

2023 국민여행조사 표본설계

1. 조사개요

「국민여행조사」는 사회가 발전하고, 여가의 중요성이 점차 증대됨에 따라 국민 여행의 중요성은 점차 증대되고 있다. 또한 UNWTO 등과 같은 국제기구에서는 관광의 행태와 국민여행의 총량을 파악하여 그 나라의 경제에 직접 또는 간접적으로 미치는 영향을 파악하는 정도를 파악하는 것을 권고하고 있다.

따라서 우리나라에서도 관광통계를 국제기준에 맞춰 작성하고, 국제적으로 경제에 미치는 영향을 파악할 수 있는 방안으로 통계를 생산할 필요성이 제기되었다. 이러한 이유로 여행의 기준과 조사표의 변경이 필요함은 물론, 조사에 의한 오차를 줄이기 위해 월별로 통계를 작성하는 것이 타당한 것으로 판단되었다.

2018년부터 월별조사 형태로 진행되는 국민여행조사는 국내여행의 정의를 광역시와 도에 있는 시와 군을 경계로 설정하여 이를 넘어 방문하는 활동을 여행으로 정의하기로 하였다.

국민여행조사의 목적은 우리나라 국민의 여행실태를 종합적으로 파악하여 국가관광에 관한 정책수립과 연구 및 분석 등을 위한 기초 자료를 제공함으로써 관광을 통해 지역발전을 도모하고, 궁극적으로 국민의 복지 및 삶의 질을 제고하는데 있다.

2018년부터 월별로 조사하는 국민여행조사는 2019년도에 월 4,000명을 조사도록 변경하였으며, 주요 조사내용은 크게 ①관광·휴양 여행 경험 여부, ②친구·친지 방문 여행 경험 여부, ③출장·업무 여행 경험 여부, ④해외 여행 경험 여부, ⑤인구통계학적 배경으로 이뤄져 있으며, 각각의 세부 항목을 살펴보면 ①여행기간, ②여행시기, ③여행지, ④여행선택이유, ⑤여행지에서의 활동, ⑥여행정보 경로 획득, ⑦주요 이동 수단, ⑧숙박시설, ⑨여행 동반자, ⑩사전 예약 서비스 이용 여부, ⑪여행 상품 구매, ⑫사전 지출비용, ⑬여행지 지출비용, ⑭지출비용 비중, ⑮전반적 만족도, ⑯재방문 의도, ⑰타인 추천 의도 등으로 구성되어 있다. 또한 조사결과의 활용을 위해 인구통계학적 배경의 질문은 직업, 학력, 결혼상태, 가구원수, 가구소득, 개인소득 등으로 구성되어 있다.

2021년도 국민여행조사는 2020년도 조사를 바탕으로 도지역(9개)에서 연간 시군구 통계가 가능하도록 조정하였다. 한편 2023년도 국민여행 실태조사의 경우 2022년도 조사의 표본규모를 일부 확대하여 4,300명을 조사하도록 하였으며 2022

년도와 마찬가지로 월간 및 연간 통계 산출이 가능하도록 변경하였다.

2. 모집단 및 표본추출틀

가. 조사대상

「국민여행조사」의 개념상 규정된 목표모집단(target population)은 조사시점 현재 대한민국 내 모든 가구의 만 15세 이상 전체 국민으로 정의된다. 이에 대한 조사모집단(survey population)은 조사의 현실적인 측면을 고려하여 섬지역, 기숙시설, 특수시설을 제외한 아파트 조사구 및 일반조사구 내 가구에 거주하는 만 15세 이상의 국민이다.

표본설계에서 사용한 표본추출틀(sampling frame)은 2020년 기준 인구주택총조사의 전체 조사구 중 섬 일부, 기숙시설, 특수사회시설 조사구를 제외한 보통조사구, 섬조사구, 아파트 조사구만으로 구성하였기 때문에 동일시점의 우리나라 전국 인구, 가구 및 거주수와는 차이가 있을 수 있다. 본 표본설계에서는 아파트조사구와 일반조사구로 구분하였는데, 일반조사구는 보통조사구와 섬조사구로 설정하였으며, 아파트조사구는 그대로 아파트조사구로 설정한 리스트를 사용하였다.

표본 조사구에서 조사대상으로 고려한 ‘가구’는 통계청의 가구 정의¹⁾에 따라 “1인 또는 2인 이상이 모여 취사, 취침 등 생계를 같이 하는 생활 단위”로 규정하고, ‘집단가구’를 제외한 ‘일반가구’를 조사대상으로 하였다. 「국민여행조사」의 조사대상은 다음과 같이 정의된 일반가구 중 만 15세 이상의 국민이 거주하고 있는 가구이다.

- ① 가족으로 이루어진 가구
- ② 가족과 5인 이하의 남남이 함께 사는 가구
- ③ 1인가구
- ④ 가족이 아닌 5인 이하의 남남이 함께 사는 가구

「국민여행조사」의 조사대상에 제외되는 집단가구는 다음과 같은 유형이 있다.

- ① 가족이 아닌 6인 이상 가구
- ② 집단시설 가구 (기숙사, 양로원, 보육원 등 사회시설)
- ③ 외국인가구 (외국인만으로 구성된 가구)

1) 2010통계청의 인구센서스 지침서

참고로 한국인과 외국인이 함께 사는 가구(다문화가정 등)는 ‘일반가구’에 해당되며, 다가구 주택의 경우는 하나의 주택에 여러 가구가 거주하지만 독립가구를 이루고 있기 때문에 별도의 가구로 간주하였다.

나. 모집단 현황

본 조사를 위한 표본설계를 위해서는 표본조사구 추출을 위한 추출틀로는 2020년 기준 센서스 결과를 활용한다. 이는 <표 1>에 정리하여 제시하였다. 전체 조사모집단 조사구는 384,652개이고, 일반가구는 21,407,212가구이다. 지역별 조사구 현황을 보면 서울시와 경기도는 각각 전체 조사구의 18.6%와 23.9%를 차지하고, 전체 가구의 약 19.2%와 24.6%를 차지하고 있다. 전반적으로 행정구역별 조사구 수와 가구 수의 구성비는 유사하다. 만 15세 이상 인구 현황을 보면 서울시와 경기도는 각각 전체의 19.1%와 25.8%를 차지하고 있다.

<표 1> 모집단 현황(2020년 기준)*주)

(단위 : 개, 가구, 명, %)

시도	조사구 수	상대빈도	가구 수	상대빈도	15세 이상 인구	상대빈도
전국	384,652	100.0	21,407,212	100.0	44,571,611	100.0
서울	71,444	18.6	4,114,963	19.2	8,516,824	19.1
부산	26,373	6.9	1,416,732	6.6	2,925,152	6.6
대구	17,018	4.4	994,896	4.6	2,088,095	4.7
인천	20,540	5.3	1,174,367	5.5	2,541,000	5.7
광주	10,518	2.7	607,060	2.8	1,254,202	2.8
대전	10,889	2.8	637,547	3.0	1,272,478	2.9
울산	8,347	2.2	451,605	2.1	960,240	2.2
세종	2,546	0.7	140,697	0.7	277,763	0.6
경기	91,944	23.9	5,272,346	24.6	11,486,631	25.8
강원	12,761	3.3	665,554	3.1	1,318,312	3.0
충북	12,697	3.3	692,681	3.2	1,385,406	3.1
충남	16,831	4.4	917,398	4.3	1,834,543	4.1
전북	14,261	3.7	765,628	3.6	1,539,276	3.5
전남	15,110	3.9	769,339	3.6	1,526,124	3.4
경북	22,032	5.7	1,145,994	5.4	2,260,930	5.1
경남	25,991	6.8	1,371,180	6.4	2,821,748	6.3
제주	5,350	1.4	269,225	1.3	562,887	1.3

*주) 통계청의 MDIS 자료를 분석한 결과로서 섬조사구와 특수시설조사구를 제외한 조사구, 가구수, 인구수임

3. 층화

「2023년 국민여행조사」 표본설계는 7개 특별·광역시와 9개 도 지역으로 1차로 층화하고, 9개 도 지역에 대해서는 각 도별 시군구가 모두 포함되도록 층화하였다. 이

조사의 1차 추출단위는 인구총조사의 조사구이고, 2차 추출단위는 가구와 가구 내 만 15세 이상 가구원이다. 1차 추출단위인 인구총조사의 조사구는 지리적으로 인접한 대개 60~70여 가구로 구성되어 있는데, 층화는 1차 추출단위인 조사구의 특성 정보를 바탕으로 이루어지게 된다.

서울은 지역의 특성을 반영하여 다음과 같이 4개 권역으로 층화하여 표본 조사구를 배분하도록 하였다. 2020년도 표본설계에서는 경기도 또한 3개 권역으로 구분하여 세부 층화하였으나, 2021년도부터는 도 지역에서 모든 시군구가 표본지역으로 포함되도록 변경하였기 때문에 경기도에서는 별도의 권역을 구분하지 않았다. 2023년도의 표본설계에서는 2021년도 표본설계를 감안하여 경기도의 경우에는 별도의 구분을 두지 않았다.

<표 2> 서울시 권역별 해당 구 현황 및 15세 이상 인구수

권역	해당 지역	15세 이상 인구수	백분율(%)
북서부	종로구, 중구, 용산구, 은평구, 서대문구, 마포구	1,444,258	17.21
북동부	성동구, 광진구, 동대문구, 중랑구, 성북구, 강북구, 도봉구, 노원구	2,591,251	30.88
남서부	양천구, 강서구, 구로구, 금천구, 영등포구, 동작구, 관악구	2,636,348	31.42
남동부	서초구, 강남구, 송파구, 강동구	1,720,032	20.50

「2023년 국민여행조사」를 위한 표본설계에서 인구총조사 조사구를 대상으로 하는 1차 층화는 총 17개 시도(7개의 특별·광역시, 세종시, 9개의 도지역)를 도시지역과 농촌지역의 특성을 반영하여 동부, 읍/면부로 세부 층화하였는데, 서울시는 조사구, 가구수, 조사대상 인구수를 고려하여 효율적으로 조사하기 위해 인구규모와 지역적 특성을 고려하여 세부 층화가 필요하여 <표 4>에서 제시한 것처럼 전체 인구 규모와 지역적 차이를 고려하여 서울시는 4개 권역(북서, 북동, 남서, 남동)으로 세부 층화하였다.

2차 층화는 행정구역에 따라 구성된 22개의 1차 층을 조사구 특성(일반조사구, 아파트조사구)에 따라 2차 층화하였으며, 그 결과는 표 5에 제시했다. 여기서 조사구가 전체 아파트로 구성되었으면 아파트조사구로, 아파트가 전혀 없다면 일반조사구로 편성하며, 아파트와 일반조사구가 섞여 있다면 아파트가 더 많으면 아파트조사구로 분배하였으며, 아파트 외의 가구가 더 많다면 일반조사구에 편성하였다.

<표 3> 층화 방법

구분	지역	1차 층화	2차 층화	총 수
인구총조사 조사구	-서울시(4개 권역) -광역시(6개) -세종특별자치시(1개)	동지역	아파트/일반조사구	22
		읍면 지역	아파트/일반조사구	10
	-도지역(9개) 지역에 대해 시군구가 모두 포함되도록 층화	-시군구 지역으로 세부층화 -군지역은 읍면, 시 및 구 지역은 동읍면지역이 혼합되어 있음		152

<표 4> 세부 권역별 모집단 조사구 현황^{주)}

(단위 : 개)

행정구역	동부		읍면부		합계
	일반	아파트	일반	아파트	
전 국	135,633	169,233	53,105	26,294	384,265 ²⁾
서울-북서	8,051	4,724	.	.	12,775
서울-북동	12,177	9,402	.	.	21,579
서울-남서	13,783	8,755	.	.	22,538
서울-남동	6,867	7,666	.	.	14,533
부산	10,765	14,344	366	885	26,360
대구	6,629	8,577	454	1,338	16,998
인천	8,752	11,069	655	55	20,531
광주	3,532	6,983	.	.	10,515
대전	4,936	5,946	.	.	10,882
울산	2,457	4,256	722	902	8,337
세종	60	1,636	574	273	2,543
경기	30,336	45,336	9,748	6,422	91,842
강원	2,861	4,332	4,002	1,543	12,738
충북	2,795	4,173	3,677	2,036	12,681
충남	2,806	4,573	6,272	3,161	16,812
전북	3,680	5,969	3,846	755	14,250
전남	2,117	3,855	7,093	2,025	15,090
경북	4,668	6,189	8,077	3,062	21,996
경남	6,035	9,953	6,231	3,703	25,922
제주	2,326	1,495	1,388	134	5,343

*주) 통계청 MDIS의 인구총조사 자료(특수시설조사구 제외)를 분석한 결과임(2020년 기준)

2) 빈집조사구와 비연계조사구를 제외하였기 때문에, 표 1의 현황과 차이가 남

<표 5> 세부 층별 모집단 일반가구 현황*주)

(단위 : 가구)

행정구역	동부		읍면부		합계 ³⁾
	일반	아파트	일반	아파트	
전 국	7,370,598	9,492,804	2,663,220	1,345,686	20,872,308
서울-북서	442,455	254,166	.	.	696,621
서울-북동	675,958	551,271	.	.	1,227,229
서울-남서	754,159	499,076	.	.	1,253,235
서울-남동	374,892	420,980	.	.	795,872
부산	574,673	761,215	20,137	46,293	1,402,318
대구	390,582	494,481	24,700	74,327	984,090
인천	471,122	636,585	34,006	2,322	1,144,035
광주	195,365	403,050	.	.	598,415
대전	280,745	349,416	.	.	630,161
울산	134,856	221,523	37,061	49,580	443,020
세종	2,125	90,678	32,087	13,904	138,794
경기	1,596,285	2,631,640	499,503	356,330	5,083,758
강원	156,472	237,342	193,874	71,058	658,746
충북	154,942	232,887	188,342	100,738	676,909
충남	150,729	259,287	322,354	156,867	889,237
전북	196,779	330,551	191,023	35,888	754,241
전남	112,555	206,558	343,319	96,770	759,202
경북	254,685	326,865	399,120	146,817	1,127,487
경남	327,082	521,854	307,386	190,490	1,346,812
제주	124,137	63,379	70,308	4,302	262,126

*주) 통계청 MDIS의 인구총조사 자료(특수시설조사구 제외)에서 일반가구를 분석한 결과임(2020년 기준)

3) 빈집조사구와 비연계조사구를 제외하였기 때문에, 표 1의 현황과 차이가 남. 조사대상이 되는 일반가구의 정보가 제공되어 일반가구수를 제시하였음.

<표 6> 세부 층별 모집단 15세 이상 인구 현황*주

(단위 : 명)

행정구역	동부		읍면부		합계
	일반	아파트	일반	아파트	
전 국	14,149,605	21,533,486	5,275,666	2,868,915	43,827,672
서울-북서	861,168	583,090	.	.	1,444,258
서울-북동	1,312,917	1,278,334	.	.	2,591,251
서울-남서	1,474,361	1,161,987	.	.	2,636,348
서울-남동	708,015	1,012,017	.	.	1,720,032
부산	1,058,779	1,688,173	38,523	97,685	2,883,160
대구	734,843	1,113,143	50,083	159,146	2,057,215
인천	941,482	1,483,118	67,831	5,164	2,497,595
광주	356,647	875,798	.	.	1,232,445
대전	487,529	765,888	.	.	1,253,417
울산	255,524	502,417	75,124	111,704	944,769
세종	3,169	184,885	55,662	30,226	273,942
경기	3,233,671	6,175,833	1,077,285	798,959	11,285,748
강원	277,433	499,511	376,957	140,092	1,293,993
충북	284,640	503,617	365,349	207,046	1,360,652
충남	267,619	563,942	640,823	326,445	1,798,829
전북	356,813	710,839	373,993	71,694	1,513,339
전남	209,874	434,526	654,071	200,044	1,498,515
경북	456,957	702,500	756,261	305,534	2,221,252
경남	611,424	1,157,590	594,173	406,652	2,769,839
제주	256,740	136,278	149,531	8,524	551,073

*주) 통계청 MDIS의 인구총조사 자료(특수시설조사구 제외)를 분석한 결과(2020년 기준)임

4. 표본크기 결정 및 표본배분

가. 표본크기 및 목표오차

「2023년 국민여행조사」의 표본크기는 연간조사 기준으로 약 51,600명이고, 월별로 는 4,300가구에서 4,300명을 추출하여 조사한다. 본 조사의 표본크기는 가용 예산과 조사 소요시간 등의 조사에 필요한 제반여건을 고려하고, 작성되는 통계의 표본오 차 수준을 검토하여 결정되었다. 본 연구에서는 월별로 전국에서 430개 표본 조사 구를 추출하고, 각 표본 조사구로부터 10가구의 표본가구를 추출하여 표본가구 내

만 15세 이상 가구원 중 한 명을 조사하는 것이 원칙으로 하였다.

주어진 목표 오차내에서 층화 추출을 적용한 최종 표본크기는 비례배분과 네이만 배분의 경우에 다음과 같은 산출식으로 구할 수 있다. 이러한 표본규모에 대한 산출값을 근거로 허용오차내에서 최종표본규모를 결정할 수 있다.

$$\text{- 비례배분 : } n = \frac{\sum_{h=1}^L N_h S_h^2}{ND + \sum_{h=1}^L W_h S_h^2}$$

$$\text{- 네이만배분 : } n = \frac{(\sum_{h=1}^L N_h S_h)^2}{N^2 D + \sum_{h=1}^L N_h S_h^2}$$

여기서 $D = B^2/4$ 이고, B 는 전국단위의 통계를 생산하는데 사용할 목표오차이며, N_h 는 h 층의 모집단 수이며, $W_h = N_h/N$ 이며, N 은 전체 모집단 수, S_h^2 은 h 층의 모집단 분산이며, 이 값은 층별 비율(15세 이상 인구비율)을 적용할 경우 $S_h^2 = P_h Q_h$ 이다.

<표 7> 목표오차에 따른 표본 크기(15세 이상 인구기준)*^{주)}

(단위 : %, 명)

목표오차	비례배분	네이만배분
1.50	1,927	1,526
1.05	3,932	3,115
1.01	4,249	3,367
1.00	4,334	3,434
0.90	5,351	4,240
0.89	5,472	4,336
0.88	5,597	4,435
0.80	6,772	5,366
0.70	8,845	7,009
0.60	12,038	9,540

*주) 목표오차를 월간단위 표본규모로 제시한 이유는 국민여행실태조사의 경우 월단위통계가 산출되기 때문에 월간 허용오차를 관리함으로써 연간 단위 오차도 관리될 수 있기 때문이다

<표 7>로부터 월간 단위로 4,300명을 조사할 경우 비례배분은 1.00% 가량으로 관리 가능하며, 네이만 배분의 경우 0.89% 가량에서 관리 가능한 것으로 나타났다.

나. 표본배분

「2023년 국민여행조사」 표본설계의 중요한 목표는 통계공표 단위별 국민여행조사에 대하여 일정 수준 이상의 정확도(precision)를 갖는 통계를 생산하는 데 있다. 이 조사의 주요 통계공표 단위는 시도 구분과 연령대와 성별 구분이다. 표본배분 단계에서는 시도별 통계작성의 정확성을 검토하여 최종 표본배분법을 결정한다.

먼저 1차 층인 각 시도에 대한 표본배분은 층 내 가구 수를 기준으로 비례배분법, 제공근 비례배분법 등을 비교하여 결정하였다.

- 비례배분법 : $n_h = n \times \frac{N_h}{\sum_{k=1}^H N_k}$, 단, N_h 는 층 h 의 모집단 가구 수

- 제공근 비례배분법 : $n_h = n \times \frac{\sqrt{N_h}}{\sum_{k=1}^H \sqrt{N_k}}$

먼저 비례배분법은 정해진 전체 표본크기를 해당 층의 가구 수에 비례하여 배분한 것이다. 비례배분법은 전체 통계작성에는 유리하지만 가구 수가 적은 시도에서 표본이 적게 배분되어 통계생산에 어려움이 발생할 수 있다. 제공근 비례배분법은 각 지역의 가구 수의 제공근에 비례하여 배분한 것이다.

한편 2023년도 조사에서는 도지역의 연간 시군구단위로 통계산출이 가능하도록 하기 위해 다음과 같은 배분 방안을 검토하였다.

- 1단계 : 기본적으로 특광역시는 100가구, 시군은 10가구를 기본배분함
- 2단계 : 기본배분 가구수를 제외한 나머지 가구수에 대해서는 비례배분 또는 제공근 배분, 제공근 배분의 경우 세대수 5만이하 지역에서는 표본지역에서 제외하거나 또는 포함하는 방안을 검토함

- (1) 방안1 : 기본배분 후 나머지 가구수는 가구수에 비례하여 배분함
- (2) 방안2 : 기본배분 후 나머지 가구수는 제공근 비례배분함
- (3) 방안3 : 조사대상 인구 5만 이하는 기본배분 후, 나머지 지역에 대해 비례 배분함
- (4) 방안4 : 조사대상 인구 5만 이하는 기본배분 후, 나머지 지역에 제공근 비례배분함

- 특광역시를 제외하고, 시도단위의 월간 통계는 동부와 읍면동별 일반조사구와 아파트조사구에서 최소 1개 이상의 조사구가 포함되도록 조정하여 배분함

최종적으로 가구 기준으로 2023년 표본배분 결과 예상 상대오차는 제주도를 제외하고 10%내외로 관리 가능하도록 하고 있으며, 이러한 기준에 따라 지역적으로 10%내외의 상대오차를 관리할 수 있도록 하였다. 결과적으로 특광역시외의 표본 규모를 현저하게 감소시키지 않으면서 나머지 지역에서 상대표준오차가 개선되는 방안으로 “방안1”을 최종 방안으로 가장 타당한 방안으로서 기본 배분 후 비례배분의 경우 지역별로 추출률을 고려하여 강제배분하기 때문에 지역별 표본규모를 유지하기 위해 제공된 비례배분보다 효과적인 방법으로 고려하였다.

<표 8> 표본배분 방안

(단위 : 가구)

시도	2023년도 표본배분			
	방안1	방안2	방안3	방안4
전국	4300	4300	4300	4300
서울	470	190	710	220
부산	230	150	250	130
대구	190	140	180	110
인천	210	150	200	120
광주	160	130	110	90
대전	160	130	110	90
울산	140	130	80	70
세종	110	120	20	40
경기	780	820	940	1240
강원	240	320	210	270
충북	170	210	170	210
충남	230	300	210	280
전북	210	260	200	240
전남	290	390	250	320
경북	340	430	300	400
경남	300	360	310	390
제주	70	70	50	80

<표 9> 지역별 예상 상대표준오차

(단위 : %)

시도	2023년도 표본배분			
	방안1	방안2	방안3	방안4
전국	1.52	1.52	1.52	1.52
서울	4.61	7.26	3.76	6.74
부산	6.59	8.16	6.32	8.77
대구	7.26	8.45	7.45	9.53
인천	6.90	8.16	7.07	9.12
광주	7.90	8.77	9.53	10.54
대전	7.90	8.77	9.53	10.54
울산	8.45	8.77	11.18	11.95
세종	9.53	9.13	22.36	15.81
경기	3.58	3.49	3.26	2.84
강원	6.45	5.59	6.90	6.08
충북	7.90	6.90	7.67	6.90
충남	6.59	5.77	6.90	5.98
전북	6.89	6.20	7.07	6.45
전남	5.87	5.06	6.32	5.59
경북	5.42	4.82	5.78	5.00
경남	5.68	5.27	5.68	5.07
제주	11.95	11.95	14.14	11.18

<표 10> 으로부터 특광역시 지역은 기존의 동부와 읍면부 및 조사구 특성에 따라 표본조사구를 배분하고, 9개 도지역은 각 도별로 시군구가 모두 포함되도록 배분한 결과 시 및 구 지역은 동부, 군지역은 읍면부로 구분하여 배분하면 경기도를 제외한 나머지 지역은 적절히 동부와 읍면부로 배분되는 효과가 나타난다. 경기도의 경우 시군구 별로 별도의 배분결과를 바탕으로 일반과 아파트 조사구에 따라 배분하도록 조정하였다. 도지역의 시군구별 표본조사구 배분현황은 <부록>에 제시하였다.

시군구 단위의 조사구 배분은 <부록>의 개별 시군구가 모두 포함되도록 1차적으로 고려하고, 다음으로 <표 10>의 일반과 아파트 조사구 특성을 고려하여 배분하였다.

<표 10> 지역별 월간 표본 조사구 배분현황

(단위 : 개)

행정구역	동부		읍면부		합계
	일반	아파트	일반	아파트	
전 국	128	159	94	49	430
서울-북서	5	3	0	0	8
서울-북동	8	6	0	0	14
서울-남서	9	6	0	0	15
서울-남동	5	5	0	0	10
부산	9	12	1	1	23
대구	7	10	1	1	19
인천	9	10	1	1	21
광주	5	11	0	0	16
대전	7	9	0	0	16
울산	4	7	1	2	14
세종	1	7	2	1	11
경기	27	35	9	7	78
강원	4	5	9	6	24
충북	3	4	6	4	17
충남	4	5	8	6	23
전북	3	6	10	2	21
전남	3	4	17	5	29
경북	6	6	17	5	34
경남	6	7	10	7	30
제주	3	1	2	1	7

5. 표본추출 및 표본대체

국민여행실태 조사를 위한 표본설계에서 1차 추출단위는 조사구이고, 2차 추출단위는 가구이다. 1차 추출단위인 표본 조사구는 조사구 내 가구 수에 비례하는 확률비례계통추출법(Probability Proportional to Size Systematic Sampling)에 의해 추출하였다. 72개의 각 층에서 조사구를 행정구역 번호에 따라 정렬한 후 확률비례계통추출법을 적용하여 추출함으로써 해당 층을 지리적으로 대표할 수 있도록 하였다. 일반적으로 계통추출법을 적용할 때 추출단위를 층화지표에 따라 정렬한 후 추출하면 모집단을 잘 대표할 수 있어 효과적이다.

표본으로 추출된 조사구에서는 10가구를 계통추출하여 조사하는 것을 원칙으로 하였다. 표본조사 대상가구 선정 방법은 표본조사구의 요도와 가구명부에서 10가구를 계통추출법으로 선정하고 순서대로 방문하여 적격 조사대상 가구를 조사하였다.

추출된 표본조사구가 조사시점 차이로 인하여 표본조사구 전체가 유고(재건축, 재개발, 기타사유 등)로 인하여 조사가 불가능하거나 특성이 변경(예: 주택에서 상가지역으로 개발된 경우)되었을 경우에는 예비조사구로 대체하였다.

표본 가구와의 접촉을 위한 방문 횟수는 최대 3회로 하며, 요일/시간대를 달리하여 방문하도록 하였다. 또한 면접원은 매회 방문할 때 <가구방문기록표> 양식에 준해 해당 가구와의 접촉 상황을 기록하였다. 3회 방문시에도 조사대상 가구/조사대상자와의 면접이 이루어지지 않을 경우에는 사유를 기록하고, 가구명부를 기준으로 정해진 접촉 순서에 따라 인접한 대체조사 가구를 선정하여 다시 접촉/방문을 시도하였다.

최종 조사대상자 선정을 위해서 가구 방문 후 ‘가구원리스트’를 작성하여 해당가구에 적격자가 있는 지 확인하고, 적격자가 2명 이상일 경우 ‘최근 생일자법’에 의거하여 조사시점과 가장 근접한 생일자로 최종 조사대상자로 선정한다. 최종 조사대상 가구원이 부재중일 때는 해당 가구원과 면접 가능한 요일/시간대를 확인하고 빠른 시일 내에 재방문하여 조사하도록 한다.

6. 가중치 산정

가. 가중치 산정 개요

통계조사에서 가중치 부여는 표본추출에 따른 추출률의 차이와 응답률 및 모집단에 대한 정보 등을 이용하여 모집단의 구조와 표본 구조를 맞추므로써 추정의 정확도를 높이는 것을 목적으로 한다. 추정단계에서 가중치를 이용하면 모집단에 대한 특성치인 모수에 대한 비편향 추정량(unbiased estimator)을 얻을 수 있다. 만약 통계 분석 과정에서 가중치를 무시하고 분석한 추정치는 심각한 편향(bias)이 발생할 수 있다. 표본의 크기가 큰 대규모 조사에서 문제가 되는 것은 추정량의 편향이기 때문에 추정과정에서 반드시 가중치를 이용해야 한다.

일반적으로 표본조사의 가중치는 ㉠ 설계가중치 산정, ㉡ 무응답에 대한 조정, ㉢ 모집단 정보를 이용한 조정 등의 세 가지 과정을 통해서 산정된다. 가중치 조정단계에서 이용하는 모집단 정보는 2023년 가구 및 인구 추계값이다.

표본설계에서는 각 시도 구분과 도 지역 내의 동 및 읍·면 구분을 이용한 세부 층화를 통해서 전체 73개 층을 구성하였다. 표본추출은 층화 2단계 확률비례계통추출법을 적용하였는데, 1차 추출단위는 조사구이고, 2차 추출단위는 가구 및 만 15세 이상 가구원이다. 각 층에서 배정된 표본 조사구 수만큼을 조사구 내의 가구 수에 비례하는 확률비례계통추출법에 따라 추출하였다. 이 조사의 표본가구로 선정된 경우에 적격 가구원이 2명 이상인 경우에는 최근 생일법에 의해서 한 명을 조사하도록 하였다.

본 연구에서 설계가중치는 표본추출 과정에서 층화2단계 확률비례계통추출법을 적용함에 따라 나타나는 조사구 내의 가구별 추출확률의 차이를 반영할 수 있도록 산정한다.

가중치 작성 과정에서 사용될 기호들을 정리하면 다음과 같다.

- L : 층의 수
- N_h : 층 h 의 모집단 조사구 수
- n_h : 층 h 의 표본 조사구 수
- S_{hi} : 층 h 의 i 번째 조사구에 대한 크기의 측도(해당 조사구의 총 가구 수)
- $S_h = \sum_{i=1}^{N_h} S_{hi}$: 층 h 에서 크기의 측도에 대한 총합
- M_{hi} : 층 h 의 i 번째 조사구 내 가구 수(조사완료+조사미완+조사미착수)

- m_{hi} : 층 h 의 i 번째 조사구 내 조사접촉 가구 수(응답+거절)
- r_{hi} : 층 h 의 i 번째 표본조사구 내 조사완료 가구 수(응답)

나. 조사구 추출에 대한 가중치 작성

(1) 설계가중치

이 조사의 설계가중치는 각 표본조사구에 대한 표본추출률의 역수와 표본조사구에서 가구조사 착수율의 역수를 곱하여 다음과 같이 산출한다.

$$\text{설계가중치} = \frac{S_h}{n_h S_{hi}} \times \frac{M_{hi}}{m_{hi}}$$

원칙적으로 각 표본 조사구에서는 10가구씩을 표본으로 조사하였기 때문에 $m_{hi} = 10$ 이다. $S_{hi} \approx M_{hi}$ 를 가정할 수 있는 경우(표본추출률 상의 조사구 내 가구 수와 실제 가구 수에 차이가 작은 경우) 각 층에서 설계가중치는 해당 층 내에서 일정한 값이 되어 설계가중치 = $\frac{S_h}{n_h \times 10}$ 으로 표현할 수 있다. 이때 각 지역 내 층에서

표본 가구들은 모두 동일한 설계가중치를 갖게 된다.

실제 조사과정에서는 표본추출률의 조사구 내 가구 수와 실제 가구 수에 차이가 있어 각 층 내에서 설계가중치는 동일하지는 않지만 비슷한 값을 갖게 된다.

(2) 무응답 조정

해당 지역 내 세부 층에서 조사 가구에 대한 설계가중치는 원칙적으로 같다. 본 조사에서 무응답 조정은 표본 조사구 단위로 진행되었다. 무응답 조정계수는 다음 식에 따라 구한다.

$$\text{무응답 조정계수} = \frac{m_{hi}}{r_{hi}}$$

(3) 표본 가구 내 추출률 반영

이 조사의 표본가구로 선정된 경우에 적격 가구원이 2명 이상인 경우에는 ‘최근 생일자법’에 의거하여 조사시점과 가장 근접하게 생일인 사람을 최종 적격자로 선정하였다. 따라서 표본 가구 내의 조사 적격자가 몇 명인가에 따라 추출률에 차이가 발생한다. 표본 가구의 적격자 추출률은 다음과 같다.

$$\text{가구 내 추출률} = \frac{1}{\text{표본 가구의 만 15세이상 전체 가구원 수}}$$

앞서 구한 가구 내 추출률의 역수를 설계가중치에 곱하여 가구 내 추출률 차이를 가중치 작성 과정에 반영한다.

(4) 「인구총조사」 결과를 이용한 조정

모집단 정보를 이용한 가중치 조정은 모집단에 대한 정보를 이용하여 모집단의 구조와 표본 구조를 유사하게 맞추므로써 추정의 정확도를 높이는 것을 목적으로 한다. 가중치 작성 단계에서 사용한 모집단에 대한 정보는 2023년 인구추계결과이다. 본 연구에서는 인구총조사 결과에 대해서 레이킹 비 접근법(Raking Ratio Method)을 적용하여 가중치를 조정하였다. 본 연구에서 사용한 모집단에 대한 정보는 2023년 기준의 인구추계 결과 중 시·도별(17)×성별(2) 구분과 지역(동부, 읍·면부)×연령대(7)×성별(2) 구분에 대한 만 15세 이상 인구 현황이다. <표 10>과 <표 11>은 각각 모집단 정보를 이용한 조정 단계에 사용된 2023년 인구추계결과의 시·도별(17)×성별(2)×연령대(7) 구분에 대한 만 15세 이상 인구 현황이다. 다만, 연령대를 세분할 것인가는 최종 조사데이터의 성 및 연령별 분포를 살펴서 최종 결정된다. 최종 가중치는 다음 식에 따라 산정되었다.

$$\text{최종 가중치} = \text{설계가중치} \times \text{무응답 조정계수} \times (1/\text{가구내추출률}) \times \text{조정계수}$$

(5) 극단 가중치 조정

일반적으로 가중치의 과도한 변동은 추정량의 분산을 크게 만들어 추정의 정확도를 떨어뜨릴 수 있다. 몇 개의 과도하게 큰 가중치를 갖는 조사값이 잘 설계되고 수행된 통계조사에서 얻어지는 조사의 정확성을 낮추게 하는 경우도 발생할 수 있다. 실제 가중치 작성 과정에서는 지나치게 큰 가중치를 표본설계가 끝난 후 사후적으로 제한하거나 조정하는 방법이 사용되고 있다. 대표적인 방법으로는 다음 두 가지로 구분할 수 있다.

첫째, 가중치 작성 과정에서 무응답 조정, 모집단 정보를 이용한 조정 등과 같은 가중치 작성 단계별로 지나치게 큰 조정값이 나오는 것을 막기 위하여 지나치게 큰 가중치의 경우에는 절단(trimming)하거나 제한을 두는 방안이다.

둘째, 가중치 작성의 전체 과정을 마친 후에 과도하게 큰 가중치를 찾아서 절단(trimming)하고, 절단으로 인하여 부족해진 가중치를 보충하는 과정을 거치는 방안이다.

「2023 국민여행실태조사」의 특이 가중치 조정은 두 번째 방안을 사용되었는데, 전체 응답자의 가중치 분포를 고려하여 상·하위 5%씩을 특이 가중치로 간주하여 조

정하였다.

7. 추정 및 추정량의 분산

가. 추정량

「2023년 국민여행실태조사」에서 주요 추정대상은 모평균이나 모비율이다. 모비율 추정 및 추정량의 분산은 모평균 추정방법과 동일하다. 다만, 모비율 추정의 경우 조사변수 값이 어떤 특성을 갖고 있는가에 따라 1 또는 0의 값을 갖는다는 점이 차이이다. 본 조사에서 각종 모평균 추정을 위해서 사용된 가중치를 이용한 추정량은 다음과 같이 정의된다.

$$\bar{y} = \frac{\sum_{h=1}^L \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} w_{hij} y_{hij}}{\sum_{h=1}^L \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} w_{hij}} = \frac{\sum_{h=1}^L \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} w_{hij} y_{hij}}{w_{\dots}} \quad (1)$$

여기서, w_{hij} 는 각 응답자에 부여된 가중치이고, y_{hij} 는 각 응답결과로 모비율 추정의 경우는 특정 속성을 갖고 있는 경우는 1, 아니면 0의 값을 갖는다. L 은 층의 수, n_h 는 층 h 에서의 1차 표본추출단위인 표본 조사구의 수, m_{hi} 는 층 h 내 i 번째 표본 조사구의 응답자 수이다. 또한 $w_{\dots} = \sum_{h=1}^L \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} w_{hij}$ 은 전체 응답자에 대한 가중치의 합계이다.

나. 분산추정

앞서 제시한 모평균 추정량에 대해서 층화와 2단 집락추출 등의 표본설계를 반영한 추정분산은 다음과 같이 계산된다.

$$\text{var}(\bar{y}) = \sum_{h=1}^L \frac{n_h(1-f_h)}{n_h-1} \sum_{i=1}^{n_h} (e_{hi.} - \bar{e}_{h..})^2 \quad (2)$$

여기서, L 은 층의 수, n_h 는 층 h 에서의 1차추출단위(PSU)인 조사구 수, m_{hi} 는 층 h 내 i 번째 표본조사구의 응답자 수, $f_h = n_h/N_h$, $e_{hi.} = \left(\sum_{j=1}^{m_{hi}} w_{hij} (y_{hij} - \bar{y}) \right) / w_{\dots}$,

$$\bar{e}_{h..} = \left(\sum_{i=1}^{n_h} e_{hi} \right) / n_h \text{ 이다.}$$

모평균 및 모비율 추정에 대한 표준오차(standard error), 상대표준오차(relative standard error: rse), 95% 신뢰수준 오차의 한계는 다음과 같다.

$$se(\bar{y}) = \sqrt{var(\bar{y})}, \quad (3)$$

$$rse(\bar{y}) = \frac{s.e(\bar{y})}{\bar{y}} \times 100(\%), \quad (4)$$

$$\text{오차의 한계} = 1.96 \times \sqrt{var(\bar{y})} \quad (5)$$

<부록> 시군별 연간 표본조사구 배분 결과

시군구			동부		읍면부		합계
시/도 코드	시군코드	시군구명	1.일반	2.아파트	1.일반	2.아파트	
31	3101	수원시	30	30	0	0	60
31	3102	성남시	24	24	0	0	48
31	3103	의정부시	12	24	0	0	36
31	3104	안양시	12	24	0	0	36
31	3105	부천시	24	24	0	0	48
31	3106	광명시	12	12	0	0	24
31	3107	평택시	12	12	6	6	36
31	3108	동두천시	6	6	0	0	12
31	3109	안산시	24	12	0	0	36
31	3110	고양시	24	36	0	0	60
31	3111	과천시	6	6	0	0	12
31	3112	구리시	12	12	0	0	24
31	3113	남양주시	3	12	9	12	36
31	3114	오산시	12	12	0	0	24
31	3115	시흥시	12	24	0	0	36
31	3116	군포시	12	12	0	0	24
31	3117	의왕시	12	12	0	0	24
31	3118	하남시	12	12	0	0	24
31	3119	용인시	12	27	6	3	48
31	3120	파주시	6	15	12	3	36
31	3121	이천시	6	6	6	6	24
31	3122	안성시	6	6	6	6	24
31	3123	김포시	6	12	12	6	36
31	3124	화성시	6	24	12	6	48
31	3125	광주시	6	6	6	6	24
31	3126	양주시	6	9	6	3	24
31	3127	포천시	6	6	6	6	24
31	3128	여주시	3	3	3	3	12
31	3135	연천군	0	0	6	6	12
31	3137	가평군	0	0	6	6	12
31	3138	양평군	0	0	6	6	12

시군구			동부		읍면부		합계
시/도 코드	시군코드	시군구명	1.일반	2.아파트	1.일반	2.아파트	
32	3201	춘천시	6	6	6	6	24
32	3202	원주시	6	18	6	6	36
32	3203	강릉시	6	6	6	6	24
32	3204	동해시	12	12	0	0	24
32	3205	태백시	6	6	0	0	12
32	3206	속초시	6	6	0	0	12
32	3207	삼척시	6	6	12	0	24
32	3231	홍천군	0	0	9	3	12
32	3232	횡성군	0	0	9	3	12
32	3233	영월군	0	0	6	6	12
32	3234	평창군	0	0	6	6	12
32	3235	정선군	0	0	6	6	12
32	3236	철원군	0	0	6	6	12
32	3237	화천군	0	0	6	6	12
32	3238	양구군	0	0	6	6	12
32	3239	인제군	0	0	6	6	12
32	3240	고성군	0	0	12	0	12
32	3241	양양군	0	0	6	6	12

시군구			동부		읍면부		합계
시/도 코드	시군코드	시군구명	1.일반	2.아파트	1.일반	2.아파트	
33	3302	충주시	12	15	6	3	36
33	3303	제천시	9	9	6	0	24
33	3304	청주시	15	24	3	6	48
33	3332	보은군	0	0	9	3	12
33	3333	옥천군	0	0	6	6	12
33	3334	영동군	0	0	6	6	12
33	3335	진천군	0	0	6	6	12
33	3336	괴산군	0	0	12	0	12
33	3337	음성군	0	0	6	6	12
33	3338	단양군	0	0	6	6	12
33	3339	증평군	0	0	6	6	12

시군구			동부		읍면부		합계
시/도 코드	시군코드	시군구명	1.일반	2.아파트	1.일반	2.아파트	
34	3401	천안시	15	21	6	6	48
34	3402	공주시	6	6	12	0	24
34	3403	보령시	3	3	6	0	12
34	3404	아산시	6	6	6	6	24
34	3405	서산시	6	6	6	6	24
34	3406	논산시	6	6	6	6	24
34	3407	계룡시	0	6	0	6	12
34	3408	당진시	6	6	6	6	24
34	3431	금산군	0	0	6	6	12
34	3433	부여군	0	0	6	6	12
34	3434	서천군	0	0	6	6	12
34	3435	청양군	0	0	12	0	12
34	3436	홍성군	0	0	6	6	12
34	3437	예산군	0	0	6	6	12
34	3438	태안군	0	0	6	6	12

시군구			동부		읍면부		합계
시/도 코드	시군코드	시군구명	1.일반	2.아파트	1.일반	2.아파트	
35	3501	전주시	12	36	0	0	48
35	3502	군산시	6	12	6	0	24
35	3503	익산시	6	12	6	0	24
35	3504	정읍시	6	6	12	0	24
35	3505	남원시	3	3	6	0	12
35	3506	김제시	3	3	6	0	12
35	3531	완주군	0	0	12	12	24
35	3532	진안군	0	0	12	0	12
35	3533	무주군	0	0	12	0	12
35	3534	장수군	0	0	12	0	12
35	3535	임실군	0	0	12	0	12
35	3536	순창군	0	0	12	0	12
35	3537	고창군	0	0	6	6	12
35	3538	부안군	0	0	6	6	12

시군구			동부		읍면부		합계
시/도 코드	시군코드	시군구명	1.일반	2.아파트	1.일반	2.아파트	
36	3601	목포시	12	12	0	0	24
36	3602	여수시	12	12	6	6	36
36	3603	순천시	6	6	6	6	24
36	3604	나주시	3	9	6	6	24
36	3606	광양시	3	9	6	6	24
36	3631	담양군	0	0	12	0	12
36	3632	곡성군	0	0	12	0	12
36	3633	구례군	0	0	12	0	12
36	3635	고흥군	0	0	24	0	24
36	3636	보성군	0	0	12	0	12
36	3637	화순군	0	0	6	6	12
36	3638	장흥군	0	0	12	0	12
36	3639	강진군	0	0	12	0	12
36	3640	해남군	0	0	6	6	12
36	3641	영암군	0	0	6	6	12
36	3642	무안군	0	0	6	6	12
36	3643	함평군	0	0	12	0	12
36	3644	영광군	0	0	6	6	12
36	3645	장성군	0	0	6	6	12
36	3646	완도군	0	0	12	0	12
36	3647	진도군	0	0	12	0	12
36	3648	신안군	0	0	12	0	12

시군구			동부		읍면부		합계
시/도 코드	시군코드	시군구명	1.일반	2.아파트	1.일반	2.아파트	
37	3701	포항시	12	12	6	6	36
37	3702	경주시	6	6	6	6	24
37	3703	김천시	6	12	6	0	24
37	3704	안동시	6	6	6	6	24
37	3705	구미시	12	12	6	6	36
37	3706	영주시	12	6	6	0	24
37	3707	영천시	6	6	12	0	24
37	3708	상주시	3	3	6	0	12
37	3709	문경시	3	3	6	0	12
37	3710	경산시	6	6	6	6	24
37	3731	군위군	0	0	12	0	12
37	3732	의성군	0	0	12	0	12
37	3733	청송군	0	0	12	0	12
37	3734	영양군	0	0	12	0	12
37	3735	영덕군	0	0	12	0	12
37	3736	청도군	0	0	12	0	12
37	3737	고령군	0	0	6	6	12
37	3738	성주군	0	0	12	0	12
37	3739	칠곡군	0	0	12	12	24
37	3740	예천군	0	0	6	6	12
37	3741	봉화군	0	0	12	0	12
37	3742	울진군	0	0	6	6	12
37	3743	울릉군	0	0	12	0	12

시군구			동부		읍면부		합계
시/도 코드	시군코드	시군구명	1.일반	2.아파트	1.일반	2.아파트	
38	3803	진주시	12	12	6	6	36
38	3805	통영시	6	6	6	6	24
38	3806	사천시	3	3	3	3	12
38	3807	김해시	12	12	6	6	36
38	3808	밀양시	6	6	6	6	24
38	3809	거제시	6	12	3	3	24
38	3810	양산시	3	9	6	6	24
38	3811	창원시	24	24	6	6	60
38	3831	의령군	0	0	12	0	12
38	3832	함안군	0	0	6	6	12
38	3833	창녕군	0	0	6	6	12
38	3834	고성군	0	0	6	6	12
38	3835	남해군	0	0	6	6	12
38	3836	하동군	0	0	6	6	12
38	3837	산청군	0	0	12	0	12
38	3838	함양군	0	0	6	6	12
38	3839	거창군	0	0	6	6	12
38	3840	합천군	0	0	12	0	12

시군구			동부		읍면부		합계
시/도 코드	시군코드	시군구명	1.일반	2.아파트	1.일반	2.아파트	
39	3901	제주시	24	6	12	6	48
39	3902	서귀포시	12	6	12	6	36