

パネルデータにおける脱落特性とサンプルの代表性の検証 「結婚と家族に関する国際比較パネル調査」の経験

国立社会保障・人口問題研究所
国際関係部 研究員 中川 雅貴

1. はじめに

パネル調査は、同一の調査対象を継続的に追跡することで、その意識や行動の変化のパターンや要因を捉える動学的分析を行うとともに、調査対象ごとの観察されない異質性を考慮した精緻な分析を行ううえで必要となる情報量の多いデータを得られるという利点がある。一方で、とりわけ個人を対象とするパネル調査においては、継続調査への協力拒否を含む何らかの理由によって生じるサンプル脱落の問題が存在する。こうしたサンプル脱落の生起は、標本調査の根幹となるサンプルの代表性を損なうとともに、分析の対象となる意識や行動とサンプルの脱落が独立でない場合には、そのデータによって得られた推定結果に深刻なバイアスをもたらす。

近年では、データ分析に必要なコンピュータの計算処理能力が向上するとともに、パーソナル・コンピュータ用のデータ解析ソフトが比較的安価に入手できるようになったこともあり、日本国内においても、経済学や社会学、人口学といった分野の実証的な研究では、パネルデータ分析によって得られた新たな知見が蓄積されている。しかし一方で、情報量の多いパネルデータセットの構築に求められる調査の大規模化および長期化に伴い、こうした社会科学の諸分野における実証研究に取り組むに際しては、それぞれの研究者による個人研究というよりはプロジェクト化された大規模な共

同研究に参加してデータを共有したり、あるいは、個別の研究者による個別の研究関心に基づく分析を行う場合であっても、すでに公開されているデータセットの二次利用に依存する傾向が顕著になりつつある。すなわち、実証研究を専門とする研究者であっても、分析に用いるデータセットの構築や整備に実際に関わる機会はますます減少しており、調査の実施やデータの構造に関する上述の諸問題についての関心や理解が薄れていると考えられる。とりわけ、日本国内におけるパネルデータの分析に関しては、データ蓄積の歴史も浅く、利用実績も限られていることから、調査としてどのような問題が発生しているのか、また、こうした実証分析における基本的な問題が、データの構造や分析結果に与える影響については十分に理解されていないことが指摘されている（北村 2005）。

本稿は、少子・高齢化と密接に関連する家族変動を「世代」と「ジェンダー」という二つの視点から社会人口学的に分析する目的で実施されている「世代とジェンダーに関する国際比較プロジェクト(JGGP)」の経験をもとに、上述のパネル調査の実施に関する基本的な問題およびそれらがデータの構造に与える影響についての一事例を報告することを主たる目的とする。具体的には、本調査研究プロジェクトにおいて2004年・2007年・2010年に実施された「結婚と家族に関する国際比較パネル調査」の回収

結果に基づき、サンプル脱落の特徴を把握するとともに、調査時点間におけるサンプル脱落が、層化二段無作為抽出法によって得られた確率サンプルの代表性に与える影響について検証する。なお、本稿で紹介する調査データは、後述の GGP 国際コンソーシアムで手続きをすれば利用できるようになった。本稿が、「結婚と家族に関する国際比較パネル調査」データを利用するうえでの参考となれば幸いである。

2 「結婚と家族に関する国際比較パネル調査」(JGGP)の概要

JGGP は、国連ヨーロッパ経済委員会(UNECE)人口部が中心になって組織し、20 近くの先進国が共同して行う国際比較研究である「世代とジェンダー・プロジェクト(Generations & Gender Project)」に参加し、日本の少子化を世代(親子・世代関係)とジェンダー(夫婦・パートナー関係)という二つの視点から国際比較分析することを目的として実施されている¹。国際比較調査研究である本プロジェクトは、プロジェクトに参加する各国の研究チームが、GGP 国際コンソーシアム²によって定められた共通の調査票を用いたパネル調査を複数回実施し、比較可能な共通のフレームでミクロ・データを収集している。先進国が共通に直面する少子・高齢化と密接に関連する家族変動を、「世代」と「ジェンダー」という二つの視点から分析する GGP プロジェクトの発足に際しては、UNECE 域外の主要な先進国である日本にも参加が呼び掛けられ、当時、国立社会保障・人口問題研究所に在籍していた研究者を中心に、GGP の日本版である「結婚と家族に関する国際比較パネル調査: JGGP」研究チームが 2002 年に立ち上げられた。

JGGP 研究チームでは、国際コンソーシアムの指針にしたがい、2004 年から 2010 年まで、3 年ごとに計 3 回にわたるパネル調査を実施してきた。同時に、法制度改革や行政統計データを含むマクロレベルでのコンテキスト・データの整備を行い、マクロのコンテキスト・データとミクロのパネルデータを有機的かつ多面的にリンクさせながら、他の先進国との国際比較という広い視野から、日本における未婚化・少子化の原因の分析と有効な対策について研究成果を発表してきた。とりわけ、GGP 国際コンソーシアムの指針にもとづく共通の調査票を用いて実施された「結婚と家族に関する国際比較パネル調査」は、パートナーシップ、出生行動、家族関係、ジェンダー、高齢者ケア、ソーシャル・ネットワーク、家計と社会保障に関する調査項目を含み、この分野ではおそらく日本では初めての包括的なパネル調査といえる。以下では、2004 年に開始された「結婚と家族に関する国際比較パネル調査」の実査方法および経緯について概説する。なお、2004 年、2007 年、2010 年の計 3 回にわたって実施された調査の実査は、いずれも、(社)新情報センターを調査委託実施機関として行われた。

2004 年に実施された JGGP の第一回調査は、国勢調査(2000 年)の調査区を基に無作為に抽出された 530 地点から層化二段確率サンプルによって抽出された 18-69 歳の日本人男女 15,000 人を対象として訪問留置法による調査が行われた。2004 年 2 月～3 月における第一次調査の結果、全調査対象者の 60.5%にあたる 9,074 名(男性 4,265 名、女性 4,809 名)から有効回答を得た。3 年後の第二次調査では、未婚化・少子化の背景および構造に関するさらなる理解と政策提言に資することを中心的な研究課題

と位置づけ、調査対象者を第一次調査時点に49歳以下であった回答者4,568名(男性2,058名、女性2,510名)に限定し、フォローアップ調査を実施した。2007年2月から5月にかけて実施された第二次調査の結果、3,082名(男性1,306名、女性1,776名)から有効回答を得た。さらに2010年の第三次調査においては、上記のフォローアップ調査において有効回答が得られた若壮年調査対象者(以下、「若壮年サンプル」とする)の第三次調査に加えて、2007年の調査では調査対象とならなかった「高齢者サンプル」の第二次調査を6年ぶりに実施し、若壮年サンプル2,306名の三時点パネルデータおよび高齢者サンプル2,611名の二時点パネルデータを構築した。

これまでJGGPプロジェクトでは、おもに2007年調査完了時点において構築された若壮年サンプル2時点パネルデータを用い、結婚や同棲、出産と子育ておよび就業などにおけるジェンダー関係、そして親世代と子世代の居住形態や交流・支援関係を中心とした世代間関係について、パネルデータ分析というツールと国際比較という視座からの新たな知見を発表してきた³。しかしながら、日本における未婚化・少子化についての社会人口学的分析テーマに関するパネルデータという希少性をもつJGGPデータにおいても、パネル調査一般が抱えるパネルの脱落によるサンプル消耗(sample attrition)や標本と母集団の関係性の歪みといった問題が存在することは否めない。また、JGGPプロジェクトにおける主たる分析テーマの一つとなっているパートナーシップや出生行動といったライフコース・イベントは、パネル調査におけるサンプル脱落確率と密接な関係があることが指摘されており、内生変数によってサンプルが脱落

する場合には、そのデータによって得られた推計結果が深刻なバイアスをもつという問題にも注意が払われる必要がある⁴。

以下では、JGGP「若壮年サンプル」および「高齢者サンプル」それぞれについて、標本を構成する基本的な属性である地域・性別・年齢の分布に関して、サンプルの代表性が維持されているか否かについて検証する。つづいて、配偶関係や婚姻関係や居住形態といった基本的な社会人口学的属性とサンプル脱落との関係を検証し、サンプル脱落の特徴の把握を試みる。

3 JGGP 若壮年調査における脱落状況とその傾向

表1は、JGGP調査における若壮年サンプルの基本的属性別にみた有効回収数の推移ならびに累積脱落率を示している。前述のとおり、若壮年サンプルについては、2004年の第一次調査における有効回収数5,013件のうち、3年後の2007年に実施された第二次調査において追跡データが得られたのは3,082件であり、この3年間に38.5%のパネル脱落が発生している。また、つづく2010年の第三次調査が完了したサンプル数は2,306件となっており、2004年の第一次調査からの累積脱落率は50%を越えている。

性別でみた累積脱落率は、女性が50%に対して男性は58%と高くなっており、この結果、若壮年サンプルの性別分布は、一次調査完了時点では男性:45%、女性:55%だったのに対し、2010年の第三次調査完了時点では、41%対59%と、男女比率の差が拡大している。年齢別にみた若壮年パネルの脱落状況については、調査開始時点での年齢が若い層ほど脱落率が高くなる傾向が確認され、「25歳未満」のサンプルの累積脱落率

が 70% 近くに達しているのをはじめ、「25-29 歳」の年齢層では 64% のサンプルにおいて 3 次調査が完了していない。一方で、「40-44 歳」および「45 歳以上」の年齢層

においては、累積脱落率がそれぞれ約 44% に留まっており、若壮年サンプルの平均累積脱落率 (54%) を 10 ポイント下回っている。

表 1 JGGP 「若壮年サンプル」における基本的属性別にみたパネル脱落状況

	第一次調査(2004年) における分布	第二次調査(2007年) における分布	第三次調査(2010年) における分布	第三次調査 完了時点での 累積脱落率
有効回収件数	5,013	3,082	2,306	若壮年サンプル計: 54.0%
性別				**
男性	45.1%	42.4%	40.8%	58.4%
女性	54.9%	57.6%	59.2%	50.3%
年齢(2004年時点)				
25歳未満	17.0%	14.0%	11.4%	69.2% **
25 - 29歳	12.8%	11.0%	10.0%	64.0% **
30 - 34歳	16.2%	16.1%	15.9%	54.8%
35 - 39歳	18.1%	18.9%	19.0%	51.7%
40 - 44歳	17.1%	18.8%	20.9%	43.7% **
45歳以上	18.8%	21.2%	22.8%	44.4% **
都市規模別				
大都市	23.6%	22.2%	21.0%	59.1% **
中都市(人口10万人以上)	39.3%	38.1%	36.8%	57.0% **
小都市(人口10万人未満)	18.1%	18.1%	18.9%	51.9%
町村	19.0%	21.6%	23.3%	43.4% **
地域別				
北海道	4.9%	5.1%	5.1%	52.8%
東北	8.3%	8.8%	8.7%	51.3%
関東	30.4%	28.3%	27.8%	58.1% **
北陸	4.1%	5.3%	6.0%	32.9% **
東山	4.4%	4.6%	4.7%	50.7%
東海	11.0%	11.4%	11.0%	53.7%
近畿	15.9%	15.0%	15.1%	56.4%
中国	5.4%	5.2%	5.5%	53.2%
四国	3.1%	3.4%	3.4%	49.7%
北九州	8.2%	8.9%	8.8%	50.6%
南九州	4.3%	4.2%	3.9%	57.7%

* $p < 0.05$, ** $p < 0.01$

都市規模別および地域別の脱落状況については、人口規模の大きい調査地区に属するサンプルにおいてパネル脱落率が高くなる傾向が確認できる。都市規模別では、「大都市」グループ(東京 23 区・横浜市・川崎市・京都市・千葉市・大阪市・北九州市・その他政令指定都市)において、一次調査完了時点から三次調査完了までの累積脱落

率が 60% 近くになっており、つづく「中規模都市」グループ(人口 10 万人以上)においても累積脱落率が 57% となっている。一方で、「町村」グループにおける累積脱落率は 43% に留まっており、若壮年サンプルの平均を 10 ポイント以上下回っている。また、地域別では、「関東」(58%)、「南九州」(58%)、「近畿」(56%) において高い脱落率が確認

される一方で、「北陸」では33%と極端に低くなっている。

パネル脱落によるサンプルの代表性への影響をより詳細に吟味するために、表2ではJGGP 標本抽出の第一次層ごとの脱落状況を示した。前述のとおり、層化二段無作為抽出法によって実施されたJGGP 調査は、全国の市町村を「地域」と「都市規模」によって47層に分け、これを標本抽出の一次層としている。表2に示される調査地区(標本抽出第一次層)ごとのパネル脱落状況を

みると、前掲表において示唆されるとおり、「関東」の「大都市」において累積脱落率が高くなっているのが確認される。また、「地域」別でみた場合に「関東」に次いで累積脱落率が高くなっている「南九州」ブロックにおいては、「大都市」カテゴリーに属する調査地区が割り振られていないものの、「人口10万人以上の都市」カテゴリーにおいて、68%という高い確率でパネル脱落が発生している。

表2 JGGP「若壮年サンプル」調査地区(標本抽出第一次層)別脱落率

	大 都 市					人口10万人以上の都市	人口10万未満の市	郡部	計
	東京23区	横浜市	川崎市・京都市	千葉市・大阪市・北九州市	その他の政令市				
北海道					8	7	4	5	24
					84	78	32	54	248
					60.7%	53.8%	50.0%	40.7%	52.8%
東北					5	12	10	13	40
					42	135	112	128	417
					47.6%	57.8%	52.7%	44.5%	51.3%
関東	36	15	6	4	5	68	23	19	176
	303	138	45	24	34	633	194	155	1,526
	64.4%	56.5%	57.8%	54.2%	70.6%	58.9%	52.6%	48.4%	58.1%
北陸						9	7	7	23
						84	53	70	207
						40.5%	28.3%	27.1%	32.9%
東山						7	6	7	20
						75	76	68	219
						56.0%	57.9%	36.8%	50.7%
東海					9	24	10	10	53
					108	255	93	93	549
					57.4%	53.7%	48.4%	54.8%	53.7%
近畿			6	10	6	39	13	11	85
			52	107	60	345	130	105	799
			53.8%	60.7%	55.0%	58.6%	55.4%	48.6%	56.4%
中国					5	13	6	7	31
					48	124	42	55	269
					45.8%	56.5%	61.9%	45.5%	53.2%
四国						7	4	6	17
						57	39	59	155
						64.9%	51.3%	33.9%	49.7%
北九州				4	6	9	7	11	37
				50	89	111	57	102	409
				56.0%	61.8%	52.3%	45.6%	39.2%	50.6%
南九州						9	7	7	23
						75	79	61	215
						68.0%	58.2%	44.3%	57.7%
合計	36	15	12	18	44	204	97	103	529
	303	138	97	181	465	1,972	907	950	5,013
	64.4%	56.5%	55.7%	58.6%	57.4%	57.0%	51.9%	43.4%	54.0%

上段：調査地点数 / 中段：第一次調査有効回収数 / 下段：累積脱落率(2010年第3次調査終了時点)

一方、「地域」別でみた累積脱落率が33%と極端に低い「北陸」ブロックにおいては、とりわけ「人口10万未満の都市」(28%)および「郡部」(27%)といった小規模調査地区において、脱落率が極端に低くなっていることが確認できる。この傾向は、「地域」別の脱落率が「北陸」ブロックに次いで低い「四国」においても確認できる。とくに、「関東」ブロックにおいて最も低い「郡部」の脱落率(48%)が、「北陸」ブロックにおいて最も高い「人口10万以上の都市」(41%)を上回っているのは、注目に値する。

2004年時点での基本的な社会人口学的属性ごとのパネル脱落率(2010年第三次調査における累積脱落率)を概観すると、以下のような傾向がみられる(表3)。まず、配偶関係については、2004年時点で有配偶であった調査対象者のパネル脱落率が、男女ともに平均よりも低くなっていることが観察できる。また、第一次調査時点での「有子」「持ち家居住」といった属性についても、男女ともに、その後のパネル脱落率が平均よりも低くなることを確認できる。就業状態については、男性のみ、2004年時点で「就業中」であったサンプルの累積脱落率が有意に低くなっている。一方で、健康状態については、男女ともにパネル脱落との有意な関係は確認されない。

表3 JGGP「若壮年サンプル」における社会人口学的基本属性ごとの累積脱落率

	男女計	
	男	女
全サンプル	54.0%	50.3%
有配偶	47.3% **	44.5% **
有子	46.4% **	43.5% **
持ち家居住	46.7% **	43.0% **
就業中	53.6%	49.6%
健康状態良好	53.6%	49.9%

* $p < 0.05$, ** $p < 0.01$

4 JGGP 高齢者調査における脱落状況とその傾向

「高齢者サンプル」についても、まず標本を構成する基本的な属性である地域・性別・年齢の分布に関して、サンプルの代表性が維持されているか否かについて検証する。前述のとおり、高齢者調査については、2004年の一次調査から6年後の2010年に二次調査が行われた。2004年調査回答者4,061人のうち、6年後の2010年調査(2次調査)への回答が得られたのは、2,611人であり、2時点調査間における脱落件数は1,450件、脱落率は35.7%となっている。約6年間のインターバルがあるため、パネルの消耗が懸念されたが、比較的良好なパネル残存率を維持していると言える。

表4に示されるとおり、性別による二時点間の脱落率については、男性:37%、女性:35%と若干の違いがあり、2010年調査における有効回収(二時点パネル完了)の性別分布も48%対52%と、その差が若干拡大しているものの、性別でみた二時点間のパネル脱落率に統計的に有意な差は認められない。年齢別にみた「高齢者サンプル」のパネル脱落状況をみると、「若壮年サンプル」の脱落パターンとは逆に、年齢層が高いほどパネル脱落率が高くなる傾向が確認できる。とくに、2004年調査開始時点で「65歳以上」だったグループのパネル脱落率が40%を超えており、他の年齢グループと比較して顕著に高くなっている。この年齢グループに観察される高いパネル脱落率については、高齢の調査対象者に特有の要因、すなわち「死亡」によって追跡調査が不可能になったという要因も考慮される必要があるが、たとえば、第二次調査の前年2009年に実施されたパネルメンテナンス作業(ニュースレターの送付)において確認さ

れた「死亡件数」は8件にとどまっており、この年齢層の調査対象者のみに集中して発生しているわけではないことも確認されて

いる。こうしたことから、高齢期に特有の何らかの別の要因に、パネル脱落との因果関係を見出す必要があると考えられる。

表4 JGGP「高齢者サンプル」における基本的属性別にみたパネル脱落状況

	第一次調査(2004年) における分布	第二次調査(2010年) における分布	第二次調査(2010年) における脱落率
有効回収件数	4,061	2,611	高齢者サンプル脱落率 計:35.7%
性別			
男性	49.3%	48.4%	36.9%
女性	50.7%	51.6%	34.5%
年齢(2004年時点)			
55歳未満	26.1%	26.4%	34.9%
55 - 59歳	25.2%	25.9%	33.8%
60 - 64歳	25.5%	26.0%	34.3%
65歳以上	23.2%	21.6%	40.1% **
都市規模別			
大都市	20.9%	19.6%	39.6% **
中都市(人口10万人以上)	36.8%	36.0%	37.0%
小都市(人口10万人未満)	19.6%	20.3%	33.5%
町村	22.7%	24.1%	31.9% **
地域別			
北海道	5.9%	5.4%	40.8%
東北	9.4%	9.3%	36.4%
関東	29.5%	28.5%	37.7% *
北陸	5.3%	6.2%	24.7% **
東山	4.5%	4.6%	33.1%
東海	11.0%	10.7%	37.4%
近畿	13.7%	13.6%	36.2%
中国	6.0%	6.4%	31.6%
四国	2.9%	3.1%	32.8%
北九州	8.0%	8.3%	33.3%
南九州	3.9%	3.9%	36.1%

* $p < 0.05$, ** $p < 0.01$

一方、都市規模別および地域別の脱落状況については、「若壮年サンプル」の脱落パターンと同様に、人口規模の大きい調査地区に属するサンプルにおいてパネル脱落率が高くなる傾向が確認できる。また、「高齢者サンプル」の「町村」グループにおける累積脱落率は32%であるが、「若壮年サンプル」の脱落パターンと比較した場合に、都市規模別でみたパネル脱落率の違いが比較

的小さくなっている点が注目される。また、「北海道」「東北」「東海」地域における脱落率の高さも、「若壮年パネル」と比較した場合の「高齢者サンプル」の特徴として指摘できる。一方で、「北陸」ブロックにおける脱落率(25%)が、他のブロックと比較して極端に低くなっている点は、「若壮年サンプル」の脱落パターンと類似している。

「高齢者サンプル」についても、標本抽

出の第一次層ごとの脱落状況を表5に示した。ここでも、前掲表において示唆されるとおり、「関東」の「大都市」において累積脱落率が高くなっているのが確認される。「都市規模」と「パネル脱落率」の正の相関関係は、「北海道」「東海」「南九州」とい

ったブロック全体での脱落率が大きい地域において共通に確認される傾向である。一方で、「北陸」ブロックのような全体的に脱落率の低い地域においては、「都市規模」と「パネル脱落率」に関する明確な関係は確認できない。

表5 JGGP「高齢者サンプル」調査地区(標本抽出第一次層)別脱落率

	大都市					人口10万人以上の都市	人口10万未満の市	郡部	計
	東京23区	横浜市	川崎市・京都市	千葉市・大阪市・北九州市	その他の政令市				
北海道					8	7	4	5	24
					67	78	50	45	240
					44.8%	43.6%	34.0%	37.8%	40.8%
東北					5	12	10	14	41
					44	107	92	139	382
					38.6%	36.4%	37.0%	35.5%	36.4%
関東	26	15	5	4	5	68	23	19	146
	198	97	11	14	25	500	191	160	1,196
	41.4%	35.1%	54.5%	28.6%	56.0%	38.6%	34.0%	33.1%	37.7%
北陸						9	7	7	23
						83	63	69	215
						24.1%	25.4%	24.6%	24.7%
東山						7	6	7	20
						64	50	67	181
						39.1%	34.0%	26.9%	33.1%
東海					9	24	10	10	53
					111	178	74	83	446
					44.1%	37.1%	33.8%	32.5%	37.4%
近畿			6	11	6	36	13	11	83
			27	90	38	219	100	82	556
			48.1%	38.9%	44.7%	33.3%	34.0%	35.4%	36.2%
中国					5	13	6	7	31
					29	84	57	74	244
					31.0%	34.5%	31.6%	28.4%	31.6%
四国						7	4	6	17
						46	25	48	119
						30.4%	36.0%	33.3%	32.8%
北九州				4	6	9	7	11	37
				51	46	74	45	108	324
				23.5%	30.4%	44.6%	40.0%	28.7%	33.3%
南九州						9	7	7	23
						61	49	48	158
						44.3%	28.6%	33.3%	36.1%
合計	26	15	11	19	44	201	97	104	498
	198	97	38	155	360	1,494	796	923	4,061
	41.4%	35.1%	50.0%	32.9%	41.7%	37.0%	33.5%	31.9%	35.7%

上段：調査地点数 / 中段：第一次調査有効回収数 / 下段：累積脱落率(2010年第2次調査終了時点)

「若壮年サンプル」に関する検証と同様に、2004年時点での基本的な社会人口学的属性とパネル脱落率(2010年第二次調査に

おける累積脱落率)の関係を、男女別にみたのが、表6である。まず、配偶関係については、「若壮年サンプル」の場合と同様に、

2004年の調査開始時点で有配偶であった調査対象者のパネル脱落率が、男女ともに平均よりも低くなっていることが観察できる。「子どもと同居」「持ち家居住」についても、同様の傾向が観察できるが、女性については、調査開始時点での「子どもと同居」とパネル脱落の関係が有意ではない点は、注目に値する。また、「就業状態」および「健康状態」についても、男女ともに、パネル脱落との間に有意な関係が認められる。とくに、高齢者を対象としたパネル調査における「就業状態」と「パネル脱落」の関係については、一般に、「退職による居住地の移動に伴うパネル脱落」という可能性が考えられるが、ここでは、調査開始時点においてフルタイムで就業していた調査回答者のパネル脱落率が低くなる傾向が示された。

表6 JGGP「高齢者サンプル」における社会人口学的基本属性ごとの累積脱落率

	男女計	性別	
		男	女
全サンプル	35.7%	36.9%	34.5%
有配偶	34.3% **	35.2% **	33.3% *
子どもと同居	34.2% *	34.9% *	33.4%
持ち家居住	32.3% **	33.4% **	31.4% **
就業中	33.3% **	34.7% **	31.2% **
健康状態良好	34.2% **	35.7% **	32.8% **

* $p < 0.05$, ** $p < 0.01$

5. おわりに

以上、足早ではあるが、本稿では、同一の調査対象を継続的に追跡するパネル調査の実施に際する「サンプル消耗 (sample attrition)」の問題について、その脱落パターンの特性と確率サンプルの代表性に与える影響を中心に、「結婚と家族に関する国際比較パネル調査：JGGP」の経験を概観してきた。冒頭にも述べたとおり、パネル調査

の長期化にともない、調査の最初に選択された標本から多くの参加者が脱落していくサンプル消耗がパネル調査から得られるデータの情報量そのものを損なうとともに、標本と母集団の関係性が歪んだデータを分析に利用することによる推計バイアスの原因となることは、パネル調査の実施が抱える共通の問題である。

2004年から継続的に調査が実施されてきたJGGPでは、調査実施委託機関である(財)新情報センターの協力を得て、調査結果の要約を掲載したニューズレターや挨拶状を定期的に調査対象者に送付し、転居状況を把握するとともに、継続調査協力への理解を得ることに努めてきた。こうしたパネルメンテナンス作業をもっても、なお、「若壮年サンプル」で54%、「高齢者サンプル」で36%という低いとは言えないパネルの累積脱落率が確認された。

調査時点間におけるサンプル脱落が、層化二段無作為抽出法によって得られた確率サンプルの代表性に与える影響については、JGGPにおいても、性別・年齢・都市規模・地域といった確率サンプルを構成する基本的な属性とパネル脱落との関係が明らかになった。パネルデータの分析に際しては、調査時点間の脱落確率を反映したウェイト付与処置を行った「重みづけデータ」を再構築する必要が示唆された。

個別の属性とパネル脱落の関係をみると、年齢については、「若壮年サンプル」においては若年層ほどその後の追跡調査におけるパネル脱落が発生しやすく、「高齢者サンプル」においては、逆の影響が表れた。配偶関係については、「若壮年サンプル」の男女、および「高齢者サンプル」男性において、パネル脱落への有意な影響が認められ、その後の配偶関係の変化がパネル脱落確率に

与える影響が示唆された。また、その他、パネル脱落への強い影響が認められた社会人口学的変数は、調査開始時点での「持ち家居住」であり、居住地の変更に伴う地理的移動に起因するパネル脱落が発生することが示唆された。ここで確認された、特定の属性、とりわけ年齢とパネル脱落確率の関係は、離家（および移動）、結婚、就職・退職といったライフコース・イベントとパネル脱落の因果関係を示唆するものであり、こうしたライフコース・イベントに関する分析に際しては、パネル脱落とイベントの発生が独立ではないことに起因するバイアス、すなわち「サンプル脱落バイアス」に留意する必要があると考えられる。

注)

- 1 GGP についての詳細は、西岡（2009）および以下のウェブサイトを参照。
(<http://www.ggp-i.org/>)
- 2 国連ヨーロッパ経済委員会人口部の GGP コンソーシアム本部については、以下のウェブサイト参照。
(<http://www.unece.org/pau/ggp/welcome.html>)
- 3 JGGP データを用いた主要な研究成果としては、阿藤ほか（2011）などがある。
- 4 パネルデータにおけるサンプル脱落の問題を包括的に議論した重要な文献としては、Fitzgerald *et al.* (1998)、Lillard and Panis (1998)、Watson and Wooden (2009) などが挙げられる。また、日本国内におけるパネル調査におけるサンプル脱落の問題を取り扱った文献としては、村上（2003）、宮内・McKenzie・木村（2005）、坂本（2006）などがある。

【参考文献】

阿藤誠・西岡八郎・津谷典子・福田亘孝（2011）『少子化時代の家族変容：パートナーシップと出生行動』東京大学出版会。
北村行伸（2005）『パネルデータ分析』岩波書店。

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 巻、55-70 頁。
西岡八郎（2009）「世代とジェンダーに関する国際共同プロジェクト（GGP）について—特集に寄せて—」、『人口問題研究』第 65 巻、第 1 号、pp.1-2。
宮内環・C.R.McKenzie・木村正一（2005）「パネルデータ継続と回答行動の分析」、樋口美雄・慶應義塾大学経商連携 21 世紀 COE（編）『日本の家計行動のダイナミズム[II]』慶應義塾大学出版会、9-52 頁。
村上あかね（2003）「なぜ脱落したのか」『消費生活に関するパネル調査』における脱落サンプル分析」財団法人家計経済研究所編『家計・仕事・暮らしと女性の現在消費生活に関するパネル調査第 10 年度』国立出版局、第 3 部第 1 章。
Fitzgerald, J., P. Gottschalk, and R. Moffitt (1998) "An Analysis of Sample Attrition in Panel Data: The Michigan Panel Study of Income Dynamics," *The Journal of Human Resources*, Vol. 33: 251-29.
Lillard, L.A. and C.W.A. Panis (1998) "Panel Attrition from the Panel Study of Income Dynamics: Household Income, Marital Status, and Mortality," *Journal of Human Resources*, Vol. 33, pp.437-457.
Watson, W. and M. Wooden (2009) "Identifying Factors Affecting Longitudinal Survey Response," Lynn, Peter ed., *Methodology of Longitudinal Surveys*, Chichester: Wiley.

筆者プロフィール

中川 雅貴（なかがわ まさたか）

1979 年生まれ。同志社大学経済学部卒業、オーストラリア国立大学人口・社会統計研究所博士課程修了、Ph.D. 取得。早稲田大学人間科学学術院助手、日本福祉大学健康社会研究センター研究員を経て、2012 年 10 月より国立社会保障・人口問題研究所国際関係部研究員。専攻は、社会人口学、人口・社会統計学。