



ビジネスレジスター 大企業のプロファイリング

発表内容

- なぜプロファイリングをするのか?
- ビジネスレジスター
 - 母集団 universe
 - 資源
 - 維持・管理戦略
- プロファイルにおける対象の選定プロセス
- プロファイリング
- 品質保証
- 意義

2

Statistics Canada • Statistique Canada

なぜプロファイリングをするのか?

- 各分野別の調査において、複雑な構造の企業が存在するがゆえに母集団名簿(フレーム)にかかる負担を軽減するため。
統計調査と公表に焦点を当てている。
- さまざまな分野別の調査担当からのフィードバックを増やすことにより、構造の品質と信頼性を向上させるため。
概念と手順の整合性を保証する。

3

Statistics Canada • Statistique Canada

ビジネスレジスターにおける活動中の事業の母集団

- 2011年12月(時点)の包括的な調査母集団ファイル
 - 企業
 - 単独 事業所企業 - 243 万レコード
 - 複数事業所企業 - 2万6千レコード
 - 事業所
 - 単独 事業所企業 - 243 万レコード
 - 複数事業所企業 - 13万2千レコード

4

Statistics Canada • Statistique Canada

資源

- 平均 38 人のプロファイラー
毎日情報を更新
- 資源の不足
2万6千社の複数事業所企業の構造の変化を十分に監視するためには不十分。

5

Statistics Canada • Statistique Canada

維持・管理戦略

- 上位5千社:
 - プロファイル
2~3年に1回
企業集団における変化が探知されたとき。
 - 維持
必要に応じてプロセス・フィードバックを実施。
 - 監視
更新は各分野別の調査担当及び調査票回収担当による。

6

Statistics Canada • Statistique Canada

維持・管理戦略

- その他の2万1千社:
 - 品質保証監視
 - 必要に応じた維持
 - 以下の関与の増大:
各分野別の調査担当の母集団名簿(フレーム)専門官
調査票回収担当

7

Statistics Canada • Statistique Canada

維持・管理戦略

- 単純な構造(の企業)
 - 1つの法的主体/1つの事業主体
- 以下によって更新:
 - 税務記録による置換
 - 各分野別の調査担当
 - 調査票回収担当

8

Statistics Canada • Statistique Canada

維持・管理戦略

- 単純な複数事業所企業
 - 1つの法的主体が複数の事業を行っているケース。
- 更新は以下のように分担している:
 - 各分野別の調査担当
 - 産業が1つの企業
 - プロファイラー
 - 複数の産業にまたがる企業

9

Statistics Canada • Statistique Canada

維持・管理戦略

- 大きく複雑な(企業)構造
 - 複数の法的主体、事業活動(から構成され)、(複数の)産業分類(にまたがっている)。
- 更新は以下の分担で行われている:
 - プロファイリング活動
 - ビジネスレジスターと企業プログラムマネージャー

10

Statistics Canada • Statistique Canada

プロフィールにおける対象の選定プロセス

- ステップ 1:
 - 順位付けシステムが以下の変数を用いてプロフィールの対象となる上位5千単位を選定:
 - 複数の法的主体(から構成されているか)
 - 複数の事業主体(から構成されているか)
 - 複数の産業分類(にまたがっているか)
 - 複数の州(にまたがっているか)
 - 連結(決算を行っているか)

11

Statistics Canada • Statistique Canada

プロフィールにおける対象の選定プロセス

- ステップ 1:
 - 特定の産業及び地域については規模変数(収入、資産、雇用者数)も考慮するが、プロフィールにおける対象の選定においては重要な要素ではない。
 - 複雑さの方がより重要な要素である。(例えば)規模変数の値が大きいが、単純な構造の企業はプロファイリング以外の方法で更新される。

12

Statistics Canada • Statistique Canada

プロフィールにおける対象の選定プロセス

■ ステップ 2:

- 企業プログラムマネージャー (EPM) プログラムの一部として単位を選定する。
 - 企業統計課に割り当てられた単位についてはプロファイリングと調査票回収活動(の両方)を徹底するために実際に訪問する。
 - 各人25~30単位。
 - 単位は上位から下位まで2~3年に1回はプロファイリングされる: 変化の度合いに依存する。
 - (システムからの)更新要求メッセージにより単位は絶えず監視される。

プロフィールにおける対象の選定プロセス

■ ステップ 2:

- 複雑な企業の局内プログラム (CIP) に選定される単位。
- ビジネスレジスターCIPチームに割り当てられた単位は調査及び電話によってデータ収集され、全てプロフィールされる。
- 各人100単位。
 - 単位は上位から下位まで2~3年に1回はプロファイリングされる - 変化の度合いに依存する。
 - (システムを経由した各分野別の調査担当、調査票回収担当からの)更新要求メッセージにより単位は絶えず監視される。

プロフィールにおける対象の選定プロセス

■ ステップ 2:

- 分野別調査担当に対して高い重要性で選定される単位 - 上位5千単位の残り
 - ビジネスレジスターのプロファイリング・チームに割り当てられた単位に対して、必要に応じて、調査及び/あるいは電話によるデータ収集を実施。
 - 企業構造の上から下まで2年に1回はプロファイリングされる - 変化の度合いに依存する。
 - (システムを経由した各分野別の調査担当、調査票回収担当からの)更新要求メッセージにより単位は絶えず監視される。

プロフィールにおける対象の選定プロセス

■ ステップ 3:

- プロファイリングを優先する単位の選定
 - プロファイラー : EPM, BR CIP, BRのプロファイリング・チームはそれぞれの責任によりプロフィールされる単位の順序を決める。
 - すでに、(一定の)方法論に従って作成されたプロセスによって、(プロフィールされる対象は)それぞれどこが担当するか選定されているので、プロファイラーが考慮するのは:
 - 最後にプロフィールされた日
 - 調査された単位数
 - 未処理の(システムからの)メッセージの数
 - 未処理の(システムからの)メッセージの性質
 - 分野別調査における(企業)構造の影響

プロファイリング

- 行政記録データ
- フィードバック
 - 調査票回収 – 調査に対する回答
 - 分野別調査担当からのフィードバック
 - 外部のデータによるシグナル
- 調査
 - インターネット
官報公示, 企業のホームページ、他

17

Statistics Canada • Statistique Canada

プロファイリング

- 報告者へのコンタクト
 - 電話照会
ビジネスレジスターのプロファイラー
 - 訪問照会
企業プログラムマネージャー (EPM) のプロファイラー

18

Statistics Canada • Statistique Canada

プロファイリング

- 頻度
 - 2~3年に1回
調査票回収活動を通じて判明した変化の数と(企業)構造に対する変化の影響に依存する。
 - 暫定処置
更新と変更は報告者へのコンタクトなしに行う。

19

Statistics Canada • Statistique Canada

プロファイリング

- 1年当たりの報告者の数
 - 企業プログラムマネージャー (EPM)
-1人 1年当たり10 単位 – 100% コンタクト
 - 複雑な企業の局内プログラム(CIP)
-1人 1年当たり30 ~35単位 – 95% コンタクト
 - ビジネスレジスターのプロファイリング・チーム
- 1人1年当たり65~70 単位 – 50% ~60% コンタクト

20

Statistics Canada • Statistique Canada

プロファイリング

- 報告による方法の採用
 - プロファイラーは全てのケースについて情報を収集し、まとめる。
 - 報告者は(企業の組織図など)自身に関する情報を要求された場合には提供しなければならない、あるいは自発的に提供することがありうる。:その情報は電子媒体あるいは紙媒体のどちらでも良く、報告者の選択に任されている。

品質保証

- 品質保証チーム

(企業)構造の変化及び(企業)構造(に関する情報)の更新における手順と概念の一貫性と妥当性を監視する。
- 合同スタッフプロセス

複数の分野別調査に影響を与える単位については、データの関連性と正確さに与える影響の大きさに応じて、影響を受けると判断された全ての関連分野別調査の担当が参加して検討される。

プロファイリング

- 報告者に対する関連する研修
 - 研修不要:プリセットの調査票ではない。
 - 必要に応じ情報が要求される。
 - プロファイラーは調査と行政記録を通じて全ての変更を確認しようとする。
 - 解消されなかった事項についてのみ報告者に要求。
 - この方法は報告者負担を大きく減らし、データ収集に良い影響をもたらす。

品質保証

- プロファイラーは通知システムを通じて割り当てられた(企業)構造の品質を保障する。
 - ビジネスレジスターシステムの利用者によって、それぞれの責任担当の(企業)構造について、処理が行われたいかなる変更も、正確性や一貫性を、以下から割り当てられたプロファイラーが検討する。
 - 各分野別の調査担当
 - 調査票回収担当
 - 外部データ処理担当

統計プログラムに対する意義

- プロファイリング目的に対する毎年の予算を配分
- 各分野別の調査担当の期待と信用
 - プロファイリングについて
 - 支援プログラムについて
 - 調査マネージャーと研修
 - 品質保証プロセス
 - 通知と月次報告

統計プログラムに対する意義

- 整合性と一貫性:
 - フレームデータ
 - 概念
 - 手順
 - 方法論



統一企業調査 (UES) の標本抽出の概観

発表の内容

1. 統一企業調査の概観
2. 母集団名簿 (フレーム)
3. 標本抽出単位
4. 層化と配分
5. ネットワークサンプリング
6. 標本選定と調整
7. データ置き換え
8. 統合産業統計プログラム

カナダ



統一企業調査 (UES)

- 統一企業調査 (UES) は1997年に開始された。
 - 最終目標：州単位の高い品質の経済統計の作成
- 主な利用者：国民経済計算
 - その他の利用者には連邦政府省庁、州政府、民間部門、組織、学術、他。

統一企業調査 (UES)

- 約60種類の年次産業調査から構成されている。
 - サービス (約30調査)
 - 製造業 (約20調査)
 - 林業
 - 小売業
 - 卸売業
 - 水産養殖業
 - 本社

母集団名簿 (フレーム)

- 統一企業調査 (UES) は全ての調査の共通の母集団名簿 (フレーム) を活用している。
- 統一企業調査 (UES) はある時点のビジネスレジスター (BR) のスナップショットから標本を選定している。
 - 例えば、参照年が2011年の場合、2011年10月1日ビジネスレジスターのスナップショットが母集団名簿 (フレーム) を取り出すために用いられている。

母集団名簿 (フレーム)

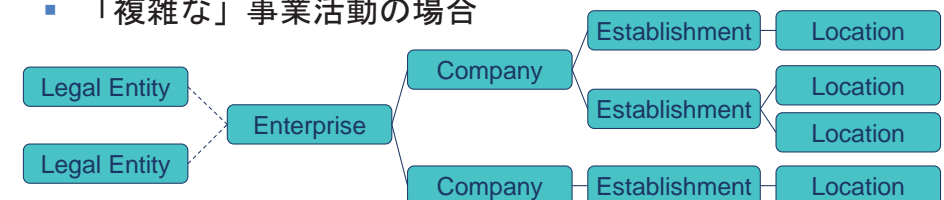
- 標本抽出に用いられるビジネスレジスター (BR) からの変数は：
 - 識別子 (事業者番号、企業識別番号、事業所識別番号、...)
 - 産業分類
 - 地理情報 (州、他...)
 - プロファイルあるいは行政記録データからの規模変数 (収入、支出、雇用者数 ...)

母集団名簿 (フレーム)

- ビジネスレジスター (BR) には 事業活動について4階層の統計構造、事業構造、法的構造がある。
- 単位間の接続はほとんどプロファイリングによって構築されている。いくつかは行政記録データによる。
- 「単純な」事業活動の場合。



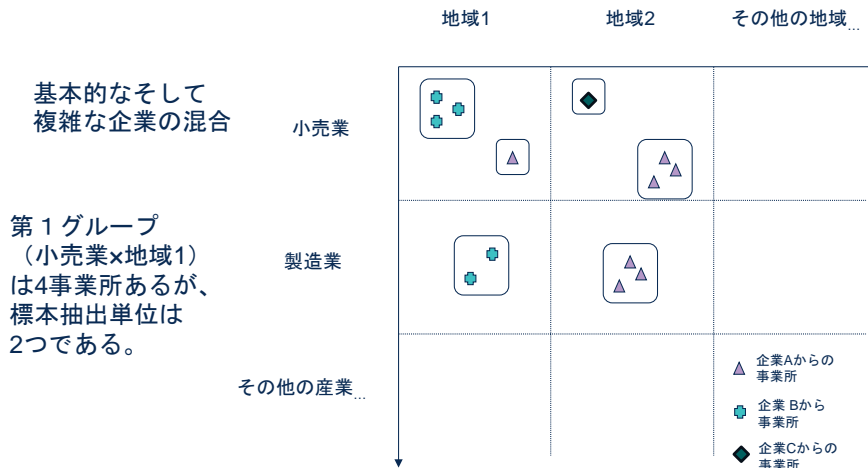
- 「複雑な」事業活動の場合



母集団名簿（フレーム）

データの種類	統計レベル
完全な一組の財務諸表	Enterprise 企業
収入、支出、投資	Company 企業
収入、支出	Establishment 事業所
収入あるいは支出	Location 区画

標本抽出単位の例



標本抽出単位

- 統一企業調査（UES）は事業所ベースの調査である。
- 標本抽出単位は同一の企業、同一の州、同一の産業（これを「セル」と呼ぶ）の中の全ての事業所から構成される。

複雑な企業

- 統一企業調査（UES）において、ビジネスレジスター（BR）の企業は、以下がどれかが当てはまると「複雑な」企業とされる。：
 - 一つ以上の法的主体と接続している。
 - 事業所は複数の州にある。
 - 事業所は異なる事業を営んでいる。
- 以上が当てはまらなければ、企業は「単純」である。

標本抽出の基礎：層化

- 悉皆層 Must-Take Stratum
- 非抽出層 Take-None Stratum
- 層化変数 Stratification variables
 - 産業（北米産業分類体系） Industry (NAICS)
 - 州 Province
 セル：産業×州 Cell: Industry by province
 - 規模 Size

悉皆層 Must-Take Stratum

- 悉皆層の単位は標本の中に（必ず）ある。
- 基準の例
 - 収入が25百万カナダドルの複雑な企業に所属する標本抽出単位（は悉皆層になる）
- 分野別の調査担当によって（調査されるべきであると）特定された単位も悉皆層になる。

非抽出層 (TN)

- 標本抽出の層化に用いられる収入がセル（産業×州）の下位10%の標本抽出単位から構成される。
- これらの単位は標本抽出にふさわしくない。
 - コントクトを取らない – この部分を推定するために行政記録データが用いられる。
- これらの単位は費用と報告者負担の観点から除外されている。

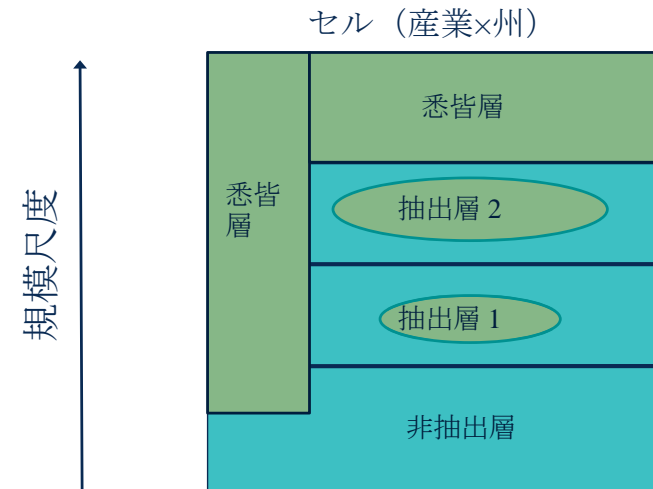
規模層化：規模尺度

- 事業所の規模の定義においては、異なる収入変数が利用可能である：
 - ビジネスレジスターからの、プロフィールによって得られた「収入」と行政記録データからの「収入」
 - 調査データからの「収入」
 - 分野別の調査担当が指定した「収入」

規模層化

- 推定精度を向上させるために規模層化を行う。
- 統一企業調査 (UES) には1つの悉皆層 (TA) と2つの抽出層 (TS) がある。
 - 悉皆層：確実に抽出される。
 - 抽出層：無作為に抽出される。小さな抽出層の抽出率は低い。
- 境界はラヴェリー＝ヒドログオウ (Lavallée-Hidiroglou(1988)) が考案した手法を用いて設定されている。
 - 所与の目標精度に対して標本数を最小化する層化と配分を定義するアルゴリズム

層化の概念図



標本数の配分

- 各セルの目標精度を入力として、層化／配分アルゴリズムは動く。
- 全統一企業調査の目標精度 (の設定) から開始する。
- それから相対的な重要性に応じて各州、産業、セルに配分する。

標本選定

- 標本は単純無作為抽出 (SRS)で選定される。
 - 必要な調整が考慮される。
- 選定後、ネットワークサンプリングを通じて追加的な単位が加わる。

標本の調整

- 標本について3種類の調整法がとられている。
 - Independent Sampling 独立標本抽出法
 - Maximum Overlap 最大重複法
 - Controlled Overlap 管理重複法
- 最大重複法は異なる時点のデータの比較とより良い時間的傾向の推定値を保証する。

ただし：

 - 選ばれた企業に重い報告者負担
 - 時間的傾向の把握は年次調査の主な目的ではない。水準である。

標本の調整

- Maximal overlap 最大重複法
 - 一度選定されると、その事業は廃業するか、セルを変えるか、新規参入により追い出されるまで標本であり続ける。



標本の調整

- **Controlled Overlap 管理重複法**
- 最大重複法で行われてきた調査の報告者負担を軽減する。
 - 抽出層にのみ適用
 - Sampling window は少しだけシフトする。



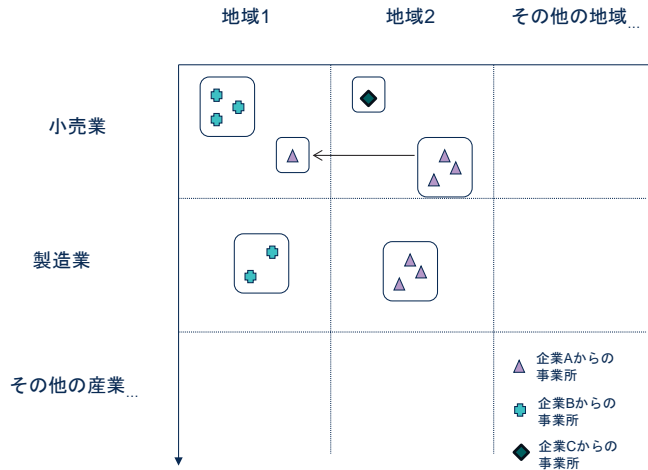
ネットワークサンプリング

- ネットワークサンプリング：もしも、ある企業が選定されたとき、この企業の、同一調査の対象となる全ての単位が選ばれる。
- 調査レベルで企業の完全なイメージを得るために行われている。
 - 追加的な情報を得るために、何ら追加的な費用がかからない。
- 標本抽出ウェイトが調整される。

ネットワークサンプリングの例

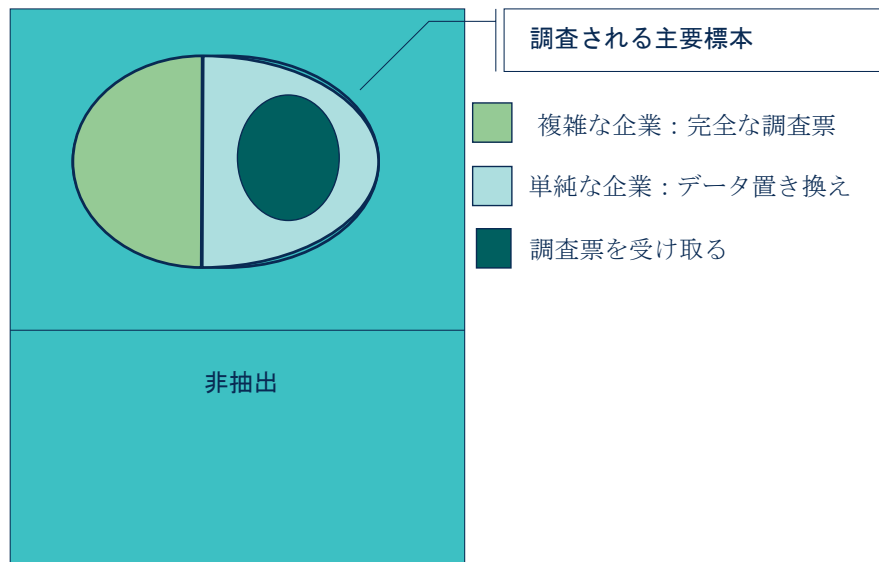
ステップ1-いくつかの標本抽出単位が標本に選定される。

ステップ2-ネットワークサンプリングは、もしも同一企業で同一の調査であれば、単位を追加する。



データの置き換え

- 報告者負担（と費用）を軽減するために単純な企業については
 - 経理変数については、調査データの代わりに行政記録データが用いられる。
 - 標本に選定された単位の副標本のみ。
 - 経理変数以外については、実査あるいは補完で得られる。



統合産業統計プログラム（IBSP）

参照年次2013年より開始

- 目的：
 - 報告者負担の軽減
 - 効率化を実現
 - プロセスの近代化
 - 統合
 - プロセスの単純化
 - 開発費と維持・管理費を削減

IBSP の主な特徴

- 異なる周期の異なる性格の調査をより統合する。
- 報告者負担を軽減するための異なる調査の調整
- 行政記録データをより活用することにより報告者負担を軽減。
- 特性及び商品の推定値を作成するための、より豊かな母集団（フレーム）情報を収集するために二層抽出を行う。

IBSP の主な特徴

- 電子媒体（ネット）でのデータ収集が主要なモードとなる。
- 電話照会を減らすために調査票回収マネジメントが行われる。
- 公表の速報化のために実査、加工、分析が並行して行われる。
- 概念は調査間で標準化される。