BAB II

MATERI PENUNJANG

2.1. Pengertian

Definisi 2.1.1:

Jika ξ sebuah percobaan yang memiliki ruang sampel ς, dan T sebuah fungsi yang dinotasikan sebuah bilangan riil $T(\zeta)$ untuk setiap kejadian $\zeta \in \varsigma$ maka $T(\zeta)$ disebut variabel random, dengan memenuhi persyaratan sebagai berikut :

1. Himpunan $\{T: T \leq t\}$ adalah peristiwa untuk setiap bilangan riil t.
2. Probabilitas peristiwa $\{T = \infty\}$ dan $\{T = -\infty\}$ sama dengan nol.

Contoh :

ξ : Sebuah tabung katoda diproduksi, dan diuji lamanya tabung tersebut sampai rusak. Waktu yang terbuang (dalam jam) atas kerusakan tersebut dicatat. Jika T adalah waktu rusak, maka $T(t) = t$.

ζ : Waktu kegagalan adalah lebih besar daripada 1000 jam.

ς : $\{t : t \geq 0\}$

Definisi 2.1.2 :

Statistik berurut variabel random T_i adalah n variabel random Y_k yang didefinisikan sebagai berikut : Untuk kejadian tertentu ζ suatu percobaan, variabel random T_i menjalani harga-harga $T_i\ (\zeta)$, $i = 1,2, ., k, k+1, .., n$

$T_1(\zeta) \leq T_2(\zeta) \leq \ldots \leq T_k(\zeta) \leq \ldots \leq T_n(\zeta)$.

Variabel random Y_k didefinisikan sebagai berikut :

$Y_1(\zeta) = T_1(\zeta) \leq Y_2(\zeta) = T2(\zeta) \leq \ldots \leq Y_k = T_k(\zeta) \leq \ldots \leq Y_n(\zeta) = T_n(\zeta)$.

Contoh :

ξ : Lima buah tabung katoda ditempatkan pada percobaan lamanya pemakaian sampai mengalami kerusakan sebelum 10 jam.

ζ : Tabung pertama mengalami kerusakan setelah pemakaian selama 4 jam, kemudian berturut-turut 3 jam; 6 jam; 5 jam dan 7 jam.

ς : $\{t : t \geq 0\}$

setelah diurutkan diperoleh barisan :

$T_1(\zeta)=3 \leq T_2(\zeta)=4 \leq T_3(\zeta)=5 \leq T_4(\zeta)=6 \leq T_5(\zeta)=7$

Variabel random Y_k adalah sebagai berikut :

$Y_1(\zeta)=T_1(\zeta)=3 \leq Y_2(\zeta)=T_2(\zeta)=4 \leq Y_3(\zeta)=T_3(\zeta)=5 \leq Y_4(\zeta)=T_4(\zeta)=6 \leq Y_5(\zeta)=T_5(\zeta)=7$

Definisi 2.1.3 :

Waktu ketahanan yaitu interval waktu yang diamati dari suatu individu saat pertama kali masuk ke dalam pengamatan hingga keluar dari pengamatan.

Definisi 2.1.4 :

Status ketahanan ialah status individu tepat pada saat ia meninggalkan pengamatan, yaitu gagal (mati), hilang (tidak terdeteksi) atau with drawn-alive.

Definisi 2.1.5 :

Individu withdrawn-alive adalah individu yang bertahan hidup meskipun telah keluar dari pengamatan.

2.2. Prinsip Dasar Metode Maksimum Likelihood

Metode maksimum likelihood merupakan salah satu cara untuk memperoleh sebuah taksiran tunggal maupun tidak tunggal pada suatu fungsi distribusi. Metode ini dapat juga digunakan untuk menentukan taksiran dari data ketahanan yang diasumsikan mempunyai pola suatu distribusi tertentu.

2.2.1. Untuk data lengkap atau data tidak tersensor

Definisi 2.2.1 :

Data lengkap atau data tidak tersensor adalah data ketahanan yang diperoleh dari suatu pengamatan dimana status ketahanan seluruh individu adalah gagal.

Diasumsikan T_1 , T_2 , ..., T_n adalah sampel random dari variabel random **T** yang mempunyai probabilitas p(t) atau fungsi kerapatan kegagalan f(t) dan tergantung pada parameter yang tidak diketahui misalnya θ , dengan $\theta \in \Omega$, dan Ω adalah ruang parameter.

Maka fungsi Likelihood sampel random tersebut dinyatakan dengan :

$$
\begin{aligned}
L(T/\theta) &= p(t_1).\,p(t_2)\ldots p(t_n) && \text{, T diskrit.}\\
&= f(t_1).\,f(t_2)\ldots f(t_n) && \text{, T kontinu.}
\end{aligned}
$$

Jika T adalah variabel random diskrit, dan $T_1,T_2,\ldots,T_n$ adalah sampel random dari T ; $t_1,t_2,\ldots,t_n$ adalah nilai-nilai pengamatan pada sampel random. Maka probabilitas nilai pengamatan yang terjadi dalam sampel adalah :

$$
\begin{aligned}
P(T_1=t_1, T_2=t_2,\ldots,T_n=t_n) &= p(T_1=t_1).\,p(T_2=t_2)\ldots p(T_n=t_n)\\
&= p(t_1).\,p(t_2)\ldots p(t_n)\\
&= L(T/\theta)
\end{aligned}
$$

Fungsi likelihood yaitu $L(T/\theta)$ menyatakan probabilitas nilai pengamatan yang dihasilkan sebagai fungsi θ.

Jika T variabel random kontinu dan sampel randomnya memuat nilai-nilai yang terletak dalam interval.

$t_1 \le T_1 \le t_1+\Delta t,\ \ t_2 \le T_2 \le t_2+\Delta t, \ldots, t_n \le T_n \le t_n+\Delta t$

probabilitas nilai pengamatan sampelnya adalah :

$$\begin{aligned} P\,(\, t_1 \le T_1 \le t_1+\Delta t,\ \ t_2 \le T_2 \le t_2+\Delta t, \ldots, t_n \le T_n \le t_n+\Delta t\,) \\ &= f(t_1)\ \Delta t .\ f(t_2)\ \Delta t \ldots f(t_n)\ \Delta t \\ &= f(t_1).\ f(t_2) \ldots f(t_n)\ .\ (\Delta t)^n \\ &= L\,(\,T/\theta\,).\ (\Delta t)^n \end{aligned}$$

Fungsi likelihood untuk sampel random kontinu secara umum yaitu :

$$L\,(\,T/\,\theta) = \prod_{i=1}^{n} f\,(t_i\,;\theta) = f(t_1\,;\theta)\ f(t_2\,;\theta)\ \ldots\ f(t_n\,;\theta) \qquad (2.2.1)$$

Prinsip dasar metode maksimum likelihood adalah mencari suatu fungsi dari $T_1,T_2,\ldots,T_n$ sedemikian sehingga jika θ diganti dengan $\hat{\theta}\,(T_1,T_2,\ldots,T_n)$ fungsi likelihood L adalah suatu maksimum, yaitu : $L\,[\,\hat{\theta}\ (\,T_1,T_2,\ldots,T_n);\ T_1,T_2,\ldots,T_n)] \ \ge\ L\ [\,\theta;\,T_1,T_2,\ldots,T_n]$ untuk setiap $\theta \in \Omega$. Jika statistik $\hat{\theta}$ $(T_1,T_2,\ldots,T_n)$ bernilai, maka $\hat{\theta}\ (T_1,T_2,\ldots,T_n)$ merupakan taksiran maksimum likelihood dari θ. Sedangkan taksiran maksimum likelihood dari θ yaitu $\hat{\theta}$ dapat diperoleh dengan mendifferensialkan fungsi likelihood $L\,(T/\theta)$ sama dengan nol, yaitu :

$$\frac{d\,L(T/\theta)}{d\theta} = 0 \qquad (2.2.2)$$

Estimator maksimum likelihood $\hat{\theta}$ adalah nilai θ yang memaksimalkan fungsi likelihood L (T/θ).

Jika terdapat beberapa parameter yang tidak diketahui dalam fungsi distribusi misalnya $\theta_1, \theta_2, \ldots, \theta_k$ maka fungsi likelihoodnya :

$$L(T/\theta_1,\theta_2,\ldots,\theta_k) = \prod_{i=1}^{n} f(t_i\,;\theta_1,\theta_2,\ldots\theta_k) \qquad , \qquad n>k$$

Estimator maksimum Likelihood untuk $\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_k$ dapat ditentukan dengan mendifferensialkan parsial pertama fungsi likelihoodnya sama dengan nol, yaitu :

$$\frac{\partial L(T/\theta_1,\theta_2,\ldots,\theta_k)}{\partial\theta_i} = 0 \quad , i = 1,2,\ldots,k \qquad (2.2.3)$$

Harga-harga $\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_k$ merupakan calon estimator maksimum likelihood yang mungkin, untuk memperoleh estimator $\theta_1, \theta_2, \ldots, \theta_k$ yang sungguh-sungguh memaksimalkan fungsi likelihood L(T / $\theta_1,\theta_2,\ldots,\theta_k$) harus memenuhi :

$$\left.\frac{\partial^2 L(T/\theta_1,\theta_2,\ldots,\theta_k)}{\partial\theta_i^2}\right|_{\theta i=\hat{\theta} i} < 0 \quad , i = 1,2,\ldots,k$$

Karena fungsi logaritma monoton naik pada [0,∞] , yang berarti bahwa L (T/θ) dan ln L (T/θ) akan mempunyai ekstrem yang sama , sehingga akan lebih mudah menggunakan differensiasi dari logaritma natural yaitu :

$$\frac{\partial \ln L(T/\theta_1,\theta_2,\ldots,\theta_k)}{\partial\theta_i} = 0 \quad , i = 1,2,\ldots,k$$

yang merupakan persamaan maksimum likelihood.

Contoh :

Pandang suatu variabel random berdistribusi eksponensial dengan parameter λ. Jika diberikan sampel random nilai data pengamatan ketahanan dengan ukuran n dari suatu populasi. Misalkan nilai-nilai data yang diperoleh n=5 adalah 7; 19; 13; 12 dan 21(dalam bulan). Maka fungsi kerapatan kegagalan sampel adalah :

$$f(t) = \lambda e^{-\lambda t}$$

Sehingga fungsi Likelihoodnya adalah :

$$L(T/\lambda) = \prod_{i=1}^{n} (\lambda e^{-\lambda t_i}) = \lambda^n e^{-\lambda \Sigma t_i}$$

$$\ln L(T/\lambda) = n \ln \lambda - \lambda \sum_{i=1}^{5} t_i$$

$$\frac{d \ln L(T/\lambda)}{d\lambda} = \frac{n}{\lambda} - \sum_{i=1}^{5} t_i = 0$$

$$\hat{\lambda} = \left[\frac{n}{\sum_{i=1}^{5} t_i}\right] = \left[\frac{5}{7+19+13+12+21}\right] = 0{,}069$$

Karena $\frac{\partial^2 \ln L(T/\lambda)}{\partial\lambda^2} = \frac{-5}{\lambda^2} < 0$

maka $\hat{\lambda}$ adalah estimator maksimum nilai λ.

2.2.2. Untuk data tidak lengkap atau data tersensor

Definisi 2.2.2 :

Data tidak lengkap atau data tersensor ialah data ketahanan yang diperoleh dari suatu pengamatan dimana tidak semua status ketahanan individu adalah gagal, tetapi ada beberapa individu yang hilang (with drawn-alive).

Untuk data yang mengalami penyensoran, fungsi kerapatan probabilitasnya adalah gabungan dari sampel random T_1 , T_2 , ..., T_d dengan k parameter dimana $\theta' = (\theta_1, \theta_2, \ldots \theta_k)$ maka fungsi likelihoodnya :

$$L(T/\ \theta') = \left[\frac{n!}{(n-d)!} \prod_{i=1}^{d} f(t_i\,;\theta')\,[\,(S(t_d\,;\theta')\,]^{\,n-d} \right] \qquad (2.2.4)$$

keterangan :

n = total data pengamatan

n-d = jumlah data pengamatan yang disensor

t_d = data pengamatan ke –d setelah diurutkan

Taksiran maksimum likelihood $\hat{\theta}_1$, $\hat{\theta}_2$,..., $\hat{\theta}_n$ diperoleh melalui persamaan :

$$\frac{\partial \ln L\ (T/\ \theta')}{\partial \theta_i} = 0 \quad , \quad i = 1,2,\ldots,k \qquad (2.2.5)$$

Selanjutnya taksiran-taksiran ($\hat{\theta}_1$, $\hat{\theta}_2$, ..., $\hat{\theta}_n$) berdistribusi normal secara asimptotik dan mempunyai matrik varian-kovarian berukuran m x m (berordo sama).

$$\hat{V}_\theta = \begin{bmatrix} -E\left(\frac{\partial^2 \ln L}{\partial \theta_1^{\,2}}\right) & \cdots & -E\left(\frac{\partial^2 \ln L}{\partial \theta_1\, \partial \theta_n}\right) \\ \vdots & \cdots & \vdots \\ -E\left(\frac{\partial^2 \ln L}{\partial \theta_1\, \partial \theta_n}\right) & \cdots & -E\left(\frac{\partial^2 \ln L}{\partial \theta_n^{\,2}}\right) \end{bmatrix}^{-1} \qquad (2.2.6)$$

$$\begin{pmatrix} \hat{\theta}_1 \\ \vdots \\ \hat{\theta}_n \end{pmatrix} \sim N\left(\begin{pmatrix} \theta_1 \\ \vdots \\ \theta_n \end{pmatrix} ;\ \hat{V}_\theta \right) \qquad (2.2.7)$$

2.3. Fungsi-fungsi Ketahanan

Fungsi-fungsi ketahanan merupakan suatu fungsi yang menggunakan variabel random waktu ketahanan. Waktu ketahanan (*survival time*) adalah data yang mengukur waktu sampai rusaknya suatu barang produksi, matinya suatu makhluk hidup, kambuhnya suatu penyakit, atau sampai terjangkitnya suatu penyakit. Secara umum, waktu ketahanan merupakan data yang mengukur waktu sampai terjadinya proses kegagalan. Waktu ketahanan merupakan variabel random yang biasanya dinotasikan dengan huruf "T", maka akan membentuk suatu distribusi.

Distribusi dari waktu ketahanan dapat disajikan oleh tiga fungsi berikut :

1. Fungsi Ketahanan Hidup(*the survivorship function*) **S(t)**
2. Fungsi Kerapatan Kegagalan atau Kematian (*the death density function*) **f(t)**
3. Fungsi Hazard (*the hazard function*) **h(t)**

Ketiga fungsi ini secara matematika adalah berhubungan, artinya jika salah satu dari ketiga fungsi tersebut diketahui, maka fungsi-fungsi yang lainnya dapat ditentukan dari fungsi tersebut.

2.3.1. Fungsi Ketahanan (the survivorship function)

Fungsi ketahanan hidup adalah probabilitas individu bertahan hidup lebih dari waktu t dengan $t > 0$. Misalkan T variabel random yang menyatakan waktu ketahanan, dimana $T \geq 0$, maka S(t) adalah probabilitas bahwa T adalah suatu waktu lebih dari waktu tetap t. Fungsi distribusi kumulatif pada waktu t untuk suatu individu dinyatakan oleh F (t). Fungsi ini merupakan distribusi waktu kegagalan yang secara statistik didefinisikan sebagai berikut :

$$F(t) = P(T \leq t) = \int_0^t f(r)dr \qquad (2.3.1)$$

Fungsi ini diinterpretasikan sebagai probabilitas bahwa suatu individu mengalami kegagalan sebelum waktu t. Fungsi ketahanan yang dinyatakan oleh S(t) didefinisikan sebagai berikut :

$$S(t) = P(T > t) = 1 - F(t) \qquad (2.3.2)$$

contoh dalam masalah kematian artinya :

S (t) = P (seseorang hidup lebih lama dari waktu t)

Kemudian S(t) ditaksir sebagai berikut :

$$\hat{S}(t) = \left[\frac{\text{jumlah pasien yang hidup lebih dari waktu t}}{\text{total jumlah pasien}}\right] \qquad (2.3.3)$$

S (t) merupakan suatu fungsi tidak naik dari waktu t dengan sifat-sifat sebagai berikut :

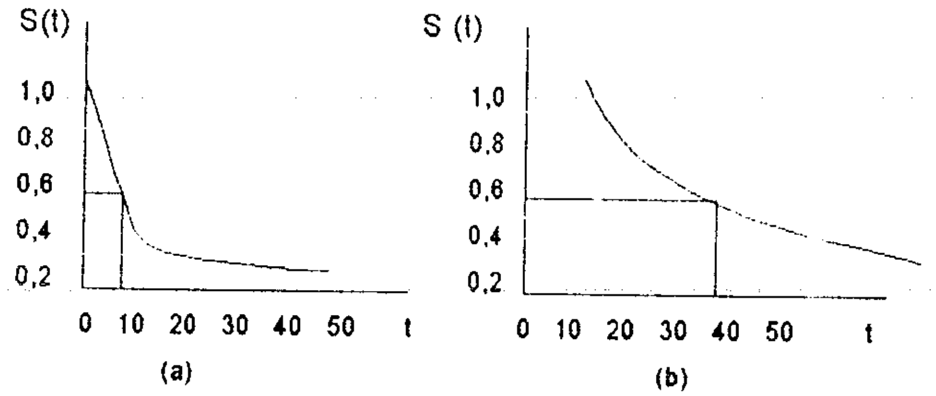
.) S (t) = 1 untuk t = 0

artinya probabilitas suatu individu bertahan lebih lama dari waktu 0 adalah satu.

.) S (t) = 0 untuk t = ∞

artinya probabilitas suatu individu bertahan pada waktu yang tidak terhingga adalah nol.

S (t) disebut juga rata-rata ketahanan kumulatif. Untuk memperjelas pengertian ketahanan, diberikan sebuah grafik yang menggambarkan S (t). Grafik S (t) ini disebut *kurva ketahanan*.



Gambar 2.1. Contoh Kurva Survival

Keterangan :

Gambar 2.1 (a) menunjukkan rata-rata ketahanan rendah atau waktu ketahanan yang terhenti sementara, sedangkan gambar 2.1 (b) menunjukkan rata-rata ketahanan tinggi atau waktu ketahanan yang lebih lama.

2.3.2. Fungsi Kerapatan Kegagalan atau Kematian (the death density function)

Fungsi kerapatan kegagalan atau kematian diinterpretasikan sebagai probabilitas kegagalan suatu individu pada suatu interval yang kecil $(t, t+\Delta t)$ persatuan waktu. Fungsi kerapatan probabilitas T untuk suatu individu yang dinyatakan oleh f(t) didefinisikan sebagai berikut :

$$f(t) = \lim_{\Delta t \to 0} \left[\frac{P\,(t \leq T < t + \Delta t)}{\Delta t} \right] \tag{2.3.4}$$

contoh dalam masalah kematian berarti :

$$f(t) = \lim_{\Delta t \to 0} \left[\frac{P\{\text{seseorang meninggal dalam interval } (t, t+\Delta t)\}}{\Delta t} \right]$$

Taksiran untuk f(t) adalah :

$$\hat{f}(t) = \left[\frac{\text{jumlah pasien meninggal dalam interval dimulai } t}{(\text{total jumlah pasien})\ (\text{panjang interval})} \right] \tag{2.3.5}$$

Fungsi kerapatan kegagalan disebut juga fungsi kerapatan probabilitas atau fungsi frekuensi variabel random waktu ketahanan, f(t) mempunyai dua sifat yaitu:

1). f(t) adalah fungsi yang tidak negatif, karena waktu ketahanan hanya mengukur nilai-nilai positif t.

$$f(t) > 0 \text{ untuk semua } t \geq 0$$

$$= 0 \text{ untuk } t < 0$$

2). $\int_0^\infty f(t)\, dt = 1$

Jika T variabel random kontinu dengan skala interval, maka probabilitas bahwa suatu individu mengalami proses kegagalan pada interval waktu [a,b] ditentukan sebagai berikut :

$$P(a < T < b) = \int_a^b f(t)\, dt \qquad (2.3.6)$$

Sedang fungsi distribusi kumulatifnya adalah :

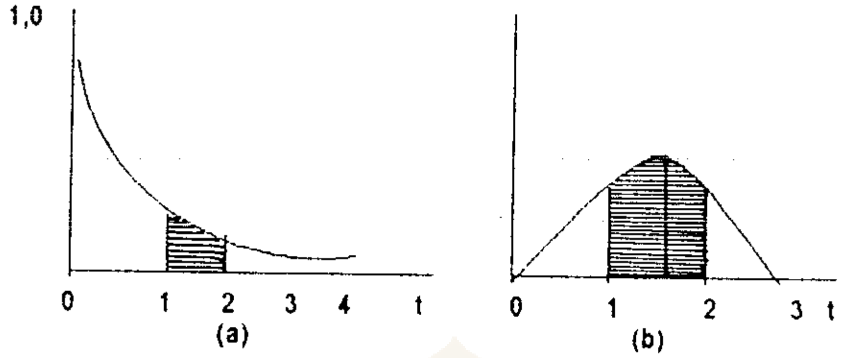
$$F(t) = P(T \leq t) = \int_0^t f(r)\, dr$$

Fungsi ini menyatakan probabilitas suatu individu gagal pada interval waktu [0,t].

Berdasarkan pengertian fungsi kerapatan kegagalan, f(t) dapat dipandang sebagai turunan dari F(t) terhadap t. Dengan demikian f (t) dapat dinyatakan sebagai berikut :

$$f(t) = \left[\frac{d\,F(t)}{dt}\right] = \left[\frac{d\,[1-S(t)]}{dt}\right] = \left[\frac{-\,d\,S(t)}{dt}\right] \qquad (2.3.7)$$

Untuk memperjelas pengertian kerapatan kegagalan, diberikan sebuah grafik yang menggambarkan f(t).



Gambar 2.2. Contoh pola-pola fungsi kerapatan

Keterangan :

Gambar 2.2 (a) adalah gambar pola rata-rata kegagalan atau kematian yang tinggi pada awal masa pengamatan dan kemudian menurun dengan semakin panjangnya waktu, gambar 2.2 (b) menggambarkan puncak frekuensi kegagalan atau kematian yang berada pada unit 1,7.

2.3.3. Fungsi Hazard (the hazard function)

Fungsi ini diinterpretasikan sebagai probabilitas bahwa suatu individu gagal didalam interval (t, t + Δt), diketahui bahwa individu tersebut telah hidup selama waktu t. Fungsi hazard yang dinyatakan oleh h(t) didefinisikan sebagai berikut :

$$h(t) = \lim_{\Delta t \to 0} \left[\frac{P\ (t \leq T < t + \Delta t \ /\ T \geq t)}{\Delta t} \right] \quad (2.3.8)$$

Contoh dalam masalah kematian berarti :

$$h\ (t) = \lim_{\Delta t \to 0} \left[\frac{P\ (\text{seseorang berumur } t \text{ meninggal pada interval } [t,\ t+\Delta t])}{\Delta t} \right]$$

Taksiran untuk h(t) adalah :

$$\hat{h}\ (t) = \left[\frac{\text{jumlah pasien meninggal tiap unit waktu dalam interval}}{\text{jumlah pasien hidup pada interval}} \right] \quad (2.3.9)$$

Fungsi hazard dapat juga diinterpretasikan sebagai resiko kegagalan pada waktu t dari individu. Fungsi hazard dikenal juga sebagai tingkat kegagalan seketika (*instantaneus failure rate*), tekanan mortalitas (*force of mortality*), tingkat kegagalan bersyarat (*conditional failure rate*), atau tingkat kegagalan umur tertentu.

Fungsi hazard dapat juga didefinisikan dalam bentuk-bentuk fungsi distribusi kumulatif F(t) dan fungsi kerapatan kegagalan f(t), sebagai berikut :

$$h(t) = \left[\frac{f(t)}{1-F(t)}\right] = \left[\frac{f(t)}{S(t)}\right] \qquad (2.3.10)$$

Dengan mensubstitusikan persamaan (2.3.7) pada persamaan (2.3.10), maka h(t) dapat dinyatakan dalam S(t), yaitu :

$$h(t) = \left[\frac{-1}{S(t)}\right] \left[\frac{d\,S(t)}{dt}\right] = \left[\frac{-\,d\ln S(t)}{dt}\right] \qquad (2.3.11)$$

Dengan mengintegralkan persamaan (2.3.11) dari 0 sampai t, dan menggunakan salah satu sifat S(t), bahwa S(0) = 1, maka S(t) dapat dinyatakan dalam h(t) sebagai berikut :

$$\int_0^t h(r)\,dr = -\left[\ln S(r)\right]_0^t$$

$$-\int_0^t h(r)\,dr = \ln S(t) - \ln S(0)$$

$$-\int_0^t h(r)\,dr = \ln S(t) - \ln 1$$

$$S(t) = \exp\left[-\int_0^t h(r)\,dr\right] \qquad (2.3.12)$$

dari persamaan (2.3.10) dan (2.3.12) diperoleh bahwa f(t) dapat dinyatakan dalam h (t) sebagai berikut :

$$f(t) = h(t) \exp\left[-\int_0^t h(r)\,dr\right] \tag{2.3.13}$$

Dari penjelasan di atas, jelas bahwa ketiga fungsi waktu ketahanan tersebut berhubungan satu dengan yang lain.

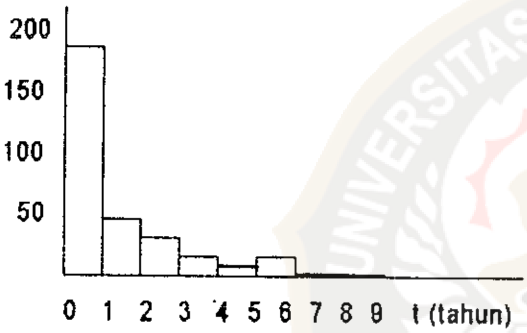
Fungsi hazard mempunyai beberapa kemungkinan, yaitu cenderung naik, cenderung turun, ataupun konstan . Untuk lebih jelasnya dalam memberikan pengertian fungsi ketahanan, fungsi kerapatan kegagalan serta fungsi hazard, akan dipaparkan melalui suatu contoh mengenai ketiga fungsi tersebut.

Contoh . Berdasarkan laporan-laporan Mac Donald (1963) tentang ketahanan hidup dari 256 pria dengan penyakit Malignant melanoma (tumor penyebab kematian) pada periode 1944 sampai 1960 dengan ijin dari klinik Tumor M.D. Anderson, yang dapat dilihat pada Tabel 2.1 dengan pola ketahanan dari pasien-pasien tersebut.

Tabel 2.1. Pola Ketahanan Data

Waktu survival (tahun)	Jumlah pasien yang hidup pada awal interval (orang)	Jumlah pasien yang mati dalam interval (orang)
0-1	256	167
1-2	89	48
2-3	41	23
3-4	18	6
4-5	12	3
5-6	9	6
6-7	3	1
7-8	2	1
8-9	1	1
9 keatas	0	0

Berdasarkan pola ketahanan data di atas , dapat dibuat suatu histogram jumlah kematian pasien dalam setiap interval data. Sehingga dapat ditentukan probabilitas pasien yang mati f(t) dengan menggunakan grafik poligon frekuensi yang diperoleh dengan cara menghubungkan titik tengah dari histogram tersebut. Dengan menggunakan tabel 2.1 di atas akan dijelaskan pula proses menentukan taksiran ketahanan hidup S(t) dan rata-rata hazard h(t).



Gambar 2.3

Gambar 2.3. Histogram dari jumlah kematian dalam setiap interval data Tabel 2.1.

Gambar 2.4. adalah grafik poligon frekuensi. Jadi f(0,5) = 167/256 = 0,652 ; f(1,5) = 48/256 = 0,188 dan seterusnya seperti terlihat pada tabel 2.2.

Sedangkan gambar 2.5. adalah grafik ketahanan taksiran $\hat{S}(t)$ untuk 256 pasien. $\hat{S}(t)$ = 1 pada saat t = 0 karena seluruh 256 pasien masih hidup. Setelah setahun menderita penyakit tumor tersebut, tinggal 89 pasien yang hidup. Jadi probabilitas seorang pasien bertahan hidup paling sedikit selama 1 tahun taksirannya adalah $\hat{S}(1)$ = 89/256 = 0,347. Setelah menderita penyakit selama 2 tahun hanya 41 pasien yang masih hidup. Taksiran probabilitas seorang pasien bertahan hidup paling sedikit selama 2 tahun adalah $\hat{S}(2)$ = 41/256 = 0,160. Dengan cara yang sama dapat diperoleh $\hat{S}(3)$ sampai $\hat{S}(9)$ seperti pada tabel 2.2.

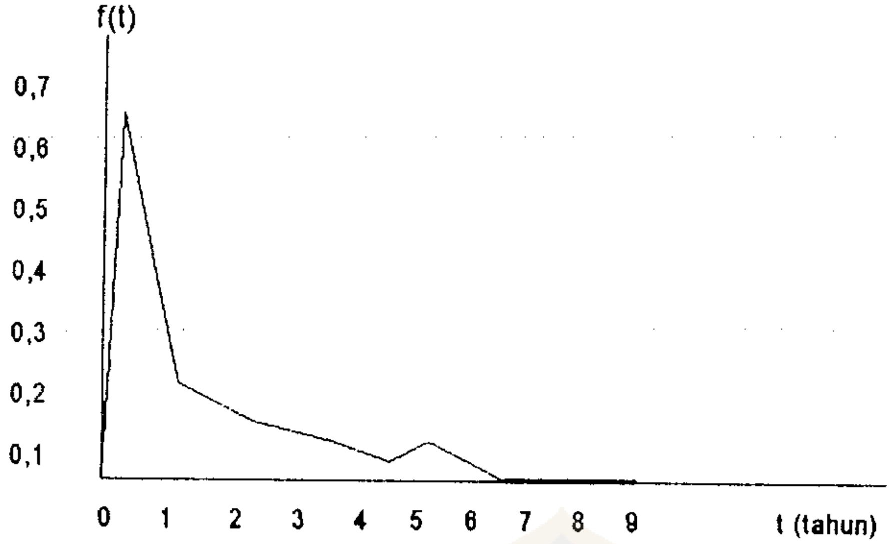
Pada gambar 2.6 grafik taksiran hazard adalah jumlah kematian per unit waktu dalam interval dibagi jumlah ketahanan awal interval. Taksiran hazard untuk 256 pasien, $\hat{h}(1)$ = 167/256 = 0,652 adalah taksiran hazard pada akhir tahun pertama. Kemudian $\hat{h}(2)$ = 48/89 = 0,539 adalah taksiran hazard pada akhir tahun kedua. Dengan cara yang sama dapat diperoleh $\hat{h}(3)$ sampai $\hat{h}(9)$ seperti pada tabel 2.2.

Hasil-hasil perhitungan disajikan pada tabel 2.2 sebagai berikut :

Tabel 2.2.

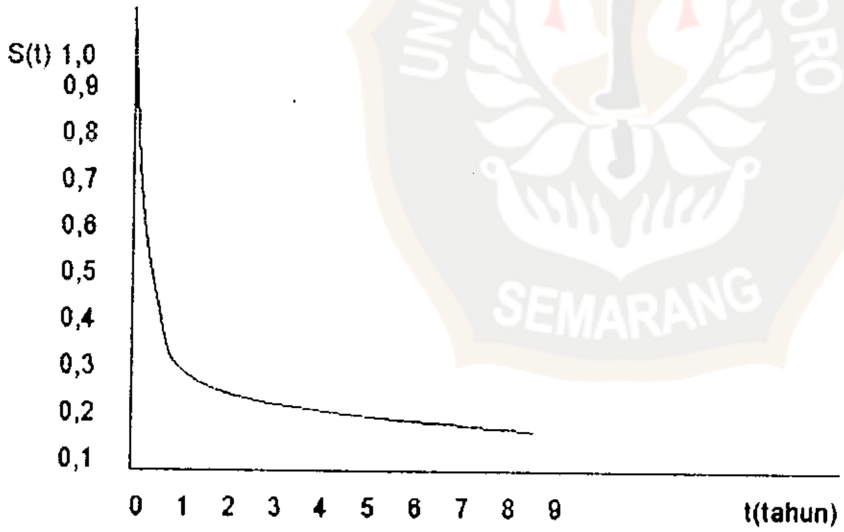
Waktu survival (tahun)	Jumlah Pasien hidup awal interval (orang)	Jumlah pasien meninggal dalam interval (orang)	Taksiran f (t)	Taksiran S (t)	Taksiran h (t)
0-1	256	167	0,652	1	0,652
1-2	89	48	0,188	0,347	0,539
2-3	41	23	0,090	0,160	0,561
3-4	18	6	0,023	0,070	0,333
4-5	12	3	0,012	0,047	0,250
5-6	9	6	0,023	0,035	0,667
6-7	3	1	0,004	0,012	0,333
7-8	2	1	0,004	0,008	0,500
8-9	1	1	0,004	0,004	1,000
9 +	0	0	0		

Berikut merupakan grafik-grafik dari poligon frekuensi, ketahanan hidup dan rata-rata hazard yang data-datanya diperoleh dari tabel.2.2

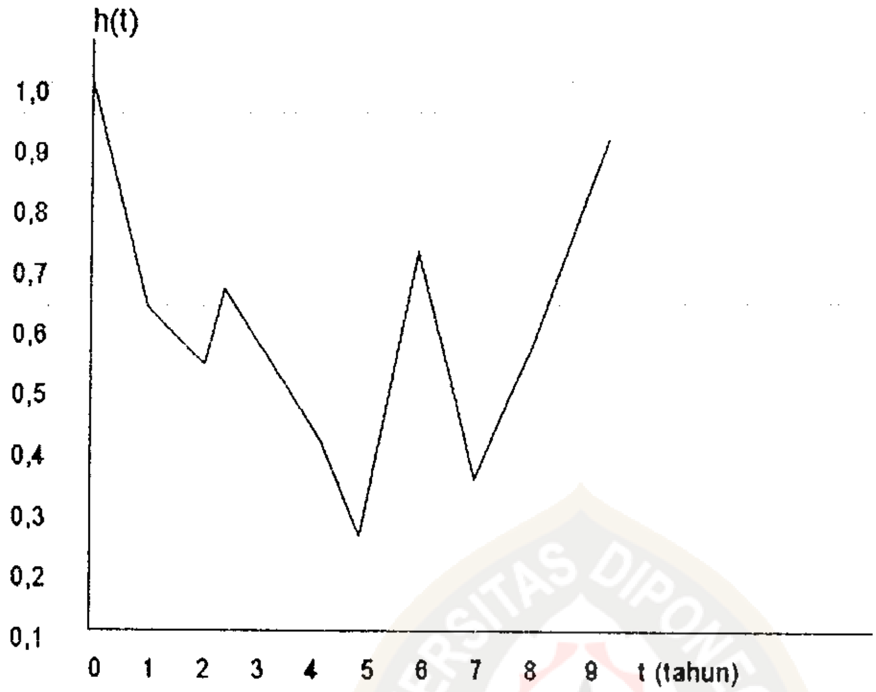


Gambar 2.4

Gambar 2.4. Grafik Poligon Frekuensi yang diperoleh dengan Menghubungkan titik tengah dari Histogram gambar 2.3



Gambar 2.5. Kurva Survival taksiran S(t) pada tabel 2.2



Gambar 2.6.

Gambar 2.6. Taksiran rata-rata Hazard untuk penderita Malignant Melanoma.

2.4. Jenis-jenis Penyensoran

Dalam analisis ketahanan dapat terjadi individu yang diamati tersensor. Masalah sensor ini merupakan suatu hal yang membedakan analisis ketahanan (*survival analysis*) dari bidang-bidang ilmu statistik lainnya. Data tersensor adalah data yang diperoleh sebelum hasil yang diinginkan dari suatu penelitian terjadi, sedangkan waktu penelitian tersebut telah berakhir oleh suatu sebab. Data yang mengalami penyensoran hanya memuat sebagian informasi mengenai variabel random yang diperhatikan, namun berpengaruh terhadap pengertian-pengertian dan perhitungan-perhitungan statistik.

Dalam tugas akhir ini hanya akan dibahas penyensoran secara progresif dan penyensoran secara tunggal yang merupakan bagian dari metode penyensoran .

2.4.1. Penyensoran Data Secara Progressif (progressively cencored data)

Definisi 2.4.1 :

Penyensoran data secara progresif merupakan penyensoran yang dilakukan dalam waktu yang berlainan, dimana tiap-tiap obyek memasuki pengamatan dalam waktu yang berbeda-beda.

Jadi setiap data ketahanan selalu memuat informasi mengenai sensor. Jika T_1, $T_2,\ldots, T_n$ berdistribusi identik dengan fungsi distribusi F, maka terdapat $C_1, C_2,\ldots, C_n$ berdistribusi identik dengan fungsi distribusi G. C_i adalah waktu sensor yang berhubungan dengan T_i. hanya dapat diperhatikan $(Y_1, \delta_1), \ldots (Y_i,\delta_i)$, $i = 1,2,3,\ldots.n$

dimana : $Y_i = \mathrm{Min}(T_i,C_i)$

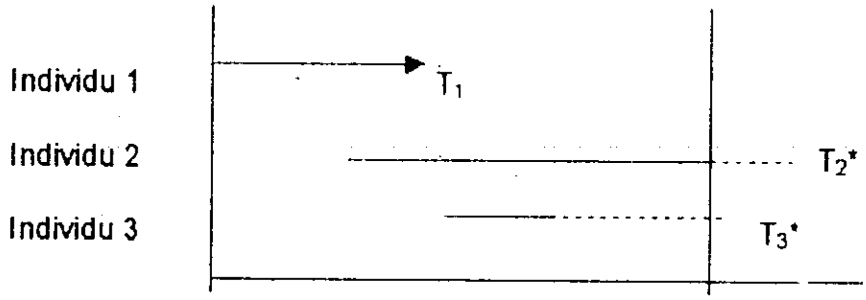
$$\delta_i = I(T_i \le C_i) \begin{cases} = 1, \text{ jika } T_i \le C_i, \text{ maka } T_i \text{ tak tersensor} \\ = 0, \text{ jika } T_i > C_i, \text{ maka } T_i \text{ tersensor} \end{cases}$$

fungsi likelihoodnya adalah :

$$L(T/\theta) = \prod_{i=1}^{n} [f(t_i ; \theta)]^{\delta i} \quad [S(C_i ; \theta)]^{1-\delta i}$$

$Y_1, Y_2, \ldots, Y_n$ adalah berdistribusi identik dengan fungsi distribusi H.

Juga $\delta_1, \delta_2, \ldots, \delta_n$ berisi informasi sensor. Dengan penyensoran progresif akan dibuat asumsi yang penting, T_i dan C_i adalah independen. Perhatikan gambar 2.7:



0 waktu ketahanan akhir pengamatan

Gambar 2.7

Keterangan :

Individu pertama diamati pada waktu t = 0 dan gagal atau mati pada waktu T_1, sehingga menghasilkan pengamatan tidak tersensor. Individu kedua diamati pada waktu tertentu dan pada akhir penelitian tetap bertahan yang menghasilkan pengamatan tersensor T_2^*. Sedangkan individu ketiga mulai diamati pada waktu tertentu dan pengamatan tidak dapat dilanjutkan atau terhenti sebelum penelitian berakhir dalam hal ini menghasilkan pengamatan tersensor T_3^*.

Penyensoran secara progresif sering terjadi pada aplikasi kedokteran dengan percobaan terhadap binatang maupun percobaan-percobaan klinik.

Hal-hal yang menyebabkan penyensoran diantaranya :

1. *Lost to Follow-up* : Data yang tidak dapat diperoleh karena obyek pengamatan hilang sebelum pengamatan berakhir.

   Misal pasien memutuskan untuk berpindah tempat, sehingga peneliti tidak dapat menemuinya lagi.

2. *Drop out* : Data yang tidak dapat diperoleh meskipun waktu pengamatan belum berakhir dan belum mencapai kegagalan.

Misal terapi yang diberikan memiliki efek sampingan sehingga tidak perlu dilanjutkan lagi, atau pasien masih dapat dihubungi tetapi menolak untuk melanjutkan pemberian perlakuan.

3. Berakhirnya masa penelitian sebelum kegagalan terjadi.

   Misalnya disebabkan pengamatan membutuhkan waktu yang sangat panjang dan biaya yang relatif besar.

2.4.2. Penyensoran Data Secara Tunggal (singly cencored data)

Definisi 2.4.2:

Penyensoran data secara tunggal adalah penyensoran yang dilakukan dalam waktu bersamaan, maksudnya tiap-tiap obyek memasuki pengamatan dalam waktu bersamaan.

$d < n$ adalah ditentukan.

$T_{(1)} < T_{(2)}, \ldots, < T_{(n)}$ adalah statistik terurut dari $T_1, T_2, \ldots, T_n$. Setelah diperiksa pada urutan kegagalan ke-d maka dapat diperhatikan $T_{(1)}, T_{(2)}, \ldots, T_{(d)}$.

Penyensoran secara tunggal sering kali terjadi dalam aplikasi teknik. Misalkan dalam situasi tertentu terdapat tumpukan transistor, semuanya diletakkan dalam suatu pengujian dengan memberikan waktu awal test pada $t = 0$, dan waktu kegagalannya dicatat.

Beberapa transistor membutuhkan waktu panjang untuk terbakar, dan tidak akan ditunggu selama waktu tersebut untuk menghentikan pengamatan. Maka diputuskan untuk menunggu sampai pada fraksi tertentu d/n dari transistor yang terbakar, dalam hal ini akan diperoleh sensor tunggal.

Urutan observasi lengkap adalah :

$Y_{(1)} = T_{(1)}$

$Y_{(2)} = T_{(2)}$

$\vdots$

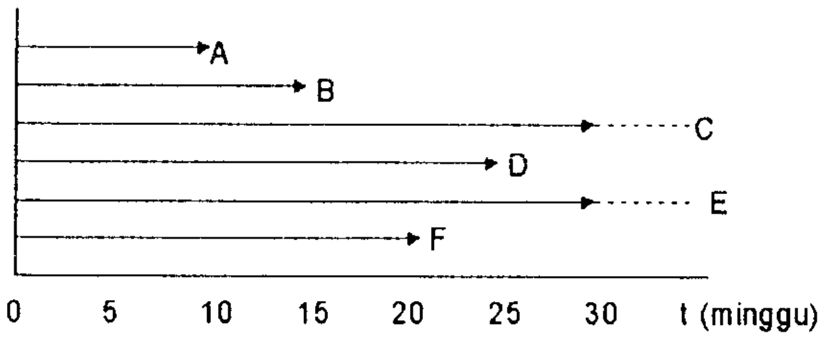
$Y_{(d)} = T_{(d)}$

$Y_{(d+1)} = T_{(d)}$

$\vdots$

$Y_{(n)} = T_{(d)}$

Contoh. Seorang peneliti melakukan percobaan terhadap enam ekor tikus percobaan dengan menyuntikkan bibit-bibit tumor melalui telapak kakinya. Pengamatan dimulai pada saat tikus-tikus itu diberi disuntikan pada waktu yang bersamaan. Karena pertimbangan tertentu, pengamat memutuskan untuk menghentikan pengamatan setelah tigapuluh minggu. Hasil pengamatan dapat dilihat pada gambar 2.8 yang menandakan waktu tumbuh tumor keenam ekor tikus tersebut.



Gambar 2.8

Keterangan :

Tikus-tikus A, B, D dan F terjangkit tumor setelah 10, 15, 25 dan 20 minggu, sedangkan tikus-tikus C dan E sampai pengamatan selesai belum terjangkit tumor. Jadi waktu ketahanan, yaitu waktu bebas tumor, keenam ekor tikus tersebut ialah 10, 15, 30*, 25, 30* dan 20 minggu (tanda "*" menyatakan data tersensor). Setelah diurutkan data waktu ketahanannya adalah 10, 15, 20, 25, 30*, 30* minggu.

2.5. Metode Perhitungan Numerik pada Data Ketahanan

Dalam suatu proses perhitungan data ketahanan terdapat permasalahan untuk mencari nilai estimasi dari suatu parameter, misalnya θ yang mempunyai persamaan :

$$g(\hat{\theta}) = 0 \qquad (2.5.1)$$

Bila $g(\hat{\theta})$ berbentuk kuadrat , pangkat tiga, atau pangkat empat maka ada rumus - rumus aljabar untuk menghitung akar-akarnya sebaliknya, bila $g(\hat{\theta})$ suatu polinom berderajat tinggi atau berbentuk transenden seperti; $1 + \cos\hat{\theta} - 5\hat{\theta}$, $\log\hat{\theta}^2$, $e^{-\hat{\theta}}$ - $\sin\hat{\theta}$ dan seterusnya, tidak tersedia metode aljabar untuk solusinya. Persamaan ini akan lebih sering dijumpai pada data ketahanan yang diasumsikan mengikuti pola suatu fungsi distribusi ketahanan yang mempunyai lebih dari satu parameter. Untuk menyelesaikan persamaan tersebut digunakan suatu metode perhitungan numerik yaitu metode Newton-Raphson dan metode interpolasi.

2.5.1. Metode Iterasi Newton-Raphson

Misal $\hat{\theta}_0$ adalah aproksimasi akar dari $g(\hat{\theta}) = 0$ dan h adalah kekeliruan dari aproksimasi tersebut yang berharga positif dan sangat kecil sedemikian hingga $\hat{\theta}_1 = \hat{\theta}_0 + h$ dan $g(\hat{\theta}_1) = 0$. Ekspansi $g(\hat{\theta}_1) = 0$ oleh deret Taylor diperoleh :

$$g(\hat{\theta}_0) + \frac{h}{1!} g'(\hat{\theta}_0) + \frac{h^2}{2!} g''(\hat{\theta}_0) + \ldots\ldots = 0$$

Karena aproksimasi yang terbaik adalah yang kekeliruannya terkecil, dalam hal ini h^2 sangat kecil sekali hingga nilainya dapat diabaikan. Maka turunan kedua atau lebih dari persamaan di atas dapat diabaikan pula sehingga :

$$g(\hat{\theta}_0) + h\, g'(\hat{\theta}_0) = 0$$

$$h = -\left[\frac{g(\hat{\theta}_0)}{g'(\hat{\theta}_0)}\right]$$

Aproksimasi yang baik dari $\hat{\theta}_0$ diberikan oleh $\hat{\theta}_1$, dengan :

$$\hat{\theta}_1 = \hat{\theta}_0 - \left[\frac{g(\hat{\theta}_0)}{g'(\hat{\theta}_0)}\right]$$

Karena $\hat{\theta}_0$ adalah nilai taksiran awal dari $\hat{\theta}$, maka iterasi ke-v dari $\hat{\theta}$ dinyatakan dengan :

$$\hat{\theta}_{v+1} = \hat{\theta}_v - \left[\frac{g(\hat{\theta}_v)}{g'(\hat{\theta}_v)}\right] \qquad (2.5.2)$$

Persamaan (2.5.2) merupakan formula iterasi Newton-Raphson.

2.5.3. Metode Interpolasi Linear

Metode interpolasi linear adalah suatu perhitungan dengan cara menaksir harga-harga tengahan diantara titik-titik data yang telah tepat. Bentuk yang paling mudah dari interpolasi ini adalah menghubungkan segitiga sebangun :

$$\frac{g_1(\hat{\theta}) - g(\hat{\theta}_0)}{\hat{\theta} - \hat{\theta}_0} = \frac{g(\hat{\theta}_1) - g(\hat{\theta}_0)}{\hat{\theta}_1 - \hat{\theta}_0}$$

Yang dapat diatur kembali, sehingga memenuhi :

$$g_1(\hat{\theta}) = g(\hat{\theta}_0) + \left(g(\hat{\theta}_1) - g(\hat{\theta}_0) \right) \left[\frac{\hat{\theta} - \hat{\theta}_0}{\hat{\theta}_1 - \hat{\theta}_0} \right] \qquad (2.5.5)$$

Persamaan (2.5.5) adalah sebuah formula interpolasi linear. Semakin kecil interval diantara titik-titik data, aproksimasinya semakin baik.

2.6. Uji Likelihood Rasio

Uji likelihood rasio merupakan suatu prosedur pengujian hipotesa terhadap alternatifnya dengan menggunakan estimasi maksimum likelihood. Misal t_1, t_2, ... , t_n adalah sampel random berukuran n dari suatu populasi dengan fungsi kerapatan $f(t; \theta_1, \theta_2, \ldots, \theta_k)$ dengan $n > k$. Dasar dari sampel ini adalah menguji hipotesa H_0 : r dari k parameter θ_1, θ_2, ... , θ_k yang ditentukan terhadap alternatifnya H_a : semua parameter yang tidak ditentukan. Bila ω mewakili (k-r) dimensi dari parameter-parameter yang tidak ditentukan di bawah H_0, dan Ω adalah k- dimensi parameter-parameter yang tidak ditentukan di bawah H_a. Diasumsikan di bawah H_0 parameter-parameter θ_1, θ_2, ... , θ_r ditentukan dan berharga $\theta_1^0, \theta_2^0, \ldots, \theta_r^0$ dengan menyusun kembali parameter yang tidak ditentukan

$\theta_{r+1}, \theta_{r+2}, \ldots, \theta_k$ dapat diperoleh taksiran – taksiran dari maksimum likelihoodnya yaitu $\hat{\theta}_{r+1}, \hat{\theta}_{r+2}, \ldots, \hat{\theta}_k$.

$$\hat{L}(\omega) = \prod_{i=1}^{n} f(t_i; \theta_1^0, \theta_2^0, \ldots, \theta_r^0, \hat{\theta}_{r+1}, \hat{\theta}_{r+2}, \ldots, \hat{\theta}_k) \tag{2.6.1}$$

$$\hat{L}(\Omega) = \prod_{i=1}^{n} f(t_i; \hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3, \ldots, \hat{\theta}_k) \tag{2.6.2}$$

Uji statistik dari likelihood rasio didefinisikan :

$$\Lambda = -2 \ln \left(\frac{\hat{L}(\omega)}{\hat{L}(\Omega)} \right) \tag{2.6.3}$$

Hipotesa dua arah :

$H_0 : \theta = \theta_0$ terhadap alternatifnya $H_1 : \theta \neq \theta_0$

Daerah kritis dari uji likelihood rasio adalah bila nilai Λ melebihi nilai $\chi^2(\alpha;1)$ pada tingkat α distribusi chi – kuadrat dengan 1 derajat bebas untuk sampel minimal 25 ($n \geq 25$).