

比例ハザードモデルを使った倒産確率推定の 問題について

大 上 慎 吾

1 はじめに

近年、金融不安を背景として、計量的な企業の倒産分析への関心が高まっている。古典的な多変量判別分析や、最近では、ニューラル・ネットワークやデータマイニングの手法を応用した倒産予測など、さまざまな試みがなされている [3] [4]。そうした中で、企業の抱える信用リスクを倒産確率という形で明示的に測定しようという試みがロジット分析やハザードモデルを使った生存分析である。

ハザードモデルは企業の生存時間の確率分布を推定するため、得られた分布を使って、例えば「ある財務状態にある企業が、1年以内には倒産しないが3年後までには倒産している確率」などさまざまな量を計算することができる。したがって、いろいろな状況を想定した倒産分析のツールとして、またクレジット・デリバティブの計算など他の分析のパーツとしても大きなポテンシャルを持っていると考えられる。

ハザードモデルの推定においては、サンプル中の各期間における倒産企業数の全体に占める割合が重要な役割を果たす。したがって、サンプル中に含まれる倒産企業数が非常に少ない場合には、財務状態の非常に悪い企業に対しても、倒産確率が過小に評価されるという問題が起こる。これまで倒産企業の絶対数が少なかった日本のような場合、ハザードモデルを使った分析の成功例が多く見られないことの大きな要因であると考えられる。

本論文では、生存分析の代表的なモデルである比例ハザードモデルを倒産

確率の推定に使う場合のこうした問題点について、実際のデータ分析例を紹介しながら考察していきたい。

2 比例ハザードモデル

ある企業が固定された一時点 ($t=0$ 時点) から倒産に到るまでの時間 T を、密度関数 f 、分布関数 F をもつ確率変数と仮定する。任意の正の数 x について、 $S(x)=1-F(x)=P(T>x)$ はこの企業が x 時間経過しても倒産していない確率であり、サバイバル関数とよばれる。さらに、生存時間 T のハザード関数は

$$\lambda(t) = \lim_{dt \rightarrow 0} \frac{P\{t \leq T < t+dt \mid t \leq T\}}{dt} = \frac{f(t)}{1-F(t)}$$

で定義され、 $\Lambda(t) = \int_0^t \lambda(u) du$ とすると、

$$S(x) = \exp\{-\Lambda(t)\}$$

である。

今、 t 時点での企業の財務状態を表わす説明変数のベクトル $\mathbf{Z}(t)$ が与えられているとする。比例ハザードモデルではハザード関数 λ が次のような特別な形を持っていると仮定する。

$$\lambda(t) = \lambda_0(t) \exp\{\boldsymbol{\beta}^T \mathbf{Z}(t)\} \quad (1)$$

ここで λ_0 は説明変数に依存しない関数であり、ベースライン・ハザード関数とよばれる。 $\boldsymbol{\beta}$ と $\lambda_0(\Lambda_0)$ の推定値については、付録1を参照されたい。

3 財務データ

今回、比例ハザード分析を行うにあたって、上場・非上場企業5,000社の91年10月から96年9月にわたる236,413期間分の財務データを利用した¹⁾。このサンプルの中にはこの期間に倒産した企業173社が含まれている。

分析に使った説明変数は、代表的な財務指標の中から、奥野・山田 [2] によって企業優劣の判別分析に有効であると認められたものを中心に次のものを選択した。これら財務指標の定義は付録2にまとめてある。

1. 収益性

- x1: 総資本利払い後事業利益率 (%)
- x2: 企業利潤率 (%)
- x3: 自己資本利払い後事業利益率 (%)
- x4: 一人当たり利払い後事業利益

2. 安全性

- x5: 売上高純金利負担率
- x6: 借入金依存度
- x7: 自己資本比率
- x8: 流動比率
- x9: 固定長期適合率
- x10: 手元流動性比率
- x11: 固定負債キャッシュフロー倍率
- x12: 有利子負債キャッシュフロー倍率
- x13: インタレスト・カバレッジ・レシオ

3. 生産性

- x14: 一人当たり付加価値額
- x15: 売上高付加価値率

4. 規模

- x16: 売上高の自然対数
- x17: 使用総資本の自然対数
- x18: 従業員数の自然対数

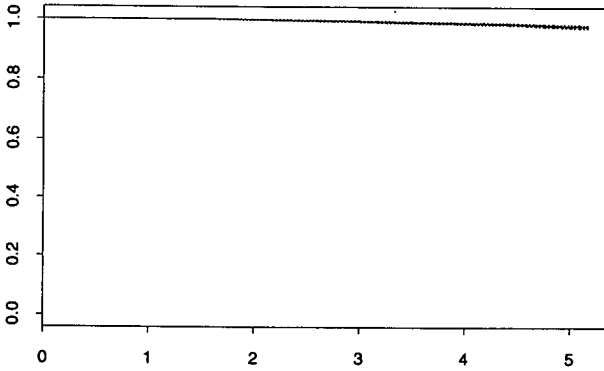
なお、使用総資産 (x17) は他の規模の指標との相関が非常に強いため最終的なモデルには使われなかった (付録 2 参照)。

4 分析結果

比例ハザードモデルのデータへのあてはめには、Splus の coxph 関数を使った。結果は表 1 にまとめられている。

分析に使われた財務指標のうち、5% 水準で有意であるとされたものは、

図1 平均的なサバイバル関数(実線)。破線は5%信頼区間。横軸の単位は年数。

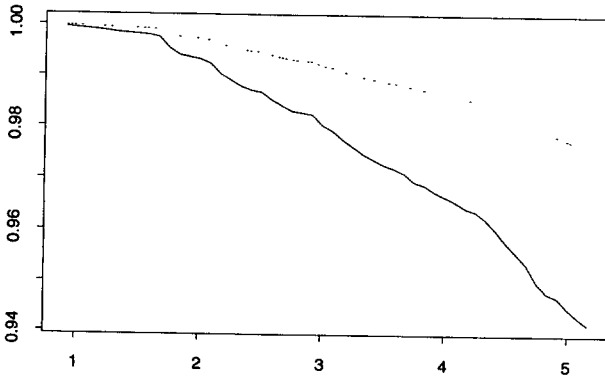


売上高純金利負担率 (x5), 自己資本比率 (x7), 手元流動性比率 (x10), 売上高付加価値率 (x15) であった。倒産の危険性が高い企業に対して融資が行われるとすれば, 当然プレミアムも高いだろうから, 売上高純金利負担率も大きな値となるだろう。また, 自己資本比率と手元流動性比率が各々長期と短期の安全性を測っているとすれば, これらの指標が有意となるのは妥当であると考えられる。一方で売上高付加価値率が有意とされたのは, 安全性が高い企業であっても, 資産を効率的に活用していなければ, 倒産の危険性が高まることを理由としているのであろう。

各々の財務指標について, サンプル中の全ての企業の平均値をとったとき, その推定されたサバイバル関数を描いたものが図1である。年数がかなり進んでも, 生存確率が一向に減少しない様子が見てとれる。

サンプル中の倒産企業のみについて, 同様に各財務指標の平均値をとった場合の推定されたサバイバル関数が図2である。全企業の平均値に比べると生存確率が大きな割合で減少しているが, その開きは5年後においても0.04以下と極めて小さな値である。実際には1年後に倒産しているような企業の財務データを与えても, その5年後までの生存確率は0.94を越えるといった結果になってしまった。

図2 倒産企業の平均的なサバイバル関数(実線)
と全企業の平均的なサバイバル関数(破線)



5 おわりに

4節では、日本のような倒産企業の絶対数が少ないような場合には、単純な比例ハザードモデルによる倒産確率の推定値は過小評価されてしまう可能性があることを分析例を使って説明した。ハザードモデルの利用を有益なものとするためにはさらに何らかの工夫が必要である。最後に、考えられるいくつかの工夫をあげてみよう。

1. さまざまな種類の企業を1つのサンプルとしてまとめるのではなく、いくつかのサブグループに分類して、各グループに別々の比例ハザードモデルをあてはめる。例えば、企業の格付けをもって分類すれば、低い格付けの企業グループにおいては、倒産企業の割合がかなり高くなるであろう。ただし、こうした分析のためには、サンプル中の企業や倒産予測の対象となる企業が格付けされていなければならない。
2. 比例ハザードモデル(式1)では、指数関数の中身が各説明変数の線形関数になっているが、これを非線型のものに置き換える。例えば、ある財務指標について、安全と思われる範囲内の値については0として、その範囲から離れた場合のみ大きな正の値を与えるような変換を行ってや

表1 比例ハザードモデルの係数推定値

	r 数	exp (係数)	標準偏差	基準化係数	p 値
x1	-1.31e-02	0.987	1.12e-02	-1.1699	0.24204
x2	4.40e-03	1.004	5.23e-03	0.8421	0.39973
x3	2.21e-07	1.000	1.17e-05	0.0188	0.98500
x4	-1.84e-05	1.000	1.31e-05	-1.4119	0.15798
x5	5.28e-02	1.054	1.52e-02	3.4694	0.00052
x6	-4.74e-03	0.995	3.87e-03	-1.2227	0.22144
x7	-4.53e-02	0.956	8.35e-03	-5.4283	0.00000
x8	-5.26e-06	1.000	4.51e-04	-0.0117	0.99067
x9	-2.95e-04	1.000	5.55e-04	-0.5315	0.59507
x10	-1.88e-01	0.828	6.62e-02	-2.8431	0.00447
x11	4.43e-05	1.000	3.19e-05	1.3882	0.16508
x12	-1.19e-05	1.000	2.40e-05	-0.4940	0.62131
x13	-1.77e-03	0.998	1.62e-03	-1.0886	0.27633
x14	-1.75e-06	1.000	1.22e-05	-0.1432	0.88613
x15	-1.61e-02	0.984	7.48e-03	-2.1488	0.03165
x16	-1.94e-01	0.824	1.13e-01	-1.7208	0.08529
x18	-1.78e-01	0.837	1.16e-01	-1.5382	0.12400

Likelihood ratio test=130 on 17 df, p=0 n=10852

る。ただし、この方法を用いた有効な推定方法については研究中である。

3. 株価やリスク・プレミアムなどのデータを用いて、異なった「倒産」の定義を採用する。実際には倒産の状態にあるような企業が、現実には倒産していないというような場合が多々あれば、こうした方法も有効であると考えられる。

付録1

今 $i=1, \dots, n$ について, $\{N_i(t); t \geq 0\}$ を企業 i が時点 t までに倒産していれば1, 倒産していなければ0をとる点過程とし, そのハザード関数が

$$\lambda_i(t) = \lambda_0(t) \exp(\beta^T Z_i(t))$$

で表わされると仮定する。ここで $Z_i(t)$ は時点 t における企業 i の説明変数のベクトルである。 $Y_i(t)$ を企業 i が時点 t においてまだ生存しており, かつ観察されている場合に1を, そうでなければ0をとる関数とする。すると,

回帰係数 β の推定値は部分尤度関数

$$L(\beta) = \prod_{i=1}^n \prod_{u \geq 0} \left[\frac{Y_i(u) \exp \beta^T Z_i(u)}{\sum_{j=1}^n Y_j(u) \exp \beta^T Z_j(u)} \right]^{\Delta N_i(u)}$$

を最大とするような $\hat{\beta}$ として求めることができる。また、 Λ_0 の推定値は

$$\hat{\Lambda}_0(t) = \int_0^t \left[\sum_i Y_i(u) \exp \hat{\beta}^T Z_i(u) \right]^{-1} \left\{ \sum_{i=1}^n dN_i(u) \right\}$$

で与えられる [6]。尤度関数を $L=L(\beta, \Lambda_0)$ とすると、 $L(\beta) = \max_{\Lambda_0} L(\beta, \Lambda_0)$ であり、 $L(\hat{\beta}, \Lambda_0)$ を最大とする Λ_0 は $\hat{\Lambda}_0$ であることが証明されている [7]。

また、一般的な条件のもとで、 $n^{1/2}(\hat{\beta} - \beta)$ が正規分布に分布収束することが知られている [5] [6]。

付録 2

今回の分析で説明変数として使われた財務指標の定義は次の通りである。ただし、以下において、(*) は 12 ヶ月換算を、(**) は前年度末と当年度末の平均値 (期首期末平均値) を意味する。期首期末平均値については、可能な限り計算することとしたが、前年度末のデータが得られない場合には当年度末の値をもって代えることとした。

1. 総資本金払い後事業利益率 (%)

$$\text{利払い後事業利益}^* / \text{負債} \cdot \text{資本合計}^{**} \times 100$$

2. 企業利潤率 (%)

$$(\text{実質当期利益} + \text{法人税等充当額}^* + \text{支払利息} \cdot \text{割引料}^*) / (\text{負債} \cdot \text{資本合計}^{**} + \text{受取手形割引高}^{**} + \text{受取手形裏書譲渡高}^{**}) \times 100$$

3. 自己資本金払い後事業利益率 (%)

$$\text{利払い後事業利益}^* / \text{自己資本}^{**} \times 100$$

4. 一人当たり利払い後事業利益

$$\text{利払い後事業利益}^* / \text{従業員数}^{**}$$

5. 売上高純金利負担率 (%)

$$\text{純金利負担} / \text{売上高} \cdot \text{営業収益} \times 100$$

ただし、純金利負担：支払利息・割引料－受取利息・割引料－受取配当金

6. 借入金依存度 (%)

$(有利子負債額 - 従業員預り金) / (負債 \cdot 資本合計 + 受取手形割引高 + 受取手形裏書譲渡高) \times 100$

7. 自己資本比率 (%)

$資本合計 / 負債 \cdot 資本合計 \times 100$

8. 流動比率 (%)

$流動資産合計 / 流動負債合計 \times 100$

9. 固定長期適合率 (%)

$固定資産合計 / (固定負債合計 + 自己資本) \times 100$

10. 手元流動性比率 (倍)

$(現金 \cdot 預金 + 一時保有の有価証券) / 売上高 \cdot 営業収益^* \times 12$

11. 固定負債キャッシュフロー倍率 (倍)

$固定負債合計 / キャッシュフロー^*$

ただし、キャッシュフロー：実質当期利益＋減価償却実施額*－配当金*－役員賞与*

12. 有利子負債キャッシュフロー倍率 (倍)

$有利子負債額 / キャッシュフロー^*$

13. インタレスト・カバレッジ (倍)

$(営業利益 + 受取利息 \cdot 割引料) / 支払利息 \cdot 割引料$

14. 一人当たり付加価値額 (労働生産性)

$付加価値額^* / 従業員数^{**}$

ただし、付加価値額：利払い後事業利益＋人件費・労務費＋賃借料＋租税公課＋支払特許料＋減価償却費＋純金利負担

15. 売上高付加価値率

$付加価値額 / 売上高 \cdot 営業収益 \times 100$

個々の指標の説明については、奥野・山田 [2]、伊藤 [1] を参照のこと。

分析に使われた説明変数の標本相関行列は次の通りである。

	x1	x2	x3	x4	x5	x6	x7	x8	x9
x1	1.00000	0.58093	0.04372	0.45131	-0.19202	-0.32643	0.35007	0.04842	-0.09791
x2	0.58093	1.00000	0.00860	0.26819	-0.04775	-0.16230	0.28044	0.04760	-0.06658
x3	0.04372	0.00860	1.00000	0.15610	-0.14861	-0.02751	0.01983	0.00302	-0.00416
x4	0.45131	0.26819	0.15610	1.00000	-0.23145	-0.18214	0.18460	0.01808	-0.05056
x5	-0.19202	-0.04775	-0.14861	-0.23145	1.00000	0.49558	-0.31395	-0.05621	0.17481
x6	-0.32643	-0.16230	-0.02751	-0.18214	0.49558	1.00000	-0.53279	-0.08189	0.12107
x7	0.35007	0.28044	0.01983	0.18460	-0.31395	-0.53279	1.00000	0.26090	-0.21463
x8	0.04842	0.04759	0.00302	0.01808	-0.05621	-0.08189	0.26090	1.00000	-0.12453
x9	-0.09791	-0.06658	-0.00416	-0.05056	0.17481	0.12107	-0.21463	-0.12453	1.00000
x10	-0.00335	0.05500	0.00505	-0.00687	0.01543	0.02383	0.13166	0.20614	-0.06555
x11	-0.01225	-0.00896	0.00020	0.00282	0.05098	0.03796	-0.03322	0.01385	-0.00207
x12	-0.01190	-0.01375	0.00005	0.00079	0.02909	0.03126	-0.03223	-0.00145	0.00579
x13	0.11358	0.09144	0.00123	0.04238	-0.03696	-0.07396	0.09063	0.02768	-0.02040
x14	0.22680	0.15710	0.01698	0.37739	0.09433	0.02176	-0.00888	0.01703	-0.02797
x15	0.24144	0.15682	0.01296	0.13797	0.13523	0.02687	0.08321	0.04317	-0.01931
x16	0.12954	-0.02010	0.00292	0.12791	-0.06585	-0.00441	-0.08794	-0.10313	0.05650
x17	0.06780	-0.06516	-0.01169	0.07653	0.13969	0.12433	-0.12191	-0.06297	0.07918
x18	0.06641	-0.01232	0.00674	0.03923	0.01368	0.04364	-0.01213	-0.08419	0.06443
	x10	x11	x12	x13	x14	x15	x16	x17	x18
x1	-0.00335	-0.01225	-0.01190	0.11358	0.22680	0.24144	0.12954	0.06780	0.06641
x2	0.05500	-0.00896	-0.01374	0.09144	0.15710	0.15682	-0.02010	-0.06516	-0.01232
x3	0.00505	0.00020	0.00005	0.00123	0.01698	0.01296	0.00292	-0.01169	0.00674
x4	-0.00687	0.00282	0.00079	0.04238	0.37739	0.13797	0.12791	0.07653	0.03923
x5	0.01543	0.05098	0.02909	-0.03696	0.09433	0.13523	-0.06585	0.13969	0.01368
x6	0.02383	0.03796	0.03126	-0.07396	0.02176	0.02687	-0.00441	0.12433	0.04364
x7	0.13166	-0.03322	-0.03223	0.09063	-0.00888	0.08321	-0.08794	-0.12191	-0.01213
x8	0.20614	0.01385	-0.00145	0.02768	0.01703	0.04317	-0.10313	-0.06297	-0.08419
x9	-0.06555	-0.00207	0.00579	-0.02040	-0.02797	-0.01931	0.05650	0.07918	0.06443
x10	1.00000	-0.00418	-0.00170	0.01476	-0.01684	-0.06957	-0.14726	0.04875	-0.04596
x11	-0.00418	1.00000	0.71933	-0.00215	0.04074	0.00854	-0.01118	0.00255	-0.01876
x12	-0.00170	0.71933	1.00000	-0.00186	0.02780	0.000622	-0.00375	0.00461	-0.01336
x13	0.01476	-0.00215	-0.00186	1.00000	0.01790	0.03333	0.00942	0.00554	0.00733
x14	-0.01684	0.04074	0.02779	0.01790	1.00000	0.38415	0.08794	0.08162	-0.21030
x15	-0.06957	0.00854	0.00062	0.03333	0.38415	1.00000	-0.22327	-0.15754	-0.11090
x16	-0.14726	-0.01118	-0.00375	0.00942	0.08794	-0.22327	1.00000	0.91354	0.81094
x17	0.04875	0.00255	0.00461	0.00554	0.08162	-0.15754	0.91354	1.00000	0.82294
x18	-0.04596	-0.01876	-0.01336	0.00733	-0.21030	-0.11090	0.81094	0.82294	1.00000

参考文献

- [1] 伊藤邦雄(1998)『ゼミナール現代会計入門』第2版 日本経済新聞社
- [2] 奥野忠一・山田文道(1978)『情報化時代の経営分析』東京大学出版会
- [3] Altman, E. I. and Saunders, A. (1996) Credit Risk Measurement: Development over the Last 20 Years. *Salomon Center, W. P.*, S-96-40.
- [4] Altman, E. I., Narco, G., and Varetto, F. (1993) Corporate Distress Diagnosis: Comparisons Using Linear Discriminant Analysis and Neural Networks, *Salomon Center, W. P.*, S-93-50.
- [5] Andersen, P. K. and Gill, R. D. (1982) Cox's RRegression Model for Counting Processes: A Large Sample Study. *The Annals of Statistics*, **10**, 4, 1100-1120.
- [6] Fleming, T. R. and Harrington, D. P. (1991) *Counting Processes and Survival Analysis*, Wiley, New York.
- [7] Johansen, S. (1983) An Extension of Cox's RRegression. *International Statistical Review*, **51**, 165-174.

1) オリジナルのデータセットは120,184社の273,289期間分の財務情報が含まれており、うち3,912社が倒産企業である。今回の分析はこの中からランダムに選ばれた5,000社のデータを使っている。

(一橋大学専任講師)