ChiSq　0.1968＝Pr{$\chi^2(1)>1.6528$} さらにハザード比 Hazard Ratio 0.632＝exp(−0.45889)/exp(0) が出力されている．なお Chi-Square の値が上の段では Wald 検定値の名前で示されている．

次に共変量に X を追加した結果が表 3.2 である．今回の検定の帰無仮説は「X の係数も TREAT の係数もともに 0」である．3 つの検定ともに帰無仮説を棄却する結果となっている．最下段に出力されている TREAT の回帰係数の推定値は −1.105 (±0.4077) であるから有意であり (p=0.0067)，治療はハザードを約 67%＝1−0.331 減少させる効果のあったことが結論される．なお薬効のハザード比の (Wald による) 95% 信頼区間は，

exp(−1.105−1.96×0.4077)＝0.1490，exp(−1.105＋1.96×0.4077)＝0.7365

により (0.1490，0.7365) と計算される．表 3.2 の結果も前章の層別ログランク検定とよく一致しているが，実は層別ログランク検定と正確に同じ計算をするのは次章で扱う層別 Cox 回帰のスコアー検定である．表 3.1 と表 3.2 での検定結果の著しい違いは冒頭のパラグラフで述べたことの重要性を示している．図 3.1 は推定生存率曲線である．生存率は共変量の値に依存するので，共変量の値を具体的に指定する必要がある．

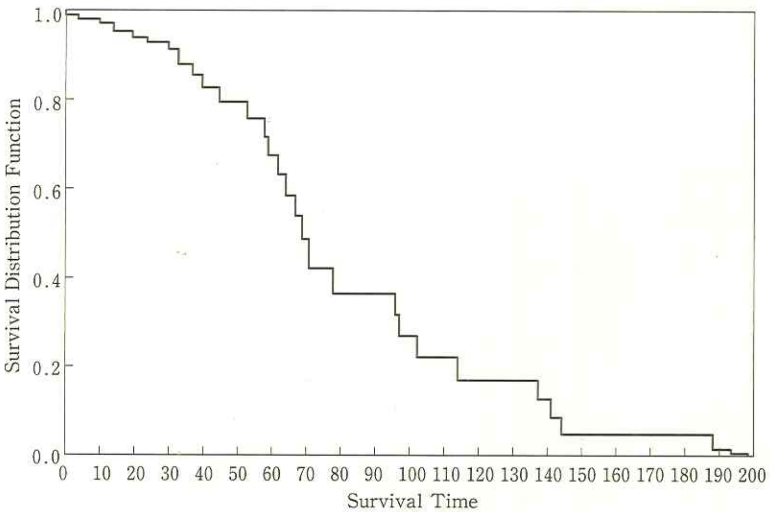

図 3.1　推定生存率曲線 (Treat=1，X=2)

以下の節において，共変量の効果の推定と検定のための Cox モデルの計算原理，およびその計算法の基礎をなす比例ハザード性と部分尤度 (partial likelihood) の定義ならびにその直感的理解，変数選択の意味と利用法，陥りやすい落とし穴，そして競合リスク要因 (competing risk) を指定したときの解釈を順次扱う．

3.2　比例ハザードモデル

Cox モデル解析法とは，比例ハザード性 (proportional hazards) を仮定して部分尤度 (partial likelihood) を用いてデータを解析する方法のことを指す．この節では関連用語の定義をしその意味を解説する．

2 つのハザード関数 $\lambda_1(t)$, $\lambda_2(t)$ の間に関係式

$$\lambda_1(t)=c\lambda_2(t) \tag{3.1}$$

がすべての可能な $t>0$ で成立するとき，2 つのハザード関数は比例するという．ただし，c は経過時間 t に依存しない定数である．

個体の生存時間に影響を与える因子を背景因子，予後因子 (prognostic factor) あるいは共変量などと呼ぶ．共変量は一般に複数あるので，共変量のベクトル $\boldsymbol{z}=(z_1, z_2, \cdots, z_m)$ を考える．共変量 $\boldsymbol{z}$ をもつ個体のハザードを $\lambda(t|\boldsymbol{z})$ と書くことにする．$\boldsymbol{z}$ の値は時間 t には依存しない定数とする．もしあるハザード関数 $\lambda_0(t)$ と $\boldsymbol{z}$ の関数 $r(\boldsymbol{z})$ が存在して，すべての $\boldsymbol{z}$ について，等式

$$\lambda(t|\boldsymbol{z})=\lambda_0(t)r(\boldsymbol{z}),\quad \text{for all } t>0 \tag{3.2}$$

が成立するとき，共変量 $\boldsymbol{z}$ の効果は比例ハザードモデルに従うという．$\lambda_0(t)$ はベースラインハザード (baseline hazard)，$r(\boldsymbol{z})$ は相対危険度 (relative risk) 関数と呼ばれる．比例ハザードモデルのもとでは，2 つの共変量 $\boldsymbol{z}$ と $\boldsymbol{z}'$ について

$$\lambda(t|\boldsymbol{z}')=\lambda(t|\boldsymbol{z})\left(\frac{r(\boldsymbol{z}')}{r(\boldsymbol{z})}\right),\quad \text{for all } t>0 \tag{3.3}$$

となるので，ハザード関数 $\lambda(t|\boldsymbol{z})$ と $\lambda(t|\boldsymbol{z}')$ は比例する．比例定数 $r(\boldsymbol{z}')/r(\boldsymbol{z})$ は $\boldsymbol{z}'$ の $\boldsymbol{z}$ に対する相対危険度 (relative risk of $\boldsymbol{z}'$ to $\boldsymbol{z}$) と呼ばれる．

式 (3.3) の対数をとった形

$$\log \lambda(t|\boldsymbol{z}')=\log \lambda(t|\boldsymbol{z})+\log\left(\frac{r(\boldsymbol{z}')}{r(\boldsymbol{z})}\right),\quad \text{for all } t>0 \tag{3.4}$$

で $r(\boldsymbol{z})$ に対数線形性 (log-linear model)

$$\log r(\boldsymbol{z})=\beta_1 z_1+\beta_2 z_2+\cdots+\beta_m z_m=\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{z} \tag{3.5}$$

を仮定すると，式 (3.4) は

$$\begin{aligned}\log\left(\frac{\lambda(t|\boldsymbol{z}')}{\lambda(t|\boldsymbol{z})}\right)&=\log\left(\frac{r(\boldsymbol{z}')}{r(\boldsymbol{z})}\right)\\&=\boldsymbol{\beta}^{\mathrm{T}}(\boldsymbol{z}'-\boldsymbol{z})\\&=\beta_1(z_1'-z_1)+\beta_2(z_2'-z_2)+\cdots+\beta_m(z_m'-z_m)\end{aligned}$$

となる．ベクトル $\boldsymbol{\beta}^{\mathrm{T}}=(\beta_1, \beta_2, \cdots, \beta_m)$ は回帰係数と呼ばれ，通常部分尤度法により推定される．たとえ $\boldsymbol{z}\neq\boldsymbol{z}'$ でも $\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{z}=\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{z}'$ ならばハザードは同じになる．$\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{z}$ は $\boldsymbol{z}=\boldsymbol{0}$ に対する相対ハザード $\lambda(t|\boldsymbol{z})/\lambda(t|\boldsymbol{0})$ の対数なので，対数相対ハザード (log relative hazard) あるいは対数相対危険度と呼ばれ，医学分野では予後指数 (prognostic index) とも呼ばれる[*1]．式 (3.5) を式 (3.2) に代入して，

$$\lambda(t|\boldsymbol{z})=\lambda_0(t)\exp(\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{z}) \tag{3.6}$$

となる．この式が一般に比例ハザードモデルとして用いられるが，実は本来の比例ハザード性 (式 (3.2)) 以外に対数線形性 (式 (3.5)) の仮定も含んでいる．

図 3.2 をみると，年度と性の効果，ならびに年齢の効果も近似的に比例ハザードモデルに従っていることがわかる．このほか，多くの薬効，毒物の効果も近似的に比例ハザードモデルに従うことが確認されている．

式 (3.2) を満たすハザード関数 $\lambda(t|\boldsymbol{z})$，$\lambda_0(t)$ の累積ハザード

$$\Lambda(t|\boldsymbol{z})=\int_0^t \lambda(u|\boldsymbol{z})du,\quad \Lambda_0(t)=\int_0^t \lambda_0(u)du$$

を，ハザード関数から生存時間関数を求める公式 (1.5) に代入して

$$S(t|\boldsymbol{z})=\exp\{-\Lambda(t|\boldsymbol{z})\},\quad S_0(t)=\exp\{-\Lambda_0(t)\}$$

となる．さらに $\Lambda(t|\boldsymbol{z})=\Lambda_0(t)r(\boldsymbol{z})$ を代入して，

$$S(t|\boldsymbol{z})=S_0(t)^{r(\boldsymbol{z})} \tag{3.7}$$

となる．両辺の対数をとると，

$$\log S(t|\boldsymbol{z})=r(\boldsymbol{z})\log S_0(t)$$

*1) $\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{z}$ を単に対数ハザードと呼ぶのは正確ではない．対数ハザードは $\log\lambda_0(t)+\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{z}$ である．

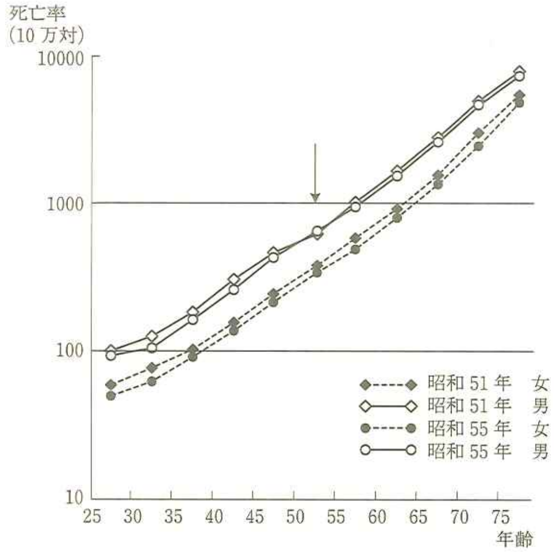

図 3.2　昭和 51 年および 55 年における日本人の死亡率
(厚生省人口動態調査より抜粋)
4 つの死亡率曲線は 35 歳以上でほぼ直線で年齢によらず一定の値だけ異なる
(矢印の年齢階級は昭和一桁世代男性の異常な死亡率による例外である).

となるが，これは負の数なので両辺に $-$ をかけてさらに対数をとると，

$$\log\{-\log S(t|\boldsymbol{z})\}=\log\{-\log S_0(t)\}+\log r(\boldsymbol{z}) \tag{3.8}$$

となる．累積ハザードを用いて表現することにより応用上重要な関係式

$$\log \Lambda(t|\boldsymbol{z})=\log \Lambda_0(t)+\log r(\boldsymbol{z})$$

を得る．さらに対数線形性 (式 (3.5)) が成立するときは，

$$\log\{-\log S(t|\boldsymbol{z})\}=\log\{-\log S_0(t)\}+\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{z} \tag{3.9}$$

となる．最後の 2 つの式は次章で Cox モデルの適合度を確認する際に用いられる．

式 (3.7) は比例ハザードモデルの別表現であるが，比例ハザードモデルが Lehmann 対立仮説 (Lehmann alternative) を満たしていることを示している．Lehmann 対立仮説とは，対立仮説の分布 G が帰無仮説の分布 F の関数 $G=h(F)$ で表される場合をいう．Lehmann 対立仮説のもとでは，観測値の順位 (rank) を用いた検定統計量の対立仮説での分布が，h のみに依存して決まり，

F, G とは無関係になるという性質がある (Razzaghi et al., 1998). いいかえると，Lehmann 対立仮説のもとでは，順位を用いた検定統計量がノンパラメトリック検定統計量となる．これは比例ハザードモデルの理論的メリットである．特にセンサー標本では観測値そのものの利用が困難で観測値の順位に頼らざるを得ない状況が一般的なので，順位の分布がもとの分布 F, G に依存しなくなる比例ハザードモデルの効用は大きい．比例ハザード性と順位は深い関係にある．ログランク検定もこれから述べる Cox 回帰法も死亡順位しか用いていない．また 6 章で述べる周辺尤度法は比例ハザードモデルのもとでの順位統計量の分布を直接計算により求めている．

2 章でログランク検定は 2 群間の違いが比例ハザードモデルに従うときは，死亡順位に重みをつけて加えた統計量 (線形ランク統計量) のなかで相対的に最大の検出力 (locally most powerful) をもつことを述べた．このことは，ログランク検定以外の検定の検出力が一般に低いことを意味しない．実際，たとえ比例ハザード性を満たさない 2 群でもそれらの累積ハザード Λ が，任意の t について $\Lambda(t|\boldsymbol{z}=1) \geqq \Lambda(t|\boldsymbol{z}=0)$ を満たし，かつ $\Lambda(t|\boldsymbol{z}=1) > \Lambda(t|\boldsymbol{z}=0)$ がある区間で成立するならば，順位に重みをつけて加えた統計量のほとんどは標本数を大きくすれば検出力は 1 になる．式 (1.5) を用いていいかえると，一方の生存率曲線が常に他方より上にあり完全に一致することがない (すなわち交わらない) ならば，たとえ比例ハザード性が成り立たないときでも，線形ランク統計量 (linear rank statistics) を用いる検定は標本数が大きければ有意な結果になることを示唆している．

3.3　回帰係数推定のための部分尤度法

比例ハザードモデルでの回帰係数は通常 Cox により提唱された部分尤度法により推定される．このため比例ハザードモデルに部分尤度法を適用することを一般に Cox 回帰法と呼ぶ．Cox 回帰法による解析結果を正しく解釈するには部分尤度を理解する必要がある．部分尤度はログランク検定の考え方に多変量ロジスティックモデルの技法を適用して多変量にしただけの簡単なものであるが，この節では数式を用いてその原理を解説する．なお部分尤度に対応する

全尤度と両者の関係については 6.3 節で解説する．

観察された死亡数を D，観察された死亡時間を $t_1, \cdots, t_i, \cdots, t_D$ とする．死亡時間はすべて異なるとする ($i \neq j$ ならば $t_i \neq t_j$)．部分尤度 L は，

$$L = L_1 \times \cdots \times L_i \times \cdots \times L_D$$

ただし L_i は，t_i に死亡した症例の番号を (i) と書き，t_i の直前まで (死亡も脱落もせず) 観察されていた症例の集合を R_i とすると，

$$L_i = \frac{\lambda(t_i|\boldsymbol{z}_{(i)})}{\sum_{j \in R_i} \lambda(t_i|\boldsymbol{z}_j)}$$

と書かれる．いいかえると，

L_i＝死亡症例のハザード/生存を確認されていた症例のハザードの総和

となる．R_i は t_i でのアットリスク (at risk) またはリスクセット (risk set) と呼ばれる．さて，$\varDelta$ を小さい正の数とすると，

$$\mathrm{P_r}\{j \text{ は } t_i + \varDelta \text{ までに死ぬ} \mid j \text{ は } t_i \text{ の直前に生きている}\} \fallingdotseq \lambda(t_i|\boldsymbol{z}_j)\varDelta$$

であるから，

$$\begin{aligned} L_i &= \frac{\lambda(t_i|\boldsymbol{z}_{(i)})\varDelta}{\sum_{j \in R_i} \lambda(t_i|\boldsymbol{z}_j)\varDelta} \\ &\fallingdotseq \mathrm{P_r}\{(i) \text{ が死亡} \mid R_i \text{ の内の 1 人だけがこれから } \varDelta \text{ の間に死ぬ}\} \end{aligned}$$

となる[*1]．死亡が発生した時点ごとに，その時点に生存している症例が与えられたとして，死亡確率を計算する考え方はログランク検定と共通している[*2]．ここで比例ハザードの仮定 (式 (3.2)) を用いると，

$$\begin{aligned} L_i &= \frac{\lambda(t_i|\boldsymbol{z}_{(i)})}{\sum_{j \in R_i} \lambda(t_i|\boldsymbol{z}_j)} \\ &= \frac{r(\boldsymbol{z}_{(i)})}{\sum_{j \in R_i} r(\boldsymbol{z}_j)} \end{aligned} \tag{3.10}$$

となり (λ_0 は分母分子で相殺) さらに対数線形性の仮定 (式 (3.5)) を用いると，

$$L_i = \frac{\exp(\boldsymbol{\beta}^{\mathrm{T}} \boldsymbol{z}_{(i)})}{\sum_{j \in R_i} \exp(\boldsymbol{\beta}^{\mathrm{T}} \boldsymbol{z}_j)}, \quad j \in R_i \tag{3.11}$$

---

*1) リスクセット R のうちの特定の個体 j だけが死ぬ確率は正確にいうと，

$$\lambda(t|\boldsymbol{z}_j)\varDelta \prod_{k \in R-j} \{1 - \lambda(t|\boldsymbol{z}_k)\varDelta\}$$

となるが，積の項はほとんど 1 で分母分子で相殺されるので，通常は本文のように書かれる．

*2) 部分尤度の正当性ならびに有効性は 1980 年以後になって，マーティンゲール理論を用いて確立された (Andersen *et al.*, 1982, 1993) が，基本のアイディアは Mantel により疫学研究のための方法として開発されたものである．

となる．これは t とも λ_0 とも無関係である．この積をとり部分尤度

$$L(\boldsymbol{\beta})=\prod_i\left\{\frac{\exp(\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{z}_{(i)})}{\sum_{j\in R_i}\exp(\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{z}_j)}\right\},\quad i=1,\cdots,D \tag{3.12}$$

を得る．これは通常 Cox の部分尤度と呼ばれる．対数部分尤度は，

$$l(\boldsymbol{\beta})=\log L(\boldsymbol{\beta})=\sum_i\{\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{z}_{(i)}-\log\sum_{j\in R_i}\exp(\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{z}_j)\} \tag{3.13}$$

ただし，$\sum_i$ は $i\in D$ の和を示す．

ベースラインハザードとセンサー例が尤度に残っていないのがポイントである．センサー例は分母にのみ寄与する．また，実際の死亡時間ではなく，死亡時間の順位 (rank) にしか依存しないのが特徴である．

対数尤度を回帰係数で微分してスコアー関数 $U(\beta)$ (ベクトル) を得る，

$$U(\boldsymbol{\beta})=\frac{\partial l(\boldsymbol{\beta})}{\partial\boldsymbol{\beta}}=\sum_i\left(\boldsymbol{z}_{(i)}-\frac{\sum_{j\in R_i}\boldsymbol{z}_j\exp(\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{z}_j)}{\sum_{j\in R_i}\exp(\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{z}_j)}\right) \tag{3.14}$$

もう 1 回微分して $-$ をつけることにより，情報量関数 (行列) を得る，

$$\begin{aligned}I(\boldsymbol{\beta})&=-\frac{\partial U}{\partial\boldsymbol{\beta}}\\&=\sum_i\left[\frac{\sum_{j\in R_i}\boldsymbol{z}_j\boldsymbol{z}_j^{\mathrm{T}}\exp(\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{z}_j)}{\sum_{j\in R_i}\exp(\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{z}_j)}-\frac{\{\sum_{j\in R_i}\boldsymbol{z}_j\exp(\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{z}_j)\}\{\sum_{j\in R_i}\boldsymbol{z}_j\exp(\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{z}_j)\}^{\mathrm{T}}}{\{\sum_{j\in R_i}\exp(\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{z}_j)\}^2}\right]\end{aligned} \tag{3.15}$$

U と I の意味を理解するために，$w_j=\exp(\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{z}_j)$，$W=\sum_j w_j$，$E=\sum_j w_j\boldsymbol{z}_j/W$ とおき (簡単のために i は省略されている) 書きかえると，

$$U(\boldsymbol{\beta})=\sum_i\{\boldsymbol{z}_{(i)}-E_i\},\quad I(\boldsymbol{\beta})=\sum_i\frac{\sum_{j\in R_i}w_j\boldsymbol{z}_j\boldsymbol{z}_j^{\mathrm{T}}}{W_i}-E_iE_i^{\mathrm{T}}$$

となる．最大部分尤度法では $U=\mathbf{0}$ になるような $\boldsymbol{\beta}$ を求める．w_j はハザードの大きさに比例しているので，死亡症例における共変量 $\boldsymbol{z}_{(i)}$ の和が，ハザードで重みづけされた共変量の平均値 E_i の和と等しくなるような $\boldsymbol{\beta}$ を求めていることになる．観察情報量 (observed information) $I(\boldsymbol{\beta})$ があらゆる $\boldsymbol{\beta}$ の値に対して正値対称行列であることは，それが通常の重み付き分散行列の形をしていることからもわかる．微積分の言葉でいうならば，対数尤度は全空間 (独立変数は $\boldsymbol{\beta}$) で凸であり，ユニークな最大値が存在する．このような状況での最大値の発見には Newton-Raphson 法が有効である．今得ている $\boldsymbol{\beta}$ が $U(\boldsymbol{\beta})=\mathbf{0}$ を満たさないときはよりよい推定値として，$\boldsymbol{\beta}+I(\boldsymbol{\beta})^{-1}U(\boldsymbol{\beta})$ を用いる，通常

は 5 回以内の繰り返しで最尤推定値 (MLE) $\boldsymbol{\beta}^*$ に達する．

特定の係数 β_k についての仮説 H_0：$\beta_k=0$ の検定は，β_k^* の推定標準誤差 $\{I(\boldsymbol{\beta}^*)_{kk}^{-1}\}^{1/2}$ を用いて，$\beta_k^*/\{I(\boldsymbol{\beta}^*)_{kk}^{-1}\}^{1/2}$ の絶対値が 1.96 以上のときに 5%有意水準で $\beta_k\neq0$ とされる．これが Wald 検定である．対数尤度 $l(\boldsymbol{\beta}^*)$ を用いて尤度比検定を行うには，$\boldsymbol{\beta}$ から β_k を除いて次元が 1 つ減ったベクトル $\boldsymbol{\gamma}$ での対数尤度 $l(\boldsymbol{\gamma}^*)$ を求め，$X^2=2\{l(\boldsymbol{\beta}^*)-l(\boldsymbol{\gamma}^*)\}$ が H_0 のもとで自由度 1 の χ^2 分布に従うことを用いて，X^2 が 3.84 以上のときは 5%有意水準で $\beta_k\neq0$ とする．k 個 ($k>1$) の変数の回帰係数が同時に 0 かどうかの検定は尤度比検定が原理的に簡単である．上で述べた β_k の検定のときと同様に，$\boldsymbol{\gamma}$ を k 次元落ちた共変量の係数として対数尤度を求め，$X^2=2\{l(\boldsymbol{\beta}^*)-l(\boldsymbol{\gamma}^*)\}$ が H_0 のもとで自由度 k の χ^2 分布に従うことを用いて検定する．H_0：$\boldsymbol{\beta}=0$ の検定 (すべての回帰係数が 0，すなわちモデル全体の検定) は，スコアー検定統計量 $U(0)I(0)^{-1/2}$ が帰無仮説が真のときに漸近的に自由度 m の χ^2 分布に従うことを用いて行える．統計ソフトの出力には以上の統計量が出力されるのが普通である．一方特定の β_k についてのスコアー検定と，複数個の回帰係数が同時に 0 かどうかという複合仮説を検定するための Wald 検定についてはさらに計算を要することと，比較的に精度の落ちることもあるので，特別な場合を除いて出力されない．

3.4　生存率曲線

回帰係数 $\boldsymbol{\beta}$ の推定値 $\boldsymbol{\beta}^*$ を得たならば，次にベースラインハザード $\lambda_0(t)$ の推定値が得られる．$\lambda_0(t)$ の推定法にはいくつかあるが，いずれも $\boldsymbol{\beta}^*$ を真値として扱い，得られる生存率曲線は KM 曲線と同じく，観察死亡時 $t_1,\cdots,t_i,\cdots,t_D$ で値が変わる．KM 曲線では，$q_i=1-d_i/n_i$ を t_i の直前の生存率にかけて t_i の直後の生存率を求めた．個体のハザードが異なるときはハザード値を用いて，

$$q_i=\left(1-\frac{w_{(i)}}{W_i}\right)^{\frac{1}{w_{(i)}}},\quad \text{for } i=1,\cdots,D \tag{3.16}$$

$$\hat{S}_0(t)=\hat{S}_0(t_i)q_i,\quad \text{for } t_i<t\leqq t_{i+1} \tag{3.17}$$