# ANALISIS REGRESI PARAMETRIK TAHAN HIDUP UNTUK DATA TAHAN HIDUP BAYI DI KABUPATEN SLEMAN YOGYAKARTA

**SKRIPSI**

**Dosen pembimbing :**

**Drs. Gunardi, M.Si**

**Disusun Oleh :**

**Titin Herlina**

**99 611 049**

**JURUSAN STATISTIKA**

**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM**

**UNIVERSITAS ISLAM INDONESIA**

**YOGYAKARTA**

**2004**

## HALAMAN PENGESAHAN DOSEN PENGUJI

**ANALISIS REGRESI PARAMETRIK TAHAN HIDUP**
**UNTUK DATA TAHAN HIDUP BAYI**
**DI KABUPATEN SLEMAN YOGYAKARTA**

**SKRIPSI**

**TITIN HERLINA**
**990051013206120047**
**99 611 049**

Telah Dipertahankan Dihadapan Panitia Penguji Tugas Akhir
Jurusan Statistika Fakultas Matematika Dan Ilmu Pengetahuan Alam
Universitas Islam Indonesia Dan dinyatakan *LULUS*
Pada Tanggal 4 Maret 2004

| **Penguji** | **Tanda Tangan** |
|---|---|
| 1. **Drs. Gunardi, M.Si** | ............................ |
| 2. **Jaka Nugraha, M.Si** | ............................ |
| 3. **Kariyam, M.Si** | ............................ |
| 4. **Rohmatul Fajriyah, M.Si** | ............................ |

**Mengetahui,**
**Dekan Fakultas Matematika Dan Ilmu Pengetahuan Alam**
**Universitas Islam Indonesia**

UNIVERSITAS ISLAM INDONESIA YOGYAKARTA FAKULTAS MIPA

**( Jaka Nugraha, M.Si )**

# *HALAMAN MOTTO*

***Musibah apa saja yang menimpa kamu, itu disebabkan oleh hasil perbuatan tangan kamu sendiri. Dan Allah Maha Memaafkan sebagian besar dari kesalahanmu.***

*(QS. Asy-Syuura: 30)*

***Dan bahwa manusia hanyalah memperoleh hasil dari apa yang dia usahakan.***

*(QS. An-Najm: 39)*

# *HALAMAN PERSEMBAHAN*

*Sebuah karya yang sederhana ini sepenuhnya dipersembahkan kepada :*

*Ayah dan Ibu Tercinta*

*Kalian adalah anugerah terindah yang Allah berikan padaku*

*Terimakasih untuk kasih sayang dan doa yang tak henti-hentinya kalian panjatkan sehingga aku selalu termotivasi untuk selalu hidup dalam jalanNya*

*My Lovely brother and sister*

*Terima kasih untuk kasih untuk kasih sayang dan doa kalian*

*I Always Love You*

# KATA PENGANTAR

Bismillahirrahmanirrahim

Assalamu'alaikum Wr.Wb.

Segala puji bagi Allah SWT yang telah melimpahkan rahmat dan karunia Nya sehingga dapat menyelesaikan tugas akhir dengan judul " **ANALISIS REGRESI PARAMETRIK TAHAN HIDUP UNTUK DATA TAHAN HIDUP BAYI DI KABUPATEN SLEMAN YOGYAKARTA"**.

Tugas akhir ini merupakan syarat kelulusan untuk mencapai gelar sarjana Science di Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Islam Indonesia.

Banyak pihak yang telah membantu pelaksanaan maupun penulisan tugas akhir ini, sehingga sudah sepantasnya mengucapkan terima kasih kepada :

- Bapak Jaka Nugraha, M.Si, selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Islam Indonesia Yogyakarta
- Bapak Drs. Gunardi, Msi selaku dosen pembimbing, yang senantiasa meluangkan waktu dan memberi arahan maupun bimbingan dalam menyelesaikan tugas akhir ini.
- Ibu Rohmatul Fajriyah, M.Si, selaku Ketua Jurusan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Islam Indonesia Yogyakarta, terima kasih atas buku-buku yang dipinjamkan.

- Kedua orang tua dan semua saudaraku dengan segala do'a dan dukungan yang telah diberikan.
- Semua teman-teman di jurusan Statistika UII 99.
- Teman-teman kos di Pandega Bhakti 12.
- Semua pihak yang tidak dapat disebutkan satu persatu yang telah membantu terselesaikannya tugas akhir ini.

Semoga amal ibadah dan kebaikan yang telah diberikan mendapatkan imbalan yang sepantasnya dari Allah SWT.

Tugas akhir ini tentu masih banyak kekurangannya, karena itu mohon saran dan kritik yang membangun demi penyempurnaannya.

Wassalamu'alaikum Wr.Wb

Yogyakarta, Maret 2004

(Titin Herlina )

# DAFTAR ISI

## DAFTAR TABEL DAN GRAFIK

**INTISARI**

**ANALISIS REGRESI PARAMETRIK TAHAN HIDUP**

**UNTUK DATA TAHAN HIDUP BAYI**

**DI KABUPATEN SLEMAN YOGYAKARTA**

**Oleh :**

**TITIN HERLINA**

**99611049**

Skripsi ini membahas tentang analisis regresi parametrik tahan hidup, dimana tahan hidup tersebut diasumsikan berdistribusi Webull. Tahan Hidup diartikan sebagai panjang waktu suatu individu maupun benda untuk bertahan hidup atau bekerja.

Data berupa kematian bayi hingga usia satu tahun di Kabupaten Sleman Yogyakarta pada tahun 2003. Dari data tersebut bisa diketahui faktor yang mempengaruhi tahan hidup tersebut.

**Kata Kunci : Model Regresi Parametrik Tahan Hidup, Distribusi Weibull.**

**ABSTRACT**

**ANALYSE REGRESSION PARAMETRIC MODELS WITH LIFETIME For**

**FOR SURVIVAL DATA OF BABY**

**REGENCY of SLEMAN YOGYAKARTA**

**By :**

**TITIN HERLINA**

**99 611 049**

This Skripsi study about analysis of regression parametrics models with life time, where holding up the life assumed by distribution Weibull. Life time interpreted by as time length of an individual and also object to live on or work the.
Data in the form of baby death till one year in Regency of Sleman Yogyakarta in the year 2003. From the data can be known by the factor influencing to life time.
***Keyword : Regression Parametric models with life time, Distribution Weibull.***

# BAB I

# PENDAHULUAN

## 1.1 Latar Belakang

Dalam suatu penelitian sering terjadi adanya ketergantungan antara variabel yang diuji dengan satu atau beberapa variabel lain. Untuk mengetahui hubungan antar variabel tersebut digunakan analisis statistik, yang dikenal dengan nama analisis regresi. Analisis regresi akan membentuk suatu model regresi yang menjelaskan ketergantungan suatu variabel (disebut variabel dependen atau respon) pada satu atau beberapa variabel lain (disebut variabel independen atau prediktor).

Dalam tugas akhir ini akan diperkenalkan suatu populasi tahan hidup. Tahan hidup merupakan panjang waktu suatu individu baik mahluk hidup maupun benda untuk bertahan hidup atau bekerja. Uji tahan hidup ini biasanya digunakan untuk menguji daya tahan hidup atau kehandalan suatu produk hasil industri, karena itu biasanya sering digunakan dalam bidang medis, industri, biologi, dan teknik. Cara uji hidup dilakukan dengan cara menguji sebanyak n benda pada kondisi normal hingga semua benda mati (sampel lengkap) atau percobaan dihentikan sebelum semua benda mati (sampel tersensor).

Jika ingin mengetahui hubungan antara variabel tahan hidup dengan variabel lain yang dianggap mempengaruhi maka analisis yang digunakan adalah analisis regresi parametrik data tahan hidup dan akan diasumsikan berdistribusi tertentu, dalam analisis ini akan diasumsikan berdistribusi Weibull.

Data yang digunakan dalam tugas akhir ini adalah tahan hidup bayi, yaitu usia bayi ketika lahir hingga meninggal dunia. Data tersebut merupakan data kematian bayi hingga usia 1 tahun di Kabupaten Sleman Propinsi Yogyakarta pada tahun 2003. Dari analisis tersebut ingin diketahui variabel apa saja yang mempengaruhi kematian bayi, dimana variabel yang dianggap mempengaruhi kematian bayi tersebut adalah usia kehamilan, keadaan bayi ketika lahir, jenis kelamin, dan lain-lain. Sehingga variabel tersebut bisa lebih diperhatikan dan diharapkan bisa menurunkan angka kematian bayi.

## 1.2 Rumusan Masalah

Berdasarkan latar belakang tersebut maka rumusan masalahnya adalah apakah usia kehamilan, jenis kelamin, banyaknya suntikan TT, pertolongan medis (dokter, bidan, perawat) sebelum meninggal, dan keadaan bayi ketika lahir bisa mempengaruhi tahan hidup bayi yang berusia hingga 1 tahun. Hubungan tersebut akan ditunjukkan oleh model regresi untuk tahan hidup bayi.

## 1.3 Batasan Masalah

Batasan masalah dalam penelitian ini adalah sebagai berikut :

1. Data berupa kematian bayi yang berusia hingga 1 tahun pada tahun 2003 di Kabupaten Sleman propinsi Yogyakarta.
2. Metode analisis yang digunakan adalah analisis regresi parametrik tahan hidup dan diasumsikan berdistribusi Weibull.
3. Software statistik yang digunakan adalah Minitab 13.

## 1.4 Tujuan Penelitian

Tujuan diadakan penelitian ini adalah untuk mengetahui model regresi tahan hidup bayi yang berusia hingga 1 tahun di Kabupaten Sleman Propinsi Yogyakarta, sehingga variabel-variabel prediktor dari model bisa lebih diperhatikan untuk mengurangi angka kematian bayi.

## 1.5 Manfaat Penelitian

Adapun manfaat dari penelitian ini adalah sebagai berikut :

1. Menambah pengetahuan baik bagi peneliti maupun pembaca dan merupakan syarat untuk memperoleh derajat sarjana di fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Islam Indonesia.

2. Dengan memperhatikan faktor-faktor yang mempengaruhi kematian bayi diharapkan bisa mengurangi angka kematian bayi di bawah usia 1 tahun di Kabupaten Sleman.

## 1.6 Sistematika Penulisan

Sistematika Penulisan dari tugas akhir ini adalah :

Bab I Pendahuluan

Berisikan uraian tentang latar belakang, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, dan sistematika penulisan.

Bab II Landasan Teori

Memuat penjelasan tentang konsep distribusi tahan hidup dan teori-teori yang dipakai dalam analisis regresi tahan hidup.

Bab III Model Regresi Untuk Data Tahan Hidup

Bab ini membahas tentang model regresi Weibull termasuk estimasi parameter regresi, yaitu menggunakan metode maksimum likelihood.

Bab IV Analisa dan Pembahasan

Menentukan model regresi yang melibatkan faktor-faktor yang diduga mempengaruhi tahan hidup bayi yang berusia hingga 1 tahun.

Bab V Kesimpulan dan Saran

Memuat tentang Kesimpulan dari analisis yang dilakukan dan saran

# BAB II

# LANDASAN TEORI

## 2.1 Konsep Dasar Distribusi Tahan Hidup

Uji tahan hidup berguna untuk melakukan pengujian tentang daya tahan hidup dan kehandalan suatu produk hasil industri dan atau pengukuran tentang lamanya tahan hidup penderita dalam pengujian yang menyangkut pengobatan suatu penyakit.

Cara uji tahan hidup adalah dengan menguji sebanyak n benda pada kondisi normal sampai semua benda mati (sampel lengkap) dan atau percobaan dihentikan sebelum semua benda mati (sampel tersensor).

Jika T variabel random non negative kontinu pada interval $[0,\infty)$ yang menunjukkan tahan hidup suatu benda ataupun individu dalam suatu populasi, maka fungsi peluang densitas f(t) akan mempunyai fungsi distribusi kumulatif yang dinyatakan dalam persamaan (Lawless, 1982)

$$F(t) = \Pr(T \leq t) = \int_0^t f(x) \qquad (2.1.1)$$

Untuk fungsi "survivor" atau fungsi tahan hidup yang merupakan peluang suatu individu bertahan hidup sampai waktu t akan didefinisikan oleh persamaan :

$$S(t) = \Pr(T \geq t) = \int_t^\infty f(x)dx$$

Persamaan diatas bisa dibawa kebentuk persamaan lain, yaitu :

$$S(t) = 1 - \int_0^t f(x)dx$$

$$= 1 - F(t) \tag{2.1.2}$$

Fungsi lain yang juga terdapat dalam distribusi tahan hidup adalah fungsi hazard. Fungsi hazard h(t) menyatakan probabilitas kegagalan pada waktu t dengan syarat individu bertahan atau tetap hidup sampai waktu t.

Probabilitas bersyarat bahwa individu akan mati pada interval [t, t+$\Delta t$] bisa didefinisikan :

$$\Pr(t \le T < t + \Delta t | T \ge t) \tag{2.1.3}$$

Jika persamaan (2.1.3) dibagi $\Delta t$ dan diambil nilai limitnya maka akan dihasilkan persamaan :

$$h(t) = \frac{f(t)}{S(t)} \tag{2.1.4}$$

Bukti :

$$h(t) = \lim_{\Delta t \to 0} \frac{\Pr(t \le T \le t + \Delta t | T \ge t)}{\Delta t}$$

$$= \lim_{\Delta t \to 0} \frac{\Pr(t \le T \le t + \Delta t, T \ge t)}{\Delta t \Pr(T \ge t)}$$

$$= \lim_{\Delta t \to 0} \frac{\Pr(t \le T \le t + \Delta t)}{\Delta t (1 - F(t))}$$

$$= \frac{f(t)}{1-F(t)} = \frac{f(t)}{S(t)}$$

Hubungan antara fungsi peluang densitas f(t), fungsi survivor S(t) dan fungsi hazard h(t) bisa dinyatakan sebagai berikut :

1. Jika f(x) diketahui maka :

   (i) $F(t) = \int_0^t f(x)dx$

   (ii) $S(t) = 1 - \int_0^t f(x)dx$

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

2. Jika F(t) diketahui maka :

   (i) $f(t) = F'(t)$

   (ii) $S(t) = 1 - F(t)$

   (iii) $h(t) = \frac{F'(t)}{1-F(t)}$

3. Jika S(t) diketahui maka :

   (i) $F(t) = 1 - S(t)$

   (ii) $f(t) = -S'(t)$

Bukti :

$$S(t) = 1 - F(t)$$

$$S'(t) = 0 - F'(t) = -f(t)$$

$$f(t) = -S'(t)$$

(iii) $h(t) = \dfrac{-S'(t)}{S(t)}$

Dari persamaan diatas bisa dihasilkan fungsi kumulatif hazard, yaitu :

-H(t) = ln S(t)

Bukti :

$$-h(t) = \frac{S'(t)}{S(t)}$$

Jika dimisalkan t = x, maka

$$-\int_0^t h(x)dx = \int_0^t \frac{S'(x)}{S(x)}dx$$

$$-H(x)\Big|_0^t = \ln S(x)\Big|_0^t$$

$$-H(t) = \ln S(t)$$

Sehingga bisa dihasilkan persamaan-persamaan :

$$S(t) = e^{-H(t)}$$

$$F(t) = 1 - e^{-H(t)}$$

$$f(t) = h(t)e^{-H(t)}$$

## 2.2 Distribusi Peluang Penting Dalam Analisis Data Tahan Hidup

Distribusi yang sering dipakai dalam analisis data tahan hidup adalah distribusi Eksponensial dan distribusi Weibull.

### 2.2.1 Distribusi Eksponensial

Distribusi Eksponensial merupakan salah satu distribusi yang sering dipakai dalam analisis data tahan hidup. Distribusi ini mempunyai ciri khusus, yaitu fungsi hazard yang konstan terhadap waktu.

Jika T variabel random kontinu berdistribusi Eksponensial, fungsi peluang densitas dari T didefinisikan sebagai :

$$f(t) = \lambda e^{-\lambda t} \qquad (2.2.1)$$

Dari persamaan (2.2.1) bisa diturunkan fungsi survivor dan fungsi hazard untuk T berturut-turut adalah :

(i) $S(t) = e^{-\lambda t}$ (2.2.2)

(ii) $h(t) = \lambda$ (2.2.3)

Bukti :

(i) $S(t) = 1 - F(t)$

$$= 1 - \int_0^t \lambda e^{-\lambda x} dx$$

$$= 1 + \left( e^{-\lambda x} \Big|_0^t \right)$$

$$= 1 + \left( e^{-\lambda t} - 1 \right)$$

$$= e^{-\lambda t}$$

(ii) $h(t) = \dfrac{f(t)}{S(t)}$

$$= \frac{\lambda e^{-\lambda x}}{e^{-\lambda x}} = \lambda$$

Jika t variabel random berdistribusi Eksponensial, *moment generating function* dari t adalah :

$$M_x(t) = E(e^{tx}) = \int_0^{\infty} e^{tx} \lambda e^{-\lambda x} dx$$

$$= \lambda \int_0^{\infty} e^{-x(\lambda - t)} dx$$

$$= \frac{-\lambda}{\lambda - t} e^{-x(\lambda - t)} \Big|_0^{\infty}$$

$$= \frac{-\lambda}{\lambda - t} (e^{-\infty(\lambda - t)} - 1)$$

$$= \frac{\lambda}{\lambda - t}$$

Turunan pertamanya adalah :

$$M'(t) = \frac{\lambda}{(\lambda - t)^2}$$

$$M'(t = 0) = \frac{1}{\lambda}$$

Turunan keduanya adalah :

$$M''(t) == \frac{2\lambda}{(\lambda - t)^3}$$

$$M''(t=0) = \frac{2}{\lambda^2}$$

Sehingga rata-rata dari distribusi Eksponensial adalah :

$$E(T) = M'(0) = \frac{1}{\lambda}$$

Variansi dari distribusi Eksponensial adalah :

$$Var(T) = E(T^{2)}) - [E(T)]^2$$

$$= \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$$

### 2.2.2 Distribusi Weibull

Distribusi peluang yang memainkan peranan penting dalam analisis data tahan hidup adalah distribusi Weibull yang diperkenalkan oleh Waloddi Weibull (1951). Jika variabel tahan hidup T berdistribusi Weibull dengan parameter α dan β maka bisa ditulis T~WEI($\alpha,\beta$).

Fungsi peluang densitas f(t) untuk distribusi Weibull didefinisikan sebagai :

$$f(t,\alpha,\beta) = \left(\frac{\beta}{\alpha}\right)\left(\frac{t}{\alpha}\right)^{\beta-1} e^{-\left(\frac{t}{\alpha}\right)^{\beta}} \quad ; \quad \alpha > 0, \beta > 0 \qquad (2.2.4)$$

Dengan parameter skala $\alpha$ dan parameter bentuk $\beta$.

Jika T suatu variabel random berdistribusi Weibull dengan fungsi peluang densitas, maka fungsi distribusi kumulatif dari T adalah :

$$F(t;\alpha,\beta) = 1 - e^{\left(-t/\alpha\right)^{\beta}} \qquad (2.2.5)$$

Bukti :

$$F(t) = \int_0^t f(x)dx$$

$$= \int_0^t \frac{\beta}{\alpha}\left(\frac{x}{\alpha}\right)^{\beta-1} e^{-\left(\frac{x}{\alpha}\right)^{\beta}} dx$$

Misalkan diambil substitusi

$$y = \left(\frac{x}{\alpha}\right)^{\beta} ; y^{1/\beta} = \frac{x}{\alpha}$$

sehingga $\frac{1}{\beta} y^{\frac{1-\beta}{\beta}} dy = \frac{dx}{\alpha}$

Untuk $x = 0 \rightarrow y = 0; x = t \rightarrow y = \left(\frac{t}{\alpha}\right)^{\beta}$

Maka akan diperoleh :

$$F(t) = \int_0^t \frac{\beta}{\alpha}\left(\frac{x}{\alpha}\right)^{\beta-1} e^{-\left(\frac{x}{\alpha}\right)} dx$$

$$= \frac{\beta}{\alpha} \int_0^{(t/\alpha)^{\beta}} y^{\frac{\beta-1}{\beta}} e^{-y} \frac{1}{\beta} y^{\frac{1-\beta}{\beta}} \alpha dy$$

$$= \int_0^{(t/\alpha)^{\beta}} e^{-y} dy = 1 - e^{-(t/\alpha)^{\beta}}$$

Selanjutnya akan dihasilkan fungsi survivor dari T adalah :

$$S(t) = e^{-(t/\alpha)^{\beta}} \tag{2.2.6}$$

Sedangkan tingkat kegagalan pada waktu t dinyatakan sebagai fungsi hazard dengan bentuk :

$$h(t) = \frac{\beta}{\alpha}\left(\frac{t}{\alpha}\right)^{\beta-1} \tag{2.2.7}$$

Rata-rata dan variansi dari tahan hidup T yang berdistribusi Weibull dinyatakan dalam bentuk :

$$E(t) = \alpha\Gamma\left(1+\frac{1}{\beta}\right)$$

$$\text{var}(t) = \alpha^2\left[\Gamma\left(\frac{2}{\beta}+1\right)-\Gamma^2\left(\frac{1}{\beta+1}\right)\right]$$

Bukti :

$$E(t) = \int_0^\infty t\left(\frac{\beta}{\alpha}\right)\left(\frac{t}{\alpha}\right)^{\beta-1} e^{-\left(t/\alpha\right)^\beta} dt$$

$$= \beta\int_0^\infty \left(\frac{t}{\alpha}\right)^{\beta} e^{-\left(t/\alpha\right)^\beta} dt$$

$$= \beta\int_0^\infty ye^{-y}\frac{1}{\beta}\alpha y^{\frac{1-\beta}{\beta}} dy$$

$$= \alpha\int_0^\infty e^{-y} y^{\frac{1-\beta+\beta}{\beta}} dy$$

$$= \alpha\Gamma\left(1+\frac{1}{\beta}\right)$$

$$\mathrm{var}(t) = E(t^2) - [E(t)]^2$$

dengan

$$E(t^2) = \int_0^\infty t^2 f(t)dt$$

$$= \int_0^\infty t^2\left(\frac{\beta}{\alpha}\right)\left(\frac{t}{\alpha}\right)^{\beta-1} e^{-\left(t/\alpha\right)^\beta} dt$$

$$= \beta\int_0^\infty t\left(\frac{t}{\alpha}\right)^{\beta} e^{-\left(t/\alpha\right)^\beta} dt$$

$$= \beta\alpha\int_0^\infty \left(\frac{t}{\alpha}\right)^{\beta+1} e^{-\left(t/\alpha\right)^\beta} dt$$

$$= \beta\alpha\int_0^\infty y^{\frac{\beta+1}{\beta}} e^{-y} \frac{1}{\beta} y^{\frac{1-\beta}{\beta}} \alpha dy$$

$$= \alpha^2\int_0^\infty y^{\frac{2}{\beta}} e^{-y} dy$$

$$= \alpha^2\Gamma\left(\frac{2}{\beta}+1\right)$$

Jadi diperoleh variansi untuk distribusi Weibull adalah :

$$\operatorname{var}(t) = \alpha^2\Gamma\left(\frac{2}{\beta}+1\right) - \alpha^2\Gamma^2\left(\frac{1}{\beta}+1\right)$$

$$\operatorname{var}(t) = \alpha^2\left[\Gamma\left(\frac{2}{\beta}+1\right) - \Gamma^2\left(\frac{1}{\beta}+1\right)\right]$$

Persentil ke-p dari distribusi weibull diperoleh dengan menyelesaikan persamaan :

$$p = F(t_p) = 1 - e^{\left(t_p/\alpha\right)^\beta}$$

$$1 - p = e^{\left(t_p/\alpha\right)\beta}$$

$$\ln(1-p) = \left(t_p/\alpha\right)^\beta$$

$$\alpha[\ln(1-p)]^{1/\beta} = t_p$$

### 2.2.3 Distribusi Nilai Ekstrem

Dalam pembicaraan selanjutnya akan diperkenalkan suatu distribusi yang sering disebut sebagai distribusi nilai ekstrem (Exterm Value Distribution). Distribusi ini merupakan salah satu cara penyelesaian dalam inferensi distribusi

Weibull dan diperoleh dengan mentransformasikan data yang berdistribusi Weibull menggunakan transformasi logaritma.

Jika T variabel random tahan hidup berdistribusi Weibull dengan fungsi peluang densitas, maka untuk x= ln T berdistribusi nilai ekstrem dengan fungsi peluang densitas, akan dihasilkan persamaan :

$$f(x) = \frac{1}{b}\exp\left(\frac{x-\mu}{b}\right)\exp\left\{-\exp\left(\frac{x-\mu}{b}\right)\right\} \tag{2.2.8}$$

Bukti :

Dengan transformasi logaritma x = ln T, μ = $\ln\alpha$ atau T = expx, $\alpha =$ $\exp\mu$, $\beta = \frac{1}{b}$ dan $|J| = \frac{dt}{dx} = e^x$ maka akan diperoleh fungsi peluang dari x adalah sebagai berikut :

$$f(x) = f(e^x)|J|$$

$$= \left(\frac{e^{-\mu}}{b}\right)\left(\frac{e^x}{e^\mu}\right)^{\frac{1}{b}-1}\exp\left(-\left(\frac{e^x}{e^\mu}\right)^{\frac{1}{b}}\right)e^x$$

$$= \frac{1}{b}\exp\left(\frac{x-\mu}{b}\right)\exp\left\{-\exp\left(\frac{x-\mu}{b}\right)\right\}$$

Karena $t \geq 0$ maka $e^x \geq 0$ atau $-\infty < x < \infty$ dengan $\mu$ adalah parameter skala $(-\infty < \mu < \infty)$ dan b adalah parameter lokasi ( b > 0 ).

Jika t variabel random berdistribusi nilai ekstrem dengan fungsi peluang seperti diatas, maka fungsi kumulatif untuk x diperoleh sebagai berikut :

$$F(x) = \int_{-\infty}^{x} f(t)dt$$

$$= \int_{0}^{x} \frac{1}{b} \exp\left(\frac{t-\mu}{b}\right) \exp\left\{-\exp\left(\frac{t-\mu}{b}\right)\right\} dt$$

diambil substitusi $r = \exp\left(\frac{t-\mu}{b}\right)$ dengan $dr = \frac{1}{b}\exp\left(\frac{t-\mu}{b}\right)dt$ .

Untuk $t = -\infty \rightarrow r = 0$ dan $t = x \rightarrow r = \exp\left(\frac{x-\mu}{b}\right)$, sehingga bentuk di atas menjadi :

$$F(x) = \int_{0}^{\exp\left(x-\mu/b\right)} e^{-r} dr = -e^{-r}\Big|_{0}^{\exp\left(x-\mu/b\right)}$$

$$= 1 - \exp\left\{-\exp\left(\frac{x-\mu}{b}\right)\right\}$$

Selanjutnya akan diperoleh persamaan untuk fungsi survivor dan fungsi hazard dari distribusi nilai ekstrem, yaitu :

$$S(x) = \exp\left\{-\exp\left(\frac{x-\mu}{b}\right)\right\} \tag{2.2.9}$$

$$h(x) = \frac{1}{b}\exp\left(\frac{x-\mu}{b}\right) \tag{2.2.10}$$

Persentil ke-p dari distribusi nilai ekstrem, yaitu $x_p$ diperoleh dengan menyelesaikan persamaan :

$$P = F(x_p) = 1 - \exp\left(-\exp\left(\frac{x_p-\mu}{b}\right)\right)$$

Sehingga diperoleh :

$$x_p = \mu + b\ln(-\ln(1-p))$$

## 2.3 Metode Pengambilan Sampel

Dalam analisis data tahan hidup terdapat dua macam sampel, yaitu sampel lengkap dan sampel tersensor.

### 2.3.1 Sampel Tersensor

Sampel tersensor adalah data yang teramati oleh seorang peneliti sampai timbulnya suatu kejadian. Kejadian ini diartikan berupa kematian bagi suatu organisme hidup atau terjadinya suatu penyakit, waktu kegagalan suatu komponen, masa inkubasi suatu penyakit dan lain-lain tergantung dari obyek yang diteliti.

Tipe-tipe sensor dalam uji tahan hidup ini ada dua macam, yaitu penyensoran tipe I dan penyensoran tipe II.

1. Penyensoran tipe I

Dalam suatu eksperimen uji tahan hidup, n item ditempatkan pada suatu uji dan diputuskan untuk mengakhiri uji setelah waktu L berlalu. Tahan hidup akan diketahui hanya untuk item yang gagal dalam selang waktu L tersebut. Jadi, suatu sampel tersensor tipe I adalah sampel yang timbul bila individu 1, 2, ..., n pengamatannya dibatasi oleh periode $L_1, L_2, ..., L_n$ . Tahan hidup individu $T_i$ teramati hanya jika $T_i \leq L_i$ dan untuk semua $L_i$ sama dikatakan data tersensor tipe I secara tungggal.

Dalam penyensoran tipe I, banyaknya individu yang teramati tahan hidupnya merupakan variabel random. Misalkan ada n individu dengan $T_i$ tahan hidup ke-i dikenai waktu penyensoran $L_i$.

Diasumsikan i.i.d dengan fungsi peluang densitas f(t) dan fungsi survivor S(t). Data yang terkumpul dapat disajikan dengan n pasang variabel random $(t_i, \delta_i)$ dengan :

$t_i = \min(T_i, L_i)$ dan

$$\delta_i = \begin{cases} 1, T_i \le L_i \\ 0, T_i > L_i \end{cases} \qquad t_i = \begin{cases} T_i \text{teramati} \\ L_i \text{tersensor} \end{cases}$$

Fungsi peluang bersama dari $t_i$ dan $\delta_i$ adalah :

$$f(t_i, \delta_i) = f(t_i)^{\delta_i} [S(L_i)]^{1-\delta_i}$$

Jika $(t_i, \delta_i)$ independen , maka fungsi likelihoodnya menjadi :

$$L = \prod_{i=1}^{n} [f(t_i)]^{\delta_i} [S(L_i)]^{1-\delta_i}$$

Atau bisa juga dinyatakan dalam persamaan :

$$L = \prod_{i \in D} f_i(t_i) \prod_{i \in} S(L_i)$$

dimana D adalah himpunan individu yang tahan hidupnya terobservasi dan C himpunan individu yang hanya diketahui waktu sensornya.

2. Penyensoran Tipe II

Suatu sampel tersensor tipe II adalah suatu keadaan dimana hanya r pengamatan terkecil dalam suatu sample random yang terdiri dari n item yang teramati $(1 \le r \le n)$. Misalkan n item ditempatkan pada suatu tes dan diputuskan untuk mengakhiri tes setelah diperoleh kegagalan ke-r. Dalam penyensoran ini n ditetapkan terlebih dahulu sebelum data dikumpulkan. Data terdiri dari r tahan hidup terkecil $T_{(1)}, T_{(2)}, \ldots, T_{(n)}$ dari sampel random $T_1, T_2, \ldots, T_n$

Jika $T_i$ i.i.d dengan f(t) dan S(t), maka fungsi peluang bersama dari $T_{(1)}, T_{(2)}, \ldots, T_{(r)}$

$$\frac{n!}{(n-r)!}\prod_{i=1}^{r} f(t_i).S(t_{(r)})^{n-r}$$

Jadi fungsi likelihoodnya adalah :

$$L = \frac{n!}{(n-r)!}\prod_{i=1}^{r} f(t_i).S(t_{(r)})^{n-r}$$

### 2.3.2 Sampel Lengkap

Jika menggunakan tipe sampel lengkap, maka semua komponen yang akan diuji akan dihentikan bila semua komponen yang diuji telah mati.

Sampel lengkap merupakan bentuk khusus dari sampel tersensor tipe II , yaitu jika r=n.

Fungsi peluang bersama untuk sampel lengkap adalah :

$$f(t_i) = n!\, f(t_i)$$

Bukti :

$$f(t_i) = \frac{n!}{(n-n)!} f(t_i).S(t_{(n)})^{n-n}$$

$$= \frac{n!}{0!} f(t_i).S(t_{(n)})^{0}$$

$$f(t_i) = n!\, f(t_i)$$

Fungsi Likelihoodnya adalah sebagai berikut :

$$L = n! \prod_{i=1}^{n} f(t_i)$$

## 2.4 Penduga Maksimum Likelihood

Misalkan $L(\theta) = f(x_1,\ldots,x_n;\theta), \theta \in \Omega$ adalah fungsi peluang bersama dari $x_1, x_2, \ldots, x_n$ Untuk suatu himpunan observasi $(x_1,\ldots,x_n)$, suatu nilai $\theta$ dalam $\Omega$ dimana $L(\theta)$ maksimum disebut *maximum likelihood estimate* (MLE) dari $\theta$. Sehingga $\hat{\theta}$ adalah nilai $\theta$ yang memenuhi

$f(x_1,\ldots,x_2;\theta) = \max_{\theta\in\Omega} f(x_1,\ldots,x_n;\theta)$, (L.j Bain-M.Engelhardt, 1992)

Maka MLE merupakan penyelesaian dari persamaan (persamaan maksimum likelihood)

$$\frac{d}{d\theta}L(\theta) = 0$$

Sebagai catatan, karena setiap nilai $\theta$ yang memaksimumkan $L(\theta)$ juga memaksimumkan log-likelihood, $\ln L(\theta)$ maka bentuk alternative untuk persamaan maksimum likelihood $\frac{d}{d\theta}\ln L(\theta) = 0$ lebih sering digunakan. Sebagai syarat maksimum yang harus dipenuhi adalah $\left.\frac{\partial^2 \ln L(\theta)}{\partial\theta^2}\right|_{\theta} < 0$.

## 2.5 Uji Anderson-Darling

Statistik Anderson-Darling adalah salah satu statistik uji kesesuaian suatu data ketahanan hidup terhadap suatu distribusi tertentu yang dikembangkan oleh Stephens tahun 1974 (Johnson et.al, 1980). Statistik Anderson-Darling didefinisikan oleh persamaan :

$$A^2 = -N - \left[ \frac{1}{N} \sum (2_i - 1) [\ln(w_i) + \ln(1 - w_{N+1-i}] \right]$$

dimana :

N = banyaknya sampel

$w_i$ = cdf dari distribusi yang akan diasumsikan, untuk distribusi Weibull adalah :

$$w_i = F(t) = 1 - \exp(-(^{t_i}/_{\alpha})^{\beta}$$

$t_i$ merupakan data observasi tahan hidup

Uji hipotesisnya adalah sebagai berikut:

$H_0$ = tahan hidup berdistribusi tertentu

$H_1$ = tahan hidup tidak berdistribusi tertentu

Jika nilai Anderson-Darling > nilai kritis maka $H_0$ akan ditolak, yaitu tahan hidup tidak berdistribusi tertentu.

Nilai statistik Anderson-Darling juga bisa digunakan untuk mengetahui tentang kecocokan/kesesuaian model, yaitu apakah model regresi yang dihasilkan sesuai. Uji hipotesis nya adalah sebagai berikut :

$H_0$ = model regresi tepat/sesuai

$H_1$ = model regresi tidat tepat/tidak sesuai

Jika nilai Anderson-Darling hitung > nilai kritis maka $H_0$ akan ditolak, yang berarti model tidak tepat/tidak sesuai.

Tabel 2.1

nilai kritis Anderson-Darling

| Function | Quantile | | | | |
|---|---|---|---|---|---|
| | 0,85 | 0,90 | 0,95 | 0,975 | 0,99 |
| $A^2$ | 1,610 | 1,933 | 2.492 | 3,070 | 3,857 |

"Reprinted,with permission, from Stephens (1974)"

## 2.6 Pemilihan Model Regresi Terbaik

Permasalahan yang sering muncul dalam pembentukan suatu model regresi yaitu apakah model yang diberikan sudah diberikan sudah sesuai. Dalam skripsi ini akan diperkenalkan salah satu cara dalam prosedur pemilihan regresi "terbaik", yaitu prosedur Eliminasi langkah mundur *(The Backward Elimination Procedure).*

Metode ini dimulai dengan regresi terbesar dengan menggunakan semua peubah, dan secara bertahap mengurangi banyaknya variabel prediktor di dalam persamaan sampai mendapatkan persamaan regresi terbaik. Langkah-langkah dalam prosedur eliminasi langkah mundur adalah sebagai berikut :

1. Menghitung persamaan regresi yang mengandung semua variabel prediktor, yaitu variabel x.
2. Dengan melihat nilai probabilitas, buanglah variabel prediktor yang memiliki nilai probabilitas terbesar.
3. Hitung kembali persamaan regresi dengan tidak mengikutsertakan variabel prediktor yang dibuang tersebut.
4. Lakukan terus kegiatan seperti diatas, hingga semua variabel prediktor yang dimasukkan ke dalam model regresi telah signifikan (probabilitas $\leq \alpha$).

# BAB III

# MODEL REGRESI PARAMETRIK

# DATA TAHAN HIDUP

Model regresi parametrik data tahan hidup memodelkan hubungan antara tahan hidup dengan variabel prediktor, dimana populasi tahan hidupnya diasumsikan mengikuti suatu distribusi tertentu dan pada analisis ini akan diasumsikan berdistribusi Weibull. Pembentukan model regresi tahan hidup adalah penentuan suatu model regresi bersyarat x, dengan T menyatakan tahan hidup dan x suatu variabel prediktor.

Bab ini menjelaskan model regresi tahan hidup yang dikenal sebagai model lokasi-skala. Dalam model ini, tahan hidup T ditransformasi dengan transformasi logaritma Y = ln T, sehingga diperoleh persamaan regresi $y = \mu(x) + \sigma\varepsilon$ untuk $\sigma > 0$ dan $\varepsilon$ berdistribusi nilai ekstrim.

## 3.1 Model Regresi Weibull

Misalkan T variabel random tahan hidup berdistribusi Weibull, maka fungsi peluang dari T dengan x sebagai variabel prediktor bisa dinyatakan dalam bentuk:

$$f(t|x)=\frac{\delta}{\alpha(x)}\left(\frac{t}{\alpha(x)}\right)^{\delta-1}\exp\left[-\left(\frac{t}{\alpha(x)}\right)^{\delta}\right] \qquad t\geq 0 \qquad (3.1.1)$$

Dimana parameter skala $\alpha$ diberikan sebagai fungsi dari x.

Fungsi distribusi kumulatif bersyarat x dari T bisa didapatkan dengan menggunakan persamaan (2.1.1), yaitu :

$$F(t|x)=1-\exp\left[-\left(\frac{t}{\alpha(x)}\right)^{\delta}\right] \qquad t\geq o \qquad (3.1.2)$$

Selanjutnya berdasarkan persamaan (2.1.2), fungsi survivor bersyarat x dari T bisa dinyatakan dalan persamaan :

$$S(t|x)=\exp\left[-\left(\frac{t}{\alpha(x)}\right)^{\delta}\right] \qquad t\geq 0 \qquad (3.1.3)$$

Fungsi hazard merupakan fungsi yang menyatakan tingkat kegagalan pada waktu t yang dipengaruhi x. Berdasarkan persamaan (2.1.3), fungsi hazard untuk T bersyarat x bisa dinyatakan dalam persamaan :

$$h(t|x)=\frac{\delta}{\alpha(x)}\left(\frac{t}{\alpha(x)}\right)^{\delta-1} \qquad t\geq 0 \qquad (3.1.4)$$

Model hazard tersebut merupakan model hazard proporsional dikarenakan fungsi hazardnya bisa dinyatakan kedalam bentuk :

$$h(t|x)=h_0(t).g(x)$$

Dimana : $h_0(t) = \delta t^{\delta-1}$

$g(x) = [\alpha(x)]^{\delta}$

Jika tahan hidup T ditransformasi dengan menggunakan transformasi logaritma Y = ln T, maka fungsi peluang dari Y bersyarat x sebagai variabel prediktor akan menghasilkan persamaan baru, yaitu :

$$f(y|x) = \frac{1}{\sigma}\exp\left[\frac{y-x\beta}{\sigma} - \exp\left(\frac{y-x\beta}{\sigma}\right)\right] \quad -\infty < y < \infty \tag{3.1.5}$$

Bukti :

Dengan transformasi logaritma Y = lnT dan $x\beta = \ln\alpha(x)$, $\frac{1}{\delta} = \sigma$ atau T = exp Y, $\alpha(x) = \exp(x\beta)$. Diperoleh Jacobian $|J| = \frac{dt}{dy} = \exp y$

sehingga fungsi peluang Y bersyarat x adalah :

$$f(y|x) = f(e^{y}|x)|J|$$

$$= \frac{1}{\sigma}\exp\left[\frac{y-x\beta}{\sigma}\right]\exp\left[-\exp\left(\frac{y-x\beta}{\sigma}\right)\right]$$

$$= \frac{1}{\sigma}\exp\left[\frac{y-x\beta}{\sigma} - \exp\left(\frac{y-x\beta}{\sigma}\right)\right]$$

Jika didefinisikan error $\varepsilon = \frac{y-x\beta}{\sigma}$, maka model regresi lokasi-skala untuk ln T bisa dinyatakan dengan persamaan linear sebagai berikut :

$$Y = x\beta + \sigma\varepsilon \tag{3.1.6}$$

dimana $x = (x_1,\ldots,x_p)$ sebagai variabel prediktor dan $\beta = (\beta_1,\ldots,\beta_p)$ sebagai vektor parameter regresi.

dalam analisis ini $\varepsilon$ berdistribusi nilai ekstrim dengan fungsi peluang densitas

$$f(\varepsilon|x) = \exp(\varepsilon - \exp(\varepsilon)) \qquad -\infty < \varepsilon < \infty$$

Selanjutnya fungsi distribusi kumulatif Y bersyarat x bisa dinyatakan dalam persamaan :

$$F(y|x) = 1 - \exp\left\{-\exp\left(\frac{y - x\beta}{\sigma}\right)\right\} \tag{3.1.7}$$

Persamaan fungsi survivor untuk log tahan hidup ini bisa dinyatakan dalam bentuk persamaan :

$$S(y|x) = \exp\left\{-\exp\left(\frac{y - x\beta}{\sigma}\right)\right\} \tag{3.1.8}$$

Fungsi hazard atau fungsi yang menyatakan kegagalan dari Y bersyarat x akan dinyatakan dalam persamaan :

$$h(y|x) = \frac{1}{\sigma}\exp\left(\frac{y - x\beta}{\sigma}\right) \tag{3.1.9}$$

## 3.2 Estimasi Parameter

Koefisien-koefisien $\beta$ dalam model regresi diatas yang merupakan parameter-parameter tak diketahui dalam model dapat diketahui dengan cara estimasi menggunakan metode maksimum likelihood. Untuk menjalankan metode ini, pertama kali dicari fungsi likelihood dari data sampel. Fungsi likelihood ini merupakan fungsi peluang bersama dari data observasi yang dianggap sebagai fungsi dari parameter-parameter tak diketahui dalam model yang diasumsikan. Estimasi maksimum likelihood adalah nilai yang memaksimumkan fungsi likelihood. Untuk memudahkan perhitungan seringkali digunakan fungsi loglikelihood. Pendekatan untuk rata-rata dari estimasi diperoleh dari turunan pertama dan variansi dapat diperoleh dari turunan kedua fungsi log-likelihood.

Misalkan $y_i$ adalah log tahan hidup atau log waktu sensoring dan $(y_i, x_i)$ (i = 1, …,n) merupakan observasi independen. Maka persamaan fungsi likelihood dari suatu sampel tersensor tipe I adalah :

$$L(\beta,\sigma) = \prod_{i\varepsilon D} \frac{1}{\sigma} \exp\left[\frac{y_i - x_i\beta}{\sigma} - \exp\left(\frac{y_i - x_i\beta}{\sigma}\right)\right] \prod_{i\varepsilon C} \exp\left[-\exp\left(\frac{y_i - x_i\beta}{\sigma}\right)\right]$$

dengan D dan C berturut-turut menyatakan himpunan individu yang $y_i$-nya adalah log tahan hidup dan log waktu sensoring. Apabila diambil logaritma dari fungsi likelihood, maka akan diperoleh persamaan baru sebagai berikut :

$$\ln L(\beta,\sigma) = -r\ln\sigma + \sum_{i\varepsilon D}\frac{y_i - x_i\beta}{\sigma} - \sum_{i=1}^{n}\exp\left(\frac{y_i - x_i\beta}{\sigma}\right) \tag{3.2.10}$$

dimana r adalah jumlah tahan hidup yang terobservasi.

Kemudian berdasarkan persamaan log-likelihood tersebut akan dicari turunan pertama dan kedua. Misalkan $z_i = \left(\frac{y_i - x_i\beta}{\sigma}\right)$, maka turunan pertama dan kedua dari fungsi log-likelihood tersebut adalah :

$$\frac{\partial \ln L}{\partial \beta_l} = -\frac{1}{\sigma}\sum_{i\varepsilon D} x_{il} + \frac{1}{\sigma}\sum_{i=1}^{n} x_{il}e^{z_i}, l = 1,\ldots,p$$

$$\frac{\partial \ln L}{\partial \sigma} = -\frac{r}{\sigma} - \frac{1}{\sigma}\sum_{i\varepsilon D} z_i + \frac{1}{\sigma}\sum_{i=1}^{n} z_i e^{z_i}$$

$$\frac{\partial^2 \ln L}{\partial \beta_l \partial \beta_s} = -\frac{1}{\sigma^2}\sum_{i=1}^{n} x_{il}x_{is}e^{z_i}, l,s = 1,\ldots,p$$

$$\frac{\partial^2 \ln L}{\partial \sigma^2} = \frac{r}{\sigma^2} + \frac{2}{\sigma^2}\sum_{i\in D} z_i - \frac{2}{\sigma^2}\sum_{i=1}^{n} z_i e^{z_i} - \frac{1}{\sigma^2}\sum_{i=1}^{n} z_i^2 e^{z_i}$$

$$\frac{\partial \ln L}{\partial \beta_l \partial \sigma} = \frac{1}{\sigma^2}\sum_{i\in D} x_{il} - \frac{1}{\sigma^2}\sum_{i=1}^{n} x_{il}e^{z_i} - \frac{1}{\sigma^2}\sum_{i=1}^{n} x_{il}z_i e^{z_i}$$

Untuk data tersensor tipe II, observasi sebanyak r diurutkan yaitu $t_i \le t_2 \le \ldots \le t_r$ dari sampel random berdistribusi Weibull sebanyak n. Atau jika

$y_i = \ln t_i$ maka data terurut menjadi $y_i \leq y_2 \leq .. \leq y_r$ dan fungsi peluang bersama untuk data tersensor tipe II dari log tahan hidup dinyatakan dalam persamaan berikut :

$$L(\beta,\sigma) = \frac{n!}{(n-r)!}\prod_{i=1}^{r} f(y_i|x_i)\left[S(y_{(r)}|x_{(r)})\right]^{-r}$$

$$= \frac{n!}{(n-r)!}\prod_{i=1}^{r}\frac{1}{\sigma}\exp\left[\frac{y_i - x_i\beta}{\sigma} - \exp\left(\frac{y_i - x_i\beta}{\sigma}\right)\right]\left\{\exp\left[-\exp\left(\frac{y_{(r)} - x_{(r)}\beta}{\sigma}\right)\right]\right\}^{n-r}$$

$$= \frac{n!}{(n-r)!}\frac{1}{\sigma^r}\exp\left[\sum_{i=1}^{r}\frac{y_i - x_i\beta}{\sigma} - \sum_{i=1}^{r}\exp\left(\frac{y_i - x_i\beta}{\sigma}\right)\right]\exp\left[-(n-r)\exp\left(\frac{y_{(r)} - x_{(r)}\beta}{\sigma}\right)\right]$$

$$= \frac{n!}{(n-r)!}\frac{1}{\sigma^r}\exp\left\{\sum_{i=1}^{r}\frac{y_i - x_i\beta}{\sigma} - \sum_{i=1}^{r}\exp\left(\frac{y_i - x_i\beta}{\sigma}\right) - (n-r)\exp\left(\frac{y_{(r)} - x_{(r)}\beta}{\sigma}\right)\right\}$$

Salah satu hal yang penting dalam analisis regresi tahan hidup adalah persentil ke-p untuk tahan hidup suatu individu. Estimasi persentil ke-p ($0 < p < 1$) dari logaritma tahan hidup Y dinyatakan dengan persamaan :

$$\hat{y}_p(x) = \hat{\beta}_0 + \hat{\beta}_1 x_1 ... + \hat{\beta}_n x_n + \hat{\sigma}\varepsilon_p \qquad (3.2.2)$$

dimana $\varepsilon_p$ adalah persentil ke-p dari distribusi error. Dalam kasus ini error berdistribusi nilai ekstrim sehingga persentilnya ke-pnya memenuhi :

$$\varepsilon_p = \ln(-\ln(1-p))$$

Untuk memperoleh persentil ke-p dari tahan hidup T dengan mengambil antilog dari $\hat{y}_p$, yaitu :

$$t_p = \exp\left(\hat{y}_p\right) \qquad (3.2.3)$$

# BAB 1V

# ANALISIS DAN PEMBAHASAN

Data berupa tahan hidup bayi hingga usia 1 tahun di Kabupaten Sleman pada tahun 2003 sebanyak 60 bayi. Tujuan dari analisis ini adalah untuk mengetahui model regresi untuk tahan hidup bayi sehingga bisa mengetahui faktor-faktor yang mempengaruhi tahan hidup tersebut. Beberapa faktor yang dianggap mempengaruhi tahan hidup bayi adalah usia kehamilan, jenis kelamin bayi, kemudian pemberian vaksin TT, yang dikelompokkan menjadi tiga yaitu tidak mendapat TT, mendapat TT satu kali dan mendapat TT dua kali. Faktor lain yang dianggap relevan adalah keadaan bayi, yaitu apakah bayi kurus dan lebih kecil dari normal (keadaan bayi kurus dengan berat badan < 2500 gram) atau tidak, dan sebelum meninggal pernah ditolong medis (dokter, bidan,perawat) atau tidak. Tidak ada data yang tersensor, yang berarti menggunakan sampel lengkap.

Akan dibentuk suatu model regresi dengan asumsi berdistribusi Weibull, dimana model tersebut akan menggambarkan hubungan antara tahan hidup bayi dengan beberapa faktor yang telah disebutkan diatas. Dalam hal ini, bertindak sebagai variabel respon (dependen variabel) adalah tahan hidup bayi (T) dan sebagai variabel prediktor adalah $x = (x_1, ..., x_6)$, yang didefinisikan :

$x_1$ = usia kehamilan (minggu)

$x_2$ = 1 untuk jenis kelamin laki-laki dan 0 untuk jenis kelamin perempuan

$x_3$ = 1 untuk mendapat TT 1 kali dan 0 untuk yang lain

$x_4$ = 1 untuk mendapat TT 2 kali dan 0 untuk yang lain

$x_5$ = 1 jika ditolong oleh medis dan 0 jika tidak

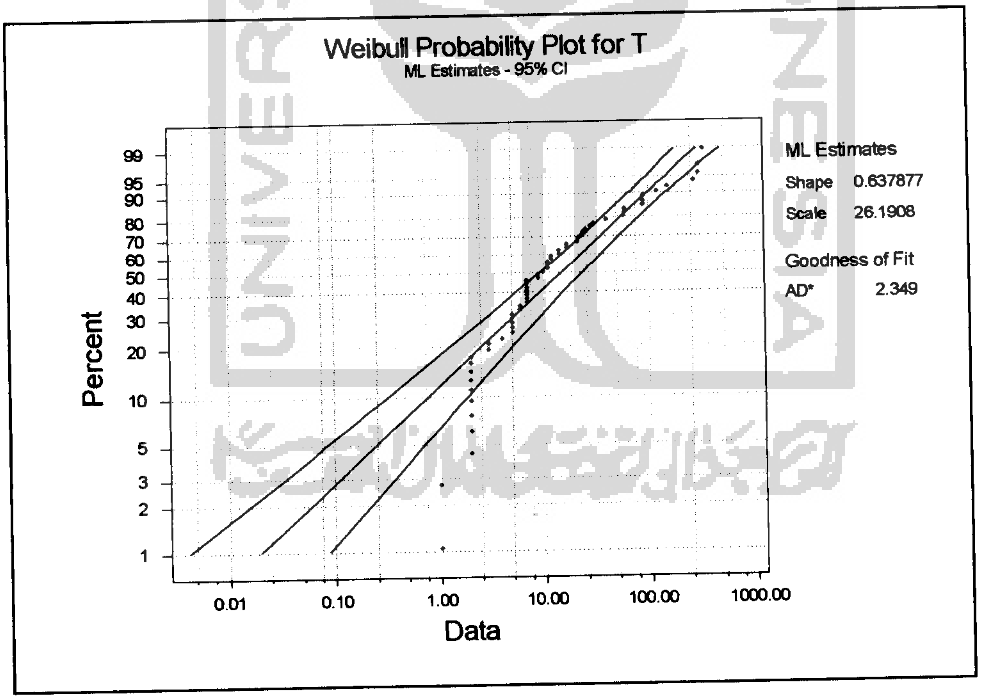
$x_6$ = 1 jika keadaan bayi kurus dan lebih kecil dari normal dan 0 jika tidak

Selanjutnya dilakukan analisis regresi dengan memasukkan semua variabel prediktor, analisis dilakukan dengan bantuan software statistik MINITAB 13 dan diperoleh output seperti dibawah ini :

## 4.1 Analisis Asumsi Distribusi

**grafik 4.1**

**Plot T untuk Distribusi Weibull**



Plot di atas bisa digunakan untuk mengetahui apakah asumsi yang digunakan sesuai atau tidak, yaitu dengan menggunakan nilai statistik Anderson-Darling.

Uji hipotesis untuk analisis ini adalah :

$H_0$ = tahan hidup berdistribusi Weibull

$H_1$ = tahan hidup tidak berdistribusi Weibull

Jika nilai Anderson-Darling > nilai kritis maka $H_0$ akan ditolak, yaitu tahan hidup tidak berdistribusi Weibull.

Dari hasil output diatas didapat bahwa dengan interval konfidensi 95%, nilai statistik Anderson-Darling (2,349) < nilai kritis (2,492), maka $H_0$ diterima yang berarti tahan hidup berdistribusi Weibull.

## 4.2 Analisis Regresi tahan hidup untuk model terbaik

Dalam analisis regresi pada sripsi ini akan digunakan metode eliminasi langkah mundur, yaitu membuang variabel independen yang memiliki probabilitas terbesar sampai semua variabel independen yang dimasukkan kedalam model telah signifikan.

Uji signifikansi konstanta dan koefisien regresi.

Hipotesis untuk analisis adalah :

$H_0$ = koefisien regresi tidak signifikan ( $\beta_0, \beta_1, \ldots, \beta_6 = 0$ )

$H_1$ = koefisien regresi signifikan ( $\beta_0, \beta_1, ..., \beta_6 \neq 0$ )

Dasar pengambilaan keputusan :

a. Dengan membandingkan statistik hitung dengan statistik tabel.

   Jika statistik hitung < Z tabel, maka $H_0$ diterima

   Jika statistik Z hitung $\geq$ Z tabel , maka $H_0$ di tolak

b. Berdasarkan probabilitas.

   Jika probabilitas > 0,05 , maka $H_0$ diterima

   Jika probabilitas $\leq$ 0,05, maka $H_0$ ditolak

c. Berdasarkan interval konfidensi

   Dengan interval konfidensi 95 %, jika interval konfidensi tidak memuat nol, maka variabel tersebut signifikan dalan model.

**Tabel 4.1**

**Deskriptif analisis**

```
Response Variable:  T

Censoring Information                   Count
Uncensored value                           60
Censoring value:  CENSOR = C

Estimation Method:  Maximum Likelihood
Distribution:  Weibull
```

Dari tabel diatas diketahui bahwa :

1. *Cencoring information*; dari 60 data tidak ada yang tersensor, yang berarti semua tahan hidupnya teramati.

2. *Estimation Method*; metode estimasi yang digunakan untuk mengestimasi koefisien regresi adalah metode maksimum likelihood.
3. *Distribution*; tahan hidup diasumsikan berdistribusi Weibull.

Dengan menggunakan prosedur eliminasi langkah mundur, maka analisisnya adalah sebagai berikut :

**Analisis 4.2.1**

**Tabel 4.2**

**Regression with Life Data: T versus x1, x2, x3, x4, x5, x6**

Regression Table

| Predictor | Coef | Standard Error | Z | P | 95.0% Normal CI Lower | Upper |
|---|---|---|---|---|---|---|
| Intercept | -2.030 | 1.782 | -1.14 | 0.255 | -5.522 | 1.463 |
| x1 | 0.14237 | 0.04312 | 3.30 | 0.001 | 0.05786 | 0.22688 |
| x2 | -0.2362 | 0.3475 | -0.68 | 0.497 | -0.9173 | 0.4450 |
| x3 | 0.7657 | 0.5190 | 1.48 | 0.140 | -0.2515 | 1.7829 |
| x4 | 0.8644 | 0.5366 | 1.61 | 0.107 | -0.1874 | 1.9162 |
| x5 | -0.2612 | 0.4816 | -0.54 | 0.588 | -1.2052 | 0.6828 |
| x6 | -0.8498 | 0.4039 | -2.10 | 0.035 | -1.6414 | -0.0582 |
| Shape | 0.84926 | 0.08429 | | | 0.69913 | 1.03162 |

Log-Likelihood = -251.717

Anderson-Darling (adjusted) Goodness-of-Fit

Standardized Residuals = 0.6710

Dengan memasukkan semua variabel independen kedalam model regresi, yaitu variabel $x_1, x_2, x_3, x_4, x_5$ dan $x_6$ diketahui bahwa tidak semua variabel independen siginifikan dan variabel $x_5$ yang memiliki probabilitas terbesar, maka variabel tersebut tidak dimasukkan dalam model regresi berikutnya.

## Analisis 4.2.2

**Tabel 4.3**

**Regression with Life Data: T versus x1, x2, x3, x4, x6**

Regression Table

| Predictor | Coef | Standard Error | Z | P | 95.0% Normal CI Lower | 95.0% Normal CI Upper |
|---|---|---|---|---|---|---|
| Intercept | -2.247 | 1.754 | -1.28 | 0.200 | -5.684 | 1.190 |
| x1 | 0.14327 | 0.04327 | 3.31 | 0.001 | 0.05846 | 0.22808 |
| x2 | -0.2850 | 0.3394 | -0.84 | 0.401 | -0.9503 | 0.3802 |
| x3 | 0.7808 | 0.5216 | 1.50 | 0.134 | -0.2417 | 1.8032 |
| x4 | 0.8282 | 0.5368 | 1.54 | 0.123 | -0.2239 | 1.8802 |
| x6 | -0.8532 | 0.4068 | -2.10 | 0.036 | -1.6506 | -0.0558 |
| Shape | 0.84553 | 0.08357 | | | 0.69662 | 1.02627 |

Log-Likelihood = -251.869

Anderson-Darling (adjusted) Goodness-of-Fit

Standardized Residuals = 0.5875

diketahui bahwa tidak semua variabel independen siginifikan dan variabel $x_2$ yang memiliki probabilitas terbesar, maka variabel tersebut tidak dimasukkan dalam model regresi berikutnya.

## Analisis 4.2.3

**Tabel 4.4**

**Regression with Life Data: T versus x1, x3, x4, x6**

Regression Table

| Predictor | Coef | Standard Error | Z | P | 95.0% Normal CI Lower | 95.0% Normal CI Upper |
|---|---|---|---|---|---|---|
| Intercept | -2.388 | 1.742 | -1.37 | 0.170 | -5.803 | 1.027 |
| x1 | 0.14502 | 0.04354 | 3.33 | 0.001 | 0.05969 | 0.23036 |
| x3 | 0.6683 | 0.5077 | 1.32 | 0.188 | -0.3267 | 1.6634 |
| x4 | 0.6918 | 0.5162 | 1.34 | 0.180 | -0.3200 | 1.7036 |
| x6 | -0.8514 | 0.4067 | -2.09 | 0.036 | -1.6485 | -0.0542 |
| Shape | 0.84177 | 0.08340 | | | 0.69320 | 1.02218 |

Log-Likelihood = -252.230

Anderson-Darling (adjusted) Goodness-of-Fit

Standardized Residuals = 0.6613

diketahui bahwa tidak semua variabel independen siginifikan dan variabel $x_3$ yang memiliki probabilitas terbesar, maka variabel tersebut tidak dimasukkan dalam model regresi berikutnya.

**Analisis 4.2.4**

**Tabel 4.5**

**Regression with Life Data: T versus x1, x4, x6**

Regression Table

| Predictor | Coef | Standard Error | Z | P | 95.0% Normal CI Lower | Upper |
|---|---|---|---|---|---|---|
| Intercept | -1.936 | 1.827 | -1.06 | 0.289 | -5.518 | 1.646 |
| x1 | 0.14807 | 0.04614 | 3.21 | 0.001 | 0.05764 | 0.23850 |
| x4 | 0.1255 | 0.3199 | 0.39 | 0.695 | -0.5014 | 0.7525 |
| x6 | -0.8744 | 0.4334 | -2.02 | 0.044 | -1.7240 | -0.0249 |
| Shape | 0.82478 | 0.08089 | | | 0.68054 | 0.99959 |

Log-Likelihood = -252.985

Anderson-Darling (adjusted) Goodness-of-Fit

Standardized Residuals = 0.7442

diketahui bahwa tidak semua variabel independen siginifikan dan variabel $x_4$ yang memiliki probabilitas terbesar, maka variabel tersebut tidak dimasukkan dalam model regresi berikutnya.

**Analisis 4.2 5**

**Tabel 4.6**

**Regression with Life Data: T versus x1, x6**

Regression Table

| Predictor | Coef | Standard Error | Z | P | 95.0% Normal CI Lower | Upper |
|---|---|---|---|---|---|---|
| Intercept | -1.897 | 1.825 | -1.04 | 0.299 | -5.475 | 1.681 |
| x1 | 0.14831 | 0.04612 | 3.22 | 0.001 | 0.05791 | 0.23871 |
| x6 | -0.8689 | 0.4355 | -2.00 | 0.046 | -1.7224 | -0.0153 |
| Shape | 0.82226 | 0.08030 | | | 0.67902 | 0.99571 |

Log-Likelihood = -253.062

Anderson-Darling (adjusted) Goodness-of-Fit

Standardized Residuals = 0.7922

diketahui bahwa variabel $x_1$ dan $x_6$ telah signifikan dalam model (probabilitas $\leq \alpha$), yang berarti telah didapatkan model terbaik untuk tahan hidup.

Dari hasil analisis diatas maka persaman model regresi untuk tahan hidup bayi hingga usia 1 tahun adalah :

$$\hat{y}_p = -1{,}897 + 0{,}14831x_1 - 0{,}8689x_6 + 1{,}21616\varepsilon_p$$

dimana :

$$\hat{\sigma} = \frac{1}{shape} \quad \text{dan} \quad \varepsilon_p = \ln(-\ln(1-p))$$

## 4.3 Prediksi Tahan Hidup

Dari persamaan regresi tersebut bisa didapatkan estimasi persentil ke-p, yaitu perkiraan waktu tahan hidup bayi hingga usia 1 tahun, dengan persamaan :

$$t_p = \exp\left(\hat{y}_p\right)$$

Sebagai contoh, akan diprediksi tahan hidup pasien dengan kombinasi variabel $x_1$ dan $x_6$ sebagai berikut :

**Tabel 4.7**
**kombinasi variabel $x_1$ dan $x_6$**

| Row number | $X_1$ | $X_6$ | $\hat{t}_p$ | Probabilitasbertahan hidup sampai 40 hari |
|---|---|---|---|---|
| 1 | 32 | 0 | 65.5783 | 0.1360 |
| 2 | 32 | 1 | 27.5050 | 0.0170 |
| 3 | 38 | 0 | 159.6686 | 0.3830 |
| 4 | 38 | 1 | 66.9686 | 0.1407 |

**Analisis 4.3.1**

**Tabel 4.8**
**Tabel Persentil**

```
Table of Percentiles
                          Standard          95.0% Normal CI
 Percent     Percentile    Error          Lower         Upper
      95        65.5783   26.1385        30.0251      143.2305
      95        27.5050    7.1292        16.5494       45.7131
      95       159.6686   35.8524       102.8226      247.9422
      95        66.9686   23.4875        33.6771      133.1707
```

Berdasarkam kolom *percentiles* pada tabel persentil diatas maka bisa dihasilkan prediksi tahan hidup bayi yang berusia hingga 1 tahun, yaitu sebagai berikut :

1. Jika usia kehamilan 32 minggu dan keadaan bayi tidak kurus maupun kecil dari ukuran normal, maka bayi tersebut akan bertahan hidup kurang dari 66 hari. Atau dengan kata lain, jika ada 100 bayi dengan usia kehamilan 32 minggu dan keadaan bayi tidak kurus maupun kecil dari ukuran normal, maka diperkirakan 95 bayi akan mempunyai tahan hidup kurang dari 66 hari dan 5 bayi akan memiliki tahan hidup lebih dari 66 hari.

2. Jika usia kehamilan 32 minggu dan keadaan bayi kurus maupun kecil dari ukuran normal, maka bayi tersebut akan bertahan hidup kurang dari 26 hari. Atau dengan kata lain, jika ada 100 bayi dengan usia kehamilan 32 minggu dan keadaan bayi kurus maupun kecil dari ukuran normal, maka diperkirakan 95 bayi akan mempunyai tahan hidup kurang dari 26 hari dan 5 bayi akan memiliki tahan hidup lebih dari 26 hari.
3. Jika usia kehamilan 38 minggu dan keadaan bayi tidak kurus maupun kecil dari ukuran normal, maka bayi tersebut akan bertahan hidup kurang dari 160 hari. Atau dengan kata lain, jika ada 100 bayi dengan usia kehamilan 38 minggu dan keadaan bayi tidak kurus maupun kecil dari ukuran normal, maka diperkirakan 95 bayi akan mempunyai tahan hidup kurang dari 160 hari dan 5 bayi akan memiliki tahan hidup lebih dari 160 hari.
4. Jika usia kehamilan 38 minggu dan keadaan bayi kurus maupun kecil dari ukuran normal, maka bayi tersebut akan bertahan hidup kurang dari 67 hari. Atau dengan kata lain, jika ada 100 bayi dengan usia kehamilan 38 minggu dan keadaan bayi kurus maupun kecil dari ukuran normal, maka diperkirakan 95 bayi akan mempunyai tahan hidup kurang dari 67 hari dan 5 bayi akan memiliki tahan hidup lebih dari 67 hari.

## Analisis 4.3.2

**Tabel 4.9**

**Tabel probabilitas waktu tahan hidup**

Table of Survival Probabilities

| | | 95.0% Normal CI | |
|---|---|---|---|
| Time | Probability | Lower | Upper |
| 40.0000 | 0.1360 | 0.0231 | 0.3479 |
| 40.0000 | 0.0170 | 0.0018 | 0.0729 |
| 40.0000 | 0.3830 | 0.2492 | 0.5153 |
| 40.0000 | 0.1407 | 0.0336 | 0.3219 |

Dari tabel diatas bisa didapatkan peluang bayi bertahan hidup hingga usia 40 hari, sehingga bisa diketahui perbandingan peluang tahan hidup bayi dengan usia kehamilan ibu 32 minggu (8 bulan )dan 38 minggu (9 bulan 14 hari) serta peluang tahan hidup bayi dengan keadaan bayi kurus dan lebih kecil dari normal atau tidak, yaitu :

1. Peluang bayi dengan usia kehamilan ibu 32 minggu bertahan hidup sampai 40 hari (1 bulan 10 hari) adalah 0,1360 sedangkan bayi dengan usia kehamilan 38 minggu memiliki peluang tahan hidup lebih besar yaitu 0,3830.
2. Perbandingan keadaan bayi (kurus dan kecil dari normal) bisa dilihat dengan membandingkan hasil estimasi fungsi survivor pada usia kehamilan yang sama. Dari "tabel waktu tahan hidup" diatas terlihat bahwa bayi dengan keadaan kurus dan lebih kecil dari normal akan memiliki peluang hidup sampai 40 hari sebesar 0,0170 sedangkan bayi yang lahir dengan

Uji hipotesis nya adalah sebagai berikut :

$H_0$ = model regresi tepat/sesuai

$H_1$ = model regresi tidak tepat/.tidak sesuai

Berdasarkan nilai Anderson-Darling (0,7922) < nilai kritis (2,492) maka dengan interval konfidensi 95% $H_0$ diterima, artinya model regresi dikatakan sesuai untuk tahan hidup.

# BAB V

# KESIMPULAN

## 5.1 Kesimpulan

Kesimpulan yang bisa diambil dari analisis ini adalah :

1. Dari analisis tahan hidup bayi hingga usia 1 tahun dan diasumsikan berdistribusi Weibull, maka dihasilkan model regresinya yaitu sebagai berikut :

   $$\hat{y}_p = -1{,}897 + 0{,}14831x_1 - 0{,}8689x_6 + 1{,}21616\varepsilon_p$$

2. Faktor-faktor yang dianggap mempengaruhi tahan hidup bayi hingga usia 1 tahun adalah usia kehamilan dan keadaan bayi apakah kurus dan ukuran kurang dari normal atau tidak.
3. Bayi yang lahir dengan usia kehamilan ibu lebih lama akan memiliki peluang hidup lebih lama dibandingkan dengan usia kehamilan lebih kecil, dan jika bayi lahir dengan keadaan tidak kurus dan lebih kecil dari normal juga akan memiliki peluang tahan hidup lebih lama dibandingkan yang lahir dengan keadaan kurus dan lebih kecil dari normal.

## DAFTAR PUSTAKA


**Ani Hendarti.** 2003. *Model Regresi Parametrik Data Tahan Hidup.* Fakultas Matematika dan Ilmu Pengetahuan Alam. Universitas Gadjah Mada

**Bain, L, J and M. Engelhardt**. 1992. *Introduction to Probability and Mathematical Statistic*. Second Edition. Duxbury Press. California

**Lawless, J. F**. 1982. *Statistical Models and Methods for Lifetime Data.* John Wiley & Sons. Canada

**Draper, Norman and Smith, Harry**. 1992. *Analisis Regresi Terapan.* PT. Gramedia Pustaka Utama Jakarta

**Ronald E Walpole and Raymond H Myers**. 1995. *Ilmu Peluang dan Statistika Untuk Insinyur dan Ilmuwan.* ITB. Bandung.

http://rac.alionscience.com. 2004. *Anderson Darling : Agoodness Of Fit Test for Small Samples Assumptions*. RAC. USA

http://www.itlnist.gov/div8981handbook/eda/section3/eda35e.htm. 2004. *Anderson-Darling test*

# LAMPIRAN

| No | T | $x_1$ | $x_2$ | $x_3$ | $x_4$ | $x_5$ |
|---|---|---|---|---|---|---|
| 1 | 2 | 34 | LK | 1 | RS | YA |
| 2 | 40 | 40 | LK | 1,2 | RS | TDK |
| 3 | 25 | 39 | LK | 1,2 | BIDAN | TDK |
| 4 | 11 | 38 | LK | - | DOKTER | TDK |
| 5 | 1 | 36 | PR | 1,2 | RS | YA |
| 6 | 3 | 28 | LK | 2 | DOKTER | YA |
| 7 | 22 | 28 | LK | 1,2 | BIDAN | YA |
| 8 | 2 | 28 | LK | 1,2 | RS | YA |
| 9 | 14 | 34 | PR | 2 | PUSKES | YA |
| 10 | 58 | 40 | LK | 2 | BIDAN | TDK |
| 11 | 7 | 28 | PR | 1,2 | BIDAN | YA |
| 12 | 2 | 28 | LK | 1,2 | - | TIDAK |
| 13 | 5 | 40 | LK | 1,2 | RS | TDK |
| 14 | 12 | 36 | LK | 2 | RS | YA |
| 15 | 11 | 41 | LK | 1,2 | DOKTER | TDK |
| 16 | 24 | 40 | PR | 1,2 | BIDAN | TDK |
| 17 | 11 | 41 | LK | 1,2 | DOKTER | TDK |
| 18 | 150 | 40 | LK | 2 | RS | TDK |
| 19 | 3 | 38 | LK | 2 | BIDAN | TDK |
| 20 | 300 | 42 | PR | 1,2 | RS | TDK |
| 21 | 5 | 42 | PR | 1 | DOKTER | YA |
| 22 | 17 | 39 | PR | 1 | BIDAN | YA |
| 23 | 5 | 37 | LK | 1,2 | BIDAN | TDK |
| 24 | 30 | 40 | LK | 2 | RS | TDK |
| 25 | 9 | 32 | LK | 2 | RS | YA |
| 26 | 300 | 39 | PR | 2 | - | TDK |

| | | | | | | |
|---|---|---|---|---|---|---|
| 27 | 14 | 37 | LK | 1,2 | BIDAN | TDK |
| 28 | 17 | 37 | PR | - | - | YA |
| 29 | 330 | 40 | LK | 1,2 | RS | TDK |
| 30 | 60 | 39 | LK | 1,2 | RS | YA |
| 31 | 12 | 40 | LK | - | RS | YA |
| 32 | 90 | 40 | PR | 2 | RS | TDK |
| 33 | 21 | 35 | LK | 1,2 | RS | YA |
| 34 | 5 | 38 | LK | 2 | - | YA |
| 35 | 2 | 39 | LK | 2 | RS | TDK |
| 36 | 1 | 33 | LK | 1,2 | BIDAN | YA |
| 37 | 6 | 34 | LK | 1 | RS | TDK |
| 38 | 60 | 40 | LK | 1,2 | BIDAN | TDK |
| 39 | 2 | 32 | PR | 1,2 | - | YA |
| 40 | 7 | 36 | PR | 2 | BIDAN | YA |
| 41 | 7 | 28 | LK | 2 | BIDAN | YA |
| 42 | 6 | 30 | PR | - | RS | TDK |
| 43 | 7 | 30 | PR | - | RS | TDK |
| 44 | 2 | 38 | LK | 2 | PUSKES | TDK |
| 45 | 9 | 40 | LK | 1,2 | BIDAN | YA |
| 46 | 270 | 39 | LK | 2 | RS | TDK |
| 47 | 7 | 40 | PR | 2 | BIDAN | TDK |
| 48 | 28 | 40 | LK | 1,2 | DOKTER | TDK |
| 49 | 10 | 43 | LK | - | RS | TDK |
| 50 | 4 | 34 | PR | 2 | RS | TDK |
| 51 | 23 | 36 | PR | 2 | DOKTER | YA |
| 52 | 2 | 39 | LK | 2 | BIDAN | TDK |
| 53 | 120 | 40 | LK | 2 | - | TDK |

| | | | | | | |
|---|---|---|---|---|---|---|
| 54 | 90 | 40 | LK | 2 | RS | TDK |
| 55 | 90 | 39 | PR | 1,2 | RS | TDK |
| 56 | 7 | 28 | LK | 2 | RS | YA |
| 57 | 7 | 32 | PR | - | BIDAN | YA |
| 58 | 2 | 32 | PR | 1 | - | YA |
| 59 | 5 | 40 | LK | 1,2 | RS | TDK |
| 60 | 2 | 33 | LK | 2 | BIDAN | YA |

*Tabel data kematian bayi di Kabupaten Sleman Yogyakarta Pada tahun 2003.*

*Sumber. Dinas kesehatan Kabupaten Sleman Yogyakarta*

T = usia kematian bayi (hari)

$x_1$ = usia kehamilan (minggu)

$x_2$ = Jenis kelamin bayi

$x_3$ = Banyak suntukan TT yang dilakukan ketika ibu hamil

$x_4$ = Sebelum meninggal apakah pernah ditangani tim medis (Dokter, Bidan ,Perawat)

$x_5$ = keadaan bayi kurus dan kecil dari normal (berat < 2500 mg) atau tidak