

SASによるWebデータの統計解析 XMLデータ入出力と不動産バリュエーションマップ作成への応用

松島純之介* 石島博** 渡部育恵

*中外製薬株式会社 臨床企画推進部

**中央大学大学院 国際会計研究科

Web Data Analysis with SAS

Input and Output of XML Data and Application to Real Estate Valuation Map

Junosuke Matsushima*, Hiroshi Ishijima**, Ikue Watanabe

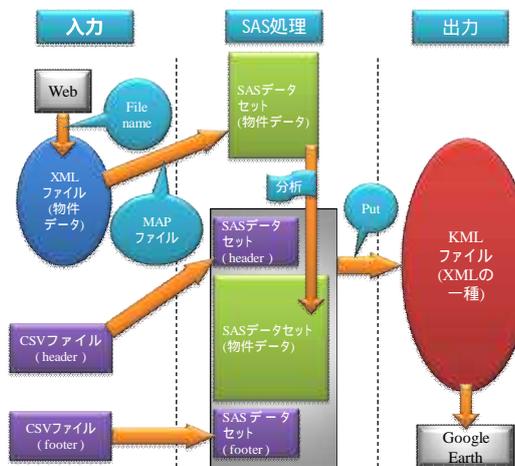
*Clinical Research Planning Department, Chugai Pharmaceutical Co., LTD.

**Graduate School of International Accounting, Chuo University

July 27, 2010

発表内容

- 本発表の要旨と背景
- Webデータの統計解析
 - Web上のデータを取り込む方法 (第2節)
 - XML形式のデータの入力 (第3節)
- 応用例: 不動産バリュエーションマップ作成
 - 不動産の価格とリスクを評価する統計モデル (第4節)
 - XML形式のデータの出力: Google Earthへの出力例 (第5節)
- まとめ



要旨

- SASによるWeb上のXMLデータの解析法を紹介し、例として不動産バリュエーションマップを作成する。
- Webデータから不動産の価格とリスクを推定し、結果をGoogle Earth上に表示する。

キーワード: XML, Filename, Valuation Map, Google Earth

背景

- 新統計法の施行など→分析可能なデータが膨大に開示。
- Web上のデータは凄まじい勢いで増加→Web上のデータを有効に利用・活用するフレームワークを構築することが重要になってきている。
- Web上のデータをSASで扱う際の2つの問題
 1. Web上のデータをどのようにしてSASで取り込んだらいいか
 2. Web上のデータ形式をどのように扱ったらいいか
 - Web上のデータの多くはXML(eXtensible Markup Language)で記述される。

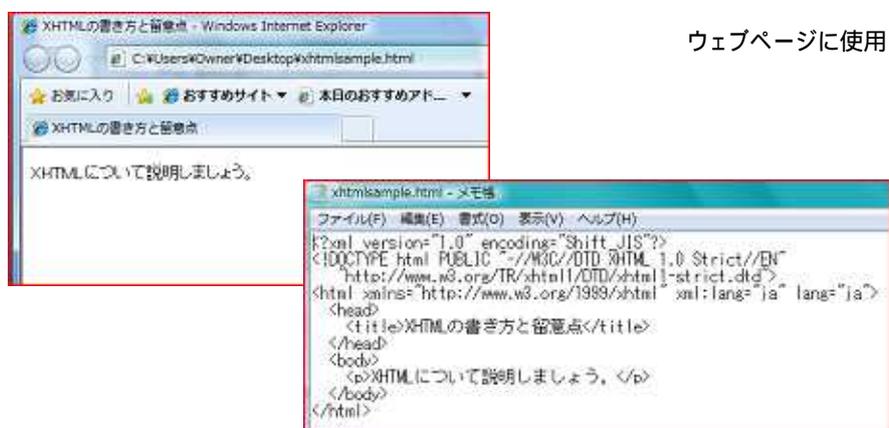
XML

- データの中にデータ構造や意味を示す情報(タグ)を含めるマークアップ言語の一つ
- 通信形式やファイルフォーマットの定義などに使われる
- XMLで記述されるWeb上のデータの例
 - XHTML
 - RSS
 - Web API

4

XHTML

ウェブページに使用



```
<?xml version="1.0" encoding="Shift_JIS"?>
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN"
  "http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd">
<html xmlns="http://www.w3.org/1999/xhtml" xml:lang="ja" lang="ja">
<head>
  <title>XHTMLの書き方と留意点</title>
</head>
<body>
  <p>XHTMLについて説明しましょう。</p>
</body>
</html>
```

<http://www.kanzaki.com/docs/html/xhtml1.html> を基に作成

5

RSS

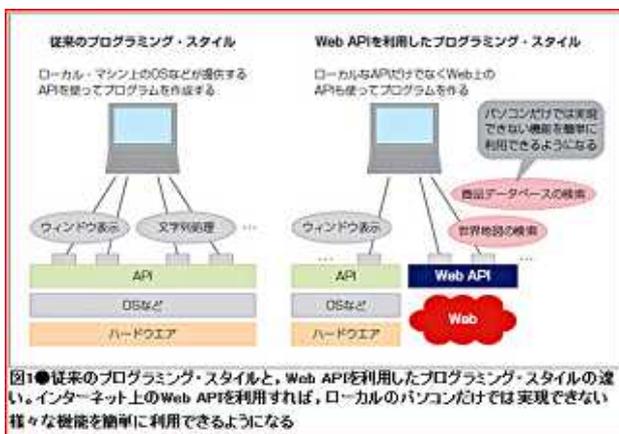


ニュースやブログのトピック配信

```
<rss version="2.0">
<channel>
<title>Yahoo!ニュース - 国内 - レスポンス</title>
<link>http://nsearch.yahoo.co.jp/bin/search?to=2&...
<description>Yahoo! JAPANのニュースに掲載されてい...
<language>ja</language>
<copyright>Copyright (C) 2010 Yahoo Japan Corpora...
<lastBuildDate>Sun, 27 Jun 2010 01:15:15 +0900</l...
<image>
<title>Yahoo!ニュース</title>
<url>http://i.yimg.jp/images/news/ynews_s.gif</ur...
<link>http://headlines.yahoo.co.jp/hk</link>
<width>101</width>
<height>18</height>
</image>
<item>
<title>【緊急寄稿】 ニュル・マイスター 成瀬弘先生...
<link>http://rd.yahoo.co.jp/rss/1/headlines/nee/rs...
<category>人</category>
<pubDate>Fri, 25 Jun 2010 20:13:28 +0900</pubDate>
</item>
</channel>
```

<http://public.news.yahoo.co.jp/rss/>
より抜粋

Web API



サーバー上のデータベースを利用

<http://itpro.nikkeibp.co.jp/article/COLUMN/20060925/248842/> より抜粋

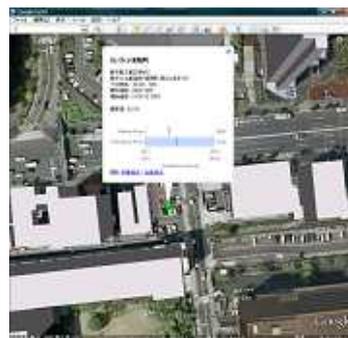
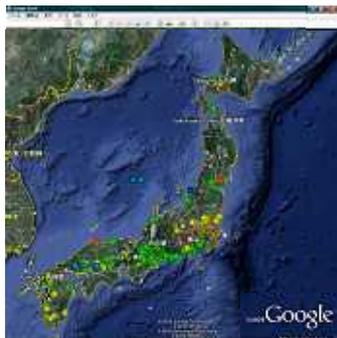
本発表のモチベーション

- Web上のデータをSASで扱う際の2つの問題
 1. Web上のデータをどのようにしてSASで取り込んだらいいか?
 2. Web上のデータ形式をどのように扱ったらいいか?
 - Web上のデータの多くは、XMLで記述される。
 - Web上のデータを扱うには、XMLの扱い方を知る必要がある。
 - 現状のSASユーザーにとって、XMLは必ずしも馴染みがあるものではない。

8

不動産バリュエーションマップ: 概要

- 「不動産バリュエーションマップ」: 不動産のXML形式のデータから、不動産の価格とリスクを推定し、高度な地球儀・地図アプリケーション上に表示するフレームワーク



9

不動産バリュエーションマップ: 背景

1. 利用可能なWebデータの増加
 - 従来、不動産のデータはごく一部のしか利用できなかったが、最近では、Webから誰もが利用できるようになった。
2. 強力な統計解析パッケージの必要性
 - 最新のファイナンス理論に基づいて不動産価格とリスクを推定するには複雑なモデルが必要であり、SAS等の強力な統計解析パッケージが不可欠。
3. 高度なWebアプリケーションの登場
 - 不動産データを表示する際、Google EarthやGoogle Mapsといった優れた地球儀・地図アプリケーションを利用できるようになった。

10

不動産バリュエーションマップ: 作成の3ステップ

1. Webデータの利用
 - マンション情報サイト「スマッチ!(株式会社リクルートのWebサービス)」のAPIを利用したXML形式のデータをSASに取り込み、分析可能なSASデータセットに変換
2. 統計解析
 - ファイナンス理論に基づいた統計モデルに基づき、不動産の価格とリスクをSASによって推定(不動産バリュエーション)
3. Webアプリケーションの活用
 - SASによる分析結果という高付加価値を加えたKML(XMLの一種)ファイルを出力し、Google EarthやGoogle Maps上に表示

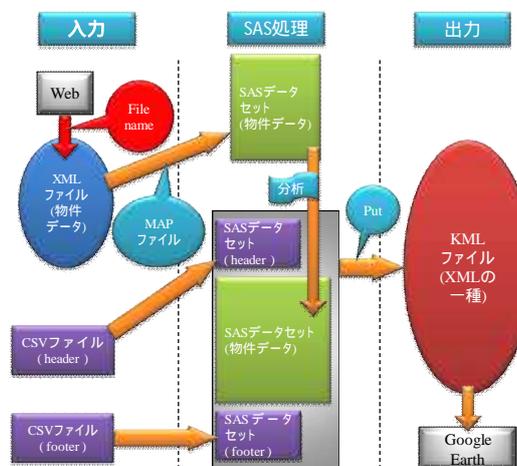
11

クラウド時代のSASの有効活用

- Web上で公開される膨大なオリジナルXMLファイルをSASに取り込む
- 分析可能なSASデータセットに変換
- ファイナンス理論に基づいた統計モデルをSASによって分析
- SASによる分析結果という高い付加価値を加えたXMLファイルを出力
- Web上で展開されるGoogle等の強力なツールと連携・利用

発表内容

- 本発表の要旨と背景
- **Webデータの統計解析**
 - Web上のデータを取り込む方法 (第2節)
 - XML形式のデータの入力 (第3節)
- 応用例: 不動産バリュエーションマップ作成
 - 不動産の価格とリスクを評価する統計モデル (第4節)
 - XML形式のデータの出力: Google Earthへの出力例 (第5節)
- まとめ



Web上のデータの取り込み

→ Filenameステートメントを用いる

filename <データセット名> **url** '<URL>' (オプション);

<URL>にデータが存在するURLを指定する。
(オプション)で、エンコード等を指定する。

リクルートWEBサービス

The screenshot shows the RECRUIT WEB SERVICE homepage. The 'WHAT'S NEW' section contains the following updates:

- 2010/04/28**
スレッドレスポンス検索APIの検索条件に「リード線(keyword)」を追加しました。
スレッドレスポンス検索APIのレスポンスフィールドに「コメント数(comment_count)」を追加しました。
スレッドレスポンス検索APIのレスポンスフィールドに「最新コメント(newly_comment)」を追加しました。
- 2010/04/28**
スレッドレスポンス検索APIのレスポンスフィールドに「検索画像(sample_image)」を追加しました。
カンボジアSAFのレスポンスフィールドに「スレッドID(thread_id)」を追加しました。
- 2010/04/28**
JSONフォーマット出力時に特殊文字のエスケープを行うように修正しました。
- 2010/03/26**
物件番号のデフォルト値を「開拓対象」に変更しました。
検索条件のフィールド名に同値化処理を追加しました。
- 2010/03/26**
検索フィールド名が「検索対象」に変更されたため、検索条件のフィールド名に同値化処理を追加しました。
- 2010/03/26**
検索フィールド名が「検索対象」に変更されたため、検索条件のフィールド名に同値化処理を追加しました。

<http://webservice.recruit.co.jp/smacth/> より抜粋

スマッチ! 物件情報検索API

1 物件情報検索API

2 リクエストURL

`http://api.smatch.jp/apartment/`

3 検索クエリ

パラメータ	項目名	説明	必須	値
key	APIキー	APIを利用するために割り当てられたキーを設定します。	○	
area	エリアコード	地域コード。関東、関西、東海、その他から選びます。エリアマスタAPIを参照できます。		000 関東 050 東海 060 関西 999 その他
pref	都道府県コード	都道府県に割り当てられたコード番号で検索します。指定できるコード番号は都道府県マスタAPIを参照。	※1	(例)0303598
city	市郡区コード	市郡区に割り当てられたコード番号で検索します。指定できるコード番号は市郡区マスタAPIを参照。	※1	(例)04

<http://webservice.recruit.co.jp/smatch/reference.html#a1to> より抜粋

16

物件情報データベース URL

4 サンプルクエリ

物件ID 030990010003243 のデータを取得する

```
http://api.smatch.jp/apartment/?key=[APIキー]&area=030&apartment=030990010003243
```

※実際にご利用いただくには新規登録をした上で、取得したご自身のAPIキーをGETパラメータに加えていただく必要があります。

<http://webservice.recruit.co.jp/smatch/reference.html#a1to> より抜粋

17

物件情報

```

<results xmlns="http://api.smatch.jp/" >
  <api_version>1.0</api_version>
  <results_available>1</results_available>
  <results_returned>1</results_returned>
  <results_start>1</results_start>
  <apartment>
    <id>000990010003249</id>
    <name>THE TOKYO TORRE</name>
    <address>東京都中央区橋どき</address>
    <area>
      <name>関東</name>
      <code>030</code>
    </area>
    <pref>
      <name>東京都</name>
      <code>0309508</code>
    </pref>
    <city>
      <name>中央区</name>
      <code>02</code>
    </city>
    <station>
      <name>橋どき</name>
      <code>50065</code>
      <line_area>
        <name>都心部 (山手線・地下鉄)</name>
        <code>09001</code>
      </line_area>
      <line>
        <name>都営大江戸線</name>
        <code>0070</code>
      </line>
    </station>
  </apartment>

```

Web上のデータの取り込み

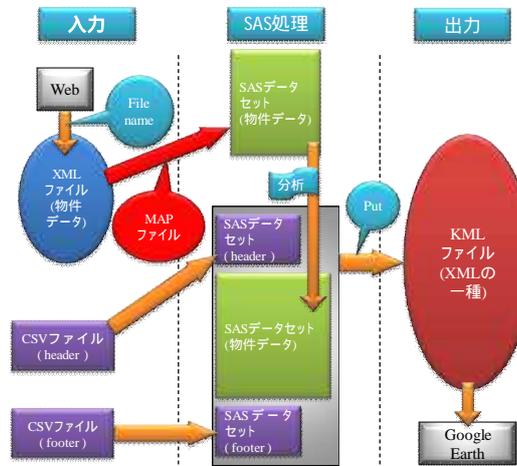
➔ Filenameステートメントを用いる

filename <データセット名> url '<URL>' (オプション);

行	SASスクリプト	コメント
1	options validvarname=any;	変数の設定
2	filename junk url '「 URL 」' encoding="UTF-8" termstr=LF debug lrecl=8192;	Filenameステートメント。「URL」にデータが存在するURLを指定する。
3	data data1;	Infileステートメントによりデータを読み込む。
4	infile junk length=len;	
5	input record \$varying8192. len;	
6	run;	
7	data;	「C:\mysasfiles\xml\in.xml」に、取得したデータを保存する。
8	set data1;	
9	file 'C:\mysasfiles\xml\in.xml' termstr=LF encoding="UTF-8" lrecl=8192;	
10	put record;	
11	run;	

発表内容

- 本発表の要旨と背景
- **Webデータの統計解析**
 - Web上のデータを取り込む方法 (第2節)
 - **XML形式のデータの入力 (第3節)**
- 応用例: 不動産バリュエーションマップ作成
 - 不動産の価格とリスクを評価する統計モデル (第4節)
 - XML形式のデータの出力: Google Earthへの出力例 (第5節)
- まとめ



XMLが特定の3層構造の場合

行	XMLスクリプト	コメント												
1	<?xml version="1.0" encoding="windows-1252" ?>	構造: 以下のような3層構造になっている。 TABLE CLASS NAME SEX AGE <table border="1" style="margin-top: 10px;"> <thead> <tr> <th>Obs</th> <th>AGE</th> <th>SEX</th> <th>NAME</th> </tr> </thead> <tbody> <tr> <td>1</td> <td>25</td> <td>Male</td> <td>Taro</td> </tr> <tr> <td>2</td> <td>24</td> <td>Female</td> <td>Hanako</td> </tr> </tbody> </table>	Obs	AGE	SEX	NAME	1	25	Male	Taro	2	24	Female	Hanako
Obs	AGE		SEX	NAME										
1	25		Male	Taro										
2	24		Female	Hanako										
2	<TABLE>													
3	<CLASS>													
4	<NAME> Taro </NAME>													
5	<SEX> Male </SEX>													
6	<AGE> 25 </AGE>													
7	</CLASS>													
8	<CLASS>													
9	<NAME> Hanako </NAME>													
10	<SEX> Female </SEX>													
11	<AGE> 24 </AGE>													
12	</CLASS>													
13	</TABLE>													

スクリプト 3.2-1

SASスクリプト(特定の3層構造)

行	SASスクリプト	コメント
1	Libname trans xml 'C:\mysas files\xml\basic.xml';	読み込むXMLファイルを指定
2	Libname myfiles 'C:\mysas files\xml';	ライブラリとするフォルダを指定
3	data myfiles.class;	XMLファイルのCLASSに含むデータをSASデータセットとして読み込む。
4	set trans.class;	
5	run;	

スクリプト 3.2-2

特定の3層構造の場合はごくまれ。

【問題】→ 一般的にはどのようにすればいいか？

XMLが様々な形式の場合

【答え】→ MAPファイルを用いる。

XMLの構造	SASデータセットの構造
results	
apartment	APARTMENTS
id	id
...	
area	
name	areaname
...	...

表 3.3

XMLスクリプトとMAPスクリプト

XMLスクリプト

```
<?xml version="1.0" encoding="UTF-8" ?>
<results xmlns="http://api.smatch.jp/">
  <api_version>1.0</api_version>
  ...
  <apartment>
    <id>030990066142534</id>
    ...
    <area>
      <name>関東</name>
      <code>030</code>
    </area>
  </apartment>
  ...
  </results>
```

↑スクリプト 3.3-1

MAPスクリプト

```
<?xml version="1.0" ?>
<SXLEMAP version="1.2">
  <TABLE name="APARTMENTS">
    <TABLE-PATH syntax="xpath">
      /results/apartment
    </TABLE-PATH>
    <COLUMN name="id">
      <PATH>/results/apartment/id</PATH>
      <TYPE>character</TYPE>
      <DATATYPE>STRING</DATATYPE>
      <LENGTH>15</LENGTH>
    </COLUMN>
    <COLUMN name="areaname">
      <PATH>/results/apartment/area/name</PATH>
      <TYPE>character</TYPE>
      <DATATYPE>STRING</DATATYPE>
      <LENGTH>8</LENGTH>
    </COLUMN>
  </TABLE>
</SXLEMAP>
```

スクリプト 3.3-2 →

SASスクリプト

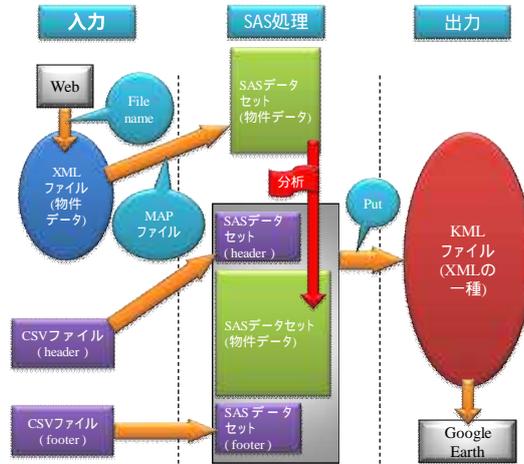
行	SASスクリプト	コメント
1	filename RESULTS 'C:\mysasfiles\XML\in.xml' ;	読み込むXMLファイルを指定
2	filename MAP 'C:\mysasfiles\XML\in.map' ;	読み込むMAPファイルを指定
3	libname RESULTS xml xmlmap=MAP xmlconcatenate=yes ;	MAPファイルの設定
4	libname MYFILES 'C:\mysasfiles\XML' ;	ライブラリとするフォルダを指定
5	data MYFILES.APARTMENTS;	XMLファイルのAPARTMENTに含むデータをSASデータセットとして読み込む。
6	set RESULTS.APARTMENTS;	
7	run;	

スクリプト 3.3-3

- MAPファイルを用いれば、XMLがどのような形でも、上記スクリプトによりSASデータセットに変換することができる。

発表内容

- 本発表の要旨と背景
- Webデータの統計解析
 - Web上のデータを取り込む方法 (第2節)
 - XML形式のデータの入力 (第3節)
- 応用例: 不動産バリュエーションマップ作成
 - 不動産の価格とリスクを評価する統計モデル (第4節)
 - XML形式のデータの出力: Google Earthへの出力例 (第5節)
- まとめ



不動産の価格とリスクの統計モデル 1

- Box-Cox変換付き線形混合効果モデル

$$H_{ij}^*(\lambda) = \sum_{l=1}^N \underbrace{x_{ij}^{(l)}}_{(1)} \underbrace{\beta^{(l)}}_{(3)} + \sum_{k=1}^K \underbrace{x_{ij}^{(k)}}_{(2)} \underbrace{(\beta^{(k)} + v_i^{(k)})}_{(2')} + \underbrace{\varepsilon_{ij}}_{(4)}$$

- (1) 被説明変数: べき乗(Box-Cox)変換した不動産価格
- (2) 第1の説明変数: 属性量(延床面積、築年数、駅徒歩)
- (2') 第1の係数: 属性単価(固定単価+変動単価という混合効果)
- (3) 第2の説明変数: 不動産クラス(地域よるクラス)を表すダミー変数
- (3') 第2の係数: 不動産クラスに応じた価格プレミアム(固定効果)
- (4) 誤差: 正規分布に従う

不動産の価格とリスクの統計モデル 2

- 不動産価格評価の理論
 - 拡張されたヘドニックモデル(石島・前田, 2009)
- 理論価格の要件1: 「不動産価格は線形でなければならない」
 - ↓
 - 動的一般均衡モデルに基づいた、完全競争下における均衡価格の条件
- 理論価格の要件2: 「属性単価は同じでなければならない」
 - ↓
 - 一物一価の原則という条件。不動産に依らず共通した同一の属性単価である必要

28

不動産の価格とリスクの統計モデル 3

~ 理論価格と市場価格とのギャップを捉える工夫 その1 ~

- 理論価格の要件1: 「不動産価格は線形でなければならない」
 - ↓原則
 - 動的一般均衡モデルに基づいた、完全競争下における均衡価格の条件
 - ↓しかし
 - 現実市場では、不動産価格は歪んでいる可能性がある
 - ↓そこで
- Box-Cox変換付き線形混合効果モデルによる対応
 - 不動産価格にべき乗(Box-Cox)変換を施す:

$$H_{ij}^*(\lambda) := \begin{cases} \frac{H_{ij}^\lambda - 1}{\lambda} & (\lambda \neq 0 \text{ のとき}) \\ \log H_{ij} & (\lambda = 0 \text{ のとき}) \end{cases}$$

【λの解釈】

λ = 1 のとき

→ 線形 = 理論価格

λ ≠ 1 のとき

→ 価格の歪みを表す

→ λ = 0 のとき = 対数価格

29

不動産の価格とリスクの統計モデル 4

～ 理論価格と市場価格とのギャップを捉える工夫 その2 ～

- 理論価格の要件2: 「属性単価は同じでなければならない」
 - ↓ 原則
 - 一物一価の原則という条件。不動産に依らず共通した同一の属性単価である必要
 - ↓ しかし
 - 現実市場では、属性単価は不動産によって異なっている可能性がある
 - ↓ そこで
- Box-Cox変換付き線形混合効果モデルによる対応
 - 属性単価を混合効果モデルで表現する:

$$\sum_{k=1}^K \underbrace{x_{ij}^{(k)}}_{\text{属性}k\text{の量}} \left(\underbrace{\beta^{(k)}}_{\text{固定単価}} + \underbrace{v_i^{(k)}}_{\text{変動単価}} \right)$$

属性kの単価を、
 → 固定単価 = 不動産によらず共通
 → 変動単価 = 不動産によって変動
 に分けて推定する。

30

データと推定法

- 不動産データ
 - スマッチ! (株式会社リクルートのWebサービス)より2010年4月初旬に取得
 - 関東地区のマンションの募集価格と属性(延床面積、築年数、駅徒歩)のデータ
 - 不動産データを地域によって、8つのクラス(不動産クラス)に分類
 - 不動産クラス*i*に属する*j*番目の不動産価格(H_{ij})と、その属性*k*の量($x_{ij}^{(k)}$)に関するデータ
- モデルの推定
 - 統計モデル: 「Box-Cox変換付き線形混合効果モデル」
 - 利用プロシージャ: PROC MIXED
 - ただし、 λ の推定は、Gurka et al (2006)の方法で行った。

31

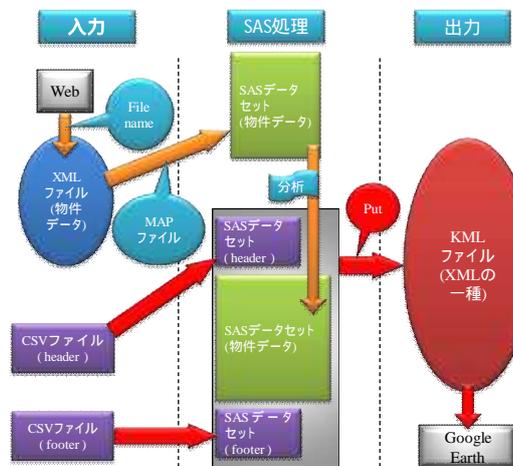
推定結果

	Box-Cox変換付き線形混合効果モデル ($\lambda = -0.05$)						
	東京都心3区	東京都区部	東京都下	神奈川県	千葉県	埼玉県	その他関東
東京都心3区 ダミー変数	41,624.73 (<.0001)						
東京都区部 ダミー変数		39,704.78 (<.0001)					
東京都下 ダミー変数			39,375.93 (<.0001)				
神奈川県 ダミー変数				38,992.04 (<.0001)			
千葉県 ダミー変数					39,216.69 (<.0001)		
埼玉県 ダミー変数						38,593.43 (<.0001)	
その他関東 ダミー変数							36,738.00 (<.0001)
延床面積 (平米)	69.07 (<.0001)	78.78 (<.0001)	64.82 (<.0001)	73.00 (<.0001)	57.44 (<.0001)	68.87 (<.0001)	72.20 (<.0001)
築年数(年)	21.14 (0.8385)	-3.21 (0.9662)	-94.52 (0.5295)	-246.52 (0.0245)	-347.16 (0.0200)	-346.69 (0.0100)	-320.73 (0.0280)
駅徒歩(分)	-19.27 (0.1709)	-20.40 (0.0106)	-33.63 (0.0267)	-49.42 (0.0012)	-33.40 (0.0205)	-31.73 (0.0329)	-23.42 (0.0852)
価格インデックス	8,552 (<.0001)	4,856 (<.0001)	3,940 (<.0001)	3,945 (<.0001)	3,518 (<.0001)	3,458 (<.0001)	2,581 (<.0001)

表 4

発表内容

- 本発表の要旨と背景
- Webデータの統計解析
 - Web上のデータを取り込む方法 (第2節)
 - XML形式のデータの入力 (第3節)
- 応用例: 不動産バリュエーションマップ作成
 - 不動産の価格とリスクを評価する統計モデル (第4節)
 - XML形式のデータの出力: Google Earthへの出力例 (第5節)
- まとめ



XML形式のデータの出力 ～ Google Earthへの出力例 ～

- 評価した不動産の価格とリスクを、Google Earth上に表示
 → [Web](#)
 → [kml](#)

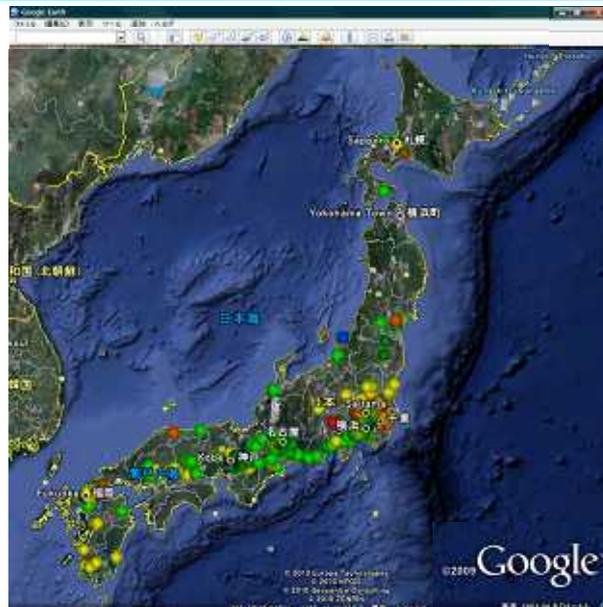
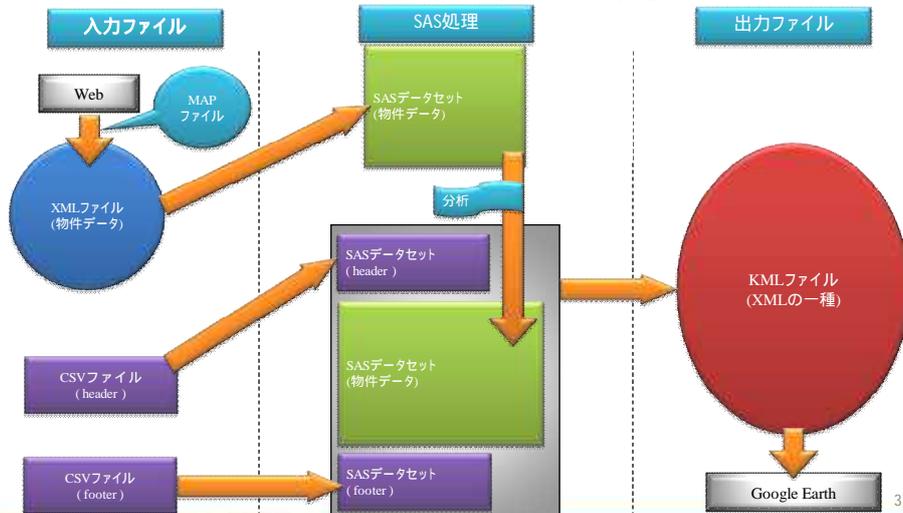


図 5-1



図 5-2 36

XMLファイルの入力 → SAS処理 →
→ XMLファイルの出力



KMLファイルの構造

行	KMLスクリプト	コメント
1	<?xml version="1.0" encoding="utf-8"?>	構造: 以下のような構造になっている。
2	<KML xmlns="http://earth.google.com/kml/2.1">	DOCUMENT
3	<DOCUMENT>	STYLE (アイコンSTYLEの定義)
4	<STYLE ID="ICON3">	ICONSTYLE
5	<ICONSTYLE >...</ICONSTYLE>	PLACEMARK
6	</STYLE>	NAME
7	<STYLE ID="ICON4">	STYLEURL (アイコンSTYLEの呼び出し)
8	<ICONSTYLE >...</ICONSTYLE>	DESCRIPTION
9	</STYLE>	(説明バルーンに表示される ユーザー定義コンテンツ)
10	<PLACEMARK>	POINT
11	<NAME> ガーラ・レジデンス浅草 </NAME>	COORDINATES (経度、緯度)
12	<STYLEURL> #ICON4 </STYLEURL>	
13	<DESCRIPTION> ... </ DESCRIPTION>	
14	<POINT>	
15	<COORDINATES>	
16	139.799965,35.717301	
17	</ COORDINATES>	
18	</POINT>	
19	</ PLACEMARK >	
(20)	...	
21	</DOCUMENT>	
22	</KML>	

header部分

物件データ部分

footer部分

また、KMLファイルは以下の3部分に分けられる。
header部分 (アイコンSTYLE等の定義部分)
<PLACEMARK>の前
物件データ部分
<PLACEMARK>の繰り返し
footer部分 (終了タグ)
<PLACEMARK>の後 スクリプト 5-1

38

SASスクリプト: 物件データ部分

- SASデータセットに、KMLスクリプトのタグを文字変数として加える。

```
data MYFILES.APARTMENTS;
set MYFILES.APARTMENTS;
keep id name ... tag1 tag2 ...;
tag1 = '<Placemark><name>';
tag2 = '</name><styleUrl>';
...
run;
```

39

SASスクリプト: Header, Footer部分

- Header部分、Footer部分を外部ファイルに記述しておく。

```
%LET kml_header_file="C:\mysasfiles\output\kml_header.csv";  
%LET kml_footer_file="C:\mysasfiles\output\kml_footer.csv";
```

- Header部分、Footer部分をSASデータセットとして読み込む。

```
proc import datafile=&kml_header_file  
  out=MYFILES.kml_header replace;  
  getnames=no;  
run;  
proc import datafile=&kml_footer_file  
  out=MYFILES.kml_footer replace;  
  getnames=no;  
run;
```

40

SASスクリプト: KMLファイルの出力

- 出力するKMLファイルを指定

```
%LET kml_file="C:\mysasfiles\out.kml";
```

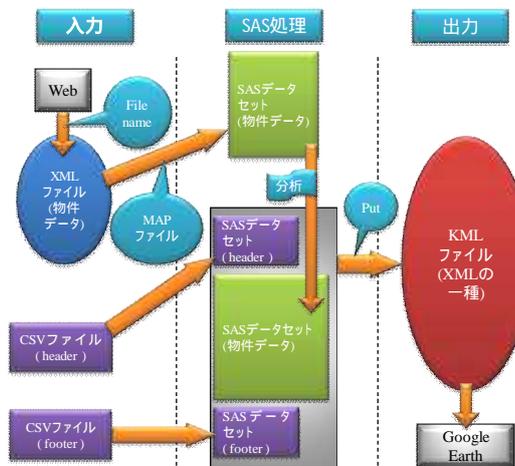
- kml_header(header部分)、APARTMENTS(物件データ部分)、kml_footer(footer部分)の順に、SASデータセットを結合しKMLファイルに出力する。

```
data;  
  set MYFILES.kml_header  
  MYFILES.APARTMENTS MYFILES.kml_footer;  
  file &kml_file termstr=LF encoding="UTF-8"  
  lrecl=8192;  
  put var1 tag1 name tag2 address ...;  
run;
```

41

発表内容

- 本発表の要旨と背景
- Webデータの統計解析
 - Web上のデータを取り込む方法 (第2節)
 - XML形式のデータの入力 (第3節)
- 応用例: 不動産バリュエーションマップ作成
 - 不動産の価格とリスクを評価する統計モデル (第4節)
 - XML形式のデータの出力: Google Earthへの出力例 (第5節)
- **まとめ**



42

まとめ

- 本発表では、今後ますます重要だがその基礎が十分に浸透しているとは言い難い、Web上のXMLデータをSASに取り込む方法と、SASにおけるXMLファイルの入出力を行う方法を示した。要点は以下の3点である。
 1. Web上のデータの取り込みには、Filenameステートメントを用いる。
 2. 任意のXMLファイルをSASデータセットとして入力するにはMAPファイルを用いる。
 3. SASデータセットからXMLファイルに出力する際には、XMLの「Header部分」「繰り返し部分」「Footer部分」という構造に着目する。
- 応用例として不動産バリュエーションマップを作成した。

43