資料２－１

擬似ミクロデータについて

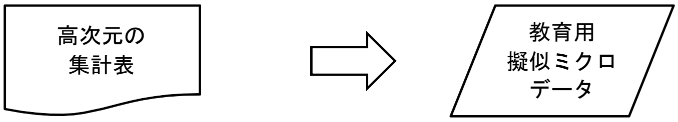
平成 24 年 12 月 19 日
独立行政法人　統計センター

１　背景

- 改正統計法の全面施行に伴い、オーダーメード集計、匿名データの提供開始（平成 21 年 4 月）。
- 一方、ミクロ統計データ活用研究会、日本学術会議報告など学術界からは、自由に使用できるミクロデータへの要望があった。
- 「統計データの二次利用促進に関する研究会」や統計委員会でも教育用・訓練用の簡易なデータに言及。
- （独）統計センターでは、全国消費実態調査のデータを用いて平成 21 年度から「擬似データ」の研究を開始。

２　擬似ミクロデータの作成

- 擬似ミクロデータとは、「調査票から作った集計表から、逆に本物の調査票と同じような分布構造を持った擬似的なミクロデータ」。



世帯人員	・・・	職業	度数	消費支出平均	・・・
2		1	10	250,000	・・・
2		2	20	280,000	・・・
:		:	:	:	:

世帯人員	・・・	職業	消費支出	・・・	
2		1	240,000		10レコード
2		1	:		
2		1	260,000		
2		2	300,000		20レコード
2		2	:		
2		2	260,000		
:		:	:		

- 集計表の各セルの中で、各変数の平均値、標準偏差と相関行列を計算。
- 対数正規分布を仮定して元データと同じ数の乱数を発生（見かけはミクロデータだが集計表と同じ情報しか持たない）。
- 小さなセルの取扱い、量的属性値が 0 のデータの取扱い 、集計用乗率の取扱い、量的属性値間の加法性の取扱いなどの課題を逐次検討。

３　擬似ミクロデータの試行提供（別紙 1）

- 大規模データ（二人以上の勤労者世帯約 32,000 世帯の 197 項目、csv 形式）と簡易データ（約 8,000 世帯の 25 項目、xls 形式及び csv 形式）。
- 利用要件は、
  - 再配布禁止。
  - 分析結果は実証研究の結果と見なすことはできないことに留意。
  - ６か月以内に利用者アンケートを提出。
- 現在までに合計約 70 件の利用申請があり、約 1000 人の学生が利用している。

資料２－１（別紙１）

業務案内
- 製表
- 公的統計基盤サービス
  - 政府統計共同利用システムの概要
  - 公的統計のミクロデータ利用
    - 匿名データの利用
    - オーダーメード集計の利用
    - 利用実績
    - イベント情報
    - オンサイト利用
    - ルクセンブルク所得研究のデータベース利用
    - 擬似ミクロデータの利用
    - 学術研究機関等との連携
    - 他府省における提供状況
- 刊行案内
- 技術研究
- 国際協力

業務案内

擬似ミクロデータの利用

統計センターでは、公的統計のミクロデータの利用を図るため、ミクロデータを用いた実証研究のプログラムテストや大学・高等学校等の授業で利用できるような擬似ミクロデータの試行提供を行っています。なお、擬似ミクロデータは、集計表から作成したミクロデータ形式の擬似的なデータセットであることを理解してご利用ください。
また、擬似ミクロデータをご利用の際は、アンケートにご協力いただきますようよろしくお願いいたします。

■擬似ミクロデータの作成方法

製表技術参考資料「平成24年7月「教育用擬似ミクロデータの開発とその利用　～平成16年全国消費実態調査を例として～」」をご覧ください。

■利用可能な擬似ミクロデータ

○全国消費実態調査

年次	調査の概要	データレイアウト及び符号表等	レコード数・項目数	擬似ミクロデータ
平成16年（二人以上の世帯・勤労世帯）	調査ページ（総務省統計局）	大規模データ (XLS)	約3万2千 197項目	CSV形式（ZIP（パスワード付）） 00001-05000 *1：約4MB 05001-10000　：約4MB 10001-15000　：約4MB 15001-20000　：約4MB 20001-25000　：約4MB 25001-30000　：約4MB 30001-32027　：約2MB
		簡易データ ・CSV形式(XLS) ・EXCEL形式(XLS)	約8千 25項目	CSV形式 *1 （ZIP（パスワード付））：約1MB EXCEL形式 *1,2 （ZIP（パスワード付））：約1MB

*1 ZIPファイルの中には、擬似ミクロデータのほか、利用の手引、アンケート等も収録されております。
*2 簡易データCSV形式を Microsoft Excel形式に変換して、変数名を付したものです。

このページのTOPへ

■利用要件

以下の事項に同意していただける方なら、どなたでもご利用いただけます。

1. 擬似ミクロデータを申出者及び利用者以外の者に利用させない。
2. 擬似ミクロデータは集計表から作成したミクロデータ形式の擬似的なデータセットであるため、これから導かれた分析結果は実証研究の結果と見なすことはできないことに留意する。
3. 申出日から6か月以内に擬似ミクロデータに同梱されている「擬似ミクロデータの利用者アンケート」を提出する。

実証研究としての分析を行う場合には、擬似ミクロデータの利用ではなく、「公的統計のミクロデータ利用」（匿名データの利用、オーダーメード集計の利用）などをご検討ください。

擬似ミクロデータのZIPファイルの解凍には、パスワードが必要です。利用を希望する方は、申出書（WORDファイル）を作成し、メールで統計センターに申出を行ってください。統計センターは、申出に対し、パスワードの発行を行います。

■利用に関する問い合わせ

擬似ミクロデータの利用に関するお問い合せは、下記連絡先までお問い合わせください。

統計センター製表部 統計作成支援課
MAIL：giji-jimu@nstac.go.jp

PDFファイルをご覧頂くためには「Adobe Reader」が必要です。

Get Adobe Reader

Copyright (c) 2008 National Statistics Center.All Rights reserved.

2012.11.9

擬似ミクロデータの試行提供

（独）統計センター　槙田 直木

平成24 年11 月16 日（金）　統計数理研究所
公的統計のミクロデータの利用に関する研究集会
（午前の部）「擬似ミクロデータの利用」

独立行政法人 統計センター NSTAC

構成

1．公的統計と擬似ミクロデータ

2．擬似ミクロデータの試行提供

3．擬似ミクロデータの特性

4．擬似ミクロデータの課題

独立行政法人 統計センター NSTAC

1．公的統計と擬似ミクロデータ

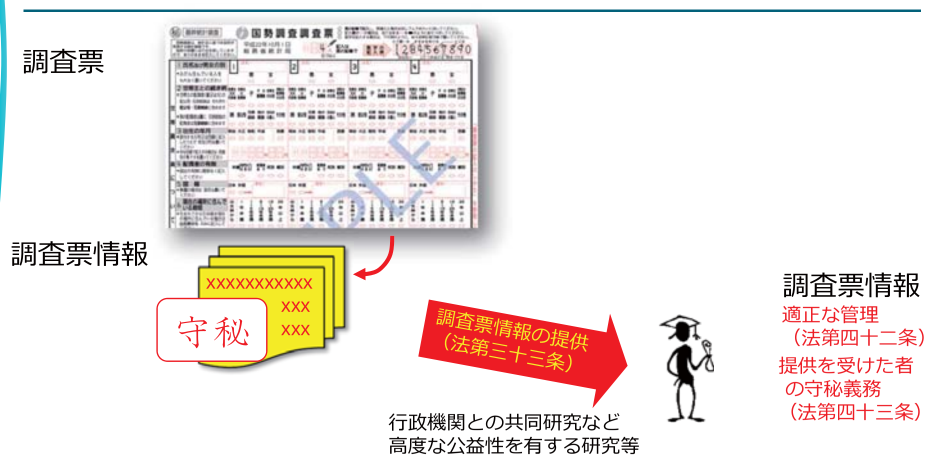
- 公的統計
  - 行政施策の一つとして作成される統計
  - 公的統計は、「公共財」
  - 国の予算による統計調査、その規模は大きい。
- 統計法
  - 公的統計や統計調査には、法令による権利･義務

独立行政法人 統計センター NSTAC

1．公的統計と擬似ミクロデータ

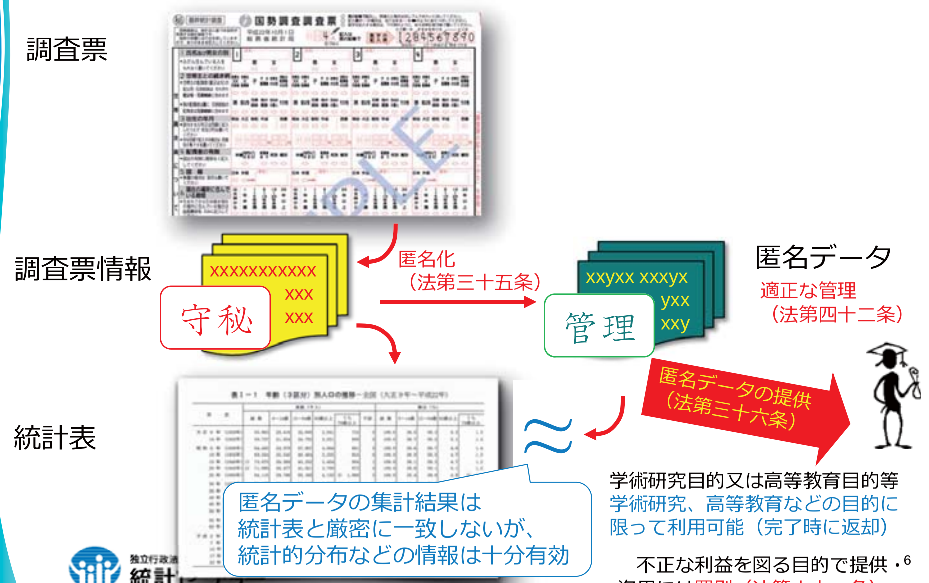
独立行政法人 統計センター NSTAC

1．公的統計と擬似ミクロデータ



秘密の漏えい 又は 不正な利益を図る目的で提供・盗用には
罰則（法第五十七条・第五十九条）

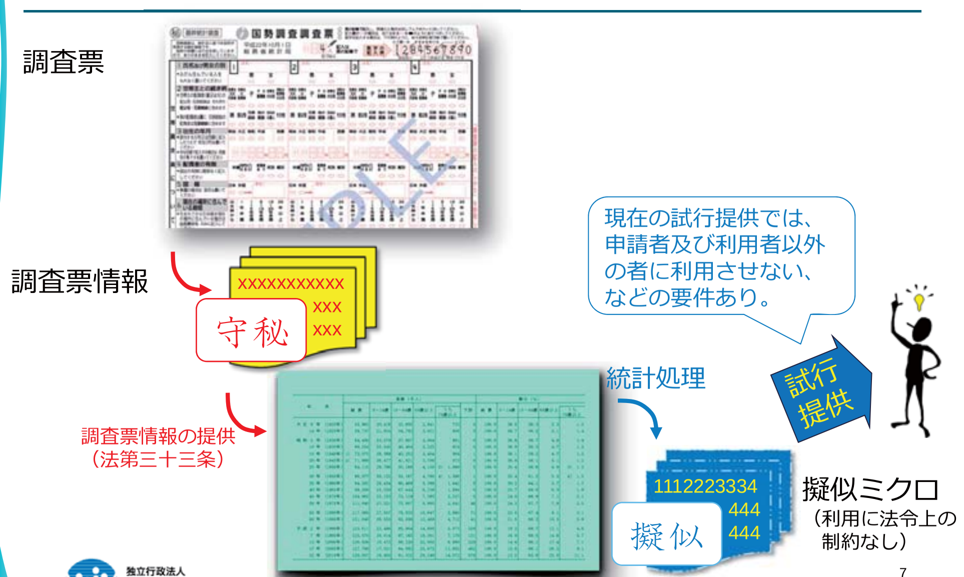
1．公的統計と擬似ミクロデータ



学術研究目的又は高等教育目的等
学術研究、高等教育などの目的に限って利用可能（完了時に返却）

不正な利益を図る目的で提供・盗用には罰則（法第六十一条）

1．公的統計と擬似ミクロデータ



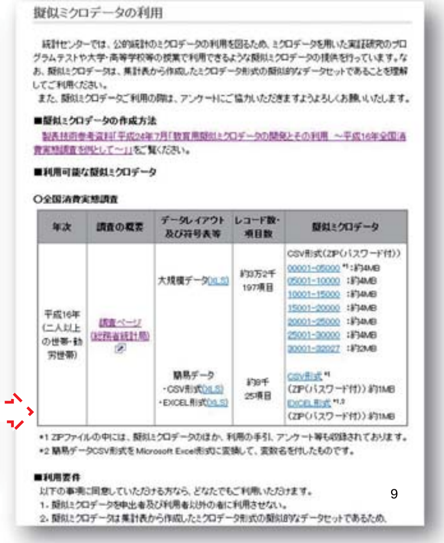
2．擬似ミクロデータの試行提供

- ミクロデータ利用者の裾野を広げ、公的統計の二次的利用の拡大を図るため、試行版の擬似ミクロデータを作成※し、試行提供（無償）。
  ※ 平成16年全国消費実態調査の特別集計結果に基づく
- 教育目的や、公的統計の二次的利用の際のテストデータなどに利用。
- 利用要件
  - 申請者及び利用者以外の者に利用させないこと
  - 集計表から擬似的に作成したデータであるので、分析結果は実証研究の結果と見なすことができないことを理解すること
  - 利用者アンケートの提出

2．擬似ミクロデータの試行提供

統計センターホームページ
http://www.nstac.go.jp

擬似ミクロデータの利用



擬似ミクロデータの利用

統計センターでは、公的統計のミクロデータの利用を図るため、ミクロデータを用いた実証研究のプログラムテストや大学・高等学校等の授業で利用できるような擬似ミクロデータの提供を行っています。なお、擬似ミクロデータは、集計表から作成したミクロデータ形式の擬似的なデータセットであることを理解してご利用ください。
また、擬似ミクロデータご利用の際は、アンケートにご協力いただきますようよろしくお願いいたします。

■擬似ミクロデータの作成方法
製表技術参考資料「平成24年7月「教育用擬似ミクロデータの開発とその利用 ～平成16年全国消費実態調査を例として～」」をご覧ください。

■利用可能な擬似ミクロデータ

○全国消費実態調査

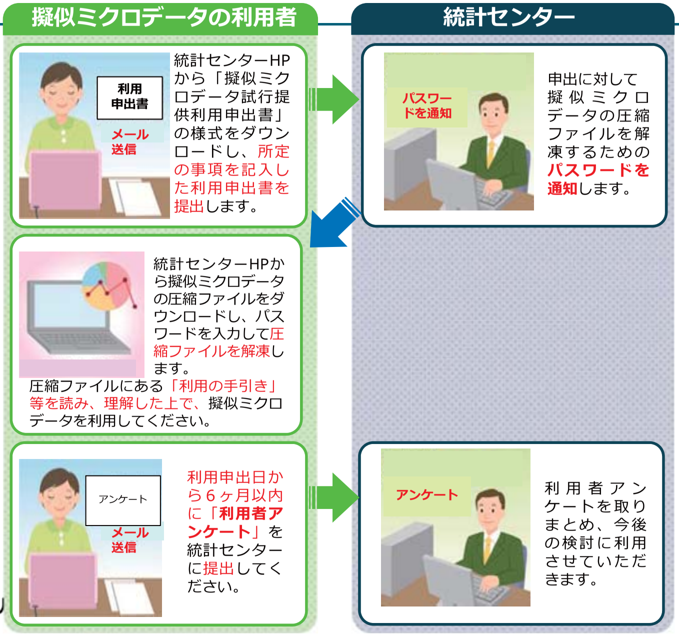
年次	調査の概要	データレイアウト及び符号表等	レコード数・項目数	擬似ミクロデータ
平成16年(二人以上の世帯・勤労世帯)	調査ページ(総務省統計局)	大規模データ(XLS)	約3万2千 197項目	CSV形式(ZIP(パスワード付)) 00001-05000 *1：約4MB 05001-10000 ：約4MB 10001-15000 ：約4MB 15001-20000 ：約4MB 20001-25000 ：約4MB 25001-30000 ：約4MB 30001-32027 ：約2MB
		簡易データ ・CSV形式(XLS) ・EXCEL形式(XLS)	約8千 25項目	CSV形式 *1 (ZIP(パスワード付)) 約1MB EXCEL形式 *1,2 (ZIP(パスワード付)) 約1MB

*1 ZIPファイルの中には、擬似ミクロデータのほか、利用の手引、アンケート等も収録されております。
*2 簡易データCSV形式をMicrosoft Excel形式に変換して、変数名を付したものです。

■利用要件
以下の事項に同意していただける方なら、どなたでもご利用いただけます。
1．擬似ミクロデータを申出者及び利用者以外の者に利用させない。
2．擬似ミクロデータは集計表から作成したミクロデータ形式の擬似的なデータセットであるため、

2．擬似ミクロデータの試行提供

手続きの流れ



2．擬似ミクロデータの試行提供

▫ 提供実績

	申出件数	利用目的				利用人数
		大学・高等学校等での授業	匿名データ提供申出前の演習	データ処理のためのプログラムテスト	その他	
2011(平成23)年度	45	45	-	1	-	363
2012(平成24)年度※	17	16	1	3	0	565

※10月末現在

2．擬似ミクロデータの試行提供

▫ 試行提供中の擬似ミクロデータ

- 平成16年全国消費実態調査 に基づくもの。

▫ 平成16年全国消費実態調査

- 世帯を対象に、家計の収入・支出及び貯蓄・負債、耐久消費財、住宅・宅地などの家計資産を総合的に調査。総務省統計局が5年ごとに実施。
- 標本の大きさ
  - 二人以上の世帯 約5万5千、単身世帯 約4千
- 調査票
  - 世帯票、家計簿、耐久財等調査票、年収・貯蓄等調査票

2．擬似ミクロデータの試行提供

- 試行提供中の擬似ミクロデータ
  平成16年全国消費実態調査 に基づくもの。

	レコード数	収録項目	ファイル形式
大規模データ（CSV形式）	約3万2千 二人以上の勤労者世帯	197項目	CSV形式 （ZIP（パスワード付））： 約４MB×6 +約2ＭＢ×1 （→解凍CSV 約90MB）
簡易データ（CSV形式 Excel形式）	約　8千 世帯人員が4名で有業人員が1～2人の世帯のみ	25項目	CSV形式 （ZIP（パスワード付））： 約1MB（→解凍CSV 約1MB） Excel形式 （ZIP（パスワード付））： 約1MB（→解凍XLS 約2MB）

ZIPファイルの中には、擬似ミクロデータのほか、利用の手引、アンケート等も収録

参考： 平16全消匿名データ 二人以上の世帯 レコード数 約4.4万、収録項目1611、サイズ971MB

2．擬似ミクロデータの試行提供

- 試行提供中の擬似ミクロデータ（続き）

	レコード数	収録項目の内訳		
		世帯属性等	支出項目	収入項目
大規模データ（CSV形式）	約3万2千 二人以上の勤労者世帯	14項目 世帯主の年齢、住居の種類など	149項目 用途分類	34項目 年間収入等
簡易データ（CSV形式 Excel形式）	約8千 世帯人員が4名で有業人員が1～2人の世帯のみ		11項目 消費支出及び十大費目のみ	なし

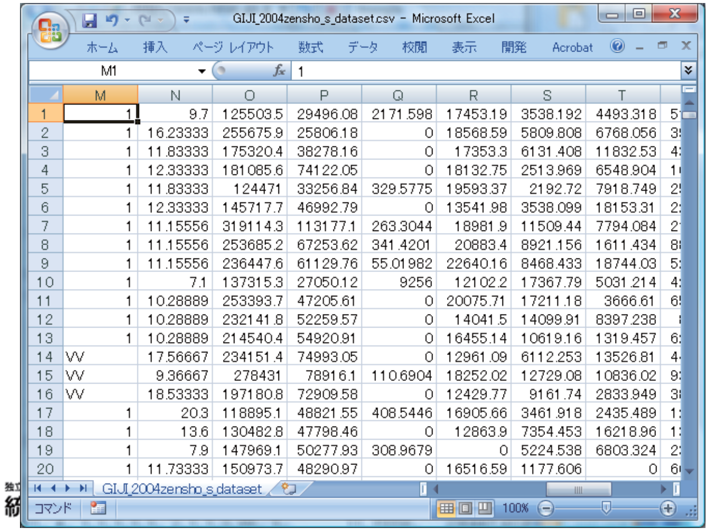
独立行政法人 統計センター NSTAC

2．擬似ミクロデータの試行提供

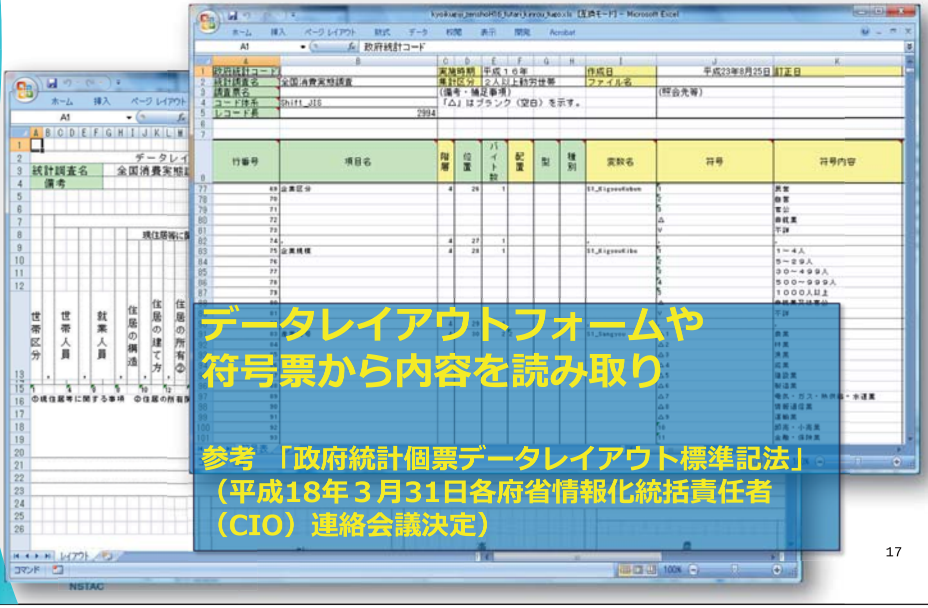
- CSV形式の内容（テキストエディタ）

2．擬似ミクロデータの試行提供

- CSV形式の内容（Excel）



2．擬似ミクロデータの試行提供

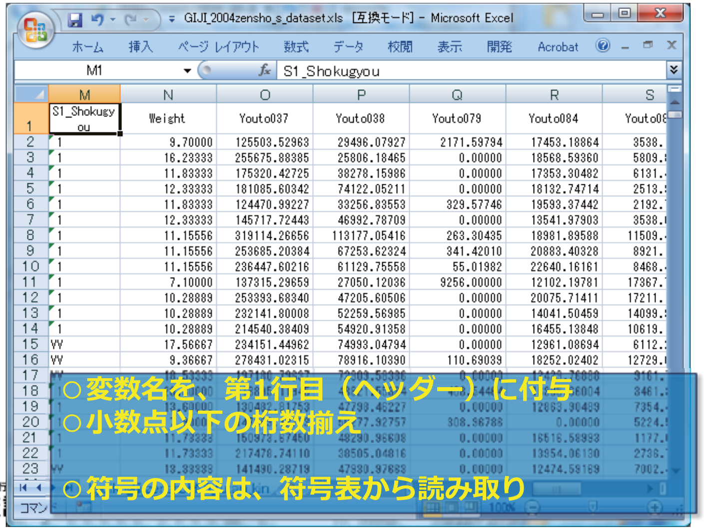


データレイアウトフォームや
符号票から内容を読み取り

参考 「政府統計個票データレイアウト標準記法」
（平成18年３月31日各府省情報化統括責任者
（CIO）連絡会議決定）

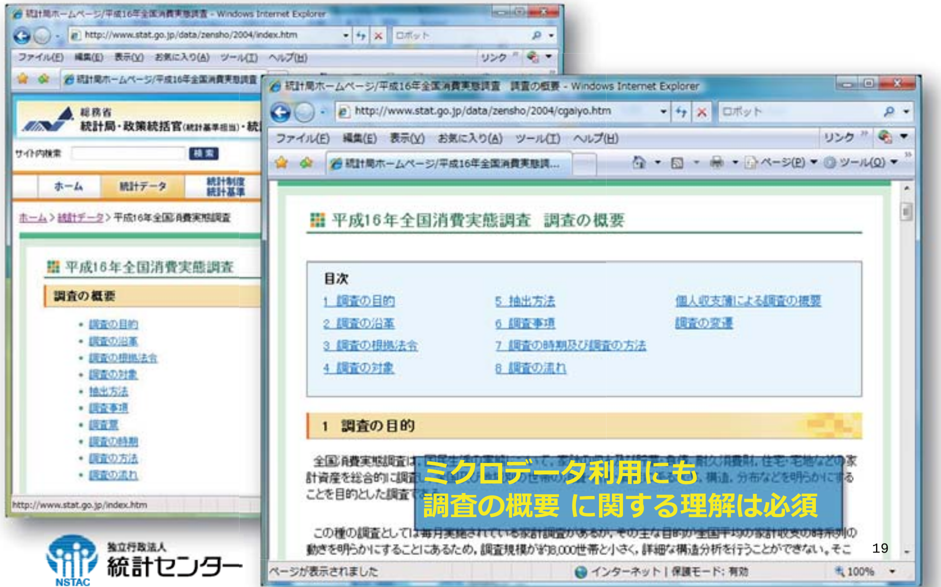
2．擬似ミクロデータの試行提供

- Excel形式の内容（Excel）



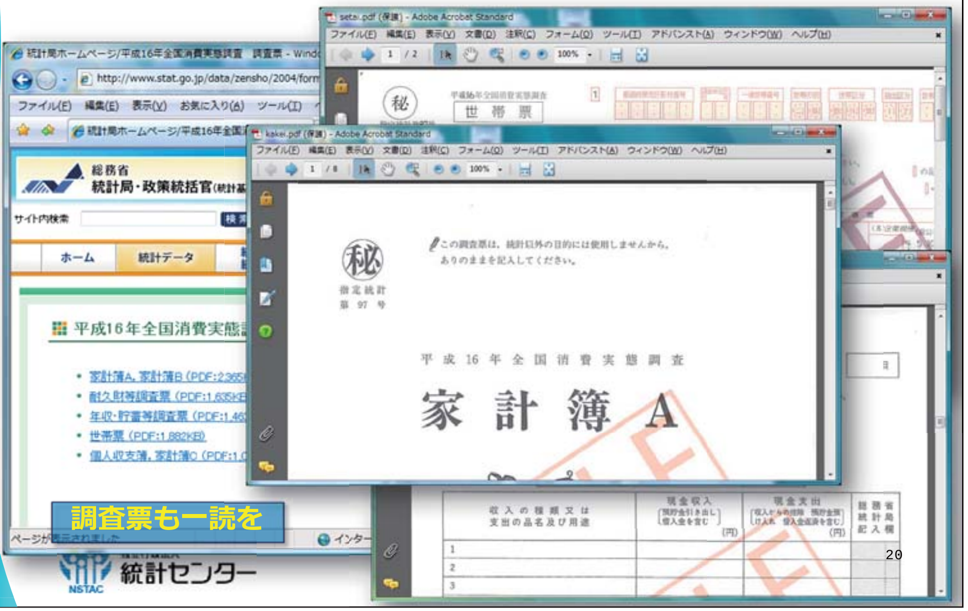
○変数名を、第1行目（ヘッダー）に付与
○小数点以下の桁数揃え
○符号の内容は、符号表から読み取り

2．擬似ミクロデータの試行提供



ミクロデータ利用にも
調査の概要 に関する理解は必須

2．擬似ミクロデータの試行提供



調査票も一読を

3．擬似ミクロデータの特性

基本統計量（平均、標準偏差）

	平均 調査票情報	平均 擬似ミクロ	平均 差率	標準偏差 調査票情報	標準偏差 擬似ミクロ	標準偏差 差率
年間収入	740. 18	729. 81	-0. 01	358. 18	337. 69	-0. 06
収入総額	971, 789. 24	946, 779. 03	-0. 03	541, 290. 74	473, 480. 73	-0. 13
実収入	502, 133. 73	497, 655. 92	-0. 01	280, 695. 92	261, 558. 27	-0. 07
実収入以外の収入	391, 823. 98	372, 130. 47	-0. 05	353, 922. 37	263, 445. 65	-0. 26
繰入金	77, 831. 53	76, 992. 65	-0. 01	87, 036. 21	98, 947. 04	0. 14
支出総額	971, 789. 24	946, 779. 03	-0. 03	541, 290. 74	473, 480. 73	-0. 13
実支出	415, 809. 39	403, 746. 63	-0. 03	224, 419. 69	219, 290. 60	-0. 02
消費支出	339, 199. 37	328, 139. 70	-0. 03	194, 501. 15	192, 447. 21	-0. 01
食料	73, 738. 54	72, 883. 42	-0. 01	30, 149. 02	28, 064. 49	-0. 07
住居	19, 387. 99	17, 687. 21	-0. 09	52, 962. 36	60, 587. 32	0. 14
光熱・水道	19, 395. 36	19, 237. 81	-0. 01	8, 009. 23	7, 690. 12	-0. 04
家具・家事用品	9, 783. 81	9, 204. 04	-0. 06	15, 977. 65	14, 933. 13	-0. 07
被服及び履物	14, 649. 44	14, 137. 63	-0. 03	18, 837. 04	19, 823. 09	0. 05
保健医療	11, 936. 01	11, 366. 36	-0. 05	19, 763. 39	19, 284. 07	-0. 02
交通・通信	50, 740. 68	47, 960. 92	-0. 05	85, 021. 69	84, 654. 38	0. 00
教育	22, 332. 15	22, 269. 65	0. 00	51, 989. 72	64, 157. 45	0. 23
教養娯楽	32, 472. 95	31, 389. 49	-0. 03	32, 161. 60	32, 723. 04	0. 02
その他の消費支出	84, 762. 44	82, 003. 18	-0. 03	95, 898. 83	102, 040. 97	0. 06
非消費支出	76, 610. 02	75, 606. 93	-0. 01	56, 199. 75	66, 378. 49	0. 18
実支出以外の支出	475, 947. 80	464, 318. 09	-0. 02	394, 805. 29	334, 227. 09	-0. 15
繰越金	80, 032. 04	78, 714. 31	-0. 02	96, 421. 45	118, 055. 82	0. 22

差率：（擬似ミクロ－調査票情報）／調査票情報
-0. 5より小、0. 5より大
-0. 5以上-0. 1より小、0. 1より大0. 5以下

単位　年間収入：万円
年間収入以外：円

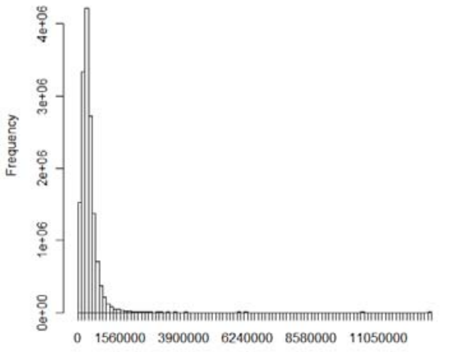
3．擬似ミクロデータの特性

ヒストグラム 標準偏差の差率が最大となる「実収入以外の収入」に関する分布の比較

実収入以外の収入

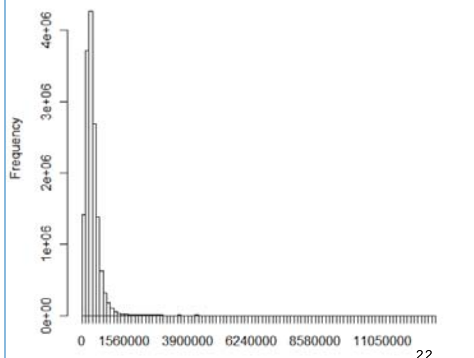
調査票情報

平均 391,823.98
標準偏差 353,922.37



擬似ミクロデータ

平均 372,130.47
標準偏差 263,455.65



3．擬似ミクロデータの特性

量的属性間の相関係数 調査票情報

赤字　符号が変化
差：教育用擬似ミクロデータ－個票データ
-0. 2より小、0. 2より大
-0. 2以上-0. 1より小、0. 1より大0. 2以下

調査票情報	年間収入	収入総額	実収入	実収入以外の収入	繰入金	支出総額	実支出	消費支出	食料	住居	光熱・水道	家具・家事用品	被服及び履物	保健医療	交通・通信	教育	教養娯楽	その他の消費支出	非消費支出	実支出以外の支出	繰越金
年間収入	1.00																				
収入総額	0.60	1.00																			
実収入	0.66	0.78	1.00																		
実収入以外の収入	0.35	0.85	0.36	1.00																	
繰入金	0.19	0.26	0.14	0.04	1.00																
支出総額	0.60	1.00	0.78	0.85	0.26	1.00															
実支出	0.60	0.73	0.56	0.63	0.17	0.73	1.00														
消費支出	0.49	0.66	0.45	0.61	0.16	0.66	0.97	1.00													
食料	0.47	0.42	0.37	0.31	0.17	0.42	0.52	0.50	1.00												
住居	-0.02	0.11	0.00	0.16	0.01	0.11	0.24	0.28	-0.03	1.00											
光熱・水道	0.32	0.24	0.22	0.16	0.11	0.24	0.28	0.27	0.44	-0.07	1.00										
家具・家事用品	0.15	0.25	0.12	0.26	0.09	0.25	0.26	0.27	0.17	0.07	0.10	1.00									
被服及び履物	0.30	0.30	0.24	0.24	0.10	0.30	0.39	0.38	0.29	0.02	0.12	0.16	1.00								
保健医療	0.11	0.16	0.10	0.15	0.07	0.16	0.24	0.25	0.15	0.01	0.07	0.08	0.09	1.00							
交通・通信	0.14	0.33	0.15	0.37	0.04	0.33	0.54	0.57	0.12	0.01	0.05	0.05	0.10	0.06	1.00						
教育	0.18	0.23	0.15	0.23	0.03	0.23	0.37	0.39	0.24	-0.03	0.19	0.02	0.09	0.04	0.07	1.00					
教養娯楽	0.32	0.35	0.27	0.30	0.12	0.35	0.44	0.42	0.32	0.02	0.10	0.15	0.26	0.10	0.10	0.09	1.00				
その他の消費支出	0.39	0.46	0.38	0.37	0.12	0.46	0.66	0.66	0.21	0.01	0.13	0.12	0.19	0.11	0.12	0.04	0.16	1.00			
非消費支出	0.70	0.63	0.70	0.38	0.12	0.63	0.62	0.43	0.35	-0.02	0.19	0.12	0.26	0.08	0.17	0.14	0.29	0.34	1.00		
実支出以外の支出	0.44	0.90	0.72	0.79	0.04	0.90	0.40	0.32	0.25	0.01	0.14	0.18	0.17	0.08	0.14	0.11	0.22	0.23	0.49	1.00	
繰越金	0.16	0.24	0.13	0.06	0.86	0.24	0.13	0.12	0.13	0.02	0.10	0.07	0.07	0.05	0.02	0.02	0.08	0.10	0.10	0.01	1.00

3．擬似ミクロデータの特性

量的属性間の相関係数 擬似ミクロ

赤字　符号が変化
差：教育用擬似ミクロデータ－個票データ
-0. 2より小、0. 2より大
-0. 2以上-0. 1より小、0. 1より大0. 2以下

擬似ミクロデータ	年間収入	収入総額	実収入	実収入以外の収入	繰入金	支出総額	実支出	消費支出	食料	住居	光熱・水道	家具・家事用品	被服及び履物	保健医療	交通・通信	教育	教養娯楽	その他の消費支出	非消費支出	実支出以外の支出	繰越金
年間収入	1.00																				
収入総額	0.58	1.00																			
実収入	0.63	0.85	1.00																		
実収入以外の収入	0.38	0.83	0.48	1.00																	
繰入金	0.12	0.32	0.15	0.05	1.00																
支出総額	0.58	1.00	0.85	0.83	0.32	1.00															
実支出	0.52	0.71	0.59	0.64	0.14	0.71	1.00														
消費支出	0.42	0.63	0.49	0.60	0.14	0.63	0.96	1.00													
食料	0.46	0.40	0.36	0.32	0.13	0.40	0.45	0.43	1.00												
住居	-0.05	0.08	0.04	0.09	0.03	0.08	0.24	0.28	-0.06	1.00											
光熱・水道	0.32	0.25	0.23	0.18	0.09	0.25	0.26	0.25	0.44	-0.07	1.00										
家具・家事用品	0.12	0.15	0.11	0.14	0.04	0.15	0.19	0.19	0.15	0.00	0.10	1.00									
被服及び履物	0.21	0.23	0.19	0.20	0.06	0.23	0.29	0.28	0.20	0.01	0.08	0.12	1.00								
保健医療	0.07	0.13	0.09	0.13	0.04	0.13	0.19	0.20	0.11	0.00	0.06	0.05	0.05	1.00							
交通・通信	0.12	0.30	0.17	0.35	0.04	0.30	0.50	0.54	0.10	-0.01	0.05	0.03	0.06	0.04	1.00						
教育	0.14	0.24	0.18	0.24	0.02	0.24	0.38	0.41	0.18	-0.02	0.16	0.01	0.04	0.02	0.04	1.00					
教養娯楽	0.26	0.30	0.24	0.28	0.06	0.30	0.35	0.34	0.26	-0.01	0.06	0.12	0.18	0.07	0.07	0.05	1.00				
その他の消費支出	0.33	0.44	0.38	0.37	0.11	0.44	0.63	0.65	0.17	-0.02	0.11	0.07	0.11	0.06	0.09	0.04	0.10	1.00			
非消費支出	0.50	0.50	0.52	0.35	0.07	0.50	0.53	0.26	0.24	-0.04	0.14	0.07	0.14	0.05	0.09	0.07	0.18	0.21	1.00		
実支出以外の支出	0.45	0.85	0.77	0.74	0.07	0.85	0.32	0.25	0.25	-0.05	0.15	0.08	0.13	0.05	0.09	0.09	0.18	0.18	0.35	1.00	
繰越金	0.10	0.28	0.14	0.05	0.82	0.28	0.07	0.07	0.09	0.00	0.08	0.03	0.03	0.02	0.01	0.00	0.02	0.06	0.04	0.00	1.00

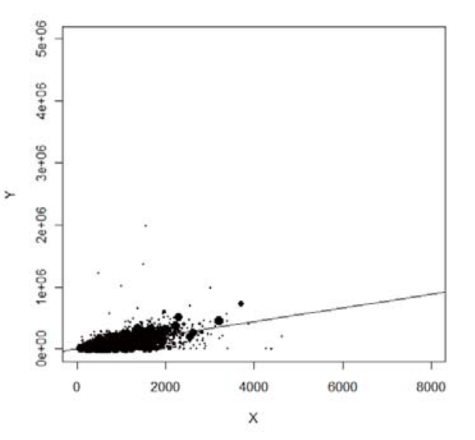
3．擬似ミクロデータの特性

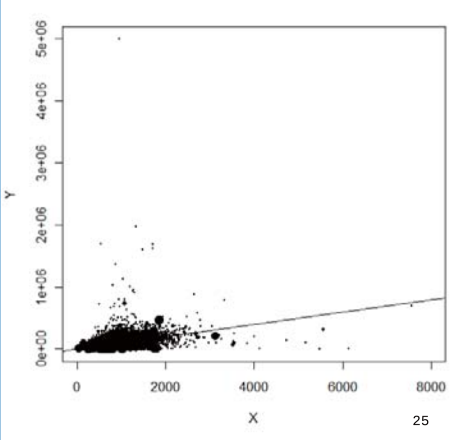
□ 散布図

調査票情報と擬似ミクロの間で、相関係数の差が最も大きい「年間収入」と「非消費支出」の散布図。

年間収入(x)×非消費支出(y)

調査票情報	擬似ミクロデータ





3．擬似ミクロデータの特性

□ クロス集計表

公表値（平成16年 全国消費実態調査報告第13表）を参考に作成

差：（教育用擬似ミクロデータ－個票データ）／個票データ
-0.5より小、0.5より大
-0.5以上-0.1より小、0.1より大0.5以下

世帯人員２人、年間収入階級別１世帯当たり１か月間の支出

調査票情報（公表値）

	平 均	年間収入階級（万円） 200 未満	200 ～ 300	300 ～ 400	400 ～ 500	500 ～ 600	600 ～ 800	800 ～1000	1000 ～1250	1250 ～1500	1500 以上
食料	60, 984	38, 439	46, 341	51, 449	55, 430	59, 147	64, 313	68, 487	71, 419	81, 540	92, 303
住居	23, 905	25, 703	22, 623	23, 331	26, 316	24, 298	22, 386	24, 143	27, 417	22, 171	15, 331
光熱・水道	15, 096	12, 304	13, 709	14, 515	14, 544	14, 654	15, 349	15, 822	16, 155	16, 640	20, 022
家具・家事用品	9, 286	4, 398	5, 549	7, 483	7, 368	8, 698	10, 255	12, 056	10, 972	13, 492	14, 796
被服及び履物	13, 654	5, 904	6, 825	7, 955	8, 821	11, 513	14, 737	17, 283	22, 484	26, 940	37, 843
保健医療	11, 134	5, 501	8, 175	9, 444	10, 006	11, 713	11, 314	13, 163	12, 805	14, 661	16, 623
交通・通信	45, 703	20, 996	25, 966	32, 570	38, 339	46, 288	46, 286	55, 709	63, 407	78, 109	81, 492
教育	1, 213	3, 641	2, 530	1, 260	836	1, 153	785	697	1, 940	887	1, 142
教養娯楽	30, 863	11, 051	16, 564	18, 288	23, 541	26, 479	34, 068	37, 126	51, 438	56, 794	67, 421
その他の消費支出	89, 539	27, 955	38, 910	51, 071	60, 584	76, 484	97, 967	120, 000	153, 095	163, 964	208, 602

擬似ミクロデータ

	平 均	年間収入階級（万円） 200 未満	200 ～ 300	300 ～ 400	400 ～ 500	500 ～ 600	600 ～ 800	800 ～1000	1000 ～1250	1250 ～1500	1500 以上
食料	60, 342	40, 970	44, 783	50, 789	55, 577	59, 450	62, 807	68, 987	72, 671	77, 010	87, 092
住居	22, 241	23, 467	25, 408	22, 322	24, 755	24, 034	19, 762	25, 761	15, 321	15, 644	9, 864
光熱・水道	15, 062	12, 565	13, 113	14, 135	14, 410	15, 090	15, 190	16, 060	16, 440	17, 481	19, 122
家具・家事用品	8, 550	4, 413	6, 072	7, 048	7, 472	8, 272	9, 069	9, 775	10, 675	13, 353	13, 103
被服及び履物	13, 215	7, 535	6, 748	7, 746	9, 586	11, 361	14, 166	18, 306	20, 846	24, 587	29, 413
保健医療	10, 386	7, 824	7, 702	8, 618	9, 970	9, 575	10, 846	12, 150	12, 488	13, 335	13, 873
交通・通信	43, 772	23, 271	24, 048	34, 239	35, 225	44, 714	46, 207	53, 821	61, 657	59, 127	79, 587
教育	1, 337	4, 640	4, 141	1, 196	1, 489	830	744	398	634	0	10, 052
教養娯楽	30, 194	11, 654	15, 620	19, 395	23, 153	28, 587	32, 610	40, 740	43, 390	50, 752	60, 672
その他の消費支出	85, 264	26, 578	38, 004	52, 043	59, 657	77, 723	90, 088	116, 778	142, 395	160, 150	216, 446

3．擬似ミクロデータの特性

□ 回帰

用途項目の１か月当たり支出金額(円) = f (消費支出(円))

調査票情報

	消費支出 係数	消費支出 標準誤差	消費支出 P値	定数 係数	定数 標準誤差	定数 P値	決定係数	調整済 決定係数	F値
食料	0. 07681	0. 00075	0. 000	47, 683. 0	294. 2	0. 000	0. 2456	0. 2456	10, 425
住居	0. 07649	0. 00146	0. 000	-6, 558. 3	571. 0	0. 000	0. 0789	0. 0789	2, 744
光熱・水道	0. 01123	0. 00022	0. 000	15, 585. 8	86. 6	0. 000	0. 0744	0. 0744	2, 574
家具・家事用品	0. 02205	0. 00044	0. 000	2, 305. 1	172. 9	0. 000	0. 0720	0. 0720	2, 486
被服及び履物	0. 03669	0. 00050	0. 000	2, 202. 6	195. 8	0. 000	0. 1436	0. 1435	5, 368
保健医療	0. 02551	0. 00055	0. 000	3, 284. 6	214. 9	0. 000	0. 0630	0. 0630	2, 153
交通・通信	0. 25071	0. 00200	0. 000	-34, 299. 9	782. 4	0. 000	0. 3289	0. 3289	15, 699
教育	0. 10323	0. 00138	0. 000	-12, 682. 5	538. 7	0. 000	0. 1491	0. 1491	5, 613
教養娯楽	0. 06986	0. 00084	0. 000	8, 775. 9	327. 5	0. 000	0. 1785	0. 1785	6, 959
その他の消費支出	0. 32741	0. 00206	0. 000	-26, 296. 4	805. 5	0. 000	0. 4410	0. 4410	25, 262

擬似ミクロデータ

	消費支出 係数	消費支出 標準誤差	消費支出 P値	定数 係数	定数 標準誤差	定数 P値	決定係数	調整済 決定係数	F値
食料	0. 06232	0. 00074	0. 000	52, 433. 5	280. 3	0. 000	0. 1826	0. 1826	7, 156
住居	0. 08949	0. 00169	0. 000	-11, 679. 6	641. 7	0. 000	0. 0808	0. 0808	2, 815
光熱・水道	0. 01002	0. 00022	0. 000	15, 951. 0	82. 2	0. 000	0. 0628	0. 0628	2, 147
家具・家事用品	0. 01509	0. 00043	0. 000	4, 251. 4	161. 8	0. 000	0. 0378	0. 0378	1, 259
被服及び履物	0. 02908	0. 00055	0. 000	4, 594. 6	210. 1	0. 000	0. 0797	0. 0797	2, 774
保健医療	0. 01958	0. 00055	0. 000	4, 940. 7	208. 9	0. 000	0. 0382	0. 0382	1, 271
交通・通信	0. 23652	0. 00207	0. 000	-29, 652. 4	788. 4	0. 000	0. 2891	0. 2891	13, 023
教育	0. 13728	0. 00170	0. 000	-22, 777. 2	645. 8	0. 000	0. 1696	0. 1695	6, 539
教養娯楽	0. 05808	0. 00089	0. 000	12, 331. 1	339. 7	0. 000	0. 1167	0. 1166	4, 230
その他の消費支出	0. 34253	0. 00226	0. 000	-30, 392. 9	860. 4	0. 000	0. 4173	0. 4173	22, 934

3．擬似ミクロデータの特性

□ 擬似ミクロデータ作成における検討課題

- 参考 「教育用擬似ミクロデータの開発とその利用
  ～平成16年全国消費実態調査を例として～」
  （製表技術参考資料 平成24年7月（独）統計センター）
  http://www.nstac.go.jp/services/reference.html

  - 多変量対数正規乱数を発生させた場合に、実際の個票データでは出現していない範囲を超えた数値が作成されてしまうこと。
  - 高次元の集計表の中で度数が少ないセルのデータについて、個票データの分布との乖離が大きくなってしまうこと。
  - セル内の度数に対応するレコードの属性値群のすべてが、基本的な量的属性値の内訳に関して同一の構成比とならないように、レコードによって構成比を変える必要性あり。
  - 新たな擬似ミクロデータの作成。
    （今度は、質的属性が多い統計調査に関するもの）

NSTAC 独立行政法人 統計センター

４．擬似ミクロデータの課題

□ 公的統計における制度上の位置付け

「統計データの二次的利用促進に関する研究会　平成23年度報告書」（抄）
平成24年７月　総務省政策統括官室（統計基準担当）

さらに、諸外国における政府統計データの活用の状況について幅広くとらえてみると、ニュージーランド統計局では、公表されている集計結果表から作成されたミクロデータが利用制限なしに提供されていることがわかった。日本では、統計センターにおいて、同様の観点から作成する擬似ミクロデータの研究が行われている。

（中略）

このため、このような新たな技術的手法による擬似ミクロデータについて、その研究を進めるとともに、総務省においてその制度上の位置づけ等について検討を行うことが求められる。

独立行政法人 統計センター NSTAC

４．擬似ミクロデータの課題

□ 公的統計における制度上の位置付け

「平成23年度統計法施行状況に関する審議結果報告書」（抄）
平成24年9月25日　内閣府 統計委員会

（イ）施策の進捗状況等に対する評価
○また、法第３３条第２号による調査票情報の利用については、調査対象の識別可能性や情報漏えいのリスクへの対応として厳格な運用が求められている中で、オンサイト利用や、匿名データ・教育用擬似ミクロデータとの関係整理も必要である。

（ウ）今後の施行の方向性等についての基本的考え方
○また、二次的利用を取り巻く諸課題については、総務省の研究会における検討情報を注視していくこととする。なお、オンサイト利用や教育用擬似ミクロデータの検討に当たっては、コストやその負担、国民の理解や研究者に対する国民の信頼感にも密接に関係することに留意が必要である。

独立行政法人 統計センター NSTAC

４．擬似ミクロデータの課題

□ 擬似ミクロデータの利活用のために

- 擬似ミクロデータを用いて行った授業の資料
- 擬似ミクロデータを利用するために
  作成したプログラム
  （CSVデータのインポート、分析 等）　など

上のようなコンテンツへ、統計センターウェブ（擬似ミクロコーナー）からリンクを張らせていただきたいと考えています。その際は、ご理解、ご協力をお願いします。

※　擬似ミクロデータ の全部又は一部などをウェブで提供する行為は、控えて下さい。
　統計センターが試行提供している擬似ミクロデータは、統計センターから直接入手してください（申出者・利用者による第三者提供は禁止）。

独立行政法人 統計センター NSTAC

ご清聴ありがとうございました。

意見、見解に及ぶものは、本人のものであり、
所属する組織を代表するものではありません。

独立行政法人 統計センター NSTAC

資料２－１（参考②）

参考２

擬似ミクロデータの利用目的（アンケート結果）

平成 24 年 12 月 19 日
独立行政法人　統計センター

平成 24 年 11 月 30 日時点

		申出件数	利用目的				利用人数
			大学・高等学校等での授業	匿名データ提供申出前の演習	データ処理のためのプログラムテスト	その他	
平成 23 年度		45	45	0	1	0	363
平成 24 年度	8月まで	8	8	0	0	0	508
	9月以降	14	11	2	5	0	98
合計		67	64	2	6	0	969

注：「利用目的」は複数回答可。
　　平成 24 年 8 月にそれまでの試行的提供の結果から提供内容を見直したため、分けて掲載している。