

---

研究論文

---

保険契約解約のパネル計数データ・モデル

神楽岡 優昌\*

2009年6月11日投稿

2010年12月24日受理

概要

保険契約者の属性ごとにグループ分けした保険契約を対象に解約件数の計数データ・モデルを構築した。理論的整合性とパラメーターの推定可能性から、分布の再生性と overdispersion がモデル選択基準となること、そして負の2項モデルの1つである NB1 モデル（分散が期待値の線形関数）だけがその条件を満たすことを明らかにした。パネル分析により NB1 モデルのパラメーターを推定し、固定効果とランダム効果モデルの両者について比較・検討をおこなった。説明変数には、保険契約者の属性に加えて、保険契約後の経過時間・季節性・失業率などのマクロ経済変数を選んだ。実証分析の結果、契約後の経過時間や季節性、とりわけ失業率変化が解約件数に影響をあたえることを明らかにした。また解約特性が保険契約申込書のドキュメンテーション・クオリティに大きく依存して異なり、それが低くなると契約後の経過時間の説明力が優位になることが判明した。

キーワード： 保険，解約，計数データ・モデル，パネル分析

---

\* 武蔵大学 経済学部，〒176-8534 東京都練馬区豊玉上 1-26-1, E-mail: kagraoka@cc.musashi.ac.jp

## 1 はじめに

保険契約は、さまざまな特約事項はもちろんのこと、保険金請求、保険料支払いの遅延・デフォルト、保険の解約・失効、他の保険契約への転換など多岐にわたるオプションのポートフォリオからなっている。保険会社におけるリスク管理の高度化・精緻化の流れのなかで、保険解約リスクの測定と管理の重要性が認識され始めている(森平(2008))。保険が解約された場合に保険会社は保険契約者に解約返戻金などを払い戻す必要があり、解約リスクに備えて一定金額を高流動性の資産での運用や現預金での留保がなされていなければならない。保険解約リスクの測定と管理は保険会社の資産負債管理や保険商品の開発においても考慮しなければならない。また今後、保険契約の証券化ビジネスの興隆が予想されており、その証券化商品のリスク評価においても解約リスクの定量的・定性的評価が必須である(Cowley and Cummings(2005), Bütow(2001))。

保険解約に関する既存研究のほとんどは解約オプション評価に数理ファイナンスのオプション評価モデルを適用した理論的な解析にとどまっている。保険契約後に金利が上昇すると既存契約の解約が経済的に合理的となることから、金利変動のもとでの解約オプション評価をおこなっている。Albizzati and Geman(1994)は金利のHeath-Jarrow-Mortonモデルを採用して解約オプションの評価をおこなった。Grosen and Jørgensen(2000)は配当付保険に組込まれたボーナス・オプションと解約オプションの評価をおこなった。Bacinello(2003a)は養老保険について、Bacinello(2003b)は配当付生命保険についてそれぞれ解約オプションの評価をおこなった。Bacinello(2005)およびShen and Xu(2005)はユニットリンク型の生命保険について解約オプションを評価した。これまでのところ保険解約の実証分析をおこなっている研究はKim(2005)のみである。この研究は韓国の保険契約の解約率・失効率にロジット関数およびcomplementary log-log関数を適用して回帰分析をおこなった。説明変数には契約時の予定利率と現行予定利率の差に加えて、保険契約後の経過時間・失業率・経済成長率・季節性などを採用した。Kim(2005)の実証分析結果は、保険契約の解約・失効をモデル化するには、金利変動以外の説明変数も重要であることを示している。

本研究は、保険契約者の属性ごとにグループ分けして別々に構成した保険契約ポートフォリオについて、保険解約件数の計数データ・モデルの構築を目的とする。計数データ・モデルの候補として代表的なPoissonモデルと一般性をもつ負の2項モデルのファミリーを検討した。最初に保険契約の解約の振舞いをモデル化するために適切な計数データ・モデルが持つべき性質を議論した。そして理論的整合性とパラメーターの推定可能性から、分布の再生性とoverdispersionがモデル選択基準であることを明らかにした。それらの条件を満たすモデルは、負の2項モデルのファミリーの1つであるNB1モデル(分散が期待値の線形関数となる)のみであった。引き続き年金払い積立傷害保険を対象に実証分析をおこなった。NB1モデルの未定パラメーターはパネル分析により推定し、固定効果モデルとランダム効果モデルの両者について比較・検討をおこなった。NB1モデルの説明変数には保険契約者の属性に加えて観測時点の共変量(保険契約後の経過時間・季節性・マクロ経済変数など)を取り込むことができる。保険解約は保険契約者の信用力が低下したときに発生すると予想されるため、保険契約者の信用力に影響をあたえる完全失業率などのマクロ経済変数を説明変数として検討した。実証結果からNB1モデルが解約の振る舞いを適切に説明することが確認され、契約後の経過時間や季節性、とりわけ完全失業率変化が保険解約件数に影響をあたえることが明らかにされた。固定効果モデルとランダム効果モデルの推定結果に大きな違いはなく、いずれのモデルを選択すべきかについてHausmanテストからは判別できなかった。また保険契約申込書のドキュメンテーション・クオリティが保険解約件数に及ぼす影響を調べたところ、ドキュメンテーション・クオリティが低くなると保険解約の経過年数の説明力が優位になることが判明した。

本論文の構成は以下の通りである。第2章で計数データ・モデルを概説し、モデルの選択基準とモデル・パラメーターの推定方法を議論する。第3章でNB1モデルを保険解約件数に適用した実証分析結果を報告し、考察する。第4章で保険契約申込書のドキュメンテーション・クオリティが保険解約件数の振舞いにあたえる影

響について分析するとともにその実証結果について議論する。第5章で結論を述べる。

## 2 保険解約モデル

### 2.1 計数データ・モデル

保険契約者の属性別にグループ分けして別々に保険契約ポートフォリオを構成し、単位時間内のそれぞれの属性別ポートフォリオの保険解約件数のモデル化をおこなう。保険解約は非常に稀なイベントであることから、解約率ではなく、計数データ・モデルによる解約件数のモデル化をおこなう。解約モデルは計数データ・モデルの中から理論的整合性とパラメータ推定の実行可能性を基準に選択する。計数データ・モデルについての一般的な解説は Cameron and Trivedi (1998) や Winkelmann (2008) に詳しい。本小節では本研究でモデル候補とする Poisson モデルと負の2項モデル (NB1, NB2) について解説する。

計数データ・モデルのなかで Poisson モデルはもっとも単純である。このモデルでは第  $i$  番目の保険契約ポートフォリオの時刻  $t$  における保険解約の発生件数  $Y_{i,t}$  の観測値を  $y_{i,t}$  とするとき

$$\Pr[Y_{i,t} = y_{i,t}] = \frac{1}{y_{i,t}!} e^{-\gamma_{i,t}} \gamma_{i,t}^{y_{i,t}}, \quad (1)$$

の確率分布にしたがう。モデル・パラメータ  $\gamma_{i,t}$  は Poisson モデルの強度パラメータとよばれる。計数の期待値の正值性を保証するために

$$\gamma_{i,t} = \exp(X_{i,t} \cdot \beta), \quad (2)$$

とパラメータ化する。ここで  $X_{i,t}$  は第  $i$  番目の保険契約ポートフォリオの時刻  $t$  における共変量ベクトルであり、 $\beta$  はパラメータ・ベクトルである。モデルの説明変数  $X_{i,t}$  は保険契約ポートフォリオの属性だけではなく、観測時点の変数を取り込むことが可能である。例えば保険契約後の経過時間・季節性に加えて、完全失業率などのマクロ経済変数を採用できる。Poisson モデルは分散と期待値が等しいという equidispersion の性質

$$E[Y_{i,t}] = \text{Var}[Y_{i,t}] = \gamma_{i,t}, \quad (3)$$

をもっている。第3節で見ると、現実の保険解約件数の振舞いは分散が期待値よりも大きくなるという overdispersion を示しており、Poisson モデルでは実際の現象を説明できない。

同じ保険契約ポートフォリオに属する保険契約者であっても、ポートフォリオの分類に採用されなかった属性に関しては互いに異なっていると考えられる。また保険契約者間に観測できない非均質性も存在する。Poisson モデルに不均質性を取り込むことができ、負の2項モデルに拡張される。最初に不均質性を具現化するためにガンマ分布にしたがう確率変数

$$\lambda_{i,t} \sim \Gamma\left(\gamma_{i,t}, \frac{1}{\theta_{i,t}}\right), \quad (4)$$

を導入する。<sup>\*1</sup>このもとで

$$E[\lambda_{i,t}] = \theta_{i,t} \gamma_{i,t}, \quad \text{Var}[\lambda_{i,t}] = \theta_{i,t}^2 \gamma_{i,t}, \quad (5)$$

となる。 $Y_{i,t}$  の周辺分布は  $\lambda_{i,t}$  について積分すると

$$\Pr(Y_{i,t} = y_{i,t} | X_{i,t}) = \frac{\Gamma(y_{i,t} + \gamma_{i,t})}{\Gamma(\gamma_{i,t}) \Gamma(y_{i,t} + 1)} \left(\frac{1}{1 + \theta_{i,t}}\right)^{\gamma_{i,t}} \left(\frac{\theta_{i,t}}{1 + \theta_{i,t}}\right)^{y_{i,t}}, \quad (6)$$

となる。ここで強度パラメータ

$$\mu_{i,t} = \theta_{i,t} \gamma_{i,t}, \quad (7)$$

<sup>\*1</sup> 第4節のパネル分析では、 $\theta_{i,t}$  ではなく  $\theta_i$  と時間に関して定数とする。

を導入すると

$$E[Y_{i,t}] = \mu_{i,t}, \quad \text{Var}[Y_{i,t}] = \mu_{i,t} (1 + \theta_{i,t}), \quad (8)$$

となって overdispersion を表現できる.

負の 2 項モデルはパラメーター化の取り方によって異なるモデルを導くことができる. ここで

$$\theta_{i,t} = \sigma_{i,t}, \quad (9)$$

とパラメーター化すると

$$\text{Var}[Y_{i,t}] = \mu_{i,t} + \mu_{i,t} \sigma_{i,t}, \quad (10)$$

と分散が期待値の線形関数となる. Cameron and Trivedi (1998) にしたがってこのモデルを NB1 とよぶ. パラメーター化は一意ではなく,

$$\theta_{i,t} = \mu_{i,t} \sigma_{i,t}^2, \quad (11)$$

と採ると

$$\text{Var}[Y_{i,t}] = \mu_{i,t} + \mu_{i,t}^2 \sigma_{i,t}^2, \quad (12)$$

と分散が期待値の 2 次関数となる. 同様に Cameron and Trivedi (1998) にしたがってこのモデルを NB2 とよぶ. さらにパラメーター化の採り方によって, 負の 2 項モデルは分散が期待値の  $p$  次関数 ( $p = 1, 2, \dots$ ) となるバリエーションがある.

## 2.2 モデルの選択基準

一般に保険契約者の属性別に構成した保険契約ポートフォリオ間で, そのポートフォリオに含まれる保険契約数は異なっている. たとえ保険契約者の属性別に構成した保険契約ポートフォリオに含まれる保険契約数を均一に揃えても, 時間の経過とともに保険解約が発生してポートフォリオに含まれる保険契約数は単調減少し, ポートフォリオ間の保険契約数にばらつきが生じる. 保険解約件数は保険契約ポートフォリオに含まれる総保険契約数に比例すると考えられるため, 保険契約の計数モデルはポートフォリオに含まれる保険契約数を補正できなければならない. すなわち第  $i$  番目のポートフォリオの時刻  $t$  における保険解約の発生件数がポートフォリオに含まれる保険契約数  $n_{i,t}$  に比例すると仮定するのは自然なことである. そこで保険解約件数の期待値を  $\mu_{i,t} = E[Y_{i,t}]$  と表すとき

$$\mu_{i,t} = n_{i,t} \exp(X_{i,t} \cdot \beta), \quad (13)$$

とパラメーター化する.\*2これは保険契約 1 件あたりの保険解約の発生確率を  $\exp(X_{i,t} \cdot \beta)$  と解釈することを意味する.

保険契約者の属性別に構成した保険契約ポートフォリオの 1 つに注目したときに, そのポートフォリオに含まれる保険契約数は時間の経過とともに, 保険解約の発生のために単調減少する. 例えばある月末の残存保険契約数を 200 件, 翌月の解約数が 20 件, その翌月末の残存保険契約数が 180 件になったとしよう. 保険契約数が 200 件のときと保険契約数が 180 件になったときで解約のしたがう確率分布が共通でなければ, その確率分布のパラメーターを推定は困難である.\*3たとえ式 (13) のパラメトリゼーションを採用していても, 解約の確率分布を記述するボラティリティなどの他のパラメーターも共通でなければパネル分析を適用してもパラメーターの推定はできない. これは負の 2 項分布モデルを適用する場合は, 第  $i$  番目のポートフォリオのボラティリティ・パラメーターが,  $\sigma_{i,t}$  と時間と共に変化するのではなく,  $\sigma_i$  と定数であることを意味する. そして, 残

\*2 式 (13) の  $\mu$  は, 前小節における Poisson モデルの  $\gamma$  あるいは負の 2 項モデルの  $\mu$  に対応する.

\*3 これは時間方向を逆に見たほうがわかり易いかもしれない. 保険契約数が 180 件のポートフォリオに, 20 件の保険契約を加えて, 保険契約数が 200 件のポートフォリオを作成した場合を想定しよう. ポートフォリオに含まれる保険契約数の調整後に, それぞれのポートフォリオに含まれる保険契約の解約が共通の確率分布にしたがっていないならばパラメーター推定は困難である.

存保険契約数に変化しても  $\sigma_i$  が定数であることを保証する確率分布でなければならない。このことはすぐ下で議論する分布の再生性を要請することにはかならない。

これまでの議論にしたがうと、1つの保険契約ポートフォリオに注目したときに、そのポートフォリオに含まれる個々の保険契約者が同一の確率分布にしたがって保険解約をおこない、その結果としてその保険契約ポートフォリオもまたその確率分布にしたがうと要請することになる。ちなみにこの要請は式 (13) のパラメータ化と両立する。言い方を変えれば、保険契約ポートフォリオを統合したときに、統合された保険契約ポートフォリオにおける保険解約のしたがう確率分布が、結合する前の元々の保険契約ポートフォリオの保険解約のしたがう確率分布と同じであることを要請する。この条件は分布の再生性と知られている。パラメータ解釈のうえでも、式 (13) のパラメトリゼーションを保証するためにも、分布の再生性が必須条件である。

確率論で再生性とはパラメータは異なるものの同じ分布にしたがう2つの確率変数の和をとるとき、その新しい変数が元々の変数と同じ分布にしたがうことをいう。Poisson 分布では

$$Y_1 \sim \text{Poisson}(\mu_1), \quad Y_2 \sim \text{Poisson}(\mu_2), \quad (14)$$

ならば

$$Y = Y_1 + Y_2 \sim \text{Poisson}(\mu_1 + \mu_2), \quad (15)$$

となり、再生性をもっている。負の2項分布モデルについては、NB1 は

$$Y_1 \sim \text{NB1}(\mu_1, \sigma), \quad Y_2 \sim \text{NB1}(\mu_2, \sigma), \quad (16)$$

ならば

$$Y = Y_1 + Y_2 \sim \text{NB1}(\mu_1 + \mu_2, \sigma), \quad (17)$$

となって再生性を満たすが、NB2 は再生性の性質をもたない。

本研究では計数データ・モデルのなかでも Poisson モデルと NB1 と NB2 の2つの負の2項モデルを比較検討対象にする。式 (13) のパラメータ化を許す再生性をもつ確率分布は、Poisson モデルと NB1 タイプの負の2項モデルである。次節で議論するように保険解約の計数データは overdispersion を示しており、NB1 タイプの負の2項モデルが最も適切なモデルと選択される。

### 3 実証分析：保険解約

本節では NB1 モデルを実際の保険解約件数データに適用して実証分析をおこなう。分析対象の保険契約は年金払い積立傷害保険契約である。第2節の選択基準に基づき NB1 モデルを適用し、パラメータ推定をパネル分析によっておこなう。パネル分析は固定効果モデルとランダム効果モデルのそれぞれについて推定結果を報告し、議論する。

#### 3.1 データ

本研究では年金払い積立傷害保険の解約の実証分析をおこなう。最初に年金払い積立傷害保険の商品説明をおこない、引き続き分析データを解説する。

年金払い積立傷害保険は、傷害保険と年金がセットになった商品である。保険契約者は、あらかじめ決めておいた期間中、保険料を定期的に支払う。保険金受取人は保険料払込みの完了時よりも将来に設定した期日から有期年金を受取る。保険料払込み完了時から年金支払い開始時までに最大5年までの据置き期間を設けることができる。保険金受取人は保険料支払い期間・据置き期間・年金受取り期間の間にわたって傷害保険に加入している。

年金払い傷害保険のデータは日本の損害保険会社から提供された。データは

- 保険契約者と保険金受取人は同一人である。
- 保険料は月払いである。
- 払込みは継続中である。
- 契約貸付はない。
- 年金払い積立傷害保険から他の保険に転換をしていない。
- 保険契約者は保険を提供している損害保険会社と関係をもたない。

の条件を満たす契約が抽出された。典型的な保険契約者は、保険加入時の年齢に関係なく、保険料の支払い完了時を 60 歳あるいは 65 歳に設定していた。それらの各契約について毎月の状況、「保険契約の継続」あるいは「保険契約の解約」のいずれにあるかが、記録されている。データ収録期間は 1993 年 3 月から 2001 年 6 月である。

損害保険会社は個人情報保護のガイドラインのため個々の契約データにアクセスすることを許可せず、保険契約年月・契約者の性別・契約者の契約時年齢階層の属性によってグループ化されたポートフォリオについての時系列データを提供した。保険契約年月・性別・契約時の年齢階層を同一とする契約者からなるポートフォリオについて、月初の継続している保険契約数（残存契約数）と月ごとの解約件数が記録されている。各月の月初における残存契約数が 100 未満のポートフォリオは分析対象から除外した。各ポートフォリオに含まれる残存契約数は 100 から 1500 までばらついている。全体でポートフォリオの数は 183 個あり、それらのポートフォリオについて毎月の解約件数を観測することにより 12771 個のオブザベーションをえた。グループ化の属性分類基準と属性別に計数したオブザベーション数を表 1 に掲載する。観測年月別にポートフォリオの解約を計測した解約件数と解約率の基本統計量を表 2 に掲げる。解約件数の分散 3.396 は平均値の 0.98998 よりも大きく、overdispersion を示している。ポートフォリオごとに毎月の解約件数を時系列として計測し、その平均値と分散に関して作成した散布図を図 1 に掲げる。この散布図からも解約件数の分散が平均値よりもずっと大きくなるという overdispersion の傾向を読み取ることができる。

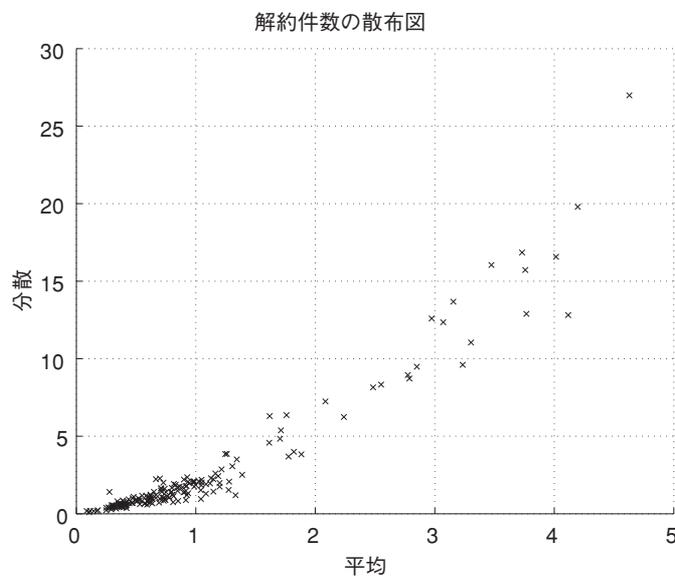


図1 解約件数の平均値と分散の散布図

保険契約者の属性と契約年月別に構成した保険契約ポートフォリオごとに毎月の保険解約件数を時系列として計測する。その平均値と分散に関して散布図を作成した。

表1 属性・観測年月ごとのオブザベーション数

属性	オブザベーション数
契約者の性別	
男性	6,626
女性	6,145
契約者の契約時の年齢	
年齢階層#2 (25歳以上, 35歳未満)	2,816
年齢階層#3 (35歳以上, 45歳未満)	4,694
年齢階層#4 (45歳以上, 55歳未満)	4,618
年齢階層#5 (55歳以上, 65歳未満)	643
契約時からの経過年数	
経過年数#1 (0年以上, 1年未満)	2,187
経過年数#2 (1年以上, 2年未満)	2,175
経過年数#3 (2年以上, 3年未満)	2,091
経過年数#4 (3年以上, 4年未満)	1,978
経過年数#5 (4年以上, 5年未満)	1,844
経過年数#6 (5年以上, 6年未満)	1,362
経過年数#7 (6年以上, 7年未満)	774
経過年数#8 (7年以上, 8年未満)	341
経過年数#9 (8年以上, 9年未満)	19
合計	12771

観測年月別にポートフォリオの解約件数を計測した総オブザベーション数の内訳を示す。ポートフォリオは保険契約者の属性ごとに構成し、契約後経過月ごとに保険解約件数を計測する。

表2 解約件数と解約率の基本統計量

	解約件数	解約率 (%)
平均値	0.9900	0.3538
標準偏差	1.8428	0.5817
分散	3.3960	0.3384
最小値	0	0.0000
最大値	17	6.7308

観測年月別にポートフォリオの解約を計測した解約件数と解約率の基本統計量を掲げる。解約率の単位は%である。総オブザベーション数は12771である。

### 3.2 説明変数

本小節ではNB1モデルの説明変数を解説する。まず、ポートフォリオのグループ分けに用いた保険契約者の性別・契約者の契約時の年齢階層を説明変数に含めた。次に、解約の契約後の経年変化を捉えるために契約後の経過年数を説明変数に加えた。さらに年度末前後に保険契約やその解約が増加することに注目して、解約の季節性を取り込むために観測月を説明変数に加えた。最後に、マクロ経済変数として、日経平均株価指数、10年スワップレート、完全失業率(季節調整値)を採用した。日経平均株価指数は景気動向の見通しを反映すると考えられる。10年スワップレートは長期金利の代表的指標であり、保険の予定利率と関連性が高いと考えられる。完全失業率は契約者の信用力の全体的な傾向を表すと考えられる。分析期間中の日経平均株価指数、

10年スワップレート，完全失業率（季節調整値）の時系列グラフをそれぞれ図2，3，4にあたえる．いずれのマクロ変数も，契約時の値を  $x_0$ ，観測時の値を  $x_t$  とするとき

$$\Delta x = x_t - x_0, \quad (18)$$

として，以下で説明する単位で経済変数の変化幅を説明変数に採択した．日経平均株価指数は，指数値を10000で除した契約時からの日経平均株価指数の変化である．10年スワップレートは，契約時からの10年スワップレートの変化であり，単位は%表記である．完全失業率は，契約時からの完全失業率（季節調整値）の変化であり，単位は%表記である．\*4マクロ経済変数の保険解約への影響は時間的なラグをとともう可能性がある．そこでマクロ経済変数については観測年月における値だけではなく，1ヶ月から6ヶ月までのラグを順にとって説明変数に入替えて採用し，最適なラグを求めた．

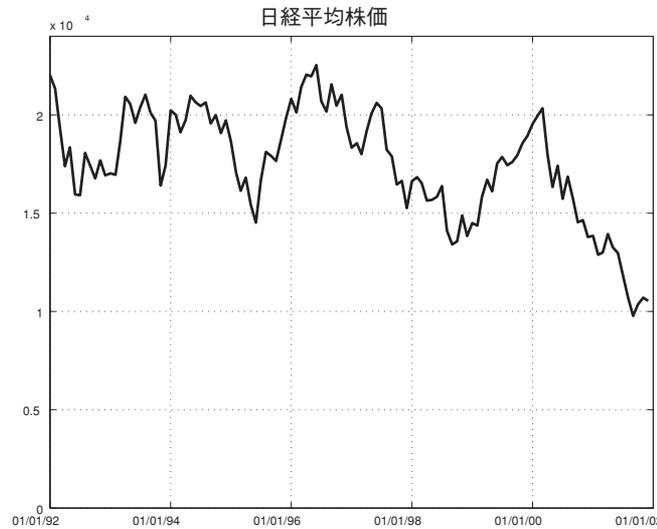


図2 日経平均株価の時系列

### 3.3 パネル分析によるパラメーター推定

第2節の選択基準に基づき NB1 モデルを適用した．保険解約件数は観測時期に大きく依存しているためパラメーター推定はパネル分析によっておこなった．計数回帰モデルのパネル分析は，Hausman, Hall and Griliches (1984) が特許件数のモデル化に適用したのが最初であり，それ以後数多くの論文が発表されている．

第  $i$  番目のポートフォリオの時刻  $t$  における解約件数を  $Y_{i,t}$  と記すとき，解約件数が  $y_{i,t}$  となる確率は

$$\Pr(Y_{i,t} = y_{i,t} | X_{i,t}) = \frac{\Gamma_{i,t}(y_{i,t} + \gamma_{i,t})}{\Gamma(\gamma_{i,t}) \Gamma(y_{i,t} + 1)} \left( \frac{1}{1 + \sigma_i} \right)^{\gamma_{i,t}} \left( \frac{\sigma_i}{1 + \sigma_i} \right)^{y_{i,t}}, \quad (19)$$

$$\mu_{i,t} = \gamma_{i,t} \sigma_i, \quad (20)$$

$$\gamma_{i,t} = \frac{n_{i,t} \exp(X_{i,t} \cdot \beta)}{\sigma_i}, \quad (21)$$

であたえられる．固定効果モデルにおいて  $\sigma_i$  は nuisance パラメーターである．ランダム効果モデルにおいて  $1/(1 + \sigma_i)$  はベータ分布

$$\frac{1}{1 + \sigma_i} \sim \text{Beta}(r, s), \quad (22)$$

\*4 本研究はマクロ経済変数の将来値の予測に基づいたリスク管理などの実務的適用を目的としており，マクロ経済変数について分散を1にするなどの標準化はおこなわなかった．

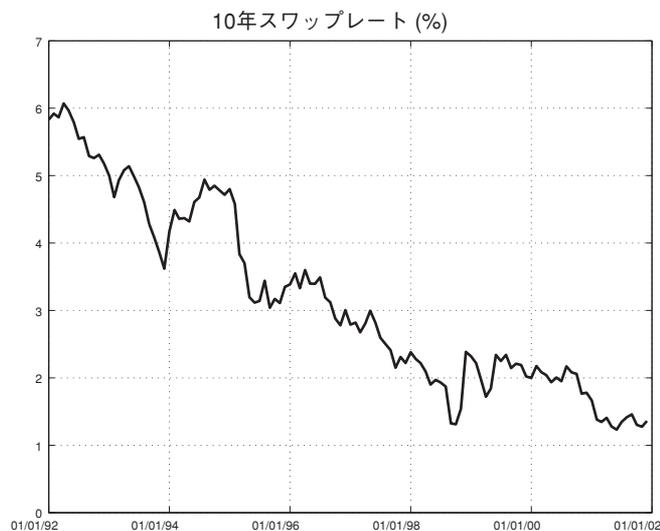


図3 10年スワップレートの時系列

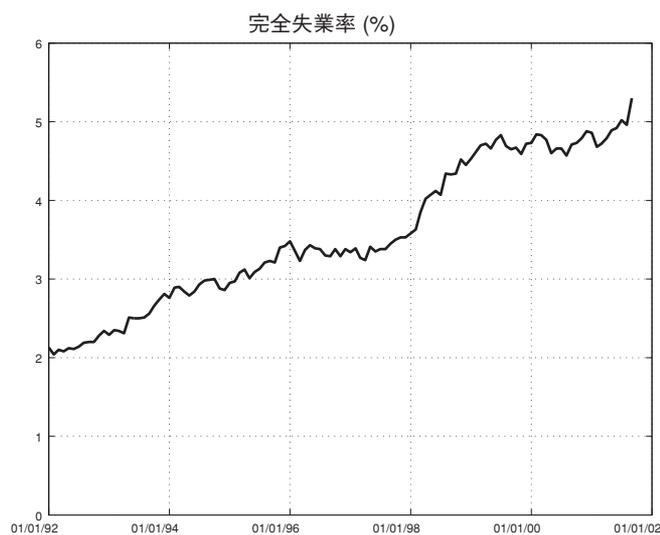


図4 完全失業率（季節調整値）の時系列

にしたがって確率分布する。この分布は区間  $(0, 1)$  上で確率密度

$$f(x; r, s) = \frac{1}{B(r, s)} x^{r-1} (1-x)^{s-1} \quad (23)$$

をもつ。ただし、 $r > 0$ 、 $s > 0$  であり、 $B(r, s)$  はベータ関数である。本分析では固定効果モデルとランダム効果モデルの2つモデルを適用し、分析結果の比較・検討をおこなった。パラメータ推定は最尤法によっておこなった。

### 3.4 パネル分析結果：固定効果モデル

固定効果モデルを採用したときのパネル分析の実証分析結果を報告する。説明変数には、ポートフォリオのグループ分けに用いた保険契約者の性別・契約時の年齢階層・契約年月に加えて、保険契約後の経過年数・観

測月・マクロ経済変数を採用した。マクロ経済変数はラグを伴って保険解約件数に影響をあたえる可能性がある。そこで個々のマクロ経済変数についてラグをとらなかった場合に加えて、1ヶ月から6ヶ月までのラグを順にとって説明変数に採用し、各々の場合の対数尤度を比較した。推定モデルの対数尤度の一覧を表3に掲げる。日経平均株価・完全失業率についてはラグをとらずに観測月の値をそのまま採用したときに対数尤度が最大になった。スワップレートは2ヶ月前の値を採用したときに対数尤度が最大になった。以下では、日経平均株価・完全失業率についてはラグをとらずに観測年月の値を説明変数に、スワップレートのみ2ヶ月前の値を説明変数に採用した場合の分析結果を考察する。

表3 固定効果モデル：マクロ経済変数のラグと対数尤度

ラグ	日経平均株価	スワップレート	失業率
0	-12003.639	-12003.639	-12003.639
1	-12012.526	-11986.353	-12165.484
2	-12028.672	-11984.279	-12289.173
3	-12031.774	-12008.367	-12378.817
4	-12034.793	-12013.247	-12436.722
5	-12038.736	-12005.852	-12551.685
6	-12035.178	-12002.495	-12581.507

保険解約件数にNB1モデルを適用し、マクロ経済変数に関してラグをとったときの対数尤度を掲げる。マクロ経済変数のラグは1ヶ月から6ヶ月まで検討した。ラグが0とは観測時点の経済変数を採用して、ラグをとらなかったことを意味する。

固定効果モデルの推定結果を表4に掲載する。固定効果モデルにおいて $\sigma_i$ はnuisanceパラメーターである。経過年数および季節性に関するパラメーターについてはそれぞれ図5、図6にグラフ化する。推定したモデルのベースラインは、

- 性別：男性
- 年齢階層：#2 (25歳以上, 35歳未満)
- 経過年数：#1 (1年未満)
- 観測年月：1月

に対応する。係数はこれらの属性をもつポートフォリオを0とするベースラインに対する相対的なものであり、負値はこのベースラインとしたポートフォリオの場合よりも解約件数が低いことを意味する。ベースラインとしたポートフォリオについて、式(13)の $X_{i,t} \cdot \beta$ は定数項に対応する。対数尤度は-11984.279となった。すべてのパラメーターが共にゼロという仮説を検定するWald検定統計量(自由度26の $\chi^2$ 分布にしたがう)は4705.12となり、帰無仮説は有意に棄却された。この統計量から解約モデルの説明変数が有効と判断できる。説明変数間に多重共線性は認められなかった。

引き続き各説明変数について考察をおこなう。性別や年齢階層の説明変数は統計的に有意ではない。経過年数とともに解約件数は上昇し、経過年数#3(2年以上, 3年未満)の1.360810をピークに減少する。経過年数が#8(7年以上, 8年未満)、#9(8年以上, 9年未満)では負値をとり、基準となる経過年数#1(1年未満)よりも低い値になる。これは「信用力の高い保険契約者」と「信用力の低い保険契約者」の2つのグループが、各ポートフォリオを構成していると仮定すると説明できる。「信用力の低い保険契約者」は契約後の保険契約者の信用力の変動とともに保険契約の解約をおこす。一方、「信用力の高い保険契約者」は契約後に信用力の変動があっても保険契約の解約までは至らない。その結果、短期的には契約後の経過時間とともに各ポートフォリオに含まれる「信用力の低い保険契約者」が解約をするが、やがては「信用力の低い保険契約者」の数が減少し、中長期的には解約件数の低下となって現れる。季節性では3月と10月にピークがあり、6

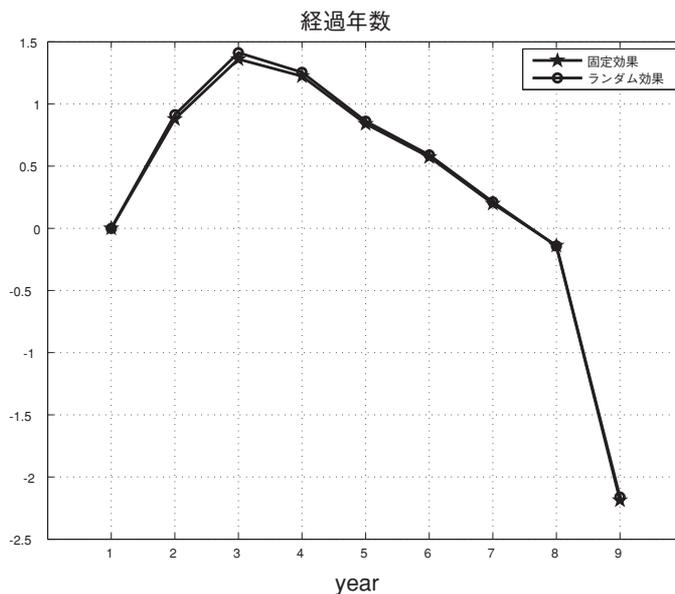


図5 保険解約件数と経過年数の関係

固定効果モデル (fixed-effects) とランダム効果モデル (random effects) の結果を比較のため重ねて表示した。

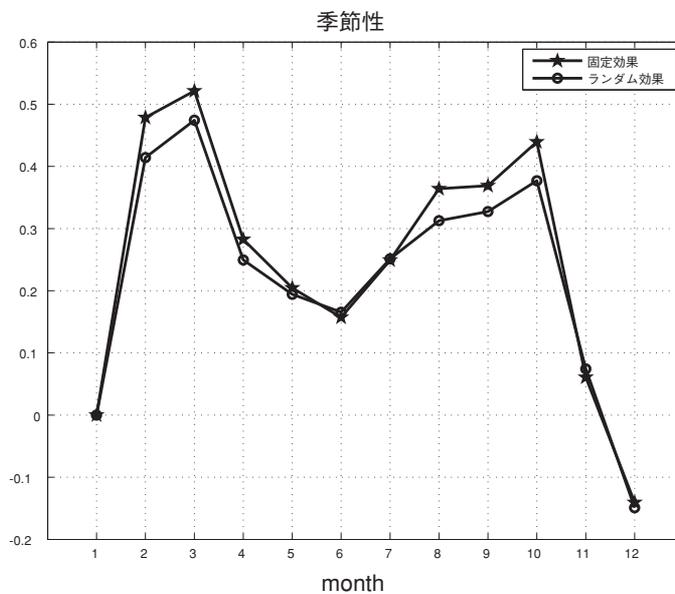


図6 保険解約件数の季節性

固定効果モデル (fixed-effects) とランダム効果モデル (random effects) の結果を比較のため重ねて表示した。

表4 固定効果モデルによるパラメータの推定結果

説明変数	係数	標準偏差	t 値	P 値
性別 (女性)	0.238251	0.117684	2.02	0.0430
年齢階層#3	0.067435	0.287209	0.23	0.8140
年齢階層#4	0.015935	0.279365	0.06	0.9550
年齢階層#5	0.293308	0.286945	1.02	0.3070
経過年数#2	0.878673	0.128551	6.84	0.0000
経過年数#3	1.360810	0.127016	10.71	0.0000
経過年数#4	1.223386	0.132431	9.24	0.0000
経過年数#5	0.839982	0.138078	6.08	0.0000
経過年数#6	0.573524	0.145416	3.94	0.0000
経過年数#7	0.198646	0.154473	1.29	0.1980
経過年数#8	-0.139503	0.168507	-0.83	0.4080
経過年数#9	-2.186992	0.528639	-4.14	0.0000
2 月	0.478672	0.053881	8.88	0.0000
3 月	0.521371	0.052581	9.92	0.0000
4 月	0.282577	0.053673	5.26	0.0000
5 月	0.204543	0.053663	3.81	0.0000
6 月	0.157189	0.053913	2.92	0.0040
7 月	0.249128	0.056581	4.40	0.0000
8 月	0.364039	0.056674	6.42	0.0000
9 月	0.368946	0.055495	6.65	0.0000
10 月	0.439359	0.055107	7.97	0.0000
11 月	0.060653	0.056863	1.07	0.2860
12 月	-0.140594	0.058662	-2.40	0.0170
△ (日経平均株価指数)	-0.583013	0.062476	-9.33	0.0000
△ (10年スワップレート)	-0.355608	0.042492	-8.37	0.0000
△ (完全失業率)	1.642591	0.052019	31.58	0.0000
定数	-8.922239	0.294225	-30.32	0.0000
対数尤度	-11984.279			
Wald $\chi^2(26)$	4705.12			
Prob > $\chi^2$	0.0000			

保険契約者の性別・契約時の年齢階層・契約年月によってグループ分けしたポートフォリオの保険解約件数に NB1 モデルを適用した固定効果モデルの推定結果を掲げる。説明変数にはポートフォリオの分類に用いた保険契約者の属性に加えて、保険契約後の経過年数・観測月・マクロ経済変数を採用した。

月と12月にボトムになる。特に3月に解約が多くなるのは、年度末に就職・転職や子供の入学・進学などの出来事が多くあるためと思われる。10月に解約が多い理由は不明である。

マクロ経済変数はいずれも統計的に有意であるが、完全失業率と日経平均株価が特に有意である。パラメータの大きさから判断すると、完全失業率の影響が大きく、完全失業率が悪化すると解約件数が上昇することがわかる。これは完全失業率が上昇するときは家計が苦しくなり、保険契約の解約につながると思われる。日経平均株価はパラメータの符号が負であり、日経平均が上昇すると心理的に経済的見通しが楽観的になり、保険契約の解約の低下になると解釈される。10年スワップレートに対する解釈は難しい。長期金利の上昇は契約している年金払い傷害保険の現在価値の低下につながり、解約件数の上昇をもたらす。一方、長期金利の上昇は景気の減速を意図しておこなうため、経済は好景気にあることが多く、契約件数の下降につながるであろう。この2つの効果は互いに逆方向にはたらき、どちらが優位になるか理論的な示唆はない。しかし、分析結

果はパラメーターの符号が負であることは、後者の影響が大きかったことを示している。

### 3.5 パネル分析結果：ランダム効果モデル

ランダム効果モデルを採用したときのパネル分析の実証分析結果を報告する。説明変数には、ポートフォリオのグループ分けに用いた保険契約者の性別・契約時の年齢階層・契約年月に加えて、保険契約後の経過年数・観測月・マクロ経済変数を採用した。固定効果モデルの場合と同様に、個々のマクロ経済変数についてラグをとらなかった場合に加えて、1ヶ月から6ヶ月までのラグを順にとって説明変数に採用し、各々の場合の対数尤度を比較した。推定モデルの対数尤度の一覧を表5に掲げる。

表5 ランダム効果モデル：マクロ経済変数のラグと対数尤度

ラグ	日経平均株価	スワップレート	失業率
0	-12937.498	-12937.498	-12937.498
1	-12952.742	-12932.246	-13097.224
2	-12980.961	-12931.151	-13225.920
3	-12989.558	-12937.785	-13326.682
4	-12996.910	-12937.277	-13385.975
5	-13006.950	-12937.568	-13496.161
6	-13006.346	-12936.996	-13534.504

保険解約件数にNB1モデルを適用し、マクロ経済変数に関してラグをとったときの対数尤度を掲げる。マクロ経済変数のラグは1ヶ月から6ヶ月まで検討した。ラグが0とは観測時点の経済変数を採用して、ラグをとらなかったことを意味する。

マクロ経済変数の最適なラグのとり方は固定効果モデルの場合と同様であった。日経平均株価・完全失業率については、ラグをとらずに観測月の値をそのまま採用したときに対数尤度が最大になった。スワップレートは2ヶ月前の値を採用したときに対数尤度が最大になった。そこで以下では、日経平均株価・完全失業率については観測年月の値を説明変数に、スワップレートのみラグをとって2ヶ月前の値を説明変数に採用した場合の分析結果を考察する。

ランダム効果モデルの推定結果を表6に掲載する。経過年数および季節性に関するパラメーターについてはそれぞれ図5、図6にグラフ化する。推定したモデルのベースラインは、固定効果モデルの場合と共通である。対数尤度は-12931.151となった。すべてのパラメーターが共にゼロという仮説を検定するWald検定統計量（自由度26の $\chi^2$ 分布にしたがう）は4577.58となり、帰無仮説は有意に棄却された。この統計量から解約モデルの説明変数が有効と判断できる。pooled推定量（分散パラメーター $\sigma_i$ をすべてのポートフォリオにわたって共通の定数としたときの推定パラメーター）と比較する対数尤度比テストは580.96となり、パネル推定量がより望ましいことがわかる。説明変数間に多重共線性は認められなかった。

引き続き各説明変数について考察をおこなう。各説明変数について固定効果モデルとランダム効果モデルの推定結果に大きな相違点はなかった。性別や年齢階層の説明変数は統計的に有意ではない。経過年数とともに解約件数は上昇し、経過年数#3（2年以上、3年未満）の1.412624をピークに減少する。季節性では3月と10月にピークがあり、6月と12月にボトムになる。マクロ経済変数はいずれも統計的に有意であるが、完全失業率と日経平均株価が特に有意である。いずれのパラメーターも固定効果モデルの推定結果と類似しており、同様の経済的解釈が可能である。

ランダム効果モデルでは式(22)にあるように $1/(1+\sigma_i)$ がベータ分布にしたがう。ベータ分布 $Beta(r, s)$ のパラメーターは $r=6.505514$ 、 $s=3.447007$ となった。推定パラメーターに基づいた確率密度関数を図7に示す。この図によると $1/(1+\sigma_i)$ が0.692付近を中心に分布が広がっている。この中心は $\sigma_i=0.4451$ に対応

表 6 ランダム効果モデルによるパラメーターの推定結果

説明変数	係数	標準偏差	t 値	P 値
性別 (女性)	0.031315	0.078136	0.40	0.6890
年齢階層#3	0.166476	0.185897	0.90	0.3710
年齢階層#4	0.125864	0.179109	0.70	0.4820
年齢階層#5	0.198928	0.180339	1.10	0.2700
経過年数#2	0.914010	0.127982	7.14	0.0000
経過年数#3	1.412624	0.126252	11.19	0.0000
経過年数#4	1.254481	0.131522	9.54	0.0000
経過年数#5	0.861885	0.137271	6.28	0.0000
経過年数#6	0.589753	0.144765	4.07	0.0000
経過年数#7	0.212505	0.153961	1.38	0.1680
経過年数#8	-0.142514	0.168103	-0.85	0.3970
経過年数#9	-2.161377	0.528524	-4.09	0.0000
2 月	0.414329	0.054875	7.55	0.0000
3 月	0.474275	0.053432	8.88	0.0000
4 月	0.249525	0.054502	4.58	0.0000
5 月	0.194284	0.054373	3.57	0.0000
6 月	0.165636	0.054708	3.03	0.0020
7 月	0.251254	0.057345	4.38	0.0000
8 月	0.312611	0.057480	5.44	0.0000
9 月	0.327401	0.056216	5.82	0.0000
10 月	0.377041	0.055919	6.74	0.0000
11 月	0.074329	0.057497	1.29	0.1960
12 月	-0.149076	0.059531	-2.50	0.0120
△ (日経平均株価指数)	-0.738158	0.061679	-11.97	0.0000
△ (10 年スワップレート)	-0.153184	0.042513	-3.60	0.0000
△ (完全失業率)	1.730233	0.051712	33.46	0.0000
定数	-8.825960	0.210780	-41.87	0.0000
ln(r)	1.872650	0.138068		
ln(s)	1.237506	0.137830		
r	6.505514	0.898201		
s	3.447007	0.475101		
対数尤度	-12931.151			
Wald $\chi^2(26)$	4577.58			
Prob > $\chi^2$	0.0000			
対数尤度比テスト	580.96			

保険契約者の性別・契約時の年齢階層・契約年月によってグループ分けしたポートフォリオの保険解約件数に NBI モデルを適用したランダム効果モデルの推定結果を掲げる。説明変数にはポートフォリオの分類に用いた保険契約者の属性に加えて、保険契約後の経過年数・観測月・マクロ経済変数を採用した。

する。

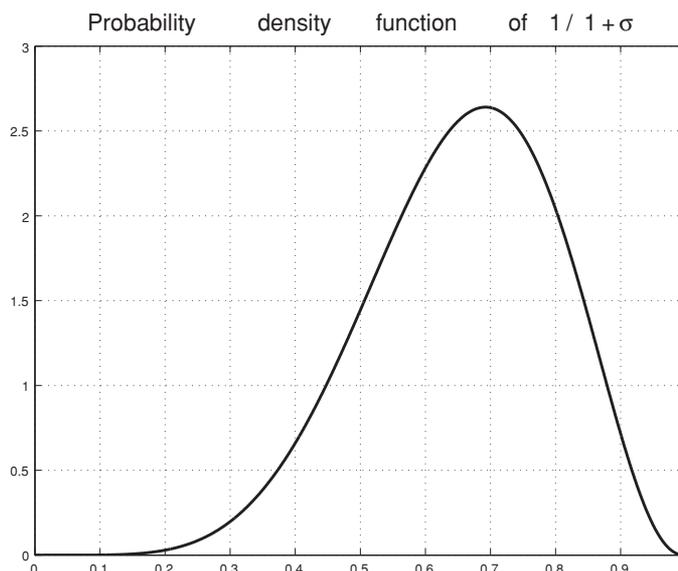


図7 ランダム効果モデルのパラメーター分布

式 (22) のパラメーター分布  $\frac{1}{1+\sigma} \sim \text{Beta}(r, s)$ .

Hausman テストにより固定効果モデルの一致推定量とランダム効果モデルの有効推定量が同一であるかについて検定を試みた。しかしながら、検定統計量を計算するための共分散行列の差が正定値行列とはならないために Hausman テストを適用できなかった。したがって固定効果モデルとランダム効果モデルのいずれを採用すべきかの結論に至らなかった。

## 4 ドキュメンテーション・クオリティが保険解約件数にあたる影響

### 4.1 データと予備分析結果

保険契約申込書のドキュメンテーション・クオリティが保険契約者の振舞いにどのような影響をあたえるかは非常に興味深い問題である。<sup>\*5</sup>本研究では郵便番号が未記入の保険契約申込書だけを抽出したデータセットを新たに保険会社に提供を依頼した。郵便番号が未記入の保険契約申込書でも現住所など他の記入欄は完全に記入されているため、郵便番号のデータは冗長である。保険契約申込書の郵便番号が未記入の保険契約数は非常に少なかった。そのため前節と同一の分類基準でポートフォリオを構築すると、個々のポートフォリオに含まれる保険契約が数十にとどまり、計数データ・モデルを適用するのは不適當となった。そこでドキュメンテーション・クオリティによる保険解約件数の分析にあたって、保険契約者の契約時の年齢ではグループ分けせずに、保険契約年月と性別のみを分類項目とした。比較のために前節で用いた保険契約申込書が完全な契約についても、保険契約年月と性別のみで分類したポートフォリオを新たに構成した。毎月の月初における残存契約数が 100 未満のポートフォリオは分析対象から除外した。それらのポートフォリオについて、毎月の解約件数を計数した。以後、保険契約申込書のドキュメンテーションが不完全なデータセットもしくはそれから構築したポートフォリオの全体を LQ、それが完全なデータセットもしくはそれから構築したポートフォリオの

<sup>\*5</sup> 欧米の住宅ローンの証券化においては、住宅ローンの申込書類のドキュメンテーション・クオリティが与信後の住宅ローン債務者の借換え・デフォルトなどのイベントの発生確率と密接に関連することが広く認識されている。

全体を HQ とよぶ。

データセット LQ について、ポートフォリオは全部で 59 あり、毎月の解約件数を観測した結果 1361 のオブザベーションをえた。データセット HQ については、ポートフォリオは全部で 129 あり、毎月の解約件数を観測した結果 8058 のオブザベーションをえた。グループ化の属性分類と属性別に計数したオブザベーション数を表 7 に掲載する。データセット LQ については、保険契約後 5 年以上経過した保険契約ポートフォリオは存在していない。観測年月別にポートフォリオの解約を計測した解約件数の基本統計量を表 8 に掲げる。データセット LQ では解約件数の平均値は 8.4276、分散は 54.6023、データセット HQ では解約件数の平均値は 2.0377、分散は 15.1804 といずれも overdispersion を示している。ポートフォリオごとに毎月の解約件数を計測し、その平均値と分散に関して作成した散布図を図 8 に掲げる。この散布図からも解約件数の分散が平均値よりもずっと大きくなるという overdispersion の傾向を読み取ることができる。

表 7 属性・観測年月ごとのオブザベーション数

属性	オブザベーション数	
	LQ	HQ
契約者の性別		
男性	757	4,009
女性	604	4,049
契約時からの経過年数		
経過年数#1 (0 年以上, 1 年未満)	623	1,548
経過年数#2 (1 年以上, 2 年未満)	444	1,486
経過年数#3 (2 年以上, 3 年未満)	225	1,313
経過年数#4 (3 年以上, 4 年未満)	63	1,181
経過年数#5 (4 年以上, 5 年未満)	6	1,023
経過年数#6 (5 年以上, 6 年未満)	0	780
経過年数#7 (6 年以上, 7 年未満)	0	492
経過年数#8 (7 年以上, 8 年未満)	0	223
経過年数#9 (8 年以上, 9 年未満)	0	12
合計	1,361	8,058

ドキュメンテーション・クオリティに基づいて別々に構成したポートフォリオのオブザベーション数を掲げる。LQ はドキュメンテーションが不完全な保険契約ポートフォリオの集合である。HQ はドキュメンテーションが完全な保険契約ポートフォリオの集合である。観測年月別にポートフォリオの解約件数を計測した総オブザベーション数の内訳を示す。ポートフォリオは保険契約者の属性ごとに構成し、契約後経過月ごとに保険解約件数を計測する。

## 4.2 パネル分析結果

前節の分析と同様にマクロ経済変数についてラグを変化させて対数尤度の検討をおこなった。LQ については、マクロ経済変数のラグをとっても対数尤度の変化に全体的な傾向を見いだせられなかった。HQ はスワップレートのみ 2 ヶ月のラグをとったときに対数尤度が最大になった。本分析は HQ と LQ 間のパラメータの比較を目的としており、LQ と HQ に同一のモデルを適用することとする。そこでマクロ経済変数はラグをとらないモデルを共通に適用した。

最初に固定効果モデルの分析結果を考察する。推定パラメータを表 9 に掲げる。固定効果モデルにおいて  $\sigma_i$  は nuisance パラメータである。推定したモデルのベースラインは、

- 性別：男性

表8 ドキュメンテーション・クオリティ別の解約件数の基本統計量

	LQ	HQ
総オブザベーション数	1361	8058
平均値	8.4276	2.0377
標準偏差	7.3893	3.8962
分散	54.6023	15.1804
歪度	1.6910	3.6235
尖度	6.4309	20.4665
最小値	0	0
最大値	48	42

観測年月別にポートフォリオの解約を計測した解約件数の基本統計量を掲げる。

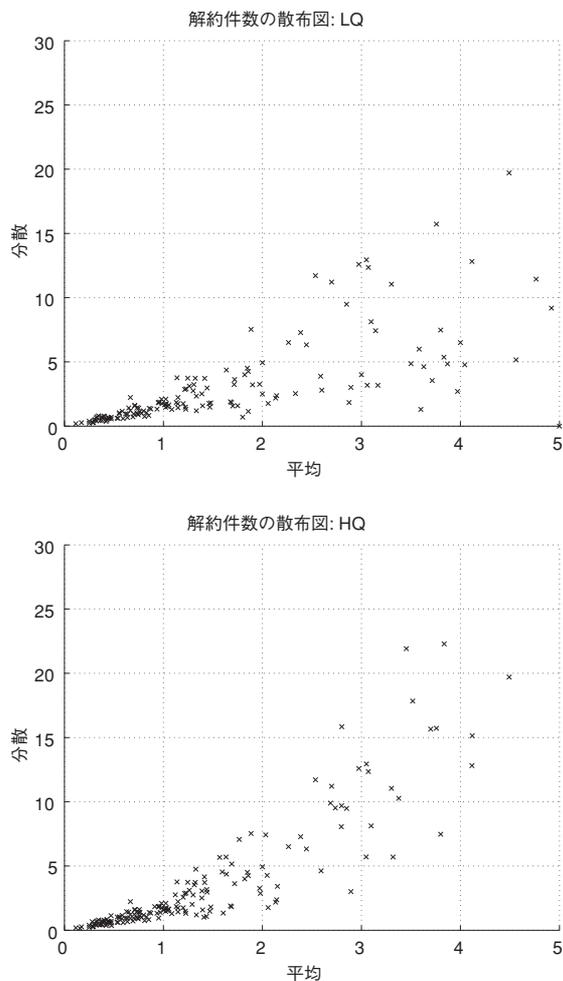


図8 ドキュメンテーション・クオリティによる保険契約ポートフォリオの解約件数の平均値と分散の散布図の比較

上図 (LQ) はドキュメンテーションが不完全な保険契約ポートフォリオの解約件数の平均値と分散の散布図である。下図 (HQ) はドキュメンテーションが完全な保険契約ポートフォリオの解約件数の平均値と分散の散布図である。

- 経過年数：#1（1年未満）
- 観測年月：1月

に対応する。LQ および HQ の対数尤度はそれぞれ  $-3321.166$  と  $-10718.960$  になった。すべてのパラメーターが共にゼロという仮説を検定する Wald 検定統計量（LQ の場合は自由度 19 の  $\chi^2$  分布に、HQ の場合は自由度 23 の  $\chi^2$  分布にしたがう）はそれぞれ 1857.17 と 5611.58 となり、いずれも帰無仮説は有意に棄却された。これらの統計量から解約モデルの説明変数が有効と判断できる。いずれも説明変数間に多重共線性は認められなかった。以下では各説明変数について考察をおこなう。LQ は経過年数が統計的に最も有意であり、保険契約後時間の経過とともに解約件数は単調増加する。HQ では完全失業率の係数が正で統計的に非常に有意であるのとは対照的に、LQ では完全失業率の係数は負となった。LQ では完全失業率が上昇すると解約件数が下落するという経済的に不可解な事実を意味する。しかし、パラメーターの大きさは HQ の 2.03 に対して LQ では  $-0.43$  で、月毎の完全失業率の変化は大きくないため、完全失業率上昇による保険解約件数の減少の影響は小さい。また、HQ では顕著な季節性が認められたが、LQ ではそのような傾向が存在しなかった。以上のことから LQ の解約動向はもっぱら保険契約後の経過時間だけに依存して、単調に増加することが判明した。

次にランダム効果モデルの分析結果を考察する。ランダム効果モデルの推定結果を表 10 に掲載する。推定したモデルのベースラインは、固定効果モデルの場合と共通である。LQ および HQ の対数尤度はそれぞれ  $-3610.959$  と  $-11464.602$  になった。すべてのパラメーターが共にゼロという仮説を検定する Wald 検定統計量（LQ の場合は自由度 19 の  $\chi^2$  分布に、HQ の場合は自由度 23 の  $\chi^2$  分布にしたがう）はそれぞれ 1829.24 と 5469.36 となり、いずれも帰無仮説は有意に棄却された。これらの統計量から解約モデルの説明変数が有効と判断できる。pooled 推定量（分散パラメーター  $\sigma_i$  をすべてのポートフォリオにわたって共通の定数としたときの推定パラメーター）と比較する対数尤度比テストは、LQ では 158.99、HQ では 870.63 となって、いずれもパネル推定量がより望ましいことがわかる。これらの統計量から解約のモデル化が有効と判断できる。いずれも説明変数間に多重共線性は認められなかった。各説明変数の推定数値は固定効果の場合と大きな相違が認められなかった。保険契約者の属性別に構築したポートフォリオは固有の  $\sigma_i$  をもっている。ランダム効果モデルでは式 (22) にあるように  $1/(1 + \sigma_i)$  がベータ分布にしたがう。ベータ分布  $Beta(r, s)$  のパラメーターは、LQ では  $r = 26.970330$ 、 $s = 23.340650$ 、HQ では  $r = 2.703208$ 、 $s = 3.067405$  となり、大きな相違がある。パラメーター分布をわかりやすく比較するためにベータ分布の確率密度関数を図 9 に描画した。図が顕著に示すように LQ では  $1/(1 + \sigma_i)$  が 0.538 付近を中心に分布が集中しているが、HQ では  $1/(1 + \sigma_i)$  が 0.451 を中心に分布がゆるやかに広がっている。属性別に構築したポートフォリオについて、解約の説明変数にそれらの属性を採用し、さらに  $\sigma_i$  を導入してポートフォリオの固有要因を考慮した。それにも関わらず保険契約者の属性が異なるポートフォリオ間で  $\sigma_i$  のばらつきが LQ の方が HQ よりも小さいことに注目されたい。LQ については経過年数以外のポートフォリオの属性が説明力をもたなかったこと、そして LQ の方が HQ よりも  $\sigma_i$  の散らばりが小さいことから、LQ は保険契約者の属性に依らない非経済的な特殊な理由で解約行動をおこなっていると推察される。言い換えれば、LQ の解約を決定づける非経済的な属性があり、その属性は LQ に共通と考えられる。このようにドキュメンテーション・クオリティが解約行動に大きく影響し、さらにドキュメンテーションが不完全な保険契約者が保険解約の意思決定において特殊な振舞いをすることは本分析で初めて明らかにされた事実である。

データセット LQ および データセット HQ それぞれについて、Hausman テストにより固定効果モデルの一致推定量とランダム効果モデルの有効推定量が同一であるかについて検定を試みた。しかしながら、検定統計量を計算するための共分散行列の差が正定値行列とはならないために Hausman テストを適用できなかった。したがって固定効果モデルとランダム効果モデルのいずれを採用すべきかの結論に至らなかった。

表9 ドキュメンテーション・クオリティの影響：固定効果モデルによるパラメーターの推定結果

説明変数	LQ				HQ			
	係数	標準偏差	t 値	P 値	係数	標準偏差	t 値	P 値
性別（女性）	0.417866	0.185911	2.25	0.0250	0.206878	0.096079	2.15	0.0310
経過年数#2	0.776610	0.039264	19.78	0.0000	0.345641	0.079005	4.37	0.0000
経過年数#3	1.238026	0.062212	19.90	0.0000	0.405301	0.085077	4.76	0.0000
経過年数#4	1.524034	0.092408	16.49	0.0000	0.186523	0.092761	2.01	0.0440
経過年数#5	1.846738	0.186538	9.90	0.0000	-0.192383	0.101367	-1.90	0.0580
経過年数#6					-0.504091	0.111279	-4.53	0.0000
経過年数#7					-0.862121	0.122531	-7.04	0.0000
経過年数#8					-1.247152	0.138776	-8.99	0.0000
経過年数#9					-3.053488	0.478351	-6.38	0.0000
2 月	-0.059507	0.059901	-0.99	0.3210	0.321087	0.051065	6.29	0.0000
3 月	-0.038860	0.060932	-0.64	0.5240	0.392927	0.051144	7.68	0.0000
4 月	0.016260	0.060599	0.27	0.7880	0.192034	0.053339	3.60	0.0000
5 月	0.054686	0.060566	0.90	0.3670	0.142012	0.053489	2.65	0.0080
6 月	-0.213397	0.064227	-3.32	0.0010	0.060110	0.053666	1.12	0.2630
7 月	-0.188926	0.062233	-3.04	0.0020	0.205988	0.055688	3.70	0.0000
8 月	-0.027183	0.058673	-0.46	0.6430	0.312684	0.054107	5.78	0.0000
9 月	-0.024406	0.060043	-0.41	0.6840	0.281737	0.054667	5.15	0.0000
10 月	-0.018733	0.060728	-0.31	0.7580	0.363661	0.052862	6.88	0.0000
11 月	-0.151795	0.063495	-2.39	0.0170	0.107960	0.056114	1.92	0.0540
12 月	-0.124616	0.062742	-1.99	0.0470	-0.180860	0.057158	-3.16	0.0020
△（日経平均株価指数）	-0.075886	0.079857	-0.95	0.3420	-0.806069	0.067557	-11.93	0.0000
△（10年スワップレート）	-0.277290	0.042793	-6.48	0.0000	-0.014515	0.042589	-0.34	0.7330
△（完全失業率）	-0.429179	0.110720	-3.88	0.0000	2.029457	0.050187	40.44	0.0000
定数	-4.032797	0.118052	-34.16	0.0000	-8.335414	0.093585	-89.07	0.0000
対数尤度	-3321.166				-10718.960			
Wald $\chi^2$	1857.17				5611.58			
Prob > $\chi^2$	0.0000				0.0000			

ドキュメンテーション・クオリティに基づいて別々にモデル・パラメーターを推定した。LQ はドキュメンテーションが不完全な保険契約ポートフォリオの集合である。HQ はドキュメンテーションが完全な保険契約ポートフォリオの集合である。保険契約者の性別・契約年月によってグループ分けしたポートフォリオの保険解約件数にNB1モデルを適用した固定効果モデルの推定結果を掲げる。説明変数にはポートフォリオの分類に用いた保険契約者の属性に加えて、保険契約後の経過年数・観測月・マクロ経済変数を採用した。

## 5 おわりに

分析期間中の長期金利は一貫して低下傾向にあり、保険解約は経済的に合理的でない。それにもかかわらず保険契約者が保険解約をおこなうのは、保険契約者の家計状況・信用力の悪化により、保険料支払いが経済的負担になるからと考えられる。保険解約のモデル化において、現時点の完全失業率が有意であったことが証左となった。しかしながら熟考すると、家計が悪化したときに最初に保険契約の解約がおこるのであり、失業してから保険解約に至るのではない。すなわち、家計が悪化した場合に、経済的節約をおこなうのは自然であり、その1つの方策として保険料支払いの負担感から保険解約に至ると考えられる。あるいは保険の解約返戻金

表 10 ドキュメンテーション・クオリティの影響：ランダム効果モデルによるパラメータの推定結果

説明変数	LQ				HQ			
	係数	標準偏差	t 値	P 値	係数	標準偏差	t 値	P 値
性別 (女性)	0.064092	0.079661	0.80	0.4210	0.095065	0.080718	1.18	0.2390
経過年数#2	0.810404	0.039149	20.70	0.0000	0.349265	0.078867	4.43	0.0000
経過年数#3	1.306506	0.061704	21.17	0.0000	0.398731	0.085055	4.69	0.0000
経過年数#4	1.614413	0.091841	17.58	0.0000	0.171327	0.092777	1.85	0.0650
経過年数#5	1.963227	0.184130	10.66	0.0000	-0.201827	0.101590	-1.99	0.0470
経過年数#6					-0.511659	0.111815	-4.58	0.0000
経過年数#7					-0.862571	0.123341	-6.99	0.0000
経過年数#8					-1.250874	0.139848	-8.94	0.0000
経過年数#9					-3.034812	0.483013	-6.28	0.0000
2 月	-0.062989	0.060121	-1.05	0.2950	0.318720	0.051858	6.15	0.0000
3 月	-0.046983	0.060898	-0.77	0.4400	0.412663	0.051924	7.95	0.0000
4 月	0.015249	0.060564	0.25	0.8010	0.223461	0.054052	4.13	0.0000
5 月	0.049993	0.060681	0.82	0.4100	0.171482	0.054247	3.16	0.0020
6 月	-0.220890	0.064333	-3.43	0.0010	0.070954	0.054534	1.30	0.1930
7 月	-0.196853	0.062419	-3.15	0.0020	0.209152	0.056593	3.70	0.0000
8 月	-0.033425	0.058772	-0.57	0.5700	0.312501	0.054902	5.69	0.0000
9 月	-0.030761	0.060028	-0.51	0.6080	0.304291	0.055342	5.50	0.0000
10 月	-0.009200	0.060732	-0.15	0.8800	0.381681	0.053501	7.13	0.0000
11 月	-0.146734	0.063573	-2.31	0.0210	0.136430	0.056821	2.40	0.0160
12 月	-0.122955	0.063106	-1.95	0.0510	-0.186125	0.057946	-3.21	0.0010
△ (日経平均株価指数)	-0.014060	0.077559	-0.18	0.8560	-0.904489	0.066254	-13.65	0.0000
△ (10 年スワップレート)	-0.262702	0.039195	-6.70	0.0000	0.087671	0.041360	2.12	0.0340
△ (完全失業率)	-0.630077	0.109782	-5.74	0.0000	2.062002	0.049991	41.25	0.0000
定数	-3.858966	0.102584	-37.62	0.0000	-8.277918	0.090451	-91.52	0.0000
ln(r)	3.294737	0.255350			0.994439	0.145016		
ln(s)	3.150197	0.257482			1.120832	0.147402		
r	26.970330	6.886862			2.703208	0.392009		
s	23.340650	6.009800			3.067405	0.452142		
対数尤度	-3610.959				-11464.602			
Wald $\chi^2$	1829.24				5469.36			
Prob > $\chi^2$					0.0000			

ドキュメンテーション・クオリティで分けてモデル・パラメータを推定した。LQ はドキュメンテーションが不完全な保険契約ポートフォリオの集合である。HQ はドキュメンテーションが完全な保険契約ポートフォリオの集合である。保険契約者の性別・契約年月によってグループ分けしたポートフォリオの保険解約件数に NB1 モデルを適用したランダム効果モデルの推定結果を掲げる。説明変数にはポートフォリオの分類に用いた保険契約者の属性に加えて、保険契約後の経過年数・観測月・マクロ経済変数を採用した。

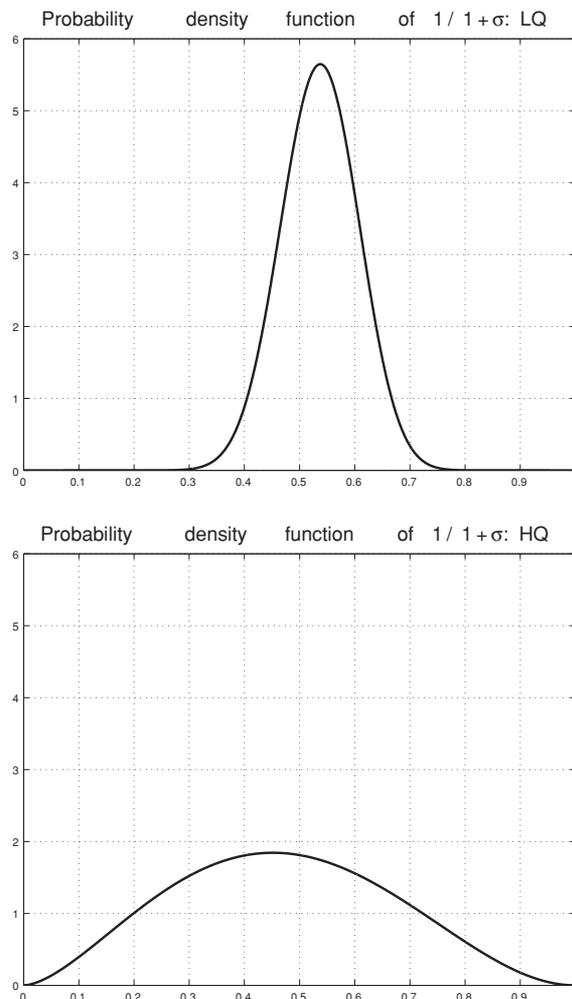


図9 ドキュメンテーション・クオリティによるランダム効果モデルのパラメータ分布の比較

式 (22) のパラメータ分布  $\frac{1}{1+\sigma} \sim \text{Beta}(r, s)$  を描画した。上図 (LQ) はドキュメンテーションが不完全な保険契約ポートフォリオのパラメータの確率分布である。下図 (HQ) はドキュメンテーションが完全な保険契約ポートフォリオのパラメータの確率分布である。

を期待しているかもしれない。もしその推論が正しいなら、保険解約は家計が悪化し始めたときに発生するため、失業の先行事象であろう。そして説明変数に採用した完全失業率がラグをとる必要がなかったのは自然な結論である。

本論文の貢献は主に3つある。第1にモデルの選択基準を明らかにした。モデルの論理的整合性から分布の再生性が要請される。現象論的な観点からは計数データの *overdispersion* の振舞いを再現できなければならない。実証分析をおこなうために、ポートフォリオに含まれる保険契約数が時間とともに変化することを補正できるモデルでなければならない。これらの基準のもとでNB1とよばれる負の2項モデルが適切であることを明らかにした。第2にパネル分析によりパラメータ推定し、実証分析をおこなったことである。保険の解約に関する実証は本研究が最初であり、数多くの新しい事実を明らかにした。第3に保険契約申込書のドキュメンテーション・クオリティによって保険解約特性が異なることを明らかにしたことである。

本モデルは保険解約件数の予測に使えるため、保険契約ポートフォリオ中の個々の契約が毎月何件ずつ解約するかを簡単にシミュレーションできる。特定のシナリオ分析はもちろんモンテカルロ・シミュレーションに

より統計的な特性を算出することも可能である。今回のモデル化では保険金額は同額として保険解約件数をモデル化しているが、保険総額が一定ではない場合にモデルを拡張することは容易である。モデルは保険契約者のカテゴリ別にグループ化して構成した保険契約ポートフォリオについて、毎月の月初の保険契約数・月中の保険解約件数がわかればモデル・パラメータの推定は可能である。すなわちモデルは保険契約ポートフォリオごとの統計量だけを入力データとしており、昨今問題になるような個人データへのアクセスを必要としないという利点もある。

## 謝辞

本研究を進めるにあたって貴重な助言を下された早稲田大学の森平爽一郎先生、広島大学の矢野順治先生に感謝する。最後に2名の匿名レフェリーからの助言・コメントに感謝する。

## 参考文献

- [1] 森平 爽一郎 (2008), 「保険負債の評価」, 『欧州の先進的な保険リスク管理システムに関する研究会報告書』第5章, 金融庁金融研究研修センター。
- [2] Albizzati, M.-O., H. Geman (1994), "Interest rate risk management and valuation of the surrender option in life insurance policies", *The Journal of Risk and Insurance*, Vol.61, pp.616–637.
- [3] Bacinello, A.R. (2003a), "Pricing guaranteed life insurance participating policies with annual premiums and surrender options", *North American Actuarial Journal*, Vol.7, No.3, pp.1–17.
- [4] Bacinello, A.R. (2003b), "Fair valuation of a guaranteed life insurance participating contract embedding a surrender option", *The Journal of Risk and Insurance*, Vol.70, pp.461–487.
- [5] Bacinello, A.R. (2005), "Endogenous model of surrender conditions in equity-linked life insurance", *Insurance: Mathematics and Economics*, Vol.37, pp.270–296.
- [6] Bütow, S. (2001), "Securitization in life, health and personal accident reinsurance", hannover re, working paper.
- [7] Cameron, A.C. and P.K. Trivedi (1998), *Regression analysis of count data*, Cambridge University Press.
- [8] Cowley, A. and J.D. Cummings (2005), "Securitization of life insurance assets and liabilities", *The Journal of Risk and Insurance*, Vol.72, pp.193–226.
- [9] Grosen, A. and P.L. Jørgensen (2000), "Fair valuation of life insurance liabilities: The impact of interest rate guarantees, surrender options, and bonus policies", *Insurance: Mathematics and Economics*, Vol.26, pp.37–57.
- [10] Hausman, J., B.H. Hall and Z. Griliches (1984), "Econometric models for count data with an application to the patents-R&D relationship", *Econometrica*, Vol.52, pp.909–938.
- [11] Kim, C. (2005), "Modeling Surrender and Lapse Rates with Economic Variables", *North American Actuarial Journal*, Vol.9, No.4, pp.56–70.
- [12] Shen, W. and H. Xu (2005), "The valuation of unit-linked policies with or without surrender options", *Insurance: Mathematics and Economics*, Vol.36, pp.79–92.
- [13] Winkelmann, R. (2008), *Econometric analysis of count data, Fifth Edition*, Springer-Verlag.

# Panel Count Data Model for Insurance Surrenders

Yusho KAGRAOKA

*Musashi University, Toyotama-kami 1-26-1, Nerima-ku, Tokyo 176-8534, Japan*

---

## Abstract

Surrender of insurance policies is modelled and empirically investigated by applying a count panel data model. Model selection criteria is a reproductive property of distribution and an overdispersion property, and a negative binomial model, NB1 (variance is a linear function of mean), is selected. Model parameters are estimated by the fixed-effects and random-effects panel data analysis. Macroeconomic variables as well as policyholders' characteristics are taken as explanatory variables. It is found that the number of surrender is mainly explained by change in unemployment rates, age of policy, and seasonality. It is also examined how documentation quality affects surrender behaviors. For insurance policies with incomplete documentation, the age of policy predominates other explanatory variables. Surrender of insurance negative binomial models count panel data model

---