengan Kaplan-Meier dari probabilita ketahanan dari beberapa waktu yang khusus merupakan hasil kali estimasi yang sama pada waktu sebelumnya dan angka ketahanan yang terobservasi dari tahun-tahun tersebut [3].

$S(t_{i+1}) = S(t_i) \,.\, S(t)$ atau

$$S(t) = \frac{S(t_{i+1})}{S(t_i)}. \qquad (2.2)$$

Fungsi hazard didefinisikan sebagai peluang suatu individu gagal dalam interval (t, ,t+Δt) dengan diketahui bahwa individu tersebut telah hidup selama waktu t.

$$h(t_i) = P\left(T = t_i \mid T \ge t_i\right)$$

$$= \frac{p(t_i)}{S(t_i)}, \qquad i = 1,2,\ldots \qquad (2.3)$$

Karena $p(t_i) = S(t_i) - S(t_{i+1})$ maka persamaan (2.3) menjadi

$$h(t_i) = \frac{p(t_i)}{S(t_i)} = \frac{S(t_i) - S(t_{i+1})}{S(t_i)}$$

$$= 1 - \frac{S(t_{i+1})}{S(t_i)}, \qquad i = 1,2,3,\ldots \quad (2.4\ )$$

Subtistusi persamaan (2.2) dan (2.3) :

$$S(t) = \frac{S(t_{i+1})}{S(t_i)} = 1 - h(t). \qquad (2.5)$$

Jika terdapat observasi yang saling bebas maka estimator nonparametrik fungsi ketahanan adalah

$$\hat{S}(t) = \prod_{i=1}^{n} \left[1 - \hat{h}(t_i)\right], \quad i = 1,2,\ldots,n, \qquad (2.6)$$

dengan $\hat{h}(t_i)$ diestimasi dengan menggunakan metode maksimum likelihood sehingga diperoleh

$$\hat{h}(t_i) = \frac{\delta_i}{R_i},\ \text{i = 1,2,…,n,}$$

dengan

$\delta_i$ adalah jumlah kematian individu pada pengamatan ke-i.,

$R_i$ adalah jumlah individu yang tetap hidup pada pengamatan ke-i..

Dengan demikian fungsi ketahanan *S(t)* dengan menggunakan estimasi Product Limit (Kapalan–Meier) adalah

$$\hat{S}(t) = \prod_{i=1}^{n}\left(1-\frac{\delta_i}{R_i}\right)$$
$$= \prod_{i=1}^{n}\left(1-\hat{q}_i\right) = \prod_{i=1}^{n}\hat{p}_i \,. \qquad (2.7)$$

Pada prakteknya estimasi Kaplan-Meier dapat dicari dengan cara membentuk tabel yang kolom-kolomnya berisi:

1. Kolom pertama ( i ), banyaknya pengamatan yang diambil sebagai sampel.
2. Kolom kedua ( $t_i$ ), semua waktu hidup baik yang termasuk observasi tersensor
3. maupun observasi tidak tersensor dengan diurutkan dari waktu yang terkecil sampai dengan waktu yang terbesar. Tanda "*" berarti termasuk observasi yang tersensor.
4. Kolom ketiga *($R_i$)*, jumlah individu yang tetap hidup pada pengamatan ke-i.
5. Kolom keempat *($\delta_i$)*, jumlah kematian individu pada pengamatan ke-i.
6. Kolom kelima ($\hat{q}_i$), estimasi peluang kematian individu pada pengamatan ke-i.
7. Kolom keenam ($\hat{p}_i$), estimasi peluang ketahanan individu pada pengamatan ke-i.
8. Kolom ketujuh ($\hat{S}(t)$), estimasi fungsi ketahanan untuk tiap-tiap individu yang masuk pada pengamatan.

Tabel 1. Tabel estimasi Kaplan-Meier

| **Amatan i** (1) | **Waktu pengamatan $t_i$** (2) | **Jumlah individu $R_i$** (3) | **Jumlah individu mati $\delta_i$** (4) | **Estimasi probabilita bersyarat kematian $\hat{q}$** (5) | **Estimasi probabilita bersyarat ketahanan $\hat{p}_i$** (6) | **Estimasi fungsi ketahanan $\hat{S}(t)$** (7) |
|---|---|---|---|---|---|---|
| 1 | $t_1$ | $R_1$ | $\delta_1$ | $q_1$ | $p_1$ | $S(t_1)$ |
| 2 | $t_2$ | $R_2$ | $\delta_2$ | $q_2$ | $p_2$ | $S(t_2)$ |
| 3 | $t_3$ | $R_3$ | $\delta_3$ | $q_3$ | $p_3$ | $S(t_3)$ |
| . | . | . | . | . | . | . |
| . | . | . | . | . | . | . |
| . | . | . | . | . | . | . |
| n | $t_n$ | $R_n$ | $\delta_n$ | $q_n$ | $p_n$ | $S(t_n)$ |

Sebagai contoh, berikut ini diberikan 21 pasien penderita leukemia yang diberi perlakuan secara acak berupa obat 6–MP dan 21 pasien yang lain tidak diberikan perlakuan (kontrol) [1].

| | |
|---|---|
| Sampel 1 ( obat 6–MP ) | 6*, 6, 6, 6, 7, 9*, 10*, 10, 11*, 13, 16, 17*, 19*, 20*, 22, 23, 25*, 32*, 32*, 34*, 35* |
| Sampel 2 ( kontrol ) | 1, 1, 2, 2, 3, 4, 4, 5, 5, 8, 8, 8, 8, 11, 11, 12, 12, 15, 17, 22, 23 |

"*" observasi tersensor.

- Untuk sampel kesatu

Tabel 2. Tabel Penderita Leukimia yang Diberi Perlakuan

| **Waktu pengamatan $t_i$** | **Jumlah individu yang diamati beresiko $R_i$** | **Jumlah individu yang mati $\delta_i$** | **Jumlah individu yang tersensor $C_i$** | **$q_i$** | **$p_i$** | **S(t)** |
|---|---|---|---|---|---|---|
| 6 | 21 | 3 | 1 | 0.1429 | 0.8571 | 0.8571 |
| 7 | 17 | 1 | 0 | 0.0588 | 0.9412 | 0.8067 |
| 9 | 16 | 0 | 1 | 0 | 1 | 0.8067 |
| 10 | 15 | 1 | 1 | 0.0667 | 0.9333 | 0.7529 |
| 11 | 13 | 0 | 1 | 0 | 1 | 0.7529 |
| 13 | 12 | 1 | 0 | 0.0833 | 0.9167 | 0.6902 |
| 16 | 11 | 1 | 0 | 0.0910 | 0.9090 | 0.6274 |
| 17 | 10 | 0 | 1 | 0 | 1 | 0.6274 |
| 19 | 9 | 0 | 1 | 0 | 1 | 0.6274 |
| 20 | 8 | 0 | 1 | 0 | 1 | 0.6274 |
| 22 | 7 | 1 | 0 | 0.1429 | 0.8571 | 0.5377 |
| 23 | 6 | 1 | 0 | 0.1667 | 0.8333 | 0.4481 |
| 25 | 5 | 0 | 1 | 0 | 1 | 0.4481 |
| 32 | 4 | 0 | 2 | 0 | 1 | 0.4481 |
| 34 | 2 | 0 | 1 | 0 | 1 | 0.4481 |
| 35 | 1 | 0 | 1 | 0 | 1 | 0.4481 |

- Untuk sampel kedua

Tabel 3. Tabel Penderita Leukimia yang Tidak Diberi Perlakuan

| **Waktu pengamatan $t_i$** | **Jumlah individu yang diamati beresiko $R_i$** | **Jumlah individu yang mati $\delta_i$** | **Jumlah individu yang tersensor $C_i$** | **$q_i$** | **$p_i$** | **S(t)** |
|---|---|---|---|---|---|---|
| 1 | 21 | 2 | 0 | 0.0952 | 0.9048 | 0.9048 |
| 2 | 19 | 2 | 0 | 0.1053 | 0.8947 | 0.8095 |
| 3 | 17 | 1 | 0 | 0.0588 | 0.9412 | 0.7619 |
| 4 | 16 | 2 | 0 | 0.1250 | 0.8750 | 0.6667 |
| 5 | 14 | 2 | 0 | 0.1429 | 0.8571 | 0.5714 |
| 8 | 12 | 4 | 0 | 0.3333 | 0.6667 | 0.3810 |
| 11 | 8 | 2 | 0 | 0.2500 | 0.7500 | 0.2858 |
| 12 | 6 | 2 | 0 | 0.3333 | 0.6667 | 0.1905 |
| 15 | 4 | 1 | 0 | 0.2500 | 0.7500 | 0.1429 |
| 17 | 3 | 1 | 0 | 0.3333 | 0.6667 | 0.0953 |
| 22 | 2 | 1 | 0 | 0.5000 | 0.5000 | 0.0477 |
| 23 | 1 | 1 | 0 | 1 | 0 | 0 |

## 4. PERBANDINGAN DUA FUNGSI KETAHANAN

Jika fungsi ketahanan hidup suatu individu diestimasi menggunakan estimasi Product Limit (Kaplan–Meier), maka untuk membandingkan dua fungsi ketahanan dapat menggunakan uji Peto dan Peto sebagai generalisasi uji Wilcoxon.. Uji ini dianggap sebagai skor untuk tiap observasinya seperti biasa pada uji logrank. Pemberian skor untuk tiap–tiap observasi adalah sebagai berikut.

$$U_i = \begin{cases} \hat{S}(t_i) + \hat{S}(t_{i-1}) - 1, & \text{untuk observasi yang tidak tersensor} \\ \hat{S}(t_i) - 1 & \text{,untuk obsservasi tersensor} \end{cases} \quad (2.8)$$

Langkah–langkah untuk membandingkan dua fungsi ketahanan adalah sebagai berikut.

1. Tiap observasi diberi skor $U_i$ dengan diberi tanda grup 1 dan grup
2. Dihitung nilai dari S untuk masing–masing grup, dimana
$$S = \sum_{i=1}^{n_1} U_i \qquad dan \qquad S = \sum_{j=1}^{n_2} U_j \ .$$
3. Dicari nilai varian dari *S* untuk semua grup,
$$Var\ (S) = \frac{n_1 \cdot n_2 \sum_{i=1}^{n_1+n_2} U_i^2}{(n_1+n_2)(n_1+n_2-1)} .$$
4. Melakukan pengujian hipotesis yang dipakai,
$H_0$ : $S_1(t)=S_2(t)$ (Perlakuan 1 dan 2 sama–sama efektif )
$H_1$ : $S_1(t)>S_2(t)$ (Perlakuan 1 lebih efektif daripada perlakuan 2 )
atau
$H_2$ : $S_1(t)<S_2(t)$ (Perlakuan 2 lebih efektif daripada perlakuan 1)
atau
$H_3$ : $S_1(t) \neq S_2(t)$ (Perlakuan 1 dan 2 efektifnya berbeda).
5. Menghitung statistik ujinya untuk masing–masing grup,
$$Z = \frac{S}{\sqrt{Var\ (S)}}$$
6. Dengan α sebagai tingkat signifikansi maka dicari daerah penolakan $H_0$.
   - Jika S ada di grup 1 maka daerah kritis penolakan $H_0$ adalah $Z < -Z_\alpha$.
   - Jika S ada di grup 2 maka daerah kritis penolakan $H_0$ adalah $Z > Z_\alpha$.

Untuk contoh kasus diatas diperoleh Tabel 4 pada Lampiran.

Grup 1 : *S* = – 6.0006 dan
Grup 2 : *S* = 6.0096 dengan
*Var (S))* = 3.189738725
Untuk sampel kesatu uji hipotesa yang dipakai adalah :
$H_o$ : $S_1(t)=S_2(t)$ (Perlakuan 1 dan 2 sama – sama efektif).
$H_1$ : $S_1(t)>S_2(t)$ (Perlakuan 1 lebih efektif daripada perlakuan 2).
Statistik uji:

$$Z = \frac{S}{\sqrt{Var\ (S)}} = \frac{-0.6006}{\sqrt{3.1897}}$$
$$= \frac{-6.0006}{1.7860} = -3.3598$$

Daerah penolakan $H_0$ untuk sampel kesatu adalah jika $Z_{hitung} < -Z_{tabel}$ dengan penggunaan α sebesar 5 %. Karena $Z_{hitung} = -3.3598 < -Z_{tabel} = -1.645$ maka $H_0$ ditolak yang berarti bahwa perlakuan pada sampel satu yang diberi obat 6–MP lebih efektif daripada tidak diberi perlakuan seperti pada sampel kedua.

## 5. KESIMPULAN

Estimator Kaplan–Meier digunakan untuk mengestimasi fungsi ketahanan hidup dari data yang tersensor dan data yang tidak tersensor untuk masing–masing observasi dengan panjang interval waktu yang bervariasi. Untuk membandingkan dua fungsi ketahanan dapat digunakan uji Peto dan Peto sebagai generalisasi uji Wilcoxon. berdasarkan hasil analisis pasien yang diberi perlakuan obat 6–MP memiliki fungsi ketahanan yang lebih baik dibandingkn dengan pasien yang tidak diberi perlakuan.

## 6. DAFTAR PUSTAKA

[1]. Cox, D. R and Oakes, D. (1983), *Analysis of Survival Data*, John Wiley and Sons, Inc.

[2]. Lawless, J. F. (1982), *Statistical Models And Methods For Lifetime Data,* John Wiley and Sons, Inc.

[3]. Lee, E. T. (1992), *Statistical Methods for Survival Data Analysis*, John Wiley and Sons, Inc.

[4]. Miller, R. G. (1981), *Survival Analysis,* John Wiley and Sons, Inc.

## Lampiran

Tabel 4. Tabel Penderita Leukimia yang Diberi Perlakuan

| No. Obser vasi | Waktu Pengamatan $t_i$ | Grup | $\hat{S}(t)$ | $U_i$ |
|---|---|---|---|---|
| 1. | 1 | 2 | 0.9048 | 1 + 0.9048 – 1 = 0.9048 |
| 2. | 1 | 2 | 0.9048 | 0.9048 + 0.9048 – 1 = 0.8096 |
| 3. | 2 | 2 | 0.8095 | 0.9048 + 0.8095 – 1 = 0.7143 |
| 4. | 2 | 2 | 0.8095 | 0.8095 + 0.8095 – 1 = 0.6190 |
| 5. | 3 | 2 | 0.7619 | 0.6667 + 0.7619 – 1 =0.5714 |
| 6. | 4 | 2 | 0.6667 | 0.6667 + 0.7619 – 1 = 0.4286 |
| 7. | 4 | 2 | 0.6667 | 0.6667 + 0.6667 – 1 =0.3334 |
| 8. | 5 | 2 | 0.5714 | 0.5714 + 0.6667 – 1 = 0.2381 |
| 9. | 5 | 2 | 0.5714 | 0.5714 + 0.5714 – 1 = 0.1428 |
| 10. | 6* | 1 | – | 0.5714 – 1 = – 0.4286 |
| 11. | 6 | 1 | 0.8571 | 0.5714 + 0.8571 – 1 = 0.4285 |
| 12. | 6 | 1 | 0.8571 | 0.8571 + 0.8571 – 1 = 0.7142 |
| 13. | 6 | 1 | 0.8571 | 0.8571 + 0.8571 – 1 = 0.7412 |
| 14. | 7 | 1 | 0.8067 | 0.8067 + 0.8571 – 1 = 0.6638 |
| 15. | 8 | 2 | 0.3810 | 0.3810 + 0.8067 – 1 = 0.1877 |
| 16. | 8 | 2 | 0.3810 | 0.3810 + 0.3810 – 1 = – 0.2380 |
| 17. | 8 | 2 | 0.3810 | 0.3810 + 0.3810 – 1 = – 0.2380 |
| 18. | 8 | 2 | 0.3810 | 0.3810 + 0.3810 – 1 = – 0.2380 |
| 19. | 9* | 1 | – | 0.3810 – 1 = – 0.6190 |
| 20 | 10* | 1 | – | 0.3810 – 1 = – 0.6190 |
| 21. | 10 | 1 | 0.7529 | 0.3810 + 0.7529 – 1 = 0.1339 |
| 22. | 11* | 1 | – | 0.7529 – 1 = – 0.2471 |
| 23. | 11 | 2 | 0.2858 | 0.7529 + 0.2858 – 1 = 0.0387 |
| 24. | 11 | 2 | 0.2858 | 0.2858 + 0.2858 – 1 = – 0.4284 |
| 25. | 12 | 2 | 0.1905 | 0.2858 + 0.1905 – 1 = – 0.5237 |
| 26. | 12 | 2 | 0.1905 | 0.1905 + 0.1905 – 1 = – 0.6190 |
| 27. | 13 | 1 | 0.6902 | 0.6902 + 0.1905 – 1 = – 0.1193 |
| 28. | 15 | 2 | 0.1429 | 0.1429 + 0.6902 – 1 = 0.1669 |
| 29. | 16 | 1 | 0.6274 | 0.6274 + 0.4129 – 1 = – 0.2297 |
| 30. | 17 | 2 | 0.0953 | 0.0953 + 0.6274 – 1 = – 0.2773 |
| 31. | 17* | 1 | – | 0.0953 – 1 = – 0.9047 |
| 32. | 19* | 1 | – | 0.0953 – 1 = – 0.9047 |
| 33. | 20* | 1 | – | 0.0953 – 1 = – 0.9047 |
| 34. | 22 | 1 | 0.5377 | 0.0953 + 0.5377 – 1 = – 0.3670 |
| 35. | 22 | 2 | 0.0477 | 0.0477 + 0.5377 – 1 = – 0.4146 |
| 36. | 23 | 2 | 0 | 0.0477 + 0 – 1 = – 0.9523 |
| 37. | 23 | 1 | 0.4481 | 0.4481 + 0 – 1 = – 0.9523 |
| 38. | 25* | 1 | – | 0.44481 – 1 = – 0.5519 |
| 39. | 32* | 1 | – | 0.44481 – 1 = – 0.5519 |
| 40. | 32* | 1 | – | 0.44481 – 1 = – 0.5519 |
| 41. | 34* | 1 | – | 0.44481 – 1 = – 0.5519 |
| 42. | 35* | 1 | – | 0.44481 – 1 = – 0.5519 |