



信用リスク計測に関する一考察

Implied Data にもとづく統計的推論と格付けごとの倒産確率の推定と推定誤差の計測方法

高橋 一

信用リスク計測に関する一考察

Implied Data にもとづく統計的推論と 格付けごとの倒産確率の推定と推定誤差の計測方法

2009年6月20日

2009年9月15日改訂

2009年9月30日改訂

一橋大学大学院経済学研究科

CFEE

高橋 一

本稿は日本商品先物振興協会2004年度研究助成及びCFEEからの援助により遂行された。同協会とCFEEのご厚意に深く感謝致します。

始めに

ある企業が倒産する確率はその企業の信用度により異なる。逆に信用度はその確率により決まるといっても良い。例えばムーディーズは企業の信用度を **Aaa, Aa, A, Baa, Ba, B, Caa** の7グループに分類、倒産状態を入れると全体では8のクラスに分けている。1970年から2003年の間でそれらの企業が6年以内に倒産する確率は **0.12% (Aaa), 0.24% (Aa), 0.54% (A), 2.16% (Baa), 11.17% (Ba), 31.9% (B), 60.36% (CCC)** となっている。ムーディーズがどのような方法でヒストリカルなデータより倒産確率を算出しているか、少なくとも筆者には不明であるが本論文では各クラス毎の倒産確率の推定方法について考察する。又、本稿では現実の確率の元での倒産確率と所謂マルチンゲール測度に於ける倒産確率の双方での推定法を提示・提案し、更に推定精度を求める。マルチンゲール測度のもとでの倒産確率は派生証券価格決定の際に必要とされる、一方現実の確率のもとでの倒産確率の必要性は言うまでもない。両者はリスクの市場価格を推定することにより統合可能ではあるが、本論文ではその部分には立ち入らないこととする。何れのケースでも推定精度を与える事によりリスクを加味した価格設定、ポートフォリオ評価等が可能となる。

本論文は、第1部”既存理論のレビュー”。新たな推定法の提案は”第2部” **Implied 倒産確率の推定と統計学的考察**、そして第3部” **RealProbability** のもとでの倒産確率の推定”より構成されている。第1部では、現代ファイナンス理論の標準理論というべき **Harrison-Pliska** による派生証券価格決定の基本定理とそれに基づくゼロクーポン債の価格決定論。又、過去30年の間、主に生物・医学統計の分野で著しく発展してきた生存時間解析についての簡単な解説を行う。第2部では **Implied 倒産確率**をデータと見立てた統計解析を、第3部では **Bootstrap** 方を中心に議論を進めて行く。

第1部 既存理論のレビュー

以下、 $(\Omega, \mathcal{F}, Pr)$ を基本となる確率空間とする。測度空間 (Ω, \mathcal{F}) はブラウン運動が定義できるだけの豊かさを持ち、 \mathcal{F} は Pr に関し完備であると仮定する。 $\{W(t), t \geq 0\}$ を $(\Omega, \mathcal{F}, Pr)$ で定義される標準ブラウン運動、 $\{\mathcal{F}_t = \sigma(W(s), s \leq t), t > 0\}$ (但し、 $\mathcal{F}_t \subset \mathcal{F}$) をブラウン運動に関する情報集合列(Brownian Filtration)とする、但し、 $\sigma(W(s), s \leq t)$ はブラウン運動 $(W(s), s \leq t)$ より生成された σ 加法族である。一方、離散時間のフレームワークでの議論の祭には $\varepsilon_1, \dots, \varepsilon_n$ を独立で標準正規分布に従う確率変数列とし、 $\{\mathcal{F}_n^{\#} = \sigma(\varepsilon_1, \dots, \varepsilon_n), n \geq 1\}$, $(\mathcal{F}_n^{\#} \subset \mathcal{F})$ を正規ランダムウォークに関する情報集合列とする。

第1-1節 金利の期間構造とスプレッド

市場には各種の利子が存在する。銀行預金の利子が一番判りやすいが、それ以外にも LIBOR、国債や社債に代表されるクーポン債やゼロクーポン債、さらにはスワップ・レーが市場利子率を決定する。さて、借り入れや貸し出し期間の長さ、借り手・貸し手の信用力や流動性によって利子率は異なる。一般に国債には倒産の危険がなく(Default Free)流動性も十分あるとされている為、国債の利子率は一番低い。これは、市場では日常的に観察されるし、直感的にもわかりやすい事実であろう。一方、社債等は発行主体の倒産可能性(Defaultable)がある為、利子率は高くなる。そこで倒産リスクを持った債券と無リスクな国債の利回りとの間に**金利スプレッド(Spread)**と呼ばれる差が生る。金利スプレッドの存在を説明し、更にそれを定量的に算出する理論の構築とは過去30年以上に渡り金融工学・数理ファイナンス世界で大きな問題の一つとなっている。本稿では取りあえず、倒産等の信用リスクがスプレッドの唯一の要因と仮定し、以上の問題に答えていこう*1。

利子率を現時点から満期日までの時間の関数と見たものを金利の期間構造(Term Structure of Interest Rate)とよび、そのグラフがイールドカーブ(Yield Curve)である。期間構造は日々変化する、その変化を分析する代表的な方法として(1)局所無裁定法(Local Method)と(2)マルチンゲール法(Martingale Method)がある。局所無裁定法に基づく派生証券価格決定理論の代表的な論文は Vasicek [1977]、マルチンゲール法では Harrison and Kreps [1979], Harrison and Pliska [1981], Heath, Jarrow and Morton [1992] 等であるが、本稿では、次節以下、マルチンゲール法に基づいた議論を紹介する。

第1-2節 マルチンゲール法に基づく無リスクな利子率の決定

本節ではまず、利子率決定のメカニズムを国債等に代表されるクーポン支払いがなく(Zero Coupone) Default Freeな割引債券の価格(安全債券価格)の決定を解説する。又、簡単化のため我々の考える経済では連続時間での取引が有限な時間内 $[0, T]$ で行われるも

*1流動性もスプレッド存在の要素となっている。

のとし、 $P^*(t,T)$ を T を満期日とするZero Coupon Default Free債の t 時点での価格とする ($0 \leq t \leq T \leq T$)。以下、各債権を満期日により区別するために記号の乱用ではあるが $P^*(t,T)$ で債権そのものをも表すことにし、それを T 債権と呼ぶことにする*2。 $\xi = T-t$ を当該債権の残存期間、 $R^*(t, \xi)$ で利回り(Yield to Maturity)を表すことにしよう。上でも述べたように $R^*(t, \xi)$ を ξ の関数と見たものが金利の期間構造である。金利の期間構造を分析する際の、もう一方の主役がスポット金利(Spot Rate)である。スポット金利 $r(t)$ は時点 t から $t+dt$ の間の利子率である(離散時間モデルでは次時点までの間の利子率)。この金利は時点 t で決定され、当該区間では既知であるという意味で安全利子率(Risk Free Rate)と呼ばれる。利子率 $r(u)$ で1円を $[0,t]$ の間連続複利で運用したときの元利合計を基準資産とよび $B(t)$ と書く。ここで、以下の関係式が成立する、

$$(1-1) \quad B(t) = \exp\left\{\int_0^t r(u) du\right\}$$

次に $P^*(t,T)$ に対する基本的な仮定を述べ、その後無裁定条件下での $P^*(t,T)$, $R^*(t, \xi)$, $r(t)$ の関係を考える。特に $P^*(t,T)$ と $r(t)$ との関係は基本的である。

[仮定1-1] 時点 t における割引債価格 $P^*(t,T)$ は以下の条件を満たすものとする。

(i) 任意の $0 \leq t \leq T$ に対し

$$(1-2) \quad P^*(t,t) = 1$$

(ii) 任意の $0 \leq t \leq T \leq T$ に対し、

$$(1-3) \quad P^*(t,T) > 0$$

(iii) 一般に $P^*(s,T)$, $t \leq s \leq T$ は満期日以外ではランダムに変動する確率過程である。即ち、時点 t で全ての $T \geq 0$ に対し、 $P^*(t,T)$ は既知であるが、 $P^*(s,T)$ は不確定要素を含む確率変数である。//

T 債権の時点 t 価値は $P^*(t,T)$ 円、時点 T では1円の価値がある。利回り $R^*(t, \xi)$ との間に裁定関係が存在しないためには以下の式が成立していなければならない。

$$(1-4) \quad P^*(t,T) \cdot \exp\{\xi R^*(t, \xi)\} = P^*(T,T) = 1, \quad \text{但し、} \xi = T-t$$

又は、

$$(1-5) \quad R^*(t, \xi) = \frac{-1}{\xi} \log P^*(t,T)$$

*2 本稿では満期日により債券を区別している。即ち、満期日が同じであれば同じ債券としている。

次に[仮定 1 – 1 (iii)]に関連し確率的な期間構造モデルを導入する。市場には一般に複数の確率変動の要素が存在するが、本節では簡単のため確率変動のソースは一つ、1 – ファクターモデルを仮定する

[仮定 1 – 2 : 確率モデル] $W^*(t)$ を (Ω, F, Pr) 上で定義された標準ブラウン運動とする。任意の $0 \leq t \leq T \leq T$ に対し、 $P^*(t, T)$, $r(t)$ の確率変動は以下の (伊藤の) 確率微分方程式で与えられる。

$$(1-6) \quad \frac{dP^*(t, T)}{P^*(t, T)} = \mu^*(t, T) dt + \sigma(t, T) dW^*(t),$$

$$(1-7) \quad dr(t) = \alpha(t) dt + \beta(t) dW^*(t) \quad //$$

さて、もしもスポットレート $r(u)$ と債券価格 $P^*(u, T)$ が全ての $0 \leq u \leq T$ に対し既知であったとすると (即ち、非確率的モデル)、市場に裁定機会が存在しない時には

$$(1-8) \quad P^*(t, T) \exp\left\{\int_t^T r(u) du\right\} = 1$$

が成立する。 $P^*(T, T) = 1$ より、(1-8)を書きかえると、

$$(1-9) \quad P^*(t, T) = \exp\left\{\int_0^t r(u) du\right\} \cdot \exp\left\{-\int_0^T r(u) du\right\} \cdot P^*(T, T)$$

$B(t)$ の定義より、

$$(1-10) \quad \frac{P^*(t, T)}{B(t)} = \frac{P^*(T, T)}{B(T)} \quad \text{for all } 0 \leq t \leq T$$

が成立する。さらに(1-10)は無裁定条件と同値な条件となる事も証明される (cf. 高橋 [2009])。ここで重要なことは一般の確率モデルに於いても、或る確率測度のもとの期待値の世界で同様の結果が証明される。それが以下で紹介する **Harrison-Pliska の定理**”無裁定条件下における派生証券価格決定の基本定理”でマルチンゲール法の基礎である。

[定理 1 – 4 : 派生証券価格決定の基本定理 **Harrison and Pliska [1982]**]市場に裁定機会が存在しない為の必要十分条件は Pr と同値な確率測度 Q が存在し、そのもとで $P^*(t, T)/B(t)$ が F_t マルチンゲールとなることである。即ち、任意の $0 \leq t \leq T \leq T$ に対し、

$$(1-11) \quad \frac{P^*(t, T)}{B(t)} = E^Q\left\{\frac{P^*(T, T)}{B(T)} \middle| F_t\right\} = E^Q\left\{\frac{1}{B(T)} \middle| F_t\right\}$$

が成立する。但し、同値な確率測度とは両確率の元での確率ゼロで起こる事象(不可能事象)を共有するという意味での同値関係である。また、 E^Q は測度 Q のもとでの期待値である。
//

これより、 T 債権の時点 t 価格は

$$(1-12) \quad P^*(t, T) = E^Q\{\exp\{-\int_t^T r(u) du | \mathcal{F}_t\}$$

となる。勿論この公式を実際に使うためには Q 測度の元でのスポット金利プロセス $\{r(t), t \geq 0\}$ とその同時分布が与えられなければならないが、幾つかの重要な特殊ケースにおいて実際に計算可能である。

これまでは債券の発行主体は倒産しないという仮定の下での議論であったが、倒産の可能性をもった一般企業が発行するゼロクーポン社債の時点 t 価格 $P(t, T)$ はどのように決定されかも同様に重要な問題である。詳しくは、本論文の主要テーマとも関連するため第2部で議論するが、まずは、ここでも $P(T, T)=1$, $P(t, T) \geq 0$ ($0 \leq t \leq T$) を仮定しよう。その上で安全債権と危険債権の時点 t から満期日 T までの利回りをそれぞれ $R^*(t, \xi)$, $R(t, \xi)$ とした時の両者の差が所謂スプレッド (Spread) である。

[定義 1-5 : 金利スプレッド] 安全債権の利回りと危険債権の利回りの間のスプレッド $\Upsilon(t) = \Upsilon(t; P^*, P)$ は

$$(1-13) \quad \begin{aligned} \Upsilon(t) &= \Upsilon(t; P^*, P) = R(t, \xi) - R^*(t, \xi) \\ &= \frac{-1}{\xi} \log \left[\frac{P(t, T)}{P^*(t, T)} \right] \end{aligned}$$

で定義される。//

倒産リスクを加味すれば $P(t, T) \leq P^*(t, T)$ 、即ち、 $R(t, \xi) \geq R^*(t, \xi)$ と予想され、一般にスプレッドは正値を取ると思われる。第2部では $P(t, T)$ の価格決定モデルとスプレッドに基づいた倒産確率の推定について考察する。その前に次節に於いてハザードモデルを紹介しよう。

第 1-3 節 ハザード (Hazard) モデルと センサーされた (Censored) データ

生存時間分析は品質管理や医学統計の世界で重要な位置を占め続けてきている。そこで研究・開発されてきた手法は信用リスク問題へ応用可能である。企業倒産時刻と死亡時刻を対応させるのみならず、信用リスク問題における他の信用クラス (各付け) への移動は

医学統計における観測個体の退出（センサーリング）に相当する。また、細かな格付けを行うことによる格付け毎のデータ数の相対的な減少はCoxの比例ハザードモデルにより改善される。以下本節ではハザード関数、比例ハザードモデルについて簡単な解説を行う。

[ハザード関数：記号と基本用語] すべての企業がある格付けクラスに入り、倒産するまで同じクラスに居つづけるならば問題は比較的単純である。しかし、実際にはそのまま倒産する場合もあれば、他の格付けへ移行するケースもある。実際、Aaaに格付けされた企業がそのまま倒産することは殆ど無いだろう。倒産する場合も、先ず下位のクラスに格付けされ、それから倒産となる場合が多い。そこで問題となるのは倒産前に他のクラスへ移行した企業の持つ情報、所謂センサーされたデータが持つ情報をどのように活用するかである。さて、センサーされたデータとは当該事象（倒産）の生起・不生起以前に対象が観測不能となる状態のことである。少なくともセンサーされた時点までは倒産していないという情報をデータは持っているが、それをどのように活用するかが生存解析の主題の一つである。正確には倒産時点が不明という意味に於いて、これを右センサーされたデータと呼ぶことがある。一方、当該格付けクラスに参入した時点が不明であるようなデータは左センサーされたデータと呼ばれる。

まず、一つの格付けクラスを選び固定する。ある企業が当該格付けに入ってから倒産する迄の時間（生存時間）を確率変数 τ で表す。一方センサーされる迄の時間は C とする。実際には、大部分の企業は観測が始まった時点では既に当該クラスに格付けされているという意味に於いて左センサーされたものと考えられるが、本稿では単純化のためこの立場は取らない。その為、例えば現時点を基準時 0 とし、 C をセンサー時刻（時点）と言うこともある。 τ の分布関数と密度関数をそれぞれ $F(t)$, $f(t)$ とする。また、センサー時点 C は生存時間 τ とは独立で分布関数 $G(t)$ 、密度関数 $g(t)$ に従うと仮定する。但し、観測終了時点 T については自動的に定数値のセンサー時点とする（ C は T で切断されている）。

さて、生存関数 $S(t)$ とハザード関数 $\lambda(t)$ はそれぞれ、

$$S(t) = 1 - F(t) = P\{\tau > t\}$$

(1-16)

$$\lambda(t) = \frac{f(t)}{S(t)}$$

で定義する。生存関数は、時点 t を超えて企業が生き残る確率、また、ハザード関数は $\lambda(t)dt \doteq P\{t < \tau \leq t+dt \mid \tau > t\}$ と書き直してみると分かりやすい。即ち、ハザード関数とは時点 t まで倒産しなかった企業が次の瞬間（微少時間内）に倒産するという条件付き確率である。簡単な積分計算より

$$(1-17) \quad S(t) = \exp\{-\int_0^t \lambda(u)du\}, \quad F(\infty)=1 \Leftrightarrow \int_0^\infty \lambda(u)du = \infty$$

が得られる。次に観測時点 Y とセンサー指示関数 δ を以下のように定義する。

$$(1-18) \quad Y = \min\{\tau, C\} = \tau \wedge C \sim H(t)$$

$$(1-19) \quad \delta = I_{[\tau \leq C]} = \begin{cases} 1 & \text{if uncensored} \\ 0 & \text{if censored} \end{cases}$$

τ の分布、又は生存関数やハザード関数に対しパラメトリック、ノンパラメトリック等色々なモデルが考えられる。以下では代表的なパラメトリックモデルを紹介するが、これらは経験的合理性を伴ったハザード関数の形から考えられたものである。

[例 1 - 6 : 指数 (Exponential) モデル] 一番簡単なパラメトリックモデルはハザード関数が時間に対して、一定となる指数分布型のモデルであろう。すべての $t > 0$ に対して、ハザード関数 $\lambda(t)$ と生存関数 $S(t)$ は、それぞれ、

$$(1-20) \quad \lambda(t) = \theta > 0 \quad (\text{定数})$$

$$S(t) = \exp\{-\theta t\} \quad t \geq 0$$

である。分布関数 $F(t) = 1 - \exp\{-\theta t\} \quad t \geq 0$, 密度関数 $f(t) = \theta \exp\{-\theta t\} \quad t \geq 0$, である。勿論、一般に θ は未知で推定の対象となる。指数分布は簡単ではあるが、その単純性から多くのケースで活用されている。指数分布の拡張として以下のガンマモデルとワイバルモデルがよく知られている

[例 1 - 7 : ガンマ (Gamma) モデル] 指数分布の自然な拡張がガンマ分布である。 θ を尺度 (Scale) パラメータとし、それに加え形状 (Shape) パラメータ α が導入される。密度関数は

$$(1-21) \quad f(t) = \frac{\theta^\alpha}{\Gamma(\alpha)} t^{\alpha-1} \exp\{-\theta t\} \quad t \geq 0$$

ただし、 Γ はガンマ関数である。ガンマ分布関数は不完全ガンマ関数を含むため生存関数、ハザード関数、分布関数の明示的な表現は得られない。

[例 1 - 8 : ワイバル (Weibull) モデル] 指数モデルにおいて、生存関数を

$$(1-22) \quad S(t) = \exp\{-(\theta t)^\alpha\}$$

と拡張したものがワイバルモデルである。簡単な計算より密度関数並びにハザード関数は

$$f(t) = \alpha \theta (\theta t)^{\alpha-1} \exp\{-(\theta t)^\alpha\}$$

(1-23)

$$\lambda(t) = \alpha \theta (\theta t)^{\alpha-1}$$

となることも直ちに示せる。//

以上、代表的な3種類のパラメトリックモデルを紹介したが、これらの他にも、ラレイ (Rayleigh) モデル、対数正規 (Lognormal) モデル、パレート (Pareto) モデル等がよく知られている。比例ハザードモデル(cf. (1-24))に於けるベースラインハザードとしてどの様なモデルを選ぶべきかについて、信用リスク問題でも、信頼性工学や医療統計同様、外的に与えられる幾つかの経済変数の影響を無視できない。特にロングランの信用リスク問題では経済の構造変化等も考えられる。一方、実装の際には指数分布型モデルの単純さも魅力的である。勿論データによる丁寧な検証は必要であるが、比較的短期期間に於いて指数分布を仮定することは一定の合理性を持つであろう。

[ランダム・センサーリングモデル(Random Censoring Model)] センサー時点について考えると、例えば観測終了時 T のように予め決まっている時点もあれば、何らかの外的理由による退場もある。ここでは、後者によるセンサー時刻 C を確率モデル化してみる (cf. Efron[1980])。ランダムセンサーリングモデルは第2節で展開するセンサーされたデータに対するパラメトリックブーツトラップを実装するために用いられる。

(1-18)で倒産時刻 τ の代わりにセンサー時刻 C に注目し、その生存関数 $1 - G(t)$ の推定問題を考える (但し、ここでは T で切断されてはいないと仮定しよう)。観測終了時刻に対応するセンサー時刻を除き C と τ の役割を入れ替えてみると、 C に取って τ はセンサー時点に他ならない。 τ を T で切断したものにする、ここでも観測値は上と同様 $Y = \min\{\tau, C\}$ であるが、センサー指示関数は $\delta' = 1 - \delta$ となる。[例1-6]~[例1-8]のパラメトリックモデルは C に関するモデルとしても有効である。//

勿論、格付けクラス毎に生存時間の分布も異なる、そこで本稿では以下のモデル、比例ハザードモデルを導入する。ポイントはハザード関数を通じて異なるモデルを特徴付けることである。

第1-4節 比例ハザードモデル(Proportional Hazard Model)

本節ではCox [1974]が提唱した比例ハザードモデルを概観する。 $\mathbf{x} = (x_1, x_2, \dots, x_p)$ を独立変数、 $\boldsymbol{\beta} = (\beta_1, \beta_2, \dots, \beta_p)$ を(未知)パラメータとし、比例ハザード関数を以下のように定義する：

$$(1-24) \quad \lambda(t; \mathbf{x}, \boldsymbol{\beta}) = \lambda_0(t) \exp\{\boldsymbol{\beta}' \mathbf{x}\}$$

以下 $\lambda_0(t)$ をベースライン (Base Line) ハザード (関数) と呼ぶ。生存関数は

$$(1-25) \quad \begin{aligned} S(t; \mathbf{x}, \boldsymbol{\beta}) &= \exp\left\{-\int_0^t \lambda(u; \mathbf{x}, \boldsymbol{\beta}) du\right\} = \exp\left\{-e^{\boldsymbol{\beta}'\mathbf{x}} \int_0^t \lambda_0(u) du\right\} \\ &= \exp\left\{-\int_0^t \lambda_0(u) du\right\} \exp\{e^{\boldsymbol{\beta}'\mathbf{x}}\} = S_0(t)^{\exp(\boldsymbol{\beta}'\mathbf{x})} \end{aligned}$$

また、分布関数は $F(t; \mathbf{x}, \boldsymbol{\beta}) = 1 - S_0(t)^{\exp(\boldsymbol{\beta}'\mathbf{x})}$ となる。例えば $p=1, \mathbf{x}=0$ or 1 であれば、 $S(t; 0, \boldsymbol{\beta})$ と $S(t; 1, \boldsymbol{\beta})$ により 2 標本問題がカバーされる。勿論これ以外にも幾つかの方法が考えられる。 $p=1$ で同一の $\boldsymbol{\beta}$ に対し異なる \mathbf{x} の値でクラスを区別する方法や $p =$ "格付けクラスの数" とし \mathbf{x} をダミー変数とする等、色々考えられる。何れにせよデータに基づいたモデル選択と $\lambda_0(t)$ と $\boldsymbol{\beta}$ の推定が必要となる。この問題については第 3 部で詳しく論じる。第 3-2 節で展開する Parametric Bootstrap の場合ベースラインハザードについてはパラメトリックモデルを仮定する。

第 1-5 節 ハザード関数のパラメトリックモデルと推定論

本節ではパラメトリックな生存時間モデル $\{H(t, \boldsymbol{\theta}) : \boldsymbol{\theta} \in \Theta\}$ を仮定し未知パラメータ $\boldsymbol{\theta}$ の最尤推定法について議論する。(1-18), (1-19) にあるように、我々が用いることの出来るデータは $(Y_1, \delta_1), (Y_2, \delta_2), \dots, (Y_n, \delta_n)$ である。ここで Y_1, Y_2, \dots, Y_n は独立で同一な分布 $H(t; \boldsymbol{\theta})$, $\boldsymbol{\theta} \in \Theta$ からの大きさ n のランダムサンプルである。先ず尤度関数の定義から始めよう、 (Y_i, δ_i) の実現値 (y_i, δ_i) の尤度関数 $L^\tau_i(\boldsymbol{\theta}) = L^\tau_\theta(y_i, \delta_i)$ は

$$(1-26) \quad \begin{aligned} L^\tau_i(\boldsymbol{\theta}) = L^\tau_\theta(y_i, \delta_i) &= \begin{cases} f(y_i) = f(y_i; \boldsymbol{\theta}) \\ S(y_i) = S(y_i; \boldsymbol{\theta}) \end{cases} \text{ if } \begin{cases} \delta_i = 1 & (\text{uncensored}) \\ \delta_i = 0 & (\text{censored}) \end{cases} \\ &= f(y_i)^{\delta_i} S(y_i)^{1-\delta_i} \quad (i = 1, \dots, n) \end{aligned}$$

従って、標本全体の尤度関数 $L^\tau(\boldsymbol{\theta})$ はセンサー時刻と倒産時刻の独立性の仮定より以下のように表しても一般性は失われない (cf. Miller [1981])

$$(1-27) \quad \begin{aligned} L^\tau(\boldsymbol{\theta}) &= L^\tau(y_1, y_2, \dots, y_n; \delta_1, \delta_2, \dots, \delta_n; \boldsymbol{\theta}) = \prod_{i=1}^n L^\tau_\theta(y_i, \delta_i) \\ &= [\prod_u f(y_i)] [\prod_c S(y_i)] \end{aligned}$$

但し、 Π_u はセンサーされていないデータに対応する i に関する積、 Π_c はセンサーされたデータに関する積である。この様に、尤度関数が求められたなら、一定の正則条件下で未知パラメータは対数尤度関数を微分し、それをゼロとおくことにより求められる：

$$0 = \frac{\partial}{\partial \theta_j} \log L^\tau(\boldsymbol{\theta}) = \sum_{i=1}^n \frac{\partial}{\partial \theta_j} \log L^\tau_{\boldsymbol{\theta}}(y_i, \delta_i)$$

(1-28)

$$= \sum_u \frac{\partial}{\partial \theta_j} \log f_{\boldsymbol{\theta}}(y_i) + \sum_c \frac{\partial}{\partial \theta_j} \log S_{\boldsymbol{\theta}}(y_i), \quad j = 1, \dots, p$$

そして、最尤推定量（ベクトル） $\hat{\boldsymbol{\theta}} = (\hat{\theta}_1, \dots, \hat{\theta}_p)$ は連立方程式

$$(1-29) \quad 0 = \frac{\partial}{\partial \theta_j} \log L^\tau(\hat{\boldsymbol{\theta}}) \quad j = 1, \dots, p$$

の解として与えられる。勿論、多くの場合上記方程式は解析的には解けず、数値的に解くしかない。数値解析の代表的方法は以下で説明する Newton-Raphson (NR) 法である。

[Newton Raphson 法] N-R 法は以下の手順で方程式(1-29)を解いていく：

1st Step: まず最尤推定値 ((1-29) の解) を予想しそれを $\hat{\boldsymbol{\theta}}^{(1)}$ とする。

2nd Step: $\frac{\partial}{\partial \theta_j} \log L^\tau(\boldsymbol{\theta})$ を $\hat{\boldsymbol{\theta}}$ の周りでテイラー展開、(1-29) 式より

$$0 = \frac{\partial}{\partial \theta_j} \log L(\hat{\boldsymbol{\theta}}) = \frac{\partial}{\partial \theta_j} \log L(\hat{\boldsymbol{\theta}}^{(1)}) + (\hat{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}^{(1)}) \frac{\partial^2}{\partial \theta^2_j} \log L^\tau(\hat{\boldsymbol{\theta}}^{(1)}) + \dots$$

(j = 1, \dots, p)

上式で二次以上の高次項を無視することより、更新された推定値 $\hat{\boldsymbol{\theta}}^{(2)}$ を得る。以下の形でのベクトル表示にすると、

$$(1-30) \quad \hat{\boldsymbol{\theta}}^{(2)} = \hat{\boldsymbol{\theta}}^{(1)} + [\mathbf{i}(\hat{\boldsymbol{\theta}}^{(1)})]^{-1} \frac{\partial}{\partial \boldsymbol{\theta}} \log L^\tau(\hat{\boldsymbol{\theta}}^{(1)})$$

但し、

$$(1-31) \quad \mathbf{i}(\hat{\boldsymbol{\theta}}^{(1)}) = \frac{-\partial^2}{\partial \theta^2} \log L^\tau(\hat{\boldsymbol{\theta}}^{(1)})$$

は標本情報行列と呼ばれる。

3rd Step: 2nd step を繰り返し、一般には以下の推定値を得る：

$$(1-32) \quad \hat{\boldsymbol{\theta}}^{(k+1)} = \hat{\boldsymbol{\theta}}^{(k)} + [\mathbf{i}(\hat{\boldsymbol{\theta}}^{(k)})]^{-1} \frac{\partial}{\partial \boldsymbol{\theta}} \log L(\hat{\boldsymbol{\theta}}^{(k)}) \quad k \geq 1$$

一方、 $L^\tau(\hat{\boldsymbol{\theta}})$ に対する適当な正則条件下で $\hat{\boldsymbol{\theta}}^{(k)} \xrightarrow{P} \hat{\boldsymbol{\theta}}$ となることが示されるため、あらかじめ定められた限界値 $a > 0$ に対し

$$(1-33) \quad \|\hat{\boldsymbol{\theta}}^{(k-1)} - \hat{\boldsymbol{\theta}}^{(k)}\| < a$$

となるところで繰り返し計算を中止し、 $\hat{\boldsymbol{\theta}}^{(k)}$ を最尤推定値とする、但し、 $\|\cdot\|$ はユークリッドノルムである。

もしも、Fisher 情報行列 $\mathbf{I}(\boldsymbol{\theta}) = \mathbf{E}\{\mathbf{i}(\boldsymbol{\theta})\}$ が利用できる場合は、標本情報行列の代わりに Fisher 情報行列を用い(1-32)の一般項の代わりに

$$(1-34) \quad \hat{\boldsymbol{\theta}}^{(k+1)} = \hat{\boldsymbol{\theta}}^{(k)} + \mathbf{I}(\hat{\boldsymbol{\theta}}^{(k)})^{-1} \frac{\partial}{\partial \boldsymbol{\theta}} \log L^{\tau}(\hat{\boldsymbol{\theta}}^{(k)}) \quad k \geq 1$$

と求めることも出来る。//

一定の正則条件下で、センサーされたデータに基づく最尤推定量も一致性と漸近正規性を持つことが知られている：例えば、

$$(1-35) \quad (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \mathbf{I}(\boldsymbol{\theta})^{1/2} \xrightarrow{d} N(0, \mathbf{I}_p) \quad \text{as } n \rightarrow \infty$$

ただし、 \mathbf{I}_p は $p \times p$ の単位行列である。これより、 $\boldsymbol{\theta}^0$ を真のパラメータとするとき以下の結果が証明される：

(1) Wald

$$(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0)' \mathbf{I}(\boldsymbol{\theta}^0) (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0) \xrightarrow{d} \chi^2_p \quad \text{as } n \rightarrow \infty$$

(2) Neyman-Pearson

$$-2 \log [L(\boldsymbol{\theta}^0) / L^{\tau}(\hat{\boldsymbol{\theta}})] \xrightarrow{d} \chi^2_p \quad \text{as } n \rightarrow \infty$$

(3) Rao

$$\frac{\partial}{\partial \boldsymbol{\theta}} \log L^{\tau}(\boldsymbol{\theta}^0)' \mathbf{I}(\boldsymbol{\theta}^0)^{-1} \frac{\partial}{\partial \boldsymbol{\theta}} \log L^{\tau}(\boldsymbol{\theta}^0) \xrightarrow{d} \chi^2_p \quad \text{as } n \rightarrow \infty \quad //$$

この様に、生存関数の推定に関してはセンサーされたデータに基づく場合でも、通常のデータ抽出の場合とほぼ同様な統計学的な結果が示される。また、比例ハザードモデルの推定等も現在では標準的な統計学手法であり、その統計学的な性質について標準的な結果が数多く示されてきている。しかしながら本論文では若干特異な立場から論じていく(パラメトリックモデルとノンパラメトリックモデルの融合)ため、この件については第3部で今一度論じることとする。

以下は Parametric Bootstrap 法適用の際に必要な。死亡時刻とセンサー時刻を入れ替えた時の尤度関数を与えておく。

ランダムセンサーリングモデルに於いても尤度関数は(1-26), (1-27)と同様に定義できる。(C, δ^c)に関する尤度関数を L^Cと置くと尤度は

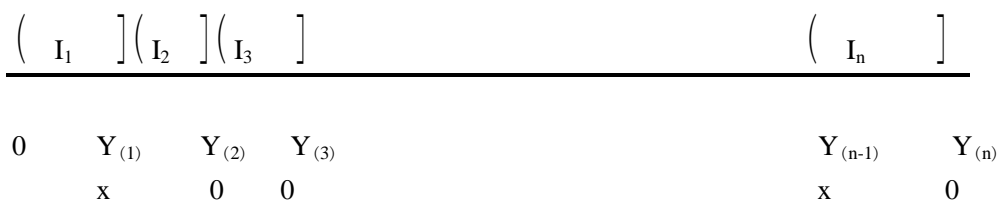
$$(1-36) \quad L^C(\theta) = L^C(y_1, y_2, \dots, y_n; \delta^{c_1}, \delta^{c_2}, \dots, \delta^{c_n}; \theta) = \prod_{i=1}^n L^C_{\theta}(y_i, \delta^{c_i}) \\ = [\prod_{u_g}(y_i)] [\prod_{c_R}(y_i)]$$

$$(1-37) \quad \delta^{c_i} = I_{[\tau \geq C]} = \begin{cases} 1 & \text{if } \text{uncensored} \\ 0 & \text{if } \text{censored} \end{cases}$$

但し、Π_uはセンサーされていないデータ(変数 C)に対応する i に関する積、Π_cはセンサーされたデータ(変数 τ)に関する積である。ここでも上と同様、尤度関数に基づいた統計学的推論が可能となる。

第 1 - 6 節 生存関数の Non-Parametric 推定

生存関数のノンパラメトリック推定法は古くから研究されてきている。代表的なものとして、例えば、Reduced Sample Method, Actuarial Method, そして Product-Limit Estimator、即ち Kaplan-Meier (K-M) 推定量等がある。以下 K-M 推定量と、その導出のアイデアを簡単に紹介する。まず、時間軸を区間 I_i = (Y_(i-1), Y_(i)], i=1, ..., n と分割する。ただし、(Y₍₁₎, Y₍₂₎, ..., Y_(n))は観測時点 (Y₁, Y₂, ..., Y_n)を順番に並べたもので δ_(i) は Y_(i) に対応するセンサー指示関数とする。



[図 1 - 1]

但し、変数 Y の下の "0" と "x" は "0 = センサー、x = センサーされていない"データということを意味する。ここで、R(t)を時点 t におけるリスク集合(Risk Set) (時点 t まで生き残っている企業の集合) とし、さらに以下の記号を導入する。

- n_i = R(Y_(i))に含まれる個体数 = 時点 Y_(i)- で倒産していない企業数
- d_i = 時点 Y_(i)で倒産した企業数
- p_i = Pr{τ > Y_(i) | τ > Y_{(i-1)}}
- q_i = 1 - p_i

時点 t における生存関数の値を求めるために、区間(0,t]を上記のような小区間に細分する。

すると、確率の乗法定理より、

$$\begin{aligned}
 S(Y_{(k)}) &= Pr\{\tau > Y_{(k)}\} \\
 (1-38) \quad &= Pr\{\tau > Y_{(1)}\} Pr\{\tau > Y_{(2)} | \tau > Y_{(1)}\} \cdots Pr\{\tau > Y_{(k)} | \tau > Y_{(k-1)}\} \\
 &= p_1 p_2 \cdots p_k
 \end{aligned}$$

となる。従って $S(Y_{(k)})$ を推定するには(1-38)式中の各 $p_i = Pr\{\tau > Y_{(i+1)} | \tau > Y_{(i)}\}$ を推定すればよい。そこで、 p_i と q_i を

$$(1-39) \quad \hat{q}_i = \frac{d_i}{n_i}$$

そして、

$$(1-40) \quad \hat{p}_i = 1 - \hat{q}_i = \begin{cases} 1 - \frac{1}{n_i} & \text{if } \begin{cases} \delta_{(i)} = 1 \\ \delta_{(i)} = 0 \end{cases} \\ 1 & \end{cases}$$

で推定する。ここで、タイのない (No Tied Data, 同時刻で観測されたという意味での) データの場合、生存関数の K-M 推定量 $\hat{S}(t)$ を

$$(1-41) \quad \hat{S}(t) = \prod_{y^{(i)} \leq t} \hat{p}_i = \prod_{y^{(i)} \leq t} \left[\frac{n-i}{n-i+1} \right]^{\delta_{(i)}}$$

で定義する。

第2部 Implied 倒産確率の推定と統計学的考察

第1部は準備として、本論文で用いられるファイナンスと統計学の標準的な手法を概観してきた。第2部では金利スプレッドより得られる情報を用い倒産確率を推定する方法と、推定誤差について考えていく。第2部の主目的は推定値が必ず持っている誤差限界をキチンと求めていくことにある。

第2-1節：Duffie-Singleton の信用リスクモデル

本節では信用リスクを加味した金利の期間構造モデルの解説・構築、およびモデルの推定法について考えていく。1970年代に Merton[1974]により提唱された構造型モデル (Structural Model) とその後、90年代に主に Duffie-Singleton[1997]等により考察された誘導型 (Reduced Model) が代表的な方法である。本稿では先ず Duffie-Singleton (D-S) モデルを紹介し、それに基づく倒産確率の推定法” Implied 倒産確率” について議論していく：

[モデルと価格式] 倒産リスクを持つ (Defaultable) 企業の社債価格を求めるために、本節では先ず離散時間モデルで基本的なアイデアを解説し、その後連続時間モデルにおける価格式を提示する。以下、離散モデルではブラウン運動の代わりに正規ランダムウォークを仮定し、フィルトレーションも正規ランダムウォークに対応したものとする。又、離散時間モデルでは記号に[#]を付けることとする。

本節ではリスク中立的世界 (同値マルチンゲール測度下) で議論を進めていく。以下 Q を原確率 Pr と同値なマルチンゲール確率測度とする。本章でもこれまで同様 Defaultable な企業の発行する債券価格等を表す時は*無しの記号を用いることとする。例えば Defaultable 債券の時点 t 価格は離散時間の元では $P^{\#}(t, T)$ 、連続時間では単に $P(t, T)$ と記す。現時点は t 、満期時点 T 等の記号はこれまで通りである。本節の前半部は一部を除き基本的には Duffie-Singleton [1999] からの抜粋である。

ある特定の企業を企業倒産に関連し ζ 以下の記号を導入する。まず、当該企業が倒産する時刻を離散型確率変数 $\tau^{\#}$ としその分布関数を $F^{\#}$ とする：

$$(2-1) \quad F^{\#}(t) = Q\{\tau^{\#} \leq t\}$$

また、離散型ハザード関数を

$$(2-2) \quad \lambda^{\#}(t) = Q\{\tau^{\#} = t \mid \tau^{\#} \geq t\}$$

で定義する。次に $\phi^{\#}(t)$ を倒産時における資金の回収率、 $r^{\#}(t) = r_t^{\#}$ を区間 $[t, t+1)$ における安全利子率とする。

以上の準備の元、満期時点 T で X_T 円償還する **Defaultable** な債権の時点 t 価格 $V^\#(t)$ を求めてみよう。時点 t 迄倒産が起こらないという条件下で債権の時点 t 価格 $V^\#(t)$ を求めるために、次の時点 $t+1$ で倒産する場合と生き残る場合に分けて考える。 E^Q を確率測度 Q のもとでの期待値とすると、

$$(2-3) \quad V^\#(t) = \lambda^\#(t) \exp\{-r^\#\} E^Q\{\phi^\#(t+1)|F^\#\} + [1-\lambda^\#(t)] \exp\{-r^\#\} E^Q\{V^\#(t+1)|F^\#\},$$

但し、債務回収は $t+1$ 時点で行われるものとする。上式を 1 期ずつ繰り返し t を T 迄増加させると、最終的には

$$(2-4) \quad V^\#(t) = E^Q[\sum_{j=0}^{T-t-1} \exp\{-\sum_{k=0}^j r^\#(t+k)\} \phi^\#(t+j+1) \prod_{k=0}^j \{1-\lambda^\#(t+k-1)\} |F^\#\] \\ + E^Q[\exp\{-\sum_{k=0}^{T-t-1} r^\#(t+k)\} X_T \prod_{m=0}^{T-t} [1-\lambda^\#(t+m-1)] |F^\#\]$$

を得るが、実際に使うには少々複雑すぎる。そこで簡単化のため時点 t に於ける回収率 $\phi^\#(t+1)$ は $V^\#(t+1)$ の一定割合で $F^\#_{t+1}$ -可予測関数 $\delta^\#_t = (1-L^\#_t)$ と仮定し、以下の関係式を得る：

$$(2-5) \quad E^Q\{\phi^\#(t+1)|F^\#\} = \delta^\#_t E^Q\{V^\#(t+1)|F^\#\}$$

上式で導入された $\delta^\#_t = (1-L^\#_t)$ は市場価値回収率 (**RMV: Recovery of Market Value**) と呼ばれる時間の関数で **D-S** モデルを特徴づけるポイントである。(2-5) より (2-3) 式は

$$(2-6) \quad \delta^\#_t \lambda^\#(t) \exp\{-r^\#\} E^Q\{V^\#(t+1)|F^\#\} + [1-\lambda^\#(t)] \exp\{-r^\#\} E^Q\{V^\#(t+1)|F^\#\}$$

となる。上と同様に t を T まで増やしていくことにより、

$$(2-7) \quad V^\#(t) = E^Q\{\exp\{-\sum_{j=0}^{T-t-1} R^\#_{t+j}\} X_T | F^\#\}$$

を得る、但し

$$(2-8) \quad \exp\{-R^\#_t\} = (1-\lambda^\#(t)) \exp\{-r^\#(t)\} + \lambda^\#(t) \exp\{-r^\#(t)\} \delta^\#_t$$

である。Taylor の定理より、

$$(2-9) \quad \exp\{-R^\#_t\} = \exp\{-r^\#(t) + \lambda^\#(t)(1-\delta^\#_t)\} [1+o(r^\#_t)] \quad \text{as } r^\#_t \rightarrow 0$$

が成立する。即ち、ハザード関数 $\lambda^\#(t)$ と回収率 $\delta^\#_t = (1-L^\#_t)$ で特徴づけられる **Defaultable** な企業の社債価格は無リスク債券価格と同様な式より求められる。**RiskFree** の場合との唯一の違いは割引率が **Risk Free** レート $r^\#(t)$ でなく $R^\#(t) = r^\#(t) + \lambda^\#(t)(1-\delta^\#_t)$

になる点である。

以下、この結果を連続時間のフレームワークで書き直す。記号は主な関数から[#]を外すことにする。離散時間モデルの結果は、幾つかの数学的な問題は生じるが基本的には観測時点間をゼロに”収束”させると、結果的にはその間における利子率もゼロの収束することより連続時間モデルに変換される。これより証明は省くが、以下の連続時間における価格決定式が求められるのである。

$$(2-10) \quad V(t) = E^Q\{\exp\{-\int_t^T R(u) du\} X_T | F_t\}$$

勿論、 $R^{\#}(t) = r(t) + \lambda(t)(1 - \delta)$ となることは言うまでもない。

第2-2節 Implied 倒産確率の推定：ノンパラメトリック法

本節では主として連続時間モデルを仮定する。仮に $r(t)$ と $L(t)$ が既知ならば、市場で観測された実際の価格 $V(t)$ を(2-10)式の左辺に代入することにより”理論的には”ハザード率を逆算することが可能である（離散時間の場合は $V^{\#}(t)$ を(2-7)式の左辺に代入）。この様にして求められたハザード関数をインプライドハザード関数（Implied hazard function）と呼ぶ。勿論、解析的に求めることはもとより、このままの形では数値的に計算することも非常に難しい。そこで、楠岡・青沼・中川 [2001] と同様な単純化をはかろう。

まず、金融資産としてゼロクーポン割引債を仮定； $X_T = 1$ （円）（即ち $V(t) = P(t, T)$ である）、 RMV $\delta(t)$ は定数 δ とする。更なる単純化のため $t=0$ とし、 $Q\{\tau > s\} = \exp\{-\int_0^s \lambda(u) du\}$ に注意し、改めて $T=t$ と書くと、

$$\begin{aligned} P(0,t) &= E^Q\{\exp\{-\int_0^t R(u) du\}\} \\ (2-11) \quad &= E^Q\{\exp\{-\int_0^t r(u) du - (1-\delta) \int_0^t \lambda(u) du\}\} \\ &= E^Q\{\exp\{-\int_0^t r(u) du\} \cdot \exp\{(1-\delta) \int_0^t \lambda(u) du\}\} = P^*(0,t) S(t)^{(1-\delta)} \end{aligned}$$

これと(1-13)式より、

$$(2-12) \quad S(t) = \exp\{-Y(t; P^*, P)t\}^{(1/\delta)}$$

ここでも、勿論回収率 δ の推定のみならず、最終的な結果は回収の方法と時期にも依存する。実際、D-S モデルでは Default の次の時点で債務回収が行われると仮定されているが、楠岡 et. al. [2001] では債務回収は満期日に支払われると仮定している為、彼等の算式は若干異なっている；

$$(2-13) \quad S(t) = [\exp\{-Y(t; P^*, P)t\} - \delta] / (1 - \delta)$$

(2-12)式に基づくにせよ(2-13)式から求めるにせよ、生存関数は満期日に対応する t の値でのみ計算される。従って、求められた関数の平滑化と云う問題が次に出てくるだろう。また、この様に計算された生存関数の精度等の統計学的な性質は一体どうなっているのか。次節以降ではこれらの問題について考えていく。

第2-3節 Implied 倒産確率の推定：パラメトリックモデルとその推定

生存関数 $S(t)$ の推定問題を考えたとき、上記のようなノンパラメトリック法ではすべての t に対し $S(t)$ を求めることは難しい。一方、生存関数の形態に関して何らかの事前情報が存在する場合もあるだろう。そのような時には生存関数（又はハザード関数）に以下のようなパラメトリックモデルを仮定することが考えられる。

[仮定2-1 確率モデル] Θ を R^p の部分集合とする、但し p はパラメータの次元。 $\theta \in \Theta$ を未知パラメータとし、 $G(t; \theta) = Q\{\tau > t; \theta\}$ をパラメトリックな生存関数とする。又、我々がスプレッド $Y(t)$ より求めることが出来る観測値 $S(t; \theta)$ は以下のモデルに従う確率的な誤差を含むと仮定する（即ち、確率変数である）；

$$(2-14) \quad S(t_i)^{1-\delta} = S(t_i; \theta)^{1-\delta} = G(t_i; \theta)^{1-\delta} + \sigma \rho_{t_i} \varepsilon_{t_i},$$

但し ε_{t_i} は i.i.d. で平均と分散は $(0,1)$ ($i=1, 2, \dots, n$)、 $\sigma > 0$ は未知の標準偏差、 ρ_t は $\rho_0 = 0$ を満足する t の“既知の”関数とする。市場で観測されるスプレッド $Y(t)$ より計算される $S(t)$ を $s(t)$ と書き実現値（“Implied 生存関数”）と呼ぶ。また、 $G(t; \theta)$ は任意の t に対して、 θ に関し2回微分可能であると仮定する。//

以下、 ρ_{t_i} は既知とし議論を進める。応用の場面では $\rho_{t_i} = [1 - P^*(0, t)]$ 又は $\rho_{t_i} = \left[\frac{1 - P^*(0, t)}{P^*(0, t)} \right]$ 等を仮定することが合理的であろう²。

所で、モデル(2-14)は可変分散モデルであるが本論文では ρ_{t_i} を既知と仮定しているため、変換 $X(t_i) = \frac{S(t_i)^{1-\delta}}{\rho_{t_i}}$ 並びに $\Xi(t_i; \theta) = \frac{G(t_i; \theta)^{1-\delta}}{\rho_{t_i}}$ を施すことにより等分散モデルへ変換でき、(2-14)は

$$(2-15) \quad X(t_i) = \Xi(t_i; \theta) + \sigma \varepsilon_{t_i} \quad \varepsilon_{t_i} \text{ are i.i.d. with } (0,1) \text{ for } i=1, 2, \dots, n,$$

となる。上と同様 $X(t)$ の実現値を $x(t)=x_t$ と書くことにし、それを擬似データと呼ぶ。モデル(2-15)のもとでの θ の推定は、インプライドな生存確率より求められた擬似データ列

² このモデルの導出については (Takahashi[2009]) を参照。

$\{x(t_i) = x_{t_i} \ i=1, \dots, n\}$ を観測データ値とし θ の"擬似最小二乗"推定値 $\hat{\theta}$ を以下の判定基準 (2-16) のもとで (非線形最小二乗法を用い) 求めることから始まる :

$$(2-16) \quad \hat{\theta} = \arg_{\theta \in \Theta} [\min \sum_{i=1}^n \{ \Xi(t_i; \theta) - x(t_i) \}^2]$$

推定量 $\hat{\theta}$ の統計学的性質については以下で展開する "直感的" な議論より、一定の正則条件下で漸近正規性が成立する。

簡単化のため $p=1$, 即ち 1-パラメータの場合を考える。まず、"擬似" スコア関数 ("Pseudo" score function) を以下のように定義する、

$$(2-17) \quad \begin{aligned} \psi(t, x(t), \theta) &= \frac{1}{2} \frac{\partial}{\partial \theta} [\Xi(t; \theta) - x(t)]^2 \\ &= \left[\frac{\partial}{\partial \theta} \Xi(t; \theta) \right] [\Xi(t; \theta) - x(t)] \end{aligned}$$

確率変数としての疑似スコア関数 $\psi(t, X(t), \theta)$ の平均と分散はそれぞれ、

$$E\{\psi(t, X(t), \theta)\} = 0$$

$$\text{Var}\{\psi(t, X(t), \theta)\} = \sigma^2 \left[\frac{\partial}{\partial \theta} \Xi(t; \theta) \right]^2$$

となることは (2-15) から直ちに判る。

さて、"データ" $\{x(t_i) = x_{t_i} \ i=1, \dots, n\}$ が与えられたとき、 $\hat{\theta}$ が Θ の内点であれば $\hat{\theta}$ は以下の方程式の解となる ;

$$(2-18) \quad \sum_{i=1}^n \psi(t_i, x(t_i), \hat{\theta}) = 0$$

また、推定量 $\hat{\theta}_n$ を $\sum_{i=1}^n \psi(t_i, X(t_i), \hat{\theta}_n) = 0$ で定義し、以下、 $\hat{\theta}_n$ の漸近正規性を示そう。その為に先ず、 $\psi(t_i, X(t_i), \hat{\theta}_n)$ を真のパラメータ値 θ の周りで Taylor 展開すると、

$$(2-19) \quad \begin{aligned} 0 &= \sum_{i=1}^n \psi(t_i, X(t_i), \hat{\theta}_n) \\ &= \sum_{i=1}^n [\psi(t_i, X(t_i), \theta) + (\hat{\theta}_n - \theta) \frac{\partial}{\partial \theta} \psi(t_i, X(t_i), \theta) + (1/2) (\hat{\theta}_n - \theta)^2 \frac{\partial^2}{\partial \theta^2} \psi(t_i, X(t_i), \theta^*)] \end{aligned}$$

但し、 $|\theta^* - \theta| < |\hat{\theta}_n - \theta|$ 。これより、

$$(2-20) \quad \sqrt{n}(\hat{\theta}_n - \theta) = \frac{\frac{1}{\sqrt{n}} \sum_{i=1}^n \psi(t_i, X(t_i), \theta)}{\frac{1}{n} \sum_{i=1}^n \frac{\partial}{\partial \theta} \psi(t_i, X(t_i), \theta)} + O_p(\sqrt{n}(\hat{\theta}_n - \theta)^2) \quad \text{as } |\hat{\theta}_n - \theta| \xrightarrow{p} 0$$

次なるステップのために以下の正則条件を考える。

[仮定 2-2 : 漸近正規性のための正則条件]

(a). $\{\psi(t_i, X(t_i), \theta), i = 1, \dots, n\}$ は Lindberg-Feller の中心極限定理に従う。

(b). $\hat{\theta}_n \rightarrow \theta$ in probability (as $n \rightarrow \infty$)

(c). $\frac{1}{n} \sum_{i=1}^n \left[\frac{\partial}{\partial \theta} \Xi(t_i, \theta) \right]^2 \rightarrow A > 0$ (as $n \rightarrow \infty$)

(d). $\frac{1}{n} \sum_{i=1}^n \left[\frac{\partial}{\partial \theta} \Xi(t_i, \theta) \right] \left[\frac{\partial^2}{\partial \theta^2} \Xi(t_i, \theta) \right] \rightarrow B \neq 0$ (as $n \rightarrow \infty$)

(e). $\sqrt{n}(\hat{\theta}_n - \theta)^2 \rightarrow 0$ in probability (as $n \rightarrow \infty$). //

ここで上記の正則条件下で次の定理が成立する。

[定理 2-3] [仮定 2-1, 2-2]のもとで

$\sqrt{n}(\hat{\theta}_n - \theta)$ は平均 0、分散 $\sigma^2 A^{-1}$ の正規分布に分布収束する。 //

[証明] 以下、証明をスケッチしていく。先ず、(2-14)より

$$(2-21) \quad E\left\{ \frac{\partial}{\partial \theta} \psi(t_i, X(t_i), \theta) \right\} = \left[\frac{\partial}{\partial \theta} \Xi(t_i, \theta) \right]^2$$

$$(2-22) \quad E\left\{ \frac{\partial^2}{\partial \theta^2} \psi(t_i, X(t_i), \theta^*) \right\} = \left[\frac{\partial}{\partial \theta} \Xi(t_i, \theta) \right] \left[\frac{\partial^2}{\partial \theta^2} \Xi(t_i, \theta) \right]$$

大数の法則と[仮定 2-2] (c),(d),(e)より $n \rightarrow \infty$ の時、 $\frac{1}{n} \sum_{i=1}^n \frac{\partial}{\partial \theta} \psi(t_i, X(t_i), \theta)$ は漸近的に

$\frac{1}{n} \sum_{i=1}^n \left[\frac{\partial}{\partial \theta} \Xi(t_i, \theta) \right]^2$ と等しくなり A に確率収束する。議論は若干ややこしくなるが、

$\frac{1}{n} \sum_{i=1}^n \frac{\partial^2}{\partial \theta^2} \psi(t_i, X(t_i), \theta^*)$ も同様にして B へ確率収束する。即ち、

$$(2-23) \quad \frac{1}{n} \sum_{i=1}^n \frac{\partial}{\partial \theta} \psi(t_i, X(t_i), \theta) \sim \frac{1}{n} \sum_{i=1}^n \left[\frac{\partial}{\partial \theta} \Xi(t_i, \theta) \right]^2 \rightarrow A$$

$$(2-24) \quad \frac{1}{n} \sum_{i=1}^n \frac{\partial^2}{\partial \theta^2} \psi(t_i, X(t_i), \theta^*) \sim \frac{1}{n} \sum_{i=1}^n \left[\frac{\partial}{\partial \theta} \Xi(t_i, \theta) \right] \left[\frac{\partial^2}{\partial \theta^2} \Xi(t_i, \theta) \right] \rightarrow B$$

一方、Lindberg-Feller の中心極限定理より

$$(2-25) \quad \frac{\sum_{i=1}^n \psi(t_i, X(t_i), \theta)}{\sigma \sqrt{\sum_{i=1}^n \left[\frac{\partial}{\partial \theta} \Xi(t_i, \theta) \right]^2}} \xrightarrow{d} N(0,1)$$

以上を組み合わせ、(2-20)より

$$\begin{aligned}
 \sqrt{n}(\hat{\theta}_n - \theta) &= \frac{\frac{1}{\sqrt{n}} \sum_{i=1}^n \Psi(t_i, X(t_i), \theta)}{\frac{1}{n} \sum_{i=1}^n \frac{\partial}{\partial \theta} \Psi(t_i, X(t_i), \theta)} + O_p(\sqrt{n}(\hat{\theta}_n - \theta)^2) \\
 (2-26) \qquad &= \frac{\sum_{i=1}^n \Psi(t_i, X(t_i), \theta) / \sigma \sqrt{\sum_{i=1}^n \left[\frac{\partial}{\partial \theta} \Xi(t_i, \theta) \right]^2}}{\frac{1}{n} \sum_{i=1}^n \frac{\partial}{\partial \theta} \Psi(t_i, X(t_i), \theta) / \sigma \sqrt{\frac{1}{n} \sum_{i=1}^n \left[\frac{\partial}{\partial \theta} \Xi(t_i, \theta) \right]^2}} \\
 &\quad + \frac{1}{2} \sqrt{n}(\hat{\theta}_n - \theta)^2 \frac{\frac{1}{n} \sum_{i=1}^n \frac{\partial^2}{\partial \theta^2} \Psi(t_i, X(t_i), \theta^*) / \frac{1}{n} \sum_{i=1}^n \left[\frac{\partial}{\partial \theta} \Xi(t_i, \theta) \right] \left[\frac{\partial^2}{\partial \theta^2} \Xi(t_i, \theta) \right]}{\frac{1}{n} \sum_{i=1}^n \frac{\partial}{\partial \theta} \Psi(t_i, X(t_i), \theta) / \frac{1}{n} \sum_{i=1}^n \left[\frac{\partial}{\partial \theta} \Xi(t_i, \theta) \right] \left[\frac{\partial^2}{\partial \theta^2} \Xi(t_i, \theta) \right]}
 \end{aligned}$$

上式最右辺第一項の分子は Lindberg-Feller の中心極限定理により標準正規分布に分布収束する。分母は(2-23)より \sqrt{A} に収束。一方、第二項は $O_p\left(0\sqrt{n}(\hat{\theta}_n - \theta)^2\right)$ となり、仮定よりゼロに確率収束する。以上より、

$$(2-27) \qquad \sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{d} N(0, \sigma^2 A^{-1})$$

が導かれる。//

この様に定理自体は標準的な手法で証明される。実際に応用する場面では[仮定2-2]のチェックが問題となる。[定理2-3]より θ の $(1-\varepsilon) \times 100\%$ 信頼区間は近似的に

$$(2-28) \qquad \left(\hat{\theta}_n - z_{\varepsilon/2} \sigma \sqrt{\frac{1}{nA}}, \hat{\theta}_n + z_{\varepsilon/2} z_{\varepsilon/2} \sigma \sqrt{\frac{1}{nA}} \right)$$

で与えられる、ここで $z_{\varepsilon/2}$ は標準正規分布の上側 $\frac{\varepsilon}{2}$ 点である。 σ の推定値は

$$\frac{1}{n-d} \sum_{i=1}^n \{ \Xi(t_i; \hat{\theta}_n) - x(t_i) \}^2 \text{ で与える事が合理的であろう。}$$

これより、観測された Implied 倒産確率を用いたパラメトリックモデルの推定と推定量の漸近正規性が示され、その意味において推定値(2-15)は正当性を主張できる。勿論、仮定2-2の十分条件を求めることが現実には必要である。

第 2 - 4 Implied 倒産確率の推定 : Bootstrap 法による推測

前節では中心極限定理を用い未知パラメータの信頼区間を求めたが、本節では Bootstrap (標本からの再標本法) を用い同様のことを行う (cf. Efron[1979],[1981])。手順は非常に単純である。ここでも前節と同様パラメトリックな確率モデルを仮定する。パラメトリックモデルへの限定は単に以下の形で Bootstrap 法を用い、未知パラメータの推定と精度の推定の議論をするためである。本節における Bootstrap 法は標準的な回帰モデルに於ける Bootstrap 法の準用である。その手順は次のステップから成る :

[Bootstrap 法]

1st Step: 先ず、与えられた疑似データ $\{s(t_i) = s_{t_i} \ i=1, \dots, n\}$ より θ の推定値 $\hat{\theta}$ を判定基準(2-15)のもとで (非線形最小二乗法を用い) 推定する : 参考のため(2-15)を再度示す。

$$(2-29) \quad \hat{\theta} = \arg_{\theta \in \Theta} [\min \sum_{i=1}^n \{ \Xi(t_i; \theta) - x(t_i) \}^2]$$

2nd Step: 次に、[仮定 2-1] (2-14)式中有る ε_{t_i} を $e_i (= e_{t_i}) = x(t_i) - \Xi(t_i; \hat{\theta})$ で推定、残差ベクトル $\mathbf{E} = \{e_1, \dots, e_n\}$ を構築する。これからの所謂 Bootstrap 法のメインパートである。

3rd Step: \mathbf{E} の要素から繰り返しを許し無作為に n 個の e_i を選び、それを

$$(2-30) \quad \mathbf{E}^{*(1)} = \{e^{*(1)}_1, \dots, e^{*(1)}_n\}$$

と書く。 $\{e^{*(1)}_1, \dots, e^{*(1)}_n\}$ はトータルでは n^n 個の組み合わせの中から選ばれる。その中で極端なケースとしては全て異なる場合 (順列を除けば元のデータと同じ中身) もあれば、他方では全ての要素が等しくなる場合があり得る。

4th Step: Bootstrap 標本を以下のように構築する :

$$(2-31) \quad \begin{aligned} x^{*(1)}(t_i) &= x^{*(1)}_i = Q\{\tau > t_i \mid \hat{\theta}\} + e^{*(1)}_i \\ &= \Xi(t_i, \hat{\theta}) + e^{*(1)}_i \quad i = 1, \dots, n \end{aligned}$$

5th Step: 疑似データ (2-31) に基づく θ の推定値を

$$(2-32) \quad \hat{\theta}^{*(1)} = \arg_{\theta \in \Theta} [\min \sum_{i=1}^n \{ \Xi(t_i; \theta) - x^{*(1)}(t_i) \}^2]$$

で求める。

6th Step: Step 3 ~ Step 5 (2-30) ~ (2-32) を M 回繰り返し θ の推定値の集合

$$(2-33) \quad \{\hat{\theta}^{*(1)}_1, \dots, \hat{\theta}^{*(M)}_n\}$$

を得る（ $M=10000$ 程度とれば、可成り精密なヒストグラムを描くことが出来る。計算時間については非線形最小自乗法の数値計算に大きく依存する）。

ここで、 θ の **Bootstrap 推定値** を以下の式で定義する：

$$(2-34) \quad \tilde{\theta}^{(\text{Boots})} = \frac{1}{M} \sum_{k=1}^M \hat{\theta}^{*(k)}$$

Bootstrap バイアス推定値を

$$(2-35) \quad \widetilde{\text{Bias}}(\hat{\theta})^{(\text{Boots})} = \tilde{\theta}^{(\text{Boots})} - \hat{\theta}$$

一方、**Bootstrap 分散推定値**は

$$(2-36) \quad \text{Var}(\hat{\theta})^{(\text{Boots})} = \frac{1}{M-1} \sum_{k=1}^M (\hat{\theta}^{*(k)} - \tilde{\theta}^{(\text{Boots})})^2$$

そして、**推定量の Bootstrap 推定された分布関数**は

$$(2-37) \quad \text{Dist}_{\hat{\theta}}(t)^{(\text{Boots})} = \frac{1}{M} \{\# \text{ of } k \text{ such that } \hat{\theta}^{*(k)} \leq t\}$$

で求められる。

しかし、 θ の信頼区間を構築するには $\{\hat{\theta}^{*(1)}, \dots, \hat{\theta}^{*(M)}\}$ のヒストグラムから直接求める方が簡単であろう。

第3部 Real Probabilityのもとでの倒産確率

第3部では現実の確率のもとでの倒産確率の推定について考える。倒産確率は当該企業の信用レベル、即ち”格付け”により異なる為、各格付けクラス毎の倒産確率を求める必要がある。又、例えばAAAに格付けされている企業がAAAのまま倒産することは通常無い。AAAに限らず、多くの高い格付けクラスの企業は同じクラスに居る時に倒産するより、先ず下のクラスへ降格し、その後倒産することが多い(勿論、逆に格上げされ倒産しないケースもある)。従って、我々の目の前にある倒産データは、所謂”センサーされたデータ(Censored Data)”となる。即ち、得られるデータは、観測期間中に倒産する場合、倒産しない場合だけでなく、他の格付けに移るという場合も考慮しなければならない。この様な、データを扱う一般的な方法として近年盛んに利用されているのが生存分析からの各種方法である。本論文では、その中でもD.R. Cox [1972]により提唱された比例ハザードモデルとその変形を中心に論じていく。第1-3節で論じた様に、比例ハザードモデルではハザード関数をベースラインと各クラスに特徴的な部分との積に分解する。以下ではベースラインについては主にパラメトリックなモデルを仮定する。そこで、次の小節では先ず始めに、パラメトリックモデルと推定量の統計学的性質について解説し、その後所謂Coxモデルに於ける統計学的諸性質をBootstrap法に基づき求める。

第3部でもデータ等に関する記号はこれまで通りとする。以下(1-24,25)式で定義された比例ハザードモデルの推定について論じていく。参考のためにモデル(1-24),(1-25)を再度示しておく：

[比例ハザードモデル] ある共変量(Covariate)ベクトル \mathbf{x} と回帰係数ベクトル $\boldsymbol{\beta}$ に対し、ハザード関数が

$$(1-24) \quad \lambda(t; \mathbf{x}, \boldsymbol{\beta}) = \lambda_0(t) \exp\{\boldsymbol{\beta}'\mathbf{x}\}, \quad \boldsymbol{\beta} = (\beta_1, \dots, \beta_p)', \quad \mathbf{x} = (x_1, \dots, x_p)'$$

である生存時間モデルを比例ハザードモデルと呼ぶ。ここで、 $\lambda_0(t)$ はベースラインハザードと呼ばれる共通項である。簡単な積分計算より生存関数は

$$(1-25) \quad \begin{aligned} S(t, \mathbf{x}) &= \exp\{-\int_0^t \lambda(u; \mathbf{x}) du\} = \exp\{-e^{\boldsymbol{\beta}'\mathbf{x}} \int_0^t \lambda_0(u) du\} \\ &= \exp\{-\int_0^t \lambda_0(u) du\} \exp\{e^{\boldsymbol{\beta}'\mathbf{x}}\} = S_0(t) \exp(\boldsymbol{\beta}'\mathbf{x}) \end{aligned}$$

で与えられる。//

第3-1節：回帰係数 β の推定：ノンパラメトリック法

Coxの興味は β に関する検定にあるようだが本稿では推定問題に焦点を合わせていく。この問題と解法に対するCoxの基本的な姿勢は「 β に関する検定問題では(尤度比検定等を用いる場合)ベースラインハザード関数 λ_0 は任意である。極端に言えば倒産時点(failure time)以外ではそれが恒等的にゼロであっても構わない。倒産時点やセンサーされた時刻のみが β に関する情報を持っている。従って β に関する推論は倒産時点やセンサー時刻が与えられたという条件のもとで行えばよい」である(cf. Cox [1972])。そこで、以下の”条件付き尤度関数(Conditional Likelihood Function)”が定義され、それに基づいた推測を行うことが正当化される。推定問題でもベースラインハザード関数と β を別々に推定するならば同様の議論が成立すると考えられる為、本稿でも条件付き尤度関数に基づいた議論を展開する。

[条件付き尤度関数]以下簡単化のため、データ $\mathbf{Y} = \{Y_1, Y_2, \dots, Y_n\}$ は全て異なる値を取るものとし(No tied data)、その順序統計量を $\{Y_{(1)}, Y_{(2)}, \dots, Y_{(n)}\}$ と書く。又、これまでと同様に小文字表現で観測値を表すこととする。 $\delta_{(i)}$ を時点 $Y_{(i)}$ に対応するセンサー指示関数、 $\mathbf{x}_{(i)}$ は対応する共変量ベクトルとする。 $Y_{(i)}$ に対応するリスク集合、 $\mathbf{R}(Y_{(i)}) = \{\text{時点 } Y_{(i)}\text{ まで倒産していない(生き残っている)企業の集合}\}$ を簡単のため $\mathbf{R}_{(i)}$ と書く。この時、任意のセンサーされていない $Y_{(i)}$ に対して

$$\begin{aligned} &Pr\{\text{区間 } [Y_{(i)}, Y_{(i)} + \Delta y] \text{ で倒産する} \mid \mathbf{R}_{(i)}\} \\ (3-1) \quad &\doteq \sum_{j \in \mathbf{R}_{(i)}} \exp\{\beta' \mathbf{x}_j\} \lambda_0(Y_{(i)}) \Delta y \end{aligned}$$

が成立している事が示される。これより、

$$\begin{aligned} &Pr\{\text{会社}(i) \text{ が時点 } y_{(i)} \text{ で倒産する} \mid \text{時点 } y_{(i)} \text{ で } \mathbf{R}_{(i)} \text{ 内の } 1 \text{ 社が倒産する}\} \\ (3-2) \quad &= \exp\{\beta' \mathbf{x}_{(i)}\} / \sum_{j \in \mathbf{R}_{(i)}} \exp\{\beta' \mathbf{x}_j\} \end{aligned}$$

そして、センサーされていない全てのデータに対する上式の積 $L_c(\beta)$ を条件付き尤度関数と呼ぶ。

$$(3-3) \quad L_c(\beta) = \prod_u \exp\{\beta' \mathbf{x}_{(i)}\} / \sum_{j \in \mathbf{R}_{(i)}} \exp\{\beta' \mathbf{x}_j\}$$

ただし、 \prod_u はセンサーされていないデータに関する積である。//

一方、センサーされたデータを含む場合はどうなるか？ そのために(3-3)の意味を別の側面から考えてみる。

Y_1, Y_2, \dots, Y_n は独立で $Y_i \sim F_i$ 、 R_i を Y_i のランク ($i=1, \dots, n$) とする。ここでも $\mathbf{Y} = (Y_1, Y_2, \dots, Y_n)$ 、 $\mathbf{R} = (R_1, R_2, \dots, R_n)$ とし、小文字で観測値 (ベクトル) を表す。 f_i を F_i の密度関数とすれば

$$(3-4) \quad p(\mathbf{r}) = Pr\{\mathbf{R} = \mathbf{r}\} = \int_{\mathbf{u}_1 < \mathbf{u}_2 < \dots < \mathbf{u}_n} \dots \int \prod_{i=1}^n f_{(i)}(\mathbf{u}_i) d\mathbf{u}_i$$

ただし、 $f_{(i)}$ は順序統計量 $Y_{(i)}$ の密度関数である。ここで以下の結果が Kalbfleisch and Prentice[1973]により与えられている

[定理 3-1] もしも、 $F_i(t) = 1 - \exp\{-\exp\{\boldsymbol{\beta}'\mathbf{x}_i \int_0^t \lambda_0(u) du\}\}$ であれば、

$$(3-5) \quad p(\mathbf{r}) = \prod_{i=1}^n \exp\{\boldsymbol{\beta}'\mathbf{x}_{(i)}\} / \sum_{\mathbf{j} \in \mathbf{R}(\mathbf{i})} \exp\{\boldsymbol{\beta}'\mathbf{x}_{\mathbf{j}}\} //$$

即ち、Cox の条件付き尤度関数はランク統計量の密度 (確率) 関数に他ならない。そこで、この考えをセンサーされたデータに拡張するために、先ず、センサーされたデータセットにおける順序統計量を以下のように定義する：

[定義 3-2 : センサーされた観測値のランク] $(Y_1, \delta_1), (Y_2, \delta_2), \dots, (Y_n, \delta_n)$ をセンサーされたデータセット、 Y_i は分布関数 F_i (密度関数 f_i) に従うとする。ここで、センサーされたデータに対するランクベクトルを以下のように定義する：

$$(3-6) \quad \mathbf{R}^{w/c} = (R^{w/c}_1, \dots, R^{w/c}_n)$$

ただし、

$$(3-7) \quad R^{w/c}_i = \begin{cases} \text{センサーされていないデータ内の } Y \text{ のランク} & \delta_i = 1 \\ \text{直前のセンサーされていないデータのランク} & \delta_i = 0 \end{cases}$$

さて、 $\boldsymbol{\delta} = (\delta_1, \delta_2, \dots, \delta_n)$ と $\mathbf{r}^{w/c} = (r^{w/c}_1, \dots, r^{w/c}_n)$ を実現値のベクトルとするとき

$$(3-8) \quad \begin{aligned} p(\mathbf{r}^{w/c}, \boldsymbol{\delta}) &= Pr\{\mathbf{R}^{w/c} = \mathbf{r}^{w/c}, \boldsymbol{\delta} = \boldsymbol{\delta}\} \\ &= \int_{\mathbf{u}_1 < \mathbf{u}_2 < \dots < \mathbf{u}_n} \dots \int \left[\left\{ \prod_{i=1}^{n(u)} f_{u(i)}(\mathbf{u}_i) d\mathbf{u}_i \right\} \left\{ \prod_{\mathbf{j} \in \mathbf{C}_{i,i+1}} [1 - F_{\mathbf{j}}(\mathbf{u}_{\mathbf{j}})] \right\} \right] d\mathbf{u}_1 \dots d\mathbf{u}_{n(u)} \end{aligned}$$

ただし、 $u(i)$ は i 番目のセンサーされていないデータ、 $n(u)$ はセンサーされていないデータの総数。 $\mathbf{C}_{i,i+1}$ はセンサーされていない i 番目と $i+1$ 番目のデータの間に含まれるセンサーされたデータの集合である。ここでも上と同様の結果が Kalbfleisch and Prentice[1973]

により証明されている：

[定理 3-3] もしも、 $F_i(t) = 1 - \exp\{-\exp\{\beta' \mathbf{x}_i \int_0^t \lambda_0(u) du\}\}$ であれば、

$$(3-9) \quad p(\mathbf{r}^{wc}, \delta) = \prod_u \exp\{\beta' \mathbf{x}_{(i)}\} / \sum_{j \in R(i)} \exp\{\beta' \mathbf{x}_j\} = L_c(\beta) //$$

即ち(3-3) で定義された条件付き尤度関数はセンサーされたデータに関してもそのまま用いることが出来るのである！Cox[1972] は上記 $L_c(\beta)$ を、あたかも通常の尤度関数と同じに扱って良いと予想し、 β の MLE も標準的な手法で求められると予想した。実際、この予想は後日、Prentice and Kalibfleisch[1979]等により厳密な形で証明されている。従って β の最尤推定量を求めるには、スコアベクトルと標本情報行列を用いた数値計算を行えばよい。具体的には先ず、スコアベクトル $\frac{\partial}{\partial \beta} \log L_c(\beta)$ と標本情報行列 $\mathbf{i}(\beta)$ を定義する。

$$(3-10) \quad \frac{\partial}{\partial \beta} \log L_c(\beta) = \left[\frac{\partial}{\partial \beta_1} \log L_c(\beta), \frac{\partial}{\partial \beta_2} \log L_c(\beta), \dots, \frac{\partial}{\partial \beta_p} \log L_c(\beta) \right]$$

$$(3-11) \quad \mathbf{i}(\beta) = - \begin{bmatrix} \frac{\partial^2}{\partial \beta_1 \partial \beta_1} \log L_c(\beta) & \cdots & \frac{\partial^2}{\partial \beta_1 \partial \beta_p} \log L_c(\beta) \\ \vdots & \ddots & \vdots \\ \frac{\partial^2}{\partial \beta_1 \partial \beta_p} \log L_c(\beta) & \cdots & \frac{\partial^2}{\partial \beta_p \partial \beta_p} \log L_c(\beta) \end{bmatrix}$$

第 1 部 5 節のパラメトリックモデルの時と同様 $\frac{\partial}{\partial \beta} \log L_c(\beta) = \mathbf{0}$ を解くことで β の最尤推定量を求めるが、ここでも一般的には Newton-Raphson 法が用いられる。

先ず $\hat{\beta}^{(1)}$ を初期値 (初期予測ベクトル) とする。そして、順次、

$$(3-12) \quad \hat{\beta}^{(k)} = \hat{\beta}^{(k-1)} + \mathbf{i}(\hat{\beta}^{(k-1)})^{-1} \frac{\partial}{\partial \beta} \log L_c(\hat{\beta}^{(k-1)}) \quad k \geq 2$$

で次の推定値を順次求め、安定的な推定値が得られるまで続ける。

β の推定のあと、ベースラインハザード関数の推定を行うが、本稿では利便性の観点よりパラメトリックモデルを仮定する。従って問題は未知パラメータの推定となり、基本的には第 1-5 節の方法で推定される。ただし、ここで利用するデータは共変量を含む $(Y_1, \mathbf{x}_1, \delta_1), (Y_2, \mathbf{x}_2, \delta_2), \dots, (Y_n, \mathbf{x}_n, \delta_n)$ である。 $(Y_i, \mathbf{x}_i, \delta_i)$ の実現値 $(y_i, \mathbf{x}_i, \delta_i)$ の尤度関数 $L_i(\theta) = L_\theta(y_i, \mathbf{x}_i, \delta_i)$ は

$$(3-13) \quad L_i(\theta) = L_\theta(y_i, \mathbf{x}_i, \delta_i) = \begin{cases} f(y_i, \mathbf{x}_i; \theta) & \text{if } \delta_i = 1 \quad (\text{uncensored}) \\ S(y_i, \mathbf{x}_i; \theta) & \text{if } \delta_i = 0 \quad (\text{censored}) \end{cases}$$

$$= f(y_i, \mathbf{x}_i; \boldsymbol{\theta})^{\delta_i} S(y_i, \mathbf{x}_i; \boldsymbol{\theta})^{1-\delta_i}$$

但し、

$$S(y, \mathbf{x}; \boldsymbol{\theta}) = \exp\left\{-\int_0^y \lambda_0(u; \boldsymbol{\theta}) du \cdot \exp(\widehat{\boldsymbol{\beta}}' \mathbf{x})\right\}$$

(3-14)

$$f(y, \mathbf{x}; \boldsymbol{\theta}) = \lambda_0(y; \boldsymbol{\theta}) \cdot \exp(\widehat{\boldsymbol{\beta}}' \mathbf{x}) \cdot \exp\left\{-\int_0^y \lambda_0(u; \boldsymbol{\theta}) du \cdot \exp(\widehat{\boldsymbol{\beta}}' \mathbf{x})\right\}$$

又ここでは $\boldsymbol{\beta}$ の代わりに(最尤)推定値 $\widehat{\boldsymbol{\beta}}$ を用いる。仮に、ベースラインとしてワイブルモデルを仮定すると、 $S_0(t, \mathbf{x}_1; \boldsymbol{\theta}) = \exp\{-(\vartheta t)^\alpha\}$ 、 $f(t, \mathbf{x}_1; \boldsymbol{\theta}) = \alpha \vartheta (\vartheta t)^{\alpha-1} \exp\{-(\vartheta t)^\alpha\}$ 、 $\lambda(t) = \alpha \vartheta (\vartheta t)^{\alpha-1}$ より、

$$L_i(\boldsymbol{\theta}) = f(y_i, \mathbf{x}_i; \boldsymbol{\theta})^{\delta_i} S(y_i, \mathbf{x}_i; \boldsymbol{\theta})^{1-\delta_i}$$

(3-15)

$$= [\alpha \vartheta (\vartheta t)^{\alpha-1} \exp\{-(\vartheta y_i)^\alpha\} \cdot \exp(\widehat{\boldsymbol{\beta}}' \mathbf{x}) \cdot \exp\{-(\vartheta y_i)^\alpha\} \cdot \exp(\widehat{\boldsymbol{\beta}}' \mathbf{x})]^{\delta_i} [\exp\{-(\vartheta t)^\alpha\} \exp(\widehat{\boldsymbol{\beta}}' \mathbf{x})]^{1-\delta_i}$$

ただし、 $\boldsymbol{\theta} = (\vartheta, \alpha)$ である。(3-13)より $\boldsymbol{\theta}$ に関する尤度関数は

$$L(\boldsymbol{\theta}) = L(y_1, y_2, \dots, y_n; x_1, x_2, \dots, x_n, \delta_1, \delta_2, \dots, \delta_n; \boldsymbol{\theta})$$

(3-16)

$$= \prod_{i=1}^n L_{\theta}(y_i, x_i, \delta_i) = [\prod_u f(y_i)^{\delta_i}] [\prod_c S(y_i)^{1-\delta_i}]$$

これより、1-5節の方法で $\boldsymbol{\theta}$ の最尤推定量 $\widehat{\boldsymbol{\theta}} = \boldsymbol{\theta}(\widehat{\boldsymbol{\beta}})$ を求めることが可能である。

第3-2節 Bootstrap 法

モデル内のパラメータを推定した後必要なことは推定値の精度をどのように評価するかである。最尤推定量は、一定の正則性の条件下で漸近正規性を持っているが、前節の様な方法で β とベースラインハザード関数のパラメータ θ を求めた場合、標準的な方法による誤差評価の正当性は疑問である。そこで、ここでも第2-4節同様 Bootstrap 法を適用する。比例ハザードモデルに於ける Bootstrap 法はこれまでも幾つかの方法が提案されてきているが本節では、①パラメトリック Bootstrap (Cf. Casellas, Tarres and Varona [2006]) の拡張と②Zelteman, Le and Louis [1996] による Bayesian Bootstrap 法の変形、の二方法を紹介する。しかしながら、Casellas et. al の方法ではセンサーされたデータに対する Bootstrap は扱えない。そこで本節では Efron [1981] によるランダムセンサーリングの考えを利用することにより、これを可能とした。

第3-2-1節 パラメトリック Bootstrap 法

パラメトリック Bootstrap 法とは、先ず原データより未知パラメータ θ を推定、推定値を $\hat{\theta}$ と記す。次に推定されたパラメータの元で推定された分布関数 $F(t)=F(t|\hat{\theta})$ から再標本となる確率変数を発生させる方法の通称である。通常 Bootstrap が原データから直接再標本するのに対し、ここでは推定された分布関数より新たなデータを生み出すところが異なる。形式的には、 T_1, T_2, \dots, T_n を $F(t|\hat{\theta})$ からの大きさ n のランダムサンプルで t_1, t_2, \dots, t_n をその実現値とする。又、 $\hat{\theta} = \theta(t_1, t_2, \dots, t_n)$ を θ の推定値とする。ここで、再標本列 $(T^*_1, T^*_2, \dots, T^*_n)$ は $F(t|\hat{\theta})$ からの大きさ n のサンプルである。具体的には n 個の独立な $(0,1)$ の一様乱数 (U_1, U_2, \dots, U_n) を発生させ $T^*_i = F^{-1}(U_i | \theta)$ とすればよい。後の手続きは2-4節と同様である。第1番目の再標本列 $(T^{*(1)}_1, T^{*(1)}_2, \dots, T^{*(1)}_n)$ から推定値 $\hat{\theta}^{*(1)}$ を求める。

以下同様に第 k 番再標本列 $(T^{*(k)}_1, T^{*(k)}_2, \dots, T^{*(k)}_n)$ から推定値 $\hat{\theta}^{*(k)}$ を順次求め $(k=1, \dots, M)$ 、bootstrap 推定値の列 $\{\hat{\theta}^{*(1)}, \dots, \hat{\theta}^{*(M)}\}$ を得る。ここでも上と同様 θ の Bootstrap 推定値を以下の式で定義する：

$$(3-17) \quad \tilde{\theta}^{(Boots)} = \frac{1}{M} \sum_{k=1}^M \hat{\theta}^{*(k)}$$

Bootstrap バイアス推定値

$$(3-18) \quad \widetilde{Bias}(\hat{\theta})^{(Boots)} = \tilde{\theta}^{(Boots)} - \hat{\theta}$$

Bootstrap 分散推定値

$$(3-19) \quad \text{Var}(\hat{\theta})^{(Boots)} = \frac{1}{M-1} \sum_{k=1}^M (\hat{\theta}^{*(k)} - \tilde{\theta}^{(Boots)})^2$$

推定量の Bootstrap 推定された分布関数

$$(3-20) \quad \text{Dist}_{\hat{\theta}}(t)^{(\text{Boots})} = \frac{1}{M} \{\# \text{ of } k \text{ such that } \hat{\theta}^{*(k)} \leq t\}$$

で求められる。

比例ハザードモデルのパラメトリック Bootstrap を行う際に、例えばワイブルハザード関数を仮定すると、倒産時刻 T の分布関数は

$$(3-21) \quad F(t; \mathbf{x}, \boldsymbol{\beta} | \vartheta, \alpha) = [1 - \exp\{- (\vartheta t)^\alpha \exp(\boldsymbol{\beta}' \mathbf{x})\}]$$

となる。 U を一様確率変数とするとき倒産確率変数 T^* は

$$(3-22) \quad \begin{aligned} T^* &= F^{-1}(U; \mathbf{x}, \boldsymbol{\beta} | \hat{\vartheta}, \hat{\alpha}) \\ &= \left(-\log[1 - U] / [\exp\{\hat{\alpha} \log(\hat{\vartheta}) + \boldsymbol{\beta}' \mathbf{x}\}] \right)^{\frac{1}{\hat{\alpha}}} \end{aligned}$$

で与えられる。従って、独立な $(0,1)$ 上の一様乱数 (U_1, U_2, \dots, U_n) より、倒産時刻の再標本列 $(T^{*(k)}_1, T^{*(k)}_2, \dots, T^{*(k)}_n)$, $k=1, \dots, M$ を発生させればよい。

一方、センサーされた再標本データの作り方は Efron のランダムセンサーリングの考え方をを用いる。まず、センサー時点 C (これは倒産時刻と観測最終時点でセンサーされている!) の分布を推定。ここでもベースラインハザードにはパラメトリックモデルを用いた比例ハザードモデルを用いると良いだろう。即ち、共変量ベクトル x_i を伴った、データ列 (C_i, x_i, δ'_i) $i=1, \dots, n$, に基づき C の分布関数 $G(c; \mathbf{x}, \phi)$ の未知パラメータ ϕ を推定する、但し δ' は観測最終時点を除き $\delta' = 1 - \delta$ である。次に、上と同様の方法で x_i を与えたうえで C^*_i を発生する。そして、上記 $\{(T^{*(k)}_1, T^{*(k)}_2, \dots, T^{*(k)}_n), k=1, \dots, M\}$ と組み合わせることによりセンサーされたデータ列 $\{(Y^{*(k)}_i, \delta^{*(k)}_i), i=1, \dots, n\}$, $k=1, \dots, M$ を発生させる、但し、

$$(3-23) \quad Y^{*(k)}_i = \min\{T^{*(k)}_i, C^{*(k)}_i\}, \quad \delta^{*(k)}_i = \begin{cases} 1 & \text{if } T^{*(k)}_i \leq C^{*(k)}_i \\ 0 & \text{if } T^{*(k)}_i > C^{*(k)}_i \end{cases}$$

第3-2-2節ノンパラメトリック Bootstrap

ノンパラメトリック比例ハザードモデルでの Bootstrap 標本発生方法は Zelterman, Le and Louis [1996]により与えられている。彼等の考え方は以下の通りである。

まず、ここでは共変量ベクトル \mathbf{x} もランダムベクトルと考え、 $(\mathbf{Y}, \delta, \mathbf{X})$ の同時分布を想定する。Bootstrap 標本についてはこれまで通り $(\mathbf{Y}^*, \delta^*, \mathbf{X}^*)$ と記すことにする。再標本変量を発生する際、 (\mathbf{Y}^*, δ^*) を先ず求め、次に \mathbf{X}^* を (\mathbf{Y}^*, δ^*) が与えられた時の条件付き分布より発生させる。

我々の手元には n 個のデータ $\{(y_{(i)}, \delta_i, \mathbf{x}_i), i=1, \dots, n\}$ がある。オリジナルのデータに関してタイ (Tie)を排除することは可能であったが、Bootstrap でタイは必然であるため、以下の議論ではタイを許容する。そこで、先ず、互いに異なるセンサーされていない y の値を大きさの順に並べたものを

$$(3-24) \quad y_{(1)} < y_{(2)} < \dots < y_{(N)}$$

とする、ここで N は互いに異なる y の数である。さらに、これまでと同様、 $d_i =$ 時点 $y_{(i)}$ で観測されたセンサーされていない観測値の数、 $\mathbf{R}_i = \mathbf{R}(y_{(i)}) =$ 時点 $y_{(i)}$ におけるリスク集合を定義しておく。また、ここでは共変量ベクトル \mathbf{x}_i も確率ベクトルと仮定、その周辺分布を

$$(3-25) \quad \rho_i = \Pr\{\mathbf{x}^* = \mathbf{x}_i\} \quad (\text{ここでは } \rho_i = 1/n \text{ である})$$

とする。Bootstrap の第一ステップは (y^*, δ^*) を n 個のペア $\{(y_1, \delta_1), \dots, (y_n, \delta_n)\}$ から抽出、ただし、当初の仮定よりここでは $\delta^* = 1$ である。次いで共変量ベクトル \mathbf{x} を条件付き確率

$$(3-26) \quad \pi_{ij} = \pi_{ij}(\boldsymbol{\beta}) = \Pr\{\mathbf{x}^* = \mathbf{x}_j \mid (y^*, \delta^*) = (y_{(i)}, 1)\}$$

に従って選択する。詳しい導出プロセスは Zelterman, Le and Louis [1996]にあるが、ベイズの定理により

$$(3-27) \quad \begin{aligned} \pi_{ij} &= \pi_{ij}(\boldsymbol{\beta}) \\ &= [\exp\{\boldsymbol{\beta}'\mathbf{x}_j\} (\delta_i + \rho_j - \sum_{l=1}^{i-1} \pi_{lj}d_l)] / [\sum_{k=1}^n \exp\{\boldsymbol{\beta}'\mathbf{x}_k\} (\delta_i + \rho_k - \sum_{l=1}^{i-1} \pi_{lk}d_l)] \end{aligned}$$

となる、但し、 $\rho_i + \sum_{j=i}^n \delta_j = \sum_{j=i}^n d_j$ である。又ここでは $\sum_{i=1}^0 = 0$ とする。実装の際には $\pi_{ij}(\boldsymbol{\beta})$ は、 $\boldsymbol{\beta}$ にその MLE (乃至は Maximal Partial Likelihood Estimate) $\hat{\boldsymbol{\beta}}$ を代入し、 $i=1$ から順次求められる； $\rho_i = 1/n$ に注意すると、例えば、

$$\begin{aligned}
\pi_{1j} &= \pi_{1j}(\hat{\boldsymbol{\beta}}) = \exp\{\hat{\boldsymbol{\beta}}' \mathbf{x}_j\} / \sum_{k=1}^n \exp\{\hat{\boldsymbol{\beta}}' \mathbf{x}_k\} \\
(3-28) \quad \pi_{2j} &= \pi_{2j}(\hat{\boldsymbol{\beta}}) \\
&= [\exp\{\hat{\boldsymbol{\beta}}' \mathbf{x}_j\} (\delta + \rho_j - d_1 \pi_{1j})] / \sum_{k=1}^n \exp\{\hat{\boldsymbol{\beta}}' \mathbf{x}_k\} / [\sum_{k=1}^n \exp\{\hat{\boldsymbol{\beta}}' \mathbf{x}_j\} (\delta + \rho_k - d_1 \pi_{1j})]
\end{aligned}$$

となる。

一方、センサーされた標本 $(y_{(i)}^*, 0)$ については

$$(3-29) \quad \pi_{ij}^C = \Pr\{x^* = x_j \mid (y^*, \delta^*) = (y_{(i)}, 0)\}$$

を求めなければならない。さて、 $(y^*, \delta^*) = (y_{(i)}, 0)$ が意味することは、時点 $y_{(i)}$ では倒産が起きたか否かは定かでない。唯一確定的なことは、倒産時刻は時点 $y_{(i)}$ より先と云うことである。従って、倒産時刻を T とすると、

$$\begin{aligned}
\pi_{ij}^C &= \Pr\{x^* = x_j \mid T \geq y_{(i)}\} \\
(3-30) \quad &= \sum_{l=u}^N \pi_{lj} d_l / \sum_{l=u}^N d_l
\end{aligned}$$

但し、 $u = u(i)$ は $\min y_u \geq y_{(i)}$ となる最小の u である。

最後に2点ほどコメントを付け加えておく。まず、 $y(N)$ がセンサーされている場合(3-30)式は定義出来ない。そこで、 π_{ij}^C をそれまでで一番生存時間の長いデータの中で、センサーされていない観測値と同じ π_{ij} とする。又、 $y(N)$ がセンサーされていない場合は対応するハザード関数値が1となる。結果として(3-27)の π_{ij} が負又は1以上の値を取る可能性が出てくる。この場合も、上と同様の工夫で π_{ij} を定義する。

[参考文献]

- Casellas, J., J. Tarres and L. Varona [2006] "Parametric Bootstrap for Testing Model Fitting in the Proportional Hazards Framework: An Application to the Survival Analysis of Brunas Dels Pirineus Beef Calves" *American Society of Animal Sciences* 84, 2609-16
- Cox, D.R. [1972] "Regression Models and Life-Tables" *Journal of the Royal Statistical Society, Ser. B* 34, 187-202
- Duffie, D. and K. Singleton [1997] "An Economic Model of the Term Structure of Interest Rate Swap Yields," *Journal of Finance* 52, 1287-1321
- Duffie, D. and K. Singleton [1999] "Modeling Term Structures of Defaultable Bonds" *Review of Financial Studies* 12, 678-720
- Efron, B. [1979] "Bootstrap Methods: Another Look at the Jackknife" *Annals of Statistics* 7, 1-26
- Efron, B. [1980] "Censored Data and the Bootstrap" *Journal of the American Statistical Association*, 76, 312-319
- Efron, B. [1982] *The Jackknife, the Bootstrap and Other Resampling Planes*, SIAM, Philadelphia
- Harrison, M. and D. Kreps [1979] "Martingales and Arbitrage in Multiperiod Security Markets" *Journal of Economic Theory*, 20, 381-408.
- Harrison, M. and S. Pliska [1981] "Martingales and Stochastic Integrals in the Theory of Continuous Trading" *Stochastic Processes and Their Applications*, 11, 215-260.
- Heath, D., R. Jarrow, and A. Morton [1992] "Bond Pricing and the Term Structure of Interest Rates: A New Methodology for Contingent Claims Valuation" *Econometrica*, 60, 77-106
- Kalbfleisch, J.D. and R.L. Prentice [1973] "Marginal Likelihoods Based on Cox's Regression and Life Model" *Biometrika* 60, 267-278
- Kaplan, E.L. and Meier, P. [1958] "Nonparametric Estimation from Incomplete Observations" *Journal of the American Statistical Association*, 53, 457-481
- Merton, R. [1974] "On the Pricing of Corporate Debt: The Risk Structure of Interest Rates" *The*

Journal of Finance 29, 449-470

Miller Jr, R. G. [1981] *Survival Analysis*, John Wiley & Sons New York

Prentice, R.L. and J.D. Kalbfleisch [1979] "Hazard Rate and Models with Covariates,"
Biometrics 35, 25-39

Takahashi H. [2009] "Statistical Analysis of Implied Data" to appear

D. Zelterman, Le, C.T. and T.A. Louis [1996] "Bootstrap Techniques for Proportional Hazards Models with Censored Observations" *Statistics and Computing* 6, 191-199

楠岡成雄、青沼君明、中川秀敏 [2001] *クレジットリスクモデル*, 金融財政事情研究会