

16 データ・クリーニングの概況

本調査を行うにあたって NFRJ03 データの実査およびデータ入力から第一次クリーニングまでを、社団法人中央調査社に委託した。2004年8月に中央調査社から納品されたデータは、実行委員に version1.0 として公開された。その後、報告書執筆時までに大幅なデータの修正が加えられ、2005年1月時点で version 2.0 が公開されている。以下に、クリーニング作業の概況と、そこからみえてきた調査票の問題点について述べたい。

15-1 クリーニング作業の概況

コード表およびラベルの仕様書、およびデータ・クリーニング仕様書は、実行委員（稲葉昭英、田中重人、澤口恵一）が中央調査社と共同で作成し、その上で中央調査社に入念なクリーニングを依頼した。

日本家族社会学会ニューズレターにて、クリーニング作業を担当していただく有志を募り、データと調査票原票の納品後に、実行委員のクリーニング担当者とチェック作業を行った。有志として作業に参加をしたのは、田中慶子・土倉玲子である。なお、クリーニング作業班全体の統括は、澤口が行った。

作業班のメンバーから指摘されたエラー・確認事項は、基本的にすべて調査票にもどって確認し、必要であれば修正を加えた。調査票原票の確認ならびにデータの大幅な修正は澤口が行った。その結果、10月に実行委員およびクリーニング担当者に公開されたデータ version 2.0 をもとに本書の分析が進められた。

15-2 主たる修正箇所

- (1) 無回答が多い4ケースを有効回答から除外した。
- (2) 世帯表【問5付問1】の内容と、後続の同別居に関する問い【問9付問1、中高年票問16付問1（ウ）、問17付問1（エ）、問18付問2（ウ）、問19付問1（ウ）】に矛盾がある場合、原票にあたり一部修正をほどこした。
- (3) 世帯表に記載された個人の年齢が、対応する続柄における年齢と明らかに矛盾する場合には、原票にあたり修正をほどこした。原則的に後者を優先させているが、他の項目との整合性から前者にあわせたケースもある。
- (4) 本人または配偶者の妊娠・出産、育児などの理由による離職・休職経験に関する問い【問6付問4】で無回答が62ケース生じた。この項目における無回答は論理エラーではないが、未婚かつ無子の場合、調査票原票を精査したうえで3、「いずれもない」に修正した。
- (5) きょうだい人数【問18】において、無回答の場合のデータ入力方法が0となってい

たため無回答に修正した。

- (6) 既婚者の義父母有無【問 19】では、離婚による無配偶者の回答に多数無回答が生じた。スキップに関する指示が「結婚したことがない方」となっていたためである。これらについては、一括して非該当に修正した。
- (7) 続柄カテゴリーの保有状況と他の項目に整合性を欠くケースが多数生じていた（例 健在の子ども数【問 18】が 0 であるのにも関わらず、意識項目の「子どものことで悩んだこと」【問 12（ア）】が無回答となっている場合）。これについては、原則として、一括して非該当のコードに修正を行った（問 12(ア)であれば 5.「子どもはいない」に修正）。
- (8) ライフイベントの経験年齢の異常値は、原票にもどりチェックをし、必要であれば修正を行った。父母の年齢差については、親が実父・実母であれば異常値は不明として処理した。子については養子の可能性もあるため異常値に修正をほどこさなかった。

上記の修正はあくまでも原則であり、一部あえて整合性が欠けるかのように見える部分も現在のデータには残している。たとえば、対象者にとって、同別居の概念が、世帯表と特定の続柄カテゴリーごとの質問で同一のものとして捉えられているとはいえない。したがって、項目間での整合性をとるためにデータを強引に修正することは危険である。そのため同別居の整合性に関しては、単純なエディティング・コーディングのミスと判断された場合を除いて、修正を控えている。上で一括して修正を加えたと述べた、続柄カテゴリーの有無と、意識項目における当該の続き柄カテゴリーに関する問いの整合性に関しても同様である。たとえば、父母の生死が不明の場合（実際に行方不明のために本人にもわからない場合がある）に、「あなたの親に対する、配偶者の接し方について」【問 9 付問 17(イ)】の質問が「5. 親はいない」となっているケースには修正をほどこしていない。

15-3 調査票の構造的問題

最後にデータ・クリーニング作業からみえてきた、NFRJ03 の調査票の課題について指摘しておこう。

- (1) 世帯表を組み込んだ結果として、同じ対象者に関する情報を複数の箇所で見つける結果となった。対象者にとっても負担が大きく、データ・クリーニングのさいに、多重に論理チェックをすべき箇所が NFRJ98 よりも増えている。
- (2) とりわけ同居の概念に関しては、対象者の主観に依存する曖昧な余地を残している。入所・入院などに関する指示については対象者が迷わないよう、明確な基準や定義を明確にしておくべきであろう。
- (3) 子どもの出産・育児に伴う離退職経験について、多数無回答が生じたことは調査票の構造に依存した問題である。本人および配偶者（すでに離死別をした配偶者を含む）の双方に関する質問となっているため、エラーの発見や修正は非常に困難であ

る。NFRJ03 の調査票では親なり経験の有無について知ることはできないため修正は論理的には不可能である。原票に戻る手間を軽減するうえでも、次回以降の調査では修正する必要があるだろう。

- (4) 離婚経験者の義父母に関する問いについては、調査票作成時に多数無回答が生じることを予測していた。今回の修正は予想の範囲内のことではあったが、対象者の負担や迷いをなくすためにも次回にはワーディングを考慮すべきであろう。

なお、NFRJ03 データは、本報告書の執筆後に、もう一度クリーニングを行ったうえで一般公開される予定である。version 1.0 から version 2.0 で修正を行った箇所とその件数については、以下の表に示す。

変数名	問番号		修正件数
r5	問5		3
r5s102b	問5付問1	2b	1
r5s103b		3b	2
r5s103c		3c	1
r5s103da		3d 年齢	1
r5s103dg		3d 元号	1
r5s103dy		3d 年	2
r5s104b		4b	1
r5s104c		4c	1
r5s104da		4d 年齢	2
r5s104dg		4d 元号	2
r5s104dy		4d 年	4
r5s105b		5b	1
r5s105c		5c	1
r5s105dg		5d 元号	1
r5s105dy		5d 年	2
r5s107b		7b	1
r5s107c		7c	1
r5s107da		7d 年齢	1
r5s107dg		7d 元号	1
r5s107dy		7d 年	1
r5s2	問5付問2		1
r5s221b		21b	2
r5s221c		21c	2
r5s221da		21d 年齢	2
r5s221dg		21d 元号	2
r5s221dy		21d 年	2
r6s3	問6付問3		1
r6s4_1	問6付問4	1	18
r6s4_2		2	18
r6s4_3		3	18
r8	問8		1

r9	問9		1
r9s1	問9付問1		1
r9s10	問9付問 10		1
r9s11h	問9付問 11	時間	2
変数名	問番号		修正件数
r9s11m		分	2
r9s12h	問9付問 12	時間	1
r9s12m		分	1
r9s13	問9付問 13		1
r9s14	問9付問 14		1
r9s15_01	問9付問 15	ア	1
r9s15_02		イ	1
r9s15_03		ウ	1
r9s16r01	問9付問 16	本人ア	1
r9s16r02		本人イ	1
r9s16r03		本人ウ	1
r9s16r04		本人エ	1
r9s16r05		本人オ	1
r9s16r06		本人カ	6
r9s16r07		本人キ	6
r9s16s01		配偶ア	1
r9s16s02		配偶イ	1
r9s16s03		配偶ウ	1
r9s16s04		配偶エ	1
r9s16s05		配偶オ	1
r9s16s06		配偶カ	6
r9s16s07		配偶キ	6
r9s17_01	問9付問 17	ア	5
r9s17_02		イ	53
r9s17_03		ウ	1
r9s17_04		エ	1
r9s17_05		オ	1
r9s17_06		カ	1

r9s18	問9付問 18		1
r9s19	問9付問 19		1
r9s2a	問9付問2	年齢	2
r9s2g		元号	1
r9s2m		月	1
r9s2y		年	1
r9s3a	問9付問3	年齢	4
r9s3g		元号	1
r9s3m		月	1
r9s3y		年	1
r9s4	問9付問4		1
r9s5	問9付問5		1
r9s6	問9付問6		1
r9s7	問9付問7		2
r9s8	問9付問8		1
r9s9	問9付問9		2
r12_01	問 12	ア	5
r12_01	問 12	ア	5
r12_02		イ	14
r12_03		ウ	35
r12_07		キ	27
r12_08		ク	28
r12_09		ケ	28
r12_10		コ	29
r16	問 16		6
r16a2a	問 16 付問第2子	年齢	1
r16a2y		年	9
r16a3p4q		エ	1
r16b2a	問 16 付問第2子	イ年齢	1
r16b2y		イ年	2
変数名	問番号		修正件数
r16b3p4q		エ	1
r16c5q		オ	1
r17f	問 17 父		6
r17f1	問 17 付問1父	ア	1
r17f10		コ	1
r17f11		サ	1
r17f12		シ	1
r17f2a		イ年齢	2
r17f2g		イ元号	3
r17f2m		イ月	3
r17f2y		イ年	3
r17f3		ウ	1

r17f4		エ	1
r17f5		オ	1
r17f6		カ	1
r17f7		キ	1
r17f8		ク	1
r17f9		ケ	1
r17m		問 17 母	6
r17m1	問 17 付問1母	ア	2
r17m10		コ	2
r17m11		サ	2
r17m12		シ	2
r17m2a		イ年齢	5
r17m2g		イ元号	5
r17m2m		イ月	4
r17m2y		イ年	5
r17m3		ウ	2
r17m4		エ	2
r17m5		オ	2
r17m6		カ	2
r17m7		キ	2
r17m8		ク	2
r17m9		ケ	2
r17s2		問 17 付問2	2
r18a2g		イ元号	1
r18a2y		イ年	1
r18a3	問 18 付問2:1番目	ウ	14
r18b3	問 18 付問2:2番目	ウ	4
r18c3	問 18 付問2:3番目	ウ	3
r18s1_11	問 18 付問1	健在兄	4
r18s1_12		健在姉	5
r18s1_14		健在弟	4
r18s1_15		健在妹	5
r19f	問 19 義父		286
r19f10		コ	20
r19f11		サ	20
r19f1a		ア年齢	20
r19f1g		ア元号	20
r19f1m		ア月	20
r19f1y		ア年	20
r19f2		イ	20
r19f3		ウ	20
r19f4		エ	20
r19f5		オ	20

r19f6		カ	20
r19f7		キ	20
r19f8		ク	20
r19f9		ケ	20
r19m	問 19 義母		286
r19m10		コ	32
r19m11		サ	32
r19m1a		ア年齢	32
r19m1g		ア元号	32
r19m1m		ア月	32
r19m1y		ア年	32
r19m2		イ	32
r19m3		ウ	32
r19m4		エ	32
r19m5		オ	32
r19m6		カ	32
r19m7		キ	32
r19m8		ク	32
r19m9		ケ	32
r19s2		イ	17
r20_04	問 20	4	1