

不動産流通経営協会研究助成報告書

東日本大震災被災地域における  
平成 25 年住宅・土地統計調査の  
被災・復興状況を考慮した調査方法の研究

2013 年 2 月

東京大学大学院 情報学環  
東京大学 空間情報科学研究センター  
山田 育穂

## 1. はじめに

本研究は、平成 25 年実施予定の住宅・土地統計調査(以下、住宅調査)において、東日本大震災の被災地域の被災・復興状況を適切に考慮して、統計の信頼性を確保するための調査方法を提案・検証することを目的としている。

総務省統計局が 5 年ごとに実施する住宅調査は、国内の住宅など人の居住する建物の実態、現住居以外の住宅・土地の保有状況等そこに居住する世帯の実態を把握するためのもので、その結果は住生活に関わる諸施策の基礎資料として利用されている(総務省統計局 <http://www.stat.go.jp/data/jyutaku/2008/index.htm>)。次回調査は平成 25 年に予定されており、平成 23 年 3 月の震災により生活の基盤となる施設が甚大な被害を受けた地域では、復興の過程を正確に把握し、また復興計画を適切に遂行していく上で、通常以上に重要な役割を担っている。一方、住宅調査は標本調査であり、その調査区の抽出は直前の国勢調査(次回住宅調査では平成 22 年国勢調査)の調査結果に基づいて行われるため、調査区の特徴が国勢調査時点から大きく変化してしまった被災地域においては、従来通りの調査手法では統計の信頼性が損なわれる可能性が高い。従って、被災地域における調査手法を整備することは緊急の課題であり、本研究では、浅見・山田(2011)が提案した手法を基に具体的な検討を行う。

住宅調査では、直前の国勢調査の調査区を基本抽出単位とする層化 2 段抽出法により、調査対象住戸が決定される(総務省統計局 [http://www.stat.go.jp/data/jyutaku/2008/pdf/sui\\_1.pdf](http://www.stat.go.jp/data/jyutaku/2008/pdf/sui_1.pdf))。これは、まず国勢調査調査区を、世帯数、住宅の所有関係など国勢調査で得られる情報に基づいていくつかの層に分類し、各層から都市規模などにより定められた抽出率で標本単位区を抽出した後、標本単位区内の住戸から標本住戸を抽出するという方法である。層別抽出法の利点は、類似した抽出単位(この場合は国勢調査調査区)をまとめて層とすることにより、特定の特徴を持つ抽出単位が適切な割合で選択されるようにして、標本誤差を抑えることにある。この標本抽出の過程においては、国勢調査から住宅調査までの約 3 年間の間に、調査区の特徴が大きく変化していないことが暗黙の前提となっている。が、東日本大震災の被災地域、特に津波による重篤な被害を受けた地域では、この前提は非現実的なものである。更に、仮設住宅の建設地など、従来の層分類の枠組みに収まらない地域も発生し、復興過程においても時々刻々と地域属性が変化していくことが予想される。

こうした調査区属性の変化による調査結果への影響を最小限に抑えるため、浅見・山田(2011)は以下の 3 つの手法を提案した。

- (1) 被災地域特有の状況を反映した新たな層を設定し、従来の層分類とは別に標本

## 抽出を行う方法

(2) 被災地域内において調査区の抽出率を上乗せする方法

(3) 平成 22 年国勢調査による地域属性データを、2 次データを用いて補正する方法

復興過程の被災地域では、ほとんどの住宅が消失し新たに再建される地域、甚大な津波被害による都市計画的な建築規制等により実質的に住宅がなくなる地域、仮設住宅が建設される地域など、従来の住宅調査の層分類では対応できない地域が多数発生する。手法(1)は、こうした地域を新たな層として定めて別途標本抽出を行うことで、被災地域の調査区の代表性を確保すると共に、被災地域外における調査区抽出への影響を避けるものである。

また、住宅調査では、調査区の抽出率は層ごとに市区町村の人口規模により定められる(総務省統計局 [http://www.stat.go.jp/data/jyutaku/2008/pdf/sui\\_1.pdf](http://www.stat.go.jp/data/jyutaku/2008/pdf/sui_1.pdf))。しかしながら被災地域においては、同一層に属する調査区であっても、その特性に従来想定されている以上のばらつきがあるものと考えられる。よって手法(2)は、被災地域内の調査区抽出率を上昇させることで、被災地域の調査結果の信頼性の確保を目指す。

一方、避難者の転出入など震災後の人口移動は、直接的・物理的な震災被害を受けていない地域の人口構成にも影響を及ぼしうするため、平成 22 年国勢調査の地域属性を用いることによる層分類の誤りは、手法(1)および(2)で対応できる範囲を超えて発生することが懸念される。手法(3)は、全国の避難者数、小中学校の児童・生徒数とその変化、転出入届など関連する 2 次データを利用して、国勢調査の人口データを補正しようというものである。利用可能なデータに応じ、単純な比例配分から空間的相互作用モデルなどの空間モデルの適用まで、具体的な手法は様々に工夫できる。

本研究では上記の手法(1)および(2)に着目し、その有効性を検証する。具体的には、総務省統計局から提供を受けた平成 20 年住宅・土地統計調査調査票データを用いて、平成 25 年住宅調査を統計的にシミュレーションする。そして、これらの手法を適用した場合と従来通りの方法で調査を行った場合とを比較することにより、提案された手法が調査結果の精度向上に有効であるかを分析する。なお、手法(3)については、利用可能なデータおよびその精度が自治体や調査時期により異なり汎用的な議論が難しいこと、実際の調査作業を受け持つ地方自治体においては複雑な手順は実用性に欠けることから、本研究では取り扱わないこととした。

## 2. 住宅・土地統計調査の概要

ここでは、住宅・土地統計調査の概要を、特に本研究に関係する層分類と標本抽出方法を中心にまとめる。参考としたのは、総務省統計局ホームページの公開情報 (<http://www.stat.go.jp/data/jyutaku/2008/index.htm>; 同/[suikai.htm](http://www.stat.go.jp/data/suikai.htm))、ならびに同局統計調査部国勢統計課の後述するデータ提供の担当者とのコミュニケーションである。

前述のように住宅調査では、直前の国勢調査の調査区を第1次抽出単位とし、抽出された標本単位区内の住戸を第2次抽出単位として、標本となる住戸を抽出する。標本単位区は原則として抽出された国勢調査調査区であるが、住戸数が70戸以上の場合には、一つの単位区の住戸数が概ね50戸となるように分割し、その中の1つが無作為に抽出される。

第1次抽出の前に、国勢調査調査区は国勢調査の結果に基づき、幾つかの層に分類される。層化の基準となる指標および層の数には変遷があるが、現在では、主に「換算世帯数」、「住宅の所有の関係別割合(持ち家、民営借家、など)」、および「65歳以上親族のいる一般世帯数割合」を用いた24層となっている(表1)。住宅調査は国勢調査の3年後に行われるため、層化の基準となる指標が変化している可能性もあるが、通常その変化は小さいと仮定し、特別な補正等は行われ<sup>1</sup>ない。

続く第1次抽出では、層化された調査区は、人口規模に応じて市区町村ごとまたは都道府県ごとに、各層に定められた抽出率に従って抽出される(表2)。同一の市区町村または都道府県内では、層による抽出率の変化は2段階で、主として世帯数の少ない層の抽出率が、それ以外の層の抽出率の1/2に抑えられた構造である。

第2次抽出では、抽出された調査区およびそれを分割した区を標本単位区とし、住宅および住宅以外で人が居住する建物、ならびにこれらに居住している世帯を、1単位区あたり17住戸(世帯)、全国で合計約350万戸抽出する。調査区内の住戸数の確認は標本単位区の抽出後に行われるため、17住戸の抽出ができない単位区も発生するが、結果の推計の際に適宜乗率を用いて補正される。ただし、住戸数が0であった場合は「無住戸調査区」として、結果の推計からは除外される<sup>2</sup>。このとき代替となる調査単位区の再抽出は行わないため、無住戸調査区が増えると、有効な調査区数が設計より減少してしまう。平成25年の住宅調査で被災地域を考慮せずに調査区の抽出を行うと、特に津波被害の深刻な地域ではこうした状況が起こりやすく、注意が必要である。

---

<sup>1</sup> 総務省統計局のデータ提供の担当者とのコミュニケーションによる。

<sup>2</sup> 同上。

表 1: 平成 20 年住宅・土地統計調査における国勢調査調査区の層別基準および層符号  
 (総務省統計局・政策統括官・統計研修所ホームページから転載;  
[http://www.stat.go.jp/data/iyutaku/2008/pdf/sui\\_2.pdf](http://www.stat.go.jp/data/iyutaku/2008/pdf/sui_2.pdf))

層別基準			調査区の層符号		
平成 17 年国勢調査調査区	換算世帯数が 16 以上の調査区	後置番号 8 の調査区又は 50 人以上の準世帯のある調査区	010		
		世帯数が 0 の調査区又は換算世帯数が 15 世帯以下の調査区で、 65 歳以上親族のいる一般世帯割合が	30% 未満	021	
			30% 以上	022	
		間借り等の世帯数が 5 % 以上の調査区で、 65 歳以上親族のいる一般世帯割合が	30% 未満	031	
			30% 以上	032	
		3 階建て以上の世帯数が 90 % 以上の調査区 住宅の世帯数が 90 % 以上の調査区 その他の調査区	給与住宅に居住の世帯数が 50% 以上の調査区	110	
			公的借家に居住の世帯数が 50% 以上の調査区で、 65 歳以上親族のいる一般世帯割合が	30% 未満	121
				30% 以上	122
			民間借家に居住の世帯数が 50% 以上の調査区	130	
			持ち家に居住の世帯数が 50% 以上の調査区	140	
			その他の調査区	150	
			給与住宅に居住の世帯数が 50% 以上の調査区	210	
			公的借家に居住の世帯数が 50% 以上の調査区で、 65 歳以上親族のいる一般世帯割合が	30% 未満	311
		30% 以上		312	
		民間借家に居住の世帯数が 50% 以上 65% 未満の調査区で、 65 歳以上親族のいる一般世帯割合が	30% 未満	411	
			30% 以上	412	
		民間借家に居住の世帯数が 65% 以上の調査区で、 65 歳以上親族のいる一般世帯割合が	30% 未満	511	
			30% 以上	512	
		持ち家に居住の世帯数が 50% 以上 80% 未満の調査区で、 65 歳以上親族のいる一般世帯割合が	30% 未満	611	
			30% 以上	612	
持ち家に居住の世帯数が 80% 以上の調査区で、 65 歳以上親族のいる一般世帯割合が	30% 未満	711			
	30% 以上	712			
その他の調査区で、 65 歳以上親族のいる一般世帯割合が	30% 未満	801			
	30% 以上	802			

表 2: 平成 20 年住宅・土地統計調査における指定調査区の抽出率

(総務省統計局・政策統括官・統計研修所ホームページから転載;

[http://www.stat.go.jp/data/iyutaku/2008/pdf/sui\\_3.pdf](http://www.stat.go.jp/data/iyutaku/2008/pdf/sui_3.pdf))

地 域		調査区の抽出率	
		層符号010, 021, 022以外の調査区	層符号010, 021, 022の調査区
市 部	人口60万以上の市区	1/10	1/20
	人口50万以上60万未満の市区	1/9	1/18
	人口40万以上50万未満の市区	1/8	1/16
	人口30万以上40万未満の市区	1/7	1/14
	人口20万以上30万未満の市区	1/6	1/12
	人口10万以上20万未満の市区	1/5	1/10
	人口5万以上10万未満の市区	1/3	1/6
	人口5万未満の市区	1/2	1/4
郡 部	人口1万5千以上の町村	1/4	1/8
	人口1万5千未満の町村	1/8	1/16

このようにして抽出された調査対象住戸に関する調査には、世帯に配布する調査表によるものと、調査員が行う建物調査表によるものがある。世帯に配布する調査表には、甲と乙の2種類があり調査単位区ごとに約6:1の比率で配布される。乙は、以下に述べる甲の調査事項に、「現住居以外の住宅および土地に関する事項」を加えたものである。本研究では共通の調査項目のみを扱うこととして、両者を区別しない。

甲調査表の調査事項(すなわち全ての調査単位区に共通の調査事項)は、「住宅等に関する事項」、「住宅に関する事項」、「世帯に関する事項」、「家計を主に支える世帯員又は世帯主に関する事項」、「住環境に関する事項」からなる。住宅調査における「住宅」は「完全に区画された建物の一部で、一つの世帯が独立して家庭生活を営むことができるように建築又は

改造されたもの」(総務省統計局 <http://www.stat.go.jp/data/jyutaku/2008/1-1.htm#01>)と定義され、調査では、会社や学校の寮・寄宿舎など住宅以外で人が居住する建物も対象とするため、「住宅等に関する事項」と「住宅に関する事項」に分かれている。前者は、居住室数や広さ、所有関係(持ち家、借家、など)といった基礎的事項を、後者は、構造(木造、鉄筋・鉄骨コンクリート造、など)、階数、建て方(一戸建、長屋建、など)、種類(専用住宅、店舗その他の併用住宅)、家賃または間代、建築時期、設備、増改築・改修工事といった、より住宅に特化した事項を、それぞれ扱っている。

「世帯に関する事項」は、世帯の種類(主世帯、同居世帯、など)、構成、年間収入を、「家計を主に支える世帯員又は世帯主に関する事項」はその者の従業上の地位や通勤時間、現住居への入居時期、前住居に関する事柄をカバーしている。「住環境に関する事項」は、敷地に接する道路の幅員など個々の住宅に関する事項に加え、調査区の建ぺい率や容積率、最寄りの都市施設までの距離など、調査区に関する事項も含む。また、居住世帯のない住宅の場合には、調査員が外観で判断し、一部の調査項目のみが調査される。

住宅調査の結果の表章地域には、全国、都道府県、市区町村、および大都市圏・都市圏・距離帯がある。なお、町村の結果については、人口1万5千人以上の町村のみが表章の対象となる。調査は標本調査であるため、調査票から得られた結果は表章地域ごとに推計されて公開される。推計は、調査単位区別の調査結果に、調査単位区内の調査対象住戸の抽出率、国勢調査調査区を分割して調査単位区とした場合にはその分割数、および標本単位区の抽出率を考慮した乗率を適用して行われる。ただし前述のように、住戸を含まない無住戸調査区はこの推計過程から除外される<sup>3</sup>。

本研究では、この従来の調査方法に、被災地域の状況を考慮した層分類と抽出率を導入することにより、被災地域の統計の信頼性を確保することができるか、シミュレーションを通じて検証する。

### 3. データおよび研究手法

#### 3.1. データ

本研究の基幹となるデータは、総務省統計局から提供を受けた平成20年住宅調査の調査票および調査単位区に関するデータである。調査票データは、調査対象住戸に関する配布調査表および調査員調査表の記入結果(ただし個人情報を除く)を示したもので、各レコー

---

<sup>3</sup> 総務省統計局のデータ提供の担当者とのコミュニケーションによる。

ド(調査対象住戸)には属する調査単位区の ID 番号も付されており、対応する平成 17 年国勢調査の調査区と関連づけることができる。調査単位区に関するデータは、この国勢調査調査区との関連づけを定義するものである。本研究では、青森、岩手、宮城、福島 の 4 県のデータを用いる。これらのデータは、総務省統計局に研究内容を申請し利用許可を受けたもので、一般には非公開である。そのため本報告書においては、レコード数、調査単位区の位置等、詳細を記述できない場合があり、概数やパーセンテージでの表示となっていることに注意を願いたい。また、平成 17 年国勢調査に関しては、GIS 形式の調査区境界線データを総務省統計局より、GIS 形式の町字境界線データと町字レベル・調査区レベルの結果データを東京大学空間情報科学研究センターより、それぞれ提供を受けた。

震災の被害は多岐に及ぶものであるが、本研究では、住宅調査に強く関連するものとして、住宅の存続に直接的かつ広範な影響を及ぼす、津波による浸水被害に着目をする。浸水区域と調査単位区との関係を把握するため、国土交通省国土地理院が提供する 2 万 5 千分 1 浸水範囲概況図(国土交通省国土地理院 <http://www.gsi.go.jp/kikaku/kikaku40014.html>)を参考にデジタイズして、浸水区域境界線の GIS データを作成した。

## 3.2. 研究手法

### 3.2.1. 概要

本研究では、第 1 節で示した提案手法(1)・(2)の有効性を、平成 25 年住宅調査を仮想的に実施するシミュレーションを通じて検証する。具体的には、平成 20 年住宅調査で抽出された調査単位区を仮想的な母集団として扱い、その一部が被災したと仮定して、調査単位区の抽出および調査項目の推計を行う。そして、従来通りの層分類・調査区抽出手法を用いた場合と、提案手法(1)・(2)を取り入れた場合の推計結果を、仮想的母集団から得られる「(仮想的な)真の値」と比較する。住宅調査は標本調査であり、国土全体に対する実際の真の値を求める、全ての国勢調査調査区または調査単位区内の全ての住居からの標本抽出過程を再現するといったことは不可能であるため、既存の調査単位区および調査対象住戸に関するデータを母集団とした限定的なシミュレーションという形式を取っている。よって、より正確に表現するならば、平成 17 年国勢調査データに基づいて行われた平成 20 年住宅調査を一部再現し、平成 17 年から 20 年の間に震災が起こっていた場合に、住宅調査の結果に生じ得た影響をシミュレーションしていることになる。

第 1 次標本抽出では、平成 20 年調査で抽出された調査単位区を母集団として、その中から層別に標本となる調査単位区を抽出する。第 2 次標本抽出も同様の手順とすることも可能



ではあるが、平成 20 年調査で調査対象住戸となった住戸は各調査単位区で最大で 17 戸であり、この中から更に標本抽出をすることはシミュレーションの精度という点でも、実際の住宅調査の再現という点でも好ましくない。そこで、第 2 次標本抽出に関しては、平成 20 年の対象住戸が今回のシミュレーションでも抽出されたものとして扱う。被災地域の状況を反映した新しい層としては、津波による浸水地域内で、住宅の全壊割合が 80%以上の地域と 80%未満の地域を設定する。既存の層に関しては、従来通りの分類法が適用されたとして、平成 20 年調査時の実際の層分類を利用する。

シミュレーションの適用地域は、宮城県の浸水による住宅被害が確認されている沿岸の市町で、平成 20 年住宅調査で調査単位区として抽出された約 1600 の国勢調査調査区からなる。ただし、平成 20 年住宅調査で無住戸調査区とされたもの等は除外する。

また、シミュレーションに先立ち、住宅調査の層の間で実際の浸水被害に偏りがあるかどうかを分析しておく。上述のように、住宅調査は層別に調査区を抽出する層化抽出法を用いるため、層ごとに浸水被害が異なれば、推計結果に各層が占める影響にも設計からの乖離が生じて、推計値の信頼性に関する問題が大きくなる。一方、各層が均等に浸水被害を受けているのであれば、統計結果への相対的影響は保たれ、従来より精度は劣るものの問題は比較的抑えられる。ここでは、カイ二乗検定を適用して、浸水被害の層ごとの相違・偏りを調べることとする。

### 3.2.2. シミュレーションの流れ

最初に、3.1 節で述べた平成 17 年国勢調査の調査区境界線データと浸水区域の境界線データを、GIS を用いて重ね合わせ、各調査区が浸水被害を受けているかを判断する。浸水区域の境界線は必ずしも調査区のそれと一致するものではないが、一部でも浸水区域と重なる調査区は被害があったものとする。また、国勢調査調査区が分割されて調査単位区とされている場合でも、分割後の境界線を示すデータは公開されていないため、分割前の調査区をそのまま用いる。町字レベルの境界線データについても同様に、町字ごとの浸水被害の有無を調べる。

次に、宮城県のうち、浸水区域を含む町村を抽出し、それらに属する平成 20 年住宅調査の調査単位区を選出してシミュレーションにおける調査区母集団とする。更に、浸水被害状況に応じて、調査区母集団内の単位区を、浸水被害大（浸水による住宅の全壊率 80%以上）、小（同 80%未満）、無（浸水被害無し）の 3 つに仮想的に分類する（表 3）。続けて、この単位区ごとの浸水被害状況に応じ、単位区内の調査対象住戸の被害状況を定める。ここでは表 3 に

示す想定(平均)全壊率を用いて、各戸に無作為に全壊であるか否かを割り当てることとする。  
表3第6列に、結果として得られた調査単位区内住戸の平均全壊率を示す。

表 3: シミュレーションにおける調査区の被害区分と住宅の全壊率

被害区分	定義	調査区数		想定 全壊率	無作為割当に よる平均全壊 率
		(概数)	(割合)		
大	浸水による住宅の 全壊率が80%以上	80	(5%)	90%	89.1%
小	浸水による住宅の 全壊率が80%未満	410	(26%)	20%	20.2%
無	浸水被害無し	1110	(69%)	0%	0%
合計		1600	(100%)	-	-

このようにして仮想的に作成した浸水による住宅被害データを用いて、平成20年住宅調査の過程をシミュレーションにより再現する。層分類は実際の平成20年調査と同様の基準で、平成17年国勢調査の結果に基づいて行われたものとし、調査区の抽出率は、層010、021、022以外の調査区で1/3、層010、021、022の調査区で1/6とする。これは、人口5万人以上10万人未満の市区に対して適用される値である(表2参照)。シミュレーションでは、以下の3種類のシナリオを扱う。

シナリオ 1: 浸水による住宅被害を考慮せず、通常通りの層分類・調査区抽出方法を用いる。

シナリオ 2: 住宅の全壊率が80%以上の調査単位区は、ほぼ住宅がないものと仮定し、層010、021、022と同様の抽出率とする。これは提案手法(1)を適用したものであり、全壊率80%以上の地域を被災地域に特有の層として新たに設定して、抽出を行う場合に相当する。

シナリオ 3: シナリオ2の対応に加え、全壊率が80%未満の調査区は復興時の変化が激しく、調査区間のばらつきが大きくなると想定して、抽出率を1/3から1/2に増加させる。(ただし、層010、021、022の調査区は除

く。) これは提案手法(1)と(2)を組み合わせたものであり、全壊率 80%以上の地域および 80%未満の地域をそれぞれ新しい層として、後者に高い抽出率を適用して、抽出を行う場合に相当する。

これら 3 種類のシナリオに基づき、調査単位区とその中の調査対象住戸を無作為抽出し、調査項目の推計を行う。更に、調査区母集団に属する全ての調査対象住戸データを用いて、仮想的な「真の値」を算出し、推計値の精度を調べる。住宅調査には約 100 の調査集計項目があるが、本研究では、「建築の時期」項目について「建築時期が昭和 55 年以前の住宅の割合」、「世帯人員」項目について「世帯人員 1 名(単身者世帯)の割合」、「世帯の年間収入階級」項目について「年収 200 万円未満の世帯の割合」の 3 変数を対象とする。

## 4. 結果

### 4.1. 浸水被害の層間分布に関する検討

3.2.1 節で述べたように、シミュレーションの前に、浸水被害の分布に住宅調査の層間で偏りがあるか否かを、カイ二乗検定により調べた。表 4 に、被災 4 県の浸水被害の状況を、平成 17 年国勢調査調査区と、それらのうち平成 20 年住宅調査の調査単位区として用いられたものについてまとめた。調査単位区の実数は非公開情報であるので、概数およびパーセンテージで示してある。表 2 にあるように、住宅調査の調査区抽出率は市郡の人口規模により異なるため、県単位での抽出率には 20%から 25%という幅ができています。青森県および福島県では、国勢調査調査区と住宅調査調査単位区の浸水率に変わりはないが、岩手県および宮城県では後者の浸水率の方が高くなっており、これらの県では浸水被害による住宅調査の調査結果への影響がより大きくなる可能性がある。

表 4: 被災 4 県の国勢調査調査区および住宅調査調査単位区の浸水被害の状況

	青森県	岩手県	宮城県	福島県	合計
平成 17 年国勢調査の調査区数（概数）	10,200	10,500	16,700	14,600	52,000
平成 20 年住宅調査における調査区抽出率	22%	25%	23%	20%	22%
国勢調査調査区の浸水率	2%	9%	14%	4%	8%
住宅調査調査単位区の浸水率	2%	14%	16%	4%	10%

国勢調査調査区の層分類は、平成 20 年住宅調査で調査単位区となったもの以外については、入手したデータからは明らかでない。国勢調査の結果を基に層分類の再現を試みたが、調査区レベルで公開になっている調査項目には制限があり、再現には不十分であった。そこで、ここでは調査区の代わりに町字レベルの地区単位を用いることとして、国勢調査結果データに基づき層分類を作成した。ただし、層 010 の判定に用いる「準世帯数」は国勢調査の集計項目ではなく、実際の住宅調査の際には国勢調査の個票（調査票）データに遡って算出されるものである<sup>4</sup>ため、層 101 は除外し、23 層のみを対象とした。

表 5 は、県ごとに町字レベルの地区数を、浸水被害の有無について集計したものである。比較を容易にするため、括弧内にそれぞれの総単位地区数（列ごとの総和）に対する各層の割合を示した。層によっては、浸水被害のある場合とない場合で 10%以上の差が生じており（表 5 の青色で示した部分；層 021、612、712）、カイ二乗検定の結果、4 県全てで、浸水被害の分布には層間で統計的に有意な相違（有意水準 0.001）があることが明らかとなった。また、差の現れる層やその大小は県の間で一定ではないことも確認された。例えば、層 021 の場合、岩手県では浸水被害のない地区に偏在しているが、福島県ではその逆である。いくつかの市を選び同様の調査をしたが、結果は同じく市ごとの偏在の仕方の相違を示した。よって、この層間の浸水被害の差を共通のモデル等で把握することは難しく、調査時点での対策が必要となってくる。

<sup>4</sup> 総務省統計局のデータ提供の担当者とのコミュニケーションによる。

表 5: 被災 4 県の町字レベルの層分類と浸水被害の関係

(数値は町字レベルの地区数、括弧内の数値は浸水被害の有無で集計した  
 総和に対する割合；青色は割合に 10%以上の差のある層)

層分類	青森県			岩手県			宮城県			福島県		
	浸水被害		合計	浸水被害		合計	浸水被害		合計	浸水被害		合計
	無	有		無	有		無	有		無	有	
021	784 (19%)	21 (16%)	805 (19%)	422 (13%)	12 (3%)	434 (12%)	313 (8%)	128 (11%)	441 (9%)	951 (15%)	126 (31%)	1077 (16%)
022	361 (9%)	10 (7%)	371 (9%)	383 (11%)	11 (3%)	394 (10%)	173 (5%)	73 (6%)	246 (5%)	662 (11%)	48 (12%)	710 (11%)
031	9 (0%)	0 (0%)	9 (0%)	12 (0%)	0 (0%)	12 (0%)	2 (0%)	2 (0%)	4 (0%)	8 (0%)	0 (0%)	8 (0%)
032	84 (2%)	4 (3%)	88 (2%)	56 (2%)	13 (3%)	69 (2%)	38 (1%)	14 (1%)	52 (1%)	72 (1%)	1 (0%)	73 (1%)
110	0 (0%)	1 (1%)	1 (0%)	1 (0%)	0 (0%)	1 (0%)	4 (0%)	3 (0%)	7 (0%)	2 (0%)	0 (0%)	2 (0%)
121	5 (0%)	0 (0%)	5 (0%)	5 (0%)	1 (0%)	6 (0%)	10 (0%)	2 (0%)	12 (0%)	17 (0%)	0 (0%)	17 (0%)
122	- (0%)	- (0%)	- (0%)	- (0%)	- (0%)	- (0%)	4 (0%)	1 (0%)	5 (0%)	2 (0%)	0 (0%)	2 (0%)
130	1 (0%)	0 (0%)	1 (0%)	2 (0%)	0 (0%)	2 (0%)	11 (0%)	0 (0%)	11 (0%)	1 (0%)	0 (0%)	1 (0%)
140	1 (0%)	0 (0%)	1 (0%)	- (0%)	- (0%)	- (0%)	12 (0%)	0 (0%)	12 (0%)	- (0%)	- (0%)	- (0%)
150	1 (0%)	0 (0%)	1 (0%)	1 (0%)	0 (0%)	1 (0%)	5 (0%)	0 (0%)	5 (0%)	- (0%)	- (0%)	- (0%)
210	16 (0%)	0 (0%)	16 (0%)	8 (0%)	1 (0%)	9 (0%)	8 (0%)	3 (0%)	11 (0%)	3 (0%)	0 (0%)	3 (0%)

表 5: 被災 4 県の町字レベルの層分類と浸水被害の関係(続き)

層分類	青森県			岩手県			宮城県			福島県		
	浸水被害		合計	浸水被害		合計	浸水被害		合計	浸水被害		合計
	無	有		無	有		無	有		無	有	
311	27 (1%)	0 (0%)	27 (1%)	20 (1%)	0 (0%)	20 (1%)	23 (1%)	2 (0%)	25 (1%)	36 (1%)	0 (0%)	36 (1%)
312	18 (0%)	0 (0%)	18 (0%)	15 (0%)	1 (0%)	16 (0%)	21 (1%)	3 (0%)	24 (0%)	40 (1%)	2 (0%)	42 (1%)
411	90 (2%)	1 (1%)	91 (2%)	78 (2%)	4 (1%)	82 (2%)	222 (6%)	24 (2%)	246 (5%)	169 (3%)	3 (1%)	172 (3%)
412	43 (1%)	1 (1%)	44 (1%)	27 (1%)	6 (1%)	33 (1%)	47 (1%)	16 (1%)	63 (1%)	37 (1%)	2 (0%)	39 (1%)
511	41 (1%)	1 (1%)	42 (1%)	33 (1%)	0 (0%)	33 (1%)	120 (3%)	15 (1%)	135 (3%)	74 (1%)	0 (0%)	74 (1%)
512	1 (0%)	0 (0%)	1 (0%)	2 (0%)	0 (0%)	2 (0%)	2 (0%)	4 (0%)	6 (0%)	2 (0%)	0 (0%)	2 (0%)
611	188 (5%)	3 (2%)	191 (5%)	104 (3%)	2 (0%)	106 (3%)	208 (5%)	56 (5%)	264 (5%)	197 (3%)	3 (1%)	200 (3%)
612	666 (17%)	20 (15%)	686 (16%)	484 (14%)	117 (28%)	601 (16%)	580 (15%)	257 (22%)	837 (17%)	860 (14%)	62 (15%)	922 (14%)
711	105 (3%)	4 (3%)	109 (3%)	65 (2%)	1 (0%)	66 (2%)	199 (5%)	15 (1%)	214 (4%)	135 (2%)	0 (0%)	135 (2%)
712	1445 (36%)	67 (50%)	1512 (36%)	1490 (44%)	229 (56%)	1719 (46%)	1622 (43%)	503 (44%)	2125 (43%)	2743 (44%)	155 (38%)	2898 (44%)
801	80 (2%)	2 (1%)	82 (2%)	90 (3%)	7 (2%)	97 (3%)	119 (3%)	14 (1%)	133 (3%)	101 (2%)	5 (1%)	106 (2%)
802	70 (2%)	0 (0%)	70 (2%)	61 (2%)	7 (2%)	68 (2%)	63 (2%)	16 (1%)	79 (2%)	94 (2%)	5 (1%)	99 (1%)
合計	4036 (100%)	135 (100%)	4171 (100%)	3359 (100%)	412 (100%)	3771 (100%)	3806 (100%)	1151 (100%)	4957 (100%)	6206 (100%)	412 (100%)	6618 (100%)

更に、シミュレーションの対象地域が属する宮城県について、平成 20 年住宅調査で抽出された調査単位区の層分類と浸水被害の有無の関連性を調べた。表 6 は、調査区数が上位 10 位までに入る層(調査区数で 50 以上)についての集計である。町字レベルと同様にカイニ乗検定を行ったところ、ここでも有意水準 0.001 以下で明らかな関連性が示された。

以上より、浸水による住宅への被害が層分類の信頼性に及ぼす影響は、層ごとに異なることが明らかとなった。従って、浸水地域か否かの区別無く通常通りの層分類と調査区抽出を行った場合には、推計結果を利用する際大きな問題が生じる可能性が高い。

表 6: 宮城県の平成 20 年住宅調査対象調査区における層分類と浸水被害の関係  
(数値は浸水被害の有無で集計した総和に対する割合)

層符号	浸水地域		合計
	無	有	
032	2%	2%	2%
121	2%	1%	2%
140	4%	1%	4%
411	6%	4%	6%
412	3%	3%	3%
511	10%	5%	9%
611	6%	4%	5%
612	17%	24%	18%
711	7%	2%	6%
712	32%	44%	34%
合計	100%	100%	100%

#### 4.2. シミュレーション結果

まず、調査区母集団に属する全ての調査対象住戸データを用いて算出した、仮想的な「真の値」を表 7 に示す。3.2.2 節で述べた 3 種類のシナリオに基づくシミュレーションで得られる調査項目の推計値は、これらの値と比較して評価される。各シナリオは 10 回ずつ試行され、シナリオが定める抽出率により決まる抽出調査区数は、シナリオ 1、2、3 でそれぞれ概ね 540、530、600 である。結果を表 8 から 10 に示す。

10回の試行で得られた推計値の平均値と、「真の値」との絶対誤差をみると、3調査項目の平均ではシナリオ2、1、3の順に大きくなるが、いずれのシナリオであっても誤差は0.5%以下に収まっている。調査項目ごとに「真の値」が異なり、同レベルの絶対誤差であっても意味・影響が違ふことを考慮して、絶対誤差の「真の値」に対する比率としてみると、シナリオ2、3、1の順に大きくなるが、いずれのシナリオでもほぼ2%以下である。従って、誤差の大きさという点ではシナリオ2が最も有利ではあるものの、誤差ならびにそのシナリオ間の差は低いレベルに留まっており、通常通りの調査手法(シナリオ1)を用いた場合であっても、住宅の浸水被害による統計値の信頼性への影響は比較的小さいと言える。しかしながら、このシミュレーションで使用した仮想的な調査区母集団では、全壊率80%以上とされる調査区は全体の5%程度であり、表章地域は市のレベルを想定しているが、浸水被害を受けた調査区を多く含む小規模の表章地域を扱う場合には、より大きな影響が現れるものと推測され、注意が必要である。

一方で10回の試行間の推計値のばらつき(標準偏差)をみると、シナリオ2が最も大きくなっている。これは、シナリオ2では住宅被害の大きい地域を無住戸調査区に近い状況であると仮定して抽出率を下げ、結果的に抽出調査区数が低く抑えられているためである。シナリオ3では、住宅被害の大きい地域はシナリオ2と同様に扱うが、被害の小さい地域は復興過程での変化が激しく、国勢調査の結果に基づく層分類の信頼度が下がることを考慮して抽出率を上げているため、全体として抽出調査区数が増加し、サンプリングによる結果のばらつきは小さくなっている。シナリオ1の抽出調査区数は両者の間に位置し、結果のばらつきも同様である。

被災地域が復興していく過程を正確に捉えることを目的とする場合、理論的にはシナリオ3が望ましい。しかしながら、本研究のシミュレーションによれば、「真の値」からの誤差を抑えるためにはシナリオ2が、推計値のばらつきの少なさ、言い換えれば、安定性を確保するためにはシナリオ3が有利であるという結果となった。住宅調査の実際の作業は、地方自治体が主体となって行うものであり、コストや制度変更による現場の混乱も配慮すべき重要な要素である。よって、結果の安定性では多少劣るものの、誤差および調査の負荷の両面で優れたシナリオ2が、この場合は適切と考えられる。



表 7: 宮城県内の調査区母集団における調査項目の仮想的な「真の値」

調査対象項目	仮想的な「真の値」
建築時期が昭和 55 年以前の住宅の割合	29.75%
世帯人員 1 名の割合	24.47%
年収 200 万円未満の世帯の割合	16.43%

表 8: シナリオ 1 の結果（通常通りの調査手法）

	建築時期が昭和 55 年以前の住宅 の割合	世帯人員 1 名の 割合	年収 200 万円未満 の世帯の割合
Trial 1	29.64%	23.79%	16.41%
Trial 2	28.65%	23.77%	15.86%
Trial 3	29.94%	24.66%	15.97%
Trial 4	29.76%	24.66%	16.49%
Trial 5	30.26%	25.36%	16.59%
Trial 6	29.82%	23.94%	16.39%
Trial 7	29.43%	22.91%	15.73%
Trial 8	29.65%	23.70%	16.10%
Trial 9	30.80%	23.80%	15.96%
Trial 10	29.43%	23.08%	15.70%
平均値	29.74%	23.97%	16.12%
標準偏差	0.53%	0.71%	0.31%
レンジ (最大値-最小値)	2.15%	2.45%	0.90%
平均値と「真の値」の 絶対誤差	0.01%	0.50%	0.31%
絶対誤差の「真の値」 に対する比率	0.04%	2.05%	1.86%

表 9: シナリオ 2 の結果  
 (全壊率 80%以上の地域を新たな層として設定する方法)

	建築時期が昭和 55 年以前の住宅 の割合	世帯人員 1 名の 割合	年収 200 万円未満の 世帯の割合
Trial 1	28.81%	25.31%	16.98%
Trial 2	29.71%	23.32%	16.50%
Trial 3	29.27%	25.08%	16.11%
Trial 4	30.84%	25.60%	17.89%
Trial 5	28.47%	23.40%	16.51%
Trial 6	30.68%	23.76%	17.59%
Trial 7	29.63%	24.50%	16.18%
Trial 8	30.98%	22.78%	15.59%
Trial 9	27.29%	24.93%	16.43%
Trial 10	30.29%	24.19%	15.40%
平均値	29.59%	24.29%	16.52%
標準偏差	1.11%	0.90%	0.75%
レンジ (最大値-最小値)	3.69%	2.82%	2.49%
平均値と「真の値」 の絶対誤差	0.16%	0.18%	0.09%
絶対誤差の「真の 値」に対する比率	0.52%	0.74%	0.57%

表 10: シナリオ 3 の結果

(全壊率 80%以上および 80%未満の地域をそれぞれ新しい層として設定し、  
更に後者の抽出率を上げる方法)

	建築時期が昭和 55 年以前の住宅 の割合	世帯人員 1 名の 割合	年収 200 万円未満 の世帯の割合
Trial 1	30.45%	25.82%	16.53%
Trial 2	29.89%	24.69%	16.51%
Trial 3	29.45%	25.98%	17.68%
Trial 4	28.86%	24.37%	16.37%
Trial 5	30.50%	24.37%	16.14%
Trial 6	28.35%	23.82%	15.68%
Trial 7	29.30%	24.58%	16.52%
Trial 8	30.18%	25.85%	16.52%
Trial 9	28.44%	25.43%	16.93%
Trial 10	28.64%	24.26%	16.16%
平均値	29.41%	24.92%	16.50%
標準偏差	0.78%	0.74%	0.50%
レンジ (最大値-最小値)	2.15%	2.16%	2.00%
平均値と「真の値」の 絶対誤差	0.34%	0.45%	0.08%
絶対誤差の「真の値」 に対する比率	1.15%	1.83%	0.47%

## 5. まとめ

本研究では、平成 25 年実施予定の住宅調査に際し、東日本大震災による被災、特に住宅の被害が統計結果に影響を及ぼす可能性を考慮して、統計の信頼性を確保することのできる調査方法を提案・検証することを目的に、過去の国勢調査・住宅調査データを用いて仮想的に次回の調査をシミュレーションした。調査方法としては、住宅が甚大な被害を受け、次回住宅調査の時点ではほぼ住宅がないと予想される地域を新たな層とするもの(シナリオ 2)、更に、住宅の被害が比較的少なく復興過程での急激な変化が予想される地域も新たな層と

して、抽出率を上げてより詳細な把握を目指すもの(シナリオ3)の2種類を提案した。シミュレーションは、宮城県の沿岸部のうち、住宅被害のあった浸水地域を含む市町を対象地域として扱い、そこに含まれる平成20年住宅調査における調査単位区および調査対象住戸を母集団と仮定して、従来通りの調査手法(シナリオ1)を含む3種類のシナリオに従って、調査単位区を無作為抽出するという方法で行った。これによりシナリオごとの推計値の精度の差は比較的小さいことが示され、調査の負荷を考慮すると、シナリオ2が最も適切という結果となった。

本研究の制約としては、シミュレーションの回数が少なく結果のばらつきが大きいこと、対象地域が被災地域の一部に限られていること、新しく層を設定する基準が住宅の浸水被害に限られ、仮設住宅の影響等を考慮できていないことなどが挙げられる。次回の住宅調査には時間的に間に合わないが、本研究で扱った問題は他の標本調査でも起こりうるものであり、今後こうした点を更に追求していく予定である。

## 6. 参考文献

浅見泰司・山田育穂、2011、「東日本大震災直後の住宅・土地統計調査の調査方法案」(総務省統計局統計調査部国勢調査課への提案)

国土交通省国土地理院、「平成23年(2011年)東日本大震災に関する情報提供」、国土交通省国土地理院ホームページ、[http://www.gsi.go.jp/BOUSAI/h23\\_tohoku.html](http://www.gsi.go.jp/BOUSAI/h23_tohoku.html) および <http://www.gsi.go.jp/kikaku/kikaku40014.html> (accessed 08/04/2011)

総務省統計局、「平成20年住宅・土地統計調査 調査の概要」、総務省統計局・政策統括官・統計研修所ホームページ、<http://www.stat.go.jp/data/jyutaku/2008/index.htm> (accessed 07/14/2011)

総務省統計局、「平成20年住宅・土地統計調査 調査の結果・推計方法」、総務省統計局・政策統括官・統計研修所ホームページ、<http://www.stat.go.jp/data/jyutaku/2008/suikai.htm> (accessed 09/08/2011)

総務省統計局、「平成20年住宅・土地統計調査 標本抽出方法及び結果の推定方法」、総務省統計局・政策統括官・統計研修所ホームページ、[http://www.stat.go.jp/data/jyutaku/2008/pdf/sui\\_1.pdf](http://www.stat.go.jp/data/jyutaku/2008/pdf/sui_1.pdf) (accessed 07/19/2011)

総務省統計局、「平成20年住宅・土地統計調査 用語の解説」、総務省統計局・政策統括官・統計研修所ホームページ、<http://www.stat.go.jp/data/jyutaku/2008/1.htm> (accessed 08/12/2011)

## 7. 謝辞

本研究は、財団法人不動産流通経営協会の研究助成を受けて行われており、ここに心からの感謝を表明する。平成20年住宅・土地統計調査調査表データ、国勢調査調査区GISデータなど、本研究の基幹となるデータは総務省統計局より提供を受けた。その利用および研究の実施に際しては、同局の平澤学氏、近藤利一氏、井岡貴司氏、大山伊知郎氏に貴重なコメントをいただいた。また、東京大学空間情報科学研究センターからは平成22年国勢調査データ等の提供を受けた。同センターの浅見泰司教授には、研究の計画段階から貴重なご助言を得ており、そのご協力なしには本研究は成し得なかった。ご協力いただいた皆様に、この場を借りて厚くお礼申し上げたい。ただし、本稿に含まれ得る誤りは全て筆者に帰するものである。