

統計学

第 120 号

『統計学』創刊60周年記念論文

特集B：政府統計マイクロデータの作成・提供における方法的展望

公的統計における統計メタデータ・アーカイブの展開可能性

..... 小林 良行 (1)

研究論文

高学歴雇用者の専攻と男女間賃金格差

..... 武内真美子 (19)

本会記事

経済統計学会第64回(2020年度)全国研究大会・会員総会に関する追記 (35)

支部だより..... (39)

投稿規程..... (41)

2021年3月

経済統計学会

創刊のことば

社会科学の研究と社会的実践における統計の役割が大きくなるにしたがって、統計にかんする問題は一段と複雑になってきた。ところが統計学の現状は、その解決にかならずしも十分であるとはいえない。われわれは統計理論を社会科学の基礎のうえにおくことによって、この課題にこたえることができると考える。このためには、われわれの研究に社会諸科学の成果をとりいれ、さらに統計の実際と密接に結びつけることが必要であろう。

このような考えから、われわれは、一昨年来経済統計研究会をつくり、共同研究を進めてきた。そしてこれを一層発展させるために本誌を発刊する。

本誌は、会員の研究成果とともに、研究に必要な内外統計関係の資料を収めるが同時に会員の討論と研究の場である。われわれは、統計関係者および広く社会科学研究者の理解と協力をえて、本誌をさらによりよいものとするを望むものである。

1955年4月

経済統計研究会

経済統計学会会則

第1条 本会は経済統計学会（JSES：Japan Society of Economic Statistics）という。

第2条 本会の目的は次のとおりである。

1. 社会科学に基礎をおいた統計理論の研究
2. 統計の批判的研究
3. すべての国々の統計学界との交流
4. 共同研究体制の確立

第3条 本会は第2条に掲げる目的を達成するために次の事業を行う。

1. 研究会の開催
2. 機関誌『統計学』の発刊
3. 講習会の開催、講師の派遣、パンフレットの発行等、統計知識の普及に関する事業
4. 学会賞の授与
5. その他本会の目的を達成するために必要な事業

第4条 本会は第2条に掲げる目的に賛成した以下の会員をもって構成する。

- (1) 正会員
- (2) 院生会員
- (3) 団体会員
- 2 入会に際しては正会員2名の紹介を必要とし、理事会の承認を得なければならない。
- 3 会員は別に定める会費を納入しなければならない。

第5条 本会の会員は機関誌『統計学』等の配布を受け、本会が開催する研究大会等の学術会合に参加することができる。

- 2 前項にかかわらず、別に定める会員資格停止者については、それを適応しない。

第6条 本会に、理事若干名をおく。

- 2 理事から組織される理事会は、本会の運営にかかわる事項を審議・決定する。
- 3 全国会計を担当する全国会計担当理事1名をおく。
- 4 渉外を担当する渉外担当理事1名をおく。

第7条 本会に、本会を代表する会長1名をおく。

- 2 本会に、常任理事若干名をおく。
- 3 本会に、常任理事を代表する常任理事長を1名おく。
- 4 本会に、全国会計監査1名をおく。

第8条 本会に次の委員会をおく。各委員会に関する規程は別に定める。

1. 編集委員会
2. 全国プログラム委員会
3. 学会賞選考委員会
4. ホームページ管理運営委員会
5. 選挙管理委員会

第9条 本会は毎年研究大会および会員総会を開く。

第10条 本会の運営にかかわる重要事項の決定は、会員総会の承認を得なければならない。

第11条 本会の会計年度の起算日は、毎年4月1日とする。

- 2 機関誌の発行等に関する全国会計については、理事会が、全国会計監査の監査を受けて会員総会に報告し、その承認を受ける。

第12条 本会会則の改正、変更および財産の処分は、理事会の審議を経て会員総会の承認を受けなければならない。

付 則 1. 本会は、北海道、東北・関東、関西、九州に支部をおく。

2. 本会に研究部会を設置することができる。
3. 本会の事務所を東京都文京区音羽1-6-9（株音羽リスマチックにおく。

1953年10月9日（2016年9月12日一部改正[最新]）

公的統計における統計メタデータ・アーカイブの 展開可能性

小林良行*

要旨

本稿では、公的統計全体を対象とする統合統計メタデータ・アーカイブ及びオン
トロジーの構築を提案し、その展開可能性について指摘している。統計業務過程を
通じて生み出される統計情報は統計データとそれに関係する統計メタデータが一体
となっている情報である。本稿では、二次利用を前提とする統計業務過程のモデル
と統計情報の性質について考察している。また、個別統計調査の統計情報アーカイ
ブがデータウェアハウスで実現できることを指摘している。統計データ・アーカイ
ブと統計メタデータ・アーカイブは統計情報アーカイブを構成する。公的統計の統
計メタデータは、大局的、半大局的、局所的の3層に分類できる。この3層構造に基
づき統合統計メタデータ・アーカイブを構築すると、公的統計の用語の標準化や統
計情報のトレーサビリティなどが実現可能である。公的統計のオントロジーを確立
することにより、統計メタデータ概念の標準化が図られる。

キーワード

統計メタデータ、統計業務過程、データウェアハウス、トレーサビリティ、オン
トロジー

1. はじめに

法制度や技術の変化に伴い、統計業務の形
態や範囲は変化していくものであり、統計業
務を構成する一連の業務過程の集まりである
統計業務過程¹⁾も変化していくものである。
新統計法（平成19（2007）年法）の二次的利用
制度の下では、統計業務過程は二次利用²⁾を
前提とする業務を最初から含むようなもの
に変化することになる。調査票や統計表を二
次利用しようとするためには、それらを適切に

保存し、いつでも利用できるようにしておく
必要がある³⁾。森（2008）は調査票を二次利用
のために保存したものを情報資産と呼んでい
る。また、森（2008）、山口（2019）は、公
的統計における調査票情報の保管・管理体制を
集中型とする利点について、データの管理と
利用の面から指摘している。

公的統計の統計業務過程では、調査票から
統計表を作成する業務過程は情報システムで
行うのが一般的である。公的統計調査では近
年、オンライン調査の導入が進んできている
が紙媒体の調査票は依然として多くの調査で
利用されている。また、統計調査結果はイン

* 正会員，総務省統計研究研修所
e-mail : ykobayashi@nstac.go.jp

ターネットを通じて提供されるとともに紙媒体の調査報告書に収録され利用されている。情報は実世界に具体的に存在するものではなく、情報を直接見るには紙媒体に記録（記述・印刷）したもの（たとえば、調査票）が必要になる⁴⁾。本稿では、調査票情報とは記述済み調査票の記録内容を情報システムで扱えるようにしたものあるいは画面に表示された調査項目に直接回答を入力して情報システムで扱えるようにしたものとする。

友安（1957）は統計実務家向けに統計表の表構造について詳しく解説している。小林（2019）では、1つの調査票情報を複数の項目の直積とみて、調査票情報の集合から集計値の集合を作成する業務過程をモデル化し、統計表の表体は各集計値を表の各セルに対応付けて配置したものであると述べている。また、対応のさせ方を変えると統計表の表現形（表の見え方または表の形式のこと）は変わるが、表体に配置する元となる集計値の集合は変わらないとも述べている。本稿では、統計表情報とは調査票情報の集合から作成した複数の集計値を表形式で表現した情報とし、統計表とは統計表情報を編集して紙媒体上に記録したものと⁵⁾。なお、本稿中で、「調査票」の利用や「統計表」の利用と表現しているのは、「調査票上の記述内容」の利用や「統計表上の統計その他の印刷された内容」の利用を意味している。

調査票、調査票情報、統計表情報及び統計表は、そこに記録されている内容（以下、「情報本体」と項目の配置のしかたのように情報本体の形式的構造を表現する仕様（以下、「様式」）から構成されている⁶⁾。通常、情報本体と様式は一体として扱われる。情報本体は複数の項目（たとえば調査票なら各調査項目）の集合と考えてよい。なお、様式は、上記でいう情報本体とは別種の情報である。調査票、調査票情報、統計表情報及び統計表の情報本体は、様式と分離独立させることによ

り、共通のデータベース構造を用いて表現可能である⁷⁾。このことは、公的統計調査全体を対象とした一元的なアーカイブが構築できることを示唆する。

一般に「アーカイブ」には、①組織的に収集し保存された記録や資料、②記録を保管する組織や機関、③記録を保管する施設という3つの意味がある（たとえば、国立国語研究所（2003）、Pearce-Moses（2005）など）。本稿では、記録媒体の種類（紙媒体、電磁的媒体、光学的媒体）にかかわらず、収集された個々の記録や資料の集積全体を①の意味のアーカイブと考え、単に「アーカイブ」としたときには①の意味で用いている。一方、②及び③の意味で用いるときは、それぞれアーカイブ組織及びアーカイブ施設と表すことにする。公的統計の統計調査の調査票、調査票情報、統計表情報及び統計表を、将来の利用に供するため、収集、保存したものは、①の意味のアーカイブと言ってよい。

統計数値は、それがどのように作られ、何を表しているのかを理解していなければ適切に利用することはできない。統計数値のようなデータは、その意味を理解するのに必要となる情報と結びつかなければ単なる抽象的な数に過ぎない。一般に、データが表す意味を理解するのに必要なデータのことを「メタデータ」という。メタデータの最も簡明な定義は「データに関するデータ」であり、統計分野の情報システムにおいてメタデータに相当する語を最初に使用したのはSundgren（1973）である。UNSC and UNECE（2000）によれば、統計メタデータとは統計データのメタデータのことである。

美添（2005）、森（2008）では、調査票情報を保存する統計データ・アーカイブの必要性とともにメタデータの整備、蓄積の必要性についても指摘している。小林（2012）は、統計データの作成過程と並行して統計メタデータを収集すること及び統計データ・アーカイブ

と統計メタデータ・アーカイブを一体的に整備することの必要性を指摘している。統計改革推進会議(2017)では、統計、統計マイクロデータ及び行政記録情報にメタデータを含めた統計等データの利活用を社会全体において促進することとされており、そのための基盤の整備に言及している。諸外国と比べ、日本の公的統計分野では統計メタデータあるいは統計メタデータ・アーカイブの整備の必要性を指摘する研究等が散見されるに留まっている。統計メタデータは、統計データを解釈する上で本質的な重要性を持っているにもかかわらず、公的統計においては従来、統計データの補助的、附属的な情報として取り扱われており、あまり重要視されてこなかった。

統計情報と統計業務過程は互いに密接に関係する存在である。本稿では、統計情報とは統計業務過程の生起、進行、終了に従って発生、変化、消滅、存続していく情報及びそれらの情報相互の関係を表す情報⁸⁾からなるものであるとする。したがって、調査票(に記述された内容)、調査票情報、統計表情報及び統計表(に印刷された内容)に限定するものではない。Radermacher et al. (2009)は、統計情報とは統計データとそれに関係する統計メタデータが一体となっている情報であるとしている。本稿における統計情報アーカイブとは、統計情報が統計データ及び統計メタデータから構成されるという視点で、統計情報を保存したものである。

各府省は、総務省政策統括官(統計基準担当)(2019a)に沿って、調査票情報及び匿名データを符号表やレイアウトフォームを含むドキュメント類と共に保存し、適正に管理することが求められている。しかし、利用者が調査票情報を利用しようとする際には符号表とデータレイアウトは電子的な情報として提供されるものの、それ以外の統計メタデータ、たとえば項目の定義、調査方法、回収数(率)や有効回答数(率)、エディティング規則、エ

ディティングによる項目値の変遷情報、標本調査における推定方法や推定精度などについては提供されず、それらは調査報告書を参照せざるを得ない⁹⁾。また、統計表情報についても、分類項目と調査項目の関係、集計項目の算出方法など、やはり調査報告書を参照せざるを得ないのが現状である。

調査報告書は、統計調査に関する情報が網羅的、一元的に集積されているものである。今日では、調査票や統計表は電子的な情報として提供されるようになったものの、それらの利用上必要な統計メタデータは、調査報告書のように網羅的、一元的な形で電子的な情報として整備されているわけではない。すなわち、調査票情報や統計表情報の利用時に必要となる統計メタデータを、わざわざ調査報告書のような別の媒体上に存在する情報源から探し出し、参照しなくてはならないという不便さが解消できていない¹⁰⁾。また、電子化されていても統計メタデータが再利用可能な形式で保存されていない場合や統計データと統計メタデータあるいは統計メタデータ相互の関係付けがされていないような場合には、やはり利便性が低いと言える。統計メタデータ・アーカイブ構築の必要性の一つは、このような問題を解決することにある。

以下、第2節では統計データ及び統計メタデータの種類並びに統計メタデータ標準の動向に触れる。第3節では二次利用を前提とする統計業務過程のモデルと統計情報が具えるべき性質について述べる。第4節では公的統計の統計情報アーカイブの構造とアーカイブ実現のための技術について述べ、次いで、第5節では統計メタデータ・アーカイブの展開可能性について示す。最後に第6節で今後の展望について述べる。

2. 統計データと統計メタデータ

2.1 統計データの種類

日本では、統計データという語の意味が集

計値から個別データや匿名化データを指すものへと変化してきた¹¹⁾。一方、UNSC and UN-ECE (2000) によると、統計データとは統計マイクロデータと統計マクロデータの両者¹²⁾を包含する概念とされており、日本と比べ明確にされていると言える。

従来、統計データは個別性の有無で統計マイクロデータと統計マクロデータに分けられていた。調査票の一次利用により統計表を作成・提供する統計業務過程では、統計作成の中間段階でサマリーデータあるいは中間集計表と呼ばれるものがある。このデータは、マイクロレベルとマクロレベルの中間（メソレベル）に位置するので、メソデータ (mesodata) と呼ばれる (Radermacher et al. (2009))。メソデータは、マイクロデータを集計したものであり、マクロデータより粒度（データの集約レベル又は詳細レベルのこと）が相対的に細かいデータである。諸外国では2000年代に統計データの提供形態が多様化し、メソデータをデータキューブの形態で提供する例がみられるようになった (小林 (2012))。しかし、管見の限りでは、日本の公的統計分野でデータキューブの提供は実現していない。表1は統計データをその特徴と粒度の違いで整理したものである。

表1 統計データの種類

名称 (略称)	特徴	データの粒度	該当する例
統計マイクロデータ (マイクロデータ)	個別的, 非集約的	小 ↑ ↓ 大 細 ↓ ↑ 粗	調査票情報, 匿名化データ
統計メソデータ (メソデータ)	集約的, 非個別的		データキューブ
統計マクロデータ (マクロデータ)			統計表情報

メソデータはマクロデータに対して相対的に決まるものである。統計調査ではさまざまな統計が作成される。本稿では、調査計画で当初公表することにしていた統計をマクロデータ、それを作成する中間段階で得られる

統計をメソデータとする。しかし、ある統計調査で公表される複数の統計表間の関係を考えると、ある粒度の統計がマクロデータでありメソデータでもあるという二重性を持つことがある。たとえば、都道府県レベルの統計表と全国レベルの統計表が公表される場合、都道府県別表にある統計は二重性を有するデータである。ある統計調査において絶対的なマクロレベルのデータを決めるのは現実には困難である。また、メソデータの粒度を細かくしていくとマイクロデータと変わらない粒度のデータが出現するようになる。メソデータを提供する場合、このような疑似マイクロ性が生じないような粒度を探索的に見出すことで、この問題は実務的には解消可能である。

2.2 統計メタデータの種類と統計メタデータ標準を巡る動向

欧米における統計メタデータの整備や統計メタデータを用いた統計情報システムの構築に関する研究は、1990年代から2000年代にかけて盛んになった。90年代半ば以降になると、統計メタデータに関する用語集やガイドラインが作られている。近年では、統計業務過程を一般化した汎化統計業務過程モデル (GSBPM: Generic Statistical Business Process Model) 及び統計メタデータ概念モデルである汎化統計情報モデル (GSIM: Generic Statistical Information Model) が提案されている。

これらの動きの中で、統計メタデータに対する認識は、当初の簡明で抽象的なものからより広範で具体性のあるものに変化している。たとえば、Dippo and Sundgren (2000) は、統計メタデータを①統計データを解釈、理解、分析するのを助けるもの、②統計データを識別し、その所在を見つけ、検索するのを助けるもの、③統計調査の設計・企画プロセス及び実施プロセスに関して記述及びフィードバックに使われるものの3つに分類している。③はパラデータ (paradata) とも呼ばれて

いる (Radermacher et al. (2009)。ただし、この語自体は Couper (1998) が作ったとされる)。パラデータには、たとえば調査票の回収数(率)、項目補完率などがある。また、Radermacher et al. (2009) は、①構造的メタデータ (structural metadata) と②参照メタデータ (reference metadata) の2つに分類している。前者は統計データを識別、形式的に記述、検索するのに利用するもの、後者は統計データの意味的視点による内容と品質を記述するものである。現在では上記の例のように、統計メタデータには統計データと直接的に結びつくものと必ずしも直接的に結びつくとは限らないものがあると考えられている。

統計メタデータの実装レベルの標準として代表的なものにマイクロデータ向けの DDI (Data Documentation Initiative) と集計データ向けの SDMX (Statistical Data and Metadata eXchange) の2つがある¹³⁾。DDIでは、従来の標準 (DDI Codebook) とは別に、統計業務過程を対象にして統計データのライフサイクルを記述する DDI Lifecycle が作成されている。DDI codebook は、CESSDA (Consortium of European Social Science Data Archives)、CRDCN (Canadian Research Data Center Network) などで利用されている。また、DDI Lifecycle は、フランス、オランダなどの統計局で利用されている。

一方、SDMXでは、EurostatがSDMXを拡張してより多くのデータ品質に関する情報を含むようにした ESMS (Euro SDMX Metadata Structure) を作成し、メタデータ標準として採用している。近年では、GSIMとGSBPMを概念モデルとして、DDIとSDMXといった既存の標準間の相互運用可能性を高めたり、統合化を進めたりしていこうという動きが出てきている¹³⁾が、管見の限りでは統計業務過程全体を通して包括的、体系的に記述可能なメタデータ標準の実現には至っていないようである。

3. 統計業務過程と統計情報

旧統計法 (昭和22 (1947) 年法) 下の統計業務過程は、調査票と統計表の一次利用を前提としており、調査票の二次利用は例外的、附随的な業務であった。法制度上、調査で使用される調査票及び作成される統計表の様式と種類は、調査設計者が調査設計時に決定しておかねばならないため固定的であり、事後的な追加や変更は認められなかった。したがって、統計業務過程の処理形態は定型的である。一方、新統計法では二次の利用制度が創設されたことから、統計業務過程は初めから二次利用を前提とする業務過程、すなわち統計情報の保存と提供の2つの業務過程、を組み入れたものに変化することになる。統計表情報の二次利用のニーズはもともと不定期に発生するものである。また、調査票情報の二次利用のニーズは不定期に発生するものであり、作成される統計表の様式と種類は統計利用者が自由に決めることも可能である。さらにこれらのニーズは内容も求められる提供形態も多様であるため、二次利用のための提供業務は一般に非定型的である。

本節では、旧統計法下の統計業務過程と新統計法下の統計業務過程を区別するため、前者を「一次利用型統計業務過程¹⁵⁾」、後者を「二次利用型統計業務過程」と呼ぶことにする (図1参照)。一次利用型統計業務過程は、法制度の変化に関わらず基幹的な役割を担っている統計業務過程である。二次利用型統計業務過程の下で保存と提供の対象となる統計情報は、一次利用型統計業務過程全般に由来するものである。以下、単に「統計情報」と言った場合は一次利用型統計業務過程に由来する統計情報のことを指すものとし、調査票、調査票情報、統計表情報及び統計表を「狭義の統計情報」と呼ぶことにする。統計情報の保存と提供は、一次利用型統計業務過程が終了した後も継続して機能していくことになる¹⁶⁾。

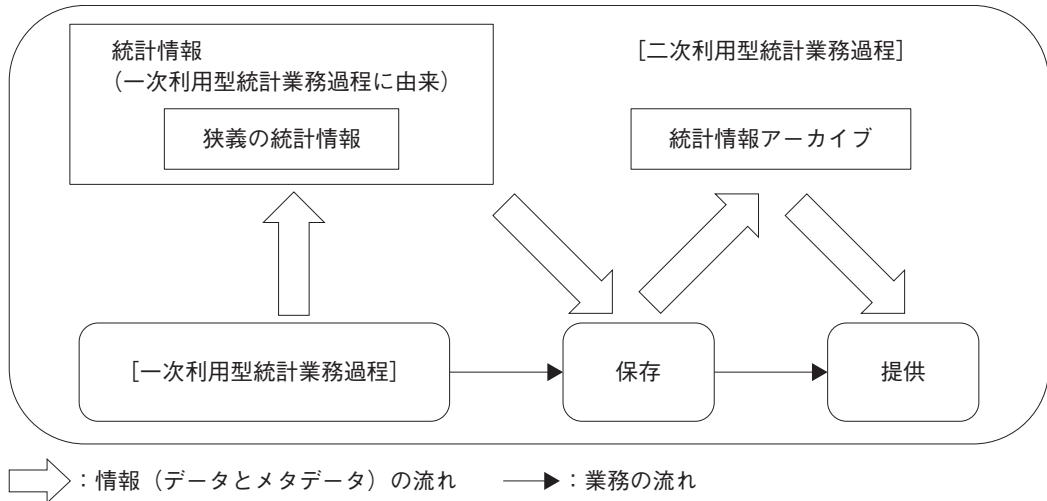


図1 二次利用型統計業務過程の概念図

実世界では調査対象が時間とともに変化するため、一度実施した統計調査とまったく同一内容の統計調査を再現することはできない。特に、調査票上の記述内容は、一部でも失われてしまうと復元ができないものなので、統計情報アーカイブの中では最も重要な情報である。統計情報アーカイブの中で狭義の統計情報は再現性を具えている必要がある。ここでいう再現性とは、保存している調査票情報又は調査票を用いて、一次利用型統計業務過程と同じ統計的方法、同じ処理手順を施すと、保存している統計表情報又は統計表と同じ成果物が得られるという性質とする。再現性は、保存している調査票情報、処理手順などの正当性を保証するものである。また、統計情報は完全性、すなわち保存している統計情報のみを用いて再現性が実現できるという性質を具えている必要がある。完全性は、再現性を実現するのに必要な情報が、保存している統計情報の範囲で充足していることを保証するものである。二次利用に当たり再現性と完全性は重要な性質と言える。保存している統計情報が再現性と完全性を備えていることにより、再集計に必要なデータがないため過去の集計結果を再現できないという事態は避けら

れることになる。過去の集計結果をいつでも再現できることで、統計作成者だけでなく統計利用者も、統計作成の業務過程と生成物の正当性を評価することが可能となる。

統計情報アーカイブの主要な関係者として統計利用者、統計作成者及びアーカイブ管理者(以下、「アーカイブ関係者」)が考えられる。以下では再現性及び完全性について、これらのアーカイブ関係者との関係の観点から見てみることにする。現行の二次の利用制度の下では、脚注15で述べているように調査票情報の二次利用で提供対象となるのは、エディティング済みのクリーンな調査票情報であるから、統計情報アーカイブから調査票情報の二次利用者に提供するのもクリーンな調査票情報となる。したがって、統計利用者にとっての再現性は、この調査票情報から統計表情報又は統計表までの範囲での実現ということになる。一方、統計作成者とアーカイブ管理者はエディティング前の調査票情報や調査票を利用できるので、再現性はこれらの情報から統計表情報又は統計表までの範囲での実現ということになる。特にアーカイブ管理者にとっての再現性は、保存している統計情報に誤りや欠落、記述のあいまいさといった

ことがないかを検証する際に不可欠な性質であると言えよう。一方、完全性はアーカイブ関係者が第1節後半で述べた統計情報利用時のメタデータ参照における不便さを解消するために必要な性質である。

4. 統計情報アーカイブとデータウェアハウス

統計表は、調査報告書の形で図書館に保存され、利用されるという形態が、長い間一般的であった。図書館は、現在も統計表のアーカイブ施設として、あるいはアーカイブ組織としての役割を担っていると考えるもよいだろう。統計審議会(1985)の提言後、情報通信技術の進展に伴い、現在では統計表(調査報告書掲載表だけでなく非掲載表も含む)は、紙媒体の印刷上の制約を受けない統計表情報として保存され、インターネットを通じて提供されるようになってきている。

一方、調査票は、伝統的な視点では「統計材料¹⁷⁾」として位置づけられており、成果物である統計表ができてしまえば不要と認識されていた。諸外国におけるデータ・アーカイブ組織の成立・発展は、社会調査や統計調査のマイクロデータの二次利用に対する社会的要請と密接に関係して進んできた¹⁸⁾。日本では、新統計法施行後、公的統計の基本的な計画を定めた統計委員会(2009)及び同(2014)の中で、各種統計調査の調査票情報の蓄積を政府として推進していくものとされている。分散型の統計機構をとる日本では、統計調査を所管する府省が第一義的に調査票情報の保存を行っている。総務省政策統括官(統計基準担当)(2019a)は、各府省が統計調査によって収集した調査票情報を国民の共有財産として将来にわたり利活用可能とするための調査票情報等の管理に関する指針を示している。その中では、符号表とレイアウトフォームに加えて調査票情報を利用するのに必要となるドキュメント類を調査票情報と併せて保存する

こととされている。

調査票の様式は統計調査によって多様であるし、統計調査によっては複数種類の調査票が用いられることがある。また、統計表の様式は一つの統計調査の中にあっても複雑・多様である。さらに、統計編成過程の情報システム設計者が設計する調査票情報と統計表情報も多様な様式で表現される。既存の統計表を様式と共に標準化して、データベース(以下、「DB」)を構築しようとする試み¹⁹⁾はあったものの、公的統計全体の範囲では実現していない。調査項目の標準化に関する研究会(2006)の検討結果に基づき、調査票情報や匿名データは政府標準レイアウト記述や符号表によって記述することが決められている。しかし、政府標準レイアウト記述では個別統計調査に由来する様式部分まで標準化されているわけではない。

二次利用型統計業務過程で統計情報をアーカイブに保存する場合、狭義の統計情報の多様性、複雑性を解決しておく必要があるが、それには発想の転換が必要であろう。第1節で述べたように狭義の統計情報は、情報本体と様式が一体の形で表現されている。様式は情報本体の見せ方を与える上で必要なものである(小林(2019)は統計表の見せ方を「表現形」と呼んでいる。この概念は狭義の統計情報に拡張できる)。様式が既定されていると、調査票の記入、統計表の集計などといった狭義の統計情報を作成する作業の操作性は向上する。しかし、調査票の情報本体は収集する調査項目が既定されていれば様式に依存せずに収集し得るし、統計表の情報本体は目的とする統計の作成方法が既定されていれば様式が既定されていなくとも作成可能である。狭義の統計情報が見せる多様性、複雑性は様式に由来するものであって、情報本体に由来するものではない。狭義の統計情報から様式という一種のメタデータを分離すること、すなわち様式独立にすることにより、情報本体を

後述する多次元DBの形で保存でき、保存時のDB構造の統一性と提供時の表現形の柔軟性を両立させることが可能となる。中間集計表の様式は通常、統計表情報の様式を想定して定めるので、上述の考え方は統計メソデータにも適用できる。DB構造の統一性はアーカイブされた様式独立な狭義の統計情報の管理容易性を高める。さらに、どの統計調査も保存時のDB構造を同一のモデルで表現できることから、公的統計全体の統計情報アーカイブ構築を進めることが期待できる。また、表現形の柔軟性は利用者自身が希望する様式で統計情報を利用することができるという利点がある。

図2は、個別統計調査の統計情報アーカイブの構造と利用に関して図示したものである。図中の両矢印は、両端に示すもの間に参照、生成などの関係があることを表す。統計情報アーカイブは、相互に関係する統計データ・アーカイブと統計メタデータ・アーカイブから構成される。統計データは、情報本体を構成する項目のうち量的項目に該当するものである。たとえば、様式独立な統計マイクロデータは、調査票情報の量的項目に該当するものである。統計メタデータには、統計データに直接的に結びつくものと間接的に結びつくものがある。図2では、統計データと直接的に結びつく統計メタデータのうち秘匿方法と様

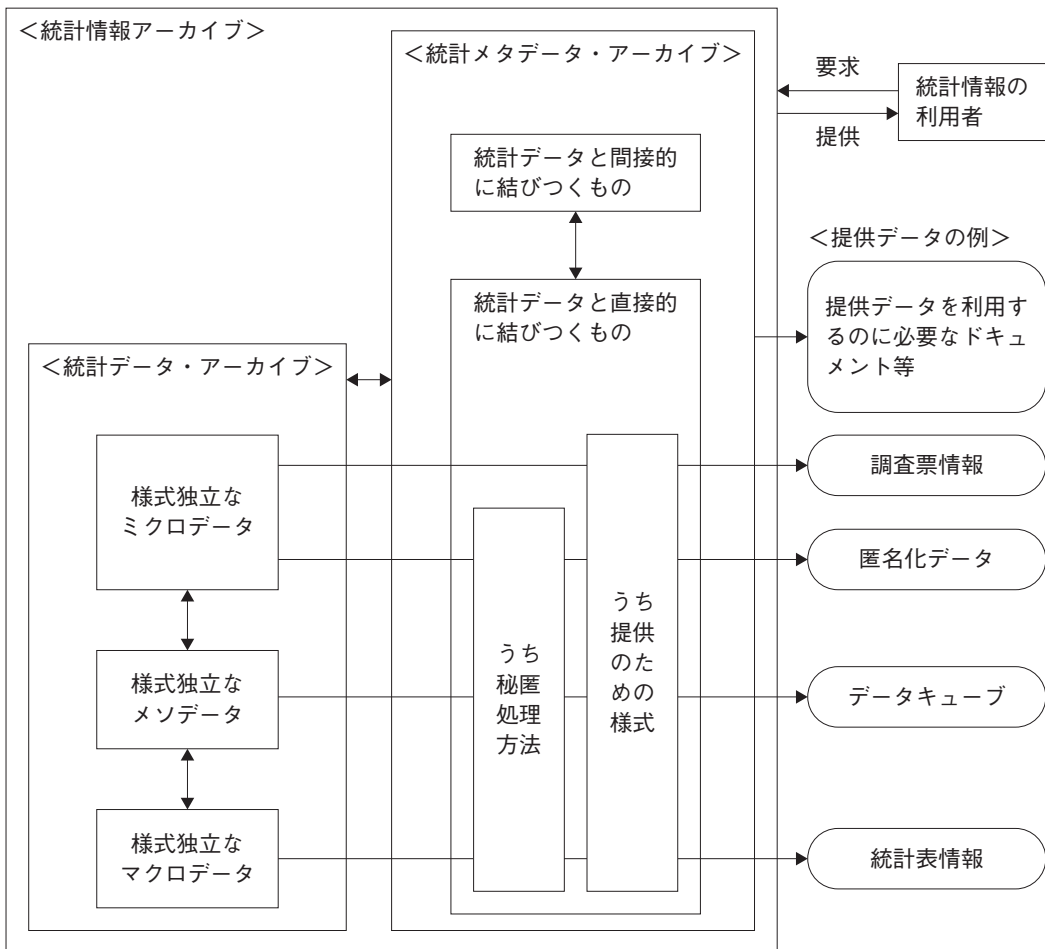


図2 統計情報アーカイブの構造と利用の概念図

式については特に明示してある。二次利用者の要求があると、統計情報提供者は必要に応じて統計データの秘匿処理を行った上で、調査設計時の様式又は二次利用者が指示する様式に従って、統計データ及び関係する統計メタデータを選択して配置した形態で統計情報を提供することになる。また、併せて提供データを利用する上で必要となるメタデータを提供する。

統計情報は一次利用型統計業務過程の進行に沿ってアーカイブに格納していくのが望ましい。特に調査票情報から統計表情報までを順次アーカイブに格納する情報システムが実現すれば、図2の統計情報の利用者には統計作成者も含むことになる。ただし統計的開示管理の観点から、利用者の立場によってアクセス可能な統計メタデータ及び/又は統計データの範囲には制約を設ける必要がある。

統計情報アーカイブを実現するには、DB技術が重要な要素技術となる。1980年代にDB技術が発展し、DBの設計、構築のシステム開発において、メタデータの有用性が認識されるようになった²⁰⁾。1990年代に入ると、企業活動の日常で発生する個別データとそれを集約したデータを多次元DBとして蓄積するデータウェアハウス(Data Warehouse。以下、「DWH」)とオンライン分析処理(OnLine Analytical Processing。以下、「OLAP」)ツールが、企業の経営意思決定支援システムとして登場した²¹⁾。多次元DBは、検索や分析の対象となる数値をセルに、検索や分析に用いる複数の属性を配列の軸にした多次元配列で実現される。統計データは多次元DBのセルに対応し、統計メタデータは多次元DBの軸になる複数の属性に対応するものと考えてよい。OLAPとは、経営意思決定のように不定期に発生しかつ非定型な処理を要するようなエンドユーザーの要求を実現するため、蓄積されているデータを多面的な視点で検索・分析し、結果を迅速に提供する処理をオンラインで行

うものである。UNSC and UNECE (1999)は、政府統計機関のクリアリング機能を実現する統計情報システムの一つとしてDWHを取り上げている。OLAPの基礎となる多次元DBは、データキューブとも呼ばれ、統計分野では統計情報提供の一形態として利用されている。

図2の右側には、統計情報アーカイブの利用と提供データの例を示している。一般的な流れとして、利用者は、利用したい統計データ又は統計メタデータを要求し、アーカイブから提供を受ける。本稿で提案する統計情報アーカイブは、第1節で述べたような狭義の統計情報の利用上の不便さを解消するものであり、提供されるのは基本的に電子的な情報になる。統計データの利用要求と提供は、技術的にはオンラインで処理可能なものの、データの特徴(表1参照)によって法制度上の扱いが異なるため、OLAPのような処理が一律に利用可能となるわけではない。一方、統計メタデータのみ利用要求と提供は、技術的にはオンラインで処理可能である。

DWHは、通常、時系列でデータを保存し、一度保存したデータは更新しないので、アーカイブとしての性質を満たすものと言える。DWHは統計データと統計メタデータの一体的な保存を実現する技術と言えよう。DWH構築の際には、統計情報の完全性を満たすために、狭義の統計情報を包含する統計情報全体をDWHとすることが必要である。

5. 統計メタデータ・アーカイブの展開可能性

5.1 統合化による展開可能性

調査項目の標準化に関する研究会(2006)では、基幹統計調査(検討当時は指定統計調査)の調査項目について定義の標準化の研究が行われ、その成果は「政府統計の総合窓口(e-Stat)」で見ることができる。しかしながら、同研究会で問題提起された統計表の表章項目の定義の標準化は、今日に至っても進ん

でない。

個別統計調査レベルの統計メタデータ・アーカイブを整備する際に考慮すべきことの一つは、調査票情報と統計表情報の各々が持つ項目間の相互参照性を確立することである。統計表の分類項目と集計項目は、もともと調査項目等に由来するものである。項目間の相互参照性の確立は、後述するトレーサビリティの確立に寄与し得るものである。項目間の相互参照性が確立していない状態であると、調査報告書を参照しなくてはならなくなる可能性があり、統計情報の完全性を満足しないことになってしまう。また、DDIとSDMXのように調査票情報と統計表情報で別々に統計メタデータを整備するのは、個別の統計調査の統計メタデータ・アーカイブ内で統計メタデータの重複や脱漏を引き起こす可能性が高く、望ましくない。調査票情報と統計表情報の項目間の相互参照性の確立は、公的統計全体を通じた統合的な統計メタデータ・アーカイブ（以下、「統合統計メタデータ・アーカイブ」）構築の際の共通語彙基盤の整備につながるものでもある。

統計メタデータは、公的統計全体に共通する大域的なメタデータ、個別の統計調査に固有の局所的メタデータ及び両者の中間的な半大域的なメタデータに分類できる（図3）。この分類に従って統計メタデータ・アーカイブも、個別統計調査レベルから公的統計全体レベルまで、階層的に構築していくことが考えられる。

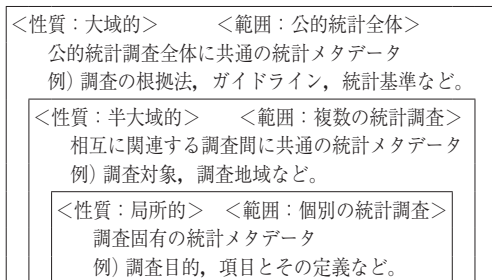


図3 統計メタデータの階層

統計調査ごとの統計メタデータが集積され、統合統計メタデータ・アーカイブが構築されることで、さらに可能となり得ることとして、下記の点が考えられる。

1点目は、公的統計における用語の標準化の実現である。個別統計調査の統計メタデータ・アーカイブの統合化を進めることにより、統計メタデータの調査横断的な相互参照性や共通化のための整合性の確立を進めることが可能となる。共通語彙基盤の構築における統計メタデータやその記述で使われる用語の標準化は、調査設計者が統計メタデータを二次利用して新規統計調査を設計する際や統計利用者が複数の統計調査の調査結果を比較する際に有用である。

2点目は、トレーサビリティの実現である。統計データのトレーサビリティとは、ある統計データに着目したときに当該データの作成元に該当する統計データは何か（トレースバック）、また当該データの利用先に該当する統計データは何か（トレースフォワード）、及びその統計データはどのような統計的方法により作成されたかといったことを明らかにし得ることである。

二次統計では、トレーサビリティは統計の品質保証、品質評価において必要なものであり、統計作成者の業務過程の透明性、成果物の適正性を示すものと言える。また、統計利用者にとっても統計の理解の向上につながるとともに、利用者自身が新たな加工統計を作成する際の参考になり得る。これは、二次統計に関する統計メタデータを包含する統合統計メタデータ・アーカイブの構築により初めて可能となる。一次統計においても、トレーサビリティは重要であって、頻度表や数量表の集計値が、どの個別データに由来するのかわかること²²⁾により、集計方法の誤りなどのチェックに利用できるなど、統計編成過程の品質評価に役立つ。もちろん、個別データにまで遡るトレーサビリティ情報へのアクセス

は、統計的開示管理の観点から統計作成者の範囲に限るのは言うまでもない。

また、統計メタデータのトレーサビリティも考えることができる。これはたとえば、ある導出項目の導出元になる調査項目は何か、その調査項目を利用した導出項目は何か、その導出項目はどのような導出方法で作成されたのかといったことを明らかにし得ることである。統計メタデータのトレーサビリティは、統計メタデータの間接的な関係を含めて考えるのがよいので、後述するオントロジーの確立が必要となる。

3点目は、統計情報アーカイブの情報セキュリティとオープンデータ化の両立の実現である。メタデータだけをドキュメントとして調査票情報と別に保存する方法では、情報の更新があった際に相互の同期を確保することが困難であるし、誤りも生じやすい。その点、DWHはデータとメタデータ相互のリンクは常に同期し、整合するようになっているという利点がある。個別統計調査の統計情報アーカイブにおいて、マイクロレベルの統計データとその統計メタデータとのリンケージ情報は秘匿して、承認された利用者のみがこのリンケージ情報を含めた統計マイクロデータへのアクセスを許可される仕組みを作ることは技術的に可能である。また、統合統計メタデータ・アーカイブにある調査項目等とその定義記述を、調査横断的、網羅的に検索し、比較できるような情報システムを構築することは技術的に可能である。これにより、統計マイクロデータの秘匿は保証しつつ、アーカイブされた統計メタデータを広く利用者にオープンデータとして公開することが可能となるであろう²³⁾。

4点目は、データリンケージのキー項目候補を調査横断的に探索することの実現である。データリンケージにおけるキー項目の候補が統計データにアクセスする前に探索可能となると、調査票情報の二次的利用の申出前にリ

ンケージ方法を検討する上で有用である。特に、リンケージデータによる実証研究では、分析のためのデータ調製作業の時間を短縮することにつながるだろう。複数の調査票情報を組み合わせたリンケージデータは、単独の統計調査の調査票情報に比べ、はるかに高い開示リスクを持っていると予想できる。統計データのリンケージキーファイルの管理と利用は厳格に行う必要がある。それにもかかわらず、リンケージデータの作成は新たに調査を行わないで有用なデータを得る可能性を持つ有効な方法と言える。データリンケージに関する先行研究において作成されたリンケージキーファイルとその作成方法を、統合統計メタデータ・アーカイブに保存²⁴⁾し、一種の共有財とすることでできるならば、リンケージ項目の探索の際に役立つだろう。そのためには、制度面の検討、リンケージキーファイル作成者の積極的貢献などが不可欠である。

5.2 オントロジー構築による展開可能性

統計情報は、実世界をどのように観測し、その観測によりとらえた実世界がどのような姿に見えるかを表すものと考えてもよいだろう。統計メタデータが表すものとは対象物を理解するのに必要な概念である。統計情報アーカイブを構築する際に筆者が最も重要と考えるのは、統計メタデータが表す概念や概念間の関係の構造体系（オントロジー）を考え、確立することである²⁵⁾。

用語の標準化は、用語の背後にある概念の標準化と密接に関係していると言ってよいであろう。現実には、本来同じである概念が調査ごとに異なる用語で呼ばれる場合や同じ用語が異なる概念を指している場合が生じ得る。既存の統計調査にはすでに長い期間用いられている用語があり、それらをすべて標準化して共通語彙基盤を実現するのは必ずしも実際的とは言い難い。この問題を解決するには、公的統計分野のオントロジーの構築が有効で

あろう。統計メタデータやその記述に使われる用語のオントロジーを整備することによって、調査内や調査間で概念の統一性、概念相互の整合性を持った統合統計メタデータ・アーカイブの整備が期待できる。公的統計全体を対象として整備されるオントロジーは、概念の標準化を実現し得るものと言えよう。

オントロジーの構築によって可能となることの一つとして、統計メタデータが表す概念に基づく統計データ検索の実現が考えられる。統計データの二次利用ではまず、どのようなデータがどこにあるかを知ることが必要である。データの内容、種類、所在などを表すものはメタデータであるから、統計データの検索は本質的には統計データに結びつく統計メタデータの検索に帰着すると考えられる。通常のキーワード検索では、人が統計メタデータや用語の意味を解釈し、適切なキーワードを考えて統計メタデータ・アーカイブを検索し、統計データを得ることになる。オントロジーが構築されると、検索に用いる統計メタデータや用語に結びつく概念間の関係を用いて、統合統計メタデータ・アーカイブから個別統計調査の統計メタデータ・アーカイブに至り、統計メタデータと結び付いている統計データを得るといった検索の実現が期待できる。

6. 今後の展望

統計データは、統計メタデータと関係付けられていなければ単なる抽象的な数に過ぎない。それゆえ、統計メタデータは、統計作成者と統計利用者の双方にとって、統計データの意味を理解して利用する上で重要な役割を担う不可欠な存在であると言えよう。さらに言えば、統合統計メタデータ・アーカイブは、公的統計の体系化、標準化だけでなく、統計調査の品質管理の視点からも重要なものとなる。

統計メタデータの収集、保存は統計データ

をアーカイブに収録する時点で行うのでは遅く、情報の散逸なく統計メタデータの収集を可能とするためには、統計調査の最初の段階から統計業務過程の進行に沿って、統計データと一体的に収集できるような仕組みを各業務過程に組み込んでおくことが必要である。

いったん作成された統計メタデータをアーカイブに収録するために組み替えたり、不足する情報を補ったりする作業をアーカイブ組織だけに委ねるべきではない。このような作業は労働集約的で時間がかかるものである。分散型の統計機構をとる日本では、統計調査を所管する府省が分担する部分は大きいと考えられるので、各府省が作業を進める上で、収集すべき統計メタデータ、統計メタデータが表す概念などに関する記述規約やガイドラインといった統一的な基準の整備が必要となる。

統合統計メタデータ・アーカイブの整備を行うには、個別統計調査の統計メタデータ・アーカイブを先に構築し、統合統計メタデータ・アーカイブの構築に向かうボトムアップ型アプローチとその逆のトップダウン型アプローチを双方向で進めていくのが実際のだろう。公的統計における統計メタデータ・アーカイブの構築は政府全体として取り組むべき事業と言える。それは、社会の情報基盤として、統計情報がより有効に利用される上で必要な事業だからである。

本稿で行ってきた議論は、行政記録情報を基に作成する業務統計のアーカイブを考える際にも適用できる。ただし、行政記録をコンピューター処理可能な行政記録情報とする取り組みの推進が必要である。また、行政記録情報自体を二次利用のためにアーカイブ化することが可能なのかについては法制度的な検討が必要である。

筆者が考える統計情報アーカイブは、DWHで保存される範囲より広いものである。DWHは、提供を考慮して、文字記号列として表現

される電子化された情報の範囲のみをカバーするものである²⁶⁾。しかし、少なくとも紙媒体の調査票と調査報告書の掲載統計表は実世界の物理的存在であり、そこに記録されている情報は電子化された情報以上の情報量を含む可能性がある。たとえば、調査票の自由記入欄や回答欄外に書き込まれた回答に関する記述、調査報告書の構成や統計表の配列順など、電子化された情報の範囲では捨象されてしまっている情報である。調査票や統計表をマイクロフィルムなどの光学的記録又はスキャナー入力したイメージデータとして保存することも統計情報アーカイブの範囲として

考えてよい。たとえば、符号化される前の産業や職業に関する自由記入は、分類基準の見直しや再構成後の基準による再集計結果を用いた実証研究といった新たな研究分野につながるだろう。調査票は調査票情報と異なり一定年数保存後に滅却されるし、統計表を収録した調査報告書も物理的なスペースが不足すればごく少数（少なくとも1セットづつは国会図書館と調査主体の元で）保存されるだけである²⁷⁾。このことは、誰もが、いつでも、どこでも、望みの統計情報にアクセス可能とする上で、将来にわたり考えていくべき課題といえよう。

謝辞

査読者からは多くの有意義なコメントをいただき、本稿の改善に資するところ大であった。ここに記して深く感謝の意を表したい。

注

- 1) 統計業務過程のモデルには必ずしも統一的な認識の合意があるわけではない。実務では、大きく分けて調査設計、実地調査、集計、公表の4業務過程からなるとの認識が比較的普及していると言える。各業務過程は、さらに下位の業務過程から構成されている。
- 2) 一般に、当初定められた利用目的（公的統計では調査の目的の中で記述）の範囲での利用は「一次利用」、一次利用以外利用（当初定められた利用目的以外の目的での利用）は「二次利用」と呼ばれる。
- 3) 本稿では公的統計の中でも中核をなす調査統計の場合を取り上げて議論を展開している。業務統計や加工統計の場合も、統計材料が行政記録や一次統計であるという違いがあるものの、統計を作成する業務過程は調査統計の場合とほぼ同様である。
- 4) アーカイブの保存対象となるものには、存在のありようによって、実世界の具体的存在物と情報がある（溝口（2012）は、実世界の実在物は具体物、抽象物及び準抽象物に分類でき、情報は準抽象物に含まれるとしている）。報告者が記述した紙媒体の調査票などは具体的存在物である。一方、情報は、通常、電子化された形態で存在する。本稿ではオントロジックの観点から踏まえて調査票と調査票情報、統計表と統計表情報のそれぞれの前者と後者を分けて考えている。調査票情報（に記録されている内容）は、紙媒体の記述済み調査票（に記録されている内容）を情報システムで扱えるように変換（一般的には入力機器を用いて電子的な媒体に入力）したもの（実際には集計で用いる付加項目、たとえば都道府県番号、標本調査のウェイトなども含まれる）である。オンライン調査では記述済み調査票から調査票情報への変換過程が省略され最初から記述済みの調査票情報が作成されていると考えられる。PDF形式の調査票もオンライン調査票も調査票と呼んでいるが、電子的な形態で存在しているので本稿の定義では調査票情報に当たる。情報を見るには何らかの表現媒体（紙媒体や電子的媒体など）が必要であるが、紙媒体に記録されている情報は、直接視認できるものである。ディスプレイに表示されている情報は、コンピューター内にある情報をディスプレイという機器の画面を通してはじめて視認できるものである。しかし、回答時には調査項目や回答記述は回答者自身が視

- 認できている点は紙媒体の調査票の場合と変わらないと言えよう。
- 5) 統計表情報は、統計原表に相当するものと考えてもよいであろう。統計表情報を紙媒体の物理的な大きさに収めるためには表頭分割や表側分割をして印刷する必要がある。また、2つの統計表情報を結合して見かけ上1つの統計表とすることがある。それゆえ、統計原表と統計表では表番号が異なる場合が生ずる。紙媒体という物理的制約を受けない統計表情報の方が、表としての様式の自由度ははるかに大きい。
 - 6) 調査票と統計表は、情報本体と様式に加えて、これらを実世界で視認可能とする紙媒体から構成される。
 - 7) 小林(2019)によれば、様式独立にすると、調査票や調査票情報は質的項目と量的項目の組で表現できる。また、統計表や統計表情報はセルを量的項目、表頭などの分類項目を質的項目と考えることができ、やはり質的項目と量的項目の組で表現できることがわかる。
 - 8) 森(2008)はアーカイブに保存すべきメタデータの例としてエディティングの処理プログラムを挙げているが、エディティングの処理プログラムを含む一連の集計プログラムは調査票情報と統計表情報の関係をプログラム言語で一連の手続きとして記述した仕様であり、統計業務過程の中で情報相互の関係を表す情報の例と言える。
 - 9) 管見の限りでは、どの統計調査もエディティングに関する情報など一部の情報は調査報告書にも掲載されていない。また、調査員事務などの実務に関するマニュアルは、一般にはアクセスが難しい。
 - 10) たとえば、符号表からは調査票情報に「労働力状態」という項目とその項目値に「完全失業者」があることはわかって、「完全失業者」の定義を知ろうとすれば調査報告書を参照する必要がある。また、分類項目に「社会経済分類」という項目が使われている統計表情報を、各府省のホームページや総務省のe-Statで提供されている統計表情報の中に見つけたとき、その分類項目がどの調査項目に該当するのかあるいはどのような調査項目どうしを組合せて導出したのかを知ろうとすると、表とは別に調査報告書にある分類項目一覧や用語の解説を参照する必要がある。用語の解説が電子的な情報としてウェブ上に存在していても、統計表情報の分類項目などと関係付いていないなら、表中の数値を見ているときに別のファイルを探して開き、参照しなければならない手間がかかるという不便さがある。
 - 11) 統計審議会(1985)では統計調査結果、すなわち集計値の意味で使用されている。統計審議会(1995)、各府省統計主管部局長等会議(2003)では、集計結果データの意味と個票データの意味の両義が混在している。統計委員会(2009b)では個票データや匿名化データの意味で使用されている。
 - 12) 統計マイクロデータ(statistical microdata)とは「ある個別オブジェクト、すなわち統計単位、に関して収集された観測データ(an observation data)」と定義されており、一方、統計マクロデータ(statistical macrodata)とは「統計的方法論に従って統計マイクロデータをある目的/意図をもって集約すること(a purposeful aggregation)により得られた観測データ」と定義されている。
 - 13) DDIについては<http://www.ddialliance.org/what>、SDMXについては<http://sdmx.org>を参照。
 - 14) DDIとSDMXの比較を行っているものとして、Gregory and Heus(2007)、Gregory(2008)、Pellegriano and Grofils(2013)などがある。
 - 15) 本稿では、一次利用型統計業務過程のモデルを調査設計、統計材料収集、統計編成、分析・公表及び評価の5過程からなるものと考えている。記述済み調査票は統計編成の過程の中でクリーンな調査票情報となる。この調査票情報が二次利用の提供対象となる。なお、本稿で考えている「評価」過程は、統計業務過程全般を対象に包括的、体系的に一連の業務過程(プロセス)とその生成物(プロダクト)の品質評価を行うものである。
 - 16) 一次利用型統計業務過程の実行とは独立して統計情報アーカイブに情報を追加することがある。例としては公的統計全体に影響する標準分類の改定が該当する。このような処理は、二次利用者に影響を与える変更だが、不定期に発生することから非定型的な業務と言える。
 - 17) 「統計材料」という用語は、統計を作成するための材料という意味で少なくとも明治期には使用され出しており、調査票に限らず行政記録も含む概念と認識されていた(呉(1892);高橋(1906a, 1906b))。
 - 18) 諸外国のマイクロデータ提供の歴史的経緯、提供形態、法制度などに関するまとまった調査研究は、

松田ほか(2000)が嚆矢であろう。また、総務省が行ったデータ・アーカイブ組織に関する調査委託研究としては、総務省政策統括官(統計基準担当)(2007)及び三菱総合研究所(2011)がある。前記2資料の中では、社会調査を中心に収集・提供しているデータ・アーカイブ組織の事例として米国のICPSR(Inter-university Consortium for Political and Social Research)、イギリスのUKDA(UK Data Archive)、ドイツのZA(Central Archive for Empirical Social Research in Cologne)のほか、日本のSSJDA(Social Science Japan Data Archive)、JEDI(Japan Educational Data-archive Initiative)などが挙げられている。一方、政府統計では政府統計機関がデータ・アーカイブ組織の役割を担っているのが一般的である。なお、統計委員会(2009)以前と以後では政府文書で使われる表記が「データ・アーカイブ」から「統計データ・アーカイブ」に変わっている。

- 19) 佐藤(1988)、佐藤(1995)は、メタデータを用いて統計表(佐藤(1995)は「要約データ」と呼んでいる)のDB設計・構築の実証研究を行っている。これらは、既存の統計表に関して様式を含めた形態でDBを構築しようとしたもので、現実の統計表様式の多様性、複雑性を解消する工夫がなされているものの、管見の限りでは政府統計分野の中で同種の研究が広がることはなかった。
- 20) DBにおけるメタデータの集りはデータディクショナリ(以下、「DD/D」と呼ばれる。DD/Dについては、Leong-Hong and Plagman(1982)、椿(1982)を参照。
- 21) DWH及び多次元DBについてはKimball(1996)、Inmon(2002)が、またOLAPについてはCodd et al.(1993)が論じている。日本では主に情報処理分野で研究が行われた(たとえば、豊島(1996)、田中ほか(1996)、田中(1997)、石井ほか(1998)など)。池田(2006)はDWH、OLAP及びOLTP(On-Line Transaction Processing)を用いた民間企業のデータ蓄積及び利用について論じている。一方、本稿ではDWHと多次元DBを用いた公的統計の統計情報アーカイブについて論じている。
- 22) このような機能はOLAPの機能の一つとして確立されており、ドリルスルーと呼ばれている。たとえば、Excel2016のピボットテーブルには、この機能が実装されている。
- 23) 無制限に公開するというのではない。たとえば、エディティング規則やそれを適用したときのデータの変遷情報は、統計作成者と統計利用者に対して、公開の可否や公開の程度と範囲に制約があるのが現実である。また、提供時に適用する秘匿方法は公開の対象に含めないのが諸外国でも一般的である。
- 24) 現行の二次的利用制度にあっても、作成したリンケージキーファイルを一定期間保存しておくことは可能である(総務省政策統括官(統計基準担当)(2019b))。しかし、データリンケージの方法論的研究やリンケージデータによる実証研究のための共有財として長期にわたり保存しておくことはできない。
- 25) オントロジーの基礎的知識、理論及び適用事例などは、溝口(2005)、溝口(2012)を参照のこと。來村(2012)は医療など各種分野への応用事例が豊富である。
- 26) 本稿で「電子化された情報」、「電子的な情報」と呼んでいるのは、電子的記録媒体上に存在する情報のことである。調査票情報や統計表情報はこれに該当する。文字には記号としての文字と図形としての文字の2つの側面があり、電子化された情報には文字記号列としての情報と文字図形としての情報の2種類がある。集計プログラムやDWHが扱う調査票情報などは電子化された文字記号列としての情報である。一方、報告者の記述済み調査票などをマイクロフィルムやイメージデータとしたものは文字図形としての情報である。本稿では、利用者に提供する電子化された情報の範囲を文字記号列としての情報に限定して議論を展開しているが、それは既存の一次利用型統計業務過程の情報システムが文字記号列としての情報を扱ってきたこと、及び統計データ・アーカイブと統計メタデータ・アーカイブの構築上の情報システム技術的な理由による。
- 27) 電子的記録媒体の保存期間は技術的に半永久的とは言えないこと、及び記録された情報を読み取る機器がないと内容を知ることができないため機器の技術的寿命に制約されることが、紙媒体が残っている理由と考えられる。それゆえ、アーカイブに保存しておくものは、電子化された情報(文字記号列として表現されたものと文字図形として表現されたもの)だけでなく紙媒体による記録も必要と言えよう。

参考文献

- 池田伸 (2006)「民間企業におけるデータの蓄積と利用 — マーケティングリサーチ, データマイニング, 統計 —」『社会科学としての統計学 第4集』第4章, pp.45-57, 産業統計研究社
- 石井義興・豊島一政・木村哲 (1998)「データウェアハウス/データマート」『電子情報通信学会誌』第81巻第10号, pp.1034-1041, 電子情報通信学会
- 各府省統計主管部局長等会議 (2003)『統計行政の新たな展開方向』
- 來村徳信 (2012)『オントロジーの普及と応用』人工知能学会編, オーム社
- 呉文聰 (1892)「統計ノ本体及問題」『統計集誌』第129号, pp.168-170
- 国立国語研究所 (2003)「第2回 「外来語」言い換え提案」国立国語研究所「外来語」委員会資料, https://www2.ninjal.ac.jp/gairaigo/Teian2/iikae_teian2.pdf (最終アクセス: 2020年9月11日)
- 小林良行 (2012)「公的統計マイクロデータ提供の現状と展望 — 一橋大学での取り組みをもとに」『日本統計学会誌』第41巻第2号, pp.401-420, 日本統計学会
- 小林良行 (2019)「統計表の構造とIPF法による教育用擬似個別データの作成方法」『公的統計情報 — その利活用と展望』坂田幸繁編著 中央大学経済研究所研究叢書 75, pp.23-38, 中央大学出版部
- 佐藤英人 (1988)『統計データベースの設計と開発 — データモデルと知識ベースの応用 —』穂鷹良介監修, 第1版, オーム社
- 佐藤英人 (1995)「要約データの基礎概念とデータベース内での推論 — 世界貿易統計データベースを例として —」『世界貿易データベースシステムの整備と利用』アジア経済研究所統計資料シリーズ 67, pp.95-106, 日本貿易振興機構アジア経済研究所
- 総務省政策統括官 (統計基準担当) (2007)「諸外国の統計データの二次的利用の状況」, 第1回統計データの二次的利用促進に関する研究会 (平成19年10月22日開催) 資料7
- 総務省政策統括官 (統計基準担当) (2019a)『調査票情報等の管理及び情報漏えい等の対策に関するガイドライン』(平成21年2月6日総務省政策統括官 (統計基準担当) 決定. 平成31年4月19日改正)
- 総務省政策統括官 (統計基準担当) (2019b)『調査票情報の提供に関するガイドライン』(平成20年12月24日総務省政策統括官 (統計基準担当) 決定. 令和元年6月27日改正)
- 高橋二郎 (1906a)「技術統計論」『統計集誌』第308号, pp.487-496
- 高橋二郎 (1906b)「技術統計論(二)」『統計集誌』第309号, pp.541-552
- 田中聡 (1997)「データウェアハウスと多次元データベース」『情報処理』Vol. 38 No. 9, pp.745-750, 情報処理学会
- 田中聡・木村哲・豊島一政・石井義興 (1996)「多次元データベース入門〜オンライン分析処理を中心として〜」情報処理学会データベースシステム研究会報告103 (1996-DBS-110), pp.1-7
- 調査項目の標準化に関する研究会 (2006)『調査項目の標準化に向けて — 統計の標準化への実践的な第一歩 —』(平成18年11月15日「統計調査等業務の業務・システム最適化計画」(各府省情報化統括責任者 (CIO) 連絡会議決定, 2006年から継続中) に基づく決定)
- <https://www.stat.go.jp/info/guide/public/hyoujun/pdf/houkoku.pdf> (最終アクセス: 2020年9月11日)
- 椿正明 (1982)「データディクショナリ/ディレクトリ」『情報処理』Vol. 23 No. 10, pp.925-930, 情報処理学会
- 統計委員会 (2009)『公的統計の整備に関する基本的な計画』(平成21年3月13日閣議決定)
- 統計委員会 (2014)『公的統計の整備に関する基本的な計画』(平成26年3月25日閣議決定)
- 統計審議会 (1985)『統計行政の中・長期構想』(昭和60年10月25日諮問第207号の答申)
- 統計審議会 (1995)『統計行政の新中・長期構想』(平成7年3月10日諮問第242号の答申)
- 統計改革推進会議 (2017)『統計改革推進会議最終取りまとめ』
- 友安亮一 (1957)『統計表の表わし方』, 一粒社
- 豊島一政 (1996)「多次元データベースとRDB (OLAPの紹介)」『情報処理学会全国大会講演論文集』, 第52回平成8年前期(4), pp.157-158
- 松田芳郎・濱砂敬郎・森博美 (編著) (2000)『講座 ミクロ統計分析 1 統計調査制度とミクロ統計

- の開示], 日本評論社
- 溝口理一郎 (2005) 『オントロジー工学』人工知能学会編, オーム社
- 溝口理一郎 (2012) 『オントロジー工学の理論と実践』人工知能学会編, オーム社
- 三菱総合研究所 (2011). 『統計データ・アーカイブの整備に関する調査研究報告書』総務省委託研究
- 森博美 (2008) 「情報資産としての統計と政府統計データ・アーカイブ」『統計学』第94号, pp.15-25, 経済統計学会
- 山口幸三 (2019) 「改正された統計法と二次的利用の現状と課題」『公的統計情報 — その利活用と展望』坂田幸繁編著 中央大学経済研究所研究叢書 75, pp.3-21, 中央大学出版部
- 美添泰人 (2005) 「統計データの保存と再利用の体制」『統計』第56巻第6号, pp.32-37, 日本統計協会
- Codd, E.F., Codd, S.B. and Salley, C.T. (1993), *Providing OLAP (On-line Analytical Processing) to User-Analysts: An IT Mandate*, Codd & Date Inc., なおCodd & Associatesからも出版されている
- Couper, M. (1998), “Measuring Survey Quality in a CASIC Environment”, *Proceedings of the Survey Research Methods Section of the ASA at JSM1998*, pp.41-49
- Dippo, C.S. and Sundgren, B. (2000), “The Role of Metadata in Statistics”, *Proceedings of the Second International Conference on Establishment Surveys*, pp.909-918
- Gregory, A. (2008), “Status on the Mapping of Metadata Standards: ISO/IEC 11179, SDMX, and Others”, WP9, Joint UNECE/Eurostat/OECD Work Session on Statistical Metadata (METIS)
- Gregory, A. and Heus, P. (2007), “DDI and SDMX: Complementary, Not Competing, Standards”, Open Data Foundation
- Inmon, W.H. (2002), *Building the Data Warehouse (Third Edition)*, New York: John Wiley & Sons,
- Kimball, R. (1996), *Data Warehouse Toolkit*, John Wiley & Sons Inc., 藤本康秀監修 (1998) 『データウェアハウス・ツールキット』, 日経BP社
- Leong-Hong, B.W. and Plagman, B.K. (1982), *Data Dictionary/Directory Systems: Administration, Implementation and Usage*, John Wiley & Sons, Inc., 成田光彰訳, 穂鷹良介監訳 (1986). 『データディクショナリ/ディレクトリシステム』第1版, オーム社
- Pearce-Moses, R. (2005), *A Glossary of Archival and Records Terminology*, the Society of American Archivists, <http://files.archivists.org/pubs/free/SAA-Glossary-2005.pdf> (最終アクセス: 2020年9月11日)
- Pellegrino, M. and Grofils, D. (2013), “DDI-SDMX Integration and Implementation”, WP5, Joint UNECE/Eurostat/OECD Work Session on Statistical Metadata (METIS)
- Radermacher, W., Baigorri, A., Delcambre, D., Kloek, W. and Linden, H. (2009), “ANNEX: TYPES OF STATISTICAL INFORMATION”, *Terminology relating to the Implementation of the Vision on the Production Method of EU Statistics*, pp.22-26, Eurostat
- Sundgren, B. (1973), *An Infological Approach to Data Base*, Doctoral Thesis, University of Stockholm
- United Nations Statistical Commission (UNSC) and United Nations Economic Commission for Europe (UNECE) (1999), *Information Systems Architecture for National and International Statistical Offices Guidelines and Recommendations*, Conference of European Statisticians Statistical Standards and Studies-No. 51, United Nations
- UNSC and UNECE (2000), *Terminology on Statistical Metadata*, Conference of European Statisticians Statistical Standards and Studies-No. 53, United Nations

【Special Section: The 60th Anniversary of the Journal】
Special Topic B: Methodological Perspectives in the Creation and Release of Official Microdata

Potentiality of Statistical Metadata Archives in the Official Statistics of Japan

Yoshiyuki KOBAYASHI*

Abstract

This study proposes the construction of integrated statistical metadata archives referred to herein as Integrated SMA and the establishment of domain ontology in official statistics and indicates the potentiality of both.

The statistical information generated through the statistical business process (SBP) is defined as the information comprising both statistical metadata and data with the associated statistical metadata. This study examines an SBP model presupposing the secondary use of statistical information and the properties of the statistical information to be satisfied. Furthermore, it is pointed out herein that the data warehouse technology plays an essential role in constructing the statistical information archives of respective surveys. The statistical data and metadata archives constitute statistical information archives.

The statistical metadata in official statistics can be classified into three layers: global, semi-global, and local metadata layers. The standardization of the terminology on the official statistics and the traceability of statistical information are feasible when the Integrated SMA based on the three-layer structure is constructed. The establishment of ontology on the official statistics can standardize statistical metadata concepts.

Key Words

Statistical metadata, Statistical business process, Data warehouse, Traceability, Ontology

* Statistical Research and Training Institute, Ministry of Internal Affairs and Communications

高学歴雇用者の専攻と男女間賃金格差

武内真美子*

要旨

本稿では2004年から2017年のパネルデータを用いて、大卒以上の高学歴者の専攻と男女間賃金格差の関連について考察した。まず、男女別の賃金(年収)関数の推計結果から、人文科学をベースとした場合に、男性は医学・歯学、薬学、工学の順に賃金プレミアムが確認され、女性は理学、医学・歯学、工学、薬学の順に賃金プレミアムが確認できた。また、男女間賃金格差の要因分解の結果、特に工学分野の女性比率が低いことが格差要因方向に寄与している一方で、工学分野の評価(係数)の差は格差縮小方向に寄与していることが明らかになった。また同一専攻内の男女間賃金格差を考察するため、専攻別の推計における女性ダミーの効果を計測したところ、理学以外の専攻では女性ダミーの負の効果が確認され、その効果は人文科学、教育、農学、社会科学の順に大きいことが確認された。

キーワード

高学歴雇用者, 専攻, 男女間賃金格差

1. はじめに

Becker (1962), Becker (1980) の人的資源理論に基づけば、大学への進学は人的資本への投資行動であり、卒業後の賃金(年収もしくは所得)はその収益(便益)と考える。海外の多くの先行研究が、この理論に基づいて学歴や教育年数が賃金に与える影響を計測している (Vikesh, 2011; Diagne and Diene, 2011; Furno, 2014; Peet et al., 2015)。これらの研究が人的資本を量的側面から分析しているのに対し、高学歴化が進むにつれ、高学歴者の質的側面からの分析も蓄積されるようになってきた。特に、専攻が年収もしくは賃金に与える影響についてはこれまでに膨大な数の先行研究の蓄積がなされている (James et al., 1989; Loury and Garman, 1995; Loury, 1997; Blundell

et al., 2000; Bratti and Mancini, 2003; Arcidiacono, 2004)。また、専攻についての男女の属性の相違が賃金格差に与える影響に関する研究の蓄積もなされてきた (Brown and Corcoran, 1997; Black et al., 2007; Lin, 2010)。

国内で大卒以上の高学歴者が取得した学位の専攻分野に踏み込んだ研究としては、橘木他 (2009) および大谷他 (2003) は、文系出身者の所得が理系より高い可能性を指摘しているが、浦坂 (2011) および浦坂他 (2012) は、文系よりも理系出身者の年収が高いことを指摘している。浦坂他 (2011) は、「日本家計パネル調査」(JHPS) データを用いて分析を行っており、特に男性の場合には国立大学の理系の年収プロファイルが高い可能性を指摘している。

このような理系の賃金プレミアムについてはより詳細な分析も行われるようになっていく。山本他 (2015) は、文系専攻に対する理

* 正会員, 愛知学院大学経済学部
e-mail : take@dpc.agu.ac.jp

系専攻のプレミアムを確認した上で、理系を医学系(医学, 歯学, 薬学, 看護学)とそれ以外に区別した場合には、理系プレミアムのほとんどが医学系によるものであることを確認している。一方、安井(2019)も男女ともに医学・薬学系の賃金プレミアムが最も高いことを確認している。さらに、山本・安井(2016)は、理工系女子の賃金が文系出身者と比較して抑制されている可能性を指摘している。しかしながら、これまでの国内の研究では専攻分類が文系・理系の2分類、もしくはそれに準ずる分類となっている。本稿では、専攻の分類についてより詳細な設問が設けられている調査データを利用し、専攻10分類が年収に与える影響を考察する。さらに、これらが男女間賃金格差に与える影響も考察する。

日本では90年代以降大学進学率が上昇してきたが、特に女性の進学率が上昇しており、1990年に約18.2ポイントあった性差は、2019年の時点では約5.9ポイントまで縮小している(文科省 2020)。また、高学歴女性は結婚、出産後に労働市場に再参入しないことが指摘され、いわゆるM字型カーブを特徴とする年齢階級別の就業率のカーブが当てはまらず、その特徴は「きりん型」と称されていた。しかしながら、1997年と2017年の就業構造基本調査では、20年間の間に学歴に関わらず有業率と仕事を主にする者の比率が上昇している。また2017年にはライフイベント後の40代から50代の大卒以上と大卒未満の女性の有業率に全く差がないことが示されている。したがって、「きりん型」と称されていた年齢階級別のカーブは2017年の時点で全く成り立っていない(総務省 1998, 2018)。

このような状況下では、学歴間の格差以上に高学歴者間における賃金格差と男女間賃金格差が格差全体に占める比重が大きくなる可能性が指摘できる。実際に海外では多くの研究がこのような指摘に基づき、高学歴者間における専攻、スキルなどによる賃金プレミア

ムを計測している(Brown and Corcoran, 1997など)。本稿では専攻による賃金プレミアムの計測に加え、専攻が男女間賃金(年収)格差に与える影響および同一専攻内の男女間賃金(年収)格差についても考察する。同時に、大学院卒、国立大学卒の収益についても併せて考察する。なお、本稿では安井(2019)に倣い、専攻をダミー変数として説明変数に挿入し、他の属性をコントロールした上で当該専攻の相対的な賃金の上昇を専攻の賃金プレミアムという言葉を用いて捉える¹⁾。

2. 使用するデータと分析モデル

使用するデータは、「慶應義塾大学家計パネル調査(KHPS)」(慶應義塾大学パネルデータ設計・解析センター)であり、この調査は2004年から全国の男女4005人を対象に実施が始まり、2007年に1447人、2012年に1012人を調査対象に加えている。また、KHPSの調査実績を生かし、2009年からは新たに全国4022人の男女を対象とした「日本家計パネル調査(JHPS)」(慶應義塾大学パネルデータ設計・解析センター)が同時に実施されており、双方の調査を使用した2017年までのデータを分析に使用する。KHPSは20歳から69歳の男女、JHPSは20歳以上の男女を調査対象としている。調査対象は層化二段無作為抽出法により選定されている。これは、行政単位(都道府県・市町村)と地域によって全国をいくつかのブロックに層化し、各層に調査時点を人口に応じて比例配分し、国勢調査における調査地域及び住民基本台帳を利用して(二段)、各地点に一定数の標本抽出を行う手法である。

本稿の主眼である専攻に関しては、短大・高専、大学、大学院の中退、卒業者を対象に各調査初年度に設問が設けられている。本調査によるデータは、同一個人を追跡調査するパネルデータとなっているため、これらの回答者の各年度の観測値をプールして、変数に

欠損値がない観測値を用いて分析を行った。分析対象は、学位取得の効果を厳密に測るため、各学歴（大卒および大学院卒）の卒業・修了者のみとし、年齢が60歳以下の雇用者としている。パネルデータの利点を生かし、ランダム効果モデル（Random Effect Model）を用いて、専攻を説明変数に加え、年取の対数値を被説明変数とする分析を行った²⁾。さらに、専攻が男女間年取格差に与える影響を考察するため、慣例に倣い男女別の最小二乗法による推計結果を利用したブラインダー・ワハカ分解（Blinder-Oaxaca Decomposition）（Blinder, 1973; Oaxaca, 1973）による男女間格差の要因分解を行っている。また、同一専攻内の男女間格差を考察するため、専攻別の最小二乗法による推計を行い、説明変数に女性ダミーを挿入してその係数について考察した。使用する変数に欠損値と外れ値を持つ観測値を除き、分析に使用した観測数は雇用者合計7497、内男性5768、女性1729であり、正社員合計6530、内男性5519、女性1011である。

3. 記述統計量

3.1 雇用者および正社員の男女別記述統計

分析は、調査対象者の産業、企業規模、職種、職位の各分類を説明変数から除いた分析と含めた分析の両方を行っているが、表1は、産業等を除いた基本属性に関する記述統計である。この表は男女別に非正規雇用者を含めた雇用者全体と正社員に限定した統計量を提示している。まず、表1から男性雇用者の主要な特徴を把握する。年取対数値の平均は6.399であるが、これは年取に換算して約661万円程度である。平均年齢は約45歳、社会科学専攻が約30%、工学専攻が合計約28%を占め、医歯学、薬学は1%に満たない。女性雇用者は、年取対数値の平均は5.536、年取に換算すると約333万円であり、非正規雇用を含めると男性の平均年取と約328万円の差となる。平均年齢は約41歳、社会科学専攻は約19%、工学専攻は約3%程度である一方、人文科学約27%、教育約17%は男性に占める比率より高く、医歯学、薬学はそれぞれ約1.5%

表1 記述統計量

変数	男性雇用者		男性正社員		女性雇用者		女性正社員	
	平均	標準偏差	平均	標準偏差	平均	標準偏差	平均	標準偏差
年取対数値	6.399	0.467	6.441	0.410	5.536	0.828	6.033	0.463
年齢	45.090	8.957	45.164	8.800	41.621	9.710	39.080	9.470
年取2乗	2113.316	796.444	2117.191	783.695	1826.543	820.664	1616.839	784.717
勤続年数	16.225	9.858	16.681	9.775	10.123	8.177	11.986	8.847
大学院	0.105	0.307	0.104	0.305	0.027	0.163	0.042	0.200
国立大学卒	0.174	0.379	0.180	0.384	0.233	0.423	0.282	0.450
人文科学	0.101	0.302	0.098	0.298	0.271	0.444	0.225	0.417
社会科学	0.309	0.462	0.308	0.462	0.195	0.396	0.182	0.386
理学	0.046	0.209	0.046	0.210	0.032	0.177	0.053	0.225
工学	0.286	0.452	0.288	0.453	0.032	0.177	0.039	0.193
農学	0.040	0.195	0.039	0.194	0.027	0.161	0.026	0.158
医学・歯学	0.006	0.074	0.006	0.075	0.017	0.128	0.027	0.161
薬学	0.008	0.089	0.008	0.089	0.015	0.122	0.017	0.129
教育	0.055	0.229	0.055	0.228	0.173	0.378	0.179	0.384
家政	0.003	0.059	0.003	0.057	0.058	0.234	0.053	0.225
その他	0.146	0.353	0.148	0.355	0.180	0.385	0.200	0.400
既婚	0.826	0.379	0.845	0.362	0.616	0.487	0.513	0.500
子供の人数	1.438	1.074	1.475	1.064	1.082	1.096	0.733	0.963
0歳児有	0.035	0.184	0.036	0.187	0.012	0.110	0.018	0.132
未就学児有	0.199	0.399	0.206	0.404	0.117	0.321	0.137	0.345
サンプルサイズ	5768		5519		1729		1011	

程度であり僅かであるが男性の比率を上回り、男女の専攻における分布の差を確認できる。また、非正規を含む雇用者全体と正社員に限定した場合を比較すると、男性は専攻の分布にほとんど差異が認められないが、女性は理学、工学、医歯薬学の理系分野では正社員に占める比率が高まり、人文科学系では正社員に占める比率が低下することが確認できる。

その他の統計については、雇用者全体に占める大学院卒者は男性が約10.5%、女性が約2.7%、国立大学出身者は男性が約17.4%、女性が約23.3%であり、大学院卒者の比率と国立大学出身者の比率でも男女差が確認できた。特に、大学院卒者の比率は男性が高いが、国立大学出身者の比率は女性が高いことは特筆すべきである。

また掲載は割愛しているが、産業、企業規模、職種、職位に関する男女の記述統計の特

徴を把握すると、男性が製造(約23.9%)、情報・調査業(約11%)の順に多く、女性は教育・学習業(約21.1%)、卸・小売業(約13.5%)の順である。企業規模では男性が相対的に大企業に偏っており、女性は事務職(約43.3%)、専門・技術職(約48.3%)で大半を占めているのに対し、男性は専門・技術職、事務職、管理職、販売職など幅広い職種に分布している。また、男性は約半数が経営者を含めた役職に就いているのに対し、女性の役職比率は約13%程度に留まっている。

3.2 専攻別の記述統計

次に各専攻別の推計に関する記述統計量を雇用者と正社員別に表2-1、表2-2に掲載している。この表では基本属性に関する記述統計のみを掲載しているが、賃金(年収)関数の分析は産業、企業規模、職種、職位を除い

表2-1 雇用者の専攻別記述統計量

変数	人文科学		社会科学		理学		工学		農学	
	平均	標準偏差	平均	標準偏差	平均	標準偏差	平均	標準偏差	平均	標準偏差
年収対数値	5.926	0.819	6.249	0.654	6.463	0.378	6.462	0.449	6.112	0.609
女性ダミー	0.444	0.497	0.159	0.366	0.175	0.381	0.033	0.178	0.168	0.374
年齢	43.922	9.709	45.022	9.429	44.438	8.098	45.131	8.736	41.515	8.792
年齢2乗	2023.335	848.133	2115.856	832.055	2040.056	721.949	2113.067	776.546	1800.471	739.976
勤続年数	13.490	10.225	15.376	10.069	15.566	9.719	16.507	9.717	15.099	8.806
大学院	0.026	0.158	0.049	0.215	0.247	0.432	0.192	0.394	0.135	0.342
国立大学卒	0.104	0.306	0.115	0.319	0.328	0.470	0.239	0.426	0.482	0.501
既婚	0.751	0.433	0.766	0.423	0.747	0.435	0.876	0.330	0.858	0.350
子供の人数	1.166	1.088	1.313	1.096	1.222	1.168	1.626	1.010	1.358	1.022
0歳児有	0.027	0.161	0.024	0.152	0.047	0.212	0.037	0.189	0.040	0.197
未就学児有	0.153	0.360	0.156	0.362	0.291	0.455	0.204	0.403	0.266	0.443
サンプルサイズ	1053		2122		320		1705		274	
変数	医学		薬学		教育		家政		その他	
	平均	標準偏差	平均	標準偏差	平均	標準偏差	平均	標準偏差	平均	標準偏差
年収対数値	6.483	0.634	6.353	0.636	6.053	0.748	5.286	0.730	6.072	0.664
女性ダミー	0.475	0.504	0.361	0.484	0.484	0.500	0.833	0.374	0.271	0.445
年齢	42.180	9.836	48.014	8.036	44.168	8.690	40.642	8.754	42.977	9.605
年齢2乗	1874.344	842.326	2369.014	776.760	2026.236	762.210	1727.742	727.520	1939.155	826.419
勤続年数	6.639	4.680	14.708	9.947	15.707	10.104	7.475	5.173	12.956	8.899
大学院	0.262	0.444	0.111	0.316	0.021	0.144	0.000	0.000	0.037	0.190
国立大学卒	0.525	0.504	0.208	0.409	0.417	0.494	0.067	0.250	0.081	0.273
既婚	0.590	0.496	0.708	0.458	0.727	0.446	0.767	0.425	0.709	0.454
子供の人数	0.836	1.036	1.556	0.991	1.429	1.117	1.292	1.260	1.229	1.077
0歳児有	0.033	0.180	0.028	0.165	0.024	0.154	0.008	0.091	0.031	0.174
未就学児有	0.197	0.401	0.194	0.399	0.157	0.364	0.133	0.341	0.178	0.383
サンプルサイズ	61		72		618		120		1152	

表2-2 正社員の専攻別記述統計量

変数	人文科学		社会科学		理学		工学		農学	
	平均	標準偏差	平均	標準偏差	平均	標準偏差	平均	標準偏差	平均	標準偏差
年収対数値	6.304	0.440	6.399	0.450	6.497	0.326	6.502	0.392	6.264	0.371
女性ダミー	0.295	0.456	0.098	0.297	0.174	0.380	0.024	0.153	0.107	0.310
年齢	44.061	9.831	44.941	9.356	44.384	7.971	45.060	8.582	41.741	8.650
年齢2乗	2037.910	855.907	2107.144	826.611	2033.255	711.510	2104.013	762.274	1816.802	730.319
勤続年数	15.631	10.500	16.417	9.979	15.926	9.649	16.972	9.584	15.905	8.761
大学院	0.022	0.147	0.055	0.227	0.252	0.435	0.188	0.391	0.148	0.356
国立大学卒	0.107	0.309	0.119	0.324	0.326	0.469	0.239	0.427	0.510	0.501
既婚	0.779	0.415	0.785	0.411	0.758	0.429	0.891	0.312	0.905	0.293
子供の人数	1.174	1.120	1.309	1.089	1.235	1.163	1.653	0.997	1.477	0.968
0歳児有	0.034	0.181	0.027	0.161	0.048	0.215	0.039	0.193	0.045	0.208
未就学児有	0.176	0.381	0.167	0.373	0.300	0.459	0.211	0.408	0.296	0.458
サンプルサイズ	769		1884		310		1631		243	
	医学		薬学		教育		家政		その他	
変数	平均	標準偏差	平均	標準偏差	平均	標準偏差	平均	標準偏差	平均	標準偏差
年収対数値	6.534	0.609	6.569	0.383	6.358	0.380	5.713	0.366	6.222	0.476
女性ダミー	0.466	0.503	0.279	0.452	0.374	0.484	0.750	0.436	0.198	0.399
年齢	42.759	9.730	48.410	8.034	43.853	8.806	36.444	7.720	42.771	9.422
年齢2乗	1921.345	836.807	2407.000	775.802	2000.502	770.684	1386.972	612.551	1918.054	811.987
勤続年数	6.879	4.676	15.885	10.077	18.242	9.753	8.597	5.980	13.686	8.950
大学院	0.276	0.451	0.131	0.340	0.014	0.120	0.000	0.000	0.042	0.201
国立大学卒	0.534	0.503	0.246	0.434	0.459	0.499	0.069	0.256	0.080	0.271
既婚	0.586	0.497	0.738	0.444	0.723	0.448	0.625	0.488	0.712	0.453
子供の人数	0.879	1.044	1.393	0.971	1.380	1.088	0.819	1.155	1.192	1.072
0歳児有	0.034	0.184	0.033	0.180	0.029	0.168	0.014	0.118	0.033	0.180
未就学児有	0.207	0.409	0.213	0.413	0.176	0.381	0.125	0.333	0.193	0.394
サンプルサイズ	58		61		484		72		1018	

た分析と含めた分析の両方を行っている。表2-1の雇用者の専攻別記述統計について、年収対数値約6.24、平均年齢約45歳である社会科学専攻を基準に他の専攻の特徴を確認すると、工学は年収対数値約6.46、平均年齢約45歳、医学は年収対数値約6.48、平均年齢約42歳であり、年収が相対的に高いことが伺える。女性比率が高いのは、家政約83%、教育約48%、医学約47%、人文科学約44%の順である。一方で工学約3%、社会科学約15%、農学約16%、理学約17%は女性の比率が低い。また、大学院の進学率は医学約26%、理学約24%、工学約19%の順に高く、大学院への進学率が低い専攻は、家政0%、人文科学約2%、教育約2%である。国立大学出身者の比率が高いのは、医学約52%、農学約48%、教育約41%であり、国立大学出身者の比率が低いのは、家政約6%、そのほか約8%、人

文科学約10%である。

また、数値は掲載していないが、各専攻の内、家政、人文科学、教育は相対的に非正規雇用者の比率が高い（数値は家政約40%、人文科学約26%、教育約21%）。したがって、これらの専攻では表2-2の正社員に限定した記述統計において、年収が相対的に改善しているのが把握できる。一方、表2-2の正社員に限定した専攻内の女性比率は、家政約75%、医学約46%、教育約37%、人文科学約29%の順に高いが、医学以外の3分野は、表2-1の雇用者の女性比率から約8%から約15%減少している。大学院進学率、国立大学進学率は正社員に限定しても大きな変化はなく、総じて若干程度比率が上昇傾向にある。

4. 分析結果

分析結果は以下の3つに分けて考察する。

まず表3および表4は、男女別の賃金(年収)関数の分析結果(雇用者および正社員別)を提示している。それぞれ産業、企業規模、職種、職位を除いた推計結果と含めた推計結果である。次に表5、表6は雇用者と正社員に限定した男女別の賃金(年収)関数の格差要因分解の結果を掲示している。最後に表7は専攻別の推計結果における女性ダミー(女性を1,男性を0とした説明変数)の効果(雇用者および正社員別)を産業、企業規模、職種、職位を除いた推計結果と含めた推計結果を掲載している。表3から表6の推計における各専攻の係数のベースカテゴリーは人文科学である。

分析結果を考察するにあたり、特に留意しなければならない点を挙げておく。まず、ブラインダー・ワハカ分解については、杉橋(2009)が指摘するとおり、説明変数別の要因分解を行う際に、各変数が独立していることを想定しているが、実際には属性間の相互関係があることは念頭に置き結果を解釈しなければならない。さらに、使用する観測値は2004年から2017年のパネルデータをプールして分析したものであるため、坂本(2006)などで指摘されている通り、特定の特徴を有するサンプルが調査から脱落する傾向にあれば推計結果にバイアスが生じている可能性があることには留意しなければならない。

4.1 男女別の賃金(年収)関数の分析結果 (表3および表4)

まず、表3から男性の雇用者、正社員別に推計した賃金関数の結果を考察する。被説明変数は年収の対数値として分析しているため、専攻のダミー変数はベースカテゴリーである人文科学(係数0)に対して、係数×100%の年収増加(もしくは減少)を示すことになる。表3の結果から、雇用者の推計(産業等を説明変数から除く)結果を見ると、年齢、年齢2乗項、勤続年数、既婚ダミー、子供の数は有

意に年収に対して正の効果を持ち多くの先行研究と一致する結果が得られている。専攻の賃金プレミアム(賃金の相対的優位)が最も大きい分野は医学・歯学であり、次に薬学が続く。これらの係数は、雇用者でそれぞれ0.446, 0.360かつ1%水準で有意であり、理学(係数0.068有意でない)および工学(係数0.106雇用者のみ有意)と比較しても高い。また、大学院および国立大学卒業も年収に有意に正の影響を与えているが、その効果は大学院卒が大きい(大学院卒係数0.181, 国立大卒0.154)。医学・歯学および薬学の係数については産業等を含めた推計、正社員に限定した推計でも依然1%水準で有意に賃金に対して正の効果を持ち、係数も大きな変化はない。一方、同じ理系分野でも農学は雇用者(産業等を含む)および正社員に限定した分析で有意に負の効果を示している。男性サンプルは少ないものの、家政専攻も有意に負の効果を持つ。その程度は、ベースカテゴリーである人文科学の年収の約31%から約50%と比較的大きい。

また、表4の女性の分析における専攻の効果は、雇用者の推計では有意な結果が得られるが、正社員に限定した推計ではほとんど有意ではない。サンプルサイズが限定されることが影響していると思われる。雇用者の推計(産業等を説明変数から除く)における専攻の賃金プレミアムは、理学、医学・歯学、工学、薬学の順に高く、係数はそれぞれ0.748, 0.666, 0.557, 0.473である。また、理学に確認できる74.8%の高いプレミアムは産業、企業規模等を細かくコントロールすると順位が逆転し、医学・歯学、薬学の方がプレミアムは高くなる。大学院と国立大学卒も男性同様に年収に正の影響を与えているが、特に女性は大学院卒のプレミアムが男性より大きい(大学院係数0.668, 国立大卒係数0.157)。ただし、この結果も正社員に限定すると10%まで有意水準が落ち、さらに産業等を細かくコ

表3 男性の賃金(年収)関数の結果

	雇用者	雇用者 産業等含む	正社員	正社員 産業等含む
年齢	0.080 *** (0.005)	0.066 *** (0.005)	0.077 *** (0.004)	0.071 *** (0.004)
年齢2乗項	-0.001 *** 5.3E-05	-0.001 *** 4.9E-05	-0.001 *** 4.5E-05	-0.001 *** 4.5E-05
勤続年数	0.016 *** (0.001)	0.010 *** (0.001)	0.012 *** (0.001)	0.010 *** (0.001)
大学院	0.181 *** (0.044)	0.141 *** (0.033)	0.164 *** (0.038)	0.139 *** (0.032)
国立大学	0.154 *** (0.035)	0.093 *** (0.026)	0.120 *** (0.030)	0.083 *** (0.026)
社会科学	0.080 * (0.044)	0.038 (0.033)	0.022 (0.039)	1.9E-04 (0.033)
理学	0.068 (0.069)	0.042 (0.052)	0.029 (0.060)	0.011 (0.051)
工学	0.106 ** (0.045)	0.051 (0.035)	0.043 (0.040)	0.032 (0.034)
農学	-0.119 (0.074)	-0.111 ** (0.055)	-0.146 ** (0.064)	-0.140 *** (0.054)
医学・歯学	0.446 *** (0.123)	0.476 *** (0.094)	0.359 *** (0.105)	0.438 *** (0.090)
薬学	0.360 *** (0.129)	0.407 *** (0.097)	0.341 *** (0.115)	0.371 *** (0.097)
教育	0.023 (0.067)	-0.023 (0.051)	-0.015 (0.058)	-0.056 (0.051)
家政	-0.507 ** (0.224)	-0.397 ** (0.166)	-0.367 * (0.191)	-0.317 ** (0.160)
その他	-0.020 (0.049)	-0.071 * (0.037)	-0.067 (0.043)	-0.081 ** (0.036)
既婚	0.079 *** (0.016)	0.075 *** (0.015)	0.056 *** (0.014)	0.064 *** (0.014)
子供の人数	0.019 *** (0.006)	0.021 *** (0.005)	0.022 *** (0.005)	0.023 *** (0.005)
出産年	0.010 (0.013)	0.007 (0.013)	0.000 (0.011)	-0.001 (0.011)
未就学児童有	-0.010 (0.009)	-0.012 (0.009)	-0.007 (0.008)	-0.008 (0.008)
定数項	4.168 *** (0.112)	3.810 *** (0.231)	4.329 *** (0.097)	3.730 *** (0.218)
σ_u	0.374	0.267	0.318	0.257
σ_e	0.158	0.147	0.130	0.128
ρ	0.848	0.769	0.858	0.801
R-squared: overall	0.385	0.581	0.372	0.493
Number of observations	5,768	5,768	5,519	5,519

注) 括弧内はロバストな標準誤差である。***1%, **5%, *10%水準で有意である。

ントロールすると有意でなくなる。女性正社員の記述統計を合わせて考察すると、プレミアムが認められる専攻の学位取得、および大学院、国立大学卒は、女性が正社員になる確率を高めており、そのことを通じて年収を上げている可能性が高い。先行研究の結果では、

山本他(2015)、安井(2019)が理系プレミアムの中でも医学・薬学系のプレミアムが最も高いことを指摘している。ただし、山本他(2015)は、特に女性の場合に理工系(医学系以外)の賃金が文系より抑制されている可能性を指摘しており、安井(2019)の結果も

表4 女性の賃金(年収)関数の結果

	雇用者	雇用者 産業等含む	正社員	正社員 産業等含む
年齢	0.047 *** (0.013)	0.035 *** (0.013)	0.054 *** (0.011)	0.038 *** (0.011)
年齢2乗項	-0.001 *** (1.5E-04)	-0.001 *** (1.4E-04)	-0.001 *** (1.3E-04)	-0.4E-04 *** (1.3E-04)
勤続年数	0.038 *** (0.003)	0.034 *** (0.003)	0.019 *** (0.004)	0.023 *** (0.003)
大学院	0.668 *** (0.241)	0.445 *** (0.175)	0.274 * (0.149)	0.170 (0.130)
国立大学	0.157 * (0.090)	0.065 (0.067)	0.104 * (0.062)	0.036 (0.055)
社会科学	0.182 * (0.105)	0.142 * (0.078)	0.064 (0.081)	0.106 (0.072)
理学	0.748 *** (0.243)	0.500 *** (0.178)	0.231 (0.143)	0.313 ** (0.128)
工学	0.557 ** (0.232)	0.372 ** (0.170)	0.175 (0.145)	0.181 (0.127)
農学	-0.011 (0.215)	-0.103 (0.159)	-0.057 (0.167)	-0.033 (0.146)
医学	0.666 ** (0.285)	0.603 *** (0.214)	0.212 (0.199)	0.300 * (0.179)
薬学	0.473 ** (0.235)	0.503 *** (0.177)	0.094 (0.180)	0.178 (0.162)
教育	0.187 * (0.113)	0.133 (0.085)	0.047 (0.088)	-0.022 (0.080)
家政	0.004 *** (0.161)	0.072 (0.120)	-0.241 ** (0.120)	-0.123 (0.107)
その他	0.182 * (0.108)	0.050 (0.080)	-0.004 (0.080)	0.023 (0.070)
既婚	-0.088 ** (0.038)	-0.060 * (0.035)	0.019 (0.031)	0.006 (0.031)
子供の人数	-0.073 *** (0.021)	-0.047 ** (0.019)	-0.045 ** (0.019)	-0.047 ** (0.018)
出産年	0.051 (0.069)	0.011 *** (0.066)	-0.028 (0.049)	-0.025 (0.049)
未就学児童有	-0.119 *** (0.033)	-0.099 *** (0.032)	-0.107 *** (0.030)	-0.108 *** (0.031)
定数項	4.295 *** (0.283)	4.735 *** (0.593)	4.543 *** (0.234)	4.412 *** (0.584)
σ_u	0.628	0.440	0.359	0.299
σ_e	0.264	0.249	0.183	0.178
ρ	0.850	0.758	0.793	0.738
R-squared: overall	0.325	0.625	0.354	0.500
Number of observations	1,729	1,729	1,011	1,011

注) 括弧内はロバストな標準誤差である。***1%, **5%, *10%水準で有意である。

女性について社会科学系が自然科学系(医学系以外)を年収で上回る可能性を示唆している。本稿では、少なくとも女性については理学、工学系で相対的な年収の低下は認められない。この点については、理学、工学と農学を区別した分析を行ったことも要因として考

えられるが、十分な大きさのサンプルを揃えた検証が必要である。

4.2 男女間賃金格差要因分解の分析結果(表5および表6)

次に男女間賃金(年収)格差を、最小二乗

法で推計した男女の賃金（年収）関数の結果を用いて、ブラインダー・ワハカ分解を行い、説明出来る要因（属性（要素）の差）と説明できない要因（係数（評価）の差）に分けて考察する。表5は雇用者の要因分解の結果を示している。表の上段は説明できる要因（属性の差）、下段は説明できない要因（係数の差）が男女間賃金格差に及ぼす影響を示しており、変数の係数が正の場合は格差拡大要因、負の場合は格差縮小要因となっていると判断できる。

まず属性の差では、大学院卒が係数0.017および社会科学が係数0.010、理学0.003、工学0.033であり、この3分野が有意に男女間賃金格差を広げる要因となっている。特に工学の係数が最も大きく、女性がこの分野で少ないことが格差を広げる要因となっていることが判る。一方で、国立大学卒の係数-0.009および医学・歯学-0.007、薬学-0.002、農学-0.002、教育-0.006の4分野は有意に男女間賃金格差を縮小する方向に働いている。表1の記述統計から医学・歯学、薬学の分野における女性の選択比率が高いことが要因と考えられる。逆に農学は男女ともに他の専攻と比較して相対的に賃金（年収）が低いことから、女性がこの分野を選択する比率が男性より低いことが格差縮小に寄与していると推察される。

一方、表5の下段に提示されている係数の差を考察すると、大学院卒が係数-0.011、理学-0.018、工学-0.032、教育-0.015、家政-0.013は、有意に格差縮小の方向に寄与しており、表4の女性雇用者の賃金（年収）関数におけるこれらの属性の係数が正に有意であることが男女間賃金格差の縮小に貢献していると考えられる。一方で国立大学卒、医学・歯学、薬学、農学の係数の差は有意ではない。女性内では医学・歯学、薬学の専攻が高年収の獲得に有利となっているが、これらの係数は表3に示される男性の賃金（年収）

関数でも係数の値が高いため、男女間の格差を縮小する要因にはなっていないと推察される。

以上から、高学歴者の学業に関わる変数は大学院への進学率および社会科学、理学、工学の選択比率の性差が男女間賃金格差を広げる要因になっているが、係数の差に起因する格差拡大の要因にはなっていないことが判る。特に理学と工学については女性比率が低いという属性の差が有意に男女間賃金格差を拡大させているが、係数の差が縮小方向に機能している。また、表5の下段より、既婚0.316、子供の人数0.204の2つの変数だけで係数の合計が0.520になり、係数の差に起因する格差全体の数値0.641（表5の最下段に記載）の約81%を説明することが判る。

表6は、正社員に限定した格差要因分解の推計結果である。表5と同様に上段には属性の差、下段には係数の差の結果を提示している。正社員に限定すると男女間の賃金格差は縮小するため、係数は縮小する変数が多い。その中で年齢については属性の差の係数が大きくなり（表5では0.069から表6では0.270）、係数の差は係数が小さくなり（表5では1.886から表6では0.115）かつ有意ではなくなっている。表1の記述統計から女性正社員の平均年齢が約2.5歳若くなること、および正社員に限定されるため係数の差が改善されていることを示唆する。また、国立大学の属性の差の係数は-0.012（表6）であり雇用者全体の場合の-0.009（表5）より負に僅かに大きくなるのは国立大学卒であることがより女性の正社員になる確率を高めていることを反映していると推察される。理学の係数が有意でなくなることも同様の解釈が可能である。一方で、係数の差は10分野の専攻の中で、理学の係数-0.010のみが負に有意であり、格差縮小の要因となっていることを示している。薬学の係数は0.003で正に有意であり、格差拡大要因となっている。表5雇用者と表6正

表5 雇用者の格差要因分解

Explained	雇用者		雇用者産業等含む	
	係数	標準誤差	係数	標準誤差
年齢	0.069	0.022***	0.107	0.019***
年齢2乗項	-0.092	0.022***	-0.089	0.018***
勤続	0.181	0.009***	0.061	0.005***
大学院	0.017	0.002***	0.013	0.001***
国立大学	-0.009	0.002***	-0.003	0.001***
社会科学	0.010	0.003***	0.001	0.002
理学	0.003	1.2E-03**	0.001	3.9E-04*
工学	0.033	0.005***	-6.3E-05	0.004
農学	-0.002	0.001**	-0.001	0.001**
医学・歯学	-0.007	0.002***	-0.005	0.002***
薬学	-0.002	1.2E-03**	-0.002	0.001**
教育	-0.006	0.003*	-0.002	0.003
家政	0.012	0.004***	0.009	0.003***
その他	-0.002	0.001*	0.001	7.7E-04*
既婚	0.031	0.004***	0.014	0.003***
子供人数	-0.012	0.003***	-0.002	0.002
出産	-4.7E-04	0.001	-0.001	0.001
未就学児童	-2.8E-04	0.001	-0.003	0.001**
Total	0.222	0.011***	0.567	0.018***
Unexplained	雇用者		雇用者産業等含む	
	係数	標準誤差	係数	標準誤差
年齢	1.886	0.749**	0.133	0.634
年齢2乗項	-0.430	0.393	0.098	0.328
勤続	-0.404	0.030***	-0.135	0.028***
大学院	-0.011	0.003***	-0.004	0.002*
国立大学	-0.005	0.010	0.001	0.008
社会科学	-0.004	0.012	0.001	0.010
理学	-0.018	0.003***	-0.005	0.002**
工学	-0.032	0.006***	-0.008	0.005*
農学	0.004	0.004	0.004	0.003
医学・歯学	-0.002	1.6E-03	0.003	0.001**
薬学	-0.002	1.9E-03	0.001	0.001
教育	-0.015	0.007**	3.5E-04	0.006
家政	-0.013	0.004***	-0.008	0.003**
その他	-6.6E-05	0.009	0.005	0.007
既婚	0.316	0.030***	0.205	0.025***
子供人数	0.204	0.026***	0.118	0.020***
出産	1.1E-04	0.003	0.002	0.002
未就学児童	-0.022	0.009**	-0.007	0.006
定数項	-0.811	0.350**	-0.164	0.372
Total	0.641	0.021***	0.296	0.015***

注) ***1%, **5%, *10%水準で有意であることを示す。

社員のいずれの推計でも産業等の変数を含むと、医学・歯学、教育、大学院等の一部の変数で係数の差における正の効果が大きくなる。サンプルサイズを増やした再検証により精緻な分析が必要であると思われるが、可能性としては、産業、職種等の追加した変数が男女間賃金格差の係数の差を縮小させる方向に機能しており、相殺する形で一部の変数が正に有意になった可能性も考えられる。

表6 正社員の格差要因分解

Explained	正社員		正社員産業等含む	
	係数	標準誤差	係数	標準誤差
年齢	0.270	0.031***	0.312	0.030***
年齢2乗項	-0.239	0.028***	-0.259	0.027***
勤続	0.078	0.006***	0.047	0.004***
大学院	0.012	0.002***	0.008	0.001***
国立大学	-0.012	0.002***	-0.006	0.001***
社会科学	0.003	1.8E-03*	0.001	0.002
理学	-0.001	0.001	-2.9E-04	3.4E-04
工学	0.008	0.003**	-2.1E-04	0.003
農学	-0.002	0.001**	-0.001	0.001**
医学・歯学	-0.007	0.002***	-0.009	0.003***
薬学	-0.002	0.001**	-0.002	1.1E-03**
教育	0.002	0.002	0.004	0.002*
家政	0.016	0.003***	0.011	0.003***
その他	0.003	0.001***	0.002	9.6E-04**
既婚	0.067	0.006***	0.057	0.005***
子供人数	0.006	0.005	0.009	0.004**
出産	-2.6E-04	3.4E-04	-2.4E-04	3.1E-04
未就学児童	-0.003	0.001***	-0.003	0.001***
Total	0.198	0.011***	0.232	0.013***
Unexplained	雇用者		雇用者産業等含む	
	係数	標準誤差	係数	標準誤差
年齢	0.115	0.554	0.156	0.485
年齢2乗項	0.112	0.281	0.072	0.241
勤続	-0.158	0.030***	-0.199	0.028***
大学院	0.002	0.003	0.006	0.003**
国立大学	-0.012	0.009	-0.005	0.008
社会科学	0.003	0.008	-0.015	0.008*
理学	-0.010	0.003***	-0.011	0.003***
工学	-0.004	0.004	-0.007	0.004*
農学	-0.002	0.002	-0.003	0.002
医学・歯学	0.003	0.002	0.004	0.002**
薬学	0.003	0.001**	0.003	0.001**
教育	0.007	0.007	0.020	0.007***
家政	0.000	0.002	-0.007	0.003***
その他	0.014	8.04E-03*	-0.001	0.007
既婚	0.103	0.021***	0.092	0.018***
子供人数	0.030	0.015**	0.061	0.014***
出産	6.9E-05	0.001	1.2E-05	0.001
未就学児童	-0.002	0.007	-0.001	0.006
定数項	0.005	0.272	-0.009	0.293
Total	0.210	0.014***	0.176	0.013***

注) ***1%, **5%, *10%水準で有意であることを示す。

また、表5の上段、下段それぞれの最下段に掲載しているTotalは、属性の差、係数の差が総合的な男女間の年収差にどのように影響しているかを示す。この数値から雇用者の推計で産業等の変数を含まない場合は、属性の差に起因する男女間格差が0.222、係数に起因する男女間格差が0.641であり、係数の差が男女間格差の約74%程度を説明することになるが、表5の右側に掲載している産業等

の変数をコントロールした場合の格差要因分解の結果では、属性の差のTotalが0.567、係数の差が0.296であり、属性の差が男女間格差の約65%程度を説明することになる。

表6の正社員に限定した場合は男女間賃金格差自体が縮小する。産業等の変数をコントロールしない場合は属性の差が0.198、係数の差が0.210であり、産業等の変数をコントロールすると、属性の差が0.232、係数の差が0.176となり、前者では係数の差が格差全体の約51%を説明するが、後者では属性の差が約56%を説明することになる。雇用者、正社員のいずれの場合にも、産業、企業規模、職種、職位における男女の属性の分布の差が格差に影響を与えているため、これらの変数を追加することにより属性の差における男女間賃金格差が大きく計測されたと推察できる。

最後に10分野の専攻がトータルで格差に与える影響について算出すると、表5の非正規雇用を含む雇用者の推計で産業等の変数を含めた場合にのみ、専攻の属性および係数の差が格差全体の約3.9%程度を説明するが、それ以外の推計では認められないか僅かな縮小要因に留まる。むしろ表5、表6から、年齢、既婚が属性の差、係数の差の両方において有意に格差拡大の要因となっている。また、勤続年数は属性の差で拡大要因となるが、係数の差がそれを相殺している。子供の数は特に係数の差で格差拡大要因となっており、ライフイベントとそれによる就業中断が男女間賃金格差を拡大させる要因となっていることが確認できる。

4.3 同一専攻内における女性へのリターンに関する分析結果(表7)

次に、同一専攻内で男女格差が生じているか否かを、専攻別に推計を行うにあたり女性ダミーを説明変数として加え考察した。表7は、専攻別の推計結果の内、説明変数として使用した女性、大学院、国立大学の各ダミー

変数の結果を抜粋して掲載している。これらの推計の記述統計量が表2-1、2-2である。これまでの推計と同様に、雇用者と正社員に分けて推計をし、それぞれに、産業、企業規模、職種、職位を変数として含めない場合と含めた場合の結果を併せて掲載している。

分析の結果、理学以外の専攻で女性ダミーの係数は有意に負の効果が認められた。雇用者(産業等を含まない)を対象とした推計では、女性ダミーの負の係数が大きい順から、人文科学-0.849、教育-0.739、農学-0.730、社会科学-0.728、医学・歯学-0.694、薬学-0.605が挙げられ、工学-0.393は相対的に負の効果は弱く、理学0.006は有意でない。概して、理系分野内の方が、女性であることの賃金ペナルティ(賃金の相対的劣位)が弱い傾向にあると言える。一方、正社員(産業等を含まない)に限定した分析では、医学・歯学-0.596、薬学-0.536、人文科学-0.329、教育-0.289の順にペナルティが大きく、雇用者の分析同様に工学-0.209と理学-0.097では負の効果は弱く、理学分野では有意にならなかった。また、大学院卒のプレミアムは雇用者(産業等を含まない)で薬学0.713、医学・歯学0.565、農学0.420の順に大きく、国立大学卒のプレミアムは、雇用者(産業等を含まない)で教育0.247、社会科学0.183、工学0.130の順に大きかった。

表5、表6の格差要因分解では、専攻を説明変数として男女別の推計を行った上で、男女間賃金(年収)格差に専攻が与える効果の比較を行ったが、表7の推計では、同一専攻内の男女間賃金(年収)格差について、専攻別の推計に女性ダミーを挿入することで格差を捉えた。相対的に年収の高い医学・歯学、薬学を専攻する女性の比率が男性より高いことは表5、表6に示されるとおり属性の差による男女間格差の縮小要因となっているが、これらの専攻分野では男性の年収も高いため、同一専攻内では依然男女間格差が大きい可能

表7 専攻別の推計結果抜粋（女性ダミー，大学院，国立大学の効果）

	雇用者		雇用者産業等含む		正社員		正社員産業等含む	
	係数	標準誤差	係数	標準誤差	係数	標準誤差	係数	標準誤差
人文								
女性	-0.849	0.096 ***	-0.516	0.075 ***	-0.330	0.064 ***	-0.314	0.059 ***
大学院	0.354	0.308	0.325	0.223	0.150	0.217	0.169	0.193
国立大学	0.026	0.152	0.005	0.110	-0.011	0.097	-0.069	0.087
社会科学								
女性	-0.728	0.068 ***	-0.438	0.049 ***	-0.278	0.065 ***	-0.212	0.053 ***
大学院	0.307	0.128 **	0.247	0.089 ***	0.198	0.097 **	0.196	0.078 **
国立大学	0.184	0.081 **	0.130	0.056 **	0.184	0.064 ***	0.138	0.051 ***
理学								
女性	0.006	0.124	-0.065	0.103	-0.098	0.105	-0.077	0.112
大学院	0.190	0.111 *	0.084	0.091	0.084	0.095	0.059	0.100
国立大学	-0.036	0.101	-0.076	0.082	-0.060	0.087	-0.085	0.091
工学								
女性	-0.393	0.112 ***	-0.169	0.093 *	-0.210	0.108 *	-0.090	0.100
大学院	0.183	0.055 ***	0.179	0.042 ***	0.175	0.050 ***	0.160	0.043 ***
国立大学	0.130	0.048 ***	0.070	0.038 *	0.128	0.044 ***	0.088	0.038 **
農学								
女性	-0.730	0.156 ***	-0.488	0.174 ***	-0.248	0.131 *	-0.397	0.136 ***
大学院	0.421	0.159 ***	0.024	0.172	0.311	0.108 ***	0.168	0.116
国立大学	-0.090	0.118	0.016	0.131	-0.204	0.086 **	-0.082	0.093
医学・歯学								
女性	-0.695	0.225 ***	-0.591	0.091 ***	-0.597	0.239 **	-0.591	0.091 ***
大学院	0.566	0.264 **	0.097	0.210	0.556	0.258 **	0.097	0.210
国立大学	-0.132	0.217	0.157	0.079 **	-0.189	0.219	0.157	0.079 *
薬学								
女性	-0.606	0.158 ***	-0.474	0.070 ***	-0.537	0.091 ***	-0.399	0.080 ***
大学院	0.713	0.362 **	0.425	0.183 **	0.485	0.202 **	0.379	0.177 **
国立大学	0.032	0.245	0.014	0.180	0.020	0.130	0.293	0.195
教育								
女性	-0.739	0.110 ***	-0.478	0.077 ***	-0.290	0.067 ***	-0.284	0.054 ***
大学院	-0.208	0.286	-0.231	0.189	-0.158	0.153	-0.277	0.124 **
国立大学	0.248	0.108 **	0.087	0.071	0.141	0.064 **	0.089	0.052 *
家政								
女性	-0.429	0.381	0.186	0.197	-0.317	0.305	-0.120	0.180
大学院	(omitted)	(omitted)	(omitted)	(omitted)	(omitted)	(omitted)	(omitted)	(omitted)
国立大学	0.164	0.358	-0.083	0.197	0.392	0.310	0.237	0.213
その他								
女性	-0.680	0.086 ***	-0.456	0.065 ***	-0.327	0.066 ***	-0.310	0.061 ***
大学院	0.062	0.251	0.076	0.174	0.012	0.157	0.068	0.141
国立大学	0.684	0.160 ***	0.499	0.114 ***	0.452	0.107 ***	0.342	0.099 ***

注) ロバストな標準誤差を提示している。***1%，**5%，*10%水準で有意であることを示す。

性が指摘できる。一方で、同じ理系分野でも工学、理学分野は女性の比率は低いが、格差要因分解の係数の差で男女間賃金格差の縮小要因となっている（表5下段）ため、同一専攻内の女性ダミーも相対的にペナルティが抑えられて示されている可能性がある。以上は、

男女間賃金（年収）格差を考察する上で、専攻間と専攻内では格差が異なる構造を持ち、区別する必要性を示唆する。

5. 最後に

本稿における分析の結果、高学歴者の専攻

の収益は男女で異なる結果が得られた。特に、男女の理系の賃金プレミアムについては異なる要因が明らかとなった。男性の場合には理系の賃金プレミアムに医歯学および薬学の学位取得が貢献しているが、女性の場合にはこれらの専攻に加えて、理学、工学も賃金プレミアムに貢献しており、女性内の賃金格差に専攻が与える影響は男性より大きい可能性が認められた。また、賃金関数の男女間格差要因分解では、特に理学・工学分野の属性の差つまり女性比率が少ないことが、格差拡大に影響を与える一方で、係数つまり評価の差は格差縮小方向に働いていることが明らかとなった。また、専攻に関わる変数以上に既婚、子供の人数といったライフイベントに関わる変数が係数の差を通じて男女間賃金格差に大きく影響していることが確認できた。最後に、各専攻別の推計における女性ダミーの効果を計測したところ、理学以外の専攻分野では女性ダミーの効果が有意に負であることが確認され、同一専攻内で女性の年収が低いという

事実が確認できた。また、負の効果は医学・歯学、薬学と比較して工学は低く抑えられており、同一専攻内における女性の賃金ペナルティは理工系で相対的に低い可能性が確認できた。

最後に残された課題を述べたい。本来であれば、専攻内についても格差要因分解を行うべきであるが、例えば工学など女性の割合が低い専攻では分析を行うための十分なサンプルが確保できないため今回は女性ダミーを説明変数に加えることで専攻内の格差を考察した。また、この研究では、海外の研究に見られるような個々人の異質性を考慮した分析を行っていない。例えば、親の学歴と専攻や大学入学以前から本人が持つ数学的能力、語学能力等を考慮した場合の専攻の賃金プレミアムを計測することも必要である。また、女性の場合はサンプルセレクションバイアスを回避するため、無業者も考慮した推計を行う必要がある。以上の点は、今後の研究課題としたい。

謝辞

本稿の執筆にあたり、本誌2名の査読者および基盤研究(A)「社会的能力の特定化とその育成適正期および教育効果の検証」研究会、社会労働研究会のメンバーから有益なコメントをいただきましたことに深く感謝申し上げます。また、データ分析にあたり慶應義塾大学パネルデータ設計・解析センターより「慶應義塾大学家計パネル調査(KHPS)」 「日本家計パネル調査(JHPS)」の提供を受けました。本研究は科研費基盤研究(C)研究課題18K01652の助成を受けています。

注

- 1) 人的資本の質に立脚して専攻分野の賃金プレミアムを考察する場合、山本・安井(2016)が指摘するとおり、理系の人的資本を特殊的、文系は一般的と見做し、分析結果からいずれの要素が日本企業で評価されているかを考えることもできるが、そのためには専攻と職業のマッチングが完全に行われていることが前提となる。本稿では、この点の判別が十分にできないため、結果の解釈については限定的になるが、安井(2019)も述べているとおり、高度人材の養成が社会的に求められている中で、高等教育における専攻の賃金プレミアム(労働市場での評価)を明らかにすることは一定の意義があると考えられる。
- 2) 男女の賃金関数の推計については、ランダム効果モデルと最小二乗法の両方の推計を行い、本稿の考察に大きな相違がないことを確認した上で、パネルデータの利点を活かしたランダム効果モデルの結果を掲載している。

参考文献

- 浦坂純子・西村和雄・平田純一・八木匡 (2011)「理系出身者と文系出身者の年収比較 — JHPS データに基づく分析結果 —」, RIETI Discussion Paper Series, 11-J-020.
- 浦坂純子・西村和雄・平田純一・八木匡 (2012)「パネルデータに基づく理系出身者と文系出身者の年収比較」*Journal of Quality Education*, 4, pp.1-10.
- 大谷剛・松繁寿和・梅崎修 (2003)「卒業生の所得とキャリアに関する学部間比較」, OSIPP Discussion Paper, DP-2003-J-007.
- 坂本和靖 (2006)「サンプル脱落に関する分析 — 消費生活に関するパネル調査」を用いた脱落の規定要因と推計バイアスの検証」『日本労働研究雑誌』第551号, pp.55-70.
- 杉橋やよい (2009)「男女間賃金格差の要因分解手法の意義と内在的限界」『経済士林』第76号, pp.53-79.
- 総務省 (1998)『平成9年就業構造基本調査』
- 総務省 (2018)『平成29年就業構造基本調査』
- 橋本俊詔・松浦司 (2009)『学歴格差の経済学』勁草書房
- 文部科学省 (2020)『令和2年度学校基本調査』
- 安井健悟 (2019)「大学と大学院の専攻の賃金プレミアム」『内閣府経済社会総合研究所経済分析』第199号, pp.42-67.
- 山本耕平・安井大輔・織田暁子 (2015)「理系の誰が高収入なのか：SSM2005 データにもとづく文系・理系の年収比較」『京都社会学年報』23, pp.35-53.
- 山本耕平・安井大輔 (2016)「大卒女性における専攻間賃金格差の分析：理工系出身女性の賃金抑制要因に着目して」『ソシオロジ』61(1), pp.63-81.
- Arcidiacono, P. (2004), “Ability sorting and the returns to college major”, *Journal of Econometrics*, 121, pp.43-375.
- Becker G.S. (1962), “Investment in human capital: a theoretical analysis”, *Journal of political economy*, 70(5), pp.9-49.
- Becker, G.S. (1980), *Human Capital*, Chicago: University of Chicago Press.
- Black D.A., Haviland A.M., Sanders S.G. and Taylor L.J. (2007), “Gender wage disparities among the highly educated”, *Journal of Human Resources*, 48(3), pp.630-659.
- Blinder, Alan S. (1973), “Wage Discrimination: Reduced Form and Structural Estimates”, *Journal of Human Resources*, 8(4), pp.436-455.
- Blundell, R., Dearden L., Goodman, A. and Reed, H. (2000), “The returns to higher education in Britain: Evidence from a British cohort”, *Economic Journal*, 110, pp.82-99.
- Bratti, M. and Mancini, L. (2003), “Differences in early occupational earnings of UK male graduates by degree subject: evidence from the 1980-1993 USR”, IZA Discussion Paper, No. 890.
- Brown, C. and Corcoran, M. (1997), “Sex-based differences in school content and the male-female wage gap”, *Journal of Labor Economics*, 14(3), pp.431-465.
- Diagne, A. and Diene, B. (2011), “Estimating returns to higher education: A survey of models, methods and empirical evidence”, *Journal of African Economies*, 20(3), pp.80-133.
- Furno, M. (2014), “Returns to education and gender gap”, *International Review of Applied Economics*, 28(5), pp.628-649.
- James, E., Alsalam, N., Conaty, J.C. and To, D.L. (1989), “College quality and future earnings: Where should you send your child to college?”, *American Economic Review*, 79(2), pp.247-252.
- Lin, E.S. (2010), “Gender wage gaps by college major in Taiwan: Empirical evidence from the 1997-2003 manpower utilization survey”, *Economics of Education Review*, 29, pp.156-164.
- Loury, L.D. (1997), “The gender gap among college-educated workers”, *Industrial and Labor Relations Review*, 50(4), pp.580-593.
- Loury, L.D. and Garman, D. (1995), “College selectivity and earnings”, *Journal of Labor Economics*, 13(2), pp.289-308.

- Oaxaca, Ronald L. (1973), "Male-Female Wage Differentials in Urban Labor Markets", *International Economic Review*, 14(3), pp.693-709.
- Peet, E.D., Fink, G., and Fawzi, W. (2015), "Returns to education in developing countries: Evidence from the living standards and measurement study surveys", *Economics of Education Review*, 49, pp.69-90.
- Vikesh, A. (2011), "Returns to education: evidence from UK Twins", *The American Economic Review*, Vol. 101(4), pp.1629-1635.

Major field of study and gender earnings gap among highly educated employees in Japan

Mamiko TAKEUCHI*

Summary

This paper examines the relationship between the major field of study among highly educated employees holding an undergraduate degree or above and the gender earnings gap using a panel data set from 2004 to 2017. First, based on the estimated earnings function by gender, the wage premium was confirmed in the order of medicine or dentistry, pharmacy, and engineering in males based on humanities, and the wage premium was confirmed in the order of physics, medicine or dentistry, engineering, and pharmacy in females. As a result of factor decomposition of each earnings function, it became clear that the low ratio of women in the engineering field positively contributed to the earnings gap, while the difference in evaluation (coefficient) in the engineering field contributed to a reduction in the earnings gap. In order to examine the gender earnings gap among men and women who majored in the same field, we examined the effect of a female dummy variable on the estimation by major field of study. The negative effect of the female dummy variable was confirmed in fields other than natural science such as humanities, education, agriculture, and social sciences in consecutive order.

Key Words

higher educated employee, major field of study, gender earnings gap

* Faculty of Economics, Aichi Gakuin University
e-mail : take@dpc.agu.ac.jp

経済統計学会第64回（2020年度）全国研究大会・会員総会に関する追記

上藤 一郎（常任理事長）

全国研究大会と会員総会は、雑誌『統計学』の発行と並び本学会の最も重要な活動であるが、周知のように、2020年度については、新型コロナウイルス蔓延の影響を考慮して対面による実施が中止された。この間、本部事務局として会務を執行する常任理事会では、2020年度の全国研究大会・会員総会及び関連事項の代替措置案を策定し、理事会の承認を得て、これら一連の措置案を実施してきた。一方、会員に対しては、代替措置案の報告や実施案内等の情報を、逐次、学会ホームページ、学会ニュースレター、東北・関東支部メーリングリスト、関西支部メーリングリスト等を通じて周知してきた。

その結果、全国研究大会については、報告集報告とWeb報告という形で進めることができ、また会員総会についても、必要最小限の事案に限定したとは言え、郵送方式によって実施し、すべての事案について会員の承認を得ることができた。そこで、今後の本学会の取り組みの参考となるよう、備忘録として、改めて2020年度の全国研究大会・会員総会に関する代替措置の概要とその経緯を本誌で追記する。

なお付言すれば、例年とは異なるこのような代替措置案を滞りなく実施できたのは、偏に会員各位の協力と関西支部並びに理事会の助力によるものである。常任理事会を代表して深く感謝する次第である。

1. 全国研究大会代替措置案

(1) 開催までの経緯

従来のスケジュールであれば、例年9月に開催される会員総会で新年度のプログラム委員会構成が承認され、次年度の全国研究大会に向けて活動が開始される。2020年度についても、当初は例年どおり、このスケジュールに従い、2019年12月16付のニュースレターNo. 56でセッション募集が開始され（2020年1月末日締め切り）、2020年4月3日付のニュースレターNo. 57でセッション設置の報告と報告者募集（2020年5月8日締め切り）が行われた。

しかしながら、新型コロナウイルスの影響により、常任理事会では、9月に神戸大学で予定していた対面による全国研究大会の開催は困難であると判断し、急遽、代替措置案を検討した。その後、常任理事会内で意見を集約し合意を得た代替措置案は、メール審議により2020年6月20日に理事会で承認され、学会ホームページ、東北・関東支部と関西支部のメーリングリスト、2020年7月5日付のニュースレター臨時増刊号を通じて全会員に報告された。

(2) 報告者募集とプログラム編成

① 報告者募集

「ニュースレター」No. 57（2010.4）の報告募集記事に記載されたセッション設置案に従って報告を募集した。（応募様式にはWeb報告の希望の有無も記載）

② プログラムの編成

プログラム案の構成等については、第64回全国研究大会実行委員会（開催支部）及びプログラム委員会が担当した。

(3) 『報告集』について

- ① 研究報告に代わるものとして『全国研究大会報告要旨集』の特別版を作成した。
- ② 名称は『第64回全国研究大会報告集』とした。
- ③ 従来の『報告要旨集』とは以下の点で差別化を図った。
 - ・ ページ数は1報告当たり2～4頁とした。
 - ・ 課題や仮説の開示だけではなく報告の結論を必ず含めた論述形式として構成とした。
 - ・ このため報告集の雛形として例文を作成した。
 - ・ オーガナイザーは1頁程度の「セッション解題」を作成した。
- ④ 報告集原稿は、セッションについてはオーガナイザー宛、自由論題はプログラム委員長宛に提出した。
- ⑤ オーガナイザー（プログラム委員長）は、提出された原稿の構成についての形式チェックを行った。
- ⑥ 提出原稿の配列は、作成されたプログラムに従うものとされた。
- ⑦ 公刊された報告については、学会が認めた正式な報告とし、希望に応じて学会長名で「報告証明書」を発行することとした。
- ⑧ 『報告集』は、個人会員（1部）団体A会員（5部）、B会員（2部）送付した。
- ⑨ 報告者からの報告掲載料は徴収しなかった。
- ⑩ 『報告集』は、公刊後すみやかに学会のHP上で公開した。
- ⑪ 上記の作業は、第64回全国研究大会実行委員会（開催支部）の責任において行われた。

(4) Web 報告について

- ① セッションのWeb報告のプログラム案（開催日時を含む）は、学会のHP上に掲載された。
- ② Web報告への参加希望者は、プログラム委員会にその旨をメールにより連絡を行うこととした。
- ③ 各セッションのWeb報告の開始については座長が執り行った。

なお、Web報告のアカウント設定、ミーティングの設定等については、本部事務局が執り行った。

(5) 財政措置

『報告集』やWeb会議ソフトの登録料等、今回の代替措置に関わる全ての経費は学会が負担した。従って、通常の全国研究大会のように参加費は徴収しなかった。

(6) 代替措置実施のスケジュール

日程	内容
2020年 7月15日～8月5日	報告者募集
8月20日	プログラム配布 (HP掲載)
9月11日	Web報告会のプログラムの開示 (学会HP)
9月12日～10月11日	Web報告参加者募集
11月7～8日	Web報告会開催 (セッション別)
12月10日	『報告集』提出締切り (セッションオーガナイザー又はプログラム委員長へ提出)
12月25日	オーガナイザーからの原稿最終提出 (プログラム委員長宛)
2021年 1月末	『報告集』刊行, 会員への配布

2. 会員総会代替措置案

(1) 理事会

① 理事会 (6月開催・メール審議)

議題 (承認事項)

- ・大会開催代替措置 (理事会承認後直ちに実行に移す)
- ・学会長を含む事務局体制の任期延長に関する承認手続きの件

② 理事会 (9月開催・Web会議)

議題 (承認事項)

- ・学会長を含む事務局体制の任期延長
- ・会計報告
- ・予算案の承認
- ・入退会の承認
- ・2022年度研究大会開催支部の決定

(2) 会員総会

① 理事会 (9月開催) による承認を経て, 2020年度会員総会を郵送方式で実施した。

議題 (承認事項)

- ・学会長任期の延長
- ・2019年度決算報告の承認
- ・2020年度の中間報告と全国研究大会代替案の経費に関する補正予算案の承認
- ・2021年度予算案の承認
- ・2022年大会開催支部の決定

② 投票葉書の経過と結果

- ・2020年11月30日 (当日消印有効) を締切日とした。
- ・2020年12月7日, 静岡大学経済統計学研究室で, 学会長立ち合いのもと, 常任理事長が開票・集計作業を行った。

- ・すべての事案について承認された。

(3) 関連事項

① 会長，常任理事長，常任理事の任期1年延長の件

本年度は，会長，常任理事長，常任理事（渉外担当，NL・HP担当，無任所の各常任理事）の交代期にあたる。しかし，全国的な移動制限等も懸念される中，対面での全国研究大会・会員総会開催を中止することになった。このため，執行部の円滑な事務引継ぎが難しい状況にあること等の事情に鑑み，現行の本部事務局体制を1年間延期することを2020年6月の理事会に提案し承認された（任期は2020年9月～2021年9月）。但し，業務負担の重さから，編集委員長担当の常任理事及びプログラム委員長担当の常任理事については，従来どおり交代することとした。なお，本事案についても郵送方式の会員総会で承認された。

② 緊急事態への対応の件

今回のような緊急事態への対応については，本会の会則等に明文規定はないが，「経済統計学会内規」の「4. 理事会について(4)その他会の運営にかかわる事項」に基づき，理事会に本事案を諮ることとし，特に会長職の任期の延長については「経済統計学会内規」により全国総会の承認事項とされていることから，会員による郵送での承認を得ることとし承認された。

支部だより (2020年4月～2021年3月)

北海道支部

今年度の北海道支部例会は新型コロナウイルス感染症拡大等のため、開催することを見合わせた。

(水野谷武志 記)

東北・関東支部

下記の通り、支部例会（リモート形式）が開催されました。

日時：2020年6月6日(土) 16:00～18:00

場所：Zoom ミーティング

報告：

1. 山野紀彦氏（非会員，国際協力開発機構（OECD））
OECDにおける産業統計を用いたSUT・IOTデータベースの整備
2. 萩野 覚会員（総務省），田原慎二氏（千葉商科大学），渡邊淳司氏（岡山県立大学）
供給使用表で組み入れるべき企業の異質性
3. 櫻本 健会員（立教大学）
COVID-19によって変化が求められる日本の統計機構

日時：2020年12月5日(土) 13:30～16:00

場所：Zoom ミーティング

報告：

1. 櫻本 健会員（立教大学）
経済統計分野でオンラインでできるフリーの教育ツールの紹介
— V-RESAS, jSTATMAP, Google による COVID-19 感染陽性者予測

日時：2021年1月9日(土) 13:30～15:30

場所：Zoom ミーティング

報告：

1. 高部 勲会員（統計データ利活用センター）
公的統計マイクロデータの更なる利活用推進に向けた取組の紹介
2. 伍賀一道会員（金沢大学名誉教授）
コロナ禍における労働力調査に関する疑問 — 調査結果が示す労働市場の姿と現実

(坂田幸繁 記)

関西支部

下記の通り，支部例会（リモート形式）が開催されました。

日時：2020年7月18日(土) 16:00～18:00

場所：Zoom ミーティング

報告：

1. 齋藤立滋（大阪産業大学）
在宅医療の推進は医療費を減らせるか
2. 泉 弘志（関西支部）・戴 艶娟（広東外語外貿大学）・李 潔（埼玉大学）
国際価値の理論と国際産業連関表による各国剰余価値率の計測

（村上雅俊 記）

九州支部

今年度の九州支部例会は新型コロナウイルス感染症拡大等のため，開催することを見合わせました。

（松川太一郎 記）

機関誌『統計学』投稿規程

経済統計学会（以下、本会）会則第3条に定める事業として、『統計学』（電子媒体を含む。以下、本誌）は原則として年に2回（9月、3月）発行される。本誌の編集は「経済統計学会編集委員会規程」（以下、委員会規程）にもとづき、編集委員会が行う。投稿は一般投稿と編集委員会による執筆依頼によるものとし、いずれの場合も原則として、本投稿規程にしたがって処理される。

1. 総則

1-1 投稿者

会員（資格停止会員を除く）は本誌に投稿することができる。

1-2 非会員の投稿

- (1) 原稿が複数の執筆者による場合、筆頭執筆者は本会会員でなければならない。
- (2) 常任理事会と協議の上、編集委員会は非会員に投稿を依頼することができる。
- (3) 本誌に投稿する非会員は、本投稿規程に同意したものとみなす。

1-3 未発表

投稿は未発表ないし他に公表予定のない原稿に限る。

1-4 投稿の採否

投稿の採否は、審査の結果にもとづき、編集委員会が決定する。その際、編集委員会は原稿の訂正を求めることがある。

1-5 執筆要綱

原稿作成には本会執筆要綱にしたがう。

2. 記事の分類

2-1 研究論文

以下のいずれかに該当するもの。

- (a) 統計およびそれに関連した分野において、新知見を含む会員の独創的な研究成果をまとめたもの。
- (b) 学術的な新規性を有し、今後の研究の発展可能性を期待できるもので、速やかな成果の公表を目的とするもの。

2-2 報告論文

研究論文に準じる内容で、研究成果の速やかな報告をとくに目的とする。

2-3 書評

統計関連図書や会員の著書などの紹介・批評。

2-4 資料

各種統計の紹介・解題や会員が行った調査や統計についての記録など。

2-5 フォーラム

本会の運営方法や統計、統計学の諸問題にたいする意見・批判・反論など。

2-6 海外統計事情

諸外国の統計や学会などについての報告。

2-7 その他

全国研究大会・会員総会記事、支部だより、その他本会の目的を達成するために有益と

思われる記事。

3. 原稿の提出

3-1 投稿

原稿の投稿は常時受け付ける。

3-2 原稿の送付

原則として、原稿は執筆者情報を匿名化したPDFファイルを電子メールに添付して編集委員長へ送付する。なお、ファイルは『統計学』の印刷レイアウトに準じたPDFファイルであることが望ましい。

3-3 原稿の返却

投稿された原稿（電子媒体を含む）は、一切返却しない。

3-4 校正

著者校正は初校のみとし、大幅な変更は認めない。初校は速やかに校正し期限までに返送するものとする。

3-5 投稿などにかかわる費用

- (1) 投稿料は徴収しない。
- (2) 掲載原稿の全部もしくは一部について電子媒体が提出されない場合、編集委員会は製版にかかる経費を執筆者（複数の場合には筆頭執筆者）に請求することができる。
- (3) 別刷は、研究論文、報告論文については30部までを無料とし、それ以外は実費を徴収する。
- (4) 3-4項にもかかわらず、原稿に大幅な変更が加えられた場合、編集委員会は掲載の留保または実費の徴収などを行うことがある。
- (5) 非会員を共同執筆者とする投稿原稿が掲載された場合、その投稿が編集委員会の依頼によるときを除いて、当該非会員は年会費の半額を掲載料として、本会に納入しなければならない。

3-6 掲載証明

掲載が決定した原稿の「受理証明書」は学会長が交付する。

4. 著作権

4-1 本誌の著作権は本会に帰属する。

4-2 本誌に掲載された記事の発行時に会員であった執筆者もしくはその遺族がその単著記事を転載するときには、出所を明示するものとする。また、その共同執筆記事の転載を希望する場合には、他の執筆者もしくはその遺族の同意を得て、所定の書面によって本会に申し出なければならない。

4-3 前項の規定にもかかわらず、共同執筆者もしくはその遺族が所在不明のため、もしくは正当な理由によりその同意を得られない場合には、本会が承認するものとする。

4-4 執筆者もしくはその遺族以外の者が転載を希望する場合には、所定の書面によって本会に願い出て、承認を得なければならない。

4-5 4-4項にもとづく転載にあたって、本会は転載料を徴収することができる。

4-6 会員あるいは本誌に掲載された記事の発行時に会員であった執筆者が記事をウェブ転載するときには、所定の書類によって本会に申し出なければならない。なお、執筆者が所属する機関によるウェブ転載申請については、本人の転載同意書を添付するものとする。

- 4-7 会員以外の者，機関等によるウェブ転載申請については，前号を準用するものとする。
- 4-8 転載を希望する記事の発行時に，その執筆者が非会員の場合には，4-4，4-5項を準用する。
1997年7月27日制定（2001年9月18日，2004年9月12日，2006年9月16日，2007年9月15日，2009年9月5日，2012年9月13日，2016年9月12日一部改正）

機関誌『統計学』の編集・発行について

『統計学』編集委員会

みなさまからの投稿を募集しています。ぜひ研究成果の本誌上での発表をご検討ください。

1. 原稿は編集委員長宛に送付して下さい(下記メールアドレス)。
2. 投稿は常時受け付けています。
なお、書評、資料および海外統計事情等の分類の記事については調整が必要になることもありますので念のため事前に編集委員長に照会して下さいをお願いします。
3. 次号以降の発行予定日は次のとおりです。
第121号：2021年9月30日
第122号：2022年3月31日
4. 原則として、すべての投稿が審査の対象となります。投稿に際しては、「投稿規程」、「執筆要綱」、および「査読要領」の確認をお願いします。最新版は、本学会の公式ウェブサイト (<http://www.jsest.jp/>) を参照して下さい。
5. 編集委員会は2021年4月から次の体制となります。引続きよろしくをお願いします。
2021年度編集委員会委員長 村上雅俊(関西)
同副委員長 佐藤智秋(東北・関東)
同委員 水野谷武志(北海道)、山口幸三(東北・関東)、西村善博(九州)

投稿、編集委員会についての問い合わせや執筆の推薦その他とも、下記編集委員長のメールアドレス宛に送付して下さい。

editorial@jsest.jp

編集後記

2020年度の日常は新型コロナ発生前とは大きく変わりました。そのような中でも『統計学』の投稿者のみなさま、そしてお忙しい中快く論文の審査をお引き受けいただきました査読者のみなさまに改めてお礼申し上げます。副編集委員長の村上先生をはじめ編集委員の水野谷先生、山田先生、松川先生には、大変お世話になりました。また、『統計学』創刊60周年記念事業委員会は特集の編集ありがとうございました。(小林良行 記)

執筆者紹介

小林良行（総務省統計研究研修所） 武内真美子（愛知学院大学経済学部）

支部名

事務局

北海道	062-8605 札幌市豊平区旭町 4-1-40 北海学園大学経済学部 (011-841-1161) mizunoya@econ.hokkai-s-u.ac.jp	水野谷武志
東北・関東	192-0393 八王子市東中野 742-1 中央大学経済学部 (042-674-3421) ysakata@tamacc.chuo-u.ac.jp	坂田幸繁(代行)
関西	580-8502 松原市天美東 5-4-33 阪南大学経済学部 (072-332-1224) m-murakami@hannan-u.ac.jp	村上雅俊
九州	890-0065 鹿児島市郡元 1-21-30 鹿児島大学法文学部 (099-285-7601) matsukawa@leh.kagoshima-u.ac.jp	松川太一郎

『統計学』編集委員

委員長 小林良行（東北・関東，総務省統計研究研修所）
副委員長 村上雅俊（関西，阪南大学）
委員 水野谷武志（北海道，北海学園大学），山田 満（東北・関東），
松川太一郎（九州，鹿児島大学）

『統計学』60周年記念事業委員会

委員長 大井達雄（和歌山大学）
副委員長 水野谷武志（北海学園大学）
委員 池田 伸（立命館大学），伊藤伸介（中央大学），
杉橋やよい（専修大学），村上雅俊（阪南大学），
金子治平（会長，神戸大学），上藤一郎（常任理事長，静岡大学）

統計学 No.120

定価 1,760円(本体1,600円)

2021年3月31日 発行	発行所	経済統計学会 〒112-0013 東京都文京区音羽1-6-9 音羽リスマチック株式会社 TEL/FAX 03(3945)3227 E-mail: office@jsest.jp http://www.jsest.jp/
	発行人	代表者 金子治平
	発売所	音羽リスマチック株式会社 〒112-0013 東京都文京区音羽1-6-9 TEL/FAX 03(3945)3227 E-mail: otorisu@jupiter.ocn.ne.jp 代表者 遠藤 誠

Statistics

No. 120

2021 March

Special Section: The 60th Anniversary of the Journal

Special Topic B: Methodological Perspectives in the Creation and Release of Official Micro-data

Potentiality of Statistical Metadata Archives in the Official Statistics of Japan

..... Yoshiyuki KOBAYASHI (1)

Articles

Major field of study and gender earnings gap among highly educated employees in Japan

..... Mamiko TAKEUCHI (19)

JSES Activities

Postscript on the 64th Session of the JSES (35)

Activities within JSES Branches (39)

Prospects for the Contribution to *Statistics* (41)

Japan Society of Economic Statistics
