

# 深層学習モデルに拡張した非線形Cox回帰モデル と東京賃貸物件市場への応用

込山 湧士<sup>†</sup>・松田 安昌<sup>†</sup>

(受付 2024 年 7 月 1 日；改訂 9 月 4 日；採択 9 月 17 日)

## 要 旨

本論文では、2019年3月から2021年3月に収集した東京賃貸物件市場データに対して、賃貸物件の広告掲載期間を生存時間とみなしてCox回帰モデルによる生存時間分析を行った。Cox回帰モデルを深層学習モデルによって非線形に拡張し、賃貸物件のもつ流動性と価格弾力性の時空間特性を非線形に表現し、COVID-19パンデミックが東京賃貸市場に及ぼした影響を評価した。流動性と価格弾力性のみをニューラルネットで表現し、解釈の可能性をなるべく残したところに本モデルの特長がある。本モデルによる分析の結果、パンデミック後には流動性の減少および価格弾力性の増加傾向が観察された。

キーワード：価格弾力性, Cox回帰モデル, 時空間モデル, ニューラルネットワーク, ハザード関数, 流動性.

## 1. はじめに

生存時間分析は、もともと治療法や薬の有効性を患者の生存時間の結果から評価する方法である。Cox (1972)が提案したCox回帰が、最も標準的な生存時間分析モデルとして広く応用されている。生存関数  $S(t)$  を時間  $t$  以上生きている確率とする。ハザード関数  $h(t)$  は対数生存関数の微分にマイナスをつけたもの  $(-d \log S(t)/dt)$  と定義され、時点  $t$  における死亡リスクを意味する。Cox回帰では、ある患者のハザード関数を、その患者の属性や特徴量をベクトル  $x$  とその線形結合  $x\beta$  を使って

$$h(t) = h_0(t) \exp(x\beta)$$

と表す。このモデルの特徴はハザード関数が時間の関数  $h_0(t)$  (基準ハザードとよぶ) と特徴量の関数  $\exp(x\beta)$  の積で与えられることである。ここで  $h_0(t)$  には特定のモデルを仮定しない。生存時間  $(t)$  と特徴量  $(x)$  のデータを収集することで、モデルのパラメータ  $\beta$  と基準ハザードを推定することが可能であり、推定したモデルによって患者の属性に応じた死亡リスクを評価することができる。データセットには打ち切りと呼ばれる生存時間が観測されないデータがしばしば存在するが、Cox回帰では、部分尤度法によってパラメータ推定に使うことができる。打ち切りを死亡リスクの評価に生かすことができる点がCox回帰の重要な特徴の一つである。

本論文は、Cox回帰を賃貸物件市場に応用して、賃貸物件市場の流動性および価格弾力性を分析することを動機とし、両特性の時空間分析を可能とする非線形モデルに拡張することを目

<sup>†</sup> 東北大学 大学院経済学研究科：〒980-8576 宮城県仙台市青葉区川内 27-1

的とする。ある賃貸物件の情報が公開されてから契約が決まって掲載が終了するまでの広告掲載期間をその賃貸物件の生存時間とみなすことにより、生存時間モデルを賃貸物件市場の分析に応用することができる。特に、賃貸物件市場の流動性(物件の成約のしやすさの度合い)と価格弾力性(価格の変動によって物件の需要が変化する度合い)が物件の掲載時期と所在地によって変動することを許容するハザード関数を提案する。ハザード関数の時変性と空間異質性を深層学習モデルを用いて学習を行い、パラメータを推定する。なお、本モデルを推定するために、2019年3月より2021年3月までの2年間にわたり、東京都23区内の広告記録を収集した。総計1,762,243件、うち487,643件が広告終了日が未記載の打ち切り物件であり、物件の所在地(緯度、経度)と様々な特徴量を含んでいる。

Cox 回帰を機械学習を使って拡張した研究は数多く行われている。例として Cox 回帰に Ridge 正則化を加えた Verweij and Van Houwelingen (1994)によるモデルや Lasso 正則化を加えた Tibshirani (1997)によるモデルがある。また、罰則を課しながらも特定の特徴量を高次元生存モデルに組み込む推定手法として Binder and Schumacher (2008)による CoxBoost がある。さらに、Faraggi and Simon (1995)は、従来のハザード関数をニューラルネットワークで非線形化する試みを行った。その後提案された Katzman et al. (2018)による DeepSurv によって従来の線形モデルを上回る成果が得られている。これを画像処理に応用した Zhu et al. (2016)による DeepConvSurv のように非構造データに対する応用も行われている。また、ニューラルネットワークによる非線形対応させた Cox 回帰を大規模データに適用するためミニバッチ処理での損失計算を提案した Kvamme et al. (2019)による Cox-Time がある。他に生存時間モデルに深層学習を応用した研究として、分布に仮定を置かず生存時間の分布を直接学習する Lee et al. (2018)による DeepHit、DeepHit に RNN と temporal attention mechanism を加えて長期にわたる、もしくは繰り返し行われた測定結果に対応できるようにした Lee et al. (2020)による Dynamic-DeepHit がある。しかし、これらのニューラルネットワークを用いたモデルは生存時間の予測精度向上を目指したものであり、特徴量の解釈性には乏しいという課題がある。本研究は DeepSurv と Cox-Time とに着想を得て、Cox 回帰の非線形化を行う。回帰関数の定数項と賃貸物件価格の偏回帰係数だけを緯度、経度、掲載時期に依存させ、ニューラルネットワークで非線形化する。それぞれを流動性と価格弾力性として解釈を与えられるという点に特徴がある。

賃貸物件市場に生存時間分析を適用した研究として、いくつかの重要な研究がある。Deng et al. (2003)は、1987年から1998年のBLS-CPI住宅サンプルからの独自データを用いて、賃貸住宅における居住期間の推定にCox回帰を使用した。また、Bhuiyan and Hasan (2016)は、観察期間内の過去の住宅販売情報をもとに、家の販売確率を予測するための生存分析に着想を得た監督付き回帰(Cox回帰)モデルを提案した。さらに、Yilmaz et al. (2022)は、イギリスの賃貸市場における流動性が時間と場所による需要の変動にどのように反応するかを探求した。Minzat et al. (2018)は、リスティング日から販売契約までの期間を予測するために生存時間分析を適用した。特に、Bhuiyan and Hasan (2016)の研究では、物件の説明文に対して Topic Modeling や LDA, Doc2Vec を適用している。しかし、賃貸物件市場に生存時間分析を適用した研究の中で、推定モデルとして深層学習を適用した例は、私たちの調べた限り、みつからない。

本論文の構成を述べる。第2章でCox回帰を時空間に拡張するため深層学習モデルを組み込んだモデルを提案する。空間への拡張は物件所在地(緯度、経度)、時間への拡張は広告公開時期から構成する月ダミー変数を用いる。第3章では本モデルを東京都23区賃貸物件市場の分析に応用し、時空間分析を行う。加えて、データはコロナ前後の区間を含んでいるので、賃貸物件市場をコロナ禍との関連に注視して分析を行う。第4章にて結論を述べる。

## 2. Cox 回帰と深層学習

### 2.1 Cox 回帰の時空間拡張

賃貸物件の募集期間を生存時間と考えて、まず従来の Cox 回帰モデルをあてはめる。ここで、募集期間は掲載終了日から広告公開日の間の日数であり、公開日と終了日が同じ物件は生存時間が 0 とみなす。基準ハザード関数を  $h_0(u)$  とおくと、Cox 回帰では、物件  $i$  のハザード関数  $h_i(u)$  は次の式で表現される。

$$(2.1) \quad \log h_i(u) = \log h_0(u) + x_i \beta, i = 1, \dots, n,$$

ここで、 $x_i$  は物件  $i$  の属性(説明変数あるいは特徴量ともよばれる)、 $\beta$  は偏回帰係数である。Cox 回帰では、基準ハザード、偏回帰係数いずれも物件  $i$  の公開時期  $t_i$ 、所在地  $s_i$  に依存せず、時間的、空間的双方に静的なハザード関数を仮定している。

賃貸市場において、募集期間を時間的、空間的に静的とするモデルは制約が極めて強い。そこで、ハザード関数が、物件の所在地  $s$  と広告公開時点  $t$  に依存することを許容するモデルを考案する。物件  $i$  の広告公開時期を  $t_i$ 、所在地を  $s_i$  とするとき、ハザード関数  $\tilde{h}_i(u)$  を、

$$(2.2) \quad \log \tilde{h}_i(u) = \log \tilde{h}_0(u) + f(x_i, s_i, t_i), i = 1, \dots, n,$$

とする。ここで、非線形関数  $f(x_i, s_i, t_i)$  のままでは表現すべき関数のクラスが広すぎるため、解釈不能になってしまう。

物件のもつ流動性、価格弾力性が時間的、地理的に変動するモデルを考える。特徴量  $x_i$  の中から物件の 1 平米あたり単価の対数値を取り出し  $\log P_i$  とおき、その他を  $z_i$  とおく。  $f(x_i, s_i, t_i)$  を線形近似したとき、定数項と  $\log P_i$  の係数だけを  $s_i, t_i$  に依存させたハザード

$$(2.3) \quad \log \tilde{h}_i(u) = \log \tilde{h}_0(u) + f_0(s_i, t_i) + f_1(s_i, t_i) \log P_i + z_i \beta, i = 1, \dots, n,$$

を考える。  $f_0(s, t)$  は、物件の条件をそろえた時の対数ハザード関数を表し、流動性の時空間変動を表すと解釈できる。  $f_1(s, t)$  は、他の条件をそろえて価格を 1% あげたときのハザードの変化率を表し、価格弾力性の時空間変動を表すと解釈できる。なお、  $f_1(s, t)$  は通常は負値をとるため、  $-f_1(s, t)$  を価格弾力性と定義することに注意する。

$f_0(s, t), f_1(s, t)$  のニューラルネットによる非線形モデリングを考える。ここで時間  $t$  を連続時間ではなく、離散時間であると仮定する。つまり、ハザード関数は  $t$  で連続的に変動するのではなく、月次や週次のように離散的に変動すると考える。全物件データを含む時間区間を  $I = [0, T]$  とおき、  $K$  個の背反な時間区間

$$I = I_1 \cup \dots \cup I_K$$

に分割する。  $I_j(t)$  を  $t$  が  $I_j$  に含まれていたら 1, 含まれなければ 0 をとるダミー変数とし、  $f_0, f_1$  に次の制約

$$(2.4) \quad \begin{aligned} f_0(s, t) &= f_0(s, I_1(t), \dots, I_K(t)), \\ f_1(s, t) &= f_1(s, I_1(t), \dots, I_K(t)), \end{aligned}$$

を課す。  $f_0, f_1$  に 4 層からなる深層学習モデルをあてはめる。緯度と経度の 2 次元からなる  $s$  と  $K$  個のダミーからなる  $t$  を入力とし、出力を 1 次元とするニューラルネットである。  $f_0$  と  $f_1$  には同じ構造のニューラルネットを使用するが、ネットワークの重みは変わることに注意する。深層学習モデルの構造を図 1 に示す。実際のノード数はデータ分析のセクションで述べる。

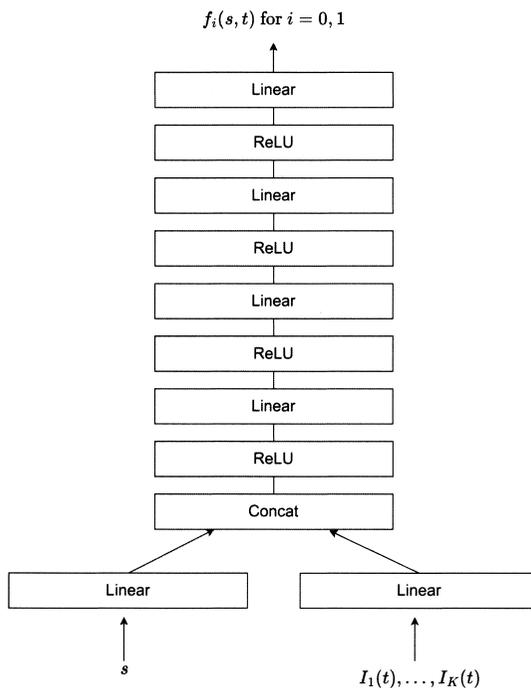


図 1. 深層学習モデルの構造.

## 2.2 推定法

前節で提案したハザード関数(2.3)では、物件の流動性と価格弾力性が広告公開時点  $t$  と所在地  $s$  に依存することを許し、 $t$  を離散化して深層学習モデルを適用した。本節では、深層学習モデルの推定法を紹介する。

本モデルでは、Cox (1972) による部分尤度法をそのまま応用することができる。物件  $i = 1, \dots, N$  の募集期間を  $u_i$  とおく。募集期間が  $u$  以上の全物件の集合を  $R(u)$  とおく。Cox (1972) は、ハザード関数が以下の線形関数

$$\log h_i(u) = \log h_0(u) + x_i \beta$$

で与えられるモデルに対して、基準ハザード関数  $h_0(u)$  を集約した対数尤度関数、つまり対数部分尤度、

$$Q(\beta) = \sum_{i=1}^N \left\{ x_i \beta - \log \left( \sum_{j \in R(u_i)} \exp x_j \beta \right) \right\}.$$

を構成し、対数部分尤度を最大化するパラメータ推定法を提案している。なお、広告期間が同じ日数となる物件が複数存在しても、物件  $i$  ごとに定義通りに部分尤度を評価するものとする。これは、Breslow (1974) によるタイデータを含む場合の部分尤度近似法に対応する。

次に、Cox 回帰モデルの部分尤度を非線形ハザード関数に拡張する。ハザード関数が  $(s, t)$  に依存するモデル(2.3)に対して、ベースラインハザードを除いた部分を  $g(s, t) = f_0(s, t) + f_1(s, t) \log P + z\beta$  とおく。このハザード関数の対数部分尤度に、Cox の対数部分尤度をそのまま拡張し、

$$(2.5) \quad Q(f) = \sum_{i=1}^N \left\{ f(s_i, t_i) - \log \left( \sum_{j \in R(u_i)} \exp f(s_j, t_j) \right) \right\}.$$

と与える．ここで  $R(u_i)$  は広告期間が  $u_i$  以上の打ち切りを含むすべての物件をあらわす．この対数尤度関数の負値を本モデルの深層学習における損失関数として，ニューラルネットワークの学習をおこなう．

### 2.3 C-index

モデルの評価指標として the concordance-index (C-index) を採用する．C-index は Harrell et al. (1982) によって提案され，生存時間解析において広く使用されている評価指標である．推定したモデルがどのくらい良く生存時間を予測できるかを示す．C-index の定義域は  $[0, 1]$  で，1 に近いほど良いとされ，ランダムな予測では値が 0.5 となる．本論文で用いるデータにはタイデータを多く含むので，Kang et al. (2015) によるタイデータに対応した C-index を使って評価する．

我々の非線形 Cox 回帰モデルを，物件の特徴量を  $x$  として，

$$h(u) = h_0(u) \exp(f(x))$$

とおくとき，C-index は以下で定義される．物件  $i$  の実測生存時間を  $T_i$ ，対数ハザードからベースラインを除いた  $f(x_i)$  に対して，

$$(2.6) \quad C = \frac{1}{2} \left( \frac{\sum_{i \neq j} \text{sgn}(f(x_i), f(x_j)) \text{csgn}(T_i, \delta_i, T_j, \delta_j)}{\sum_{i \neq j} \text{csgn}(T_i, \delta_i, T_j, \delta_j)^2} + 1 \right),$$

ここで， $\delta_i$  は物件  $i$  が打ち切りであれば 0 をとるダミー変数， $I$  を指示関数として，

$$\begin{aligned} \text{sgn}(f(x_i), f(x_j)) &= I(f(x_i) \geq f(x_j)) - I(f(x_i) \leq f(x_j)), \\ \text{csgn}(T_i, \delta_i, T_j, \delta_j) &= \delta_j I(T_i \geq T_j) - \delta_i (T_i \leq T_j) \end{aligned}$$

である．タイデータを含む場合の定義なので意味が分かりにくい，直観的には，物件  $i, j$  の全ての組み合わせに対し，実測生存時間の大小と同じハザード関数の大小をもつ組み合わせの割合を意味している．

## 3. 東京賃貸物件市場の分析

### 3.1 データ

東京都 23 区内の賃貸物件を対象に，2019 年 3 月より 2021 年 3 月までの 2 年間の広告記録を収集した．広告記録には，物件名，住所，緯度，経度，公開日，広告終了日，賃料(円)，面積(平米)，最寄り駅からの徒歩距離(分)，物件の位置する階数，物件の入る建物の階数，近辺の便利施設(例えばショッピングセンターや小学校)等の情報を含んでいる．本節では広告の公開日から掲載終了日までの日数をその物件の生存時間とみなし，賃貸物件に生存時間分析を応用して賃貸市場の地域特性とその時間変化を分析する．

本データには，総計 1,762,243 件の賃貸物件が収集され，そのうち 487,643 件が広告終了日が未記載の打ち切りデータとなっている．図 2 に，広告期間の分布を，図 3 には，23 区ごとの物件数を，打ち切りなしと打ち切りの各場合に集計したものを，図 4 には，物件数を広告公開月と掲載終了月の月別に集計したものを，それぞれ表している．広告期間の分布をみると，7 日，14 日，21 日，... の 1 週間ごとにジャンプがみられる．広告の掲載を 1 週間単位で契約している

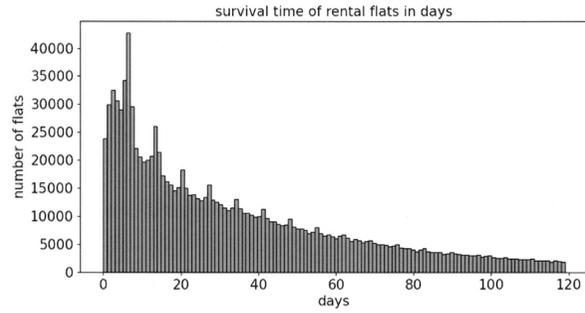


図 2. 東京 23 区賃貸物件の広告期間の分布.

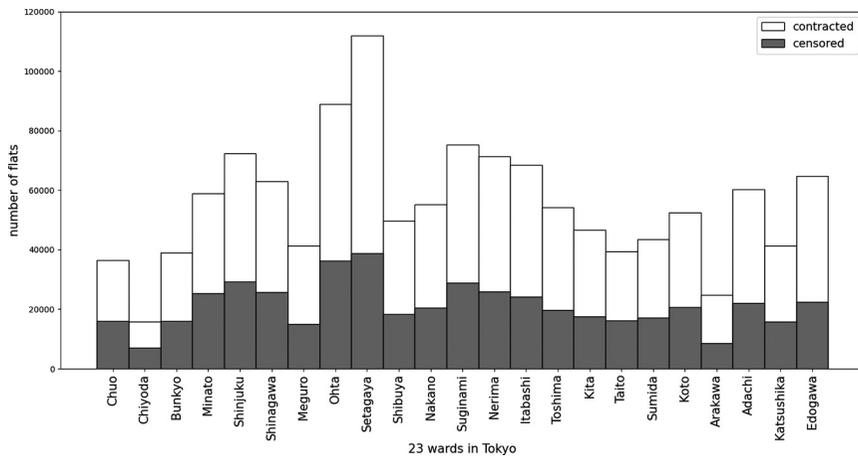


図 3. 東京 23 区賃貸物件数の分布. 区ごとの契約件数と打ち切り件数.

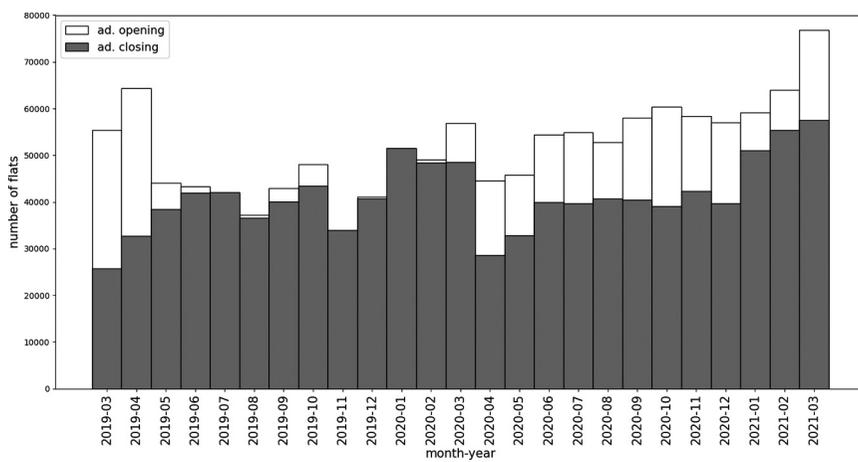


図 4. 東京 23 区賃貸物件数の分布. 月別の広告スタートと終了の賃貸物件数.

貸主がいるためであると予想される。広告期間が 0 日、つまり公開したその日に掲載が終了する人気物件も 20,000 件以上あることに注意を要する。23 区ごとの分布では、世田谷区が最大の物件数で 10 万件を超え、大田区、新宿区が続き、逆に千代田区が最小で荒川区が続き、ともに 2 万件未満の物件数しかない。最後に月別の分布では、1 月、2 月、3 月に公開、終了ともに物件数が増加する傾向にあり、年度前の 3 か月が賃貸物件取引が最も活発な時期であることを示している。

### 3.2 ハザード関数の時空間分析

2019 年 3 月から 2021 年 3 月まで全 25 か月間にわたり収集し、打ち切りを含め総計 1,762,243 件の賃貸物件の広告期間に対して、ハザード関数の流動性と価格弾力性が時空間的に変動するモデル (2.3) をあてはめ、流動性と価格弾力性のもつ地域特性の時間変化を分析する。ここで、 $z$  (対数家賃  $\log P$  以外の説明変数) として、物件の築年数、最寄り駅からの徒歩距離 (分単位)、1F ダミー、物件の入る建物の総階数を選択した。

流動性と弾力性にはそれぞれ 4 層からなるニューラルネット (図 1) を応用し、時間 (月別ダミー) と所在地 (緯度と経度) によるハザード関数の時空間変動を把握することを目指している。ここで、レイヤーごとにノード数を以下の式で設定した。ReLU( $x$ ) は活性化関数 ReLU による変換を表し、Linear $_{a \rightarrow b}(x)$  は  $a$  次元入力から  $b$  次元へ出力する線形変換とし、2 次元の  $s$ 、 $K$  次元 (本データでは 38 次元) からなる月ダミー変数  $I_1, \dots, I_K$  を入力とし、 $f_0$  を出力とする関数を以下の 4 層からなるニューラルネットにあてはめた。

$$\begin{aligned} h_1 &= \text{Linear}_{2 \rightarrow 128}(s), \\ h_2 &= \text{Linear}_{K \rightarrow 10}(I_1(t), \dots, I_K(t)), \\ h_3 &= \text{Concatenate}(h_1, h_2), \\ h_4 &= \text{ReLU}(\text{Linear}_{128+K \rightarrow 256}(h_3)), \\ h_5 &= \text{ReLU}(\text{Linear}_{256 \rightarrow 128}(h_4)), \\ h_6 &= \text{ReLU}(\text{Linear}_{128 \rightarrow 64}(h_5)), \\ f_0 &= \text{Linear}_{64 \rightarrow 1}(h_6). \end{aligned}$$

$f_1$  を出力とするニューラルネットの構造も全く同じであるが、重みは異なることに注意する。

本モデルの 4 層ニューラルネットを記述するために必要な 154,896 個のパラメータの推定を行う。全物件 1,762,243 件に対し、本モデルの対数部分尤度 (2.5) の負値を損失関数とし、損失関数を最小化するパラメータ推定を行う。

全サンプル 1,762,243 件の 60% を学習用データセット、20% を評価用データセット、残りの 20% をテスト用データセットにランダムに分割する。学習用データセットを用いて損失関数を最小化する深層学習モデルのパラメータを学習し、そのパラメータを使ってテスト用データセットの損失関数および C-index を評価し当てはまりの良さの基準とする。学習用データセットに対し確率的勾配降下法によって損失関数を最小化する推定を行った。評価用データセットは、学習時のエポック数やハイパーパラメータを決めるために用いられ、ここでは評価用データセットに対する C-index を最小化するように選択した。

従来の Cox 回帰モデル (2.1) を本モデル (2.3) のベンチマークとする。従来の Cox 回帰モデルは、本モデルにおいて流動性と弾力性をともに定数に仮定したときに一致する。本モデルと Cox 回帰モデルの損失関数 (負の対数部分尤度) と C-index を評価し、表 1 に示す。本モデルは、Cox 回帰モデルに比べて、テストデータに対する損失関数、C-index とともに改善している。損失

表 1. 損失関数(負の対数部分尤度)と C-index による Cox 回帰と深層学習モデルの比較.

| モデル          | 損失関数 (負の対数部分尤度) |           |           | C-index |       |       |
|--------------|-----------------|-----------|-----------|---------|-------|-------|
|              | 学習              | 評価        | テスト       | 学習      | 評価    | テスト   |
| Cox 回帰 (2.1) | 5,232,433       | 1,590,255 | 1,588,879 | 0.547   | 0.547 | 0.546 |
| 深層学習 (2.3)   | 5,202,859       | 1,579,736 | 1,581,290 | 0.612   | 0.608 | 0.607 |

関数および C-index の基準からみて、本モデルは線形モデルを改善していることがわかる。ここで、本モデルの学習データとテストデータに対する損失関数、C-index にわずかな差しかめられないことに注意する。過学習することなくパラメータ推定が行われたことを示している。

式(2.4)で定義した流動性  $f_0(s, t)$  と価格弾力性  $f_1(s, t)$  の推定値を、特定の地域  $s_0$  に固定した場合と 23 区全体に分けて示す。なお、流動性は(2.3)式のベースラインハザードを除いた定数項であり、他の説明変数の単位によって変化する量であるから、流動性自体より流動性の差に意味があることに注意する。言い換えると、流動性は、同条件下での賃貸物件の契約の決まりやすさを表す。逆に、価格弾力性は、(2.3)式において、対数家賃  $\log P_i$  の偏回帰係数の負値で定義されるため、値自体に意味があり、他の条件をそろえて家賃が 1% 上昇したときのハザードの減少率を表すことに注意する。

まず、 $s_0$  を杉並区荻窪 1 丁目に固定した場合の流動性  $f_0(s_0, t)$  と価格弾力性  $f_1(s_0, t)$  の月次時系列プロットを、図 5 と図 6 に示す。破線は、ベンチマークである Cox 回帰モデルによる流動性であり、時空間変化しない定数で評価される。なお、価格弾力性は  $f_1(s, t)$  の負値であるから、値が小さい(高い)ほど弾力性が高い(低い)ことに注意する。図 5 より、荻窪の流動性は新型コロナウイルスが周知され始めた 2020 年 1 月以降に急減し、10 月以降に回復が始まっていることがわかる。また、通常の Cox 回帰による流動性とほぼ同じ水準である。荻窪の物件は、ほぼ平均的な流動性を持つことを示している。図 6 より、価格弾力性は 2020 年 2 月に 0.6 から 1.7 に急増した。コロナ禍による賃貸需要の急減のため、家賃に敏感に反応するようになったと考えられる。2021 年 1 月より回復傾向がみられるが、2019 年以前の値までには戻していない。また、Cox 回帰で評価された弾力性に比べて、荻窪の弾力性は高く評価されていることがわかる。荻窪では、平均からみて家賃の設定に強く反応する傾向があることを示している。

次に流動性  $f_0(s, t)$  と価格弾力性  $f_1(s, t)$  を、月別に東京都 23 区地図上に、それぞれ図 7 と図 8 に示す。なお、図は 2015 年度の国勢調査の境界データを基に加工して作成した。

図 7 の流動性の推移より、コロナ前(2019/9, 2019/12)とその後(2020/3, 2020/9, 2020/12, 2021/3)を比較する。全体的に 2019 年 3 月から 2020 年 9 月に大きく減少した流動性が 2020 年 12 月以降、徐々に回復していく様子がみられる。

図 8 の価格弾力性の推移において、コロナ禍の 2020 年 3 月以降、全体的に価格弾力性の上昇が観察される。ただし、都心のやや北に位置する浅草近辺がコロナ禍前後によらず高く評価されている。浅草エリアは都心でありながら比較的求めやすい家賃で賃貸物件が出されている地域であるといわれており、浅草では価格の設定に賃貸需要が影響されやすいことが考えられる。2020 年 9 月に江東区、大田区の湾岸地域の弾力性が広く上昇している。東京オリンピックの 2021 年 8 月の開催が危ぶまれた時期であり、湾岸地域の賃貸需要がオリンピック延期リスクに影響を受けたものと考えられる。

#### 4. 終わりに

本論文では、大量の賃貸物件データを収集し、広告期間を生存時間とみなして Cox 回帰を応用した。古典的な Cox 回帰では、ハザード関数は物件特徴量の線形関数であらわされるため、



図 5. 杉並区荻窪 1 丁目における流動性の時間変動。

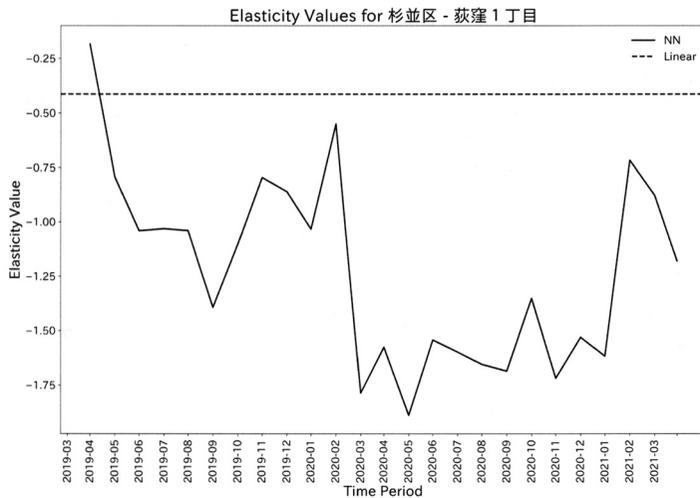


図 6. 杉並区荻窪 1 丁目における価格弾力性の時間変動。

定数項にあたる流動性と対数家賃の偏回帰係数にあたる価格弾力性は、時間と所在地によらない。そこで、ハザード関数の流動性と価格弾力性が時空間変動することをゆるすハザード関数を考案した。ここで、時間を物件広告公開月の月ダミー変数、空間を物件所在地(緯度、経度)であらわし、流動性と価格弾力性を、時間と所在地を入力とするニューラルネットワークモデルで表現した。本モデルを学習させ、流動性および価格弾力性を評価し、東京都 23 区の賃貸物件市場の特徴をコロナ禍との関連を見ながら分析を行った。コロナ禍の発生後、都心部において賃貸物件市場は急速に流動性を下げ、価格弾力性を上げたが、徐々に回復する傾向がみられた。以上のように、広告掲載期間を生存時間とみなす生存時間分析により、賃貸物件市場を数理的な分析対象とすることができる。

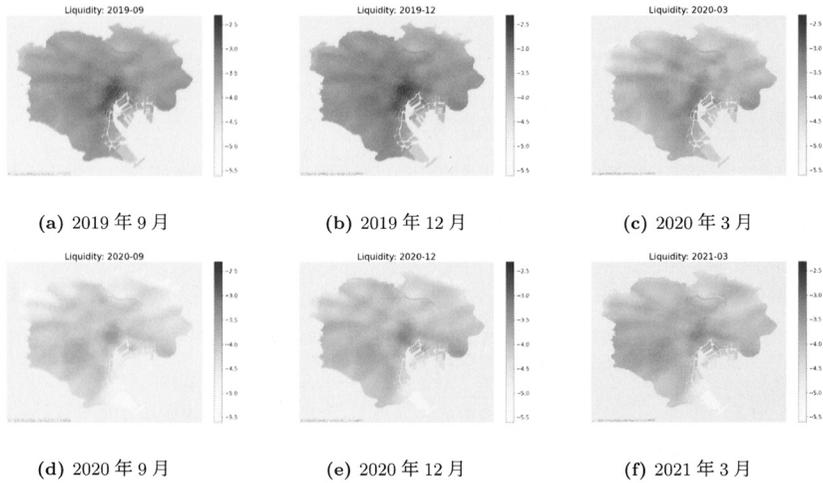


図7. 東京賃貸物件市場の流動性の推移.

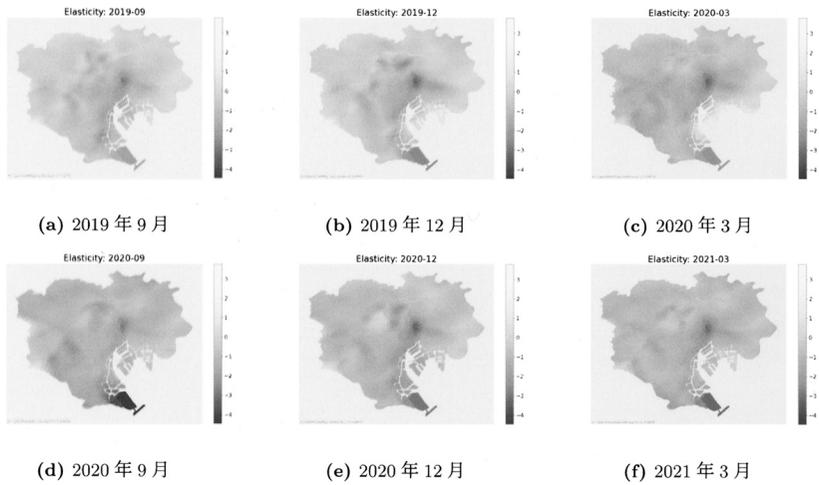


図8. 東京賃貸物件市場の価格弾力性の推移.

## 謝 辞

2名の査読者よりいただいた貴重なコメントに感謝します。

## 参 考 文 献

- Bhuiyan, M. and Hasan, M. A. (2016). Waiting to be sold: Prediction of time-dependent house selling probability, *2016 IEEE International Conference on Data Science and Advanced Analytics (DSAA)*, 468–477, <https://doi.org/10.1109/DSAA.2016.58>.
- Binder, H. and Schumacher, M. (2008). Allowing for mandatory covariates in boosting estimation of

- sparse high-dimensional survival models, *BMC Bioinformatics*, **9**(1), <https://doi.org/10.1186/1471-2105-9-14>.
- Breslow, N. (1974). Covariance analysis of censored survival data, *Biometrics*, **30**(1), 89–99, <https://doi.org/10.2307/2529620>.
- Cox, D. R. (1972). Regression models and life-tables, *Journal of the Royal Statistical Society, Series B (Methodological)*, **34**(2), 187–220, <https://doi.org/10.1111/j.2517-6161.1972.tb00899.x>.
- Deng, Y., Gabriel, S. A. and Nothaft, F. E. (2003). Duration of residence in the rental housing market, *The Journal of Real Estate Finance and Economics*, **26**(1), 267–285, <https://doi.org/10.1023/A:1022987010545>.
- Faraggi, D. and Simon, R. (1995). A neural network model for survival data, *Statistics in Medicine*, **14**(1), 73–82, <https://doi.org/10.1002/sim.4780140108>.
- Harrell, J., Frank, E., Califf, R. M., Pryor, D. B., Lee, K. L. and Rosati, R. A. (1982). Evaluating the yield of medical tests, *JAMA*, **247**(18), 2543–2546, <https://doi.org/10.1001/jama.1982.03320430047030>.
- Kang, L., Chen, W., Petrick, N. A. and Gallas, B. D. (2015). Comparing two correlated C indices with right-censored survival outcome: A one-shot nonparametric approach, *Statistics in Medicine*, **34**(4), 685–703, <https://doi.org/10.1002/sim.6370>.
- Katzman, J. L., Shaham, U., Cloninger, A. et al. (2018). DeepSurv: Personalized treatment recommender system using a Cox proportional hazards deep neural network, *BMC Medical Research Methodology*, **18**(1), 24, <https://doi.org/10.1186/s12874-018-0482-1>.
- Kvamme, H., Borgan, Ø. and Scheel, I. (2019). Time-to-event prediction with neural networks and Cox regression, *Journal of Machine Learning Research*, **20**, 1–30.
- Lee, C., Zame, W., Yoon, J. and van der Schaar, M. (2018). DeepHit: A Deep learning approach to survival analysis with competing risks, *Proceedings of the AAAI Conference on Artificial Intelligence*, **32**(1), 2314–2321, <https://doi.org/10.1609/aaai.v32i1.11842>.
- Lee, C., Yoon, J. and van der Schaar, M. (2020). Dynamic-DeepHit: A deep learning approach for dynamic survival analysis with competing risks based on longitudinal data, *IEEE Transactions on Bio-medical Engineering*, **67**(1), 122–133, <https://doi.org/10.1109/TBME.2019.2909027>.
- Minzat, D., Breaban, M. and Luchian, H. (2018). Modeling real estate dynamics using survival analysis, *2018 20th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC)*, 215–222, <https://doi.org/10.1109/SYNASC.2018.00042>.
- Tibshirani, R. (1997). The lasso method for variable selection in the Cox model, *Statistics in Medicine*, **16**(4), 385–395, [https://doi.org/10.1002/\(sici\)1097-0258\(19970228\)16:4<385::aid-sim380>3.0.co;2-3](https://doi.org/10.1002/(sici)1097-0258(19970228)16:4<385::aid-sim380>3.0.co;2-3).
- Verweij, P. J. M. and Van Houwelingen, H. C. (1994). Penalized likelihood in Cox regression, *Statistics in Medicine*, **13**(23-24), 2427–2436, <https://doi.org/10.1002/sim.4780132307>.
- Yilmaz, O., Talavera, O. and Jia, J. Y. (2022). Rental market liquidity, seasonality, and distance to universities, *International Journal of the Economics of Business*, **29**(2), 223–239, <https://doi.org/10.1080/13571516.2022.2033078>.
- Zhu, X., Yao, J. and Huang, J. (2016). Deep convolutional neural network for survival analysis with pathological images, *2016 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 544–547, <https://doi.org/10.1109/BIBM.2016.7822579>.

## Deep Learning Extensions of Cox Regression and Their Applications to Rental Property Market in Tokyo

Yuji Komiyama and Yasumasa Matsuda

Graduate School of Economics and Management, Tohoku University

This paper applies Cox regression models to dataset in Tokyo rental property market collected from March, 2019 till March, 2021. We try a deep learning extension of Cox regression models to let liquidity and price elasticity of rental properties be dependent on space and time in a nonlinear way, and analyze the effects by the pandemic of COVID-19 to the rental market. We apply neural networks only to express liquidity and elasticity from which interpretability retains in our model. We find by the model applications that the pandemic makes the tendencies of decreasing liquidity increasing and elasticity all over the Tokyo 23 wards.