

2019年7月

## 建築着工統計調査の補正調査に関する 標本設計移行期における推定方法の提案

---

澤谷 秀之\*

増成 敬三\*\*

上田 聖\*\*

\* 元総務省統計委員会担当室（平成31年3月31日まで在籍）

\*\* 総務省統計委員会担当室

---

総務省 統計委員会担当室

〒162-8668 東京都新宿区若松町19番1号

総務省統計委員会担当室ワーキングペーパーは、統計委員会担当室スタッフ又はスタッフと外部研究者との共同による調査・研究の成果をまとめたもので、公的統計の整備に係る各種施策に役立てることを企図としている。ただし、ワーキングペーパーの内容や意見は、執筆者個人に属し、総務省の公式見解を示すものではない。

---

# 建築着工統計調査の補正調査に関する 標本設計移行期における推定方法の提案

## 要旨

総務省が実施している統計精度向上の取組の対象としている調査の一つとして、国土交通省が実施している建築着工統計調査の補正調査がある。この補正調査は、建築着工前の工事費予定額と竣工後に確定した工事実施額について、そのかい離率（補正率）を把握するための標本調査である。

その調査結果の推定精度向上に向けて、総務省は新しい標本設計を提案している。本稿は標本設計を変更した際に、旧来の方法で提案された標本と新しい方法で抽出された標本とが混在するために複雑な計算が必要となる移行期において、偏りなく推定を行う二つの方法を示す。一つは、偏りなく復元できる厳密な推定を行う計算方法であり、もう一つは簡易に推定を行う計算方法である。なお、ここで提案する簡易な方法は経験上の方法であるため、国土交通省より過去に実施された建築着工統計調査の補正調査のデータ提供を受けて、リサンプリングシミュレーションを実施し、概ね偏りが無いことを示している。

キーワード：新規着工統計調査、補正調査、公的統計、ネイマン配分、  
標本設計、リサンプリングシミュレーション

本稿は、平成 28 年度統計法施行状況報告審議結果報告書（統計精度検査編）における提言への対応に資するため、総務省統計委員会担当室において国土交通省から調査票情報の提供を受けて行ったリサンプリングシミュレーションの結果をとりまとめたものである。本稿の作成にあたっては、櫻川幸恵氏（総務省）から有益な助言を頂いた。ただし、本稿の内容と意見は筆者ら個人に属し、所属組織の公式見解を示すものではない。また、ありうべき誤りはすべて筆者ら個人に属する。

## 1 はじめに

統計改革に資する統計の精度向上の取組として平成 29 年度上半期に総務省では三つの検査を実施し、「平成 28 年度 統計法施行状況報告 統計精度検査編」として平成 29 年 11 月に統計委員会に報告した。この三つの検査のうちの一つが、国土交通省の実施している建築着工統計調査の補正調査の標本設計の検査である。

当該検査結果の報告を受け統計委員会で審議が行われ、平成 30 年 3 月に審議結果報告書がとりまとめられた。その報告書において補正調査については、令和 3 年（2021 年）1 月に完成する建築物から見直し後の標本設計に基づき抽出された対象に切り替えて調査を開始することが必要であり、「この切り替えの際、ア）見直し後の標本設計に基づき抽出された調査対象と、イ）現行の抽出方法に基づき抽出され都道府県から国土交通省に移管された調査対象、の両方を混在させて調査を行うことが必要となることから、その移行期の結果推定には統計技術的な対応が必要である。そのため、総務省は移行時期の結果推定の方法の開発について国土交通省に協力し、円滑な調査の移行を支援することが必要である」と報告されたのである。

本稿は、新旧の標本が混在する移行期に実施する補正調査について、厳密な復元性を重視した推計方法と、簡易な推計方法の二つを提案する。この簡易な方法は、精度を担保しているか確認が必要である。そのため、国土交通省より建築着工統計調査の調査票情報の提供を受け、リサンプリングによるシミュレーション実験を実施したので、その結果を示す。

本稿の構成は次のとおりである。2 節においては補正調査の概要と標本設計見直しにあたっての課題を示し、3 節では、調査方法の切り替え時期における新旧標本の混在した状態における厳密な推計方法と、簡易な推計方法及び簡易な方法におけるリサンプリングシミュレーションの結果を示す。4 節では、これらの結果をまとめ結びにかえる。

## 2 建築着工統計調査の補正調査について

### 2.1 本体調査と補正調査の概要

建築着工統計調査は、

- ア) 全国の建築物（着工建築物）の着工状況（建築物の数、床面積合計、工事費予定額等）を延べ床面積 10 m<sup>2</sup>超の建築物に届出が義務付けられている「建築工事届」から集計する本体調査（全数調査）
- イ) 本体調査の建築物のうち、住宅の着工状況（戸数及び床面積合計）を構

造、建て方、利用関係、資金等に分類して集計する住宅着工統計（全数調査）

ウ) 本体調査の建築物のうち、建築物の竣工時に実際にかかった費用（工事実施額）を実地に調査し、着工時における工事費予定額とのかい離を明らかにする補正調査（標本調査）から構成されている。

このうち補正調査の標本抽出は、①調査実施市区を抽出し、②更に調査実施市区内の建築物に関する「建築工事届」を調査実施市区ごとに定められた抽出率で抽出する方法（比例配分法）が採用されている。

建築物着工前の「建築工事届」には、予定床面積と工事費予定額が含まれていることから、面積あたりの予定単価を求めることができる。補正調査とは、建築物が竣工した後に実施床面積と工事実施額を標本調査して実施単価を求めて、予定単価からのかい離率（補正率＝実施単価／予定単価）を求める調査である。

## 2.2 新旧方式の抽出作業と移行期間

実際の補正調査の標本抽出は、旧方式においては都道府県の該当部署において建築工事届を受け付けた時に、あらかじめ設定された抽出率に従って、例えば10分の1に定められた場合には建築工事届を10枚に1枚を抜き出して複写して保存し、複写後に用紙は全て倉庫に保管するようになっている。

一方、新方式では、国土交通省に届いた建築工事届から、対象となる建築物の書類番号を抽出しておき、その番号を都道府県に通知して、都道府県に保管されている建築工事届の用紙を複写して送付してもらうという手順が想定されている。

補正調査の都道府県から国土交通省への移管は2021年1月からである。よって2021年1月の時点において、新旧の方式による標本が揃っていないなければならない。標本抽出の新方式への移行期は明確に定められていないが統計精度の向上に資するため、2020年1月以降、できるだけ早期に新方式の導入が望まれる。

## 2.3 補正調査の見直しにあたっての課題

建築着工統計調査の補正調査は、精度向上に向け標本設計が見直されることになっている。その方法の概要は以下ア) からカ) のとおりであり、この見直しを実施することで同じ標本数で標本誤差を 1 / 5 に縮小できることが見込まれている。

ア) 木造・非木造別に工事費予定額階級 3 区分 (20 億円以上、1 ～20 億円、1 億円未満) の 6 層に標本を区分する。

イ) 20 億円以上の 2 層は全数を抽出する。

ウ) 残りの 4 層はネイマン配分により標本を配分する。

エ) 合計の標本サイズを 5000 件とする。

オ) 2021 年 1 月に完成する建築物から見直し後の標本設計に基づき抽出された対象に切り替えて調査を開始する (調査切り替え)。

カ) 可能な限り早期に新方式での抽出を開始する (新抽出方法の導入)。

このような対応を行うことにより、新標本抽出への移行期を仮に 2020 年 1 月とすれば、2021 年中に完成する調査対象は、表 1 に示すように旧抽出方法で抽出された標本と見直し後の抽出方法で抽出された標本が混在した状態で調査が行われる。

表 1 提案された標本配分法による標本配分イメージ

(仮に抽出方法の切り替えを 2020 年 12 月とした場合のイメージ)

		2021 年(報告年)	
		旧方式(比例配分法)	新方式(ネイマン配分法)
着工年 (届出年)	2019 年	約 120 件	—
	2020 年	約 2200 件	約 2500 件
	2021 年		約 2200 件

一方、2021 年の調査切り替え後、しばらく経過した後完成する調査対象は、見直し後の抽出方法で抽出された標本のみで調査が行われることから、集計方法は単純化される。

このため、短期の移行期間のみのために特別な集計を行う必要があることから、効率的かつ偏りのない方法での集計が望まれる。

このように移行期における調査では 2 種類の標本抽出法による標本が混在することになり、複雑な復元作業が発生することが課題となっているので、厳密な推計式に加えて、より簡素化した方法も提案する必要があると考えた。

## 2.4 補正調査のデータ概要

国土交通省より貸与された建築着工統計調査・補正調査の調査票情報について、異常値を除外した上で報告年（完成年）と工事費予定額及び完成予定年の関係を見てみる。まず、工期が長くなるほど工事費予定額の高い建築物の割合が高くなり、工期と工事費予定額規模の相関関係が観察される（表2）。

表2 工期別に見た工事費予定額別の建築物数およびその構成費  
(2010年～2016年着工建築物合計)

木造・非木造合計	合計		0年後		1年後		2年後		3年以上	
	件数	構成比	件数	構成比	件数	構成比	件数	構成比	件数	構成比
1千万円未満	2744	8.7	1883	9.9	849	7.1	12	4.9	0	0.0
1千万円以上～2千万円未満	10359	33.0	6565	34.3	3747	31.2	41	16.8	6	37.5
2千万円以上～3千万円未満	10332	32.9	6383	33.4	3894	32.4	52	21.3	3	18.8
3千万円以上～5千万円未満	4796	15.3	2910	15.2	1860	15.5	26	10.7	0	0.0
5千万円以上～1億円未満	1716	5.5	893	4.7	793	6.6	30	12.3	0	0.0
1億円以上～5億円未満	1192	3.8	432	2.3	716	6.0	42	17.2	2	12.5
5億円以上～20億円未満	208	0.7	42	0.2	130	1.1	32	13.1	4	25.0
20億円以上	28	0.1	6	0.0	12	0.1	9	3.7	1	6.3
合計	31375	100.0	19114	100.0	12001	100.0	244	100.0	16	100.0

また、完成予定年と報告年の関係については、工事遅延等の原因によって概ね2割程度が完成予定年の翌年以降に遅れて完成することが判明した。

この遅れは毎年発生するので、ある年の完成が遅れても、前年から遅れて完成する件数が完成実績として計上されるので、結果として予定と実績の差は大きくなり、平均で4.6%程度の差であると確認した（表3）。

表3 完成予定年別・報告年別の建築物数（件）

		報告年							完成 予定 合計	
		2010年	2011年	2012年	2013年	2014年	2015年	2016年		2017年
完成 予定 年	2010年	3991	520	1	1	0	0	0	0	
	2011年	0	3571	669	31	1	0	0	0	4272
	2012年	0	0	3757	744	4	3	1	0	4509
	2013年	0	0	0	4589	542	9	2	1	5143
	2014年	0	0	0	0	4335	794	11	0	5140
	2015年	0	0	0	0	0	3872	691	1	4564
	2016年	0	0	0	0	0	0	4050	249	4299
	2017年	0	0	0	0	0	0	0	24	
完成実績合計			4091	4427	5365	4882	4678	4755		
予定と実績の差			4.4%	1.9%	-4.1%	5.3%	-2.4%	-9.6%		

### 3 移行期における2つの推計方法（厳密な方法と簡易な方法）

前節で指摘したように、補正調査の結果を厳密な推計法で復元する方法に加えて、簡易な方法について以下に示す。

#### 3.1 厳密な方法（標本抽出から厳密に場合分けを行い復元する方法）

建築着工統計調査の補正調査については、本体調査において「工事費予定額」及び「完成予定時期」が全数で調査されており、工期は1か月から36か月で完成したものが調査対象となっている。このため、例えば新方式の抽出開始時期を2020年1月とすれば、2021年調査では、2019年12月まで旧標本設計で抽出し、2020年1月以降は新標本設計で抽出を行い、これらのうち2021年に完成した建築物に対し調査が開始されることとなる。

このような標本構成を表示すると表4に示すとおりとなる。理論上はこの表のセル毎に全数集計を行い、該当建築物数を把握した上でセル毎の標本数に対応したウエイトを作成することが必要である。

表4 層別報告年（完成年）別の場合分け

工事予定額	2021年（報告年）	
	木造	非木造
1億円未満	2019年に旧方式抽出	2019年に旧方式抽出
	2020年に新方式抽出	2020年に新方式抽出
	2021年に新方式抽出	2021年に新方式抽出
1億から20億円	2019年に旧方式抽出	2019年に旧方式抽出
	2020年に新方式抽出	2020年に新方式抽出
	2021年に新方式抽出	2021年に新方式抽出
20億円以上	2019年に旧方式抽出	2019年に旧方式抽出
	2020年に新方式抽出	2020年に新方式抽出
	2021年に新方式抽出	2021年に新方式抽出

具体的には2021年調査の場合、

- ア) 届出年（2019年～2021年）の調査対象のうち完成予定年が2021年の調査対象のみを抽出し、その対象を構造（2区分）及び工事費予定額階級（3区分）、届出年区分を場合分けして全数集計を行い、それぞれの部分母集団の総数を計算する。
- イ) 構造（2区分）、工事費予定額階級（3区分）及び届出年（2019年～2021年）の区分毎に回収標本数を計算する。
- ウ) 上記ア)及び上記イ)で計算された部分母集団数及び部分標本数に基づき、以下の【推計式1】で総額等の推計を行う。

【推計式 1】

$$\sum_{i,j,l} \sum_m N(i,j,k=t+2,l) \frac{X(i,j,k=t+2,l,m)}{n(i,j,k=t+2,l)}$$

$$\equiv \sum_{i,j,l} \sum_m N'(i,j,k'=t+2,l) \frac{X(i,j,k=t+2,l,m)}{n(i,j,k=t+2,l)}$$

i : 構造 (木造、非木造)

j : 工事費予定額階級 (20 億円以上、1~20 億円、1 億円未満)

k : 完成年 (2021 年) ⇔ k' : 完成予定年 (2021 年)、t : 旧方式による抽出最終年、t+2 : 調査年

l : 届出年 (着工年) (2019 年、2020 年、2021 年)

m : 標本インデックス

N : 構造 i、工事費予定額階級 j、完成年 k、届出年 l の建築物の総数 → **【計測不能】**

N' : 構造 i、工事費予定額階級 j、完成予定年 k'、届出年 l の建築物の総数

n : 構造 i、工事費予定額階級 j、完成年 k、届出年 l の抽出標本数

X : 構造 i、工事費予定額階級 j、完成年 k、届出年 l、標本インデックス m の標本の観測値

### 3.2 簡易な方法 (抽出時の抽出率の逆数をウェイトにする方法)

3.1 の推計式で集計することは、理論上は正しいが、過渡的な期間のために複雑な場合分けを要する算式を用いて集計することは非効率的と考えられる。そのため、本来は完成時期の条件付確率の逆数をウェイトとして加重平均する算式とすべきところ完成時期の条件を無視し、単に抽出時の抽出率の逆数をウェイトとする算式を提案する**【推計式 2】**。この簡易な算式は、漸近的に不偏であることについて、以降のリサンプリングシミュレーションにより確認する。

【推計式 2】

$$\sum_{i,j,l} \sum_m N'(i,j,l) \frac{X(i,j,k=t+2,l,m)}{n(i,j,l)}$$

i : 構造 (木造、非木造)

j : 工事費予定額階級 (20 億円以上、1~20 億円、1 億円未満)

k : 完成年 (2021 年) ⇔ k' : 完成予定年 (2021 年)、t : 旧方式による抽出最終年、t+2 : 調査年

l : 届出年 (着工年) (2019 年、2020 年、2021 年)

m : 標本インデックス

N' : 構造 i、工事費予定額階級 j、届出年 l の建築物の総数

n : 構造 i、工事費予定額階級 j、届出年 l の建築物の標本数

X : 構造 i、工事費予定額階級 j、完成年 k、届出年 l、標本インデックス m の標本の観測値



### 3.3 簡易な方法の偏りを計測するリサンプリングシミュレーションの方法

建築着工統計調査の補正調査は、建築工事届の届出時に標本抽出され、完成時期に調査を行うものであるが、建築物の工期が長いものほど工事実施額が高額となる相関関係が見られる。そのため、提案した簡易な推計式のように抽出率の逆数を乗じるだけでは、推定結果に偏りが生じる可能性も否定できない。

このため、2010年から2016年までの間の建築着工統計調査の補正調査の調査票情報の提供を受け、次のようなシミュレーションを実施し、簡易な方法のバイアスを計測した。なお、提供を受けた調査票情報は従前の抽出法により抽出されていることから単純無作為抽出と見なせる方法で抽出されたものである。

#### ア) 2010年～2012年に届出された調査票情報でのシミュレーション

2019年12月まで旧抽出法で抽出し2020年1月から新抽出法で抽出する。調査は2021年から切り替える場合を仮置きしたシミュレーション。

- i) 2010年に届出調査票情報から、重複を許して単純無作為抽出（旧標本抽出法）で5000件をリサンプリングする。
- ii) 2011年、2012年に届出された調査票情報から、重複を許してネイマン配分で配分された標本数をそれぞれの層からリサンプリングする。
- iii) リサンプリングされた調査票情報から、2012年に完成するもののみを選定し、上記【推計式2】の推計式により計算される値から補正率を計算する。これを1000回繰り返す。
- iv) 2010年～2012年に届出された調査票情報のうち2012年に完成するもの全体で計算された工事実施額等の統計量を推定する（標本空間上の平均）。
- v) 上記iii)の試算は上記iv)を推定していることとなるため、iii)の結果がiv)の値を偏り無く推定しているか確認する。

イ) 上記ア)を2011～2013年、2012年～2014年、・・・、2014年～2016年、の調査票情報の組で繰り返し計算し、いずれの年でも偏りが無いことを確認する。

### 3.4 簡易な集計方法のシミュレーションの結果

簡易な推計方法を用いた推計によるリサンプリングシミュレーションの結果は表5のとおりであり、前述3.2の【推計式2】は概ね偏りなく推計できていると判断され、簡易な方法による集計においても、ほぼ偏りなく単価及び補正率は推計されるものと判断される。

表5 シミュレーションの結果  
単位面積当たり工事費予定額単価(万円/㎡)

木造			非木造		
リサンプリングデータ	標本空間の平均	試算結果1000件の平均(標準偏差)	リサンプリングデータ	標本空間の平均	試算結果1000件の平均(標準偏差)
2010-2012年	15.8	15.80(0.143)	2010-2012年	16.6	16.71(0.504)
2011-2013年	15.8	15.77(0.113)	2011-2013年	17.3	17.25(0.280)
2012-2014年	16.7	16.76(0.127)	2012-2014年	17.8	17.76(0.176)
2013-2015年	16.8	16.76(0.140)	2013-2015年	18.7	18.49(0.240)
2014-2016年	16.8	16.79(0.173)	2014-2016年	20.5	20.35(0.385)

単位面積当たり工事実施額単価(万円/㎡)

木造			非木造		
リサンプリングデータ	標本空間の平均	試算結果1000件の平均(標準偏差)	リサンプリングデータ	標本空間の平均	試算結果1000件の平均(標準偏差)
2010-2012年	16.1	16.02(0.154)	2010-2012年	17.7	17.86(0.565)
2011-2013年	16.2	16.11(0.133)	2011-2013年	17.5	17.50(0.271)
2012-2014年	17.2	17.25(0.144)	2012-2014年	18.5	18.63(0.245)
2013-2015年	17.3	17.27(0.150)	2013-2015年	19.5	19.29(0.257)
2014-2016年	17.2	17.19(0.184)	2014-2016年	21.5	21.12(0.369)

補正率

木造			非木造		
リサンプリングデータ	標本空間の平均	試算結果の1000件の平均(標準偏差)	リサンプリングデータ	標本空間の平均	試算結果の1000件の平均(標準偏差)
2010-2012年	1.01	1.014(0.004)	2010-2012年	1.07	1.069(0.009)
2011-2013年	1.02	1.022(0.002)	2011-2013年	1.02	1.014(0.007)
2012-2014年	1.03	1.029(0.003)	2012-2014年	1.04	1.049(0.015)
2013-2015年	1.03	1.031(0.004)	2013-2015年	1.04	1.043(0.005)
2014-2016年	1.03	1.023(0.003)	2014-2016年	1.05	1.037(0.003)

#### 4 まとめ

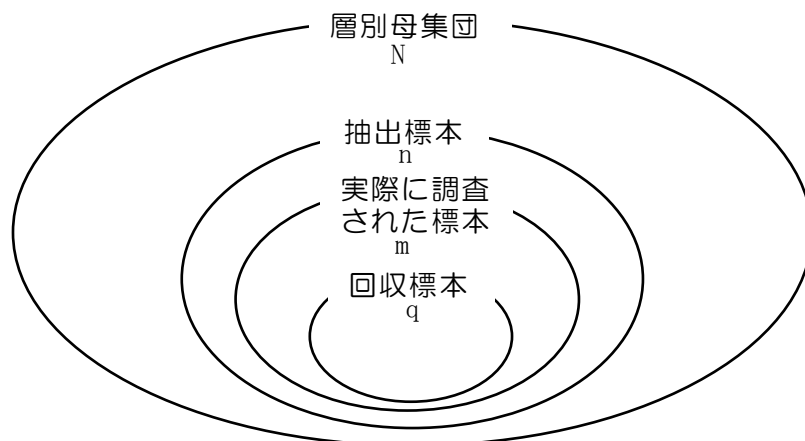
本稿においては、新旧の標本が混ざる移行期において、厳密な方法と簡易な方法を示した上で、簡易な方法であっても精度を保持していることをシミュレーションで示した。このため複雑な手順を要する理論上の推計方法のかわりに、簡易な方法で労力をかけず推計して問題はないものとする。このことは、仮に抽出法の切り替えが遅れた場合も成立するものと判断される。

ただし、シミュレーションにおいては 5000 件の標本を抽出し回収率 100% を想定して実施していることから、実際の集計にあたり、予備標本も含めた抽出が行われて、その一部に対し調査が実施されること、更に回収率も考慮することが必要であると考えられる。

具体的には、構造別工事費予定額階級の 6 層毎に抽出率の逆数に回収率の逆数を乗じること、更に予備標本を確保している場合には、予備標本に対する実際の調査を行った率の逆数を乗じることが必要となる(図 1 参照)。

抽出時の抽出率： $n/N$  簡易集計時のウェイト： $(n/N)^{-1} \times (m/n)^{-1} \times (p/m)^{-1}$

図 1 標本構成のイメージ



## <参考資料>

本稿では国土交通省から提供を受けた建築着工統計調査・補正調査の個票データを集計して利用した。その項目を下記に示す。提供されたデータは2010年1月～2016年12月の期間であり、全データ数は約3.3万件である。

### 国土交通省から提供を受けた調査票情報の掲載項目

項目名	符号	説明
着工予定年月	20YYMM	西暦の年月
竣工予定年月	20YYMM	西暦の年月
構造	1	木造
	2	鉄骨鉄筋コンクリート造
	3	鉄筋コンクリート造
	4	鉄骨造
	5	コンクリートブロック造
	6	その他
予定床面積(m <sup>2</sup> )	000011～999999	
工事予定額(万円)	0000001～9999999	
実施床面積(m <sup>2</sup> )	000011～999999	
工事実施額(万円)	0000001～9999999	
報告年月	20YYMM	西暦の年月

## <参考文献>

- ・ 総務省（2017）平成28年度統計法施行状況報告統計精度検査編  
平成29年11月21日総務省
- ・ 総務省（2018）平成28年度統計法施行状況に関する審議結果報告書  
（統計精度検査関連分）平成30年3月30日総務省統計委員会