

Panel Survey of Employment for the persons with Disabilities  
(PSED)

**장애인고용패널조사**  
**2차웨이브 1~8차년도 조사자료**

**User's Guide**

한국장애인고용공단 고용개발원





# 목 차

<b>제1장 장애인고용패널조사 소개</b> .....	<b>1</b>
1. 조사배경 .....	1
2. 조사연혁 .....	1
3. 조사개요 .....	6
4. 패널현황 .....	9
<b>제2장 표본설계 및 가중치</b> .....	<b>13</b>
1. 표본설계 .....	13
2. 표본추출 .....	14
3. 1차조사 가중치 산출 및 추정 .....	18
4. 2차조사 가중치 산출 및 추정 .....	22
5. 3차조사 가중치 산출 및 추정 .....	29
6. 4차조사 가중치 산출 및 추정 .....	41
7. 5차조사 가중치 산출 및 추정 .....	51
8. 6차조사 가중치 산출 및 추정 .....	67
9. 7차조사 가중치 산출 및 추정 .....	81
10. 8차조사 가중치 산출 및 추정 .....	95
<b>제3장 자료의 주요 특성</b> .....	<b>109</b>
1. 조사표 체계 및 구성 .....	109
2. 주요 설문항목 구성 .....	111
<b>제4장 데이터 및 변수 가이드</b> .....	<b>124</b>
1. 데이터 구조 .....	124
2. 변수 가이드 .....	128
<b>제5장 주제별 가이드</b> .....	<b>139</b>
1. 개인 자료 .....	139
2. 가구 자료 .....	160
3. 직업력 자료 .....	167
<b>[부록] 데이터의 활용</b> .....	<b>172</b>

## 표 목차

〈표 1-1〉 조사연혁 .....	4
〈표 1-2〉 장애인고용패널 1차웨이브, 2차웨이브 비교 .....	5
〈표 1-3〉 주요 조사내용 .....	7
〈표 1-4〉 조사연도별 패널유지율 현황 (단위: 명, %) .....	9
〈표 1-5〉 연속참여패널 현황 (단위: 명, %) .....	10
〈표 1-6〉 연속참여패널 현황 (단위: 명, %) (계속) .....	11
〈표 1-7〉 조사연도별 패널유지율 세부현황 .....	12
〈표 2-1〉 읍면동 표본추출을 위한 모집단 층화 .....	14
〈표 2-2〉 장애유형, 장애정도별 표본크기 및 실업률의 오차한계 .....	15
〈표 2-3〉 층별 표본배분 결과 .....	16
〈표 2-4〉 읍면동 표본추출 결과 .....	17
〈표 2-5〉 장애유형 및 장애등급별 표본분포 현황(2상 추출 대상) .....	17
〈표 2-6〉 가중치에 대한 기술통계량 .....	19
〈표 2-7〉 주요 변수별 벤치마킹 모집단 분포 - 성/연령별 .....	19
〈표 2-8〉 주요 변수별 벤치마킹 모집단 분포 - 시도별 .....	19
〈표 2-9〉 주요 변수별 벤치마킹 모집단 분포 - 경제활동상태별 .....	20
〈표 2-10〉 주요 변수별 벤치마킹 모집단 분포 - 장애유형/장애등급별 .....	20
〈표 2-11〉 주요 비율에 대한 추정치와 상대표준오차 .....	20
〈표 2-12〉 최종가중치에 대한 기술통계량 .....	21
〈표 2-13〉 경제활동상태 변수 모집단 분포와 표본 분포 .....	21
〈표 2-14〉 무응답 사유 빈도표 .....	22
〈표 2-15〉 성별 무응답 현황 .....	22
〈표 2-16〉 시도별 무응답 현황 .....	23
〈표 2-17〉 연령대별 무응답 현황 .....	23
〈표 2-18〉 경제활동상태별 무응답 현황 .....	24
〈표 2-19〉 장애유형별 무응답 현황 .....	24
〈표 2-20〉 장애정도별 무응답 현황 .....	25
〈표 2-21〉 최종 로지스틱 회귀분석 결과 .....	25
〈표 2-22〉 주요 변수별 벤치마킹 모집단 분포 - 성/연령별 .....	26
〈표 2-23〉 주요 변수별 벤치마킹 모집단 분포 - 시도별 .....	27
〈표 2-24〉 주요 변수별 벤치마킹 모집단 분포 - 경제활동상태별 .....	27
〈표 2-25〉 주요 변수별 벤치마킹 모집단 분포 - 장애유형/장애정도별 .....	28
〈표 2-26〉 최종가중치에 대한 기술통계량 .....	28
〈표 2-27〉 경제활동상태 변수의 모집단 분포와 표본 분포 .....	28
〈표 2-28〉 1차년도, 2차년도 가중치 분포 현황 .....	29

〈표 2-29〉 1~3차조사 응답 분류 현황 .....	29
〈표 2-30〉 무응답 사유 빈도표(3차조사 중단면 기준) .....	30
〈표 2-31〉 무응답 사유 빈도표(3차조사 횡단면 기준) .....	30
〈표 2-32〉 1~3차조사 응답 분류 현황 .....	30
〈표 2-33〉 성별 무응답 현황 .....	31
〈표 2-34〉 연령대별 무응답 현황 .....	31
〈표 2-35〉 시도별 무응답 현황 .....	32
〈표 2-36〉 경제활동상태별 무응답 현황 .....	32
〈표 2-37〉 장애유형별 무응답 현황 .....	33
〈표 2-38〉 장애정도별 무응답 현황 .....	33
〈표 2-39〉 로지스틱 회귀분석 결과 .....	34
〈표 2-40〉 3차조사 최종 중단 가중치에 대한 기술통계량 .....	34
〈표 2-41〉 1~3차조사의 중단가중치 분포 .....	35
〈표 2-42〉 성별 무응답 현황 .....	35
〈표 2-43〉 연령대별 무응답 현황 .....	36
〈표 2-44〉 시도별 무응답 현황 .....	36
〈표 2-45〉 경제활동상태별 무응답 현황 .....	37
〈표 2-46〉 장애유형별 무응답 현황 .....	37
〈표 2-47〉 장애정도별 무응답 현황 .....	37
〈표 2-48〉 로지스틱 회귀분석 결과 .....	38
〈표 2-49〉 3차조사 최종 횡단가중치에 대한 기술통계량 .....	40
〈표 2-50〉 1~3차조사의 횡단가중치 분포 .....	40
〈표 2-51〉 성별 무응답 현황 .....	41
〈표 2-52〉 연령대별 무응답 현황 .....	41
〈표 2-53〉 시도별 무응답 현황 .....	42
〈표 2-54〉 경제활동상태별 무응답 현황 .....	42
〈표 2-55〉 장애유형별 무응답 현황 .....	42
〈표 2-56〉 장애정도별 무응답 현황 .....	43
〈표 2-57〉 로지스틱 회귀분석 결과 .....	44
〈표 2-58〉 4차조사 최종 중단 가중치에 대한 기술통계량 .....	44
〈표 2-59〉 1~4차조사의 중단 가중치 분포 .....	44
〈표 2-60〉 성별 무응답 현황 .....	45
〈표 2-61〉 연령대별 무응답 현황 .....	45
〈표 2-62〉 시도별 무응답 현황 .....	46
〈표 2-63〉 경제활동상태별 무응답 현황 .....	46
〈표 2-64〉 장애유형별 무응답 현황 .....	46
〈표 2-65〉 장애정도별 무응답 현황 .....	47
〈표 2-66〉 로지스틱 회귀분석 결과 .....	47
〈표 2-67〉 1~4차년도 장애정도 모집단 분포 현황 .....	48

<표 2-68> 1~4차년도 지체 3급내 중증 장애 비율 .....	48
<표 2-69> 1~4차년도 장애등급 모집단 분포 현황 .....	49
<표 2-70> 1~4차년도 지체외/내 장애정도 모집단 분포 현황 .....	49
<표 2-71> 4차조사 최종 횡단 가중치에 대한 기술통계량 .....	50
<표 2-72> 1~4차조사의 횡단 가중치 분포 .....	50
<표 2-73> 시도별 종단면 무응답 현황 .....	51
<표 2-74> 성별 종단면 무응답 현황 .....	52
<표 2-75> 연령대별 종단면 무응답 현황 .....	52
<표 2-76> 경제활동상태별 종단면 무응답 현황 .....	52
<표 2-77> 장애유형별 종단면 무응답 현황 .....	53
<표 2-78> 장애정도별 종단면 무응답 현황 .....	53
<표 2-79> 모형별 종단면 로지스틱 회귀분석 결과 .....	54
<표 2-80> 5차년도 종단 가중치에 대한 통계량 .....	55
<표 2-81> 1-5차 조사의 종단 가중치 분포 .....	55
<표 2-82 시도별 횡단면 무응답 현황 .....	56
<표 2-83> 성별 횡단면 무응답 현황 .....	57
<표 2-84> 연령대별 횡단면 무응답 현황 .....	57
<표 2-85> 경제활동상태별 횡단면 무응답 현황 .....	57
<표 2-86> 장애유형별 횡단면 무응답 현황 .....	58
<표 2-87> 장애정도별 횡단면 무응답 현황 .....	58
<표 2-88> 모형별 횡단면 로지스틱 회귀분석 결과 .....	59
<표 2-89> 1차-4차년도 고용노동부 기준 장애정도 모집단 현황 .....	60
<표 2-90> 1차-4차년도 장애등급 모집단 분포 현황 .....	60
<표 2-91> 1차-4차년도 보건복지부 기준 장애정도 모집단 현황 .....	61
<표 2-92> 5차년도 장애정도 모집단 현황(2019년) .....	61
<표 2-93> 전체 장애인과 3급 장애인의 장애유형 분포 .....	61
<표 2-94> 전체 장애인과 3급 장애인의 성별 분포 .....	62
<표 2-95> 전체 장애인과 3급 장애인의 연령 분포 .....	62
<표 2-96> 전체 장애인과 3급 장애인의 경제활동상태 분포 .....	62
<표 2-97> 횡단 가중치에 대한 통계량 .....	63
<표 2-98> 성별 모집단 및 표본 분포 .....	63
<표 2-99> 연령별 모집단 및 표본 분포 .....	63
<표 2-100> 경제활동상태 변수의 모집단 및 표본 분포 .....	63
<표 2-101> 장애유형별 변수의 모집단 및 표본 분포 .....	64
<표 2-102> 시도별 모집단 및 표본 분포 .....	64
<표 2-103> 장애정도별 변수의 모집단 및 표본 분포 .....	64
<표 2-104> 1-5차년도 횡단가중치 분포 .....	65
<표 2-105> 장애유형 (15개) 별 경제활동상태(2개 구분) 빈도/비율 추정 및 상대표준오차 분석 .....	65
<표 2-106> 장애유형 (15개) 별 경제활동상태(2개 구분) 빈도/비율 추정 및 상대표준오차 분석(계속) .....	66

<표 2-107> 시도별 중단면 무응답 현황 .....	67
<표 2-108> 성별 중단면 무응답 현황 .....	68
<표 2-109> 연령대별 중단면 무응답 현황 .....	68
<표 2-110> 경제활동상태별 중단면 무응답 현황 .....	68
<표 2-111> 장애유형별 중단면 무응답 현황 .....	69
<표 2-112> 장애정도별 중단면 무응답 현황 .....	69
<표 2-113> 모형별 중단면 로지스틱 회귀분석 결과 .....	71
<표2-114> 6차년도 중단 가중치에 대한 통계량 .....	71
<표 2-115> 2-6차조사의 중단 가중치 분포 .....	71
<표 2-116> 시도별 횡단면 무응답 현황 .....	72
<표 2-117> 성별 횡단면 무응답 현황 .....	73
<표 2-118> 연령대별 횡단면 무응답 현황 .....	73
<표 2-119> 경제활동상태별 횡단면 무응답 현황 .....	73
<표 2-120> 장애유형별 횡단면 무응답 현황 .....	74
<표 2-121> 장애정도별 횡단면 무응답 현황 .....	74
<표 2-122> 모형별 횡단면 로지스틱 회귀분석 결과 .....	75
<표 2-123> 횡단 가중치에 대한 통계량 .....	76
<표 2-124> 성별 모집단 및 표본 분포 .....	77
<표 2-125> 연령별 모집단 및 표본 분포 .....	77
<표 2-126> 경제활동상태 변수의 모집단 및 표본 분포 .....	77
<표 2-127> 장애유형별 변수의 모집단 및 표본 분포 .....	77
<표 2-128> 시도별 모집단 및 표본 분포 .....	78
<표 2-129> 장애정도별 변수의 모집단 및 표본 분포 .....	78
<표 2-130> 1-6차년도 횡단가중치 분포 .....	79
<표 2-131> 장애유형 (15개) 별 경제활동상태(2개 구분) 빈도/비율 추정 및 상대표준오차 분석 .....	79
<표 2-132> 장애유형 (15개) 별 경제활동상태(2개 구분) 빈도/비율 추정 및 상대표준오차 분석(계속) .....	80
<표 2-133> 시도별 중단면 무응답 현황 .....	81
<표 2-134> 성별 중단면 무응답 현황 .....	82
<표 2-135> 연령대별 중단면 무응답 현황 .....	82
<표 2-136> 경제활동상태별 중단면 무응답 현황 .....	82
<표 2-137> 장애유형별 중단면 무응답 현황 .....	83
<표 2-138> 장애정도별 중단면 무응답 현황 .....	83
<표 2-139> 모형별 중단면 로지스틱 회귀분석 결과 .....	85
<표 2-140> 7차년도 중단 가중치에 대한 통계량 .....	85
<표 2-141> 3-7차조사의 중단 가중치 분포 .....	85
<표 2-142> 시도별 횡단면 무응답 현황 .....	86
<표 2-143> 성별 횡단면 무응답 현황 .....	87
<표 2-144> 연령대별 횡단면 무응답 현황 .....	87
<표 2-145> 경제활동상태별 횡단면 무응답 현황 .....	87

<표 2-146> 장애유형별 횡단면 무응답 현황 .....	88
<표 2-147> 장애정도별 횡단면 무응답 현황 .....	88
<표 2-148> 모형별 횡단면 로지스틱 회귀분석 결과 .....	89
<표 2-149> 횡단 가중치에 대한 통계량 .....	90
<표 2-150> 성별 모집단 및 표본 분포 .....	91
<표 2-151> 연령별 모집단 및 표본 분포 .....	91
<표 2-152> 경제활동상태 변수의 모집단 및 표본 분포 .....	91
<표 2-153> 장애유형별 변수의 모집단 및 표본 분포 .....	91
<표 2-154> 시도별 모집단 및 표본 분포 .....	92
<표 2-155> 장애정도별 변수의 모집단 및 표본 분포 .....	92
<표 2-156> 1-7차년도 횡단가중치 분포 .....	93
<표 2-157> 장애유형 (15개) 별 경제활동상태(2개 구분) 빈도/비율 추정 및 상대표준오차 분석 .....	93
<표 2-158> 장애유형 (15개) 별 경제활동상태(2개 구분) 빈도/비율 추정 및 상대표준오차 분석(계속) ..	94
<표 2-159> 시도별 종단면 무응답 현황 .....	95
<표 2-160> 성별 종단면 무응답 현황 .....	96
<표 2-161> 연령대별 종단면 무응답 현황 .....	96
<표 2-162> 경제활동상태별 종단면 무응답 현황 .....	96
<표 2-163> 장애유형별 종단면 무응답 현황 .....	97
<표 2-164> 장애정도별 종단면 무응답 현황 .....	97
<표 2-165> 모형별 종단면 로지스틱 회귀분석 결과 .....	99
<표 2-166> 8차년도 종단 가중치에 대한 통계량 .....	99
<표 2-167> 4-8차조사의 종단 가중치 분포 .....	99
<표 2-168> 시도별 횡단면 무응답 현황 .....	100
<표 2-169> 성별 횡단면 무응답 현황 .....	101
<표 2-170> 연령대별 횡단면 무응답 현황 .....	101
<표 2-171> 경제활동상태별 횡단면 무응답 현황 .....	101
<표 2-172> 장애유형별 횡단면 무응답 현황 .....	102
<표 2-173> 장애정도별 횡단면 무응답 현황 .....	102
<표 2-174> 모형별 횡단면 로지스틱 회귀분석 결과 .....	103
<표 2-175> 횡단 가중치에 대한 통계량 .....	104
<표 2-176> 성별 모집단 및 표본 분포 .....	105
<표 2-177> 연령별 모집단 및 표본 분포 .....	105
<표 2-178> 경제활동상태 변수의 모집단 및 표본 분포 .....	105
<표 2-179> 장애유형별 변수의 모집단 및 표본 분포 .....	105
<표 2-180> 시도별 모집단 및 표본 분포 .....	106
<표 2-181> 장애정도별 변수의 모집단 및 표본 분포 .....	106
<표 2-182> 1-8차년도 횡단가중치 분포 .....	107
<표 2-183> 장애유형 (15개) 별 경제활동상태(2개 구분) 빈도/비율 추정 및 상대표준오차 분석 .....	107
<표 2-184> 장애유형 (15개) 별 경제활동상태(2개 구분) 빈도/비율 추정 및 상대표준오차 분석(계속) ..	108



〈표 3-1〉	차수별 조사표 구성 .....	110
〈표 3-2〉	설문 항목 구성 - A. 패널 기본정보 .....	111
〈표 3-3〉	설문 항목 구성 - PB. 지난조사 이후 일자리 현황 .....	112
〈표 3-4〉	설문 항목 구성 - B. 경제활동상태 판별 .....	112
〈표 3-5〉	설문 항목 구성 - B. 경제활동상태 판별(계속) .....	113
〈표 3-6〉	개정 한국중사상지위분류 체계(통계청, 2021.12.29., 보도자료) .....	114
〈표 3-7〉	설문 항목 구성 - Ca. 임금근로자 .....	115
〈표 3-8〉	설문 항목 구성 - Cb. 비임금근로자 .....	116
〈표 3-9〉	설문 항목 구성 - Cc. 미취업자 .....	117
〈표 3-10〉	설문 항목 구성 - D. 취업을 위한 노력 .....	118
〈표 3-11〉	설문 항목 구성 - E. 직업적 능력 .....	118
〈표 3-12〉	설문 항목 구성 - F. 취업관련 태도 및 환경 .....	119
〈표 3-13〉	설문 항목 구성 - G. 일상생활과 삶의 질 .....	120
〈표 3-14〉	설문 항목 구성 - H. 가구 일반사항 .....	121
〈표 3-15〉	직업력 조사표 유형 구분 .....	122
〈표 3-16〉	설문 항목 구성 - JOB. 직업력 .....	122
〈표 3-17〉	설문 항목 구성 - I. 코로나19 지속 .....	123
〈표 4-1〉	장애인고용패널조사 데이터 제공 유형별 특징 .....	127
〈표 4-2〉	주요 변수 구성 .....	128
〈표 4-3〉	경제활동상태 및 종사상 지위 변수값 구성 .....	130
〈표 4-4〉	경제활동상태 및 종사상 지위 변수값 구성(계속) .....	130
〈표 4-5〉	경제활동상태 및 종사상 지위 분류기준 .....	131
〈표 4-6〉	일자리 ID 부여 규칙 .....	132
〈표 4-7〉	일자리 ID 현황 .....	132
〈표 4-8〉	주요 key 변수 현황 .....	133
〈표 4-9〉	조사표 파트 구분 .....	134
〈표 4-10〉	WORK2 변수 내용 .....	136
〈표 4-11〉	월평균 소득 변수 일람(연속형/범주형/범주형 가공/무응답 대체) .....	137
〈표 4-12〉	한국표준산업분류 코드(제10차 개정, 대분류 기준) .....	138
〈표 4-13〉	한국표준직업분류 코드(제7차 개정, 대분류 기준) .....	138
[참고사항]	.....	139
표 15	.....	139
〈표 5-1〉	패널 기본정보 변수 .....	140
〈표 5-2〉	패널 장애정보 변수 .....	141
〈표 5-3〉	일자리 정보 확인 및 경제활동상태 판별 변수 .....	142
〈표 5-4〉	취업자의 근로시간 관련 변수 .....	143
〈표 5-5〉	취업자의 소득 관련 변수 .....	144
〈표 5-6〉	취업자의 근로환경 관련 변수 .....	145
〈표 5-7〉	취업자의 직무만족도 변수 .....	146

〈표 5-8〉	취업자의 전반적 직무만족도 변수	146
〈표 5-9〉	임금근로자 주요 변수	147
〈표 5-10〉	비정규직 관별 변수	148
〈표 5-11〉	비임금근로자 주요 변수	149
〈표 5-12〉	미취업자 주요 변수	149
〈표 5-13〉	취업 및 취업유지활동 관련 변수	150
〈표 5-14〉	공공 취업알선기관 및 직업교육훈련 이용경험 변수	151
〈표 5-15〉	취업관련 필요사항 및 지원사항, 추가 소요비용 관련 변수	152
〈표 5-16〉	직업적 능력 주요변수	152
〈표 5-17〉	자격증 유형 분류 기준	153
〈표 5-18〉	자격증 주요 변수	153
〈표 5-19〉	자격증 주요 변수(계속)	154
〈표 5-20〉	자기효능감 변수	155
〈표 5-21〉	자아존중감 변수	156
〈표 5-22〉	장애수용정도 변수	156
〈표 5-23〉	일상생활 관련 주요 변수	157
〈표 5-24〉	패널 개인소득 관련 변수	158
〈표 5-25〉	노후준비 관련 변수	158
〈표 5-26〉	코로나19 관련 주요 변수	159
〈표 5-27〉	기존 가구원 관련 변수*	161
〈표 5-28〉	신규 가구원 관련 변수*	162
〈표 5-29〉	최종 가구원 현황 가공변수*	163
〈표 5-30〉	가구 근로소득 관련 변수	164
〈표 5-31〉	가구 근로 외 소득 관련 변수	165
〈표 5-32〉	생활비 관련 변수	166
〈표 5-33〉	일자리 종류(jobtype) 구성 체계	167
〈표 5-34〉	직업력 KEY 변수(WORK1)	168
〈표 5-35〉	직업력 주제별 변수(임금근로자, WORK1)	169
〈표 5-36〉	직업력 주제별 변수(비임금근로자, WORK1)	170
〈표 5-37〉	직업력 일자리별 핵심 정보 변수(WORK2)	171

## 그림 목차

[그림 3-1]	조사표 구성도 .....	109
[그림 3-2]	임금근로자 종사상지위 개정안 비교 - 기간별 .....	114
[그림 4-1]	메인 데이터 구성 .....	124
[그림 4-2]	통합 데이터 구성 .....	125
[그림 4-2]	직업력 데이터 구성(WORK2) .....	126

## 제1장

# 장애인고용패널조사(PSED) 소개

### 1. 조사배경

패널조사는 동일한 개인 또는 집단을 추적 조사함으로써 일반적인 횡단면 조사가 제공할 수 없는 심층적이고 다차원적인 정보를 제공한다. 패널 데이터는 특히 노동시장 진입 및 이동과정, 고용패턴 및 선호의 변화, 생애주기별 경제활동특성 등 노동시장 분석에 유용한 자료로써 고용정책의 수립 및 평가에 필요한 필수적인 정보를 제공한다. 특히 전체 인구에 비해 경제활동상태가 현저히 열악한 상태에서 관련 정보마저 부재한 장애인의 경우에는 이러한 특성과 제반 요인들을 종합적으로 정확하게 파악하게 하는 패널 조사의 실시는 필수적이라 할 수 있다.

이러한 필요성에 의해 고용노동부와 한국장애인고용공단은 매년 동일 장애인을 대상으로 반복하여 조사하는 장애인고용패널 1차웨이브를 2008년 도입하였다. 장애인고용패널 1차웨이브는 2008년 1월 1일 만 15~75세의 등록장애인 5,092명을 대상으로 1년에 1회 경제활동 및 직업적 능력, 고용서비스, 직업 능력개발, 일상생활, 소득/소비 등의 내용에 대해 추적조사하는 종단면 조사로서 2015년 8차조사까지 진행하였다.

### 2. 조사연혁

#### 가. 1차웨이브 조사 결과

장애인고용패널 1차웨이브의 연도별 진행경과를 살펴보면, 먼저 2006년 장애인고용패널 구축 및 패널 조사를 위한 기초연구를 수행하였다. 2007년에는 패널구축조사를 통해 본 패널과 예비 패널을 구축하였으며, 경제활동상태와 관련 요인들을 파악하기 위한 조사표 개발과 동시에 정확한 자료수집을 위해 컴퓨터를 활용한 대인면접방법인 CAPI(Computer-Assisted Personal Interviewing) 프로그램을 개발하고 세 차례 시범조사를 실시하였다.

1차조사는 2008년 6월부터 8월까지 실시되었으며 이를 통해 총 5,092명의 등록장애인이 장애인고용패널로 최종 확정되었다. 1차조사는 비록 패널조사일지라도 여타의 횡단면 조사와 다를 바가 없이 진행되었으며, 조사내용은 패널의 경제활동 관련 현재 상황 및 과거 경력 등이었다. 2009년 2차조사부터는 매년 5월부터 7월까지 실시되었으며 지난 1년간의 변화와 경험, 직업력 등에 대한 조사가 추가적으로 이뤄짐으로써 종단면 자료 축적이 가능해지고, 패널 데이터로서의 가치를 갖게 되었다.

패널조사에서는 최초 추출된 표본을 이후 조사에서도 얼마나 많이 조사하는가가 중요한데 장애인고용패널 1차웨이브는 장애인 패널과의 유대감 증진을 위해 본 조사 전 조사원이 직접 방문하여 안부를 묻고 선물을 전달하는 등 적극적인 패널관리 정책을 사용하였다. 그 결과, 8차조사에서는 원표본 5,092명 중 총 3,983명을 조사하는 데 성공해 국내외 패널조사 중 최고 수준의 패널유지율 78.2%를 기록하였다. 이는 사망, 장애등록 취소 등으로 영구 탈락한 패널을 포함하여 산출한 수치로, 이러한 영구 탈락자를 제외하면 패널유지율(유효패널 유지율)은 86.3%까지 높아진다.

또한 장애인고용패널 자료의 정책적, 학문적 활용도를 높이고자 대학 및 연구기관의 전문 연구자와 대학원생들에게 데이터를 제공하고 그 연구 성과물을 중심으로 2009년부터 매년 장애인고용패널 학술대회를 개최하는 등 다각적인 공유, 확산 노력을 실시하였으며, 이러한 노력으로 장애인고용패널 1차웨이브는 장애인고용정책의 수립·평가에 직·간접적으로 활용되었고, 정책발전을 뒷받침하는 학계의 연구활성화에도 긍정적인 역할을 수행해 왔다.

한편 1차웨이브가 성공적으로 진행되었지만 장애인 개인을 추적조사하는 중단설계가 안고 있는 구조적인 문제점 역시 차수가 진행되면서 누적되었다. 아무리 패널관리를 철저히 하더라도 사망 등 불가항력적인 이유로 이탈이 발생하게 되고, 연령의 자연 증가로 청년층이 소실되는 등 조사결과가 전체 장애인을 대표하기 어렵게 되었다. 또한 전체 인구보다 심각한 장애인의 고령화 문제는 장애인고용패널 1차웨이브에도 똑같이 반영되어 패널의 경제활동 관련 움직임을 둔화시키는 요인으로 작용하였고, 그로 인해 노동시장 이동, 고용서비스 이용, 직업능력 배양 등 경제활동과 관련된 충분한 분석 표본 수의 확보를 어렵게 만들었다. 이는 경제활동 관련 다양한 경험과 변화에 대한 정보를 충실히 제공해야 하는 패널조사로서의 역할 수행을 재검토할 필요가 있음을 의미한다.

또한 2008년 최초 구축 당시에는 장애인 경제활동과 관련된 통계자료의 부재로 횡단면 및 패널 정보 생산이라는 두 가지 목적이 공존하였고, 그에 따라 패널이 우리나라 전체 장애인을 대표해야 할 필요가 있었으며, 조사내용도 두 가지 목적을 동시에 달성할 수 있도록 설계되었다. 그러다보니 2013년부터 매년 실시하고 있는 장애인경제활동실태조사 등 타 조사와의 명확한 차별화와 역할분담도 필요하게 되었다.

이러한 한계와 문제를 극복하고자, 2014년부터 관계 기관과 전문가들과 함께 치밀한 논의과정을 거치면서 효과성과 효율성 등을 감안할 때 기존 패널을 일부 보완해서 유지하는 것보다 2차웨이브를 새롭게 구축하는 것이 여러모로 적절하다는 결론에 도달하였다. 그에 따라 기존 1차웨이브는 2015년 8차 조사로 종료하고, 2016년 새로운 패널과 새로운 설계로 장애인고용패널 2차웨이브를 구축하게 되었다.

## 나. 2차웨이브 조사 목적 및 방향성

장애인고용패널 2차웨이브는 다음과 같은 방향으로 구축하였다. 첫째, 1차웨이브의 기본틀을 유지하여 시계열적 연속성과 비교 가능성을 제고하고자 하였다. 그에 따라 장애인의 노동시장 이동과정을 파악하고 이에 영향을 주고받는 내·외부 환경요인을 도출하고자 하는 기본 목적과 방향은 동일하다.

둘째, 장애인고용패널 2차웨이브는 정책연구형 패널조사로 구축하여 데이터의 활용도를 제고하고 기존 조사와 명확한 차별화를 꾀하였다. 그에 따라 선택과 집중을 통해 패널조사에서만 파악할 수 있는 내용들과 정책연구에 활용할 수 있는 쓸만한 변수 위주로 구성하고자 하였다.

셋째, 무엇보다도 경제활동, 고용과 관련된 분석 표본 수를 최대한 확보할 수 있는 방안을 접목시켰다. 그에 따라 청장년 장애인과 취업자, 실업자, 취업을 희망하는 비경제활동인구 등을 실제 표본 비율보다 다수 구축하였으며, 다양한 주제, 영역에 대해 누구나 보편타당하게 응답할 수 있는 내용들로 구성하였다.

#### 다. 2차웨이브 패널구축 및 1차조사(2016년)

2차웨이브의 패널구축을 위해 모집단은 2016년 5월 15일 기준 우리나라에 거주하는 「장애인복지법」상 등록장애인으로, 핵심 생산가능인구를 대변할 수 있도록 만 15~64세 연령을 대상으로 하였다.

사용한 표집틀은 2015년 12월 31일 기준 보건복지부의 등록장애인 명부이며, 표본배분 및 추출은 표본 설계 전문가에 의한 연구결과에 따라 지역, 장애유형, 장애등급, 연령, 경제활동상태 등을 고려하여 실시하였고, 추출 읍면동 수를 조절함과 더불어 장애유형 및 장애등급별 적절한 표본을 추출하기 위해 이 상추출(two phase sampling)을 활용하였다.

최종적으로 4,577명의 패널이 구축되어 1차조사에 참여하였다.

#### 라. 2차웨이브 2차조사(2017년)

2017년에는 패널조사의 특징이라고 할 수 있는 종단면 구조를 설계 및 구축하여 2차웨이브 2차조사를 실시하였다. 2차조사는 지난조사의 응답결과를 재확인하고 이후 변화를 질문하여 현재 최종 상태를 확인하는 종단조사표를 구성하여 시계열 자료 수집의 정확도를 제고하였다.

#### 마. 2차웨이브 3차조사(2018년)

2018년 3차조사는 상기한 종단조사표 설계를 적용하여 실시되었으며, 이에 따라 매년마다 축적되는 시계열적 자료를 활용하여 장애인 노동시장의 역동적인 흐름과 현황을 직접적으로 살펴볼 수 있는 기반이 마련되었다.

#### 바. 2차웨이브 4차조사(2019년)

2019년 4차조사는 2, 3차조사와 동일하게 종단조사표 형태를 유지하였고, 최근 노동시장 이슈인 최저임금 관련 문항과 향후 고용 정책 근거자료 활용을 위해 중증장애인 고용유지 관련 문항을 추가하였다. 이는 노동현안 이슈를 반영하고 수요자 요구 중심의 항목을 추가하고자 하는 노력이라고 할 수 있다.

#### 사. 2차웨이브 5차조사(2020년)

2020년 5차조사는 코로나19바이러스감염증-19로 인해 조사 시점을 기 5월~8월에서 6월~9월로 순연하였고 조사표 또한 2~4차조사와 동일하게 종단조사표 형태를 유지하되 코로나19바이러스감염증-19와 관련한 전사회적 대응과 장애등급제 폐지 등 관련 문항을 추가 및 변경하여 사회변화와 고용현안을 반영하였다.

### 아. 2차웹이브 6차조사(2021년)

2021년 6차조사는 코로나19바이러스감염증-19의 지속으로 일상생활 및 인식·삶의 질 변화 등 코로나19의 영향력과 이와 관련한 전사회적 대응 현안을 보다 심도 있게 파악하기 위한 코로나19 부가조사 파트를 신설하였다.

### 자. 2차웹이브 7차조사(2022년)

2022년 7차조사는 중대재해처벌법의 도입에 따라 근로 시 발생 가능한 안전관련 위험 요소, 재난 안전 대응을 위해 필요한 요소 등 안전 관련한 문항을 추가하였다. 더불어 코로나바이러스감염증-19 전후 영향력을 파악하기 위해 6차조사 시 신설된 코로나 부가조사 파트를 유지하였다.

### 차. 2차웹이브 8차조사(2023년)

2023년 8차조사는 개정된 한국중사상지위분류의 반영을 위해 임금근로자 파트에 일부 문항을 추가하였으며, 일·가정 양립 측면에서 살펴볼 수 있도록 여가와 가사, 돌봄노동 시간 문항을 신설하였다. 더불어 6차조사 시 신설된 코로나바이러스감염증-19 파트를 유지하였다.

〈표 1-1〉 조사연혁

구 분	연 혁
장애인고용패널 1차웹이브	<ul style="list-style-type: none"> <li>· 2006년: 장애인고용패널 1차웹이브 구축을 위한 기초연구 실시</li> <li>· 2007년: 장애인고용패널 1차웹이브 구축조사 실시</li> <li>· 2008년 ~ 2015년: 장애인고용패널 1차웹이브 1차조사 ~ 8차조사 실시</li> <li>· 2015년: 장애인고용패널 1차웹이브 종료</li> </ul>
장애인고용패널 2차웹이브	<ul style="list-style-type: none"> <li>· 2015년: 장애인고용패널 2차웹이브 구축계획 수립</li> <li>· 2016년: 장애인고용패널 2차웹이브 구축 및 1차조사 실시</li> <li>· 2017년: 2017년 장애인고용패널조사(2차웹이브 2차조사) 실시</li> <li>· 2018년: 2018년 장애인고용패널조사(2차웹이브 3차조사) 실시</li> <li>· 2019년: 2019년 장애인고용패널조사(2차웹이브 4차조사) 실시</li> <li>· 2020년: 2020년 장애인고용패널조사(2차웹이브 5차조사) 실시</li> <li>· 2021년: 2021년 장애인고용패널조사(2차웹이브 6차조사) 실시</li> <li>· 2022년: 2022년 장애인고용패널조사(2차웹이브 7차조사) 실시</li> <li>· 2023년: 2023년 장애인고용패널조사(2차웹이브 8차조사) 실시</li> </ul>

### 카. 장애인고용패널 1차웨이브, 2차웨이브 비교

2023년은 작년과 동일하게 5월~8월에 조사를 하였으며, 조사 방법은 7차년도와 동일하게 방문 면접조사에서 대면 접촉을 거절하는 경우 온라인 및 전화 조사를 병행하였다.

〈표 1-2〉 장애인고용패널 1차웨이브, 2차웨이브 비교

구 분	1차웨이브(2008년~2015년)	2차웨이브(2016년 이후)
도입배경	<ul style="list-style-type: none"> <li>· 장애인 고용 관련 통계 부족</li> <li>· 노동시장 분석 시 패널조사 효용성 대두</li> <li>- 장애인에게 패널조사 형식은 더 적합</li> </ul>	<ul style="list-style-type: none"> <li>· 장애인고용패널 1차웨이브의 성공적 수행</li> <li>- 장애인고용 관련 핵심 데이터로 자리매김</li> <li>- 최고의 패널유지율 기록</li> <li>· 기존 패널 노후화 및 장애인의 고령화</li> <li>· 경찰조사, 기업체조사 등 분야별 전문 조사 도입</li> <li>· 패널조사의 정책적 활용도 제고 필요</li> </ul>
조사목적	<ul style="list-style-type: none"> <li>· 장애인의 노동시장 이동과 그에 영향을 주고받는 요인 도출</li> <li>· 장애인 경제활동 관련 실태 전반 파악</li> <li>- 고용률/실업률 등 경제지표 도출</li> <li>=&gt; 장애인 고용 관련 통계 갈증 해소</li> <li>=&gt; 제반 연구 활성화</li> </ul>	<ul style="list-style-type: none"> <li>· 장애인의 노동시장 이동과 그에 영향을 주고받는 요인 도출</li> <li>· 패널조사를 통해서만 제공 가능한 정책 통계정보 제공(정책의 취업/유지 효과 등)</li> <li>=&gt; 통계의 정책적 활용도 제고, 차별화</li> <li>=&gt; 정책연구 활성화</li> </ul>
조사대상	<ul style="list-style-type: none"> <li>· 우리나라 15~75세 등록장애인 5,092명</li> <li>※15년 조사완료 패널 수는 3,983명</li> <li>- 15~60세, 61~75세 두 개 층 내에서 표본수 할당(9:1) 후 연령비례 추출</li> </ul>	<ul style="list-style-type: none"> <li>· 우리나라 15~64세 등록장애인 4,577명</li> <li>· 연령, 장애유형의 편차 등을 고려</li> <li>· 경제활동 및 경제활동 희망 장애인 과대할당 (경제활동 참여 및 희망 3, 비희망 1)</li> </ul>
추출틀	<ul style="list-style-type: none"> <li>· 읍면동별 등록장애인 명부</li> <li>+ 인구주택총조사 활동계약자 명부</li> </ul>	<ul style="list-style-type: none"> <li>· 보건복지부 등록장애인 명부</li> </ul>
표본설계	<ul style="list-style-type: none"> <li>· 전문가에 의한 용역 실시</li> </ul>	<ul style="list-style-type: none"> <li>· 전문가에 의한 용역 실시</li> </ul>
패널구축	<ul style="list-style-type: none"> <li>· 별도의 패널구축조사 실시 후 이듬해 1차 조사 실시</li> <li>- 패널유지율 제고에 도움</li> </ul>	<ul style="list-style-type: none"> <li>· 패널구축과 동시에 1차조사 실시</li> <li>- 표본추출의 정합성 제고 및 비용 절감</li> </ul>
조사내용	<ul style="list-style-type: none"> <li>· 인적 특성, 장애 특성, 경제활동상태 판별, 임금근로자/자영업자/무급가족종사자/실업자/비경제활동인구, 직업적 능력, 고용서비스, 직업능력개발, 일상생활과 삶의 질, 개인소득, 가구 일반사항 등</li> </ul>	<ul style="list-style-type: none"> <li>· 인적 특성, 장애 특성, 경제활동상태 판별, 임금근로자/비임금근로자/미취업자, 취업을 위한 노력과 지원, 직업적 능력, 취업 태도 및 환경, 일상생활과 삶의 질, 가구 일반사항 등</li> </ul>
조사방법	<ul style="list-style-type: none"> <li>· 숙련된 조사원 방문에 의한 면접조사</li> <li>- 전문 실사업체 용역 관리</li> </ul>	<ul style="list-style-type: none"> <li>· 숙련된 조사원 방문에 의한 면접조사</li> <li>- 전문 실사업체 용역 관리</li> <li>· 코로나19로 인해 방문 면접조사가 어려운 경우, 전화 및 온라인을 활용한 비대면조사</li> </ul>
CAPI	<ul style="list-style-type: none"> <li>· 네덜란드 CAPI 소프트웨어 매년 임대</li> <li>- 노트북 컴퓨터를 활용한 조사</li> </ul>	<ul style="list-style-type: none"> <li>· 국내 조사업체 소프트웨어 및 기술 활용</li> <li>- 태블릿 PC를 활용한 조사</li> </ul>
조사기간	<ul style="list-style-type: none"> <li>· 매년 5~7월</li> </ul>	<ul style="list-style-type: none"> <li>· 매년 5~8월</li> </ul>
결과공표	<ul style="list-style-type: none"> <li>· 조사 차년도 하반기</li> </ul>	<ul style="list-style-type: none"> <li>· 조사 차년도 상반기</li> </ul>



### 3. 조사개요

#### 가. 법적근거 및 승인번호

본 조사는 「장애인고용촉진 및 직업재활법」 제26조(장애인 실태조사)에 근거하여 실시하고 있으며, 「통계법」 제18조(통계작성의 승인)에 의한 국가 승인통계(제383003호)이다.

#### 나. 조사목적

본 조사의 목적은 크게 두 가지다. 첫 번째는 2016년에 구축된 장애인고용패널 2차웨이브의 패널 유지율을 높은 수준으로 유지하여 종단조사로서 양질의 시계열 데이터를 구축하는 것이다. 두 번째는 본 조사를 통해 활용도 높은 패널데이터를 제공하여 장애인의 노동시장(일자리) 이동과정(입직-이동-은퇴)을 파악하고 이에 영향을 주고받는 내·외부 환경요인을 도출하는 데 있다. 이를 통해 합리적인 장애인고용정책 수립 및 평가, 관련 연구를 수행하기 위한 기초자료를 제공하고자 한다.

#### 다. 조사대상

지난 2016년 장애인고용패널조사(2차웨이브 1차조사)에서 구축한 본 패널 4,577명 중 「장애인복지법」에서 규정하고 있는 15개 유형의 장애를 가지고 있는 등록장애인 4,577명을 조사 대상으로 한다. 본 조사는 장애인의 경제활동상태를 파악하는 것이 목적이므로 이에 부합되도록 패널의 연령을 2016년 5월 15일 기준 만 15~64세 연령을 대상으로 한다.

1차웨이브(5,092명)에 비해 패널 수는 다소 감소하였지만, 청장년층 패널 증대, 경제활동 참여 및 희망 패널(참여자 및 희망자, 비희망자를 3:1의 비율로 할당) 증대 등으로 시간이 지날수록 분석 표본 수는 오히려 증가할 수 있도록 설계, 구축하였다.

#### 라. 조사내용

1차조사는 패널의 경제활동 상태에 따라 해당하는 경제활동 유형 문항에 응답하는 형태로 구성하였고, 세부적으로는 '패널 기본 정보', '경제활동상태 판별', '임금근로자/비임금근로자/미취업자 특성', '취업을 위한 노력과 지원', '직업적 능력', '취업 태도 및 환경', '일상생활과 삶의 질', '가구 일반사항' 등 총 10개의 내용으로 이루어져 있다.

1차조사 이후 2차조사부터는 지난조사 응답 내용을 재확인하고 이후의 변화여부를 파악하는 문항이 추가되었고, 'B. 경제활동상태 판별'에서 판별된 패널의 경제활동 상태에 따라 세부 항목을 응답하도록 구성하였으며, C파트에서 응답한 일자리 외에 부업이 있거나, 현재 지속하지 않더라도 지난조사 이후 직업 활동 경험이 있는 경우 'JOB. 직업력' 파트를 부가조사의 형태로 구성하였다. 또한, 6차조사에는 코로나 19바이러스감염증-19의 영향력 파악을 위한 부가조사 파트를 신설하였으며 7차 및 8차조사에서도 운영하였다.

종단분석을 용이하게 하기 위해 대부분의 문항은 2차조사와 동일하나, 일부 문항은 3~8차조사에 새로 추가되었다.

〈표 1-3〉 주요 조사내용

영역	조사내용
A. 패널 기본 정보	· 인적 특성 및 장애 특성
B. 경제활동상태 판별	· 경제활동상태 판별, 취업자 기본 특성, 미취업자 기본 특성 · 최근 1년 월별 경제활동상태
C-a. 임금근로자	· 일자리 기본 정보, 취업 과정, 일자리 특성, 근로조건 및 실태, 업무특성, 근로환경, 차별경험, 직장생활 및 이직, 일자리 만족 등
C-b. 비임금근로자	· 일자리 기본 정보, 창업(취업) 과정, 일자리 특성, 근로실태, 업무특성, 근로환경, 사업체 운영 및 애로사항, 일자리 만족 등
C-c. 미취업자	· 향후 취업 희망 및 희망하는 일자리 정보, 구직활동 경험 등
JOB. 직업력	· 일자리 특성, 일자리 종류 등 일자리 주요 현황, 업무 및 일자리 만족도 등
D. 취업을 위한 노력과 지원	· 취업 관련 활동 경험 및 성과, 취업 관련 필요한 지원사항, 일자리에서 근무 시 필요한 지원사항 등
E. 직업적 능력	· 근로능력, 경력, 기술(자격증), 학력, 인적 네트워크, 대인관계 유지 등
F. 취업 태도 및 환경	· 취업에 대한 생각, 가족의 지지, 심리 환경, 정책 환경 등
G. 일상생활과 삶의 질	· 전반적 건강상태, 생활습관, 일상생활 도움, 일과/여가/사회참여, 차별경험 및 만족 등
H. 가구 일반사항	· 가구 기본 정보, 소득/소비, 경제생활, 거주환경, 노후준비 등
I. 코로나	· 코로나19 경험, 코로나19 전후 변화 사항, 코로나19 인식 등

## 마. 조사방법

조사방법은 태블릿 PC를 이용한 대인면접조사(Tablet PC-Assisted Personal Interviewing, 이하 TAPI)를 기본으로 진행했다. 사전에 설계된 논리설계에 따라 TAPI 프로그램을 작성, 잘못 응답된 내용을 조사원이 현장에서 바로 확인할 수 있도록 하고, 복잡한 구조의 설문 문항의 경우 자동 분기 등을 설정하여 조사원에 의한 실수 및 오차를 최소화할 수 있도록 했다.

조사원은 패널과 사전 방문일정을 조율한 후 약속된 시간에 방문하여 타계식 면접조사를 실시하고, 조사 내용은 현장에서 태블릿 PC에 입력하는 방식으로 조사가 진행되었다.

단, 2020년부터는 코로나19바이러스감염증-19로 인해 대면 접촉을 거절하는 경우 온라인 및 전화조사를 병행하였고 전화조사는 패널이 응답하는 내용을 조사원이 직접 태블릿에 입력하며 전화가 길어져 패널이 힘들어하는 경우 조사 중단 후 이후 일정에 맞춰 다시 진행하였다. 온라인 조사는 패널에게 조사에 참여할 수 있는 개별 URL을 제공하고 조사를 진행하는 동안 조사원은 패널이 문의 시 즉각 응대하도록 대기하는 형식으로 진행하였다.

## 바. 조사기간 및 조사주기

본 조사 기준은 5월 15일이 포함된 주간으로 5월부터 8월 말까지 약 3개월 동안 진행된다. 이는 고용노동부, 한국장애인고용공단의 장애인경제활동실태조사, 통계청의 경제활동인구조사와 동일한 기준이다. 본 조사의 주기는 1년으로 매년 동일 장애인을 추적 조사한다.

## 사. 결과공표

본 조사의 공표 범위는 전국을 포괄하며, 매년 기초분석보고서 등의 간행물 또는 전자통계표 등의 형태로 한국장애인고용공단 홈페이지, 한국장애인고용공단 고용개발원 홈페이지, 통계청 국가통계포털 등을 통해 공개된다.

#### 4. 패널현황

장애인고용패널 2차웨이브 구축패널 4,577명 중 8차조사에 응답한 패널은 총 3,736명으로, 패널 유지율은 원표본 대비 81.6%, 유효표본 대비 84.9%이다. 1~8차조사에 모두 참여한 연속참여 패널은 3,396명이며, 원표본 대비 74.2% 수준이다.

〈표 1-4〉 조사연도별 패널유지율 현황 (단위: 명, %)

조사차수	원표본	유효표본	조사성공	원표본 유지율	유효표본 유지율
1차조사(2016년)	4,577	4,577	4,577	100.0	100.0
2차조사(2017년)	4,577	4,560	4,214	92.1	92.4
3차조사(2018년)	4,577	4,531	4,104	89.7	90.6
4차조사(2019년)	4,577	4,504	3,995	87.3	88.7
5차조사(2020년)	4,577	4,480	3,907	85.4	87.2
6차조사(2021년)	4,577	4,447	3,848	84.1	86.5
7차조사(2022년)	4,577	4,431	3,763	82.2	84.9
<b>8차조사(2023년)</b>	4,577	4,400	3,736	81.6	84.9

\* 유효표본이란 원표본 중 사망, 장애등록 취소로 조사대상에서 영구 탈락되는 패널을 제외한 표본을 의미함.  
단순 접촉불가 또는 응답거절 등은 제외되지 않음

〈표 1-5〉 연속참여패널 현황 (단위: 명, %)

구분	2차조사	3차조사	4차조사	5차조사	6차조사	7차조사	8차조사	사례수	원패널 대비 비율
연속참여패널	○	○	○	○	○	○	○	3,396	74.2
	○	○	○	○	○	X	○	19	0.4
	○	○	○	○	X	○	○	13	0.3
	○	○	○	X	○	○	○	35	0.8
	○	○	X	○	○	○	○	33	0.7
	○	X	○	○	○	○	○	38	0.8
	X	○	○	○	○	○	○	65	1.4
	X	X	○	○	○	○	○	18	0.4
	X	○	X	○	○	○	○	0	0.0
	X	○	○	X	○	○	○	1	0.0
	X	○	○	○	X	○	○	1	0.0
	X	○	○	○	○	X	○	1	0.0
	○	X	X	○	○	○	○	11	0.2
	○	X	○	X	○	○	○	1	0.0
	○	X	○	○	X	○	○	2	0.0
	○	X	○	○	○	X	○	1	0.0
	○	○	X	X	○	○	○	4	0.1
	○	○	X	○	X	○	○	1	0.0
	○	○	X	○	○	X	○	0	0.0
	○	○	○	X	X	○	○	7	0.2
	○	○	○	X	○	X	○	3	0.1
	○	○	○	○	X	X	○	10	0.2
	X	X	X	○	○	○	○	7	0.2
	X	X	○	X	○	○	○	1	0.0
	X	X	○	○	○	○	○	0	0.0
	X	X	○	○	○	X	○	2	0.0
	X	○	X	X	○	○	○	0	0.0
	X	○	X	○	X	○	○	0	0.0
	X	○	X	○	○	X	○	0	0.0
	X	○	○	X	X	○	○	1	0.0
	X	○	○	X	○	X	○	0	0.0
	X	○	○	X	○	X	○	0	0.0
	X	○	○	○	X	○	○	0	0.0
	X	○	○	○	X	X	○	0	0.0
	X	○	○	X	X	X	○	0	0.0
	X	○	X	○	X	○	○	10	0.2
	X	X	X	X	○	○	○	2	0.0
	X	X	X	○	X	○	○	1	0.0
	X	X	X	○	○	X	○	1	0.0
	X	X	○	X	X	○	○	0	0.0
	X	X	○	○	X	X	○	0	0.0
	X	X	○	X	X	○	○	1	0.0
	X	○	X	X	○	X	○	0	0.0
	X	○	X	○	X	X	○	0	0.0
	X	○	○	X	X	X	○	1	0.0
	○	X	X	X	○	X	○	1	0.0
	○	X	X	○	X	X	○	0	0.0
	○	X	○	X	X	X	○	2	0.0
	○	○	X	X	X	X	○	10	0.2
	X	X	X	X	X	○	○	1	0.0
	X	X	X	X	○	X	○	3	0.1
	X	X	X	○	X	X	○	0	0.0
	X	X	○	X	X	X	○	1	0.0
	X	○	X	X	X	X	○	3	0.1
	○	X	X	X	X	X	○	4	0.1
	X	X	X	X	X	X	○	15	0.3

〈표 1-6〉 연속참여패널 현황 (단위: 명, %) (계속)

구분	2차조사	3차조사	4차조사	5차조사	6차조사	7차조사	8차조사	사례수	원패널 대비 비율
	○	○	○	○	○	○	X	85	1.9
	○	○	○	○	○	X	X	70	1.5
	○	○	○	○	X	○	X	4	0.1
	○	○	○	X	○	○	X	4	0.1
	○	○	X	○	○	○	X	5	0.1
	○	X	○	○	○	○	X	1	0.0
	X	○	○	○	○	○	X	6	0.1
	X	X	○	○	○	○	X	1	0.0
	X	○	X	○	○	○	X	0	0.0
	X	○	○	X	○	○	X	0	0.0
	X	○	○	○	X	○	X	0	0.0
	X	○	○	○	○	X	X	2	0.0
	○	X	X	○	○	○	X	1	0.0
	○	X	○	X	○	○	X	0	0.0
	○	X	○	○	X	○	X	1	0.0
	○	X	○	○	○	X	X	1	0.0
	○	○	X	X	○	○	X	0	0.0
	○	○	X	○	X	○	X	0	0.0
	○	○	○	X	X	○	X	3	0.1
	○	○	○	X	X	○	X	1	0.0
	○	○	○	X	○	X	X	6	0.1
	○	○	○	○	X	X	X	83	1.8
	X	X	X	○	○	○	X	1	0.0
	X	X	○	X	○	○	X	0	0.0
	X	X	○	○	X	○	X	0	0.0
	X	X	○	○	○	X	X	2	0.0
	X	○	X	X	○	○	X	0	0.0
	X	○	X	○	X	○	X	0	0.0
	X	○	X	○	○	X	X	1	0.0
	X	○	○	X	X	○	X	0	0.0
	X	○	○	X	○	X	X	0	0.0
	X	○	○	○	X	X	X	5	0.1
	○	X	X	X	○	○	X	0	0.0
	○	X	X	○	X	○	X	0	0.0
	○	X	X	○	○	X	X	1	0.0
	○	X	○	X	X	○	X	0	0.0
	○	X	○	X	○	X	X	0	0.0
	○	X	○	X	○	X	X	0	0.0
	○	X	○	X	X	X	X	3	0.1
	○	○	X	X	X	○	X	1	0.0
	○	○	X	X	○	X	X	1	0.0
	○	○	X	○	X	X	X	1	0.0
	○	○	○	X	X	X	X	72	1.6
	X	X	X	X	○	○	X	2	0.0
	X	X	X	○	X	○	X	0	0.0
	X	X	X	○	○	X	X	0	0.0
	X	X	○	X	X	○	X	0	0.0
	X	X	○	X	○	X	X	0	0.0
	X	X	○	X	X	X	X	1	0.0
	X	○	X	X	X	○	X	1	0.0
	X	○	X	X	○	X	X	0	0.0
	X	○	X	○	X	X	X	0	0.0
	X	○	○	X	X	X	X	8	0.2
	○	X	X	X	X	○	X	0	0.0
	○	X	X	X	○	X	X	1	0.0
	○	X	X	○	X	X	X	3	0.1
	○	X	○	X	X	X	X	6	0.1
	○	○	X	X	X	X	X	114	2.5
	X	X	X	X	X	○	X	1	0.0
	X	X	X	X	○	X	X	6	0.1
	X	X	X	○	X	X	X	4	0.1
	X	X	○	X	X	X	X	3	0.1
	X	○	X	X	X	X	X	14	0.3
	○	X	X	X	X	X	X	136	3.0
	X	X	X	X	X	X	X	179	3.9
	전체							4,577	100.0

8차조사 기준 유지패널의 특성별 현황을 살펴보면 성별로는 여성(82.8%), 연령별로는 50~64세(83.9%)의 패널 유지율이 상대적으로 높았다. 권역별로는 전라권(87.7%)이 가장 높은 반면, 수도권(77.6%)의 유지율은 낮았으며, 경찰상태별로는 취업자(84.0%)의 유지율이 실업자(구직자)(79.7%)에 비해 높았다. 장애정도별로는 경증장애인(82.0%)이 중증장애인(81.0%)보다 높았고 장애유형별로는 지체장애(82.9%)의 유지율이 가장 높은 반면, 신체내부장애(72.9%)는 상대적으로 유지율이 낮았다.

〈표 1-7〉 조사연도별 패널유지율 세부현황

구분		1차조사 (2016년)		2차조사 (2017년)		3차조사 (2018년)		4차조사 (2019년)		5차조사 (2020년)		6차조사 (2021년)		7차조사 (2022년)		8차조사 (2023년)	
		n	%	n	%	n	%	n	%	n	%	n	%	n	%	n	%
성별	남성	3,010	100.0	2,746	91.2	2,673	88.8	2,609	86.7	2,549	84.7	2,516	83.6	2,454	81.5	<b>2,438</b>	<b>81.0</b>
	여성	1,567	100.0	1,468	93.7	1,431	91.3	1,386	88.4	1,358	86.7	1,332	85.0	1,309	83.6	<b>1,298</b>	<b>82.8</b>
연령별	15~29세	805	100.0	733	91.1	707	87.8	681	84.6	665	82.6	649	80.6	631	78.4	<b>655</b>	<b>78.9</b>
	30~39세	1,105	100.0	1,024	92.7	1,000	90.5	963	87.1	955	84.6	939	85.0	918	83.1	<b>913</b>	<b>82.6</b>
	40~49세	1,329	100.0	1,199	90.2	1,161	87.4	1,139	85.7	1,117	84.0	1,094	82.3	1,070	80.5	<b>1,066</b>	<b>80.2</b>
	50~64세	1,338	100.0	1,258	94.0	1,236	92.4	1,212	90.6	1,190	88.9	1,166	87.1	1,144	85.5	<b>1,122</b>	<b>83.9</b>
권역별	수도권	2,187	100.0	1,961	89.7	1,892	86.5	1,835	83.9	1,790	81.8	2,817	80.1	1,706	78.0	<b>1,697</b>	<b>77.6</b>
	경남권	700	100.0	677	96.7	635	90.7	654	93.4	645	92.1	700	91.0	628	89.7	<b>612</b>	<b>87.4</b>
	경북권	380	100.0	343	90.3	341	89.7	328	86.3	319	83.9	380	84.5	312	82.1	<b>314</b>	<b>82.6</b>
	전라권	512	100.0	484	94.5	471	92.0	467	91.2	460	89.8	512	88.9	452	88.3	<b>449</b>	<b>87.7</b>
	충청권	528	100.0	497	94.1	487	92.2	473	89.6	464	87.9	528	87.1	449	85.0	<b>451</b>	<b>85.4</b>
	강원/제주	270	100.0	252	93.3	247	91.5	238	88.1	225	83.3	270	83.0	216	80.0	<b>213</b>	<b>78.9</b>
경찰 상태별	취업자	2,270	100.0	2,080	91.6	2,036	89.7	1,984	87.4	1,950	85.9	2,270	85.2	1,899	83.7	<b>1,907</b>	<b>84.0</b>
	구직자	167	100.0	150	89.8	143	85.6	163	84.9	158	82.3	192	80.2	154	80.2	<b>153</b>	<b>79.7</b>
	미취업자 (취업 희망)	1,118	100.0	1,013	90.6	987	88.3	945	86.1	910	82.9	1,098	82.7	880	80.1	<b>871</b>	<b>79.3</b>
	미취업자 (취업 미희망)	1,022	100.0	971	95.0	938	91.8	933	88.8	889	87.4	1,017	83.8	830	81.6	<b>815</b>	<b>79.2</b>
	장애 정도별	중증	1,477	100.0	1,383	93.6	1,338	90.6	1,294	87.6	1,253	84.8	1,218	82.5	1,180	79.9	<b>1,155</b>
경증	3,100	100.0	2,831	91.3	2,766	89.2	2,701	87.1	2,654	85.6	2,630	84.8	2,583	83.3	<b>2,301</b>	<b>82.0</b>	
장애 유형별	지체장애	2,251	100.0	2,066	91.8	1,991	88.4	1,953	86.8	1,925	85.5	2,251	84.9	1,872	83.2	<b>1,867</b>	<b>82.9</b>
	지체 외 신체외부장애	332	100.0	312	94.0	307	92.5	294	88.6	283	85.2	332	84.0	274	82.5	<b>271</b>	<b>81.6</b>
	시각장애	591	100.0	541	91.5	530	89.7	510	86.3	512	86.6	591	84.4	494	83.6	<b>499</b>	<b>82.7</b>
	사각외감각장애	416	100.0	384	92.3	381	91.6	370	88.9	356	85.6	416	83.2	343	82.5	<b>338</b>	<b>81.3</b>
	정신적장애	613	100.0	571	93.1	561	91.5	541	88.3	525	85.6	613	83.2	495	80.8	<b>487</b>	<b>79.4</b>
	신체내부장애	374	100.0	340	90.9	334	89.3	327	87.4	306	81.8	374	80.7	285	76.2	<b>284</b>	<b>72.9</b>
	<b>전체</b>	<b>4,577</b>	<b>100.0</b>	<b>4,214</b>	<b>92.1</b>	<b>4,104</b>	<b>89.7</b>	<b>3,995</b>	<b>87.3</b>	<b>3,907</b>	<b>85.4</b>	<b>3,848</b>	<b>84.1</b>	<b>3,763</b>	<b>82.2</b>	<b>3,736</b>	<b>81.6</b>

\* 특성은 2016년 1차조사 기준이며, 비율은 원패널 대비 비율을 의미함

## 제2장

### 표본설계 및 가중치

#### 1. 표본설계

장애인고용패널조사의 품질을 높이며, 조사의 효율성(시간 및 비용)을 제고하기 위해 효과적인 장애인 표본 추출을 통한 패널 구축이 중요하다. 이를 위해 우리나라 만 15~64세 장애인 집단을 대표하고, 장기 패널조사 운용에 최적화된 표본을 설계하여 장애인고용패널 2차웨이브 패널을 구축하는 것이 표본설계의 목표이다.

장애인고용패널 2차웨이브의 표본설계를 위해서는 다음의 사항들을 고려해야 한다. 조사 기간이 길어질 경우 응답자들의 기억의 오차(recall bias)가 발생하게 되며 이는 조사의 신뢰도에 영향을 미치게 된다. 따라서 조사대행사의 실사능력에 맞는 표본추출 방법이 고려되어야 한다. 이와 더불어 본 패널조사의 주목적인 장애인의 경제활동상태에 대한 분석을 위하여 경제활동 장애인 및 이를 희망하는 장애인을 충분히 포함할 수 있어야 하며, 또한 장애유형별 그리고 장애등급별 모집단 대표성을 확보하기 위한 표본설계가 이루어져야 한다. 마지막으로 장애인의 경제활동 관련 자료 분석에 중요한 설명변수인 장애유형 및 장애등급 등 관심변수들의 신뢰도 높은 통계량을 생산할 수 있는 표본이 추출되어야 한다.

목표 모집단은 2016년 5월 15일 기준 대한민국에 거주하며, 15개 법정 장애유형 중 하나 이상의 장애를 지니고 있는 만 15~64세의 등록장애인이다. 표본추출을 위한 표집틀로는 2015년 12월 31일 기준 보건복지부 등록장애인 명부를 확보하여 사용하였다. 사용된 표집틀에는 등록장애인의 장애유형, 장애등급, 주소, 성별, 연령 등에 대한 정보가 포함되어 있다. 표집틀 구성 시 집락에 해당하는 읍면동을 구성하는 장애인의 장애유형 종류가 8개 미만인 읍면동은 제외하였으며, 제외된 읍면동에 속한 장애인은 12,635명이다.

모집단에 대해 분석하고, 지난 몇 해의 장애인경제활동실태조사 결과보고서를 통해 볼 때 장애인고용패널 2차웨이브 구축을 위해 표본설계 변수로 사용하고자 하는 장애유형, 장애등급 그리고 지역별 장애인 구성에 있어 지난 수년간 뚜렷한 변화는 없는 것을 파악하였다. 따라서 장애인고용패널 2차웨이브 구축을 위한 표본설계로는 지난 여러 해를 통해 표본설계의 안정성과 타당성이 검증된 기존 장애인경제활동실태조사 표본설계 방법의 틀과 유사하게 추출 읍면동 수를 조절함과 더불어 각 장애유형, 장애등급 그리고 연령별 적절한 수의 표본이 추출될 수 있는 이상추출법(two-phase sampling)을 고려하였다. 조사기간 및 비용을 고려하여 먼저 집락인 읍면동을 추출하는 1단계 집락추출법을 통하여 1상(first phase) 장애인 표본을 추출하고, 이를 장애유형, 장애등급 그리고 연령을 기준으로 층화하여 목표오차를 충족하는 적정 수준의 표본 장애인을 층화추출하였다.

덧붙여, 장애인고용패널 2차웨이브는 주목적인 장애인 경제활동상태에 대한 분석과 동태적 변화를 파악하기 위하여 주어진 예산의 제약 하에서 경제활동 장애인 및 이를 희망하는 장애인을 충분히 포함할 수 있도록 표본크기를 4,400명으로 정하였다.



## 2. 표본추출

### 가. 집락 및 층화

1상 표본추출을 위해 읍면동이 집락으로 고려되었다. 시간 및 비용 등 주어진 현실적인 제약을 감안하여 조사의 효율성을 높이기 위해 해당 읍면동의 장애유형 수가 8개 이하인 읍면동(전체 3,555개 읍면동 중 113개)은 제외하였다. 아울러 비율이 낮은 간, 뇌전증, 심장, 안면, 언어, 자폐성, 장루·요루, 호흡기장애 등이 표본에서 누락되지 않도록 희귀 장애유형에 속하는 장애인 수가 최소 2~5명 이상 거주하고 있는 51개 읍면동은 전수층으로 분류하여 1상 표본에 포함되도록 하였다. 다음은 1상 표본 추출을 위한 표집들의 지역별 표본층 읍면동 수와 전수층 읍면동 수이다.

〈표 2-1〉 읍면동 표본추출을 위한 모집단 층화

지역	읍, 면/동수	지역	읍, 면/동수
표본층	3,391	강원	187
서울	398	충북	150
부산	199	충남	204
대구	140	전북	227
인천	148	전남	290
광주	94	경북	327
대전	78	경남	302
울산	55	제주	41
세종	13	전수층	51
경기	538	전국	3,442

이상 표본추출을 위하여 각 상별 모집단 층화가 고려되었다. 1상에서 추출된 읍면동 내의 모든 장애인으로 구성된 표본으로부터 2상 표본을 추출하기 위한 층화변수는 장애유형, 장애등급, 연령을 사용하였다. 15개 법정 장애유형을 6개로 통합하였으며, 장애등급은 기존 방법과 같이 중증과 경증으로 구분하였다(장애정도). 그리고 연령은 15~29세, 30~39세, 40~49세, 50~64세 4개로 통합하였다. 따라서 2상 표본추출을 위한 층화는 장애유형(6) X 장애정도(2) X 연령(4) 기준이며, 정신적 장애(지적장애, 자폐성장애, 정신장애)는 중증만 해당되므로 총 44개의 층으로 층화하였다. 장애인경제활동실태조사 보다는 그 표본크기가 적기 때문에 동일한 층화기준을 적용할 경우 각 층별 배분되는 표본의 수가 적게 되어 비현실적인 배분의 문제가 발생할 수 있고, 또한 경제활동에 참여하거나 참여를 희망하는 장애인을 주요 조사대상으로 하는 본 조사의 특성상 이를 결정하는 주요 요인인 연령을 층화변수로 포함하였다.

### 나. 표본크기 결정 및 표본배분

주어진 예산의 제약 하에서 표본크기는 전국 약 4,400여명 내외의 장애인을 목표로 하여 각 장애유형별, 장애정도별 모집단을 대표할 수 있도록 구성하였다. 유한모집단 수정항(finite population correction

factor)을 무시하고 계산되는 95% 신뢰수준 하에서의 각 장애유형별 혹은 장애정도별 추정량의 최대 한계허용오차는 대략적으로 1.8% ~ 5.8%로 나타나도록 표본배분을 실시하였다.

패널마모가 발생한 최근(2015년)의 장애인고용패널조사 결과 대신 장애인경제활동실태조사 결과를 이용하여 산출한 실업률 기준 오차한계에 대한 정보를 보면, 각 장애유형별 실업률 기준 오차한계는 대략적으로 1% ~ 3% 수준일 것으로 예측된다. 본 패널조사에서는 각 장애유형별 혹은 장애정도별 통계산출이 주된 목적이 아닌 점을 감안하여 가능한 수준에서 모든 유형 및 정도의 장애인이 포함되도록 구성하였다. 참고로 실업률의 계산식은 다음과 같다.

$$\text{실업률} = (\text{실업자 수} / \text{경제활동인구}) \times 100$$

〈표 2-2〉 장애유형, 장애정도별 표본크기 및 실업률의 오차한계

항목	장애정도			실업률 기준 오차한계
	중증	경증	합계	
지체장애	225	1,982	2,207	0.01
지체 외 신체외부장애	125	169	294	0.04
시각장애	117	448	565	0.03
시각 외 감각장애	106	344	450	0.02
정신적장애	561	0	561	0.03
신체내부장애	153	170	323	0.03
총합계	1,287	3,113	4,400	-

본 조사에서는 이상추출에 의한 표본추출법을 이용하기 때문에 1상 표본으로 추출된 읍면동 내의 모든 장애인으로 2상 표본추출을 위한 장애인 표집틀을 구성하고 언급된 층화변수를 이용하여 구성된 각 층에 장애인 표본배분이 이루어져야 한다. 최종 분석단위가 장애인인 것을 감안하여, 장애인의 경제활동정보를 활용하고, 장애유형, 장애정도 그리고 연령 변수를 기준으로 한 층별 장애인 표본배분이 고려되어야 한다.

통합된 장애유형(6), 장애정도(2) 그리고 연령(4) 변수를 기준으로 총 44개의 층을 구성하였으며, 각 층별 표본배분을 위해서는 비례배분법이 사용되었다. 단, 이때 경제활동 참여 비율(경제활동인구 비율) 및 참여 희망 비율(비경제활동인구 중 취업 희망 비율)을 활용하여 전체 표본 장애인 중에서 일정 수준 이상의 경제활동 참여 및 참여 희망 장애인이 포함될 수 있도록 하였다.

이를 위해서는 층 변수인 장애유형, 장애정도 그리고 연령으로 정의된 각 부모집단별 경제활동 참여 비율 및 참여 희망 비율을 고려하여 경제활동에 참여하거나 참여를 희망하는 장애인과 희망하지 않는 장애인의 표본 비율이 각 층에서 평균적으로 3:1이 되도록 배분을 시행하였다. 다음은 층별 표본배분 결과이다.

〈표 2-3〉 층별 표본배분 결과

장애유형·정도	연령	15~29세	30~39세	40~49세	50~64세	총합계	(구성비)
		지체장애	중증	15	37	65	153
	경증	45	473	751	1,171	2,440	(55.5)
지체 외 신체외부장애	중증	15	15	20	40	90	(2.0)
	경증	10	15	20	59	104	(2.4)
시각장애	중증	15	15	15	27	72	(1.6)
	경증	21	82	122	194	419	(9.5)
시각 외 감각장애	중증	15	15	21	47	98	(2.2)
	경증	15	36	54	146	251	(5.7)
정신적장애	중증	150	107	101	89	447	(10.2)
	경증	-	-	-	-	-	-
신체내부장애	중증	10	17	20	48	95	(2.2)
	경증	10	17	30	57	114	(2.6)
총합계 (구성비)		321 (7.3)	829 (18.8)	1,219 (27.7)	2,031 (46.2)	4,400 (100.0)	

#### 다. 읍면동 및 장애인 표본추출

읍면동 추출 및 배분된 크기의 장애인 표본을 추출하기 위하여 각각 확률비례계통추출법과 층화계통추출법을 사용하였다. 실제 표본추출 과정에서는 1상에서 집락인 읍면동 추출이 장애인 추출보다 선행되어 이루어진다. 1상 표본인 읍면동 추출에 앞서 내재적 층화(implicit stratification)의 효과를 얻기 위하여 지역 변수를 이용하여 읍면동을 정렬하였다. 동시에 장애인 수와 장애유형 수를 크기 변수로 고려하여 모집단에 대한 대표성을 갖춘 집락이 뽑힐 수 있도록 하였다. 단, 이때 2016년 장애인경제활동실태조사를 위하여 기 추출된 표본에 대해서는 크기 변수에 대하여 1/2배를 하여 본 패널조사 표본과 겹치지 않도록 하였다.

장애인고용패널 2차웨이브 구축을 위한 표본설계에서는 장애인을 추출하는 대신 읍면동을 추출하는 집락 추출법이 1상 표본추출을 위하여 사용되었다. 아울러 앞서 설명한 바와 같이 희귀 장애유형이 표본에서 누락되지 않도록 하게 하기 위해 희귀 장애유형의 장애인을 일정수준 이상 포함한 51개 읍면동은 전수층으로 분류하여 1상 표본에 포함되도록 하였다.

1상 표본에 대한 17개 지역별 표본 읍면동 크기와 지역별 전수층 읍면동 수를 살펴보면, 전수층과 표본층을 포함하여 1상 표본인 읍면동 표본 수는 모두 492개이다.

1상 표본 추출을 위한 추출단위로 장애인이 아닌 읍면동이 사용된 이유는 언급된 대로 장애유형, 장애등급 그리고 연령만을 고려하여 장애인을 직접 추출할 경우 지나치게 많은 읍면동이 추출되어 조사비용이 증가할 뿐 아니라 조사기간이 장기화됨에 따라 비표본오차가 커질 수 있기 때문이다. 참고로 2상 추출법이 아닌 층화계통추출을 통해 직접 추출한 경우 표본에 포함되는 읍면동 수는 약 2,300개로 현재 보다 5배 가까이 늘어난다.

〈표 2-4〉 읍면동 표본추출 결과

지역	읍면동 수	전수층		지역	읍면동 수	표본층	
		전수층	표본층			전수층	표본층
서울	84	23	61	강원	22	2	20
부산	33	3	30	충북	18	0	18
대구	22	1	21	충남	24	0	24
인천	23	0	23	전북	25	1	24
광주	14	1	13	전남	27	1	26
대전	13	1	12	경북	34	0	34
울산	10	1	9	경남	36	1	35
세종	1	0	1	제주	7	2	5
경기	99	14	85	전국	492	51	441

2상에서는 1상에서 추출된 읍면동에 거주하는 모든 장애인의 리스트를 표집틀로 정의하여 이로부터 표본을 추출하였다. 1상에서 추출된 읍면동에 거주하는 모든 장애인의 장애유형 및 장애등급별 분포는 다음과 같다.

〈표 2-5〉 장애유형 및 장애등급별 표본분포 현황(2상 추출 대상)

장애유형	장애등급							계
	1급	2급	3급	4급	5급	6급		
간	38	60	100	33	1,359	-	1,590	
뇌진증	16	68	241	843	189	-	1,357	
뇌병변	4,820	3,923	4,604	2,499	2,522	2,907	21,275	
시각	3,213	597	1,057	1,194	2,011	18,761	26,833	
신장	492	6,512	9	102	3,141	-	10,256	
심장	20	77	396	6	102	-	601	
안면	18	87	164	175	13	-	457	
언어	24	242	732	1,145	-	-	2,143	
자폐성장애	1,135	760	479	-	-	-	2,374	
장루.요루	2	6	87	556	303	-	954	
정신	355	4,254	11,265	-	-	-	15,874	
지적장애	5,562	9,046	11,540	-	-	-	26,148	
지체(변형)	19	22	49	77	226	1,415	1,808	
지체(상지관절)	40	170	1,066	2,119	1,434	4,813	9,642	
지체(상지기능)	511	1,512	4,194	2,861	3,402	4,889	17,369	
지체(상지절단)	105	829	2,329	2,741	2,507	5,907	14,418	
지체(전환대상)	15	18	44	38	41	52	208	
지체(척추)	212	1,045	1,638	1,601	15,813	18,585	38,894	
지체(하지관절)	97	288	1,535	5,575	9,764	17,014	34,273	
지체(하지기능)	4,198	4,246	7,507	6,192	5,193	1,096	28,432	
지체(하지절단)	100	217	1,000	857	186	378	2,738	
청각(전환대상)	2	8	3	4	4	3	24	
청각(청력)	919	4,719	2,586	2,782	3,232	3,159	17,397	
청각(평형기능)	0	5	12	21	29	5	72	
호흡기	167	309	558	-	21	-	1,055	
계	22,080	39,020	53,195	31,421	51,492	78,984	276,192	

## 라. 표본대체

조사과정에서 무응답 및 거절, 기타 여러 가지 이유로 인하여 부득이하게 표본대체가 발생한다. 표본 대체로 인한 비표본오차를 줄이기 위해서 각각의 원표본에 대응되는 대체표본을 5명씩 추가로 추출하였다.

# 3. 1차조사 가중치 산출 및 추정

## 가. 모집단 분포 산출

가중치 보정 시 취업 희망자 비중이 지나치게 높게 나타나는 문제를 해결하기 위하여 경제활동 상태 변수를 고려하고자 한다. 경제활동상태에 대한 모집단 정보는 2016년 장애인경제활동실태조사의 경제활동 상태에 대한 추정 결과를 사용하였다. 이 추정결과의 모집단 수와 2015년 12월 기준 장애인 등록현황의 모집단 수가 일치하지 않기 때문에 추정 결과를 직접 사용할 수는 없다. 이를 해결하기 위해 2016년 「장애인경제활동실태조사」의 경제활동상태별 비율 추정 결과를 2015년 12월 기준 장애인 등록현황의 모집단 수에 곱하여 경제활동상태에 대한 모집단 분포를 산출하였다.

## 나. 레이킹 비 가중

일반적으로 표본 가중치는 무응답, 표본설계 시 사용한 자료의 시의성 문제, 현실적인 제약(조사여건), 그리고 표집틀과 모집단의 불일치 등으로 인하여 적절한 보정을 거치게 되고 이를 통해 최종가중치가 산출된다.

본 조사에서는 무응답 처리를 위하여 효과적인 표본의 대체방안이 적용되었다. 그리고 표본추출을 위한 표집틀 구성에서는 이용 가능한 최신자료인 2015년 12월 31일 기준 보건복지부 명부를 사용하여 시의성의 문제는 크게 발생하지 않는 것으로 판단하였다. 그러나 모집단에는 포함되지만 표집틀 구성에서는 시간 및 비용의 현실적인 제약으로 2015년 표본설계 시 고려한 것처럼 장애유형의 종류가 8개 이하인 113개의 읍면동을 제외하여 표집틀과 모집단의 불일치가 발생하였다. 따라서 본 조사에서는 레이킹 비 기법을 이용하여 표본 가중치를 보정하고자 한다. 참고로, 보건복지부의 최신 등록장애인 명부를 사용할 수 없는 경우(시의성 문제) 표본 가중치는 반드시 보정이 필요한데, 이러한 경우를 고려하면 일관성 있는 최종가중치 산출과정을 위해서 본 조사에서도 레이킹 비 기법을 이용하여 표본 가중치를 보정하는 것은 적절하리라 판단된다.

레이킹 비 기법을 적용하기 위해서 사용할 변수는 성, 연령, 지역, 장애유형, 장애정도, 경제활동상태 변수 등을 벤치마킹 변수로 하여 레이킹 비 방법을 적용하여 보정하였다. 경제활동상태 변수에 대한 모집단 정보로는 동일 시점에 이루어진 장애인경제활동실태조사 결과를 이용하였다.

따라서 1차조사의 가중치는 성, 연령, 지역, 장애유형, 장애정도, 경제활동상태변수 등을 벤치마킹 변수로 하여 레이킹 비 방법을 적용하여 보정하였으며, 그 결과는 다음과 같다.

〈표 2-6〉 가중치에 대한 기술통계량

통계량			
관측값 수	4,577	관측값 합	1,387,802
평균	303.21	표준편차	443.05
중위수	159.87	분산	196,291.70
사분위 범위	295.05	범위	9231.10
최솟값	3.29	최댓값	9234.38

벤치마킹 변수에 대한 모집단 분포와 가중치를 사용하여 얻어진 표본 자료의 분포는 다음과 같다.

〈표 2-7〉 주요 변수별 벤치마킹 모집단 분포 - 성/연령별

모집단 분포				표본 분포			
성별	빈도	연령	빈도	성별	빈도	연령	빈도
남자	927,388	15~29세	130,641	남자	927,388	15~29세	130,654
여자	460,414	30~39세	156,382	여자	460,414	30~39세	156,385
계	1,387,802	40~49세	310,386	계	1,387,802	40~49세	310,387
		50~64세	790,393			50~64세	790,377
		계	1,387,802			계	1,387,803

〈표 2-8〉 주요 변수별 벤치마킹 모집단 분포 - 시도별

모집단 분포		표본 분포	
시도	빈도	시도	빈도
서울	218,512	서울	218,512
부산	95,587	부산	95,586
대구	68,980	대구	68,980
인천	81,080	인천	81,079
광주	39,191	광주	39,192
대전	41,414	대전	41,414
울산	30,378	울산	30,378
경기	303,186	경기	303,184
강원	51,219	강원	51,220
충북	51,620	충북	51,621
충남	70,627	충남	70,627
전북	64,349	전북	64,349
전남	66,159	전남	66,159
경북	87,057	경북	87,058
경남	100,460	경남	100,461
제주	17,983	제주	17,983
계	1,387,802	계	1,387,803

〈표 2-9〉 주요 변수별 벤치마킹 모집단 분포 - 경제활동상태별

모집단 분포		표본 분포	
경제활동상태	빈도	경제활동상태	빈도
취업자	696,375	취업자	696,376
미취업자	49,760	미취업자	49,759
취업희망자	144,286	취업희망자	144,286
취업미희망자	497,381	취업미희망자	497,381
계	1,387,802	계	1,387,802

〈표 2-10〉 주요 변수별 벤치마킹 모집단 분포 - 장애유형/장애등급별

모집단 분포				표본 분포			
장애유형	빈도	장애등급	빈도	장애유형	빈도	장애등급	빈도
지체장애	713,280	중증	528,223	지체장애	713,309	중증	528,219
지체 외	108,062	경증	859,579	지체 외	108,057	경증	859,583
신체외부장애		계	1,387,802	신체외부장애		계	1,387,802
시각장애	131,608			시각장애	131,612		
시각 외	96,343			시각 외	96,345		
감각장애				감각장애			
정신적 장애	261,550			정신적 장애	261,522		
신체내부장애	76,959			신체내부장애	76,957		
계	1,387,802			계	1,387,802		

#### 다. 주요 변수의 추정치 및 상대표준오차

본 보고서에서는 장애인의 경제활동상태(취업자, 실업자, 비경제활동인구 중 취업희망자와 미희망자)와 종사상지위(임금근로자, 비임금근로자, 무급가족종사자, 실업자, 비경제활동인구)를 고려하였으며, 경제활동상태 및 종사상 지위에 따른 추정치와 그에 따른 상대표준오차를 제시하였다. 관련 추정치를 산출하거나 해석할 때 참고할 필요가 있다.

〈표 2-11〉 주요 비율에 대한 추정치와 상대표준오차

항목	설계 가중치		최종가중치		
	추정량	S.E	추정량	S.E	
경활 상태	취업자	601,234	18,250.5	696,375	4,157.7
	실업자	51,807	6,973.6	49,759	798.8
	비경제_취업희망	337,116	15,486.0	144,286	1,223.0
	비경제_취업미희망	369,249	16,366.4	497,381	3,050.7
종사상 지위	임금근로자	417,633	15,045.1	492,925	11,829.1
	자영업자	153,680	10,937.7	170,105	10,697.5
	무급가족종사자	29,921	4,755.6	33,345	5,393.2
	실업자	51,807	6,973.6	49,759	798.8
	비경제활동인구	706,366	21,290.2	641,667	3,688.4

### 라. 최종가중치 산출(극단값 처리)

일반적으로 몇 개의 과도하게 큰 가중치를 갖는 조사값은 추정량의 분산을 크게 만들어 추정의 정확성을 낮추게 하는 경우가 발생할 수 있다.

가중치 중에서 5,000을 넘는 경우(5079, 7088, 9234)에 대해서는 이를 극단값으로 조정하였다. 이들 극단가중치는 다음 최대값인 4343으로 조정해 주고, 성, 연령, 장애유형, 장애정도 등의 기본적인 특성 변수에 대해서 레이킹 비 보정을 실시하여 1차조사 최종가중치를 얻었다. 최종가중치의 분포와, 최종가중치를 사용한 경제활동상태 변수의 표본 분포는 아래와 같다. 극단값 보정을 위해 가중치를 조정하였으나 위의 경제활동상태의 표본 분포에서 볼 수 있듯이 주요변수에 대한 주변분포는 모집단 분포와 큰 차이가 없음을 알 수 있다.

〈표 2-12〉 최종가중치에 대한 기술통계량

통계량			
관측값 수	4,577	관측값 합	1,387,802
평균	303.21	표준편차	422.522
중위수	160.07	분산	178,524.8
사분위 범위	295.00	범위	4,484.28
최솟값	3.24	최댓값	4,487.52

〈표 2-13〉 경제활동상태 변수 모집단 분포와 표본 분포

모집단 분포		표본 분포	
지역	빈도	지역	빈도
취업자	696,375	취업자	695,780
미취업자	49,760	미취업자	49,386
취업희망자	144,286	취업희망자	145,148
취업미희망자	497,381	취업미희망자	497,488
계	1,387,802	계	1,387,803



## 4. 2차조사 가중치 산출 및 추정

1차년도(2016) 조사대상자들이 여러 이유로 응답을 거부하거나 패널에서 탈락하는 경우가 발생할 수 있다. 이와 같은 무응답이나 패널 탈락은 추후 조사자료를 이용한 모수 추정에 있어 편향이 생기는 원인이 된다. 따라서 무응답자나 패널 탈락자에 대한 가중치를 보정해 주어야 한다.

2차조사를 위한 가중치를 위하여 중단 가중치와 횡단 가중치를 동시에 고려해야 한다. 중단 가중치는 조사차수별로 발생할 수 있는 무응답자에 대한 보정 가중치로써 차수 간의 비교 분석을 위해 필요하다. 조사 차수별로 변동이 생기는 모집단의 분포를 반영하여 해당 시점의 모집단 분포와 표본의 분포가 같아지도록 가중치를 보정하여야 하는데 이때 사용되는 가중치를 횡단 가중치라고 한다. 2차조사의 가중치 산출을 위하여 먼저 무응답 보정을 통한 중단 가중치를 계산한 후 표본의 분포와 모집단의 분포가 같아지도록 보정계수를 산출한 후 이를 중단 가중치에 곱하여 횡단 가중치를 작성하도록 한다.

### 가. 중단가중치 산출

#### 1) 무응답 특성 분석

1차년도 조사대상자 4,577명 중 2차년도에도 응답한 대상자는 4,214명이며 응답률은 약 93%이다. 응답을 얻지 못한 363명의 무응답 사유는 다음과 같다.

〈표 2-14〉 무응답 사유 빈도표

방문시 부재	방문시 이사	거절	강력 거절	단기 부재	장기 부재	사망	장애등록 취소	합계
62	13	131	98	26	16	14	3	363
17.1%	3.6%	36.1%	27.0%	7.2%	4.4%	3.9%	0.8%	100.0%

다음으로 응답여부와 주요 특성 변수들 간의 연관성을 살펴봄으로써 응답여부에 중요하게 영향을 미치는 특성 변수를 찾고자 한다.

첫 번째로 성별 무응답률 현황이다. 성별 무응답률은 큰 차이는 아니지만 통계적으로 유의한 차이가 있음을 확인할 수 있다. 남성의 무응답 비율이 여성보다 상대적으로 높다.

〈표 2-15〉 성별 무응답 현황

구분	무응답		응답		계
	빈도	행%	빈도	행%	
남성	264	8.8%	2,746	91.2%	3,010
여성	99	6.3%	1,468	93.7%	1,567
계	363	7.9%	4,214	92.1%	4,577

$$\chi^2 = 8.16, df = 1, P\text{-value}=0.004$$

두 번째로 시도별 무응답률 현황이다. 무응답률의 시도 편차는 유의미하게 존재하고, 대구, 인천,

울산, 경기, 제주에서 무응답률이 높게 나타나며 부산, 광주, 대전, 경남 등의 무응답 비율이 상대적으로 낮다.

〈표 2-16〉 시도별 무응답 현황

시도	무응답		응답		계
	빈도	행%	빈도	행%	
서울	66	7.4%	823	92.6%	889
부산	12	2.9%	409	97.1%	421
대구	27	10.4%	232	89.6%	259
인천	31	13.1%	205	86.9%	236
광주	5	3.9%	123	96.1%	128
대전	6	3.3%	176	96.7%	182
울산	4	11.4%	31	88.6%	35
경기	129	12.1%	933	87.9%	1,062
강원	11	5.4%	192	94.6%	203
충북	10	5.6%	169	94.4%	179
충남	15	9.0%	152	91.0%	167
전북	13	7.4%	163	92.6%	176
전남	10	4.8%	198	95.2%	208
경북	10	8.3%	111	91.7%	121
경남	7	2.9%	237	97.1%	244
제주	7	10.4%	60	89.6%	67
계	363	7.9%	4,214	92.1%	4,577

$\chi^2 = 76.15, P\text{-value}=0.000$

세 번째로 연령대별 무응답률 현황이다. 응답여부와 연령의 분포는 유의한 관계가 있음을 확인할 수 있다. 연령대에 따른 무응답 비율을 살펴보면 50세 미만의 연령층에서 무응답률이 상대적으로 높게 나타나고 있음을 알 수 있다. 특히 40대의 무응답률이 가장 높으며 50세 이상의 무응답률이 가장 낮다.

〈표 2-17〉 연령대별 무응답 현황

구분	무응답		응답		계
	빈도	행%	빈도	행%	
15-29세	72	8.9%	733	91.1%	805
30-39세	81	7.3%	1,024	92.7%	1,105
40-49세	130	9.8%	1,199	90.2%	1,329
50-64세	80	6.0%	1,258	94.0%	1,338
계	363	7.9%	4,214	92.1%	4,577

$\chi^2 = 14.89, df = 3, P\text{-value}=0.002$

네 번째로 경제활동상태별 무응답률 현황이다. 응답여부와 경제활동상태의 관련성은 유의한 관계가 있음을 확인할 수 있다. 경제활동상태에 따른 무응답률을 살펴보면 실업자의 무응답률이 가장 높고 취업미희망자의 무응답률이 가장 낮음을 알 수 있다.

〈표 2-18〉 경제활동상태별 무응답 현황

구분	무응답		응답		계
	빈도	행%	빈도	행%	
취업자	190	8.4%	2,080	91.6%	2,270
실업자	17	10.2%	150	89.8%	167
취업희망자	105	9.4%	1,013	90.6%	1,118
취업미희망자	51	5.0%	971	95.0%	1,022
계	363	7.9%	4,214	92.1%	4,577

$$\chi^2 = 17.13, df = 3, P\text{-value}=0.001$$

마지막으로 장애유형 및 장애정도별 무응답률을 살펴보면 지체 외 신체외부장애의 무응답률이 가장 낮고, 신체내부장애의 무응답률이 가장 높으나 장애유형에 따른 무응답률 차이는 통계적으로 유의하지 않다. 장애정도별로는 경증의 무응답률이 중증의 무응답률보다 다소 높으나 장애정도에 따른 무응답률 차이는 유의하지 않다.

〈표 2-19〉 장애유형별 무응답 현황

구분	무응답		응답		계
	빈도	행%	빈도	행%	
지체장애	185	8.2%	2,066	91.8%	2,251
지체 외 신체 외부장애	20	6.0%	312	94.0%	332
시각장애	50	8.5%	541	91.5%	591
시각외감각장애	31	7.5%	385	92.5%	416
정신적장애	44	7.2%	569	92.8%	613
신체내부장애	33	8.8%	341	91.2%	374
계	363	7.9%	4,214	92.1%	4,577

$$\chi^2 = 3.15, df = 5, P\text{-value}=0.677$$

〈표 2-20〉 장애정도별 무응답 현황

구분	무응답		응답		계
	빈도	행%	빈도	행%	
중증	104	7.0%	1,373	93.0%	1,477
경증	259	8.4%	2,841	91.6%	3,100
계	363	7.9%	4,214	92.1%	4,577

$\chi^2 = 2.19, df = 1, P\text{-value}=0.140$

## 2) 최종 종단가중치 산출

무응답 보정 가중치를 산출하기 위해서는 조사대상자들의 다양한 속성(주로 인구사회학적 특성)을 반영한 응답확률 모형이 필요하다. 응답확률 모형으로 주로 로지스틱 모형을 사용한다. 추정된 모형을 통해 조사대상자 각각의 응답확률을 계산하고 그 역수를 무응답 보정계수로 하여 1차년도 가중치에 곱한 후 이를 2차년도 종단가중치로 사용한다. 즉, 2차조사에 대하여 1차조사의 표본응답자가 응답했는지 여부를 반응변수로 하고, 성별, 연령, 지역, 장애유형, 장애정도, 경제활동상태 등의 변수를 설명변수로 하는 로지스틱회귀모형을 적합하여 응답률의 차이를 보정하여야 한다. 응답확률은 로지스틱 회귀모형 하에서 다음과 같이 표현할 수 있다.

$$\hat{p}_i = \frac{\exp\left(\sum_{j=0}^k \hat{\beta}_j x_{ij}\right)}{1 + \exp\left(\sum_{j=0}^k \hat{\beta}_j x_{ij}\right)}$$

1차년도와 2차년도 모두 응답한  $i$ 번째 사람의 종단가중치는 다음과 같이 계산한다.

$$\text{2차년도 종단가중치} = \text{1차년도 최종가중치} \times \frac{1}{\hat{p}_i}$$

여기서 사용된 설명변수는 성별, 연령, 지역 등의 변수와 장애유형, 장애정도, 경제활동상태 변수이다. 모든 설명변수를 사용한 포화모형을 사용하기보다는 응답여부와 유의한 상관성이 없는 장애유형과 장애정도를 제외한 모형을 적용하였다. 최종 로지스틱 회귀분석 결과는 다음과 같다.

〈표 2-21〉 최종 로지스틱 회귀분석 결과

구분	df	모형2(최종모형)	
		$\chi^2$	유의확률
성별	1	8.789	0.003
연령대	3	13.894	0.003
지역	15	79.821	0.000
경제활동상태	3	5.217	0.157

AIC = 2474.7

## 나. 횡단가중치 산출

### 1) 모집단 분포 정의 및 레이킹 비 가중치 산출

2차년도 횡단 가중치 산출을 위해 모집단 분포에 대한 벤치마킹 보정을 사용한다. 1차조사의 레이킹 비 가중치 산출에서 사용하였던 성, 연령, 지역, 장애유형, 장애정도, 경제활동상태 변수를 통하여 이들의 모집단 분포를 이용한 가중치 합 비율을 통해 벤치마킹 보정을 실시한다. 성, 연령, 지역, 장애유형, 장애정도에 관한 모집단 정보로는 2016년 12월 등록기준 16세에서 65세까지의 보건복지부 장애인 등록현황을 사용하였다. 경제활동상태에 대한 모집단 정보는 2017년 「장애인경제활동실태조사」의 경제활동상태에 대한 추정 결과를 사용하였다. 이 추정결과의 모집단 수와 2016년 12월 기준 장애인 등록 현황의 모집단 수가 일치하지 않기 때문에 추정 결과를 직접 사용할 수는 없다. 이를 위해 2017년 「장애인경제활동실태조사」의 경제활동상태에 대한 비율 추정 결과를 2016년 12월 기준 장애인 등록 현황의 모집단 수에 곱하여 경제활동상태에 대한 모집단 분포를 정의하였다.

레이킹 비 방법을 적용하여 가중치 보정 후 최종가중치를 산출하였다. 레이킹 비 기법을 이용하여 보정된 가중치는 아래와 같다.

$$w_{ij}^{rak} = w_{ij} \exp(\lambda' x_{ij})$$

여기서  $x_{ij}$ 는 레이킹 비 기법을 위해 사용된 변수들을 나타내는 지시변수 벡터이며,  $\lambda$ 는  $\sum w_{ij}^{rak} x_{ij} = t_x$ 의 해이며,  $t_x$ 는 각 레이킹 비 기법에 사용된 벤치마킹 변수들의 범주에 해당하는 모집단 장애인 분포를 나타낸다. 2차년도 자료를 통한 벤치마킹 변수들에 대한 모집단 분포와 최종가중치를 사용하여 얻어진 표본 자료의 분포는 아래와 같다.

〈표 2-22〉 주요 변수별 벤치마킹 모집단 분포 - 성/연령별

모집단 분포				표본 분포(횡단 가중치 사용)			
성별	빈도	연령	빈도	성별	빈도	연령	빈도
남자	940,519	16~29세	124,730	남자	940,518	16~29세	124,733
여자	471,294	30~39세	149,219	여자	471,295	30~39세	149,220
계	1,411,813	40~49세	296,597	계	1,411,813	40~49세	296,597
		50~65세	841,267			50~65세	841,263
		계	1,411,813			계	1,411,813

〈표 2-23〉 주요 변수별 벤치마킹 모집단 분포 - 시도별

모집단 분포		표본 분포(횡단 가중치 사용)	
지역	빈도	지역	빈도
서울	219,084	서울	219,084
부산	97,020	부산	97,020
대구	70,065	대구	70,065
인천	82,424	인천	82,424
광주	39,786	광주	39,786
대전	41,911	대전	41,911
울산	30,720	울산	30,720
세종	5,605	세종	5,605
경기	311,003	경기	311,003
강원	52,105	강원	52,105
충북	52,631	충북	52,631
충남	67,006	충남	67,006
전북	65,441	전북	65,441
전남	67,142	전남	67,142
경북	88,859	경북	88,859
경남	102,416	경남	102,416
제주	18,595	제주	18,595
계	1,411,813	계	1,411,813

〈표 2-24〉 주요 변수별 벤치마킹 모집단 분포 - 경제활동상태별

모집단 분포		표본 분포 (횡단 가중치 사용)	
지역	빈도	지역	빈도
취업자	687,797	취업자	687,797
미취업자	41,386	미취업자	41,386
취업희망자	163,894	취업희망자	163,894
취업미희망자	518,736	취업미희망자	518,736
계	1,411,813	계	1,411,813

〈표 2-25〉 주요 변수별 벤치마킹 모집단 분포 - 장애유형/장애정도별

모집단 분포				표본 분포 (횡단 가중치 사용)			
장애유형	빈도	장애등급	빈도	장애유형	빈도	장애등급	빈도
지체장애	716,552	중증	537,272	지체장애	716,548	중증	537,271
지체 외 신체의부 장애	112,362	경증	874,541	지체 외 신체의부 장애	112,362	경증	874,542
시각장애	134,250	계	1,411,813	시각장애	134,249	계	1,411,813
시각 외 감각장애	100,873			시각 외 감각장애	100,873		
정신적 장애	265,873			정신적 장애	265,877		
신체내부 장애	81,903			신체내부 장애	81,903		
계	1,411,813			계	1,411,812		

## 2) 최종가중치 산출(횡단가중치 극단값 처리)

횡단가중치에서 5000을 넘는 가중치(5034, 7731)가 존재하여, 이러한 극단값을 조정해주기 위해 두 개의 가중치를 그다음 최대값인 4720으로 조정해 주고 성, 연령, 장애유형, 장애정도 등의 변수에 대해서 다시 레이킹 보정을 실시한다. 이를 최종가중치로 사용한다. 최종가중치에 대한 기술통계량과 경제활동상태 변수의 모집단 및 표본 분포는 다음과 같다.

〈표 2-26〉 최종가중치에 대한 기술통계량

통계량			
관측값 수	4,214	관측값 합	1,411,813
평균	335.03	표준편차	462.328
중위수	179.85	분산	213,747.1
사분위 범위	327.22	범위	4,753.41
최솟값	3.35	최댓값	4,756.76

〈표 2-27〉 경제활동상태 변수의 모집단 분포와 표본 분포

모집단 분포		표본 분포 (최종가중치 사용)	
지역	빈도	지역	빈도
취업자	687,797	취업자	689,694
미취업자	41,386	미취업자	38,480
취업희망자	163,894	취업희망자	164,218
취업미희망자	518,736	취업미희망자	519,421
계	1,411,813	계	1,411,813

1차년도와 2차년도 가중치에 대한 기술통계량을 비교하면, 1차년도 최종가중치에 비하여 2차년도의 최종가중치는 전반적으로 증가하였지만 극단값 조정을 통하여 심각한 특이 가중치는 없는 것으로 판단되고 변동계수(CV)도 1차년도와 유사한 값을 보이고 있다.

〈표 2-28〉 1차년도, 2차년도 가중치 분포 현황

	설계 가중치	1차년도 최종가중치	2차년도 최종가중치
최대값	4548.8	4487.5	4756.8
99%	1851.4	2129.5	2260.2
95%	959.2	1054.3	1166.0
90%	697.1	722.5	772.3
75%	364.6	363.2	405.6
50%	167.1	160.2	179.8
25%	83.0	68.0	78.4
10%	37.4	30.5	34.2
5%	27.1	19.8	22.4
1%	11.6	9.5	10.6
최소값	7.7	3.2	3.4
평균	297.0	303.2	335.0
CV	123.9%	139.3%	138.0%

## 5. 3차조사 가중치 산출 및 추정

### 가. 무응답 특성 분석

#### 1) 차수별 응답 현황

2차년도와 3차년도에 응답자는 각각 4,214명과 4,104명이다. 1차년도 조사대상자 4,577명 중 2차조사에서 무응답자 338명과 사망 및 장애등록취소 등의 사유에 의한 부적격자 25명이 발생하였으며, 3차조사에서는 무응답자 427명과 부적격자 46명으로 나타났다.

〈표 2-29〉 1~3차조사 응답 분류 현황

구분	1차조사	2차조사	3차조사
응답	4,577	4,214	4,104
무응답		338	427
부적격		25	46
합계	4,577	4,577	4,577



〈표 2-30〉은 3차조사의 종단면 기준 응답유형의 분포를 나타내고 있다. 2차년도 응답자 4,214명 중 3,996명이 응답하였고 무응답은 195명, 부적격자는 25명으로 최종적으로 220명의 무응답 현상이 발생하였다.

〈표 2-30〉 무응답 사유 빈도표(3차조사 종단면 기준)

방문시 부재	방문시 이사	거절	강력거절	단기부재	장기부재	사망	장애등록 취소	합계
22	30	84	43	8	8	19	6	220
10.0%	13.6%	38.2%	19.5%	3.6%	3.6%	8.6%	2.7%	100.0%

1차년도 조사대상자 4,577명 중 3차년도에도 응답한 대상자는 4,104명이며 응답률은 약 89.7%이다. 응답을 얻지 못한 무응답자는 427명이며 부적격자는 46명으로 나타나 최종적으로 473명의 무응답 현상이 발생하였다.

〈표 2-31〉 무응답 사유 빈도표(3차조사 횡단면 기준)

방문시 부재	방문시 이사	거절	강력거절	단기부재	장기부재	사망	장애등록 취소	합계
47	61	167	131	8	13	36	10	473
9.9%	12.9%	35.3%	27.7%	1.7%	2.7%	7.6%	2.1%	100.0%

## 2) 가중치 대상

〈표 2-32〉은 1~3차조사 응답여부에 따른 종단면 응답자와 횡단면 응답자에 대한 현황을 보여준다. 3차 조사의 종단면 가중치의 대상은 1~3차의 모든 조사에 응답한 3,994명이며 횡단면 가중치 대상은 2차조사에서는 응답하지는 않았지만 3차조사에서 응답한 110명이 추가되어 4,104명이다.

〈표 2-32〉 1~3차조사 응답 분류 현황

종류	1차조사	2차조사	3차조사	대상자	비중(%)	종단면 응답	횡단면 응답
1	○	○	○	3,994	87.3	√	√
2	○	○	X	220	4.8		
3	○	X	○	110	2.4		√
4	○	X	X	253	5.5		
합계	-	-	-	4,577	100.0	3,994	4,104

## 나. 종단가중치 산출

## 1) 단위무응답 분석

3차조사 종단가중치 산출에 앞서 각 특성별로 단위무응답 현황을 살펴보고자 한다.

첫 번째로 성별에서 남성과 여성의 무응답률은 유사하여 유의한 차이가 나타나지 않았다.

〈표 2-33〉 성별 무응답 현황

구분	무응답		응답		계
	빈도	행%	빈도	행%	
남성	130	4.8%	2,600	95.2%	2,730
여성	65	4.5%	1,394	95.5%	1,459
계	195	4.7%	3,994	95.3%	4,189

$\chi^2 = 0.14$ , P-value=0.710, Cramér's V=0.007

두 번째로 장애인의 연령대에 따른 무응답률의 편차는 유의한 것으로 보인다. 무응답 비율을 살펴보면 50세 이상의 무응답률이 가장 낮은 것으로 나타났고 50세 미만의 연령층에서 무응답률이 상대적으로 높게 나타났다.

〈표 2-34〉 연령대별 무응답 현황

구분	무응답		응답		계
	빈도	행%	빈도	행%	
15-29세	36	5.5%	621	94.5%	657
30-39세	47	5.0%	891	95.0%	938
40-49세	63	5.3%	1128	94.7%	1191
50-64세	49	3.5%	1354	96.5%	1403
계	195	4.7%	3994	95.3%	4189

$\chi^2 = 3.63$ , P-value=0.09, Cramér's V=0.040

세 번째로 시도별 응답현황을 정리하여 시도 분류 및 응답여부와의 독립성 검정을 위한 카이제곱 통계와 크래머 V 통계를 산출하였을 때 시도간 무응답률의 편차는 유의미하게 나타났다. 특히 서울, 광주, 경기, 경북 지역의 무응답률이 6%보다 높게 나타났으며 부산, 대전, 울산, 충남, 전남 등의 지역의 무응답률은 상대적으로 낮게 나타났다.

〈표 2-35〉 시도별 무응답 현황

시도	무응답		응답		계
	빈도	행%	빈도	행%	
서울	51	6.4%	748	93.6%	799
부산	8	2.0%	396	98.0%	404
대구	8	3.4%	225	96.6%	233
인천	11	5.4%	193	94.6%	204
광주	11	8.9%	113	91.1%	124
대전	5	2.8%	172	97.2%	177
울산	0	0.0%	31	100.0%	31
경기	0	0.0%	2	100.0%	2
강원	63	6.8%	870	93.2%	933
충북	7	3.6%	185	96.4%	192
충남	5	2.9%	165	97.1%	170
전북	3	1.9%	152	98.1%	155
전남	6	3.6%	159	96.4%	165
경북	4	2.1%	191	97.9%	195
경남	7	6.3%	104	93.7%	111
제주	3	1.3%	230	98.7%	233
계	3	4.9%	58	95.1%	61

$\chi^2 = 44.23$ , P-value=0.000, Cramér's V=0.103

네 번째로 장애인의 경제활동상태의 무응답률 차이는 유의한 것으로 나타났고 취업희망자의 무응답률이 가장 높은 반면, 취업미희망자의 무응답률이 가장 낮았다.

〈표 2-36〉 경제활동상태별 무응답 현황

구분	무응답		응답		계
	빈도	행%	빈도	행%	
취업자	98	4.7%	1,999	95.3%	2,097
실업자	6	5.3%	108	94.7%	114
취업희망자	33	7.2%	424	92.8%	457
취업미희망자	58	3.8%	1,463	96.2%	1,521
계	195	4.7%	3,994	95.3%	4,189

$\chi^2 = 9.30$ , P-value=0.026, Cramér's V=0.047

다섯 번째로 장애인의 장애유형별 응답 현황을 살펴볼 때, 지체 외 신체외부장애의 무응답률이 가장 낮고 지체 장애의 무응답률이 가장 높지만 통계적으로 유의하지 않았다.

〈표 2-37〉 장애유형별 무응답 현황

구분	무응답		응답		계
	빈도	행%	빈도	행%	
지체장애	116	5.6%	1,941	94.4%	2,057
지체 외 신체 외부장애	9	2.9%	302	97.1%	311
시각장애	23	4.3%	516	95.7%	539
시각외감각장애	15	3.9%	367	96.1%	382
정신적장애	17	3.0%	547	97.0%	564
신체내부장애	15	4.5%	321	95.5%	336
계	195	4.7%	3,994	95.3%	4,189

$\chi^2 = 10.75$ , P-value=0.057, Cramér's V=0.051

마지막으로 장애정도에 따른 무응답률을 살펴보면 경증의 무응답률이 중증 무응답률보다 다소 높으나 장애정도에 따른 무응답률 차이는 통계적으로 유의하지 않다.

〈표 2-38〉 장애정도별 무응답 현황

구분	무응답		응답		계
	빈도	행%	빈도	행%	
중증	56	4.1%	1,314	95.9%	1,370
경증	139	4.9%	2,680	95.1%	2,819
계	195	4.7%	3,994	95.3%	4,189

$\chi^2 = 1.29$ , P-value=0.26, Cramér's V=0.019

## 2) 응답성향모형 적합

앞서 인구사회학적 특성과 장애유형 및 장애정도 특성별 무응답의 연관성 분석을 토대로 3차년도 종단 가중치를 산출하기 위해서는 조사대상자들의 다양한 속성(주로 인구사회학적 특성)을 반영한 응답성향모형을 적합하였다. 응답성향모형은 3차년도 응답여부는 과거에 관측된 보조변수들로 설명할 수 있다는 임의무응답(missing at random, 이하 MAR)을 가정한다. 따라서 임의무응답의 가정 하에서는 2차조사와 3차조사에서 모두 응답한 패널의 응답성향모형은 2차년도의 주요변수들로 설명할 수 있게 된다.

응답성향모형은 로지스틱 회귀모형을 사용하였다. 추정된 모형을 통해 조사대상자 각각의 응답성향점수(혹은 응답확률)를 계산하고 그 역수를 통해 무응답에 대한 보정계수로 산출하여 2차년도 종단 가중치에 곱한 후 3차년도 종단가중치를 계산하였다. 즉, 3차조사에 대하여 2차조사의 표본응답자가 응답했는지 여부를 반응변수로 하고, 보조변수를 설명변수로 하는 로지스틱회귀모형을 적합하여 응답률의 차이를 파악하고

이를 통해 보정하였다. 응답확률은 로지스틱회귀모형 하에서 다음과 같이 표현할 수 있다.

$$\hat{p}_i = \frac{\exp(\sum_{j=0}^k \hat{\beta}_j x_{ij})}{1 + \exp(\sum_{j=0}^k \hat{\beta}_j x_{ij})}$$

여기서  $\hat{p}_i$ 는  $i$ 번째 패널개체에 대한 응답확률의 추정값이고  $x_{ij}$ 는  $i$ 번째 패널개체의  $j$ 번째 보조변수 특성이며  $\hat{\beta}_j$ 는 회귀계수를 각각 나타낸다.

### 3) 최종 종단 가중치 산출

3차년도 종단면 응답자  $i$ 에 대한 종단 가중치는 2차년도 종단 가중치에 앞서 계산한 응답확률  $\hat{p}_i$ 를 나눠 다음과 같이 계산하였다.

$$\text{3차년도 종단가중치} = \text{2차년도 종단가중치} \times \frac{1}{\hat{p}_i}$$

여기서 사용된 설명변수는 성별, 연령, 지역 등의 변수와 장애유형, 장애정도, 경제활동상태 등의 변수들이며, 로지스틱 회귀모형에 대한 적합결과를 모델별로 산출한 후 최종적으로 연령대, 지역, 장애유형을 사용한 응답 성향 모형을 채택하였다.

〈표 2-39〉 로지스틱 회귀분석 결과

구분	최종 모형		
	df	$\chi^2$	유의확률
연령대	3	6.913	0.075
지역	16	48.304	0.000
장애유형	3	15.404	0.009
AIC = 1556.37			

종단가중치에 대한 기술통계량을 살펴보면 응답자 총 3,994명의 종단 가중치 평균은 346.3이고 최솟값 3.9에서 최댓값 5,032.7의 범위를 가지게 된다.

〈표 2-40〉 3차조사 최종 종단 가중치에 대한 기술통계량

통계량			
관측값 수	3,994	관측값 합	1,390,456
평균	346.3	표준편차	468.8
중위수	186.9	분산	219,761.0
사분위 범위	342.8	범위	5,028.8
최솟값	3.9	최댓값	5,032.7

〈표 2-41〉는 1~3차조사까지의 종단 가중치 분포의 현황을 정리하고 있다. 1차년도 최종가중치에 비하여 3차년도의 종단 가중치는 전반적으로 증가하였지만 변동계수(CV)는 1차년도에 비하여 다소 줄어들었음을 알 수 있다.

〈표 2-41〉 1~3차조사의 종단가중치 분포

통계량	설계 가중치	1차년도 최종가중치	2차년도 종단가중치	3차년도 종단가중치
최대값	4,548.8	4,487.5	4,768.5	5,032.7
99%	1,851.4	2,129.5	2,291.8	2,387.3
95%	959.2	1,054.3	1,140.2	1,189.2
90%	697.1	722.5	785.9	817.9
75%	364.6	363.2	402.4	422.0
50%	167.1	160.2	176.8	186.9
25%	83.0	68.0	74.8	79.3
10%	37.4	30.5	32.9	35.4
5%	27.1	19.8	21.8	23.3
1%	11.6	9.5	10.1	11.0
최소값	7.7	3.2	3.7	3.9
평균	297.0	303.2	329.9	346.3
CV	123.9%	139.3%	137.1%	135.4%

#### 다. 횡단가중치 산출

##### 1) 단위무응답 분석

3차조사 횡단가중치는 3차에 응답한 4,104명이 대상이며, 무응답 현황에 대해 분석한 결과 성별에서 남성과 여성의 무응답률은 유의미하고 남성이 여성보다 무응답률이 높다.

〈표 2-42〉 성별 무응답 현황

구분	무응답		응답		계
	빈도	행%	빈도	행%	
남성	303	10.2%	2,673	89.8%	2,976
여성	124	8.0%	1,431	92.0%	1,555
계	427	9.4%	4,104	90.6%	4,531

$$\chi^2 = 5.57, P\text{-value}=0.018, \text{Cramér's } V=0.036$$

두 번째로 연령대에 따른 무응답 비율을 살펴보면 15~29세 미만의 연령층과 40대 연령층에서 무응답률이 상대적으로 높고, 50세 이상의 무응답률이 가장 낮았다.

〈표 2-43〉 연령대별 무응답 현황

구분	무응답		응답		계
	빈도	행%	빈도	행%	
15-29세	86	10.8%	707	89.2%	793
30-39세	100	9.1%	1,000	90.9%	1,100
40-49세	153	11.6%	1,161	88.4%	1,314
50-64세	88	6.6%	1,236	93.4%	1,324
계	427	9.4%	4,104	90.6%	4,531

$\chi^2 = 21.57, P\text{-value}=0.000, \text{Cramér's } V=0.069$

세 번째로 무응답률의 시도 차이는 유의미하며, 서울, 대구, 인천, 광주, 울산, 경기 지역에서 무응답률이 약 10~14%로 높게 나타나며, 부산, 대전, 경남 지역에서 무응답 비율이 상대적으로 낮았다.

〈표 2-44〉 시도별 무응답 현황

시도	무응답		응답		계
	빈도	행%	빈도	행%	
서울	105	12.0%	773	88.0%	878
부산	16	3.8%	402	96.2%	418
대구	27	10.5%	231	89.5%	258
인천	25	10.7%	208	89.3%	233
광주	15	11.7%	113	88.3%	128
대전	10	5.5%	171	94.5%	181
울산	4	11.4%	31	88.6%	35
경기	140	13.3%	911	86.7%	1,051
강원	17	8.5%	184	91.5%	201
충북	13	7.3%	164	92.7%	177
충남	12	7.3%	152	92.7%	164
전북	13	7.5%	160	92.5%	173
전남	10	4.8%	198	95.2%	208
경북	10	8.3%	110	91.7%	120
경남	6	2.5%	233	97.5%	239
제주	4	6.0%	63	94.0%	67
계	427	9.4%	4,104	90.6%	4,531

$\chi^2 = 68.00, P\text{-value}=0.000, \text{Cramér's } V=0.123$

네 번째로 경제활동상태는 응답여부와 유의한 관계가 있는 것으로 나타났다. 취업희망자의 무응답률이 가장 높았고, 취업미희망자의 무응답률이 가장 낮았다.

〈표 2-45〉 경제활동상태별 무응답 현황

구분	무응답		응답		계
	빈도	행%	빈도	행%	
취업자	221	9.8%	2,036	90.2%	2,257
실업자	22	13.3%	143	86.7%	165
취업희망자	116	10.5%	987	89.5%	1,103
취업미희망자	68	6.8%	938	93.2%	1,006
계	427	9.4%	4,104	90.6%	4,531

$\chi^2 = 13.22$ , P-value=0.004, Cramér's V=0.054

다섯 번째로 장애유형별 무응답률의 현황을 보여준다. 장애유형별로는 지체 외 신체외부장애의 무응답률이 가장 낮고, 지체장애의 무응답률이 가장 높았다. 장애유형에 따른 무응답률 차이는 통계적으로 유의하게 나타난다.

〈표 2-46〉 장애유형별 무응답 현황

구분	무응답		응답		계
	빈도	행%	빈도	행%	
지체장애	241	10.8%	1,991	89.2%	2,232
지체 외 신체 외부장애	21	6.4%	307	93.6%	328
시각장애	58	9.9%	530	90.1%	588
시각외감각장애	33	8.0%	381	92.0%	414
정신적장애	42	7.0%	561	93.0%	603
신체내부장애	32	8.7%	334	91.3%	366
계	427	9.4%	4,104	90.6%	4,531

$\chi^2 = 14.07$ , P-value=0.015, Cramér's V=0.05

마지막으로 장애정도에 따른 무응답률을 살펴보면 경증의 무응답률이 중증의 무응답률보다 다소 높으며 장애정도에 따른 무응답률 차이는 유의하게 나타난다.

〈표 2-47〉 장애정도별 무응답 현황

구분	무응답		응답		계
	빈도	행%	빈도	행%	
중증	117	8.0%	1,338	92.0%	1,455
경증	310	10.1%	2,766	89.9%	3,076
계	427	9.4%	4,104	90.6%	4,531

$\chi^2 = 4.57$ , P-value=0.033, Cramér's V=0.033



## 2) 응답성향모형 적합

3차년도 횡단면 조사의 응답여부는 과거에 관측된 변수들로 설명할 수 있다는 임의응답(MAR)을 가정하였고 1차년도 조사 기준의 설명변수는 물론 2차조사와 3차조사의 응답여부도 함께 다음의 확률을 고려한 로지스틱 모형을 적합하였다.

$$P(3차\ 응답여부\ |\ 2차\ 응답여부,\ 1차\ 설명변수)$$

$$\hat{p}_i = \frac{\exp(\sum_{j=0}^k \hat{\beta}_j x_{ij})}{1 + \exp(\sum_{j=0}^k \hat{\beta}_j x_{ij})}$$

반응변수로는 3차조사 응답여부를 나타내는 지시자이며 2차년도 조사의 응답여부와 1차년도 기준의 성별, 연령, 지역, 장애유형, 장애정도, 경제활동상태 등의 변수를 설명변수로 하는 로지스틱회귀모형을 다음의 형태로 적합하였다.

3차년도의 응답확률 모형에서 모든 설명변수를 사용한 포화모형을 포함하여 유의하지 않은 설명변수를 하나씩 제거하여 모형을 적합한 결과를 정리해보았을 때 성별, 경제활동상태와 장애정도를 제외한 최종 모형이 적합도(AIC) 측면에서 더 좋음을 알 수 있다. 따라서 최종적으로 선택한 모형으로는 2차년도 응답여부, 연령대, 지역, 장애유형을 설명변수로 사용한 로지스틱 회귀모형이다.

〈표 2-48〉 로지스틱 회귀분석 결과

구분	최종 모형		
	df	$\chi^2$	유의확률
2차 응답여부	1	822.9	0.000
연령대	3	8.3	0.041
지역	16	44.9	0.000
장애유형	3	17.6	0.003

AIC = 1985.81

### 3) 레이킹 비 가중치

앞서 산출한 (추정된) 응답확률을  $\hat{p}_i$ 로 정의하면 3차년도에 응답한  $i$ 번째 개체의 종단 가중치는 다음과 같이 계산하였다.

$$\text{3차년도 무응답 보정 가중치} = \text{1차년도 최종가중치} \times \frac{1}{\hat{p}_i}$$

다음 단계로 무응답 보정 가중치에 대하여 모집단 분포에 대한 벤치마킹 보정을 적용하였다. 벤치마킹 보정을 위해 성별, 연령, 지역, 장애유형, 장애정도, 경제활동상태 변수를 보조정보로 사용하였고 이들의 모집단 분포와 조정된 가중치 합의 비율이 같아지도록 반복적 적합을 실시한다.

모집단의 보조정보 합계로는 2017년 12월 등록기준 17세에서 66세까지의 보건복지부 장애인 등록현황을 사용하였다. 2016년까지 장애인 등록명부에는 지체 3급의 중증 장애 비율이 모집단 값으로 있었던 반면 2017년 자료에서는 지체 3급의 중증 장애 비율이 주어지지 않았다.

따라서 지체 3급의 중증 장애 비율을 2018년 장애인경제활동실태조사를 통하여 추정해서 사용하였다. 레이킹 비 조정 방법을 적용하여 가중치를 보정한 후 횡단 가중치를 다음과 같이 산출하였다.

$$w_{ij}^{rak} = w_{ij} \exp(\lambda' x_{ij})$$

$x_{ij}$ 는 레이킹 비 조정 방법을 위해 사용된 보조변수들을 나타내는 지시변수 벡터이며,  $\lambda$ 는  $\sum w_{ij}^{rak} x_{ij} = t_x$ 의 해이며,  $t_x$ 는 각 레이킹 비 조정 방법에 사용된 벤치마킹 변수들의 범주에 해당하는 모집단 장애인 분포를 나타낸다.

### 4) 극단값 처리 및 최종가중치 산출

레이킹 비 조정 횡단가중치에는 5,000을 넘는 것이 총 4개가 존재하였다. 이들은 조사변수에 곱해져 합산되므로 경우에 따라서 통계량에 과도한 영향력을 갖게 된다. 따라서 이러한 과도한 가중치 값을 추가로 조정해 주기 위해 5,000보다 작은 최대값인 4,792으로 절사하여 조정해 준 다음, 앞서 고려한 바와 같이 성, 연령, 장애유형, 장애정도의 보조변수들에 대해서 다시 한번 레이킹 비 보정을 실시하여 이를 최종가중치로 산출하였다.

횡단면 최종가중치에 대한 기술통계량을 살펴보면, 평균 350.5이며 최솟값 3.0에서 최댓값 4,941의 4,938.1의 범위를 갖는다.

〈표 2-49〉 3차조사 최종 횡단가중치에 대한 기술통계량

통계량			
관측값 수	4,104	관측값 합	1,438,604
평균	350.5	표준편차	498.6
중위수	179.1	분산	248,636.3
사분위 범위	335.5	범위	4,938.1
최솟값	3.0	최댓값	4,941

〈표 2-50〉는 1~3차조사까지의 횡단가중치 분포의 현황을 정리하고 있다. 1차년도 최종가중치에 비하여 3차년도의 종단 가중치는 전반적으로 증가하였지만 변동계수(CV)는 1차년도에 비하여 다소 줄어들었음을 알 수 있다.

〈표 2-50〉 1~3차조사의 횡단가중치 분포

통계량	설계 가중치	1차년도 최종가중치	2차년도 횡단가중치	3차년도 횡단가중치
최대값	4,548.8	4,487.5	4,756.8	4,941.0
99%	1,851.4	2,129.5	2,260.2	2,396.5
95%	959.2	1,054.3	1,166.0	1,220.5
90%	697.1	722.5	772.3	849.5
75%	364.6	363.2	405.6	413.4
50%	167.1	160.2	179.8	179.1
25%	83.0	68.0	78.4	77.9
10%	37.4	30.5	34.2	35.3
5%	27.1	19.8	22.4	22.8
1%	11.6	9.5	10.6	10.9
최소값	7.7	3.2	3.4	3.0
평균	297.0	303.2	335.0	350.5
CV	123.9%	139.3%	138.0%	142.2%

## 6. 4차조사 가중치 산출 및 추정

### 가. 종단 가중치 산출

#### 1) 차수별 응답 현황

4차년도 종단 가중치는 1차부터 4차에 모두 응답한 3,818명을 대상으로 산출한다. 4차 조사 종단 가중치의 산출에 앞서 단위무응답의 현황을 살펴보고 응답성향분석을 통해 단위무응답에 영향을 미치는 조사대상자 특성을 우선적으로 파악하였다.

첫 번째로 남성과 여성 간의 무응답률은 서로 유사하여 유의한 차이가 아님을 알 수 있다.

〈표 2-51〉 성별 무응답 현황

구분	무응답		응답		계
	빈도	행%	빈도	행%	
남성	103	4.0%	2,484	96.0%	2,587
여성	53	3.8%	1,334	96.2%	1,387
계	156	3.9%	3,818	96.1%	3,974

참고:  $\chi^2 = 0.03$ , P-value=0.871, Cramér's V=0.004

두 번째로 연령대에 따른 무응답률의 편차는 유의하며, 무응답 비율을 살펴보면 50세 이상의 무응답률이 가장 낮은 것으로 나타났고 40세 미만의 연령층에서 무응답률이 상대적으로 높게 나타났다.

〈표 2-52〉 연령대별 무응답 현황

구분	무응답		응답		계
	빈도	행%	빈도	행%	
15-29세	33	5.8%	536	94.2%	569
30-39세	46	5.6%	776	94.4%	822
40-49세	39	3.5%	1,064	96.5%	1,103
50-64세	38	2.6%	1,442	97.4%	1,480
계	156	3.9%	3,818	96.1%	3,974

참고:  $\chi^2 = 19.06$ , P-value=0.000, Cramér's V=0.069

세 번째로 시도간 무응답률의 편차는 유의미하게 나타났다. 특히 인천, 충북지역에서 무응답률이 6%보다 높았고 부산, 울산, 세종, 경남 등의 지역에서 무응답률은 상대적으로 낮았다.

〈표 2-53〉 시도별 무응답 현황

시도	무응답		응답		계
	빈도	행%	빈도	행%	
서울	32	4.4%	699	95.6%	731
부산	9	2.3%	387	97.7%	396
대구	10	4.5%	212	95.5%	222
인천	12	6.3%	180	93.8%	192
광주	3	2.7%	109	97.3%	112
대전	4	2.3%	172	97.7%	176
울산	0	0.0%	28	100.0%	28
세종	0	0.0%	3	100.0%	3
경기	42	4.9%	820	95.1%	862
강원	9	4.9%	176	95.1%	185
충북	11	6.6%	155	93.4%	166
충남	5	3.3%	146	96.7%	151
전북	4	2.6%	152	97.4%	156
전남	5	2.6%	189	97.4%	194
경북	5	4.6%	103	95.4%	108
경남	3	1.3%	231	98.7%	234
제주	2	3.4%	56	96.6%	58
계	156	3.9%	3,818	96.1%	3,974

참고:  $\chi^2 = 21.27$ , P-value=0.168, Cramér's V=0.073

네 번째로 경제활동상태에 따른 무응답률의 차이는 유의하지 않은 것으로 나타났다. 실업자의 무응답률이 가장 높은 반면, 취업(미)희망자의 무응답률이 가장 낮았다.

〈표 2-54〉 경제활동상태별 무응답 현황

구분	무응답		응답		계
	빈도	행%	빈도	행%	
취업자	85	4.2%	1,959	95.8%	2,044
실업자	5	6.7%	70	93.3%	75
취업희망자	13	3.4%	372	96.6%	385
취업미희망자	53	3.6%	1,417	96.4%	1,470
계	156	3.9%	3,818	96.1%	3,974

참고:  $\chi^2 = 2.50$ , P-value=0.476, Cramér's V=0.025

다섯 번째로 장애인의 장애유형별 응답 현황을 살펴볼 때, 시각장애의 무응답률이 가장 높고 지체 외 신체 외부장애의 무응답률이 그다음으로 높았다.

〈표 2-55〉 장애유형별 무응답 현황

구분	무응답		응답		계
	빈도	행%	빈도	행%	
지체장애	73	3.8%	1861	96.2%	1934
지체 외 신체 외부장애	13	4.3%	286	95.7%	299
시각장애	27	5.3%	485	94.7%	512
시각외감각장애	14	3.8%	354	96.2%	368
정신적장애	18	3.3%	524	96.7%	542
신체내부장애	11	3.4%	308	96.6%	319
계	156	3.9%	3818	96.1%	3974

참고:  $\chi^2 = 3.46$ , P-value=0.630, Cramér's V=0.029

마지막으로 장애정도에 따른 무응답률을 살펴보면 중증과 경증의 무응답률은 거의 비슷하게 나타남을 알 수 있다.

〈표 2-56〉 장애정도별 무응답 현황

구분	무응답		응답		계
	빈도	행%	빈도	행%	
중증	52	4.0%	1249	96.0%	1301
경증	104	3.9%	2569	96.1%	2673
계	156	3.9%	3818	96.1%	3974

참고:  $\chi^2 = 0.01$ , P-value=0.940, Cramér's V=0.003

## 2) 응답성향모형 적합

앞서 인구사회학적 특성과 장애유형 및 장애정도 특성별 무응답의 연관성 분석을 토대로 4차년도 종단 가중치를 산출하기 위해서는 조사대상자들의 다양한 속성(주로 인구사회학적 특성)을 반영한 응답성향모형을 적합하였다. 응답성향모형은 3차년도 응답여부는 과거에 관측된 보조변수들로 설명할 수 있다는 임의무응답(missing at random, 이하 MAR)을 가정한다. 따라서 임의무응답의 가정 하에서는 2차 조사와 3차 조사에서 모두 응답한 패널의 응답성향모형은 2차년도의 주요변수들로 설명할 수 있게 된다.

응답성향모형은 로지스틱 회귀모형을 사용하였다. 추정된 모형을 통해 조사대상자 각각의 응답성향점수(혹은 응답확률)를 계산하고 그 역수를 통해 무응답에 대한 보정계수로 산출하여 3차년도 종단 가중치에 곱한 후 이를 4차년도 종단가중치를 계산하였다.

$$\hat{p}_i = \frac{\exp\left(\sum_{j=0}^k \hat{\beta}_j x_{ij}\right)}{1 + \exp\left(\sum_{j=0}^k \hat{\beta}_j x_{ij}\right)}$$

여기서  $\hat{p}_i$ 는  $i$ 번째 패널개체에 대한 응답확률의 추정값이고  $x_{ij}$ 는  $i$ 번째 패널개체의  $j$ 번째 보조변수 특성이며  $\hat{\beta}_j$ 는 회귀계수를 각각 나타낸다.

## 3) 최종 종단 가중치 산출

4차년도 종단면 응답자  $i$ 에 대한 종단 가중치는 3차년도 종단가중치에 앞서 계산한 응답확률  $\hat{p}_i$ 를 나눠 다음과 같이 계산하였다.

$$4\text{차년도 종단가중치} = 4\text{차년도 종단가중치} \times \frac{1}{\hat{p}_i}$$

여기서 사용된 설명변수는 성별, 연령, 지역 등의 변수와 장애유형, 장애정도, 경제활동상태 등의 변수들이다. 그러나 모든 변수를 사용한 모형보다는 연령과 지역만을 설명변수로 사용한 모형이 적합도(AIC) 측면에서 더 좋을 수 있다. 따라서 응답성향모형의 적합과정에서 고려한 모든 설명변수를 사용한 포화모형을 사용하기보다는 최종 모형으로 연령대와 지역을 사용한 응답성향 모형을 사용한다.

〈표 2-57〉 로지스틱 회귀분석 결과

구분	최종 모형		
	df	$\chi^2$	유의확률
연령대	3	18.710	0.000
지역	16	23.597	0.099
AIC = 1313.64			

4차년도 중단 가중치에 대한 기술 통계량을 보여주고 있다. 중단면 응답자 총 3,818명의 중단 가중치 평균은 360.1이고 최솟값 4.1에서 최댓값 5,181.9의 범위를 갖는다.

〈표 2-58〉 4차조사 최종 중단 가중치에 대한 기술통계량

통계량			
관측값 수	3,818	관측값 합	1,374,856
평균	360.1	표준편차	482.7
중위수	196.2	분산	232,960.9
사분위 범위	355.8	범위	5,181.9
최솟값	4.1	최댓값	5,186.0

다음으로 1~4차조사까지의 중단 가중치 분포의 현황을 정리하고 있다. 1차년도 최종가중치에 비하여 4차년도의 중단 가중치는 전반적으로 증가하였지만 변동계수(CV)는 1차년도에 비하여 다소 줄어들었음을 알 수 있다.

〈표 2-59〉 1~4차조사의 중단 가중치 분포

통계량	설계 가중치	1차년도 최종가중치	2차년도 중단 가중치	3차년도 중단 가중치	4차년도 중단 가중치
최대값	4,548.8	4,487.5	4,768.5	5,032.7	5,186.0
99%	1,851.4	2,129.5	2,291.8	2,387.3	2,467.8
95%	959.2	1,054.3	1,140.2	1,189.2	1,250.0
90%	697.1	722.5	785.9	817.9	847.2
75%	364.6	363.2	402.4	422.0	440.0
50%	167.1	160.2	176.8	186.9	196.2
25%	83.0	68.0	74.8	79.3	84.3
10%	37.4	30.5	32.9	35.4	37.5
5%	27.1	19.8	21.8	23.3	25.2
1%	11.6	9.5	10.1	11.0	11.9
최소값	7.7	3.2	3.7	3.9	4.1
평균	297.0	303.2	329.9	346.3	360.1
CV	123.9%	139.3%	137.1%	135.4%	134.0%

## 나. 횡단 가중치 산출

### 1) 단위무응답 분석

4차년도 횡단 가중치는 4차에 응답한 3,995명을 대상으로 산출한다. 4차년도 응답여부와 4차년도 응답자 모두가 응답한 1차년도의 주요 특성 변수들 간의 연관성을 살펴봄으로써 응답여부에 중요하게

영향을 미치는 특성 변수를 파악하고자 하였다.

첫 번째로 성별 무응답률의 현황이다. 남성과 여성의 무응답 비율은 비슷하게 나타났다.

〈표 2-60〉 성별 무응답 현황

구분	무응답		응답		계
	빈도	행%	빈도	행%	
남성	349	11.8%	2,609	88.2%	2,958
여성	160	10.3%	1,386	89.7%	1,546
계	509	11.3%	3,995	88.7%	4,504

$\chi^2 = 1.985$ , P-value=0.159, Cramér's V=0.022

두 번째로 연령의 분포는 유의한 관계가 있다. 연령대에 따른 무응답 비율을 살펴보면 15~29세 미만의 연령층의 무응답률이 가장 높으며 50세 이상의 무응답률이 가장 낮았다.

〈표 2-61〉 연령대별 무응답 현황

구분	무응답		응답		계
	빈도	행%	빈도	행%	
15-29세	109	13.8%	681	86.2%	790
30-39세	134	12.2%	963	87.8%	1,097
40-49세	165	12.7%	1,139	87.3%	1,304
50-64세	101	7.7%	1,212	92.3%	1,313
계	509	11.3%	3,995	88.7%	4,504

$\chi^2 = 25.263$ , P-value=0.000, Cramér's V=0.075

세 번째로 시도별 무응답률 현황을 보여주고 있다. 무응답률의 시도 차이는 유의미하며, 서울, 대구, 인천, 광주, 울산, 경기, 강원, 충북 지역에서 무응답률이 약 10~14%로 높게 나타나며, 부산, 대전, 충남, 전북, 전남, 경남, 제주 지역에서 무응답 비율이 상대적으로 낮았다.



〈표 2-62〉 시도별 무응답 현황

시도	무응답		응답		계
	빈도	행%	빈도	행%	
서울	117	13.4%	756	86.6%	873
부산	24	5.8%	393	94.2%	417
대구	36	14.0%	221	86.0%	257
인천	29	12.4%	204	87.6%	233
광주	14	11.0%	113	89.0%	127
대전	12	6.7%	167	93.3%	179
울산	4	11.8%	30	88.2%	34
경기	169	16.2%	875	83.8%	1,044
강원	23	11.5%	177	88.5%	200
충북	19	10.9%	155	89.1%	174
충남	11	6.8%	151	93.2%	162
전북	13	7.6%	158	92.4%	171
전남	12	5.8%	196	94.2%	208
경북	13	10.8%	107	89.2%	120
경남	8	3.3%	231	96.7%	239
제주	5	7.6%	61	92.4%	66
계	509	11.3%	3,995	88.7%	4,504

$\chi^2 = 75.512$ , P-value=0.000, Cramér's V=0.129

네 번째로 경제활동상태는 응답여부와 유의한 관계가 있는 것으로 나타났다. 실업자의 무응답률이 가장 높았고, 취업미희망자의 무응답률이 가장 낮았다.

〈표 2-63〉 경제활동상태별 무응답 현황

구분	무응답		응답		계
	빈도	행%	빈도	행%	
취업자	268	11.9%	1,984	88.1%	2,252
실업자	24	14.5%	141	85.5%	165
취업희망자	134	12.2%	962	87.8%	1,096
취업미희망자	83	8.4%	908	91.6%	991
계	509	11.3%	3,995	88.7%	4,504

$\chi^2 = 11.938$ , P-value=0.008, Cramér's V=0.051

다섯 번째로 장애유형별로는 지체 외 신체 외부장애, 정신적 장애, 신체 내부장애의 무응답률이 낮은 편이고, 시각장애와 지체장애의 무응답률이 높게 나타난다.

〈표 2-64〉 장애유형별 무응답 현황

구분	무응답		응답		계
	빈도	행%	빈도	행%	
지체장애	270	12.1%	1,953	87.9%	2,223
지체 외 신체 외부장애	31	9.5%	294	90.5%	325
시각장애	74	12.7%	510	87.3%	584
시각외감각장애	44	10.6%	370	89.4%	414
정신적장애	57	9.5%	541	90.5%	598
신체내부장애	33	9.2%	327	90.8%	360
계	509	11.3%	3,995	88.7%	4,504

$\chi^2 = 7.374$ , P-value=0.194, Cramér's V=0.040

마지막으로 장애정도에 따른 무응답률을 살펴보면 경증의 무응답률이 중증의 무응답률보다 다소 높게 나타난다.

〈표 2-65〉 장애정도별 무응답 현황

구분	무응답		응답		계
	빈도	행%	빈도	행%	
중증	144	10.0%	1,294	90.0%	1,438
경증	365	11.9%	2,701	88.1%	3,066
계	509	11.3%	3,995	88.7%	4,504

$\chi^2 = 3.305$ , P-value=0.069, Cramér's V=0.028

## 2) 무응답 조정 가중치 산출

무응답 보정을 위해서는 조사대상자들의 다양한 속성(주로 인구사회학적 특성)을 반영한 응답성향 모형을 적합하였다. 4차년도 횡단면 조사의 응답여부는 과거에 관측된 변수들로 설명할 수 있다는 임의응답(MAR)을 가정하였고 1차년도 조사 기준의 설명변수는 물론 2차 조사와 3차 조사의 응답여부도 함께 다음의 확률을 고려한 로지스틱 모형을 적합하였다.

$P(4차\ 응답여부 | 3차\ 응답여부, 2차\ 응답여부, 1차\ 설명변수)$

$$\hat{p}_i = \frac{\exp\left(\sum_{j=0}^k \hat{\beta}_j x_{ij}\right)}{1 + \exp\left(\sum_{j=0}^k \hat{\beta}_j x_{ij}\right)}$$

반응변수로는 4차조사 응답여부를 나타내는 지시자이고 2차년도와 3차년도 조사의 응답여부와 1차년도 기준의 성별, 연령, 지역, 장애유형, 장애정도, 경제활동상태 등의 변수를 설명변수로 하는 로지스틱회귀모형을 위의 형태로 적합하였다.

4차년도의 응답확률 모형에서 모든 설명변수를 사용한 포화모형을 포함하여 유의하지 않은 설명변수를 하나씩 제거하여 모형을 적합한 결과를 정리했을 때, 모든 설명변수를 사용하는 모형보다는 성별, 경제활동상태와 장애유형, 장애정도를 제외한 모형이 적합도(AIC) 측면에서 더 좋음을 알 수 있다. 따라서 최종적으로 선택한 모형으로는 3차년도와 2차년도 응답여부, 연령대, 지역을 설명변수로 사용한 로지스틱 회귀 모형이다.

〈표 2-66〉 로지스틱 회귀분석 결과

구분	최종 모형		
	df	$\chi^2$	유의확률
3차 응답여부	1	1312.8	0.000
2차 응답여부	1	40.9	0.000
연령대	3	12.7	0.005
지역	15	24.0	0.065

AIC = 1829.28

### 3) 레이킹 비 가중치

앞서 산출한 (추정된) 응답확률을  $\hat{p}_i$ 로 정의하면 4차년도에 응답한  $i$ 번째 개체의 중단 가중치는 다음과 같이 계산하였다.

$$\text{4차년도 무응답 보정 가중치} = \text{1차년도 최종가중치} \times \frac{1}{\hat{p}_i}$$

4차년도 횡단 가중치 산출을 위해 무응답을 보정한 가중치에 대하여 모집단 분포에 대한 벤치마킹 보정을 사용한다. 성별, 연령, 지역, 장애유형, 장애정도, 경제활동상태 변수를 통하여 이들의 모집단 분포를 이용한 가중치 합 비율을 통해 벤치마킹 보정을 실시한다. 성, 연령, 지역, 장애유형, 장애정도에 관한 모집단 정보로는 2018년 12월 등록기준 18세에서 67세까지의 보건복지부 장애인 등록현황을 사용하였다.

중/경증 장애정도 분포 추정에 있어 지체 3급 중증 장애 비율은 2015년, 2016년에는 장애인 등록 명부에 모집단 값이 제공되었으나 2017년과 2018년에는 장애인 등록명부에 모집단 값이 제공되지 않았다. 따라서 2017년과, 2018년에는 지체 3급 중증 비율은 “장애인경제활동실태조사”로부터 추정된 값을 사용하였다. 모집단 분포의 변화를 살펴보기 위해 2015년부터 2018년까지 장애정도 모집단 분포 현황을 분석하였다.

〈표 2-67〉 1~4차년도 장애정도 모집단 분포 현황

구분	2015년		2016년		2017년		2018년	
	빈도	비율	빈도	비율	빈도	비율	빈도	비율
중증	528,223	38.06%	537,272	38.06%	549,943	38.23%	514,522	35.71%
경증	859,579	61.94%	874,541	61.94%	888,661	61.77%	927,781	64.29%

중/경증 장애정도 비율은 2015년부터 2017년까지 중증 비율이 약 38%대였으나 2018년에는 약 35% 정도로 감소하였다. 이러한 감소 원인이 지체 3급내 중증 장애비율의 변화인지 살펴보기 위하여 2015년부터 2018년까지 지체 3급내 중증 장애 비율을 살펴보았다. 지체 3급내 중증 장애 비율은 2017년을 제외하면 약 40%대로 비슷하게 나타난다.

〈표 2-68〉 1~4차년도 지체 3급내 중증 장애 비율

구분	2015년	2016년	2017년	2018년
장애인등록명부(모집단)	39.8%	40.0%	44.1%(추정)	39.7%(추정)

2015년부터 2018년까지 장애등급의 모집단 분포 현황을 살펴보면, 중증 장애등급인 1, 2급이 2015년부터 2017년에 비하여 2018년에 감소하였고 경증 장애등급인 4, 5, 6급이 증가하였음을 알 수 있다.

〈표 2-69〉 1~4차년도 장애등급 모집단 분포 현황

구분	2015년		2016년		2017년		2018년	
	빈도	비율	빈도	비율	빈도	비율	빈도	비율
1급	126,847	9.1%	127,787	9.1%	128,933	9.0%	120,109	8.3%
2급	206,750	14.9%	210,365	14.9%	214,000	14.9%	197,045	13.7%
3급	268,845	19.4%	273,479	19.4%	278,139	19.3%	263,825	18.3%
4급	152,783	11.0%	155,190	11.0%	158,302	11.0%	166,983	11.6%
5급	248,722	17.9%	254,218	18.0%	260,901	18.1%	271,834	18.8%
6급	383,855	27.7%	390,774	27.7%	398,329	27.7%	422,506	29.3%

2015년부터 2018년까지 지체 외/내 장애정도의 모집단 분포 현황을 보면 지체 외, 지체 내 장애 모두에서 2015년부터 2017년에 비하여 2018년에 중증 장애 비율이 감소하였음을 알 수 있다. 따라서 2018년의 중증 장애정도 감소는 지체 3급 내 중증 장애 비율의 변화라기보다는 모집단 자체의 변화라고 판단할 수 있다.

〈표 2-70〉 1~4차년도 지체외/내 장애정도 모집단 분포 현황

구분		2015년		2016년		2017년		2018년	
		빈도	비율	빈도	비율	빈도	비율	빈도	비율
지체외	중증	425,264	63.0%	434,553	62.5%	443,681	61.8%	438,260	59.9%
	경증	249,258	37.0%	260,708	37.5%	274,144	38.2%	293,106	40.1%
	합계	674,522	100.0%	695,261	100.0%	717,825	100.0%	731,366	100.0%
지체내	중증	102,959	14.4%	102,719	14.3%	106,262	14.7%	76,262	10.7%
	경증	610,321	85.6%	613,833	85.7%	614,517	85.3%	634,676	89.3%
	합계	713,280	100.0%	716,552	100.0%	720,779	100.0%	710,938	100.0%

경제활동상태에 대한 모집단 정보는 2019년 「장애인경제활동실태조사」의 경제활동상태에 대한 추정 결과를 사용하였다. 이 추정 결과의 모집단 수와 2018년 12월 기준 장애인 등록 현황의 모집단 수가 일치하지 않기 때문에 추정 결과를 직접 사용할 수는 없다. 이를 위해 2019년 「장애인경제활동실태조사」의 경제활동상태에 대한 비율 추정 결과를 2018년 12월 기준 장애인 등록 현황의 모집단 수에 곱하여 경제활동상태에 대한 모집단 분포를 정의하였다. 레이킹 비 조정 방법을 적용하여 가중치를 보정한 후 횡단 가중치를 다음과 같이 산출하였다.

$$w_{ij}^{rak} = w_{ij} \exp\left(\lambda \frac{x_{ij}}{t_x}\right)$$

여기서  $x_{ij}$ 는 레이킹 비 조정 방법을 위해 사용된 보조변수들을 나타내는 지시변수 벡터이며,  $\lambda$ 는  $\sum w_{ij}^{rak} x_{ij} = t_x$ 의 해이며,  $t_x$ 는 각 레이킹 비 조정 방법에 사용된 벤치마킹 변수들의 범주에 해당하는 모집단 장애인 분포를 나타낸다.

#### 4) 극단값 처리 및 최종가중치 산출

레이킹 비 조정 횡단가중치에는 설계 가중치의 최댓값이었던 약 5,000을 넘는 가중치들이 존재하였다. 이들은 조사변수에 곱해져 합산되므로 경우에 따라서 통계량에 과도한 영향력을 갖게 된다. 따라서 이러한 과도한 가중치 값을 추가로 조정해 주기 위해 5,000보다 작은 최댓값인 4,632으로 절사하여 조정해 준 다음, 앞서 고려한 바와 같이 성, 연령, 장애유형, 장애정도의 보조변수들에 대해서 다시 한 번 레이킹 비 보정을 실시하여 이를 최종가중치로 산출하였다.

횡단면 최종가중치에 대한 기술통계량을 보면 평균이 361.0이며 최솟값은 3.2이고 최댓값은 4,536.3이다.

〈표 2-71〉 4차조사 최종 횡단 가중치에 대한 기술통계량

통계량			
관측값 수	3,995	관측값 합	1,442,303
평균	361.0	표준편차	511.1
중위수	184.4	분산	261,198.4
사분위 범위	350.7	범위	4,533.1
최솟값	3.2	최댓값	4,536.3

벤치마킹 보정을 위해 고려하였던 보조변수인 성별, 연령대, 지역, 장애유형, 장애정도, 경제활동상태 변수의 모집단 총합과 표본가중합의 분포를 살펴보았으며, 모집단 총합과 거의 차이가 없음을 알 수 있었다.

1~4차 조사들의 횡단 가중치의 기술통계량 및 분포를 비교하고 있다. 1차년도 최종가중치에 비하여 4차년도의 횡단 가중치는 전반적으로 증가하였지만, 극단값 조정을 통하여 심각한 특이 가중치는 없는 것으로 판단되고 변동계수(CV)도 1차년도에 비하여 약 2% 정도 증가하였다.

〈표 2-72〉 1~4차조사의 횡단 가중치 분포

통계량	설계 가중치	1차년도 최종가중치	2차년도 횡단 가중치	3차년도 횡단 가중치	4차년도 최종가중치
최대값	4,548.8	4,487.5	4,756.8	4,941.0	4536.3
99%	1,851.4	2,129.5	2,260.2	2,396.5	2631.4
95%	959.2	1,054.3	1,166.0	1,220.5	1264.9
90%	697.1	722.5	772.3	849.5	862.7
75%	364.6	363.2	405.6	413.4	428.0
50%	167.1	160.2	179.8	179.1	184.4
25%	83.0	68.0	78.4	77.9	77.3
10%	37.4	30.5	34.2	35.3	33.8
5%	27.1	19.8	22.4	22.8	22.8
1%	11.6	9.5	10.6	10.9	10.7
최소값	7.7	3.2	3.4	3.0	3.2
평균	297.0	303.2	335.0	350.5	361.0
CV	123.9%	139.3%	138.0%	142.2%	141.6%

## 7. 5차조사 가중치 산출 및 추정

### 가. 종단 가중치 산출

#### 1) 차수별 응답 현황 및 단위 무응답 분석

5차년도 종단 가중치는 1차부터 5차에 모두 응답한 3,680명을 대상으로 산출한다. 5차 조사 종단 가중치의 산출에 앞서 단위무응답의 현황을 살펴보고 응답성향분석을 통해 단위무응답에 영향을 미치는 조사대상자 특성을 우선적으로 파악하고자 하였다.

〈표 2-73〉은 장애인 거주 시도별 응답현황을 정리하였고 더불어 시도 분류와 응답여부와의 독립성 검정을 위한 카이제곱 통계와 크래머 V 통계를 산출하였다. 시도간 무응답률의 편차는 유의미하게 나타났다. 특히 제주 지역에서 무응답률이 약 14%로 높게 나왔고, 그 외 지역에서 무응답률은 약 5% 미만으로 상대적으로 낮게 나타났다.

〈표 2-73〉 시도별 종단면 무응답 현황

시도	무응답		응답		계
	빈도	행%	빈도	행%	
서울	32	4.6%	661	95.4%	693
부산	5	1.3%	378	98.7%	383
대구	3	1.4%	206	98.6%	209
인천	8	4.5%	171	95.5%	179
광주	1	0.9%	108	99.1%	109
대전	4	2.4%	165	97.6%	169
울산	0	0.0%	29	100.0%	29
세종	0	0.0%	5	100.0%	5
경기	30	3.7%	790	96.3%	820
강원	4	2.3%	171	97.7%	175
충북	4	2.6%	149	97.4%	153
충남	3	2.0%	144	98.0%	147
전북	3	2.0%	149	98.0%	152
전남	5	2.7%	181	97.3%	186
경북	6	5.8%	97	94.2%	103
경남	5	2.1%	229	97.9%	234
제주	8	14.5%	47	85.5%	55
계	121	3.2%	3,680	96.8%	3,801

참고:  $\chi^2 = 44.26$ , P-value=0.000, Cramér's V=0.108

<표 2-74> 은 장애인의 성별 응답현황을 보여주고 있다. 남성과 여성 간의 무응답률은 서로 유사하여 유의한 차이가 아님을 알 수 있다.

<표 2-74> 성별 종단면 무응답 현황

구분	무응답		응답		계
	빈도	행%	빈도	행%	
남성	82	3.3%	2,389	96.7%	2,471
여성	39	2.9%	1,291	97.1%	1,330
계	121	3.2%	3,680	96.8%	3,801

참고:  $\chi^2 = 0.30$ , P-value=0.582, Cramér's V=0.009

<표 2-75>은 장애인의 연령대별 응답률 현황을 나타내고 있다. 연령대에 따른 무응답률의 편차는 유의한 것으로 보인다. 무응답 비율을 살펴보면 50세 이상의 무응답률이 약 2%로 가장 낮은 것으로 나타났고 30세 미만의 연령층에서 무응답률이 약 4%로 상대적으로 높게 나타났다.

<표 2-75> 연령대별 종단면 무응답 현황

구분	무응답		응답		계
	빈도	행%	빈도	행%	
18-29세	19	3.9%	471	96.1%	490
30-39세	31	4.5%	663	95.5%	694
40-49세	35	3.4%	1,009	96.6%	1,044
50-67세	36	2.3%	1,537	97.7%	1,573
계	121	3.2%	3,680	96.8%	3,801

참고:  $\chi^2 = 8.66$ , P-value=0.034, Cramér's V=0.048

<표 2-76>은 장애인의 경제활동상태별 응답률의 현황을 정리하고 있다. 경제활동상태에 따른 무응답률의 편차는 유의하지 않은 것으로 나타났다. 실업자의 무응답률이 가장 높은 반면, 취업미희망자의 무응답률이 가장 낮았다.

<표 2-76> 경제활동상태별 종단면 무응답 현황

구분	무응답		응답		계
	빈도	행%	빈도	행%	
취업자	66	3.4%	1,888	96.6%	1,954
실업자	3	6.1%	46	93.9%	49
취업희망자	15	4.0%	357	96.0%	372
취업미희망자	37	2.6%	1,389	97.4%	1,426
계	121	3.2%	3,680	96.8%	3,801

참고:  $\chi^2 = 4.09$ , P-value=0.252, Cramér's V=0.033

<표 2-77> 은 장애인의 장애유형별 응답률의 현황을 보여준다. 장애유형에 따른 무응답률의 편차는 유의미하게 나타났다. 신체내부장애의 무응답률이 가장 높고 시각외감각장애의 무응답률이 그다음 높으며, 시각장애의 무응답률이 가장 작게 나타났다. 그 외 장애유형의 무응답률은 비슷하게 나타났다.

<표 2-77> 장애유형별 종단면 무응답 현황

구분	무응답		응답		계
	빈도	행%	빈도	행%	
지체장애	54	2.9%	1,803	97.1%	1,857
지체 외 신체 외부장애	11	3.9%	274	96.1%	285
시각장애	6	1.2%	478	98.8%	484
시각외감각장애	16	4.5%	338	95.5%	354
정신적장애	18	3.5%	502	96.5%	520
신체내부장애	16	5.3%	285	94.7%	301
계	121	3.2%	3,680	96.8%	3,801

참고:  $\chi^2 = 13.46$ , P-value=0.020, Cramér's V=0.059

<표 2-78>은 장애인의 장애정도별 응답률 현황이다. 장애정도에 따른 무응답률의 편차는 유의하지 않은 것으로 나타났다. 장애정도에 따른 무응답률을 살펴보면 중증과 경증의 무응답률은 거의 비슷하게 나타남을 알 수 있다.

<표 2-78> 장애정도별 종단면 무응답 현황

구분	무응답		응답		계
	빈도	행%	빈도	행%	
중증	39	3.1%	1,200	96.9%	1,239
경증	82	3.2%	2,480	96.8%	2,562
계	121	3.2%	3,680	96.8%	3,801

참고:  $\chi^2 = 0.01$ , P-value=0.940, Cramér's V=0.003

## 2) 응답성향모형 적합

앞서 인구사회학적 특성과 장애유형 및 장애정도 특성별 무응답의 연관성 분석을 토대로 5차년도 종단 가중치를 산출하기 위해서는 조사대상자들의 다양한 속성(주로 인구사회학적 특성)을 반영한 응답성향모형을 적합하였다. 응답성향모형은 5차년도 응답여부는 과거에 관측된 보조변수들로 설명할 수 있다는 임의무응답(missing at random, 이하 MAR)을 가정한다. 따라서 임의무응답의 가정 하에서는 2차 조사부터 5차 조사에서 모두 응답한 패널의 응답성향모형은 4차년도의 주요 변수들로 설명할 수 있게 된다.

응답성향모형은 로지스틱 회귀모형을 사용하였다. 추정된 모형을 통해 조사대상자 각각의 응답성향점수(혹은 응답확률)를 계산하고 그 역수를 통해 무응답에 대한 보정계수로 산출하여 4차년도 종단 가중치에



곱한 후 이를 5차년도 종단가중치를 계산하였다. 즉, 5차 조사의 표본응답자가 응답했는지 여부를 반응 변수로 하고, 4차 조사의 표본응답자의 특성변수를 설명변수로 하는 로지스틱회귀모형을 적합하여 응답률의 차이를 파악하고 이를 통해 보정하였다. 응답확률은 로지스틱회귀모형 하에서 다음과 같이 표현할 수 있다.

$$\hat{p}_i = \frac{\exp\left(\sum_{j=0}^k \hat{\beta}_j x_{ij}\right)}{1 + \exp\left(\sum_{j=0}^k \hat{\beta}_j x_{ij}\right)}$$

여기서  $\hat{p}_i$ 는  $i$ 번째 패널개체에 대한 응답확률의 추정값이고  $x_{ij}$ 는  $i$ 번째 패널개체의  $j$ 번째 보조변수 특성이며  $\hat{\beta}_j$ 는 회귀계수를 각각 나타낸다.

### 3) 최종 종단 가중치 산출

5차년도 종단면 응답자  $i$ 에 대한 종단 가중치는 4차년도 종단 가중치에 앞서 계산한 응답확률  $\hat{p}_i$ 를 나눠 다음과 같이 계산하였다.

$$\text{5차년도 종단가중치} = \text{4차년도 종단가중치} \times \frac{1}{\hat{p}_i}$$

여기서 사용된 설명변수는 성별, 연령, 지역 등의 변수와 장애유형, 장애정도, 경제활동상태 등의 변수들이다.

<표 2-79>는 세 가지 로지스틱 회귀모형에 대한 적합결과를 정리하여 주고 있다. 앞서 고려한 모든 변수를 사용한 모형1보다는 연령, 지역, 장애유형을 설명변수로 사용한 모형3이 적합도(AIC) 측면에서 더 좋을 수 있다.

<표 2-79> 모형별 종단면 로지스틱 회귀분석 결과

설명변수	모형1		모형2		모형3(최종모형)		
	자유도	$\chi^2$	유의 확률	$\chi^2$	유의 확률	$\chi^2$	유의 확률
성별	1	0.424	0.515				
연령대	3	8.438	0.038	8.638	0.035	8.638	0.035
지역	16	36.509	0.002	36.381	0.003	36.381	0.003
경제활동상태	3	1.064	0.786				
장애유형	5	13.485	0.019	12.749	0.026	12.749	0.026
장애정도	1	0.529	0.467	0.883	0.347		
		AIC = 1071.88		AIC = 1065.69		AIC = 1064.57	

연령, 지역, 장애유형, 장애정도를 고려한 모형2를 기반으로 산출한 중단 가중치와 최종모형을 기반으로 산출한 중단 가중치를 비교해보면 두 가중치는 매우 유사한 값을 지님을 알 수 있다. 또한 두 가중치의 상관관계수는 0.99 이상으로 나타난다. 장애유형과 장애정도는 중단 가중치 산출에 있어서 유의미한 차이를 만들지 않는다고 판단된다. 따라서 응답성향모형의 적합 과정에서 고려한 모든 설명변수를 사용한 포화 모형을 사용하기 보다는 최종 모형으로 연령대와 지역, 장애유형을 사용한 응답성향 모형을 사용한다.

〈표 2-80〉은 5차년도 중단 가중치에 대한 기술 통계량을 보여주고 있다. 중단면 응답자 총 3,680명의 중단 가중치 평균은 371.1이고 최솟값 4.3에서 최댓값 5,365.7의 범위를 갖는다.

〈표 2-80〉 5차년도 중단 가중치에 대한 통계량

통계량			
응답수	3,680	합계	1,365,613
평균	371.1	표준편차	491
중위수	203.7	분산	241102.3
사분위 범위	365.2	범위	5361.4
최솟값	4.3	최댓값	5365.7

〈표 2-81〉은 1~5차 조사까지의 중단 가중치 분포의 현황을 정리하고 있다. 1차년도 중단 가중치에 비하여 5차년도의 중단 가중치는 전반적으로 증가하였지만, 변동계수(CV)는 1차년도에 비하여 다소 줄어들었음을 알 수 있다.

〈표 2-81〉 1-5차 조사의 중단 가중치 분포

통계량	1차년도 최종가중치	2차년도 중단 가중치	3차년도 중단 가중치	4차년도 중단 가중치	5차년도 중단 가중치
최대값	4,487.5	4,768.5	5,032.7	5,186.0	5,365.7
99%	2,129.5	2,291.8	2,387.3	2,467.8	2,518.6
95%	1,054.3	1,140.2	1,189.2	1,250.0	1,289.0
90%	722.5	785.9	817.9	847.2	874.7
75%	363.2	402.4	422.0	440.0	453.5
50%	160.2	176.8	186.9	196.2	203.7
25%	68.0	74.8	79.3	84.3	88.3
10%	30.5	32.9	35.4	37.5	39.6
5%	19.8	21.8	23.3	25.2	26.6
1%	9.5	10.1	11.0	11.9	12.8
최소값	3.2	3.7	3.9	4.1	4.3
평균	303.2	329.9	346.3	360.1	371.1
CV	139.3%	137.1%	135.4%	134.0%	132.3%

## 나. 횡단 가중치 산출

### 1) 단위무응답 분석

5차년도 횡단 가중치는 5차에 응답한 3,907명을 대상으로 산출한다. 가중치 산출을 위하여 5차년도 응답자 3,907명과 무응답자는 573명을 대상으로 무응답 특성을 분석한다. 5차년도 응답여부와 5차년도 응답자 모두가 응답한 1차년도의 주요 특성 변수들 간의 연관성을 살펴봄으로써 응답여부에 중요하게 영향을 미치는 특성 변수를 파악하고자 하였다.

〈표 2-82〉은 시도별 무응답률 현황을 보여주고 있다. 무응답률의 시도 차이는 유의미하며, 서울, 대구, 인천, 광주, 울산, 경기, 강원, 충북, 경북, 제주 지역에서 무응답률이 약 10~21%로 높게 나타나며, 부산, 대전, 충남, 전북, 전남, 경남 지역에서 무응답 비율이 약 10% 미만으로 상대적으로 낮았다.

〈표 2-82 시도별 횡단면 무응답 현황〉

시도	무응답		응답		계
	빈도	행%	빈도	행%	
서울	129	14.8%	741	85.2%	870
부산	27	6.5%	389	93.5%	416
대구	35	13.8%	219	86.2%	254
인천	42	18.0%	191	82.0%	233
광주	13	10.3%	113	89.7%	126
대전	14	8.0%	162	92.0%	176
울산	4	11.8%	30	88.2%	34
경기	178	17.1%	862	82.9%	1,040
강원	25	12.6%	174	87.4%	199
충북	19	11.0%	154	89.0%	173
충남	13	8.1%	148	91.9%	161
전북	14	8.2%	157	91.8%	171
전남	16	7.8%	190	92.2%	206
경북	18	15.3%	100	84.7%	118
경남	12	5.0%	226	95.0%	238
제주	14	21.5%	51	78.5%	65
계	573	12.8%	3,907	87.2%	4,480

참고:  $\chi^2 = 75.388$ , P-value=0.000, Cramér's V=0.130

<표 2-83>은 성별 무응답률의 현황이다. 무응답률의 성별 차이는 유의미하지 않으며, 남성의 무응답률이 다소 높게 나타난다.

<표 2-83> 성별 횡단면 무응답 현황

구분	무응답		응답		계
	빈도	행%	빈도	행%	
남성	391	13.3%	2,549	86.7%	2,940
여성	182	11.8%	1,358	88.2%	1,540
계	573	12.8%	3,907	87.2%	4,480

참조:  $\chi^2 = 1.857$ , P-value=0.173, Cramér's V=0.020

<표 2-84>은 연령대별 무응답률의 현황이다. 응답여부와 연령의 분포는 유의한 관계가 있다. 연령대에 따른 무응답 비율을 살펴보면 15~29세 미만의 연령층의 무응답률이 가장 높으며 50세 이상의 무응답률이 가장 낮았다.

<표 2-84> 연령대별 횡단면 무응답 현황

구분	무응답		응답		계
	빈도	행%	빈도	행%	
15-29세	123	15.6%	665	84.4%	788
30-39세	156	14.3%	935	85.7%	1,091
40-49세	182	14.0%	1,117	86.0%	1,299
50-64세	112	8.6%	1,190	91.4%	1,302
계	573	12.8%	3,907	87.2%	4,480

참조:  $\chi^2 = 30.048$ , P-value=0.000, Cramér's V=0.082

<표 2-85>은 경제활동상태에 따른 무응답 현황이다. 경제활동상태는 응답여부와 유의한 관계가 있는 것으로 나타났다. 실업자의 무응답률이 가장 높았고, 취업미희망자의 무응답률이 가장 낮았다.

<표 2-85> 경제활동상태별 횡단면 무응답 현황

구분	무응답		응답		계
	빈도	행%	빈도	행%	
취업자	298	13.3%	1,950	86.7%	2,248
실업자	27	16.5%	137	83.5%	164
취업희망자	162	14.9%	927	85.1%	1,089
취업미희망자	86	8.8%	893	91.2%	979
계	573	12.8%	3,907	87.2%	4,480

참조:  $\chi^2 = 20.752$ , P-value=0.000, Cramér's V=0.068

〈표 2-86〉은 장애유형별 무응답률의 현황을 보여준다. 무응답률의 장애유형별 차이는 유의미하지 않으며, 모든 장애유형에 대해서 약 11~14%로 비슷한 무응답률을 보이고 있다.

〈표 2-86〉 장애유형별 횡단면 무응답 현황

구분	무응답		응답		계
	빈도	행%	빈도	행%	
지체장애	291	13.1%	1925	86.9%	2216
지체 외 신체 외부장애	40	12.4%	283	87.6%	323
시각장애	69	11.9%	512	88.1%	581
시각외감각장애	58	14.0%	356	86.0%	414
정신적장애	68	11.5%	525	88.5%	593
신체내부장애	47	13.3%	306	86.7%	353
계	573	12.8%	3907	87.2%	4480

참조:  $\chi^2 = 2.284$ , P-value=0.809, Cramér's V=0.023

〈표 2-87〉은 장애정도별 무응답률의 현황이다. 20년도 5차 조사의 중/경증 장애정도는 기존 고용노동부 기준에서 보건복지부 기준으로 변경된 중/경증 장애정도이다. 기준 변경에 관한 추가적인 논의는 가중치 벤치마킹에서 다시 논의한다. 무응답률의 장애정도별 차이는 유의미하지 않으며, 장애정도에 따른 무응답률을 살펴보면 경증의 무응답률이 중증의 무응답률보다 다소 높게 나타난다.

〈표 2-87〉 장애정도별 횡단면 무응답 현황

구분	무응답		응답		계
	빈도	행%	빈도	행%	
중증	171	12.0%	1253	88.0%	1424
경증	402	13.2%	2654	86.8%	3056
계	573	12.8%	3907	87.2%	4480

참조:  $\chi^2 = 1.043$ , P-value=0.307, Cramér's V=0.015

## 2) 무응답 조정 가중치 산출

무응답 보정을 위해서는 조사대상자들의 다양한 속성(주로 인구사회학적 특성)을 반영한 응답성향모형을 적합하였다. 이를 위해 5차년도 횡단면 조사의 응답여부는 과거에 관측된 변수들로 설명할 수 있다는 임의 응답(MAR)을 가정하였고 1차년도 조사 기준의 설명변수는 물론 2차 조사부터 4차 조사까지 응답여부도 함께 다음의 확률을 고려한 로지스틱 모형을 적합하였다.

$$P(5차\ 응답여부 \mid 2, 3, 4차\ 응답여부, 1차\ 설명변수)$$

즉, 반응변수로는 5차조사 응답여부를 나타내는 지시자이고 2차년도부터 4차년도 조사의 응답여부와 1차년도 기준의 성별, 연령, 지역, 장애유형, 장애정도, 경제활동상태 등의 변수를 설명변수로 하는 로지스틱 회귀모형을 다음의 형태로 적합하였다.

$$\hat{p}_i = \frac{\exp\left(\sum_{j=0}^k \hat{\beta}_j x_{ij}\right)}{1 + \exp\left(\sum_{j=0}^k \hat{\beta}_j x_{ij}\right)}$$

<표 2-88>은 5차년도의 응답확률 모형에서 모든 설명변수를 사용한 포화모형을 포함하여 유의하지 않은 설명변수를 하나씩 제거하여 모형을 적합한 결과를 정리하고 있다. 모든 설명변수를 사용하는 모형1보다는 성별, 경제활동상태를 제외한 모형3이 적합도(AIC) 측면에서 더 좋음을 알 수 있다. 따라서 최종적으로 선택한 모형으로는 2차, 3차, 4차년도 응답여부, 연령대, 지역, 장애유형을 설명변수로 사용한 로지스틱 회귀 모형이다.

<표 2-88> 모형별 횡단면 로지스틱 회귀분석 결과

구분	모형1			모형2		모형3(최종모형)	
	df	$\chi^2$	유의 확률	$\chi^2$	유의 확률	$\chi^2$	유의 확률
4차 응답여부	1	1763.7	0.000	1763.7	0.000	1763.7	0.000
3차 응답여부	1	43.6	0.000	43.6	0.000	43.6	0.000
2차 응답여부	1	14.7	0.000	14.7	0.000	14.7	0.000
성별	1	0.0	0.891				
연령대	3	7.4	0.060	7.4	0.059	7.4	0.059
지역	15	31.1	0.008	31.1	0.008	31.1	0.008
경제활동상태	3	5.2	0.160	5.2	0.161		
장애유형	5	12.4	0.029	12.3	0.031	12.1	0.034
		AIC = 1609.88		AIC = 1608.01		AIC = 1607.39	

최종모형으로부터 추정한 응답확률을  $\hat{p}_i$ 로 정의하면 5차년도에 응답한  $i$ 번째 개체의 종단 가중치는 다음과 같이 계산된다.

$$5\text{차년도 무응답 보정 가중치} = 1\text{차년도 최종가중치} \times \frac{1}{\hat{p}_i}$$

### 3) 벤치마킹 보정

5차년도 횡단 가중치 산출을 위해 무응답을 보정한 가중치에 대하여 모집단 분포에 대한 벤치마킹 보정을 사용한다. 성별, 연령, 지역, 장애유형, 장애정도, 경제활동상태 변수를 통하여 이들의 모집단 분포를 이용한 가중치 합 비율을 통해 벤치마킹 보정을 실시한다. 성, 연령, 지역, 장애유형, 장애정도에 관한 모집단 정보는 2019년 12월 등록기준 19세에서 68세까지의 보건복지부 장애인 등록현황을 사용하였다.

경제활동상태에 대한 모집단 정보는 2019년 「장애인경제활동실태조사」의 경제활동상태에 대한 추정 결과를 사용하였다. 이 추정 결과의 모집단 수와 2018년 12월 기준 장애인 등록현황의 모집단 수가 일치하지 않기 때문에 추정 결과를 직접 사용할 수는 없다. 이를 위해 2019년 「장애인경제활동실태조사」의 경제활동상태에 대한 비율 추정 결과를 2018년 12월 기준 장애인 등록현황의 모집단 수에 곱하여 경제활동상태에 대한 모집단 분포를 정의하였다.

레이킹-비 조정 방법을 적용하여 가중치를 보정한 후 횡단 가중치를 다음과 같이 산출하였다.

$$w_{ij}^{rak} = w_{ij} \exp(\lambda' x_{ij}).$$

여기서  $x_{ij}$ 는 레이킹 비 조정 방법을 위해 사용된 보조변수들을 나타내는 지시변수 벡터이며,  $\lambda$ 는  $\sum w_{ij}^{rak} x_{ij} = t_x$ 의 해이며,  $t_x$ 는 각 레이킹 비 조정 방법에 사용된 벤치마킹 변수들의 범주에 해당하는 모집단 장애인 분포를 나타낸다.

#### 4) 장애정도 기준 변경

벤치마킹 보정에 사용되는 중/경증 장애정도는 2019년 4차조사까지 고용노동부 기준으로 구분되어 모집단 정보가 제공되었다. 2020년 부터는 고용노동부 정보를 사용할 수 없고 보건복지부 새로운 기준으로 장애정도가 제공된다. 보건복지부 장애정도는 기존 장애등급 1-3급은 심한 장애(고용노동부 기준 중증), 기존 장애등급 4-6급은 심하지 않은 장애(고용노동부 기준 경증)로 구분한다.

이와 같은 상황에서 장애정도와 장애등급의 모집단 분포의 변화를 살펴 보고자 2015년부터 2018년까지 고용노동부 기준 장애정도 모집단 분포 현황을 분석하였고 <표 2-89>에 그 결과가 나타나 있다. <표 2-90>에는 장애등급 모집단 분포 현황분석 결과가 나타나 있다. 또한 <표 2-91>에는 장애등급을 사용하여 보건복지부 기준의 중증(심한 장애)와 경증(심하지 않은 장애)를 구분하여 정리한 2015년부터 2018년까지 모집단 분포 현황이 나타나 있다.

<표 2-89> 1차-4차년도 고용노동부 기준 장애정도 모집단 현황

장애 정도	2015년		2016년		2017년		2018년	
	빈도	비율	빈도	비율	빈도	비율	빈도	비율
중증	528,223	38.1%	537,272	38.1%	549,943	38.2%	514,522	35.7%
경증	859,579	61.9%	874,541	61.9%	888,661	61.8%	927,781	64.3%

<표 2-90> 1차-4차년도 장애등급 모집단 분포 현황

장애 등급	2015년 기준		2016년 기준		2017년 기준		2018년 기준	
1급	126,847	9.1%	127,787	9.1%	128,933	9.0%	120,109	8.3%
2급	206,750	14.9%	210,365	14.9%	214,000	14.9%	197,045	13.7%
3급	268,845	19.4%	273,479	19.4%	278,139	19.3%	263,825	18.3%
4급	152,783	11.0%	155,190	11.0%	158,302	11.0%	166,983	11.6%
5급	248,722	17.9%	254,218	18.0%	260,901	18.1%	271,834	18.8%
6급	383,855	27.7%	390,774	27.7%	398,329	27.7%	422,506	29.3%

〈표 2-91〉 1차-4차년도 보건복지부 기준 장애정도 모집단 현황

장애 정도	2015년		2016년		2017년		2018년	
	빈도	비율	빈도	비율	빈도	비율	빈도	비율
중증	602,442	43.4%	611,631	43.4%	621,072	43.2%	580,979	40.3%
경증	785,360	56.6%	800,182	56.7%	817,532	56.8%	861,323	59.7%

\* 중증: 심한 장애, 경증: 심하지 않은 장애

고용노동부 기준 장애정도 모집단 현황을 살펴보면 2015년부터 2017년까지 중증 비율이 약 38%대였으나 2018년에는 약 35%로 감소하였다. 중증 장애등급인 1, 2급이 2015년부터 2017년에 비하여 2018년에 감소하였고 경증 장애등급인 4, 5, 6급이 증가하였음을 알 수 있다. 보건복지부 기준 장애정도 모집단 현황을 살펴보면 2015년부터 2017년까지 중증 비율이 약 43%대였으나 2018년에는 약 40%로 감소하였다. 보건복지부 기준 중증 비율은 2015년부터 2018년까지 고용노동부 기준과 비교하였을 때 약 5% 정도 크게 나타남을 알 수 있다.

〈표 2-92〉에는 5차년도 조사 모집단에 대한 보건복지부 기준과 고용노동부 기준 장애정도 분포 현황이 나타나 있다. 보건복지부 기준 중증 비율은 43%, 고용노동부 기준 중증 비율은 38%로 이전 모집단 분포와 유사하게 분포하고 있음을 알 수 있다. 두 기준 간의 차이는 약 5%로 이전 연도의 차이들과 유사하게 나타난다.

〈표 2-92〉 5차년도 장애정도 모집단 현황(2019년)

장애정도	보건복지부 기준		고용노동부 기준	
	빈도	비율	빈도	비율
중증	641,064	43.0%	568,863	38.1%
경증	850,308	57.0%	921,847	61.8%

보건복지부 기준 장애정도와 고용노동부 기준 장애정도의 차이가 발생하는 이유는 장애등급 3급이 고용노동부 기준으로는 중증과 경증으로 구분되는 반면 보건복지부 기준으로는 모두 중증으로 구분되기 때문이다. 따라서 장애등급 3급 장애인의 특성을 장애유형, 성별, 연령대, 경제활동상태등의 주요 변수의 분포를 통하여 전체 장애인과 비교 분석한다. 분석에 사용한 자료는 4차년도 조사 데이터이며 가중치를 사용하여 가중 빈도와 비율을 구하였다. 〈표 2-93〉 - 〈표 2-96〉에는 전체 장애인과 3급 장애인에 대하여 장애유형, 성별, 연령대, 경제활동 상태의 분포가 나타나 있다.

〈표 2-93〉 전체 장애인과 3급 장애인의 장애유형 분포

장애유형	전체 장애인		3급 장애인	
	가중빈도	비율	가중빈도	비율
지체장애	710,940	49%	76,748	29%
지체 외 신체외부장애	123,128	9%	24,453	9%
시각장애	137,511	10%	1,650	1%
시각 외 감각장애	110,037	8%	24,046	9%
정신적장애	272,151	19%	135,746	50%
신체내부장애	88,536	6%	6,481	2%



〈표 2-94〉 전체 장애인과 3급 장애인의 성별 분포

성별	전체 장애인		3급 장애인	
	가중빈도	비율	가중빈도	비율
남성	969,762	67%	183,159	68%
여성	472,541	33%	85,966	32%

〈표 2-95〉 전체 장애인과 3급 장애인의 연령 분포

연령	전체 장애인		3급 장애인	
	가중빈도	비율	가중빈도	비율
18-29세	108,447	8%	34,700	13%
30-39세	131,942	9%	25,194	9%
40-49세	257,989	18%	53,689	20%
50-67세	943,925	65%	155,542	58%

〈표 2-96〉 전체 장애인과 3급 장애인의 경제활동상태 분포

경제활동상태	전체 장애인		3급 장애인	
	가중빈도	비율	가중빈도	비율
취업자	709,072	49%	114,825	43%
실업자	48,239	3%	9,821	4%
취업 희망자	134,056	9%	24,761	9%
취업 미희망자	550,935	38%	119,717	44%

3급 장애인의 장애유형은 전체 장애인과 비교하였을 때 지체장애와 시각장애의 비율은 적고 정신적 장애의 비율은 많이 분포하고 있음을 알 수 있다. 전체 장애인과 3급 장애인의 성, 연령, 경제활동 상태의 비율은 전체적으로 유사하게 분포하고 있음을 알 수 있다. 3급 장애인은 장애유형에서 전체 장애인과 분포가 다르게 나타나기는 하지만 벤치마킹에 사용되는 다른 주요 변수들에서 전체 장애인과 유사하게 분포하고 있으므로 추후 3급 장애인으로 인하여 생기는 장애정도 비율의 변화는 벤치마킹을 통한 횡단가중치 산출에 있어 기존 방법과 큰 차이를 보이지 않을 것으로 판단된다.

다음 절에서 횡단 가중치 극단값 조정을 통한 최종가중치를 사용한 표본 자료 분포와 모집단 분포 비교 결과를 살펴보기로 한다.

### 5) 극단값 처리 및 최종가중치 산출

앞서 산출한 레이킹 비 조정 횡단가중치에는 설계 가중치의 최댓값이었던 약 4,500을 넘는 가중치들이 존재하였다. 이들은 조사변수에 곱해져 합산되므로 경우에 따라서 통계량에 과도한 영향력을 갖게 된다. 따라서 이러한 과도한 가중치 값을 추가로 조정해 주기 위해 4,500보다 작은 최댓값인 4,402으로 절사하여 조정해 준 다음, 앞서 고려한 바와 같이 성, 연령, 장애유형, 장애정도의 보조변수들에 대해서 다시 한번 레이킹 비 보정을 실시하여 이를 최종가중치로 산출하였다.

<표 2-97>는 횡단면 최종가중치에 대한 기술통계량을 나타내고 있다. 평균 381.7이며 최솟값은 2.9이고 최댓값은 4,644.4이다.

<표 2-97> 횡단 가중치에 대한 통계량

통계량			
응답수	3,907	합계	1,491,378
평균	381.7	표준편차	562.7
중위수	188.2	분산	316687.4
사분위 범위	374.5	범위	4641.5
최솟값	2.9	최댓값	4644.4

벤치마킹 보정을 위해 고려하였던 보조변수인 성별, 연령대, 지역, 장애유형, 장애정도, 경제활동상태 변수의 모집단 총합과 표본가중합의 분포는 <표 2-98>로부터 <표 2-103>에 각각 정리하고 있다. 횡단면 최종가중치를 사용한 표본가중합이 모집단 총합과 거의 차이가 없음을 알 수 있다.

<표 2-98> 성별 모집단 및 표본 분포

성	표본가중합		모집단 총합	
	인구	비중	인구	비중
남자	978,076	65.6%	978,077	65.6%
여자	513,302	34.4%	513,301	34.4%
총합	1,491,378	100.0%	1,491,378	100%

<표 2-99> 연령별 모집단 및 표본 분포

연령	표본가중합		모집단 총합	
	인구	비중	인구	비중
18~29세	103,751	7.0%	103,749	7.0%
30~39세	129,056	8.7%	129,055	8.7%
40~49세	253,470	17.0%	253,470	17.0%
50~67세	1,005,101	67.4%	1,005,104	67.4%
총합	1,491,378	100.0%	1,491,378	100.0%

<표 2-100> 경제활동상태 변수의 모집단 및 표본 분포

경제활동상태	표본가중합		모집단 총합	
	인구	비중	인구	비중
취업자	708,854	47.5%	708,854	47.5%
미취업자	45,406	3.0%	45,406	3.0%
취업희망자	159,777	10.7%	159,777	10.7%
취업미희망자	577,342	38.7%	577,342	38.7%
총합	1,491,378	100.0%	1,491,378	100.0%

<표 2-101> 장애유형별 변수의 모집단 및 표본 분포

장애유형	표본가중합		모집단 총합	
	인구	비중	인구	비중
지체장애	714,375	47.9%	714,376	47.9%
지체 외 신체 외부장애	129,926	8.7%	129,926	8.7%
시각장애	148,312	9.9%	148,312	9.9%
시각 외 감각장애	126,951	8.5%	126,951	8.5%
정신적 장애	271,296	18.2%	271,295	18.2%
신체 내부장애	100,518	6.7%	100,518	6.7%
총합	1,491,378	100.0%	1,491,378	100.0%

<표 2-102> 시도별 모집단 및 표본 분포

지역	표본가중합		모집단 총합	
	인구	비중	인구	비중
서울	223,928	15.0%	227,039	15.0%
부산	96,221	6.5%	95,338	6.5%
대구	75,033	5.0%	74,278	5.0%
인천	90,551	6.1%	89,229	6.1%
광주	40,724	2.7%	40,334	2.7%
대전	51,109	3.4%	50,406	3.4%
울산	31,749	2.1%	32,399	2.1%
세종	7,507	0.5%	7,319	0.5%
경기	322,687	21.6%	329,718	21.6%
강원	58,951	4.0%	58,079	4.0%
충북	53,630	3.6%	53,086	3.6%
충남	73,813	4.9%	72,878	4.9%
전북	69,201	4.6%	68,577	4.6%
전남	68,693	4.6%	67,769	4.6%
경북	96,281	6.5%	95,037	6.5%
경남	111,009	7.4%	109,807	7.4%
제주	20,291	1.4%	20,085	1.4%
총합	1,491,378	100.0%	1,491,378	100.0%

<표 2-103> 장애정도별 변수의 모집단 및 표본 분포

장애정도	표본가중합		모집단 총합	
	인구	비중	인구	비중
중증	641,066	43.0%	641,064	43.0%
경증	850,312	57.0%	850,308	57.0%
총합	1,491,378	100.0%	1,491,378	100.0%

〈표 2-104〉는 1~5차 조사들의 횡단 가중치의 기술통계량 및 분포를 비교하고 있다. 1차년도 최종 가중치에 비하여 5차년도의 횡단 가중치는 전반적으로 증가하였지만, 극단값 조정을 통하여 심각한 특이 가중치는 없는 것으로 판단되고 변동계수(CV)도 1차년도에 비하여 약 8% 정도 증가하였다.

〈표 2-104〉 1-5차년도 횡단가중치 분포

	1차년도 최종가중치	2차년도 횡단가중치	3차년도 횡단가중치	4차년도 횡단가중치	5차년도 횡단가중치
최대값	4,487.5	4,756.8	4,941.0	4,536.3	4,644.4
99%	2,129.5	2,260.2	2,396.5	2,631.4	3,040.7
95%	1,054.3	1,166.0	1,220.5	1,264.9	1,363.2
90%	722.5	772.3	849.5	862.7	909.2
75%	363.2	405.6	413.4	428.0	453.7
50%	160.2	179.8	179.1	184.4	188.2
25%	68.0	78.4	77.9	77.3	79.3
10%	30.5	34.2	35.3	33.8	33.7
5%	19.8	22.4	22.8	22.8	22.7
1%	9.5	10.6	10.9	10.7	11.0
최소값	3.2	3.4	3.0	3.2	2.9
평균	303.2	335.0	350.5	361.0	381.7
CV	139.3%	138.0%	142.2%	141.6%	147.4%

해당 가중치를 사용하여 주요 변수인 장애유형, 경제활동상태에 대한 빈도 및 비율과 상대표준오차(CV) 값을 살펴본 결과는 〈표 2-105〉 및 〈표 2-106〉에 제시하였다.

〈표 2-105〉 장애유형 (15개) 별 경제활동상태(2개 구분) 빈도/비율 추정 및 상대표준오차 분석

장애유형	경제활동상태 (2개 구분)	가중빈도	비중(%)	표준오차(%)	상대표준오차 (%)
지체장애	취업자	442,641	62.3	3.0	4.8
	미취업자	268,299	37.7	2.3	6.1
	합계	710,940	100.0		
뇌병변장애	취업자	31,213	26.9	5.3	19.6
	미취업자	84,689	73.1	9.4	12.9
	합계	115,902	100.0		

<표 2-106> 장애유형 (15개) 별 경제활동상태(2개 구분) 빈도/비율 추정 및 상대표준오차 분석(계속)

시각장애	취업자	70,084	51.0	4.2	8.2
	미취업자	67,427	49.0	5.5	11.3
	합계	137,511	100.0		
청각장애	취업자	53,069	54.1	6.9	12.7
	미취업자	45,065	45.9	5.9	12.8
	합계	98,134	100.0		
언어장애	취업자	4,789	40.2	16.6	41.3
	미취업자	7,113	59.8	18.8	31.4
	합계	11,903	100.0		
지적(정신지체)장애	취업자	46,092	25.5	3.9	15.1
	미취업자	134,474	74.5	6.9	9.3
	합계	180,566	100.0		
자폐성(발달)장애	취업자	3,183	30.3	10.8	35.6
	미취업자	7,336	69.7	21.3	30.6
	합계	10,519	100.0		
정신장애	취업자	12,672	15.6	4.7	30.3
	미취업자	68,394	84.4	11.7	13.9
	합계	81,066	100.0		
신장장애	취업자	20,342	39.1	10.0	25.4
	미취업자	31,636	60.9	9.0	14.7
	합계	51,978	100.0		
심장장애	취업자	3,342	76.7	51.6	67.2
	미취업자	1,016	23.3	8.9	38.2
	합계	4,357	100.0		
호흡기장애	취업자	1,373	41.4	18.7	45.2
	미취업자	1,941	58.6	23.7	40.5
	합계	3,314	100.0		
간장애	취업자	9,443	57.9	21.3	36.8
	미취업자	6,880	42.1	26.8	63.6
	합계	16,323	100.0		
안면장애	취업자	6,841	94.7	54.4	57.5
	미취업자	385	5.3	3.3	61.3
	합계	7,226	100.0		
장루요루장애	취업자	1,307	26.9	15.2	56.6
	미취업자	3,546	73.1	31.8	43.6
	합계	4,854	100.0		
뇌전증(간질)장애	취업자	2,683	34.8	14.8	42.5
	미취업자	5,027	65.2	19.5	29.9
	합계	7,710	100.0		

## 8. 6차조사 가중치 산출 및 추정

### 가. 종단 가중치 산출

#### 1) 차수별 응답 현황 및 단위 무응답 분석

6차년도 종단 가중치는 1차부터 6차에 모두 응답한 3,570명을 대상으로 산출한다. 6차조사 종단 가중치의 산출에 앞서 단위무응답의 현황을 살펴보고 응답성향분석을 통해 단위무응답에 영향을 미치는 조사대상자 특성을 우선적으로 파악하고자 하였다.

〈표 2-107〉은 장애인 거주 시도별 응답현황을 정리하였고 더불어 시도 분류와 응답여부와의 독립성 검정을 위한 카이제곱 통계와 크래머 V 통계를 산출하였다. 시도간 무응답률의 편차는 유의미하게 나타났다. 특히 제주 지역에서 무응답률이 약 6.5%로 높게 나왔고, 그 외 지역에서 무응답률은 약 5% 미만으로 상대적으로 낮게 나타났다.

〈표 2-107〉 시도별 종단면 무응답 현황

시도	무응답		응답		계
	빈도	행%	빈도	행%	
서울	23	3.6%	624	96.4%	647
부산	2	0.5%	364	99.5%	366
대구	1	0.5%	202	99.5%	203
인천	4	2.3%	168	97.7%	172
광주	1	0.9%	106	99.1%	107
대전	2	1.2%	162	98.8%	164
울산	1	3.2%	30	96.8%	31
세종	0	0.0%	6	100.0%	6
경기	32	4.0%	760	96.0%	792
강원	1	0.6%	169	99.4%	170
충북	1	0.7%	149	99.3%	150
충남	2	1.4%	138	98.6%	140
전북	1	0.7%	150	99.3%	151
전남	3	1.7%	175	98.3%	178
경북	1	1.0%	96	99.0%	97
경남	1	0.4%	228	99.6%	229
제주	3	6.5%	43	93.5%	46
계	79	2.2%	3,570	97.8%	3,649

참고:  $\chi^2 = 41.72$ , P-value=0.000, Cramér's V=0.107

〈표 2-108〉은 장애인의 성별 응답현황을 보여주고 있다. 남성과 여성 간의 무응답률은 서로 유사하여 유의한 차이가 아님을 알 수 있다.

〈표 2-108〉 성별 종단면 무응답 현황

구분	무응답		응답		계
	빈도	행%	빈도	행%	
남성	51	2.2%	2,316	97.8%	2,367
여성	28	2.2%	1,254	97.8%	1,282
계	79	2.2%	3,570	97.8%	3,649

참고:  $\chi^2 = 0.00$ , P-value=0.999, Cramér's V=0.000

〈표 2-109〉은 장애인의 연령대별 응답률 현황을 나타내고 있다. 연령대에 따른 무응답률의 편차는 유의한 차이가 아닌 것으로 나타난다. 30대 미만의 무응답 비율이 다소 높으나 전 연령층에서 무응답 비율은 2~3%로 유사하게 나타난다.

〈표 2-109〉 연령대별 종단면 무응답 현황

구분	무응답		응답		계
	빈도	행%	빈도	행%	
18-29세	12	2.8%	410	97.2%	422
30-39세	14	2.3%	594	97.7%	608
40-49세	19	1.9%	988	98.1%	1,007
50-68세	34	2.1%	1,578	97.9%	1,612
계	79	2.2%	3,570	97.8%	3,649

참고:  $\chi^2 = 1.36$ , P-value=0.714, Cramér's V=0.019

〈표 2-110〉은 장애인의 경제활동상태별 응답률의 현황을 정리하고 있다. 경제활동상태에 따른 무응답률의 편차는 유의하지 않은 것으로 나타났다. 취업자의 무응답률이 가장 높은 반면, 취업 희망자의 무응답률이 가장 낮았다.

〈표 2-110〉 경제활동상태별 종단면 무응답 현황

구분	무응답		응답		계
	빈도	행%	빈도	행%	
취업자	43	2.3%	1,813	97.7%	1,856
실업자	1	1.7%	58	98.3%	59
취업희망자	3	1.1%	272	98.9%	275
취업미희망자	32	2.2%	1,427	97.8%	1,459
계	79	2.2%	3,570	97.8%	3,649

참고:  $\chi^2 = 1.77$ , P-value=0.622, Cramér's V=0.022

<표 2-111>은 장애인의 장애유형별 응답률의 현황을 보여준다. 장애유형에 따른 무응답률의 편차는 유의미하지 않게 나타났다. 시각외감각장애의 무응답률이 가장 높고 신체내부장애의 무응답률이 가장 작게 나타났다. 그 외 장애유형의 무응답률은 비슷하게 나타났다.

<표 2-111> 장애유형별 종단면 무응답 현황

구분	무응답		응답		계
	빈도	행%	빈도	행%	
지체장애	34	1.9%	1,763	98.1%	1,797
지체 외 신체 외부장애	5	1.8%	266	98.2%	271
시각장애	11	2.3%	464	97.7%	475
시각외감각장애	13	3.9%	320	96.1%	333
정신적장애	12	2.4%	483	97.6%	495
신체내부장애	4	1.4%	274	98.6%	278
계	79	2.2%	3,570	97.8%	3,649

참고:  $\chi^2 = 6.42$ , P-value=0.268, Cramér's V=0.042

<표 2-112>은 장애인의 장애정도별 응답률 현황이다. 장애정도에 따른 무응답률의 편차는 유의하지 않은 것으로 나타났다. 장애정도에 따른 무응답률을 살펴보면 중증과 경증의 무응답률은 거의 비슷하게 나타남을 알 수 있다.

<표 2-112> 장애정도별 종단면 무응답 현황

구분	무응답		응답		계
	빈도	행%	빈도	행%	
중증	35	2.4%	1,397	97.6%	1,432
경증	44	2.0%	2,173	98.0%	2,217
계	79	2.2%	3,570	97.8%	3,649

참고:  $\chi^2 = 0.66$ , P-value=0.415, Cramér's V=0.013

## 2) 응답성향모형 적합

앞서 인구사회학적 특성과 장애유형 및 장애정도 특성별 무응답의 연관성 분석을 토대로 6차년도 종단 가중치를 산출하기 위해서는 조사대상자들의 다양한 속성(주로 인구사회학적 특성)을 반영한 응답성향모형을 적합하였다. 응답성향모형은 6차년도 응답여부는 과거에 관측된 보조변수들로 설명할 수 있다는 임의무응답(missing at random, 이하 MAR)을 가정한다. 따라서 임의무응답의 가정 하에서는 2차 조사부터 6차조사에서 모두 응답한 패널의 응답성향모형은 5차년도의 주요 변수들로 설명할 수 있게 된다.

응답성향모형은 로지스틱 회귀모형을 사용하였다. 추정된 모형을 통해 조사대상자 각각의 응답성향점수(혹은 응답확률)를 계산하고 그 역수를 통해 무응답에 대한 보정계수로 산출하여 5차년도 종단 가중치에 곱한 후 이를 6차년도 종단가중치를 계산하였다. 즉, 6차조사의 표본응답자가 응답했는지 여부를 반응변수로



하고, 5차조사의 표본응답자의 특성변수를 설명변수로 하는 로지스틱회귀모형을 적합하여 응답률의 차이를 파악하고 이를 통해 보정하였다. 응답확률은 로지스틱회귀모형 하에서 다음과 같이 표현할 수 있다.

$$\hat{p}_i = \frac{\exp\left(\sum_{j=0}^k \hat{\beta}_j x_{ij}\right)}{1 + \exp\left(\sum_{j=0}^k \hat{\beta}_j x_{ij}\right)}$$

여기서  $\hat{p}_i$ 는  $i$ 번째 패널개체에 대한 응답확률의 추정값이고  $x_{ij}$ 는  $i$ 번째 패널개체의  $j$ 번째 보조변수 특성이며  $\hat{\beta}_j$ 는 회귀계수를 각각 나타낸다.

### 3) 최종 종단 가중치 산출

6차년도 종단면 응답자  $i$ 에 대한 종단 가중치는 5차년도 종단 가중치에 앞서 계산한 응답확률  $\hat{p}_i$ 를 나눠 다음과 같이 계산하였다.

$$\text{6차년도 종단가중치} = \text{5차년도 종단가중치} \times \frac{1}{\hat{p}_i}$$

여기서 사용된 설명변수는 성별, 연령, 지역 등의 변수와 장애유형, 장애정도, 경제활동상태 등의 변수들이다.

로지스틱 회귀모형에 대하여 초기 모형과 최종 모형에 대한 적합결과를 정리하여 <표 3-12>에 제시하였다. 앞서 고려한 모든 변수를 사용한 초기 모형보다는 지역을 설명변수로 사용한 최종 모형이 적합도(AIC) 측면에서 더 좋음을 알 수 있다.

지역 외 연령, 장애유형, 장애정도를 설명변수로 고려한 회귀모형을 고려할 수 있으나 이 모형을 기반으로 산출한 종단가중치와 최종모형을 기반으로 산출한 종단 가중치를 비교해보면 두 가중치는 매우 유사한 값을 지니며 두 가중치의 상관계수는 0.99 이상으로 나타난다. 따라서 장애유형과 장애정도는 종단가중치 산출에 있어서 유의미한 차이를 만들지 않는다고 판단된다. 응답성향모형의 적합과정에서 고려한 모든 설명변수를 사용한 포화모형을 사용하기보다는 최종 모형으로 연령대와 지역, 장애유형을 사용한 응답성향 모형을 사용한다. 6차조사의 응답 성향 모형에서 설명변수가 지역 하나로 나오는 것은 다른 차수 조사에 비하여 무응답자 수와 비율이 현저히 작아서 불균등 자료에서 발생하는 특별한 상황이라고 생각된다. 추후 7차 이후의 조사에서도 이러한 현상이 계속 발생된다면 로지스틱 회귀와 같은 분포 가정하의 응답성향 모형 보다는 비모수적 방법을 고민할 필요가 있다.

<표 2-113>은 6차년도 종단 가중치에 대한 기술 통계량을 보여주고 있다. 종단면 응답자 총 3,570명의 종단 가중치 평균은 378.7이고 최솟값 4.3에서 최댓값 5,421.6의 범위를 갖는다.

<표 2-114>는 1~6차조사까지의 종단 가중치 분포의 현황을 정리하고 있다. 1차년도 종단 가중치에 비하여 6차년도의 종단 가중치는 전반적으로 증가하였지만 변동계수(CV)는 1차년도에 비하여 다소 줄어들었음을 알 수 있다.

〈표 2-113〉 모형별 종단면 로지스틱 회귀분석 결과

설명변수	포화모형			최종모형	
	자유도	$\chi^2$	유의 확률	$\chi^2$	유의 확률
성별	1	0.003	0.953		
연령대	3	1.296	0.730		
지역	16	43.322	0.000	44.009	0.000
경제활동상태	3	2.047	0.563		
장애유형	5	5.163	0.396		
장애정도	1	0.247	0.619		
		AIC = 769.78		AIC = 751.84	

〈표2-114〉 6차년도 종단 가중치에 대한 통계량

통계량			
응답수	3,570	합계	1,352,062
평균	378.7	표준편차	498
중위수	208.5	분산	248015
사분위 범위	371.8	범위	5417.3
최솟값	4.3	최댓값	5421.6

〈표 2-115〉 2-6차조사의 종단 가중치 분포

7통계량	2차년도 종단가중치	3차년도 종단가중치	4차년도 종단가중치	5차년도 종단가중치	6차년도 종단가중치
최대값	4,768.5	5,032.7	5,186.0	5,365.7	5421.6
99%	2,291.8	2,387.3	2,467.8	2,518.6	2506.4
95%	1,140.2	1,189.2	1,250.0	1,289.0	1314.8
90%	785.9	817.9	847.2	874.7	897.5
75%	402.4	422.0	440.0	453.5	462.0
50%	176.8	186.9	196.2	203.7	208.5
25%	74.8	79.3	84.3	88.3	90.2
10%	32.9	35.4	37.5	39.6	40.6
5%	21.8	23.3	25.2	26.6	26.9
1%	10.1	11.0	11.9	12.8	13.2
최소값	3.7	3.9	4.1	4.3	4.3
평균	329.9	346.3	360.1	371.1	378.7
CV	137.1%	135.4%	134.0%	132.3%	131.5%

## 나. 횡단 가중치 산출

### 1) 단위무응답 분석

6차년도 횡단 가중치는 6차에 응답한 3,848명을 대상으로 산출한다. 가중치 산출을 위하여 5차년도 응답자 3,848명과 무응답자는 599명을 대상으로 무응답 특성을 분석한다. 5차년도 응답여부와 5차년도 응답자 모두가 응답한 1차년도의 주요 특성 변수들 간의 연관성을 살펴봄으로써 응답여부에 중요하게 영향을 미치는 특성 변수를 파악하고자 하였다.

<표 2-116>은 시도별 무응답률 현황을 보여주고 있다. 무응답률의 시도 차이는 유의미하며, 서울, 대구, 인천, 광주, 울산, 경기, 강원, 충북, 경북, 제주 지역에서 무응답률이 약 10~21%로 높게 나타나며, 부산, 대전, 충남, 전북, 전남, 경남 지역에서 무응답 비율이 약 10% 미만으로 상대적으로 낮았다.

<표 2-116> 시도별 횡단면 무응답 현황

시도	무응답		응답		계
	빈도	행%	빈도	행%	
서울	141	16.3%	724	83.7%	865
부산	26	6.4%	382	93.6%	408
대구	32	12.7%	220	87.3%	252
인천	44	19.0%	188	81.0%	232
광주	13	10.4%	112	89.6%	125
대전	13	7.4%	163	92.6%	176
울산	5	14.7%	29	85.3%	34
경기	197	19.0%	839	81.0%	1,036
강원	23	11.7%	173	88.3%	196
충북	19	11.0%	153	89.0%	172
충남	15	9.4%	144	90.6%	159
전북	13	7.6%	157	92.4%	170
전남	17	8.4%	186	91.6%	203
경북	17	14.4%	101	85.6%	118
경남	11	4.6%	226	95.4%	237
제주	13	20.3%	51	79.7%	64
계	599	13.5%	3,848	86.5%	4,447

참고:  $\chi^2 = 95.27$ , P-value=0.000, Cramér's V=0.146

<표 2-117>은 성별 무응답률의 현황이다. 무응답률의 성별 차이는 유의미하지 않으며, 남성의 무응답률이 다소 높게 나타난다.

<표 2-117> 성별 횡단면 무응답 현황

구분	무응답		응답		계
	빈도	행%	빈도	행%	
남성	401	13.7%	2,516	86.3%	2,917
여성	198	12.9%	1,332	87.1%	1,530
계	599	13.5%	3,848	86.5%	4,447

참조:  $\chi^2 = 0.49$ , P-value=0.483, Cramér's V=0.011

<표 2-118>은 연령대별 무응답률의 현황이다. 응답여부와 연령의 분포는 유의한 관계가 있다. 연령대에 따른 무응답 비율을 살펴보면 15~29세 미만의 연령층의 무응답률이 가장 높으며 50세 이상의 무응답률이 가장 낮았다.

<표 2-118> 연령대별 횡단면 무응답 현황

구분	무응답		응답		계
	빈도	행%	빈도	행%	
15-29세	135	17.2%	649	82.8%	784
30-39세	150	13.8%	939	86.2%	1,089
40-49세	194	15.1%	1,094	84.9%	1,288
50-64세	120	9.3%	1,166	90.7%	1,286
계	599	13.5%	3,848	86.5%	4,447

참조:  $\chi^2 = 31.24$ , P-value=0.000, Cramér's V=0.084

<표 2-119>은 경제활동상태에 따른 무응답 현황이다. 경제활동상태는 응답여부와 유의한 관계가 있는 것으로 나타났다. 실업자의 무응답률이 가장 높았고, 취업미희망자의 무응답률이 가장 낮았다.

<표 2-119> 경제활동상태별 횡단면 무응답 현황

구분	무응답		응답		계
	빈도	행%	빈도	행%	
취업자	307	13.7%	1,934	86.3%	2,241
실업자	31	18.9%	133	81.1%	164
취업희망자	157	14.5%	925	85.5%	1,082
취업미희망자	104	10.8%	856	89.2%	960
계	599	13.5%	3,848	86.5%	4,447

참조:  $\chi^2 = 10.98$ , P-value=0.012, Cramér's V=0.051

<표 2-120>은 장애유형별 무응답률의 현황을 보여준다. 무응답률의 장애유형별 차이는 유의미하지 않으며, 모든 장애유형에 대해서 약 11~14%로 비슷한 무응답률을 보이고 있다.

<표 2-120> 장애유형별 횡단면 무응답 현황

구분	무응답		응답		계
	빈도	행%	빈도	행%	
지체장애	296	13.4%	1,912	86.6%	2,208
지체 외 신체 외부장애	41	12.8%	279	87.2%	320
시각장애	79	13.7%	499	86.3%	578
시각외감각장애	63	15.4%	346	84.6%	409
정신적장애	76	13.0%	510	87.0%	586
신체내부장애	44	12.7%	302	87.3%	346
계	599	13.5%	3,848	86.5%	4,447

참조:  $\chi^2 = 1.752$ , P-value=0.882, Cramér's V=0.023

<표 2-121>은 장애정도별 무응답률의 현황이다. 무응답률의 장애정도별 차이는 유의미하지 않으며, 장애정도에 따른 무응답률을 살펴보면 경증의 무응답률이 중증의 무응답률보다 다소 높게 나타난다.

<표 2-121> 장애정도별 횡단면 무응답 현황

구분	무응답		응답		계
	빈도	행%	빈도	행%	
중증	186	13.2%	1,218	86.8%	1,404
경증	413	13.6%	2,630	86.4%	3,043
계	599	13.5%	3,848	86.5%	4,447

참조:  $\chi^2 = 0.061$ , P-value=0.805, Cramér's V=0.004

## 2) 무응답 조정 가중치 산출

무응답 보정을 위해서는 조사대상자들의 다양한 속성(주로 인구사회학적 특성)을 반영한 응답성향모형을 적합하였다. 이를 위해 6차년도 횡단면 조사의 응답여부는 과거에 관측된 변수들로 설명할 수 있다는 임의응답(MAR)을 가정하였고 1차년도 조사 기준의 설명변수는 물론 2차 조사부터 5차 조사까지 응답여부도 함께 다음의 확률을 고려한 로지스틱 모형을 적합하였다.

$$P(6차\ 응답여부 \mid 2, 3, 4, 5차\ 응답여부, 1차\ 설명변수)$$

즉, 반응변수로는 6차조사 응답여부를 나타내는 지시자이고 2차년도부터 5차년도 조사의 응답여부와 1차년도 기준의 성별, 연령, 지역, 장애유형, 장애정도, 경제활동상태 등의 변수를 설명변수로 하는 로지스틱회귀모형을 다음의 형태로 적합하였다.

$$\hat{p}_i = \frac{\exp\left(\sum_{j=0}^k \hat{\beta}_j x_{ij}\right)}{1 + \exp\left(\sum_{j=0}^k \hat{\beta}_j x_{ij}\right)}$$

<표 2-122>은 6차년도의 응답확률 모형에서 모든 설명변수를 사용한 포화모형을 포함하여 유의하지 않은 설명변수를 하나씩 제거하여 모형을 적합한 결과를 정리하고 있다. 모든 설명변수를 사용하는 모형1 보다는 성별, 경제활동상태를 제외한 모형3이 적합도(AIC) 측면에서 더 좋음을 알 수 있다. 따라서 최종적으로 선택한 모형으로는 2차, 3차, 4차, 5차년도 응답여부, 연령대, 지역, 장애유형을 설명변수로 사용한 로지스틱 회귀 모형이다.

<표 2-122> 모형별 횡단면 로지스틱 회귀분석 결과

구분	모형1(초기모형)			모형2(최종모형)	
	df	$\chi^2$	유의 확률	$\chi^2$	유의 확률
5차 응답여부	1	2109.4	0.000	2109.4	0.000
4차 응답여부	1	92.1	0.000	92.1	0.000
3차 응답여부	1	8.5	0.003	8.5	0.000
2차 응답여부	1	3.7	0.055	3.7	0.055
성별	1	1.2	0.269		
연령대	3	6.8	0.080	6.7	0.082
지역	15	27.4	0.026		
경제활동상태	3	6.8	0.078		
장애유형	5	3.4	0.632		
		AIC = 1319.74		AIC = 1310.67	

최종모형으로부터 추정된 응답확률을  $\hat{p}_i$ 로 정의하면 6차년도에 응답한  $i$ 번째 개체의 종단 가중치는 다음과 같이 계산된다.

$$6\text{차년도 무응답 보정 가중치} = 1\text{차년도 최종가중치} \times \frac{1}{\hat{p}_i}$$

### 3) 벤치마킹 보정

6차년도 횡단 가중치 산출을 위해 무응답을 보정한 가중치에 대하여 모집단 분포에 대한 벤치마킹 보정을 사용한다. 성별, 연령, 지역, 장애유형, 장애정도, 경제활동상태 변수를 통하여 이들의 모집단 분포를 이용한 가중치 합 비율을 통해 벤치마킹 보정을 실시한다. 성, 연령, 지역, 장애유형, 장애정도에 관한 모집단 정보로는 2020년 12월 등록기준 19세에서 68세까지의 보건복지부 장애인 등록현황을 사용하였다.

경제활동상태에 대한 모집단 정보는 2021년 「장애인경제활동실태조사」의 경제활동상태에 대한 추정 결과를 사용하였다. 이 추정결과의 모집단 수와 2019년 12월 기준 장애인 등록 현황의 모집단 수가 일치하지 않기 때문에 추정 결과를 직접 사용할 수는 없다. 이를 위해 2021년 「장애인경제활동실태조사」의 경제활동상태에 대한 비율 추정 결과를 2020년 12월 기준 장애인 등록 현황의 모집단 수에 곱하여 경제활동상태에 대한 모집단 분포를 정의하였다.

레이킹 비 조정 방법을 적용하여 가중치를 보정한 후 횡단 가중치를 다음과 같이 산출하였다.

$$w_{ij}^{rak} = w_{ij} \exp(\lambda' x_{ij}).$$

여기서  $x_{ij}$ 는 레이킹 비 조정 방법을 위해 사용된 보조변수들을 나타내는 지시변수 벡터이며,  $\lambda$ 는  $\sum w_{ij}^{rak} x_{ij} = t_x$ 의 해이며,  $t_x$ 는 각 레이킹 비 조정 방법에 사용된 벤치마킹 변수들의 범주에 해당하는 모집단 장애인 분포를 나타낸다. 다음 절에서 횡단 가중치 극단값 조정을 통한 최종가중치를 사용한 표본 자료 분포와 모집단 분포 비교 결과를 살펴보기로 한다.

#### 4) 극단값 처리 및 최종가중치 산출

앞서 산출한 레이킹 비 조정 횡단가중치에는 설계 가중치의 최댓값이었던 약 4,500을 넘는 가중치들이 존재하였다. 이들은 조사변수에 곱해져 합산되므로 경우에 따라서 통계량에 과도한 영향력을 갖게 된다. 따라서 이러한 과도한 가중치 값을 추가로 조정해 주기 위해 4,500보다 작은 최댓값인 4,435으로 절사하여 조정해 준 다음, 앞서 고려한 바와 같이 성, 연령, 장애유형, 장애정도의 보조변수들에 대해서 다시 한번 레이킹 비 보정을 실시하여 이를 최종가중치로 산출하였다.

<표 2-123>는 횡단면 최종가중치에 대한 기술통계량을 나타내고 있다. 평균 394.3이며 최솟값은 2.9이고 최댓값은 4,989.6이다.

<표 2-123> 횡단 가중치에 대한 통계량

통계량			
응답수	3,848	합계	1,517,358
평균	394.3	표준편차	590.4
중위수	196.3	분산	348,538.5
사분위 범위	383.5	범위	4,986.7
최솟값	2.9	최댓값	4,989.6

벤치마킹 보정을 위해 고려하였던 보조변수인 성별, 연령대, 지역, 장애유형, 장애정도, 경제활동상태 변수의 모집단 총합과 표본가중합의 분포는 <표 2-124>로부터 <표 2-129>에 각각 정리하고 있다. 횡단면 최종가중치를 사용한 표본가중합이 모집단 총합과 거의 차이가 없음을 알 수 있다.

&lt;표 2-124&gt; 성별 모집단 및 표본 분포

성	표본가중합		모집단 총합	
	인구	비중	인구	비중
남자	999,332	65.9%	999,332	65.9%
여자	518,026	34.1%	518,026	34.1%
총합	1,517,358	100.0%	1,517,358	100.0%

&lt;표 2-125&gt; 연령별 모집단 및 표본 분포

연령	표본가중합		모집단 총합	
	인구	비중	인구	비중
20~29세	97,861	6.4%	97,860	6.4%
30~39세	122,303	8.1%	122,303	8.1%
40~49세	242,691	16.0%	242,691	16.0%
50~69세	1,054,503	69.5%	1,054,504	69.5%
총합	1,517,358	100.0%	1,517,358	100.0%

&lt;표 2-126&gt; 경제활동상태 변수의 모집단 및 표본 분포

경제활동상태	표본가중합		모집단 총합	
	인구	비중	인구	비중
취업자	720,437	47.5%	720,437	47.5%
미취업자	54,913	3.6%	54,913	3.6%
취업희망자	173,518	11.4%	173,518	11.4%
취업미희망자	568,490	37.5%	568,490	37.5%
총합	1,517,358	100.0%	1,517,358	100.0%

&lt;표 2-127&gt; 장애유형별 변수의 모집단 및 표본 분포

장애유형	표본가중합		모집단 총합	
	인구	비중	인구	비중
지체장애	731,833	48.2%	731,833	48.2%
지체 외 신체 외부장애	130,781	8.6%	130,781	8.6%
시각장애	144,949	9.6%	144,949	9.6%
시각 외 감각장애	128,745	8.5%	128,745	8.5%
정신적 장애	276,579	18.2%	276,579	18.2%
신체 내부장애	104,471	6.9%	104,471	6.9%
총합	1,517,358	100.0%	1,517,358	100.0%



<표 2-128> 시도별 모집단 및 표본 분포

지역	표본가중합		모집단 총합	
	인구	비중	인구	비중
서울	225,856	14.9%	224,133	14.8%
부산	102,992	6.8%	105,547	7.0%
대구	74,478	4.9%	73,485	4.8%
인천	89,296	5.9%	91,338	6.0%
광주	42,063	2.8%	42,779	2.8%
대전	43,905	2.9%	44,685	2.9%
울산	32,487	2.1%	32,961	2.2%
세종	7,342	0.5%	7,400	0.5%
경기	344,718	22.7%	337,391	22.2%
강원	56,496	3.7%	57,988	3.8%
충북	56,734	3.7%	57,747	3.8%
충남	72,904	4.8%	72,163	4.8%
전북	69,371	4.6%	71,135	4.7%
전남	71,633	4.7%	72,607	4.8%
경북	96,127	6.3%	95,775	6.3%
경남	110,656	7.3%	110,587	7.3%
제주	20,300	1.3%	19,638	1.3%
총합	1,517,358	100.0%	1,517,358	100.0%

<표 2-129> 장애정도별 변수의 모집단 및 표본 분포

장애정도	표본가중합		모집단 총합	
	인구	비중	인구	비중
중증	642,314	42.3%	642,314	42.3%
경증	875,044	57.7%	875,044	57.7%
총합	1,517,358	100.0%	1,517,358	100.0%

<표 2-130>는 1~6차조사들의 횡단 가중치의 기술통계량 및 분포를 비교하고 있다. 1차년도 최종가중치에 비하여 6차년도의 횡단 가중치는 전반적으로 증가하였지만, 극단값 조정을 통하여 심각한 특이 가중치는 없는 것으로 판단되고 변동계수(CV)도 1차년도에 비하여 약 8% 정도 증가하였다.

<표 2-130> 1-6차년도 횡단가중치 분포

	1차년도	2차년도	3차년도	4차년도	5차년도	6차년도
최대값	4,487.5	4,756.8	4,941.0	4536.3	4644.4	4989.6
99%	2,129.5	2,260.2	2,396.5	2631.4	3040.7	3098.7
95%	1,054.3	1,166.0	1,220.5	1264.9	1363.2	1417.9
90%	722.5	772.3	849.5	862.7	909.2	937.6
75%	363.2	405.6	413.4	428.0	453.7	463.4
50%	160.2	179.8	179.1	184.4	188.2	196.3
25%	68.0	78.4	77.9	77.3	79.3	79.8
10%	30.5	34.2	35.3	33.8	33.7	33.9
5%	19.8	22.4	22.8	22.8	22.7	22.6
1%	9.5	10.6	10.9	10.7	11.0	11.3
최소값	3.2	3.4	3.0	3.2	2.9	2.9
평균	303.2	335.0	350.5	361.0	381.7	394.3
CV	139.3%	138.0%	142.2%	141.6%	147.4%	149.7%

해당 가중치를 사용하여 주요 변수인 장애유형, 경제활동상태에 대한 빈도 및 비율과 상대표준오차(CV) 값을 살펴본 결과는 <표 2-131> 및 <표2-132>에 제시하였다.

<표 2-131> 장애유형 (15개) 별 경제활동상태(2개 구분) 빈도/비율 추정 및 상대표준오차 분석

장애유형	경제활동상태 (2개 구분)	가중빈도	비중(%)	표준오차(%)	상대표준오차 (%)
지체장애	취업자	287,231	40.2%	2.6%	6.5%
	미취업자	427,145	59.8%	3.0%	5.1%
	합계	714,375	100.0%		
뇌병변장애	취업자	92,355	75.0%	9.8%	13.1%
	미취업자	30,762	25.0%	5.5%	21.9%
	합계	123,117	100.0%		

<표 2-132> 장애유형 (15개) 별 경제활동상태(2개 구분) 빈도/비율 추정 및 상대표준오차 분석(계속)

시각장애	취업자	70,875	47.8%	5.3%	11.0%
	미취업자	77,437	52.2%	5.1%	9.8%
	합계	148,312	100.0%		
청각장애	취업자	59,878	52.5%	6.8%	13.0%
	미취업자	54,265	47.5%	6.6%	13.8%
	합계	114,144	100.0%		
언어장애	취업자	6,568	51.3%	20.5%	39.9%
	미취업자	6,240	48.7%	19.7%	40.5%
	합계	12,807	100.0%		
지적(정신지체)장애	취업자	128,757	71.5%	6.7%	9.4%
	미취업자	51,361	28.5%	4.7%	16.6%
	합계	180,118	100.0%		
자폐성(발달)장애	취업자	4,925	61.6%	19.7%	31.9%
	미취업자	3,065	38.4%	13.8%	36.1%
	합계	7,990	100.0%		
정신장애	취업자	67,061	80.6%	11.5%	14.2%
	미취업자	16,127	19.4%	6.4%	32.9%
	합계	83,188	100.0%		
신장장애	취업자	38,400	65.4%	10.2%	15.6%
	미취업자	20,358	34.6%	7.0%	20.3%
	합계	58,758	100.0%		
심장장애	취업자	7,482	82.0%	53.4%	65.1%
	미취업자	1,641	18.0%	12.8%	71.4%
	합계	9,123	100.0%		
호흡기장애	취업자	2,296	63.9%	29.0%	45.3%
	미취업자	1,294	36.1%	17.9%	49.5%
	합계	3,590	100.0%		
간장애	취업자	5,454	34.4%	17.2%	50.0%
	미취업자	10,408	65.6%	24.3%	37.1%
	합계	15,863	100.0%		
안면장애	취업자	1,780	26.1%	21.4%	81.8%
	미취업자	5,029	73.9%	58.8%	79.6%
	합계	6,809	100.0%		
장루요루장애	취업자	3,627	72.0%	36.4%	50.6%
	미취업자	1,413	28.0%	16.4%	58.5%
	합계	5,039	100.0%		
뇌전증(간질)장애	취업자	5,835	71.6%	19.2%	26.8%
	미취업자	2,310	28.4%	15.6%	54.9%
	합계	8,146	100.0%		

## 9. 7차조사 가중치 산출 및 추정

### 가. 종단 가중치 산출

#### 1) 차수별 응답 현황 및 단위 무응답 분석

7차년도 종단 가중치는 1차부터 7차에 모두 응답한 3,481명을 대상으로 산출한다. 7차조사 종단 가중치의 산출에 앞서 단위무응답의 현황을 살펴보고 응답성향분석을 통해 단위무응답에 영향을 미치는 조사대상자 특성을 우선적으로 파악하고자 하였다.

<표 2-133>은 장애인 거주 시도별 응답현황을 정리하였고 더불어 시도 분류와 응답여부와의 독립성 검정을 위한 카이제곱 통계와 크래머 V 통계를 산출하였다. 시도간 무응답률의 편차는 유의미하게 나타났다. 그렇지만 각 시도에서 무응답률은 모두 약 5% 미만으로 낮게 나타난다.

<표 2-133> 시도별 종단면 무응답 현황

시도	무응답		응답		계
	빈도	행%	빈도	행%	
서울	18	2.9%	598	97.1%	616
부산	9	2.5%	354	97.5%	363
대구	4	2.0%	197	98.0%	201
인천	4	2.5%	159	97.5%	163
광주	2	1.9%	106	98.1%	108
대전	0	0.0%	161	100.0%	161
울산	0	0.0%	31	100.0%	31
세종	0	0.0%	8	100.0%	8
경기	22	2.9%	737	97.1%	759
강원	0	0.0%	164	100.0%	164
충북	7	4.8%	140	95.2%	147
충남	2	1.4%	136	98.6%	138
전북	0	0.0%	147	100.0%	147
전남	0	0.0%	176	100.0%	176
경북	4	4.1%	94	95.9%	98
경남	2	0.9%	232	99.1%	234
제주	1	2.4%	41	97.6%	42
계	75	2.1%	3,481	97.9%	3,556

참고:  $\chi^2 = 28.395$ , P-value=0.028, Cramér's V=0.089

〈표 2-134〉은 장애인의 성별 응답현황을 보여주고 있다. 남성과 여성 간의 무응답률은 유의한 차이가 아니며, 남성의 무응답률이 다소 높게 나타난다.

〈표 2-134〉 성별 종단면 무응답 현황

구분	무응답		응답		계
	빈도	행%	빈도	행%	
남성	51	2.2%	2,256	97.8%	2,307
여성	24	1.9%	1,225	98.1%	1,249
계	75	2.1%	3,481	97.9%	3,556

참고:  $\chi^2 = 0.203$ , P-value=0.652, Cramér's V=0.008

〈표 2-135〉은 장애인의 연령대별 응답률 현황을 나타내고 있다. 연령대에 따른 무응답률의 편차는 유의한 차이가 아닌 것으로 나타난다. 30~39세 미만의 연령층의 무응답 비율이 가장 높고 40~49세 미만의 연령층의 무응답 비율이 가장 낮지만 전 연령층에서 무응답 비율은 2~3%로 유사하게 나타난다.

〈표 2-135〉 연령대별 종단면 무응답 현황

구분	무응답		응답		계
	빈도	행%	빈도	행%	
19-29세	8	2.1%	367	97.9%	375
30-39세	17	3.2%	519	96.8%	536
40-49세	15	1.5%	953	98.5%	968
50-70세	35	2.1%	1,642	97.9%	1677
계	75	2.1%	3,481	97.9%	3,556

참고:  $\chi^2 = 4.404$ , P-value=0.221, Cramér's V=0.035

〈표 2-136〉은 장애인의 경제활동상태별 응답률의 현황을 정리하고 있다. 경제활동상태에 따른 무응답률의 편차는 유의미하게 나타났다. 실업자의 무응답률이 6.2%로 가장 높게 나왔고, 그 외 유형에서 무응답률은 약 2% 정도로 상대적으로 낮게 나타났다.

〈표 2-136〉 경제활동상태별 종단면 무응답 현황

구분	무응답		응답		계
	빈도	행%	빈도	행%	
취업자	34	1.8%	1,816	98.2%	1,850
실업자	4	6.2%	61	93.8%	65
취업희망자	5	2.5%	198	97.5%	203
취업미희망자	32	2.2%	1,406	97.8%	1,438
계	75	2.1%	3,481	97.9%	3,556

참고:  $\chi^2 = 6.027$ , P-value=0.110, Cramér's V=0.041

<표 2-137>은 장애인의 장애유형별 응답률의 현황을 보여준다. 장애유형에 따른 무응답률의 편차는 유의하지 않게 나타났다. 정신적장애의 무응답률이 가장 높고 시각 외 감각장애의 무응답률이 가장 작게 나타났다. 그 외 장애유형의 무응답률은 비슷하게 나타났다.

<표 2-137> 장애유형별 종단면 무응답 현황

구분	무응답		응답		계
	빈도	행%	빈도	행%	
지체장애	37	2.1%	1,723	97.9%	1,760
지체 외 신체 외부장애	5	1.9%	262	98.1%	267
시각장애	11	2.4%	452	97.6%	463
시각외감각장애	4	1.3%	315	98.7%	319
정신적장애	12	2.5%	466	97.5%	478
신체내부장애	6	2.2%	263	97.8%	269
계	75	2.1%	3,481	97.9%	3,556

참고:  $\chi^2 = 1.754$ , P-value=0.882, Cramér's V=0.022

<표 2-138>은 장애인의 장애정도별 응답률 현황이다. 장애정도에 따른 무응답률의 편차는 유의하지 않은 것으로 나타났다. 장애정도에 따른 무응답률을 살펴보면 중증과 경증의 무응답률은 거의 비슷하게 나타남을 알 수 있다.

<표 2-138> 장애정도별 종단면 무응답 현황

구분	무응답		응답		계
	빈도	행%	빈도	행%	
중증	34	2.4%	1,354	97.6%	1388
경증	41	1.9%	2,127	98.1%	2168
계	75	2.1%	3,481	97.9%	3556

참고:  $\chi^2 = 1.022$ , P-value=0.312, Cramér's V=0.017

## 2) 응답성향모형 적합

앞서 인구사회학적 특성과 장애유형 및 장애정도 특성별 무응답의 연관성 분석을 토대로 7차년도 종단 가중치를 산출하기 위해서는 조사대상자들의 다양한 속성(주로 인구사회학적 특성)을 반영한 응답성향모형을 적합하였다. 응답성향모형은 7차년도 응답여부는 과거에 관측된 보조변수들로 설명할 수 있다는 임의무응답(missing at random, 이하 MAR)을 가정한다. 따라서 임의무응답의 가정 하에서는 2차 조사부터 7차 조사에서 모두 응답한 패널의 응답성향모형은 6차년도의 주요변수들로 설명할 수 있게 된다.

응답성향모형은 로지스틱 회귀모형을 사용하였다. 추정된 모형을 통해 조사대상자 각각의 응답성향점수(혹은 응답확률)를 계산하고 그 역수를 통해 무응답에 대한 보정계수로 산출하여 6차년도 종단 가중치에 곱한 후 이를 6차년도 종단 가중치를 계산하였다. 즉, 7차 조사의 표본응답자가 응답했는지 여부를 반응변

수로 하고, 6차 조사의 표본응답자의 특성변수를 설명변수로 하는 로지스틱회귀모형을 적합하여 응답률의 차이를 파악하고 이를 통해 보정하였다. 응답확률은 로지스틱회귀모형 하에서 다음과 같이 표현할 수 있다.

$$\hat{p}_i = \frac{\exp\left(\sum_{j=0}^k \hat{\beta}_j x_{ij}\right)}{1 + \exp\left(\sum_{j=0}^k \hat{\beta}_j x_{ij}\right)}$$

여기서  $\hat{p}_i$ 는  $i$ 번째 패널개체에 대한 응답확률의 추정값이고  $x_{ij}$ 는  $i$ 번째 패널개체의  $j$ 번째 보조변수 특성이며  $\hat{\beta}_j$ 는 회귀계수를 각각 나타낸다.

### 3) 최종 종단 가중치 산출

7차년도 종단면 응답자  $i$ 에 대한 종단 가중치는 6차년도 종단가중치에 앞서 계산한 응답확률  $\hat{p}_i$ 를 나눠 다음과 같이 계산하였다.

$$7\text{차년도 종단가중치} = 6\text{차년도 종단가중치} \times \frac{1}{\hat{p}_i}$$

여기서 사용된 설명변수는 성별, 연령, 지역 등의 변수와 장애유형, 장애정도, 경제활동상태 등의 변수들이다.

로지스틱 회귀모형에 대하여 초기 모형과 최종 모형에 대한 적합결과를 정리하여 <표 2-137>에 제시하였다. 앞서 고려한 모든 변수를 사용한 초기 모형보다는 지역 및 경제활동상태를 설명변수로 사용한 최종 모형이 적합도(AIC) 측면에서 더 좋음을 알 수 있다. 경제활동상태 중 취업자 대비 실업자의 효과가 유의하여 경제활동상태를 모형에 포함시켰다.

최종모형을 기반으로 산출한 종단가중치와 모든 변수를 사용한 초기 모형을 기반으로 산출한 종단 가중치를 비교해보면 두 가중치는 매우 유사한 값을 지니며 두 가중치의 상관계수는 0.99 이상으로 나타난다. 따라서 연령 및 장애유형과 장애정도는 종단 가중치 산출에 있어서 유의미한 차이를 만들지 않는다고 판단된다.

7차 조사의 응답 성향 모형에서 설명변수가 지역 하나로 나오는 것은 다른 차수 조사에 비하여 무응답자 수와 비율이 현저히 작아서 불균등 자료에서 발생하는 특별한 상황이라고 생각된다. 추후 8차 이후의 조사에서도 이러한 현상이 계속 발생된다면 로지스틱 회귀와 같은 분포 가정하의 응답성향 모형보다는 비모수적 방법을 고민할 필요가 있다.

<표 2-139>은 7차년도 종단 가중치에 대한 기술 통계량을 보여주고 있다. 종단면 응답자 총 3,481명의 종단 가중치 평균은 385.5이고 최솟값 4.4에서 최댓값 5,621.5의 범위를 갖는다.

<표 2-141>는 3~7차 조사까지의 종단 가중치 분포의 현황을 정리하고 있다. 3차년도 종단 가중치에 비하여 7차년도의 종단 가중치는 전반적으로 증가하였지만 변동계수(CV)는 3차년도에 비하여 다소 줄어들었음을 알 수 있다.

〈표 2-139〉 모형별 종단면 로지스틱 회귀분석 결과

설명변수	포화모형			최종모형	
	자유도	$\chi^2$	유의 확률	$\chi^2$	유의 확률
성별	1	0.333	0.564		
연령대	3	4.179	0.243		
지역	16	40.283	0.001	41.274	0.000
경제활동상태	3	4.351	0.226	4.455	0.216
장애유형	5	1.863	0.868		
장애정도	1	0.911	0.340		
		AIC = 735.323		AIC = 721.513	

〈표 2-140〉 7차년도 종단 가중치에 대한 통계량

통계량			
응답수	3,481	합계	1,341,881
평균	385.5	표준편차	508.9
중위수	212.3	분산	259,018.7
사분위 범위	377.8	범위	5,617.1
최솟값	4.4	최댓값	5,621.5

〈표 2-141〉 3-7차조사의 종단 가중치 분포

통계량	3차년도 종단가중치	4차년도 종단가중치	5차년도 종단가중치	6차년도 종단가중치	7차년도 종단가중치
최대값	5,032.7	5,186.0	5,365.7	5421.6	5621.5
99%	2,387.3	2,467.8	2,518.6	2506.4	2541.9
95%	1,189.2	1,250.0	1,289.0	1314.8	1328.1
90%	817.9	847.2	874.7	897.5	917.0
75%	422.0	440.0	453.5	462.0	470.0
50%	186.9	196.2	203.7	208.5	212.3
25%	79.3	84.3	88.3	90.2	92.2
10%	35.4	37.5	39.6	40.6	41.2
5%	23.3	25.2	26.6	26.9	27.7
1%	11.0	11.9	12.8	13.2	13.4
최소값	3.9	4.1	4.3	4.3	4.4
평균	346.3	360.1	371.1	378.7	385.5
CV	135.4%	134.0%	132.3%	131.5%	132.0%



## 나. 횡단 가중치 산출

### 1) 단위무응답 분석

7차년도 횡단 가중치는 7차에 응답한 3,763명을 대상으로 산출한다. 가중치 산출을 위하여 7차년도 응답자 3,763명과 무응답자는 668명을 대상으로 무응답 특성을 분석한다. 7차년도 응답여부와 7차년도 응답자 모두가 응답한 1차년도의 주요 특성 변수들 간의 연관성을 살펴봄으로써 응답여부에 중요하게 영향을 미치는 특성 변수를 파악하고자 하였다.

〈표 2-142〉은 시도별 무응답률 현황을 보여주고 있다. 무응답률의 시도 차이는 유의미하며, 서울, 대구, 인천, 광주, 울산, 경기, 강원, 충북, 충남, 경북, 제주 지역에서 무응답률이 약 10~25%로 높게 나타나며, 부산, 대전, 전북, 전남, 경남 지역에서 무응답 비율이 약 10% 미만으로 상대적으로 낮았다.

〈표 2-142〉 시도별 횡단면 무응답 현황

시도	무응답		응답		계
	빈도	행%	빈도	행%	
서울	152	17.6%	711	82.4%	863
부산	33	8.1%	374	91.9%	407
대구	38	15.1%	214	84.9%	252
인천	50	21.6%	182	78.4%	232
광주	15	12.1%	109	87.9%	124
대전	14	8.0%	162	92.0%	176
울산	5	14.7%	29	85.3%	34
경기	220	21.3%	813	78.7%	1,033
강원	24	12.4%	169	87.6%	193
충북	25	14.8%	144	85.2%	169
충남	16	10.1%	143	89.9%	159
전북	11	6.5%	158	93.5%	169
전남	17	8.4%	185	91.6%	202
경북	20	16.9%	98	83.1%	118
경남	12	5.1%	225	94.9%	237
제주	16	25.4%	47	74.6%	63
계	668	15.1%	3,763	84.9%	4,431

참고:  $\chi^2 = 111.428$ , P-value=0.000, Cramér's V=0.159

<표 2-143>은 성별 무응답률의 현황이다. 무응답률의 성별 차이는 유의미하지 않으며, 남성의 무응답률이 다소 높게 나타난다.

<표 2-143> 성별 횡단면 무응답 현황

구분	무응답		응답		계
	빈도	행%	빈도	행%	
남성	452	15.6%	2,454	84.4%	2,906
여성	216	14.2%	1,309	85.8%	1,525
계	668	15.1%	3,763	84.9%	4,431

참조:  $\chi^2 = 1.403$ , P-value=0.236, Cramér's V=0.018

<표 2-144>은 연령대별 무응답률의 현황이다. 응답여부와 연령의 분포는 유의한 관계가 있다. 연령대에 따른 무응답 비율을 살펴보면 15~29세 미만의 연령층의 무응답률이 가장 높으며 50세 이상의 무응답률이 가장 낮았다.

<표 2-144> 연령대별 횡단면 무응답 현황

구분	무응답		응답		계
	빈도	행%	빈도	행%	
15-29세	152	19.4%	631	80.6%	783
30-39세	168	15.5%	918	84.5%	1,086
40-49세	212	16.5%	1,070	83.5%	1,282
50-64세	136	10.6%	1,144	89.4%	1,280
계	668	15.1%	3,763	84.9%	4,431

참조:  $\chi^2 = 33.576$ , P-value=0.000, Cramér's V=0.087

<표 2-145>은 경제활동상태에 따른 무응답 현황이다. 경제활동상태는 응답여부와 유의한 관계가 있는 것으로 나타났다. 실업자의 무응답률이 가장 높았고, 취업미희망자의 무응답률이 가장 낮았다.

<표 2-145> 경제활동상태별 횡단면 무응답 현황

구분	무응답		응답		계
	빈도	행%	빈도	행%	
취업자	340	15.2%	1,899	84.8%	2,239
실업자	31	18.9%	133	81.1%	164
취업희망자	180	16.7%	897	83.3%	1,077
취업미희망자	117	12.3%	834	87.7%	951
계	668	15.1%	3,763	84.9%	4,431

참조:  $\chi^2 = 9.863$ , P-value=0.020, Cramér's V=0.047

<표 2-146>은 장애유형별 무응답률의 현황을 보여준다. 무응답률의 장애유형별 차이는 유의미하지 않으며, 모든 장애유형에 대해서 약 14~16%로 비슷한 무응답률을 보이고 있다.

<표 2-146> 장애유형별 횡단면 무응답 현황

구분	무응답		응답		계
	빈도	행%	빈도	행%	
지체장애	330	15.0%	1,872	85.0%	2,202
지체 외 신체 외부장애	46	14.4%	274	85.6%	320
시각장애	83	14.4%	494	85.6%	577
시각외감각장애	66	16.1%	343	83.9%	409
정신적장애	88	15.1%	495	84.9%	583
신체내부장애	55	16.2%	285	83.8%	340
계	668	15.1%	3,763	84.9%	4,431

참조:  $\chi^2 = 1.033$ , P-value=0.960, Cramér's V=0.015

<표 2-147>은 장애정도별 무응답률의 현황이다. 무응답률의 장애정도별 차이는 유의미하지 않으며, 장애정도에 따른 무응답률을 살펴보면 중증의 무응답률이 경증의 무응답률보다 다소 높게 나타난다.

<표 2-147> 장애정도별 횡단면 무응답 현황

구분	무응답		응답		계
	빈도	행%	빈도	행%	
중증	216	15.5%	1,180	84.5%	1,396
경증	452	14.9%	2,583	85.1%	3,035
계	668	15.1%	3,763	84.9%	4,431

참조:  $\chi^2 = 0.208$ , P-value=0.648, Cramér's V=0.007

## 2) 무응답 조정 가중치 산출

무응답 보정을 위해서는 조사대상자들의 다양한 속성(주로 인구사회학적 특성)을 반영한 응답성향모형을 적합하였다. 이를 위해 7차년도 횡단면 조사의 응답여부는 과거에 관측된 변수들로 설명할 수 있다는 임의응답(MAR)을 가정하였고 1차년도 조사 기준의 설명변수는 물론 2차 조사부터 6차 조사까지 응답여부도 함께 다음의 확률을 고려한 로지스틱 모형을 적합하였다.

$$P(7차\ 응답여부 \mid 2, 3, 4, 5, 6차\ 응답여부, 1차\ 설명변수)$$

즉, 반응변수로는 7차조사 응답여부를 나타내는 지시자이고 2차년도부터 6차년도 조사의 응답여부와 1차년도 기준의 성별, 연령, 지역, 장애유형, 장애정도, 경제활동상태 등의 변수를 설명변수로 하는 로지스틱회귀모형을 다음의 형태로 적합하였다.

$$\hat{p}_i = \frac{\exp\left(\sum_{j=0}^k \hat{\beta}_j x_{ij}\right)}{1 + \exp\left(\sum_{j=0}^k \hat{\beta}_j x_{ij}\right)}$$

<표 2-148>은 7차년도에 응답확률 모형에서 모든 설명변수를 사용한 포화모형을 포함하여 유의하지 않은 설명변수를 하나씩 제거하여 모형을 적합한 결과를 정리하고 있다. 모든 설명변수를 사용하는 모형1보다는 성별, 연령대, 장애유형 및 장애정도와 경제활동상태를 제외한 모형2가 적합도(AIC) 측면에서 더 좋음을 알 수 있다. 따라서 최종적으로 선택한 모형으로는 2차 - 6차년도 응답여부, 지역을 설명변수로 사용한 로지스틱 회귀 모형이다.

<표 2-148> 모형별 횡단면 로지스틱 회귀분석 결과

구분	모형1(초기모형)			모형2(최종모형)	
	df	$\chi^2$	유의 확률	$\chi^2$	유의 확률
6차 응답여부	1	2,471.4	0.000	2,471.4	0.000
5차 응답여부	1	97.1	0.000	97.1	0.000
4차 응답여부	1	21.2	0.000	21.2	0.000
3차 응답여부	1	7.4	0.007	7.4	0.007
2차 응답여부	1	6.0	0.014	6.0	0.014
성별	1	0.5	0.498		
연령대	3	3.5	0.321		
지역	15	30.6	0.010	30.7	0.010
경제활동상태	3	2.2	0.543		
장애유형	5	4.4	0.493		
		AIC = 1179.550		AIC = 1165.939	

최종모형으로부터 추정된 응답확률을  $\hat{p}_i$ 로 정의하면 7차년도에 응답한  $i$ 번째 개체의 중단 가중치는 다음과 같이 계산된다.

$$7\text{차년도 무응답 보정 가중치} = 1\text{차년도 최종가중치} \times \frac{1}{\hat{p}_i}$$

### 3) 벤치마킹 보정

7차년도 횡단 가중치 산출을 위해 무응답을 보정한 가중치에 대하여 모집단 분포에 대한 벤치마킹 보정을 사용한다. 성별, 연령, 지역, 장애유형, 장애정도, 경제활동상태 변수를 통하여 이들의 모집단 분포를 이용한 가중치 합 비율을 통해 벤치마킹 보정을 실시한다. 성, 연령, 지역, 장애유형, 장애정도에 관한 모집단 정보로는 2021년 12월 등록기준 21세에서 70세까지의 보건복지부 장애인 등록현황을 사용하였다.

경제활동상태에 대한 모집단 정보는 2021년 「장애인경제활동실태조사」의 경제활동상태에 대한 추정 결과를 사용하였다. 이 추정결과의 모집단 수와 2021년 12월 기준 장애인 등록 현황의 모집단 수가 일치하지 않기 때문에 추정 결과를 직접 사용할 수는 없다. 이를 위해 2021년 「장애인경제활동 실태조사」의 경제활동상태에 대한 비율 추정 결과를 2020년 12월 기준 장애인 등록 현황의 모집단 수에 곱하여 경제활동상태에 대한 모집단 분포를 정의하였다.

레이킹 비 조정 방법을 적용하여 가중치를 보정한 후 횡단 가중치를 다음과 같이 산출하였다.

$$w_{ij}^{rak} = w_{ij} \exp(\lambda' x_{ij}).$$

여기서  $x_{ij}$ 는 레이킹 비 조정 방법을 위해 사용된 보조변수들을 나타내는 지시변수 벡터이며,  $\lambda$ 는  $\sum w_{ij}^{rak} x_{ij} = t_x$ 의 해이며,  $t_x$ 는 각 레이킹 비 조정 방법에 사용된 벤치마킹 변수들의 범주에 해당하는 모집단 장애인 분포를 나타낸다. 다음 절에서 횡단 가중치 극단값 조정을 통한 최종가중치를 사용한 표본 자료 분포와 모집단 분포 비교 결과를 살펴보기로 한다.

#### 4) 극단값 처리 및 최종가중치 산출

앞서 산출한 레이킹 비 조정 횡단가중치에는 설계 가중치의 최댓값이었던 약 4,500을 넘는 가중치들이 존재하였다. 이들은 조사변수에 곱해져 합산되므로 경우에 따라서 통계량에 과도한 영향력을 갖게 된다. 따라서 이러한 과도한 가중치 값을 추가로 조정해 주기 위해 4,500으로 절사하여 조정해 준 다음, 앞서 고려한 바와 같이 성, 연령, 장애유형, 장애정도의 보조변수들에 대해서 다시 한번 레이킹 비 보정을 실시하여 이를 최종가중치로 산출하였다.

<표 2-149>는 횡단면 최종가중치에 대한 기술통계량을 나타내고 있다. 평균 410.2이며 최솟값은 2.6이고 최댓값은 5653.5이다.

<표 2-149> 횡단 가중치에 대한 통계량

통계량			
응답수	3,763	합계	1,543,464
평균	410.2	표준편차	629.8
중위수	196.5	분산	396638.0
사분위 범위	377.9	범위	5650.9
최솟값	2.6	최댓값	5653.5

벤치마킹 보정을 위해 고려하였던 보조변수인 성별, 연령대, 지역, 장애유형, 장애정도, 경제활동상태 변수의 모집단 총합과 표본가중합의 분포는 <표 2-150>로부터 <표 2-155>에 각각 정리하고 있다. 횡단면 최종가중치를 사용한 표본가중합이 모집단 총합과 거의 차이가 없음을 알 수 있다.

&lt;표 2-150&gt; 성별 모집단 및 표본 분포

성	표본가중합		모집단 총합	
	인구	비중	인구	비중
남자	1,014,306	65.7%	1,014,307	65.7%
여자	529,158	34.3%	529,157	34.3%
총합	1,543,464	100.0%	1,543,464	100.0%

&lt;표 2-151&gt; 연령별 모집단 및 표본 분포

연령	표본가중합		모집단 총합	
	인구	비중	인구	비중
21~29세	89,066	5.8%	89,059	5.8%
30~39세	116,912	7.6%	116,907	7.6%
40~49세	231,390	15.0%	231,388	15.0%
50~70세	1,106,094	71.7%	1,106,110	71.7%
총합	1,543,464	100.0%	1,543,464	100.0%

&lt;표 2-152&gt; 경제활동상태 변수의 모집단 및 표본 분포

경제활동상태	표본가중합		모집단 총합	
	인구	비중	인구	비중
취업자	743,696	48.2%	743,696	48.2%
미취업자	39,471	2.6%	39,471	2.6%
취업희망자	150,319	9.7%	150,319	9.7%
취업미희망자	609,977	39.5%	609,977	39.5%
총합	1,543,464	100.0%	1,543,464	100.0%

&lt;표 2-153&gt; 장애유형별 변수의 모집단 및 표본 분포

장애유형	표본가중합		모집단 총합	
	인구	비중	인구	비중
지체장애	719,152	46.6%	719,148	46.6%
지체 외 신체 외부장애	134,768	8.7%	134,769	8.7%
시각장애	150,294	9.7%	150,294	9.7%
시각 외 감각장애	147,903	9.6%	147,903	9.6%
정신적 장애	277,602	18.0%	277,606	18.0%
신체 내부장애	113,744	7.4%	113,744	7.4%
총합	1,543,464	100.0%	1,543,464	100.0%

<표 2-154> 시도별 모집단 및 표본 분포

지역	표본가중합		모집단 총합	
	인구	비중	인구	비중
서울	220,543	14.3%	233,091	15.1%
부산	70,114	4.5%	107,399	7.0%
대구	85,453	5.5%	81,076	5.3%
인천	98,969	6.4%	91,833	5.9%
광주	43,699	2.8%	40,469	2.6%
대전	45,217	2.9%	41,945	2.7%
울산	38,476	2.5%	35,549	2.3%
세종	8,051	0.5%	6,995	0.5%
경기	375,031	24.3%	351,426	22.8%
강원	55,553	3.6%	53,254	3.5%
충북	61,034	4.0%	58,702	3.8%
충남	80,042	5.2%	73,002	4.7%
전북	74,401	4.8%	68,576	4.4%
전남	74,589	4.8%	67,748	4.4%
경북	68,754	4.5%	100,445	6.5%
경남	122,917	8.0%	113,234	7.3%
제주	20,622	1.3%	18,718	1.3%
총합	1,543,464	100.0%	1,517,358	100.0%

<표 2-155> 장애정도별 변수의 모집단 및 표본 분포

장애정도	표본가중합		모집단 총합	
	인구	비중	인구	비중
중증	651,006	42.2%	651,006	42.2%
경증	892,458	57.8%	892,458	57.8%
총합	1,543,464	100.0%	1,543,464	100.0%

<표 2-156> 은 1~7차 조사들의 횡단 가중치의 기술통계량 및 분포를 비교하고 있다. 1차년도 최종가중치에 비하여 7차년도의 횡단 가중치는 전반적으로 증가하였다. 극단값 조정을 통하여 극단 가중치를 조정하였으나 최대값이 이전 6차조사까지의 횡단가중치에 비하여 크게 나타나는 경향이 보인다. 이러한 이유는 2차부터 6차까지 응답하지 않았으나 7차에 응답을 한 응답자가 존재하여 이러한 경우 응답확률이 극히 낮게 계산되어 무응답 보정 가중치가 상당히 크게 계산되어 벤치마킹 보정 가중치도 큰 값으로 나타나게 된다.

<표 2-156> 1-7차년도 횡단가중치 분포

	1차년도	2차년도	3차년도	4차년도	5차년도	6차년도	7차년도
최대값	4,487.5	4,756.8	4,941.0	4536.3	4644.4	4989.6	5653.5
99%	2,129.5	2,260.2	2,396.5	2631.4	3040.7	3098.7	3477.8
95%	1,054.3	1,166.0	1,220.5	1264.9	1363.2	1417.9	1508.5
90%	722.5	772.3	849.5	862.7	909.2	937.6	999.8
75%	363.2	405.6	413.4	428.0	453.7	463.4	459.0
50%	160.2	179.8	179.1	184.4	188.2	196.3	196.5
25%	68.0	78.4	77.9	77.3	79.3	79.8	81.1
10%	30.5	34.2	35.3	33.8	33.7	33.9	34.8
5%	19.8	22.4	22.8	22.8	22.7	22.6	22.7
1%	9.5	10.6	10.9	10.7	11.0	11.3	9.7
최소값	3.2	3.4	3.0	3.2	2.9	2.9	2.6
평균	303.2	335.0	350.5	361.0	381.7	394.3	410.2
CV	139.3%	138.0%	142.2%	141.6%	147.4%	149.7%	153.5%

해당 가중치를 사용하여 주요 변수인 장애유형, 경제활동상태에 대한 빈도 및 비율과 상대표준오차(CV) 값을 살펴본 결과는 <표 2-157> 및 <표 2-158>에 제시하였다.

<표 2-157> 장애유형 (15개) 별 경제활동상태(2개 구분) 빈도/비율 추정 및 상대표준오차 분석

장애유형	경제활동상태 (2개 구분)	가중빈도	비중(%)	표준오차(%)	상대표준오차 (%)
지체장애	취업자	301,117	41.1%	2.7%	6.5%
	미취업자	430,716	58.9%	2.9%	4.9%
	합계	731,833	100.0%		
뇌병변장애	취업자	88,241	74.4%	10.7%	14.4%
	미취업자	30,398	25.6%	5.6%	22.0%
	합계	118,639	100.0%		



<표 2-158> 장애유형 (15개) 별 경제활동상태(2개 구분) 빈도/비율 추정 및 상대표준오차 분석(계속)

시각장애	취업자	70,919	48.9%	5.8%	11.8%
	미취업자	74,030	51.1%	5.1%	10.0%
	합계	144,949	100.0%		
청각장애	취업자	50,689	45.1%	7.0%	15.5%
	미취업자	61,714	54.9%	7.8%	14.2%
	합계	112,403	100.0%		
언어장애	취업자	9,316	57.0%	31.2%	54.8%
	미취업자	7,026	43.0%	17.3%	40.2%
	합계	16,342	100.0%		
지적(정신지체)장애	취업자	130,005	70.9%	6.8%	9.6%
	미취업자	53,270	29.1%	4.8%	16.5%
	합계	183,274	100.0%		
자폐성(발달)장애	취업자	5,438	54.0%	19.4%	35.9%
	미취업자	4,626	46.0%	16.9%	36.7%
	합계	10,063	100.0%		
정신장애	취업자	68,957	82.8%	11.8%	14.2%
	미취업자	14,284	17.2%	5.4%	31.6%
	합계	83,241	100.0%		
신장장애	취업자	43,690	66.6%	12.5%	18.8%
	미취업자	21,893	33.4%	6.5%	19.4%
	합계	65,583	100.0%		
심장장애	취업자	1,254	28.3%	14.6%	51.8%
	미취업자	3,183	71.7%	55.3%	77.0%
	합계	4,437	100.0%		
호흡기장애	취업자	1,492	39.4%	18.6%	47.3%
	미취업자	2,296	60.6%	24.7%	40.8%
	합계	3,789	100.0%		
간장애	취업자	1,971	12.6%	5.8%	46.1%
	미취업자	13,721	87.4%	30.1%	34.4%
	합계	15,692	100.0%		
안면장애	취업자	10,473	86.3%	57.4%	66.5%
	미취업자	1,669	13.7%	7.2%	52.7%
	합계	12,142	100.0%		
장루요루장애	취업자	2,695	82.3%	41.8%	50.8%
	미취업자	580	17.7%	8.1%	45.7%
	합계	3,276	100.0%		
뇌전증(간질)장애	취업자	10,663	91.2%	28.9%	31.7%
	미취업자	1,032	8.8%	6.9%	78.6%
	합계	11,694	100.0%		

## 10. 8차조사 가중치 산출 및 추정

### 가. 종단 가중치 산출

#### 1) 차수별 응답 현황 및 단위 무응답 분석

8차년도 종단가중치는 1차부터 8차에 모두 응답한 3,396명을 대상으로 산출한다. 8차조사 종단가중치의 산출에 앞서 단위무응답의 현황을 살펴보고 응답성향분석을 통해 단위무응답에 영향을 미치는 조사대상자 특성을 우선적으로 파악하고자 하였다.

<표 2-159>은 장애인 거주 시도별 응답현황을 정리하였고 더불어 시도 분류와 응답여부와의 독립성 검정을 위한 카이제곱 통계와 크래머 V 통계를 산출하였다. 시도간 무응답률의 편차는 유의미하게 나타났다. 그렇지만 각 시도에서 무응답률은 제주 7.5%를 제외하고 모두 약 5% 미만으로 낮게 나타난다.

〈표 2-159〉 시도별 종단면 무응답 현황

시도	무응답		응답		계
	빈도	행%	빈도	행%	
서울	16	2.7%	570	97.3%	586
부산	7	2.0%	340	98.0%	347
대구	3	1.5%	193	98.5%	196
인천	0	0.0%	159	100.0%	159
광주	1	1.0%	103	99.0%	104
대전	0	0.0%	163	100.0%	163
울산	1	3.3%	29	96.7%	30
세종	0	0.0%	7	100.0%	7
경기	17	2.3%	723	97.7%	740
강원	1	0.6%	162	99.4%	163
충북	1	0.7%	140	99.3%	141
충남	0	0.0%	135	100.0%	135
전북	1	0.7%	147	99.3%	148
전남	1	0.6%	173	99.4%	174
경북	4	4.1%	93	95.9%	97
경남	3	1.3%	222	98.7%	225
제주	3	7.5%	37	92.5%	40
계	59	1.7%	3,396	98.3%	3,455

참고:  $\chi^2 = 30.124$ , P-value=0.017, Cramér's V=0.093

<표 2-160>은 장애인의 성별 응답현황을 보여주고 있다. 남성과 여성 간의 무응답률은 유의한 차이가 아니며, 남성의 무응답률이 다소 높게 나타난다.

<표 2-160> 성별 종단면 무응답 현황

구분	무응답		응답		계
	빈도	행%	빈도	행%	
남성	44	2.0%	2,194	98.0%	2,238
여성	15	1.2%	1,202	98.8%	1,217
계	59	1.7%	3,396	98.3%	3,455

참고:  $\chi^2 = 2.109$ , P-value=0.146, Cramér's V=0.025

<표 2-161>은 장애인의 연령대별 응답률 현황을 나타내고 있다. 연령대에 따른 무응답률의 편차는 유의한 차이가 아닌 것으로 나타난다. 30~39세 미만의 연령층의 무응답 비율이 가장 높고 40~49세 미만의 연령층의 무응답 비율이 가장 낮지만 전 연령층에서 무응답 비율은 약 1.5~2.5%로 유사하게 나타난다.

<표 2-161> 연령대별 종단면 무응답 현황

구분	무응답		응답		계
	빈도	행%	빈도	행%	
19-29세	6	1.9%	315	98.1%	321
30-39세	11	2.3%	465	97.7%	476
40-49세	14	1.5%	929	98.5%	943
50-70세	28	1.6%	1,687	98.4%	1,715
계	59	1.7%	3,396	98.3%	3,455

참고:  $\chi^2 = 1.419$ , P-value=0.701, Cramér's V=0.020

<표 2-162>은 장애인의 경제활동상태별 응답률의 현황을 정리하고 있다. 경제활동상태에 따른 무응답률의 편차는 유의한 차이가 아닌 것으로 나타난다. 실업자와 취업미희망자의 무응답률이 2%로 가장 높게 나왔고, 그 외 유형인 취업희망자와 취업자의 무응답률은 각 1.2%, 1.5% 정도로 상대적으로 낮게 나타났다.

<표 2-162> 경제활동상태별 종단면 무응답 현황

구분	무응답		응답		계
	빈도	행%	빈도	행%	
취업자	27	1.5%	1,782	98.5%	1,809
실업자	1	2.0%	49	98.0%	50
취업희망자	2	1.2%	165	98.8%	167
취업미희망자	29	2.0%	1,400	98.0%	1,429
계	59	1.7%	3,396	98.3%	3,455

참고:  $\chi^2 = 1.664$ , P-value=0.645, Cramér's V=0.022

<표 2-163>은 장애인의 장애유형별 응답률의 현황을 보여준다. 장애유형에 따른 무응답률의 편차는 유의미하지 않게 나타났다. 시각 외 감각장애의 무응답률이 가장 높고 지체 외 신체 외부장애의 무응답률이 가장 작게 나타났다. 그 외 장애유형의 무응답률은 비슷하게 나타났다.

<표 2-163> 장애유형별 종단면 무응답 현황

구분	무응답		응답		계
	빈도	행%	빈도	행%	
지체장애	26	1.5%	1,691	98.5%	1,717
지체 외 신체 외부장애	3	1.2%	255	98.8%	258
시각장애	7	1.6%	442	98.4%	449
시각외감각장애	10	3.2%	302	96.8%	312
정신적장애	8	1.7%	451	98.3%	459
신체내부장애	5	1.9%	255	98.1%	260
계	59	1.7%	3,396	98.3%	3,455

참고:  $\chi^2 = 5.141$ , P-value=0.399, Cramér's V=0.039

<표 2-164>는 장애인의 장애정도별 응답률 현황이다. 장애정도에 따른 무응답률의 편차는 유의미하지 않은 것으로 나타났다. 장애정도에 따른 무응답률을 살펴보면 중증과 경증의 무응답률은 거의 비슷하게 나타남을 알 수 있다.

<표 2-164> 장애정도별 종단면 무응답 현황

구분	무응답		응답		계
	빈도	행%	빈도	행%	
중증	24	1.8%	1,318	98.2%	1,342
경증	35	1.7%	2,078	98.3%	2,113
계	59	1.7%	3,396	98.3%	3,455

참고:  $\chi^2 = 0.025$ , P-value=0.875, Cramér's V=0.003

## 2) 응답성향모형 적합

앞서 인구사회학적 특성과 장애유형 및 장애정도 특성별 무응답의 연관성 분석을 토대로 8차년도 종단가중치를 산출하기 위해서는 조사대상자들의 다양한 속성(주로 인구사회학적 특성)을 반영한 응답성향모형을 적합하였다. 응답성향모형은 8차년도 응답여부는 과거에 관측된 보조변수들로 설명할 수 있다는 임의무응답(missing at random, 이하 MAR)을 가정한다. 따라서 임의무응답의 가정 하에서 8차년도 응답여부는 2차 조사부터 7차조사에서 모두 응답한 패널의 응답성향모형은 7차년도의 주요변수들로 설명할 수 있게 된다.

응답성향모형은 로지스틱 회귀모형을 사용하였다. 추정된 모형을 통해 조사대상자 각각의 응답성향점수(혹은 응답확률)를 계산하고 그 역수를 통해 무응답에 대한 보정계수로 산출하여 7차년도 종단가중치에 곱한 후 이를 8차년도 종단가중치를 계산하였다. 즉, 8차조사의 표본응답자가 응답했는지 여부를 반응변수로 하고,

7차조사의 표본응답자의 특성변수를 설명변수로 하는 로지스틱회귀모형을 적합하여 응답률의 차이를 파악하고 이를 통해 보정하였다. 응답확률은 로지스틱회귀모형 하에서 다음과 같이 표현 할 수 있다.

$$\hat{p}_i = \frac{\exp\left(\sum_{j=0}^k \hat{\beta}_j x_{ij}\right)}{1 + \exp\left(\sum_{j=0}^k \hat{\beta}_j x_{ij}\right)}$$

여기서  $\hat{p}_i$ 는  $i$ 번째 패널개체에 대한 응답확률의 추정값이고  $x_{ij}$ 는  $i$ 번째 패널개체의  $j$ 번째 보조변수 특성이며  $\hat{\beta}_j$ 는 회귀계수를 각각 나타낸다.

### 3) 최종 종단 가중치 산출

8차년도 종단면 응답자  $i$ 에 대한 종단가중치는 7차년도 종단가중치에 앞서 계산한 응답확률  $\hat{p}_i$ 를 나눠 다음과 같이 계산하였다.

$$\text{8차년도 종단가중치} = \text{7차년도 종단가중치} \times \frac{1}{\hat{p}_i}$$

여기서 사용된 설명변수는 성별, 연령, 지역 등의 변수와 장애유형, 장애정도, 경제활동상태 등의 변수들이다.

로지스틱 회귀모형에 대하여 초기 모형과 최종 모형에 대한 적합결과를 정리하여 <표 2-165>에 제시하였다. 앞서 고려한 모든 변수를 사용한 초기 모형 보다는 지역 및 장애유형을 설명 변수로 사용한 최종 모형이 적합도(AIC) 측면에서 더 좋음을 알 수 있다. 장애유형 중 지체장애 대비 감각장애 효과가 유의하여 장애유형을 모형에 포함 시켰다.

최종모형을 기반으로 산출한 종단가중치와 모든 변수를 사용한 초기 모형을 기반으로 산출한 종단가중치를 비교해보면 두 가중치는 매우 유사한 값을 지니며 두 가중치의 상관계수는 0.99 이상으로 나타난다. 따라서 연령 및 경제활동상태와 장애정도는 종단가중치 산출에 있어서 유의미한 차이를 만들지 않는다고 판단된다.

7차조사에서와 유사하게 8차조사의 응답 성향 모형에서 유의한 설명 변수가 지역 하나로 나오는 것은 다른 차수 조사에 비하여 무응답자 수와 비율이 현저히 작아서 불균등 자료에서 발생하는 특별한 상황이라고 생각된다. 추후 8차 이후의 조사에서도 이러한 현상이 계속 발생된다면 로지스틱 회귀와 같은 분포 가정 하의 응답성향 모형 보다는 비모수적 방법을 고민할 필요가 있다.

〈표 2-165〉 모형별 종단면 로지스틱 회귀분석 결과

설명변수	자유도	포화모형		최종모형	
		$\chi^2$	유의확률	$\chi^2$	유의확률
성별	1	2.660	0.103		
연령대	3	1.314	0.726		
지역	16	33.776	0.006	34.094	0.000
경제활동상태	3	3.263	0.353		
장애유형	5	5.293	0.381	4.507	0.479
장애정도	1	0.081	0.776		
		AIC = 610.864		AIC = 602.651	

〈표 2-166〉은 8차년도 종단가중치에 대한 기술 통계량을 보여주고 있다. 종단면 응답자 총 3,396명의 종단가중치 평균은 391.3이고 최솟값 4.5에서 최댓값 5,838.8의 범위를 갖는다.

〈표 2-167〉은 4~8차 조사까지의 종단가중치 분포의 현황을 정리하고 있다. 4차년도 종단가중치에 비하여 8차년도의 종단가중치는 전반적으로 증가하였지만 변동계수(CV)는 4차년도에 비하여 다소 줄어들었음을 알 수 있다.

〈표 2-166〉 8차년도 종단 가중치에 대한 통계량

통계량			
응답수	3,396	합계	1,328,783
평균	391.3	표준편차	522.2
중위수	214.5	분산	272,671.8
사분위 범위	380.8	범위	5,834.3
최솟값	4.5	최댓값	5,838.8

〈표 2-167〉 4-8차조사의 종단 가중치 분포

통계량	4차년도 종단가중치	5차년도 종단가중치	6차년도 종단가중치	7차년도 종단가중치	8차년도 종단가중치
최대값	5,186.0	5,365.7	5,421.6	5,621.5	5,838.8
99%	2,467.8	2,518.6	2,506.4	2,541.9	2,637.4
95%	1,250.0	1,289.0	1,314.8	1,328.1	1,361.0
90%	847.2	874.7	897.5	917.0	921.6
75%	440.0	453.5	462.0	470.0	474.1
50%	196.2	203.7	208.5	212.3	214.5
25%	84.3	88.3	90.2	92.2	93.3
10%	37.5	39.6	40.6	41.2	41.7
5%	25.2	26.6	26.9	27.7	27.9
1%	11.9	12.8	13.2	13.4	13.7
최소값	4.1	4.3	4.3	4.4	4.5
평균	360.1	371.1	378.7	385.5	391.3
CV	134.0%	132.3%	131.5%	132.0%	133.5%

## 나. 횡단 가중치 산출

### 1) 단위무응답 분석

8차년도 횡단가중치는 8차에 응답한 3,736명을 대상으로 산출한다. 가중치 산출을 위하여 8차년도 응답자 3,736명과 무응답자는 664명을 대상으로 무응답 특성을 분석한다. 8차년도 응답여부와 8차년도 응답자 모두가 응답한 1차년도의 주요 특성 변수들 간의 연관성을 살펴봄으로써 응답 여부에 중요하게 영향을 미치는 특성 변수를 파악하고자 하였다.

<표 2-168>는 시도별 무응답률 현황을 보여주고 있다. 무응답률의 시도 차이는 유의미하며, 서울, 대구, 인천, 광주, 울산, 경기, 강원, 충북, 경북, 제주 지역에서 무응답율이 약 10.0~21.0%로 높게 나타나며, 부산, 대전, 충남, 전북, 전남, 경남 지역에서 무응답 비율이 약 10% 미만으로 상대적으로 낮았다.

<표 2-168> 시도별 횡단면 무응답 현황

시도	무응답		응답		계
	빈도	행%	빈도	행%	
서울	158	18.4%	699	81.6%	857
부산	36	9.0%	364	91.0%	400
대구	34	13.5%	218	86.5%	252
인천	34	14.8%	196	85.2%	230
광주	13	10.7%	109	89.3%	122
대전	13	7.4%	162	92.6%	175
울산	6	17.6%	28	82.4%	34
경기	224	21.8%	802	78.2%	1,026
강원	21	11.0%	170	89.0%	191
충북	23	13.6%	146	86.4%	169
충남	15	9.5%	143	90.5%	158
전북	14	8.3%	155	91.7%	169
전남	16	8.0%	185	92.0%	201
경북	22	18.6%	96	81.4%	118
경남	15	6.4%	220	93.6%	235
제주	20	31.7%	43	68.3%	63
계	664	15.1%	3,736	84.9%	4,400

참고:  $\chi^2 = 115.490$ , P-value=0.000, Cramér's V=0.162

<표 2-169>은 성별 무응답률의 현황이다. 무응답률의 성별 차이는 유의미하지 않으며, 남성의 무응답률이 다소 높게 나타난다.

<표 2-169> 성별 횡단면 무응답 현황

구분	무응답		응답		계
	빈도	행%	빈도	행%	
남성	446	15.5%	2,438	84.5%	2,884
여성	218	14.4%	1,298	85.6%	1,516
계	664	15.1%	3,736	84.9%	4,400

참조:  $\chi^2 = 0.830$ , P-value=0.362, Cramér's V=0.014

<표 2-170>은 연령대별 무응답률의 현황이다. 응답 여부와 연령의 분포는 유의한 관계가 있다. 연령대에 따른 무응답 비율을 살펴보면 15~29세 미만의 연령층의 무응답률이 가장 높으며 50세 이상의 무응답률이 가장 낮았다.

<표 2-170> 연령대별 횡단면 무응답 현황

구분	무응답		응답		계
	빈도	행%	빈도	행%	
15-29세	147	18.8%	635	81.2%	782
30-39세	170	15.7%	913	84.3%	1,083
40-49세	205	16.1%	1,066	83.9%	1,271
50-64세	142	11.2%	1,122	88.8%	1,264
계	664	15.1%	3,736	84.9%	4,400

참조:  $\chi^2 = 24.439$ , P-value=0.000, Cramér's V=0.075

<표 2-171>은 경제활동상태에 따른 무응답 현황이다. 경제활동상태는 응답여부와 유의한 관계가 있는 것으로 나타났다. 실업자의 무응답률이 가장 높았고, 취업미희망자의 무응답률이 가장 낮았다.

<표 2-171> 경제활동상태별 횡단면 무응답 현황

구분	무응답		응답		계
	빈도	행%	빈도	행%	
취업자	326	14.6%	1,907	85.4%	2,233
실업자	31	18.9%	133	81.1%	164
취업희망자	183	17.1%	888	82.9%	1,071
취업미희망자	124	13.3%	808	86.7%	932
계	664	15.1%	3,736	84.9%	4,400

참조:  $\chi^2 = 7.931$ , P-value=0.047, Cramér's V=0.042



<표 2-172>은 장애유형별 무응답률의 현황을 보여준다. 무응답률의 장애유형별 차이는 유의미하지 않으며, 모든 장애유형에 대해서 약 14.0~16.0%로 비슷한 무응답률을 보이고 있다.

<표 2-172> 장애유형별 횡단면 무응답 현황

구분	무응답		응답		계
	빈도	행%	빈도	행%	
지체장애	328	14.9%	1,867	85.1%	2,195
지체 외 신체 외부장애	45	14.2%	271	85.8%	316
시각장애	84	14.7%	489	85.3%	573
시각외감각장애	68	16.7%	338	83.3%	406
정신적장애	87	15.2%	487	84.8%	574
신체내부장애	52	15.5%	284	84.5%	336
계	664	15.1%	3,736	84.9%	4,400

참조:  $\chi^2 = 1.211$ , P-value=0.944, Cramér's V=0.017

<표 2-173>은 장애정도별 무응답률의 현황이다. 무응답률의 장애정도별 차이는 유의미하지 않으며, 장애 정도에 따른 무응답률을 살펴보면 중증의 무응답률이 경증의 무응답률보다 높게 나타난다.

<표 2-173> 장애정도별 횡단면 무응답 현황

구분	무응답		응답		계
	빈도	행%	빈도	행%	
중증	210	15.2%	1,170	84.8%	1,380
경증	454	15.0%	2,566	85.0%	3,020
계	664	15.1%	3,736	84.9%	4,400

참조:  $\chi^2 = 0.013$ , P-value=0.910, Cramér's V=0.002

## 2) 무응답 조정 가중치 산출

무응답 보정을 위해서는 조사대상자들의 다양한 속성(주로 인구사회학적 특성)을 반영한 응답성향모형을 적합하였다. 이를 위해 8차년도 횡단면 조사의 응답여부는 과거에 관측된 변수들로 설명할 수 있다는 임의 응답(MAR)을 가정하였고 1차년도 조사 기준의 설명변수는 물론 2차조사부터 7차조사까지 응답여부도 함께 다음의 확률을 고려한 로지스틱 모형을 적합하였다.

$$P(8차\ 응답여부 \mid 2, 3, 4, 5, 6, 7차\ 응답여부, 1차\ 설명변수)$$

즉, 반응변수로는 8차조사 응답여부를 나타내는 지시자이고 2차년도부터 7차년도 조사의 응답여부와 1차년도 기준의 성별, 연령, 지역, 장애유형, 장애정도, 경제활동상태 등의 변수를 설명변수로 하는 로지스틱회귀 모형을 다음의 형태로 적합하였다.

$$\hat{p}_i = \frac{\exp\left(\sum_{j=0}^k \hat{\beta}_j x_{ij}\right)}{1 + \exp\left(\sum_{j=0}^k \hat{\beta}_j x_{ij}\right)}$$

<표 2-174>은 8차년도에 응답확률 모형에서 모든 설명변수를 사용한 포화모형을 포함하여 유의하지 않은 설명변수를 하나씩 제거하여 모형을 적합한 결과를 정리하고 있다. 모든 설명 변수를 사용하는 모형1 보다 는 성별, 연령대, 장애유형 및 장애정도와 2차, 3차 응답여부를 제외한 모형2가 적합도(AIC) 측면에서 더 좋은 결과를 알 수 있다. 따라서 최종적으로 선택한 모형으로는 4차 - 7차년도 응답여부, 지역, 경제활동상태를 설명 변수로 사용한 로지스틱 회귀 모형이다.

<표 2-174> 모형별 횡단면 로지스틱 회귀분석 결과

구분	모형1(초기모형)			모형2(최종모형)	
	df	$\chi^2$	유의 확률	$\chi^2$	유의 확률
7차 응답여부	1	2387.8	0.000	2387.8	0.000
6차 응답여부	1	39.0	0.000	39.0	0.000
5차 응답여부	1	2.9	0.088	2.9	0.088
4차 응답여부	1	13.4	0.000	13.4	0.000
3차 응답여부	1	0.0	0.873		
2차 응답여부	1	0.0	0.942		
성별	1	0.0	0.977		
연령대	3	0.4	0.942		
지역	15	48.2	0.000	48.3	0.000
경제활동상태	3	7.1	0.070	6.5	0.089
장애유형	5	1.6	0.901		
		AIC = 1301.260		AIC = 1281.789	

최종모형으로부터 추정된 응답확률을  $\hat{p}_i$ 로 정의하면 8차년도에 응답한  $i$ 번째 개체의 종단가중치는 다음과 같이 계산된다.

$$8\text{차년도 무응답 보정 가중치} = 1\text{차년도 최종가중치} \times \frac{1}{\hat{p}_i}$$

### 3) 벤치마킹 보정

8차년도 횡단가중치 산출을 위해 무응답을 보정한 가중치에 대하여 모집단 분포에 대한 벤치마킹 보정을 사용한다. 성별, 연령, 지역, 장애유형, 장애정도, 경제활동상태 변수를 통하여 이들의 모집단 분포를 이용한 가중치 합 비율을 통해 벤치마킹 보정을 실시한다. 성, 연령, 지역, 장애유형, 장애정도에 관한 모집단 정보는 2022년 12월 등록기준 22세에서 71세까지의 보건복지부 장애인 등록현황을 사용하였다.

경제활동상태에 대한 모집단 정보는 2023년 「상반기 장애인경제활동실태조사」의 경제활동상태에 대한

추정 결과를 사용하였다. 이 추정결과의 모집단 수와 2022년 12월 기준 장애인 등록 현황의 모집단 수가 일치하지 않기 때문에 추정 결과를 직접 사용할 수는 없다. 이를 위해 2023년 「상반기 장애인경제활동실태조사」의 경제활동상태에 대한 비율 추정 결과를 2022년 12월 기준 장애인 등록 현황의 모집단 수에 곱하여 경제활동상태에 대한 모집단 분포를 정의하였다.

레이킹-비 조정 방법을 적용하여 가중치를 보정한 후 횡단가중치를 다음과 같이 산출하였다.

$$w_{ij}^{rak} = w_{ij} \exp(\lambda' x_{ij})$$

여기서  $x_{ij}$ 는 레이킹-비 조정 방법을 위해 사용된 보조변수들을 나타내는 지시변수 벡터이며,  $\lambda$ 는  $\sum w_{ij}^{rak} x_{ij} = t_x$ 의 해이며,  $t_x$ 는 각 레이킹-비 조정 방법에 사용된 벤치마킹 변수들의 범주에 해당하는 모집단 장애인 분포를 나타낸다. 다음 절에서 횡단가중치 극단값 조정을 통한 최종가중치를 사용한 표본 자료 분포와 모집단 분포 비교 결과를 살펴보기로 한다.

#### 4) 극단값 처리 및 최종가중치 산출

앞서 산출한 레이킹-비 조정 횡단가중치에는 설계 가중치의 최댓값이었던 약 4,500을 넘는 가중치들이 존재하였다. 이들은 조사변수에 곱해져 합산되므로 경우에 따라서 통계량에 과도한 영향력을 갖게 된다. 따라서 이러한 과도한 가중치 값을 추가로 조정해 주기 위해 4,500으로 절사하여 조정해 준 다음, 앞서 고려한 바와 같이 성, 연령, 장애유형, 장애정도의 보조변수들에 대해서 다시 한 번 레이킹-비 보정을 실시하여 이를 최종 가중치로 산출하였다.

<표 2-175>는 횡단면 최종 가중치에 대한 기술통계량을 나타내고 있다. 평균 420.3이며 최솟값은 1.0이고 최댓값은 4,400.2이다. 최대값은 이전 7차조사까지의 횡단가중치에 비하여 크게 차이나지 않음을 알 수 있다.

<표 2-175> 횡단 가중치에 대한 통계량

통계량			
응답수	3,736	합계	1,570,068
평균	420.3	표준편차	591.8
중위수	215.6	분산	336,739.5
사분위 범위	410.7	범위	4,399.1
최솟값	1.0	최댓값	4,400.2

벤치마킹 보정을 위해 고려하였던 보조변수인 성별, 연령대, 지역, 장애유형, 장애정도, 경제활동상태 변수의 모집단 총합과 표본가중합의 분포는 <표 2-176>부터 <표 2-181>에 각각 정리하고 있다. 횡단면 최종 가중치를 사용한 표본가중합이 모집단 총합과 거의 차이가 없음을 알 수 있다.

<표 2-176> 성별 모집단 및 표본 분포

성	표본가중합		모집단 총합	
	인구	비중	인구	비중
남자	1,030,546	65.6%	1,030,546	65.6%
여자	539,522	34.4%	539,522	34.4%
총합	1,570,068	100.0%	1,570,068	100.0%

<표 2-177> 연령별 모집단 및 표본 분포

연령	표본가중합		모집단 총합	
	인구	비중	인구	비중
22~29세	80,304	5.1%	80,304	5.1%
30~39세	112,321	7.2%	112,321	7.2%
40~49세	220,897	14.1%	220,897	14.1%
50~71세	1,156,545	73.7%	1,156,546	73.7%
총합	1,570,068	100.0%	1,570,068	100.0%

<표 2-178> 경제활동상태 변수의 모집단 및 표본 분포

경제활동상태	표본가중합		모집단 총합	
	인구	비중	인구	비중
취업자	746,118	47.5%	746,118	47.5%
미취업자	30,039	1.9%	30,039	1.9%
취업희망자	104,953	6.7%	104,953	6.7%
취업미희망자	688,959	43.9%	688,959	43.9%
총합	1,570,068	100.0%	1,570,068	100.0%

<표 2-179> 장애유형별 변수의 모집단 및 표본 분포

장애유형	표본가중합		모집단 총합	
	인구	비중	인구	비중
지체장애	734,290	46.8%	734,290	46.8%
지체 외 신체 외부장애	144,710	9.2%	144,710	9.2%
시각장애	144,485	9.2%	144,485	9.2%
시각 외 감각장애	145,401	9.3%	145,401	9.3%
정신적 장애	281,827	17.9%	281,827	17.9%
신체 내부장애	119,356	7.6%	119,356	7.6%
총합	1,570,068	100.0%	1,570,068	100.0%

<표 2-180> 시도별 모집단 및 표본 분포

지역	표본가중합		모집단 총합	
	인구	비중	인구	비중
서울	227,727	14.5%	226,723	14.4%
부산	107,853	6.9%	107,474	6.8%
대구	82,653	5.3%	82,309	5.2%
인천	95,993	6.1%	97,068	6.2%
광주	44,456	2.8%	44,306	2.8%
대전	48,605	3.1%	48,440	3.1%
울산	31,718	2.0%	31,602	2.0%
세종	6,694	0.4%	6,647	0.4%
경기	376,151	24.0%	376,452	24.0%
강원	52,903	3.4%	52,738	3.4%
충북	52,584	3.3%	52,378	3.3%
충남	76,766	4.9%	76,507	4.9%
전북	71,597	4.6%	71,325	4.5%
전남	66,170	4.2%	65,899	4.2%
경북	97,442	6.2%	99,872	6.4%
경남	108,863	6.9%	108,517	6.9%
제주	21,894	1.4%	21,811	1.4%
총합	1,570,068	100.0%	1,570,068	100.0%

<표 2-181> 장애정도별 변수의 모집단 및 표본 분포

장애정도	표본가중합		모집단 총합	
	인구	비중	인구	비중
중증	667,976	42.5%	667,976	42.5%
경증	902,092	57.5%	902,092	57.5%
총합	1,570,068	100.0%	1,570,068	100.0%

<표 2-182>은 1~8차조사들의 횡단가중치의 기술통계량 및 분포를 비교하고 있다. 1차년도 최종 가중치에 비하여 8차년도의 횡단가중치는 전반적으로 증가하였다. 극단값 조정을 통하여 극단 가중치를 조정하여 최대값이 이전 6차조사까지의 횡단가중치와 큰 차이가 없음을 알 수 있다.

<표 2-182> 1-8차년도 횡단가중치 분포

	1차년도	2차년도	3차년도	4차년도	5차년도	6차년도	7차년도	8차년도
최대값	4,487.5	4,756.8	4,941.0	4,536.3	4,644.4	4,989.6	5,653.5	4,400.2
99%	2,129.5	2,260.2	2,396.5	2,631.4	3,040.7	3,098.7	3,477.8	3,094.2
95%	1,054.3	1,166.0	1,220.5	1,264.9	1,363.2	1,417.9	1,508.5	1,512.6
90%	722.5	772.3	849.5	862.7	909.2	937.6	999.8	1,014.1
75%	363.2	405.6	413.4	428.0	453.7	463.4	459.0	505.2
50%	160.2	179.8	179.1	184.4	188.2	196.3	196.5	215.6
25%	68.0	78.4	77.9	77.3	79.3	79.8	81.1	94.5
10%	30.5	34.2	35.3	33.8	33.7	33.9	34.8	38.5
5%	19.8	22.4	22.8	22.8	22.7	22.6	22.7	23.0
1%	9.5	10.6	10.9	10.7	11.0	11.3	9.7	1.4
최소값	3.2	3.4	3.0	3.2	2.9	2.9	2.6	1.0
평균	303.2	335.0	350.5	361.0	381.7	394.3	410.2	420.3
CV	139.3%	138.0%	142.2%	141.6%	147.4%	149.7%	153.5%	138.1%

해당 가중치를 사용하여 주요 변수인 장애유형, 경제활동상태에 대한 빈도 및 비율과 상대표준오차(CV) 값을 살펴본 결과는 <표 2-183> 및 <표 2-184>에 제시하였다.

<표 2-183> 장애유형 (15개) 별 경제활동상태(2개 구분) 빈도/비율 추정 및 상대표준오차 분석

장애유형	경제활동상태 (2개 구분)	가중빈도	비중(%)	표준오차(%)	상대표준오차 (%)
지체장애	취업자	285,328	39.7%	2.7%	6.9%
	미취업자	433,820	60.3%	3.2%	5.4%
	합계	719,148	100.0%		
뇌병변장애	취업자	96,745	78.0%	10.8%	13.9%
	미취업자	27,311	22.0%	5.5%	24.9%
	합계	124,056	100.0%		

<표 2-184> 장애유형 (15개) 별 경제활동상태(2개 구분) 빈도/비율 추정 및 상대표준오차 분석(계속)

시각장애	취업자	69,904	46.5%	5.1%	11.0%
	미취업자	80,390	53.5%	5.4%	10.2%
	합계	150,294	100.0%		
청각장애	취업자	58,526	44.7%	6.9%	15.4%
	미취업자	72,310	55.3%	6.5%	11.8%
	합계	130,837	100.0%		
언어장애	취업자	12,337	72.3%	36.8%	50.9%
	미취업자	4,730	27.7%	8.7%	31.5%
	합계	17,066	100.0%		
지적(정신지체)장애	취업자	124,686	69.2%	7.1%	10.3%
	미취업자	55,385	30.8%	4.6%	15.0%
	합계	180,071	100.0%		
자폐성(발달)장애	취업자	7,321	70.5%	22.5%	31.8%
	미취업자	3,060	29.5%	11.2%	37.8%
	합계	10,381	100.0%		
정신장애	취업자	71,329	81.8%	12.2%	14.9%
	미취업자	15,825	18.2%	6.0%	33.2%
	합계	87,154	100.0%		
신장장애	취업자	40,773	61.1%	10.1%	16.5%
	미취업자	25,960	38.9%	7.8%	20.0%
	합계	66,733	100.0%		
심장장애	취업자	4,299	80.0%	61.2%	76.5%
	미취업자	1,074	20.0%	11.6%	58.2%
	합계	5,373	100.0%		
호흡기장애	취업자	2,162	53.6%	23.0%	43.0%
	미취업자	1,871	46.4%	23.1%	49.8%
	합계	4,033	100.0%		
간장애	취업자	2,394	12.0%	5.8%	48.3%
	미취업자	17,578	88.0%	31.5%	35.8%
	합계	19,972	100.0%		
안면장애	취업자	8,816	82.3%	57.1%	69.3%
	미취업자	1,897	17.7%	8.2%	46.6%
	합계	10,713	100.0%		
장루요루장애	취업자	3,345	79.5%	40.1%	50.4%
	미취업자	865	20.5%	9.6%	46.5%
	합계	4,209	100.0%		
뇌전증(간질)장애	취업자	11,803	87.9%	26.3%	29.9%
	미취업자	1,620	12.1%	9.2%	76.5%
	합계	13,423	100.0%		

## 제3장 자료의 주요 특성

### 1. 조사표 체계 및 구성

장애인고용패널조사 조사표는 총 11개의 파트로 이루어져 있으나, 6차조사 이후 코로나19 상황 지속으로 부가조사 파트 1개를 신설하였다. [A. 패널 기본정보] 파트는 인구통계학적 변수들과 장애정보, 현 장애상태 등 응답자의 기본적인 정보를 확인하며, [B. 경제활동상태 판별] 파트는 응답자의 지난조사 이후 일자리 현황을 파악하고 경제활동상태를 판별하는 문항으로 구성되어 있다. A, B파트의 주요 내용들은 1차조사 응답내용을 재확인하고, 이후의 변화여부를 파악하도록 구성하였다.

각 패널은 취업여부와 주업의 종사상 지위에 따라 [C-a 임금근로자], [C-b 비임금근로자], [C-c 미취업자] 3개 파트 중 본인의 경제활동 상태와 일치하는 하나의 영역을 응답한다. 응답 기준시점 당시 취업자의 경우 C파트에서 현재 근무 중인 일자리에 대한 전반적인 현황을 응답하며, 미취업자의 경우 취업을 희망하는 정도, 향후 희망하는 일자리, 구직활동 여부 등의 내용을 응답한다.

C파트에서 응답한 일자리 외에 부업이 있거나, 현재 지속하지 않더라도 지난조사 이후 직업활동 경험이 있는 패널의 경우, [JOB 직업력] 파트를 부가조사의 형태로 응답한다. 직업력 파트에서는 응답자가 경험한 일자리의 개괄적인 현황을 누적으로 응답하도록 하였다.

D파트부터 H파트는 모든 패널이 응답하며, 각 파트별로 취업을 위한 노력, 패널의 직업능력 현황, 고용과 관련된 패널의 특성과 주변 환경, 일상생활, 소득 및 가구 현황을 순차적으로 응답한다. 코로나19 상황 지속으로 한시 운영 예정인 I파트는 코로나19로 인한 경험, 인식 및 변화에 대해 응답하도록 하였다.

응답 시 각 파트의 주요 문항은 TAPI 프로그램을 통해 이전조사 당시의 응답내용을 참고용으로 확인할 수 있도록 제시하여, 응답내용의 시계열적 안정성을 제고하고 응답자의 편의성을 확보하도록 하였다.

전체 조사표의 구성은 다음과 같다.

[그림 3-1] 조사표 구성도





〈표 3-1〉 차수별 조사표 구성

구 분	조사내용	조사차수							
		8차	7차	6차	5차	4차	3차	2차	1차
패널 기본 정보 확인(A)	- 패널의 인적사항 및 장애정보 재확인, 지난조사 이후의 변화내역 파악 - 인적특성, 장애정보, 등록여부, 장애상태, 보장구 등	○	○	○	○	○	○	○	○
지난조사 이후 일자리 현황(PB)	- 지난조사 이후 계속 하고 있는 일자리, 새로 시작한 일자리, 그만둔 일자리 등 주업 및 직업력 영역에서 조사해야 할 조사표의 수와 종류를 결정함 - 지난조사 일자리 지속 여부, 변화내역, 지난조사 이후 새로 시작한 일자리 유무 및 지속여부, 주업/부업 등	○	○	○	○	○	○	○	X
경제활동상태 판별(B)	- 조사시점의 취업자, 실업자, 비경제활동인구를 구분함 - 지난 주 일자리 유무, 지난 주 또는 지난 달 구직여부, 일할 능력 유무, 취업자의 일 관련 희망내역 등	○	○	○	○	○	○	○	○
임금근로자(Ca)	- PB파트에서 응답한 현재 주업인 일자리의 종사상 지위가 임금근로자(상용근로자, 임시근로자, 일용근로자)인 경우 일자리 특성 파악 - 일자리 기본정보, 취업현황, 일자리 주요 특성 등	○	○	○	○	○	○	○	○
비임금근로자(Cb)	- PB파트에서 응답한 현재 주업인 일자리의 종사상 지위가 자영업자 혹은 무급가족종사자인 경우 일자리 특성 파악 - 일자리 기본정보, 취(창)업현황, 일자리 주요 특성, 근무시간 등	○	○	○	○	○	○	○	○
미취업자(Cc)	- PB파트에서 현재 지속중인 일자리가 없는 경우 향후 취업의향과 구직활동 여부 등을 파악 - 과거 취업경험, 향후 취업의사, 구직활동 등	○	○	○	○	○	○	○	○
취업을 위한 노력과 지원(D)	- 취업알선기관 이용, 직업교육훈련 등 취업과 관련된 주요 활동	○	○	○	○	○	○	○	○
직업적 능력(E)	- 취업관련 업무능력 및 업무경력, 자격증, 취업관련 사회적지지 정도 등 직업적 능력 전반 확인	○	○	○	○	○	○	○	○
취업 태도 및 환경(F)	- 취업 및 스스로에 대한 생각과 가족의 취업지지 정도, 장애인 고용정책에 대한 인식 등 취업관련 환경적 요인 파악	○	○	○	○	○	○	○	○
일상생활과 삶의 질(G)	- 건강상태 및 생활습관, 일상생활 시 도움 필요 여부 및 도움제공자, 일상 및 여가활동, 차별경험 등 일상생활 관련 문항	○	○	○	○	○	○	○	○
가구 일반사항(H)	- 가구원 현황, 소득 및 지출 현황, 자산 및 부채, 주거환경, 노후 대비 등 전반적인 가구 현황 파악	○	○	○	○	○	○	○	○
직업력 부가조사(JOB)	- 일자리별로 지난조사 당시 응답여부, 종사상 지위, 현재 지속 여부 등	○	○	○	○	○	○	○	X
코로나(I)	- 코로나19 백신 접종, 코로나19 경험, 전후 변화 사항, 삶의 질 변화 코로나19 관련 지원 필요 사항 등	○	○	○	X	X	X	X	X

## 2. 주요 설문항목 구성

장애인고용패널조사는 개인의 경제활동과 노동시장의 동태적 변화를 관찰하기 위해 설계된 조사이므로 기본적으로 매년 동일한 내용의 설문이 반복되는 형식을 갖추고 있다. 시계열적 분석을 위하여 차수 간 조사 문항의 변화는 최소화하였으며, 일자리 관련 주요 문항은 지난 조사 시 응답한 내용을 확인하고 지난 조사 이후의 변화를 구체적으로 질문하도록 구성되어 있다.

다만 매 조사마다 문항의 적절성을 검토하고, 장애인고용 관련 최신 이슈 등을 반영하기 위해 부분적으로 조사문항의 수정 및 보완이 적용되었다. 따라서 조사 차수별로 조사되지 않았거나 내용이 수정된 문항이 있다. 각 파트별 세부 설문 구성과 조사차수별 문항 포함 여부는 다음과 같다.

### A. 패널 기본정보

- 지난조사 시 응답받은 패널 기본 인적사항 및 장애정보에 대해 지난 조사 결과를 확인 및 수정하고 지난 조사 이후 변동사항이 있는지를 확인한다.
- 지난조사 이후 장애등록이 취소된 경우에는 취소시점과 그 이유를 확인한 후 조사를 종료하며, 패널에서 제외한다.
- 5차조사에서는 이원화된 장애등급제를 문항에 반영하였고 특히 장애를 판별할 수 있는 문항을 한시적으로 추가하였다.

〈표 3-2〉 설문 항목 구성 - A. 패널 기본정보

영역	세부 영역 및 항목	8차	7차	6차	5차	4차	3차	2차	1차
패널 기본정보 확인 (지난조사 응답내용 확인)	성명, 성별, 생년월일	○	○	○	○	○	○	○	○
	혼인상태	○	○	○	○	○	○	○	○
	학력사항(최종학력, 특수교육, 전공과)	○	○	○	○	○	○	○	○
	장애정보(장애유형, 등급, 부위, 중복장애)	○	○	○	○	○	○	○	○
주장애 관련 사항	지난조사 대비 장애정보 변화시기 및 이유	○	○	○	○	○	○	○	×
	장애유형별 특성	○	○	○	○	○	○	○	○
	장애유형별 특성(지적·자폐성 장애) <sup>1)</sup> - 한글 독해 가능여부 및 대중교통수단 이용가능 정도	○	○	○	○	○	○	○	×
	주장애 발생시기 및 인지시기	○	○	○	○	○	○	○	○
	최초 장애등록시기	×	×	×	×	×	×	×	○
	주장애 발생원인	○	○	○	○	○	○	○	○
	장애로 인한 퇴직여부	○	○	○	○	○	○	○	○
	주장애의 현재 상태	○	○	○	○	○	○	○	○
장애 판별 문항	안경/보청기/ 사용 시 어려움	×	×	×	○	×	×	×	×
	걸거나 계단을 오르는데 어려움	×	×	×	○	×	×	×	×
	기억하거나 집중하는데 어려움	×	×	×	○	×	×	×	×
	몸을 씻거나 옷을 입는데 어려움	×	×	×	○	×	×	×	×
	의사소통에 어려움	×	×	×	○	×	×	×	×

1) 지적·자폐성 장애의 한글 독해 가능여부 및 대중교통수단 이용가능 정도는 2차조사에 신규 추가됨

**PB. 지난조사 이후 일자리 현황**

- 지난조사 당시 응답한 일자리 중 현재 지속하고 있는 일자리, 그만둔 일자리, 지난조사 이후 새로 시작한 일자리 현황을 파악하여 Ca, Cb, Cc 조사표 및 직업력 조사표 중 응답받을 조사표 종류와 개수를 결정한다. 응답자는 PB파트에서 응답한 현재 취업여부 및 주업인 일자리의 종사상 지위에 따라 Ca(임금근로자), Cb(비임금근로자), Cc(미취업자) 파트 중 하나를 응답하며, 그만둔 일자리 또는 주업이 아닌 일자리가 있는 경우 직업력 파트에서 해당 일자리 내용을 응답하게 된다.
- 지난조사 시 파악한 일자리 각각에 대해 당시 응답한 정보가 지난 조사시점을 기준으로 맞는지 확인하고, 현재 조사시점을 기준으로 변화된 내용이 없는지, 해당 일자리를 지금도 계속하고 있는지 등에 대해 일자리별로 파악한다.
- 지난 조사 이후 새로 시작해서 1주일 이상 지속한 일자리가 있거나, 1주일 이상 지속하지는 않았지만 지난조사 이후 새로 시작해서 지난주에 1시간이라도 일한 적이 있는 일자리가 있는 경우 신규 일자리가 있는 것으로 간주하여 해당 일자리 정보를 응답한다.

〈표 3-3〉 설문 항목 구성 - PB. 지난조사 이후 일자리 현황

영역	세부 영역 및 항목	8차	7차	6차	5차	4차	3차	2차	1차
일자리 현황 파악	지난조사 일자리 정보 확인 및 수정	○	○	○	○	○	○	○	×
	지난조사 이후 변화사항	○	○	○	○	○	○	○	×
	지난조사 이후 새로 시작한 일자리 정보	○	○	○	○	○	○	○	×
취업자	주업의 종사상 지위	○	○	○	○	○	○	○	×
	사업내용, 주로 하는 일, 주 사용장비나 도구*	○	○	○	○	○	○	○	×
	지난주 부업여부 및 개수	○	○	○	○	○	○	○	×
	부업을 하는 이유	○	○	○	○	○	○	○	×

**B. 경제활동상태 판별**

- 경제활동상태 판별 파트에서는 통계청 경제활동인구조사에서의 판별 기준을 적절하게 변용하여 패널의 경제활동상태(취업자, 실업자, 비경제활동인구)를 파악한다.
- 패널이 유지하고 있는 일자리 각각에 대한 종사상 지위와 주업 이외에 가지고 있는 일자리(부업) 현황은 1차조사에 한하여 B파트에서 응답하였으며, 2차조사 이후에는 PB파트에서 모든 일자리 현황을 응답한다.

〈표 3-4〉 설문 항목 구성 - B. 경제활동상태 판별

영역	세부 영역 및 항목	8차	7차	6차	5차	4차	3차	2차	1차
경제활동상태 판별	지난주 취업여부, 일시휴직 여부 및 이유	○	○	○	○	○	○	○	○
	지난주 일을 하지 않은 이유	○	○	○	○	○	○	○	○
	지난주 일시휴직에 코로나지속의 영향도	x	x	x	○	x	x	x	x

〈표 3-5〉 설문 항목 구성 - B. 경제활동상태 판별(계속)

영역	세부 영역 및 항목	8차	7차	6차	5차	4차	3차	2차	1차
취업자	평소(지난주) 36시간 미만 근로여부 및 이유	○	○	○	○	○	○	○	○
	지난주 실제 근로시간	○	○	○	○	○	○	○	○
	추가근로 또는 이직 희망여부, 가능성 및 이직준비 여부	○	○	○	○	○	○	○	○
	주업의 종사상 지위	×	×	×	×	×	×	×	○
	지난주 부업여부 및 개수	×	×	×	×	×	×	×	○
	부업을 하는 이유	×	×	×	×	×	×	×	○
실업자 및 비경제활동인구	지난 4주간 구직활동 여부	○	○	○	○	○	○	○	○
	지난 4주간 구직활동을 하지 않은 이유에 코로나19 지속의 영향도	○	○	○	○	×	×	×	×
	지난주 근로 가능여부 및 취업 희망여부	○	○	○	○	○	○	○	○
	지난주 취업을 희망하지 않은 이유에 코로나19 지속의 영향도	○	○	○	○	×	×	×	×
	지난주 주된 활동	○	○	○	○	○	○	○	○
경제활동상태	월별 주된 경제활동상태	○	○	○	○	○	○	○	○

### Ca. 임금근로자

- PB파트에서 응답한 현재 주업인 일자리의 종사상 지위가 임금근로자(상용근로자, 임시근로자, 일용근로자)인 경우 Ca. 임금근로자 파트에서 일자리 특성을 파악한다.
- 2차조사 이후부터는 취업현황 문항(근무시작시기, 일자리 선택 이유, 채용정보 습득경로, 실질적 취업경로, 공공 취업알선기관 도움 정도, 취업에 도움이 된 사람 및 요소 등)을 해당 일자리가 지난 조사 이후 새로 시작한 일자리인 경우에만 질의한다.
- 근무환경 및 차별경험 문항은 2차조사 이후부터는 지난조사 이후 경험에 대해 응답한다.
- 3차조사에는 임금근로자 중 비정규직 응답자를 대상으로 현 일자리 내 동일(유사)업무 수행 정규직 유무, 전환 가능여부 및 방법 문항이 새로 도입되었으며, 정규직 대비 업무수준의 차이 및 비정규직 처우개선 필요 분야 등의 문항이 추가되었다.
- 4차조사에서는 노동시장 이슈(최저임금)를 반영하여 최저임금 관련 문항이 추가되었고, 장애인 차별금지법에 따른 정당한 편의제공을 하고 있는지 내부 수요를 반영한 문항과 중증장애인 고용유지를 위해 장애인 근로자가 필요한 사항을 추가적으로 조사하여 향후 고용 정책의 근거자료로 활용하는데 도움이 되도록 하였다.
- 5차조사에는 코로나19가 장애인 임금근로자의 근로 및 소득 등에 미치는 영향을 파악하기 위한 문항과 직장 내 장애인 인식개선교육 여부 및 방법, 효과 등을 묻는 문항을 추가하였다.
- 6차조사에는 코로나19 상황의 지속으로 인해 근무하면서 겪는 어려움, 마스크 착용 여부 등을 묻는 문항을 추가하였다.
- 7차조사에는 근무 시 발생할 수 있는 안전관련 위험요소, 재난 안전 대응을 위해 장애 관련 필요 요소 등을 묻는 문항을 추가하였다.

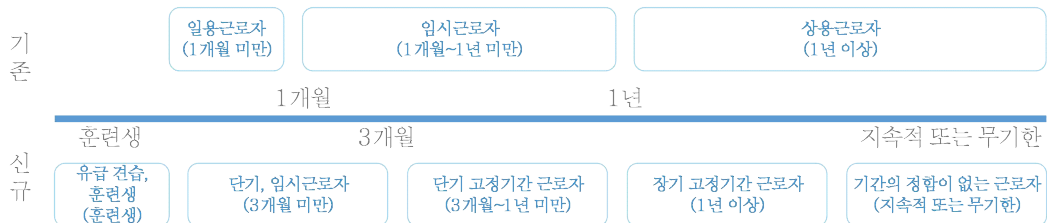
- 8차조사에서는 개정된 한국중사상지위분류의 반영을 위해 일부 문항을 추가하였으며, 일·가정 양립 측면에서 살펴볼 수 있도록 여가와 가사, 돌봄노동 시간 문항을 신설하였다.
- 통계청에서는 다양한 근로 형태(의존계약자 등)를 반영하기 위해 임금근로자와 자영업자 항목을 보다 세분화하여 개정하였다.

〈표 3-6〉 개정 한국중사상지위분류 체계(통계청, 2021.12.29., 보도자료)

지휘 권한 유형 (KCSE-A)	경제적 위험 유형* (KCSE-R)
일하는 사람과 일이 수행되는 경제단위 사이의 지휘 권한 관계에 따른 분류	어떤 일자리에서 일이 수행되는 조건에 수반되는 경제적 위험에 따른 분류
<b>◆ 독립 취업자</b> <b>A 고용주</b> 11 법인 고용주 12 개인기업 고용주 <b>B 고용원이 없는 자영업자</b> 21 고용원이 없는 법인 자영업자 22 고용원이 없는 개인기업 자영업자	<b>◆ 이윤 목적 취업자</b> <b>F 개인기업 독립 취업자</b> 12 개인기업 고용주 22 고용원이 없는 개인기업 자영업자 <b>C 의존 계약자</b> 30 의존 계약자 <b>E 무급 가족 종사자</b> 50 무급 가족 종사자
<b>◆ 의존 취업자</b> <b>C 의존 계약자</b> 30 의존 계약자 <b>D 임금 근로자</b> 41 기간의 정함이 없는 근로자 42 장기 고정 기간 근로자(1년 이상) 43 단기 고정 기간 근로자(3개월 ~ 1년 미만) 44 단기·임시 근로자(3개월 미만) 45 유급 견습, 훈련생 및 인턴 <b>E 무급 가족 종사자</b> 50 무급 가족 종사자	<b>◆ 임금 목적 취업자</b> <b>G 법인 소유 경영자</b> 11 법인 고용주 21 고용원이 없는 법인 자영업자 <b>D 임금 근로자</b> 41 기간의 정함이 없는 근로자 42 장기 고정 기간 근로자(1년 이상) 43 단기 고정 기간 근로자(3개월 ~ 1년 미만) 44 단기·임시 근로자(3개월 미만) 45 유급 견습, 훈련생 및 인턴

- 우선적으로 임금근로자 개정안만 반영하여 임금근로자는 기존 분류(상용/임시/일용)와 신분류(기간의 정함이 없는 근로자/고정기간(장기, 단기) 근로자/ 단기·임시 근로자/유급 견습, 훈련생 및 인턴) 모두 사용이 가능하다.

[그림 3-2] 임금근로자 종사상지위 개정안 비교 - 기간별



〈표 3-7〉 설문 항목 구성 - Ca. 임금근로자

영역	세부 영역 및 항목	8차	7차	6차	5차	4차	3차	2차	1차
일자리 기본정보	사업체(사업장)명 및 소재지	○	○	○	○	○	○	○	○
	사업내용, 주로 하는 일, 주 사용장비나 도구*	×	×	×	×	×	×	×	○
	부서명 및 직책 또는 직위	○	○	○	○	○	○	○	○
취업현황	근무시작시기	○	○	○	○	○	○	○	○
	일자리 선택 이유	○	○	○	○	○	○	○	○
	채용정보 습득경로, 실질적 취업경로, 공공 취업알선기관 취업에 도움이 된 사람 및 요소	○	○	○	○	○	○	○	○
일자리 주요 특성	일자리 분류	○	○	○	○	○	○	○	○
	근로자 수, 장애인 근로자 수	○	○	○	○	○	○	○	○
	장애인 일자리 여부, 계절적/주기적 일자리 여부	○	○	○	○	○	○	○	○
고용형태	비정규직 여부 판별(한시적/시간제/파견·용역/특수형태/일일/가정 내 근로/자기판단식 비정규직)	○	○	○	○	○	○	○	○
	일자리 목적(유급전습, 훈련생 및 인턴) 및 구체적인 근로계약기간	○	×	×	×	×	×	×	×
	파견/용역근로 내용	○	○	○	○	○	○	○	○
	비정규직 근로 사유 및 장애의 영향	○	○	○	○	○	○	○	○
	현 일자리 내 동일(유사)업무 수행 정규직 유무	○	○	○	○	○	○	×	×
	현 일자리의 정규직 전환기회 제공여부 및 전환방법	○	○	○	○	○	○	×	×
	동일(유사)업무 수행 정규직 대비 본인의 업무수준	○	○	○	○	○	○	×	×
근무시간	비정규직 근로자 처우 개선을 위해 조치해야 할 사항	○	○	○	○	○	○	×	×
	정규 근무일 및 근무시간	○	○	○	○	○	○	○	○
	실질 근무일 및 근무시간	○	○	○	○	○	○	○	○
임금 및 복리후생	지난달 결근여부, 교대제 여부	○	○	○	○	○	○	○	○
	임금 산정 방식 및 월평균 임금	○	○	○	○	○	○	○	○
	체불임금 여부	○	○	○	○	○	○	○	○
업무특성	최저임금 인상에 따른 일자리(직장) 영향	×	×	×	×	○	×	×	×
	4대보험 가입여부, 복리후생제도 혜택여부	○	○	○	○	○	○	○	○
	업무 시작시기	○	○	○	○	○	○	○	○
근무환경 및 차별경험	업무에 필요한 능력 및 기술수준	○	○	○	○	○	○	○	○
	응답자 학력 대비 업무수준, 적성 및 전공 부합정도	○	○	○	○	○	○	○	○
	업무능력 활용도, 주업무에 대한 평소 생각 및 만족도	○	○	○	○	○	○	○	○
직장 내 장애인 인식개선 교육	일자리(직장)에서의 안전 정도, 산재 경험 여부	○	○	○	○	○	○	○	○
	안전관련 위험요소, 장애 관련 필요 사항	○	○	×	×	×	×	×	×
	장애편의시설 설치여부	○	○	○	○	○	○	○	○
에로사항, 이직의사 및 만족도	노동조합 존재여부 및 가입여부	○	○	○	○	○	○	○	○
	출퇴근 교통수단 및 소요시간	○	○	○	○	○	○	○	×
	일자리에서의 응답자 장애 인지여부	○	○	○	○	○	○	○	○
코로나19 관련 사항	일자리에서의 (장애로 인한) 차별경험 여부	○	○	○	○	○	○	○	○
	정당한 편의제공 요청 및 제공 경험	×	○	○	○	○	×	×	×
	직장 내 장애인 인식개선 교육 실시 여부 및 방법, 효과	×	○	○	○	×	×	×	×
부업	직장 내 장애인 인식개선 교육을 실시하지 않은 이유	×	○	○	○	×	×	×	×
	일자리에서의 어려움 및 장애로 인한 업무 지장 정도	○	○	○	○	○	○	○	○
	지속 근무 희망여부 및 이직준비 여부	○	○	○	○	○	○	○	○
부업	직장을 계속 다니기 위해 필요한 사항	○	○	○	○	○	×	×	×
	일자리에 대한 항목별 만족도, 전반적 만족도	○	○	○	○	○	○	○	○
	지난조사 대비 일자리의 변화사항	○	○	○	○	○	○	○	×
코로나19 관련 사항	코로나19 지속으로 인한 일자리(직장) 내 어려움	○	○	○	○	×	×	×	×
	코로나19 지속으로 인한 일자리(직장) 내 고용안정 변화	○	○	○	○	×	×	×	×
	코로나19 지속으로 인해 현 일자리(직장)에서 근무하면서 겪는 어려움	○	○	○	×	×	×	×	×
부업	근무 시 마스크 착용 여부	×	×	○	×	×	×	×	×
	일자리(직장)에서 코로나19 감염 위험으로부터 안전한 정도	×	×	○	×	×	×	×	×
부업	임금근로 부업 현황	×	×	×	×	×	×	○	

\* 사업내용, 주로 하는 일, 주 사용장비나 도구 등 일자리 기본정보는 1차조사 시 Ca파트에서 응답하였으며, 2차조사부터는 PB. 지난조사 이후 일자리 현황에서 해당 내용을 확인함

**Cb. 비임금근로자**

- PB파트에서 응답한 현재 주업인 일자리의 종사상 지위가 자영업자 혹은 무급가족종사자인 경우 Cb. 비임금근로자 파트에서 일자리 특성을 파악한다.
- 2차조사 이후부터는 취(창)업현황 문항(근무시작시기, 일자리 시작 이유, 창업과정 참여 여부(무급가족종사자), 창업 준비기간 및 투자비용 현황, 대출현황, 창업관련 정보 습득경로 및 창업과정의 어려움(자영업자) 등)을 해당 일자리가 지난조사 이후 새로 시작한 일자리인 경우에만 질의한다.
- 근무환경 문항은 2차조사 이후부터는 지난조사 이후 경험에 대해 응답한다.

〈표 3-8〉 설문 항목 구성 - Cb. 비임금근로자

영역	세부 영역 및 항목	8차	7차	6차	5차	4차	3차	2차	1차
일자리 기본정보	사업체(사업장)명 및 소재지	○	○	○	○	○	○	○	○
	사업내용, 주로 하는 일, 주 사용장비나 도구*	×	×	×	×	×	×	×	○
	사업주와의 관계(무급가족종사자)	○	○	○	○	○	○	○	○
취(창)업현황	근무시작시기	○	○	○	○	○	○	○	○
	일자리 시작 이유, 창업과정 참여 여부(무급가족종사자)	○	○	○	○	○	○	○	○
	창업 준비기간 및 투자비용 현황	○	○	○	○	○	○	○	○
	대출 및 융자상환 현황	×	×	×	×	×	○	○	○
	창업관련 정보 습득경로 및 창업과정의 어려움(자영업자)	○	○	○	○	○	○	○	○
일자리 주요 특성	일자리 분류, 사업자등록 여부, 일자리의 주된 장소	○	○	○	○	○	○	○	○
	유급 종업원 수, 무급 종사자 수	○	○	○	○	○	○	○	○
근무시간	규칙적 근로 여부(계절적 일자리 여부)	○	○	○	○	○	○	○	○
	정규 근무일 및 근무시간	○	○	○	○	○	○	○	○
	실질 근무일 및 근무시간	○	○	○	○	○	○	○	○
	지난달 결근여부	○	○	○	○	○	○	○	○
매출액 및 소득	연간 매출액	○	○	○	○	○	○	○	○
	소득유무 및 월평균 소득	○	○	○	○	○	○	○	○
업무특성	업무 시작시기	○	○	○	○	○	○	○	○
	업무에 필요한 능력 및 기술수준	○	○	○	○	○	○	○	○
	응답자 학력 대비 업무수준, 적성 및 전공 부합정도	○	○	○	○	○	○	○	○
	업무능력 활용도, 주업무에 대한 평소 생각 및 만족도	○	○	○	○	○	○	○	○
근무환경	일자리 안전 정도, 쾌적한 정도	○	○	○	○	○	○	○	○
	안전관련 위험요소, 장애 관련 필요 사항	○	○	×	×	×	×	×	×
	산재보험 가입여부	○	○	○	○	○	○	○	○
	장애편의시설 설치여부	○	○	○	○	○	○	○	○
	출퇴근 교통수단 및 소요시간	○	○	○	○	○	○	○	×
에로사항, 이직의사 및 만족도	일자리에서의 어려움 및 컨설팅 희망 여부	○	○	○	○	○	○	○	○
	장애로 인한 업무 지장 정도	○	○	○	○	○	○	○	○
	사업체 지속 희망여부 및 이직준비 여부	○	○	○	○	○	○	○	○
	일자리에 대한 항목별 만족도, 전반적 만족도	○	○	○	○	○	○	○	○
	사업체의 운영성과에 대한 평가	○	○	○	○	○	○	○	○
코로나19 관련 사항	코로나19 지속으로 인한 사업체 운영의 어려움	○	○	○	○	×	×	×	×
	코로나19 지속으로 인한 사업체 운영상황 변화	○	○	○	○	×	×	×	×
	자영업자 대상 지원금이 사업체를 운영하는 데 도움이 되는 정도	×	×	○	×	×	×	×	×
	코로나19 지속으로 인해 현 일자리(사업체)에서 근무하면서 겪는 어려움	○	○	○	×	×	×	×	×
	근무 시 마스크 착용 여부	×	×	○	×	×	×	×	×
일자리(사업체)에서 코로나19 감염 위험으로부터 안전한 정도	×	×	○	×	×	×	×	×	
부업	비임금근로 부업 현황	×	×	×	×	×	×	×	

\* 사업내용, 주로 하는 일, 주 사용장비나 도구 등 일자리 기본정보는 1차조사 시 Cb파트에서 응답하였으며, 2차조사부터는 PB. 지난조사 이후 일자리 현황에서 해당 내용을 확인함

## Cc. 미취업자

- PB파트에서 현재 지속중인 일자리가 없는 경우, Cc. 미취업자 파트에서 향후 취업 의향과 구직 활동 여부 등을 응답한다.
- 1차조사에 한하여 과거 취업 경험 여부 및 마지막 일자리 현황을 응답한다.
- 구직활동 문항은 1차조사의 경우 최근 3년간의 경험을, 2차조사 이후부터는 지난조사 이후 경험을 질의한다.

〈표 3-9〉 설문 항목 구성 - Cc. 미취업자

영역	세부 영역 및 항목	8차	7차	6차	5차	4차	3차	2차	1차
과거 취업경험	과거 취업경험 여부	×	×	×	×	×	×	×	○
	마지막 일자리 현황	×	×	×	×	×	×	×	○
	마지막 일자리 그만둔 이유 및 장애 영향 정도	×	×	×	×	×	×	×	○
향후 취업의사	향후(1년 이내) 취업의사	○	○	○	○	○	○	○	○
	향후 취업의사 및 취업을 희망하는 이유	○	○	○	○	○	○	○	○
	취업을 희망하는 일자리 특성	○	○	○	○	○	○	○	○
구직활동	구직활동 경험 여부	○	○	○	○	○	○	○	○
	구직활동 시작시기 및 구직기간, 구직활동 현황	○	○	○	○	○	○	○	○
	구직활동 관련 애로사항 및 장애 영향 정도	○	○	○	○	○	○	○	○
코로나19 관련 사항	코로나19 지속이 구직활동에 미친 영향	○	○	○	○	×	×	×	×
	향후 취업여건 전망	○	○	○	×	×	×	×	×

## D. 취업을 위한 노력

- 취업알선기관 이용이나 직업교육훈련 등 취업 및 취업유지 관련 활동, 취업활동 관련 지원사항, 장애로 인해 직업생활 중 추가로 소요되는 비용 등 취업과 관련된 주요 활동에 대한 내용으로 구성되었다.
- 취업 및 취업유지 관련 활동 문항은 1차조사의 경우 최근 3년간의 경험을, 2차조사 이후부터는 지난조사 이후 경험을 질의한다.
- 직업생활을 위한 추가 비용의 경우, 1차조사에는 개인 지출과 외부 지원 여부를 구분하지 않고 응답하였으나, 2, 3차조사는 개인 지출과 외부 지원(국가, 민간 등) 비용으로 구분하여 응답하였다. 이후 4차조사에서는 문항 응답률 제고 및 응답 부담 감소를 위해 개인, 국가 단위로 구분하여 받았던 지출비용을 단일화 및 범주화하였다.
- 장애로 인해 근무 시 필요한 편의시설의 경우, 질문 문항을 명확하게 변경하였고 보기 항목 또한 법령에 따라 용어를 재정비하였다.



〈표 3-10〉 설문 항목 구성 - D. 취업을 위한 노력

영역	세부 영역 및 항목	8차	7차	6차	5차	4차	3차	2차	1차
취업 및 취업유지 관련 활동	취업 및 취업유지 관련 활동 경험여부 및 성과 정도	○	○	○	○	○	○	○	○
	공공 채용정보 인터넷 사이트 이용경험	○	○	○	○	○	○	○	○
	공공 취업알선기관 이용경험 여부 및 내용	○	○	○	○	○	○	○	○
	직업교육훈련 이용경험 여부 및 내용	○	○	○	○	○	○	○	○
취업활동 관련 지원사항	향후 취업/창업을 위해 필요한 사항(필요정도/시급정도)	○	○	○	○	○	○	○	○
	현 일자리에서 장애로 인해 근무 시 필요한 지원사항 <sup>1)</sup> (필요정도/지원·활용여부/도움정도)	○	○	○	○	○	○	○	○
직업생활을 위한 추가 비용	장애로 인한 추가 개인지출 유무 및 금액 <sup>2)</sup>	○	○	○	○	○	○	○	○
	장애로 인한 추가 외부 비용지원 유무 및 금액 <sup>3)</sup>						○	○	○
	추가 비용의 부담정도	○	○	○	○	○	○	○	○
	직업생활 유지를 위해 지원받기 희망하는 항목	×	×	×	×	×	○	×	×

- 1) 도움정도는 2차조사에 신규 추가됨
- 2) 장애로 인한 추가 개인지출 유무 및 금액, 장애로 인한 추가 외부 비용지원 유무 및 금액 문항은 4차조사에 단일화 및 범주화 된 문항으로 재구성함

### E. 직업적 능력

- 취업관련 업무능력 및 업무경력, 자격증, 취업관련 사회적지지 정도 등 응답자의 직업적 능력 전반을 확인한다.
- 자격증 문항의 경우 1차조사시에는 조사 당시 기준 보유한 자격증 전체에 대해 응답하였으며, 2차조사 이후부터는 지난조사 이후 취득한 자격증 정보만을 응답한다.
- 특히 4차조사에서는 「한국표준교육분류」 제정, 시행에 따른 분류기준을 적용하였다.

〈표 3-11〉 설문 항목 구성 - E. 직업적 능력

영역	세부 영역 및 항목	8차	7차	6차	5차	4차	3차	2차	1차
업무능력	취업관련 업무능력	○	○	○	○	○	○	○	○
	주관적 근로능력 평가	○	○	○	○	○	○	○	○
업무경력	일자리에서 일을 한 기간(누적)	○	○	○	○	○	○	○	○
	가장 오래한 직무를 수행한 기간	○	○	○	○	○	○	○	○
자격증	자격증 취득 현황(종류, 분야, 직업활동 도움 정도)	○	○	○	○	○	○	○	○
학력사항	(고등학교) 학교 유형	○	○	○	○	○	○	○	○
	(대학교) 전공 및 대학 소재지, 장애학생 특별전형 입학여부	○	○	○	○	○	○	○	○
사회적 관계	취업에 도움을 줄 수 있는 지인의 수	○	○	○	○	○	○	○	○
	주변 사람과 관계를 유지하는 정도	○	○	○	○	○	○	○	○

## F. 취업관련 태도 및 환경

- 취업 및 스스로에 대한 생각과 가족의 취업지지 정도, 장애인 고용정책에 대한 인식 등 취업 관련 환경적 요인을 전반적으로 파악하였다.
- 3차조사 이후 '좋은 일자리에 대한 인식' 문항 및 '취업하지 않으려는 일자리' 문항에서 근로시간 준수 관련 항목이 추가되었다.
- 상기와 같이 선다형 문항에서 이전조사 대비 항목이 변경(추가 또는 삭제)된 경우, 데이터상의 코드와 설문지의 항목 번호가 일치하지 않을 수 있다. 각 문항의 조사차수별 코드는 별첨된 코드북에 기재되어 있으며, 분석 시 유의해야 한다.
- 8차조사에서는 한국장애인고용공단에서 제공하는 고용서비스 인지도를 조사하는 문항을 추가하였다.

〈표 3-12〉 설문 항목 구성 - F. 취업관련 태도 및 환경

영역	세부 영역 및 항목	8차	7차	6차	5차	4차	3차	2차	1차
취업에 대한 생각	자기효능감	○	○	○	○	○	○	○	○
	좋은 일자리에 대한 인식 <sup>1)</sup>	○	○	○	○	○	○	○	○
	취업하지 않으려는 일자리 <sup>2)3)</sup>	○	○	○	○	○	○	○	○
가족	가족의 취업지지 정도	○	○	○	○	○	○	○	○
	부모의 생존여부, 동거여부, 최종학력	×	×	×	×	×	×	×	○
	만 14세 당시 부모님 취업여부 및 가구의 경제적 상황	×	×	×	×	×	×	×	○
스스로에 대한 인식	자아존중감	○	○	○	○	○	○	○	○
	장애수용정도	○	○	○	○	○	○	○	○
장애인 고용정책	장애인 의무고용제도 인지여부 및 도움 정도	○	○	○	○	○	○	○	○
	장애인 서비스에 대한 신뢰 정도 <sup>4)</sup>	○	○	○	○	○	○	○	○
	장애인 고용 서비스에 대한 신뢰 정도	○	○	○	○	○	○	○	×
	장애인 고용증진 주체에 대한 인식	○	○	○	○	○	○	○	○
최저임금 수준에 대한 생각	최저임금 수준에 대한 생각	×	×	×	○	○	×	×	×
공단 지원제도	알고 있는 장애인 고용 지원제도	○	×	×	×	×	×	×	×

- 1) 좋은 일자리에 대한 인식 세부 항목 중 '8. 출퇴근 시간이 준수되는 일자리'는 3차조사 신규 추가됨
- 2) 취업하지 않으려는 일자리 세부 항목 중 '6. 나는 초과근무(업무로 인한 잦은 야근, 회식 등)가 많은 곳에는 취업하지 않겠다'는 3차조사 신규 추가됨
- 3) 취업하지 않으려는 일자리 세부 항목 중 '7. 장애인에 대해 차별하는 일자리에 취업하지 않겠다.'는 2차조사 신규 추가됨
- 4) 장애인 대상 서비스에 대한 신뢰 정도는 2차조사에서 문항이 분야별로 세분화되어 일부 수정됨
- 5) 이전조사 대비 세부항목이 추가된 문항은 설문지와 데이터의 코드가 상이하므로 분석 시 유의해야 함

### G. 일상생활과 삶의 질

- 건강상태 및 생활습관, 일상생활 시 도움 필요여부 및 도움제공자, 일상 및 여가활동, 차별경험 등 일상생활 관련 문항으로 구성되었다.
- 8차조사에서는 여가 시간 외에도 가사 및 돌봄 노동에 대한 현황을 파악하기 위하여 하루 평균 노동 시간을 질문하는 문항이 추가되었다.

〈표 3-13〉 설문 항목 구성 - G. 일상생활과 삶의 질

영역	세부 영역 및 항목	8차	7차	6차	5차	4차	3차	2차	1차
건강상태	진반적 건강상태	○	○	○	○	○	○	○	○
	만성질환 여부	○	○	○	○	○	○	○	○
	만성질환 치료활동 여부	○	○	○	○	○	○	○	×
생활습관	운동, 수면, 식사습관	○	○	○	○	○	○	○	○
	흡연 및 음주 여부	○	○	○	○	○	○	○	○
스트레스 경험	일상생활 중 스트레스, 우울경험 여부	○	○	○	○	○	○	○	○
	주관적 행복감 정도	○	○	○	○	○	○	○	○
일상생활의 도움	일상생활 중 타인의 도움 필요 여부	○	○	○	○	○	○	○	○
	도움제공자 유무 및 주요 도움제공자	○	○	○	○	○	○	○	○
일상, 여가 및 사회참여	평소 주로 하는 활동	○	○	○	○	○	○	○	○
	하루 평균 여가, 노동 시간(평일/주말)	○	○	○	○	○	○	○	○
	배우자의 분담정도 만족도	○	×	×	×	×	×	×	×
	여가활동 및 사회참여 활동 경험여부 <sup>1)</sup>	○	○	○	○	○	○	○	○
	외출 시 이동수단 및 이동시 어려움	○	○	○	○	○	○	○	○
차별경험 및 만족도	일상에서의 차별경험	○	○	○	○	○	○	○	○
	일상생활 만족도, 주관적 사회경제적 지위	○	○	○	○	○	○	○	○

1) 사회참여활동 경험여부 중 투표 참여의 경우 1차조사는 제20대 국회의원 선거(2016년 4월 13일), 2차조사는 제19대 대통령 선거(2017년 5월 9일) 참여 여부를 의미함. 3~4차조사에는 선거 참여 여부를 질문하지 않았고 5차조사에서는 제21대 국회의원 선거(2020년 4월 15일) 참여 여부를 의미함. 6차조사에는 선거 참여 여부를 질문하지 않았고 7차조사에서는 제20대 대통령 선거(2022년 3월 9일) 참여 여부를 의미함.

## H. 가구 일반사항

- 가구 일반사항 파트는 가구원 현황, 소득 및 지출 현황, 자산 및 부채, 주거환경, 노후대비 등 전반적인 가구 현황을 파악하는 문항으로 구성되었다.
- 가구원 현황의 경우 지난조사 당시 응답한 가구원의 지속여부 및 정보 수정 여부를 파악하고, 지난 조사 이후 신규 가구원(출생, 전입 등) 현황을 구분하여 응답받아 조사시점 기준 가구원 현황을 응답받았다.

〈표 3-14〉 설문 항목 구성 - H. 가구 일반사항

영역	세부 영역 및 항목	8차	7차	6차	5차	4차	3차	2차	1차
가구원 정보	가구원 현황	○	○	○	○	○	○	○	○
	패널의 가구주 여부	○	○	○	○	○	○	○	○
가구소득 및 지출	가구 근로소득 현황(패널/가구주/가구 전체)	○	○	○	○	○	○	○	○
	가구 근로 외 소득 현황(가구 전체)	○	○	○	○	○	○	○	○
	가구 월평균 생활비	○	○	○	○	○	○	○	○
	경제생활 관련 어려움	○	○	○	○	○	○	○	○
자산 및 부채	자산 및 부채금액	○	○	○	○	○	○	○	○
	자동차 보유 여부 및 대수	○	○	○	○	○	○	○	×
국민기초생활보장 수급현황	국민기초생활보장 수급가구 여부 및 수급시기	○	○	○	○	○	○	○	○
	수급중인 급여 내용	○	○	○	○	○	○	○	○
주거환경	지난조사 이후 이사 여부 및 시기	○	○	○	○	○	○	○	×
	현 거주지에 살게 된 동기	○	○	○	○	○	○	○	×
	주택 종류 및 전용면적, 입주형태	○	○	○	○	○	○	○	○
	주택 가격(시가, 보증금, 월세)	○	○	○	○	○	○	○	○
고령자 <sup>1)</sup>	60세 이후 노후생활 관련 염려사항	○	○	○	○	○	○	○	○
	노후를 위한 경제적 준비 여부, 준비내용, 준비 정도	○	○	○	○	○	○	○	○

1) 고령자 문항의 경우 조사시점 기준 만 40세 이상인 패널에게만 응답받음

## JOB. 직업력

- 2차조사 이후부터는 응답자가 경험한 다수의 일자리에 대해 직업력 조사를 실시하였다. Ca 혹은 Cb파트에서 응답한 주업을 제외하고 경험한 일자리가 있는 경우, 해당 일자리 내용을 직업력 조사표에서 응답하였다. 일자리별로 지난조사 당시 응답여부, 종사상 지위, 현재 지속 여부에 따라 직업력 조사표 유형이 결정된다.
- 직업력은 개인단위가 아닌 일자리 단위로 응답하는 설문으로, 개인의 일자리 경력에 따라 복수의 직업력 문항 응답이 가능하다. 경험한 일자리의 개수만큼 직업력 문항을 응답하도록 구성하였으며, 직업력 조사표 유형구분 및 유형별 세부 문항 내용은 다음과 같다.

〈표 3-15〉 직업력 조사표 유형 구분

기존 일자리 여부	종사상 지위	현재 지속 여부	질문셋 유형
기존 일자리	임금근로자	지속	유형1
		그만됨	유형2
	비임금근로자	지속	유형3
		그만됨	유형4
신규 일자리	임금근로자	지속	유형5
		그만됨	유형6
	비임금근로자	지속	유형7
		그만됨	유형8

\* 2차조사에 한하여 지난조사에서 부업으로 조사된 일자리는 새로운 일자리(유형5~유형8)로 간주하여 조사하였음

〈표 3-16〉 설문 항목 구성 - JOB. 직업력

질문셋 유형	세부 문항 내용
유형1	일자리 기본 정보, 일자리 특성, 근로조건 및 실태, 근로환경
유형2	일자리 기본 정보, 일자리 특성, 근로조건 및 실태, 근로환경, 퇴직현황
유형3	사업체 기본 정보, 근로조건 및 실태
유형4	사업체 기본 정보, 근로조건 및 실태, 퇴직현황
유형5	취업 정보, 일자리 기본 정보, 일자리 특성, 근로조건 및 실태, 근로환경
유형6	취업 정보, 일자리 기본 정보, 일자리 특성, 근로조건 및 실태, 근로환경, 퇴직현황
유형7	취(창)업 정보, 사업체 기본 정보, 근로조건 및 실태
유형8	취(창)업 정보, 사업체 기본 정보, 근로조건 및 실태, 퇴직현황

## I. 코로나19 지속

- 코로나19 지속 파트는 코로나19 전후 변화사항, 코로나19 관련 인식 및 변화를 파악하는 문항으로 구성되었다.
- 코로나19 상황 속에서 코로나19의 영향력과 이와 관련한 전사회적 대응 현안을 파악하기 위한 파트로, 한시적으로 운영 예정이다.

〈표 3-17〉 설문 항목 구성 - I. 코로나19 지속

영역	세부 영역 및 항목	8차	7차	6차	5차	4차	3차	2차	1차
코로나19 예방 접종 및 확진	패널 본인의 코로나19 예방(백신) 접종 횟수	○	○	×	×	×	×	×	×
	패널 본인의 코로나19 확진 여부 및 시기	○	○	×	×	×	×	×	×
	확진 이후 건강 회복 정도	○	○	×	×	×	×	×	×
코로나19 경험	패널 본인 및 가구원 코로나19 검사 여부 및 결과	×	×	○	×	×	×	×	×
	패널 본인 및 가구원 자가격리 여부	×	×	○	×	×	×	×	×
코로나19 전후 변화사항 및 지원 필요 사항	코로나19 전후 일상생활 변화	○	○	○	×	×	×	×	×
	코로나19 전후 일상생활 어려움	×	×	○	×	×	×	×	×
	코로나19 관련 지원 필요 사항	○	○	×	×	×	×	×	×
코로나19 인식	코로나19 감염 위험도	○	○	○	×	×	×	×	×
	코로나19 관련 장애인 개인 인식 동의도	×	×	○	×	×	×	×	×
	코로나19 전후 삶의 질 변화	○	○	○	×	×	×	×	×
코로나19 백신	코로나19 백신 접종 여부 및 접종 의사	×	×	○	×	×	×	×	



## 제4장 데이터 및 변수 가이드

### 1. 데이터 구조

장애인고용패널조사는 데이터의 수록내용과 활용 목적에 따라 메인 데이터, 통합 데이터, 직업력 데이터로 구분하여 제공된다. 각 데이터의 특성은 다음과 같다.

#### 가. 메인 데이터(MAIN DATA)

- 메인 데이터는 직업력을 제외한 해당년도 조사 데이터 전체가 수록되어 있다. 모든 조사년도마다 만들어지므로, 조사 차수별로 각각의 메인 데이터가 존재한다.
- 일자리 정보는 현재 주업으로 영위하고 있는 1개 일자리에 대한 정보만 포함한다. 주업 외의 일자리가 있는 경우 해당 일자리 정보는 직업력 데이터에 수록되어 있다.
- 데이터의 변수 구성은 조사 기본 정보와 조사 데이터로 구분할 수 있다. 기본 정보 데이터는 패널 ID, 조사차수별 응답여부, 인구통계학적 변수(성별, 생년월일, 거주지), 가중치, 경제활동상태 가공변수 등 데이터 분석 시 활용할 수 있는 구성변수가 포함되어 있으며, 조사 데이터 앞에 배치되어 있다. 해당 변수는 통합 데이터와 직업력 데이터에도 마찬가지로 포함되어 있다.
- 조사 데이터는 조사표의 파트 순서와 동일하게 구성되어 있으며, 데이터 분석의 편의를 제고하기 위해 조사 데이터를 기반으로 산출한 가공변수가 포함되어 있다.

[그림 4-1] 메인 데이터 구성



## 나. 통합 데이터(WIDE TYPE/LONG TYPE DATA)

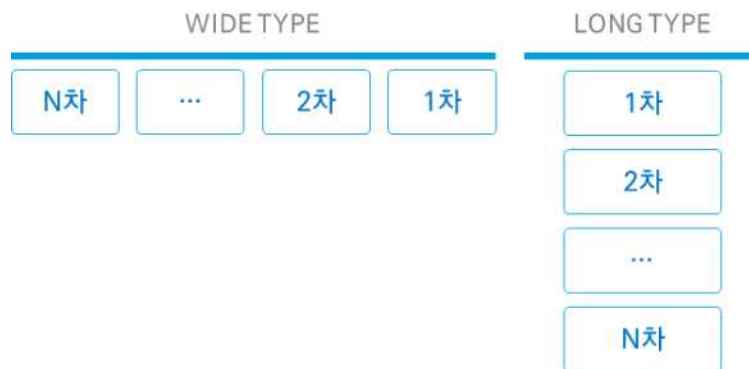
- 통합 데이터는 각 조사년도별 자료를 하나의 데이터로 통합한 자료로, 통합 방식에 따라 WIDE TYPE과 LONG TYPE 두 가지 방식으로 구분하여 제공된다.

WIDE TYPE: 각 차수별 조사 데이터를 횡(MERGE)으로 통합한 자료로, 모든 차수의 데이터를 수록하고 있다. 변수명은 메인 데이터와 동일하게 조사차수가 포함된 변수명을 적용하였다(변수 가이드 참조). 또한 7차조사부터 직업력과의 데이터 활용성 강화를 위해 WIDE TYPE 데이터에 패널별로 일자리 개수만큼의 행 변수를 추가하였다.

LONG TYPE: 각 차수별 조사 데이터를 종(APPEND)으로 누적한 자료로, WIDE TYPE과 마찬가지로 모든 차수의 데이터를 수록하고 있다. 개별 패널은 응답한 조사 차수에 따라 조사차수 변수(wave)로 구분되는 데이터 행(row)을 갖는다. 변수명 구조는 조사차수가 포함되지 않은 통합 변수명을 적용하였다.

- 통합 데이터 역시 데이터 분석의 편의성 확보를 위한 가공변수가 포함되었다. 단, 2차조사 이후 일부 문항은 다수의 가공변수가 산출되어, 지나치게 많은 변수가 오히려 데이터 이용 시 어려움을 야기할 우려가 있다. 이에 메인 데이터에는 실제 응답 변수와 가공변수를 예외 없이 모두 수록하되, 통합 데이터에는 일부 가공변수가 많은 문항에 한하여 응답 변수를 제외하고 분석용 가공변수만을 수록하였다. E파트의 자격증 보유 현황, H파트의 가구원 현황이 이에 해당한다(관련 상세 설명은 5장 주제별 가이드를 참조할 것).

[그림 4-2] 통합 데이터 구성





### 다. 직업력 데이터(WORK HISTORY)

- 직업력(work history) 데이터는 1차조사 이후 조사된 패널의 모든 일자리 정보가 수록되어 있다. 매년 조사를 거듭하며 패널의 모든 직업력이 하나의 데이터에 누적되는 형태의 자료이므로, 연도별 자료가 별도로 제공되지 않는다.
- 데이터 이용자의 분석 편의성 제고를 위해 분석변수(성별, 연령, 학력, 장애구분, 경활상태 등), 일자리 고유번호(Jobseq변수 : 패널별 동일 일자리를 확인할 수 있는 고유번호), 누적 일자리 개수(jobtotal) 등이 있고, 가공방식에 따라 WORK1, WORK2 두 가지 형태로 구분하여 제공된다.
 

**WORK1:** 직업력 데이터를 중(APPEND)으로 누적한 자료로, 패널의 일자리가 여러 개인 경우 현재 지속하고 있는 일자리, 그만둔 일자리 등 모든 주업, 부업의 일자리 정보를 직업력 데이터에서 확인이 가능하다.

**WORK2:** 패널별 직업력을 1개의 행(Raw)에 모두 포함할 수 있도록 구성하였으며, 첫 번째 일자리로부터 마지막 일자리(8차년도의 경우 7번째 일자리 정보까지 제공하고 있음) 정보를 확인할 수 있다.

[그림 4-1] 직업력 데이터 구성(WORK1)



[그림 4-2] 직업력 데이터 구성(WORK2)



〈표 4-1〉 장애인고용패널조사 데이터 제공 유형별 특징

구분	특징
메인 데이터	<ul style="list-style-type: none"> <li>· 조사차수별로 각각 데이터가 존재함</li> <li>· 직업력을 제외한 모든 조사표 정보 수록</li> <li>· 조사표의 파트 구조와 동일하게 구성</li> <li>· 일자리 정보는 조사시점 당시 주업에 대한 1개 일자리 정보만 수록</li> <li>· 데이터 분석 시 활용을 위한 가공변수 제공</li> </ul>
통합데이터 WIDE TYPE	<ul style="list-style-type: none"> <li>· 각 차수별 조사 데이터를 횡(MERGE)으로 통합, 매 조사년도별로 누적하여 모든 차수의 데이터를 수록</li> <li>· 메인 데이터와 동일한 변수명 적용(조사차수 포함)</li> <li>· 일자리 정보는 각 차수별로 당시 주업에 대한 1개 일자리 정보만 수록</li> <li>· 데이터 분석 시 활용을 위한 가공변수 제공</li> <li>· 8차조사 자료에는 직업력과 데이터 활용성 강화를 위해 WIDE 통합 데이터에 패널별로 일자리 개수만큼의 행 변수를 추가함(WORK2)</li> </ul>
통합데이터 LONG TYPE	<ul style="list-style-type: none"> <li>· 각 차수별 조사 데이터를 종(APPEND)으로 통합, 매 조사년도별로 누적하여 모든 차수의 데이터를 수록</li> <li>· 조사차수가 포함되지 않은 통합 변수명 적용</li> <li>· 일자리 정보는 각 차수별로 당시 주업에 대한 1개 일자리 정보만 수록</li> <li>· 데이터 분석 시 활용을 위한 가공변수 제공</li> </ul>
직업력 데이터	<ul style="list-style-type: none"> <li>· 1차조사 이후 조사된 패널의 모든 일자리 정보 수록하여 패널의 일자리 역사를 알 수 있는 자료임</li> <li>· 메인 데이터에 포함되지 않은 부업 및 그만둔 일자리에 대한 이력은 직업력 데이터에서 확인할 수 있음</li> <li>· 가공방식에 따라 WORK1, WORK2 두 가지 형태로 구분되어 제공</li> </ul> <p>(WORK1)</p> <ul style="list-style-type: none"> <li>- 직업력 조사표와 동일한 형태로 구성하여 8개 직업력 조사표 유형 순서대로 각 유형에 해당하는 일자리를 기재함</li> <li>- 패널 당 각 일자리를 종(APPEND)으로 누적인 자료</li> <li>- 패널별로 일자리 개수만큼의 행이 존재함</li> </ul> <p>(WORK2)</p> <ul style="list-style-type: none"> <li>- 패널별로 일자리 개수만큼의 행이 존재함</li> <li>- 패널당 1개 행에 모든 일자리 정보 포함</li> </ul>

## 2. 변수 가이드

### 가. 변수 개요

- 장애인고용패널조사의 변수는 키(Key)변수와 주제별 변수로 구분된다. 키 변수는 패널 고유 번호와 데이터 분석 시 빈번하게 사용되는 인구통계 정보 등이 포함되어 있다. 주제별 변수는 조사표의 내용을 반영한 실제 응답 변수와, 응답내용을 분석의 편의를 위해 가공한 가공변수로 구성되어 있다.

〈표 4-2〉 주요 변수 구성

구분	주요 내용
Key 변수	<ul style="list-style-type: none"> <li>· 패널 일련번호</li> <li>· 차수별 응답여부 및 조사실패 사유(미응답시)</li> <li>· 성별, 생년월일, 거주지역, 조사일시</li> <li>· 횡단면/종단면 가중치</li> <li>· 경제활동상태 및 종사상 지위 구분</li> <li>· 일자리 개수 현황</li> </ul>
주제별 변수	<ul style="list-style-type: none"> <li>· 조사표 응답 내용 및 관련 가공변수</li> </ul>

### 나. Key 변수의 구성

#### 1) 패널 ID(pid)

- 장애인고용패널조사는 매년 동일 개인을 반복 추적조사하기 때문에 패널 구성원을 구분하기 위한 고유번호인 패널 ID(pid)가 key변수로 부여된다. 패널 ID는 조사차수와 무관하게 불변하는 고유 번호이므로, 개별 패널을 구분하거나 자료 병합 시 정렬기준으로 활용할 수 있다.

#### 2) 조사 참여 여부(p\_\_ / p01\_\_) 및 미응답 사유(npreason\_\_)

- 'p\_\_' 변수는 해당 조사차수에 응답하였는지 구분하는 변수로써 '응답'(1), '미응답'(2)으로 구분된다. 조사에 응답하지 않은 경우, 응답하지 않은 사유는 'npreason\_\_' 변수에서 확인할 수 있다.
- 'p01\_\_' 변수는 연속응답여부를 확인하는 가공변수로 모든 조사차수에 응답한 패널을 구분하기 위해 사용된다. 예를 들어, 1~8차조사에 모두 응답한 패널은 '연속응답'(1), 나머지 패널은 '불연속응답'(2)으로 구분된다.

**3) 장애등록 취소시기 및 사유(y\_ce\_\_ / m\_ce\_\_ / r\_ce\_\_)**

- 장애등록이 취소된 경우, 실제 취소시기와 사유를 확인하여 패널에서 제외된다. y\_ce\_\_, m\_ce\_\_, r\_ce\_\_ 변수에는 장애등록 취소시기와 사유가 기록되어 있다.

**4) 성별 및 생년월일(gender / birth(y, m, d))**

- 성별, 생년월일 등 패널 기본정보는 매년 조사 항목에 포함되어 있으며, 패널 기본정보의 수정 혹은 변화 여부를 매년 재확인한다. 데이터에는 해당 조사년도에 수정된 최종 정보가 들어가 있다.

**5) 거주지(area\_\_ / aarea\_\_)**

- 조사년도별 패널의 거주지는 광역시/도 단위로 분류하여 'area\_\_' 변수에 기재되어 있다. 'aarea\_\_' 변수에는 아래 기준에 따라 수도권(1), 광역시권(2), 기타 시도(3)로 재분류하였다.  
 수도권: 서울, 인천, 경기  
 광역시권: 부산, 대구, 광주, 대전, 울산, 세종  
 기타 시도: 강원, 충북, 충남, 전북, 전남, 경북, 경남, 제주

**6) 조사일시(year\_\_ / month\_\_ / day\_\_)**

- 조사에 참여한 날짜가 기록된 변수이다. 차수별 조사참여 여부에 따라 조사 참여시기는 결측치로 나타날 수 있다.

**7) 가중치(wt\_01 / wt\_02)**

- 'wt\_01'은 횡단면 가중치, 'wt\_02'는 종단면 가중치 변수이다. 종·횡단면 가중치는 분석의 목적과 방법에 따라 타당하게 사용되어야 한다. 가중치에 대한 상세한 설명은 2장 '표본설계 및 가중치'에 기술되어 있다.

**8) 경제활동상태 및 종사상 지위(emp\_\_ / emp\_\_type)**

- 경제활동상태와 종사상 지위는 고용 관련 주제로 데이터 분석 시 빈번하게 사용되므로, 세부 기준에 따라 다양하게 분류한 변수를 제공한다.
- 'emp\_\_02~emp\_\_07' 변수는 패널의 차수별 주업 기준 경제활동상태를 분류한 변수이며, 'emp\_\_type01~emp\_\_type03' 변수는 취업자로 분류되는 패널의 종사상 지위를 구분한 변수이다. 변수별 항목 현황 및 구분 기준은 아래와 같다.

〈표 4-3〉 경제활동상태 및 종사상 지위 변수값 구성

변수명	emp_02	emp_03	emp_04	emp_05
항목	1. 취업자	1. 취업자	1. 취업자	1. 임금근로자
				2. 자영업자
				3. 무급가족종사자
	2. 미취업자	2. 실업자	2. 실업자	4. 실업자
		3. 비경제활동인구	3. 취업 희망자	5. 비경제활동인구
	4. 취업 비희망자			

〈표 4-4〉 경제활동상태 및 종사상 지위 변수값 구성(계속)

변수명	emp_06	emp_type01	emp_type02	emp_07	emp_type03
항목	1. 임금근로 기존 일자리	1. 임금근로자	1. 상용근로자	1. 기간의 정함이 없는 근로자	1. 기간의 정함이 없는 근로자
	2. 임금근로 신규 일자리		2. 임시근로자	2. 장기 고정 기간 근로자	2. 장기 고정 기간 근로자
	3. 비임금근로 기존 일자리		3. 일용근로자	3. 일용근로자	3. 단기 고정 기간 근로자
		4. 단기·임시 근로자			4. 단기·임시 근로자
				5. 유급 견습, 훈련생 및 인턴	5. 유급 견습, 훈련생 및 인턴
	4. 비임금근로 신규 일자리	2. 비임금근로자	4. 자영업자	6. 자영업자	
5. 실업자	5. 무급가족종사자		7. 무급가족종사자		
6. 비경제활동인구					

〈표 4-5〉 경제활동상태 및 종사상 지위 분류기준

구분	내용
경제활동 상태 분류기준	<p>① 취업자 가. 조사대상기간에 소득, 이익, 임금 등의 본인 수입을 목적으로 1시간 이상 일한 자 나. 동일가구 내 가구원이 운영하는 사업체나 농장의 수입을 위하여 주당 18시간 이상 일한 무급 가족종사자(18시간 미만 일한 무급가족종사자는 실업자 또는 비경제활동인구) 다. 직업 또는 사업체를 가지고 있으나 조사대상기간에 일시적인 병 또는 사고, 휴가·연가, 교육·훈련, 노사분규 등의 사유로 일하지 못한 경우 중 일시휴직 사유가 해소되면 즉시 복귀 가능한 자</p> <p>② 실업자: 조사대상기간에 수입이 있는 일을 하지 않았지만, 지난 4주 내 일자리를 찾아 적극적으로 구직활동을 하였던 사람으로서, 일자리가 주어지면 즉시 취업이 가능한 자</p> <p>③ 비경제활동인구: 만 15세 이상 인구 중 조사대상기간에 취업도 실업도 아닌 상태에 있는 자 - 취업 희망자: 현재 비경제활동인구이면서 향후 1년 이내 혹은 그 이후라도 취업할 의사가 있는 자 - 취업 비희망자: 현재 비경제활동인구이면서 향후 취업할 의사가 없는 자</p>
종사상 지위 분류기준	<p>① 상용근로자: 임금근로자로서 다음에 해당되는 사람 가. 고용계약 설정자는 고용계약기간이 1년 이상인 경우 나. 고용계약 미설정자는 소정의 채용절차에 의해 입사하여 인사관리 규정을 적용받거나 상여금 및 퇴직금 등 각종 수혜를 받는 사람</p> <p>② 임시근로자: 임금근로자로서 다음에 해당되는 사람 가. 고용계약 설정자는 고용계약기간이 1개월 이상 1년 미만인 경우 나. 고용계약 미설정자는 일정한 사업(완료 1년 미만)의 필요에 의해 고용된 경우</p> <p>③ 일용근로자: 임금근로자로서 다음에 해당되는 사람 가. 고용계약기간이 1개월 미만인 경우 나. 매일매일 고용되어 근로의 대가로 일급 또는 일당제 급여를 받고 일하는 자</p> <p>④ 자영업자: 고용원이 있는 자영업자와 고용원이 없는 자영업자를 합친 개념 가. 고용원이 있는 자영업자: 한 사람 이상의 유급 고용원을 두고 사업을 경영하는 사람 나. 고용원이 없는 자영업자: 자기 혼자 또는 무급가족종사자와 함께 자기 책임 하에 독립적인 형태로 전문적인 일을 수행하거나 사업체를 운영하는 사람 ☞ '사장'이라 하더라도 법인기업이면 '임금근로자'로 분류</p> <p>⑤ 무급가족종사자: 동일가구 내 가족이 경영하는 사업체, 농장에서 무보수로 18시간 이상 일하는 자 ☞ 동일가구 내 가족이라도 일정한 봉급을 받는 경우는 '임금근로자'로 분류 ☞ 가구를 달리하는 가족 또는 친척의 일을 돕는 것은 무급가족종사자에 해당되지 않음</p>
종사상 지위 (임금 근로자) 신분류	<p>① 기간의 정함이 없는 근로자 가. 정년이 있거나, 고용이 종료되는 날짜 또는 사건(프로젝트 및 휴직 대체 근무 종료 등)이 정해져 있지 않은 근로자로서 지속적 또는 무기한으로 고용된 근로자 나. 근로 계약이 무기한 계속될 것으로 예상되는 일자리에 최근 발령받은 근로자 또는 수습 신분의 근로자와 무기 계약 근로자</p> <p>② 장기 고정 기간 근로자(1년 이상) 가. 정년 이외 고용이 종료되는 날짜 또는 사건이 정해져 있는 자로서 고용(계약)이 시작된 날부터 종료되는 날짜 또는 사건이 발생할 것으로 예상되는 시기가 1년 이상인 근로자</p> <p>③ 단기 고정 기간 근로자(3개월~1년 미만) 가. 고용이 종료되는 날짜 또는 사건이 정해져 있는 자로서 고용(계약) 기간이 3개월 이상 1년 미만인 근로자</p> <p>④ 단기·임시 근로자(3개월 미만) 가. 고용(계약)이 시작된 날로부터 종료되는 날짜 또는 사건이 발생할 것으로 예상되는 시기가 3개월 미만인 근로자</p> <p>⑤ 유급 견습, 훈련생 및 인턴 가. 전문 기능이나 직무를 습득하기 위해 현장 교육의 일환으로 타인을 위해 재화를 생산하거나 서비스를 제공하는 활동을 수행하고 그 대가로 보수를 받는 근로자 ☞ 단, 새로운 일자리 시작 단계의 수습 기간, 직장 내 직무교육연수(OJT) 중인 경우, 평생교육이나 직무훈련을 받는 근로자, 무급으로 일하는 경우, 가족 또는 가구 구성원이 운영하는 사업체에서 일하며 훈련받는 무급가족종사자는 제외</p>

### 9). 일자리 ID

- 조사를 통해 파악된 패널의 일자리 현황은 조사시기와 해당 조사시기에 파악된 일자리 개수, 일자리를 구분할 수 있는 변수로 일정한 패턴을 가지고 구성되며, 일자리 ID 구성은 <표 4-6>와 같다.

**조사차수(AA):** 조사 차수는 p\_year과 동일한 정보이며 두 자리로 표시한다.

**패널id(BBBBB):** 패널 id는 pid와 동일한 정보이며 최소 5자리에서 최대 7자리까지로 구성되어 있다.

**일자리 종류(CC):** 일자리 종류는 일자리의 종사상 지위 및 지속여부를 1~8까지의 숫자로 나타낸 정보로, 상세한 설명은 <표 5-31>에서 다루기로 한다.

**일자리 개수(DD):** 일자리 개수는 해당 차수에 종사한 일자리 개수로 개수 시작은 주업부터 표기하며, 주업(01), 부업(02, 03 ...)으로 작성한다. 만약 모두 주업이라면 그만두는 경우를 01, 새로 시작하는 경우를 02로 한다.

<표 4-6> 일자리 ID 부여 규칙

jobid	AA	BBBBB	CC	DD
-	조사 차수	패널id	일자리 종류	일자리 개수

<표 4-7> 일자리 ID 현황

변수명	내용
emp_jobid	(메인/통합) 차수별 주업 일자리 id
jobid_0101~jobid_0103	(메인/통합) 차수별 기존 일자리 id
jobid_0201~jobid_0203	(메인/통합) 차수별 신규 일자리 id
cajobid_	(메인/통합) 차수별 임금근로 주업 일자리 id
cbjobid_	(메인/통합) 차수별 비임금근로 주업 일자리 id
casjobid_01, casjobid_02	(메인/통합) 1차조사 임금근로 부업 일자리 id
cbsjobid_01, cbsjobid_02	(메인/통합) 1차조사 비임금근로 부업 일자리 id
jobid	(직업력 work1) 직업력 유형별(1~8) 일자리 id
jobpreid	(직업력 work1) 직전 차수 조사 일자리 id
job00_id	(직업력 work2) 패널이 갖고 있는 일자리 개수별 id
job00_preid	(직업력 work2) 패널이 갖고 있는 일자리 직전 차수 id

## 10) 일자리 개수 (jobold\_\_ / jobtotal\_\_ / jobctotal\_\_ / joballtotal\_\_ / jobatotal\_\_)

- 패널이 경험한 일자리 개수를 조사시점에 따른 누적 일자리 개수와 현재 기준 지속하고 있는 일자리 개수, 현재까지 경험한 전체 일자리 개수로 구분하여 제공한다.
- 주요 key 변수 현황을 요약하면 다음과 같다.

〈표 4-8〉 주요 key 변수 현황

변수명	내용
pid	패널 id
p__	조사차수별 응답여부
p01__	연속응답여부
npreason__	조사차수별 미응답 사유
y_ce__, m_ce__, r_ce__	장애등록 취소시기 및 사유
gender	성별
birthy, birthm, birthd	생년월일
wave__	조사 차수
year__, month__, day__	차수별 조사일시
wave_01	차수별 조사년도
wt_01	차수별 횡단면 가중치
wt_02	차수별 종단면 가중치
area__, aarea__	거주지역
emp_02~emp_06	경제활동상태
emp_type01, emp_type02	종사상 지위
emp_jobid	주업 일자리 id
jobold__	지난조사 시 일자리 개수(주업 및 부업)
jobtotal__	지난조사 이후 총 일자리 개수
jobctotal__	현재 지속 일자리 개수
joballtotal__	현재까지 누적된 일자리 총 개수
jobatotal__	(직업력) 지난조사 시 누적 일자리 개수



## 다. 주제별 변수의 구성

### 1) 메인 및 통합 데이터 변수명 부여 규칙

- 조사표 내용을 반영한 주제별 변수는 모든 조사차수에 일관된 변수명 부여 규칙이 적용되었다. 변수명이 부여되는 규칙은 다음과 같다.

예시) a080310

파트명	조사차수	변수번호
a	08	0310

- **파트명**: 변수명의 파트명 구분은 조사표의 각 영역별 구분과 동일하며, 세부적인 내용은 다음과 같다.

〈표 4-9〉 조사표 파트 구분

파트명	내용
a	패널 기본 정보
b	경제활동상태 판별
ca	임금근로자
cb	비임금근로자
cc	미취업자
d	취업을 위한 노력과 지원
e	직업적 능력
f	취업 관련 태도 및 환경
g	일상생활과 삶의 질
h	가구 일반사항

- **조사차수**: 파트명 다음 두 자리 숫자는 조사 차수를 의미한다(01은 1차조사, 02는 2차조사, 03은 3차조사, 04는 4차조사, 05는 5차조사, 06 6차조사, 07 7차조사 08 8차조사). 단, LONG TYPE 통합 데이터의 경우 조사차수 구분 숫자가 포함되지 않은 통합 변수명을 사용한다. 예를 들어, 메인 데이터 및 WIDE TYPE 통합데이터에서 'ca012604'에 해당하는 LONG TYPE 데이터에서의 변수명은 "ca2604"이다.
- **변수번호**: 해당 변수의 일련번호를 의미한다.
- **개방형 응답**: 선다형 문항에서 '기타'로 응답한 후 구체적인 내용을 개방형으로 응답받은 경우, 해당 변수는 기존 변수명 뒤에 'f'를 추가하여 구분한다. 예를 들어, 비임금근로자의 근무 사업체 유형 문항(cb081201)에서 '기타'에 응답한 후 개방형으로 사업체 유형을 응답한 경우, 개방형으로 응답받은 구체적 내용은 'cb081201f'이다.

## 2) 직업력 데이터 변수명 부여 규칙

- 직업력 데이터는 매년 조사를 거듭하며 패널의 모든 직업력이 하나의 데이터에 누적되는 형태의 자료이므로, 연도별 자료가 별도로 제공되지 않아 조사차수에 따라 변수명이 부여되는 메인 및 통합 데이터 변수명 부여 규칙과는 차이가 있다.
- WOKR1의 변수명 부여 규칙은 다음과 같다.

예시) job05080101

파트명	일자리 종류	조사차수	변수번호
job	05	08	0101

- **파트명:** 변수명의 파트명 구분은 job으로 통일한다.
  - **일자리 종류:** 직업력 조사표 유형으로 총 8가지 유형이다. 예를 들어 05는 직업력 조사표 유형 중 5번째 유형으로 [임금근로 - 신규 - 지속] 즉, 해당 차수에서 신규로 지속하고 있는 임금근로 형태의 일자리를 의미한다.
  - **조사차수:** 조사 차수는 메인 및 통합 데이터에서 설명한 조사 차수 개념과 동일하다(01은 1차 조사, 02는 2차조사, 03는 3차조사, 04 4차조사, 05 5차조사, 06 6차조사, 07 7차조사, 08 8차조사)
  - **변수번호:** 해당 변수의 일련번호를 의미한다.
  - **개방형 응답:** 메인 및 통합 데이터 개념과 동일하다.
- WORK2는 패널별로 갖고 있는 일자리를 하나의 행(row)으로 가공한 데이터로 일자리별로 변수를 확인할 수 있도록 구성되어 다음으로 WOKR2의 변수명 부여 규칙은 다음과 같다.

예시) job002\_continue

파트명	일자리 개수	변수내용
job	002	_continue

- **파트명:** 변수명의 파트명 구분은 job으로 통일한다.
- **일자리 개수:** 현재까지 패널이 경험한 일자리 횟수에 따라 숫자가 부여된다. 첫 번째 일자리는 001, 두 번째 일자리는 002, 세 번째 일자리는 003으로 표기하였으며, 패널이 현재까지 경험한 누적 일자리가 총 7개로 확인되어 job007까지만 변수가 생성되어 있다.
- **변수 내용:** 해당 변수의 내용을 집약적으로 표현한 단어로 구성되어 있다.

〈표 4-10〉 WORK2 변수 내용

변수	내용
_id	일자리 ID
_type	일자리 종류
_firyear	일자리 최초 조사 차수
_seq	일자리 고유 번호
_preid	일자리 이전 차수 ID
_main	일자리 주업여부
_continue	일자리 지속여부
_industrial	일자리 사업내용
_industrial_code1	(가공) 일자리 사업내용_산업코드(대)
_industrial_code2	일자리 사업내용 산업코드(중)
_job	일자리 본인이 주로 하는 일
_job_code1	일자리 (가공) 본인이 주로 하는 일_직업코드(대)
_job_code2	일자리 본인이 주로 하는 일 직업코드(중)
_equipment	일자리 주로 사용하는 장비 또는 기계
_status	일자리 종사상 지위
_start_year	일자리 시작한 연도
_start_month	일자리 시작한 월
_start_season	일자리 시작한 월(계절)
_finish_year	일자리 그만둔 연도
_finish_month	일자리 그만둔 월
_finish_season	일자리 그만둔 월(계절)
_period	일자리 근속기간(조사월 기준, 개월 수)
_period_t	일자리 근속기간(조사월 기준, 4개 구분)
_working_7days	일자리 실 근무일수(일주일 평균)
_working_1day	일자리 실 근무시간(일 평균)
_type1_income	일자리 월 평균 임금_임금근로
_type1_income_t	일자리 월 평균 임금(범주)_임금근로
_type2_income	일자리 연 평균 매출액_비임금근로
_type2_income_t	일자리 연 평균 매출액(범주)_비임금근로
_impact	일자리 장애가 일하는 데 미치는 영향 정도
_satisfaction_1	일자리 업무 만족도
_satisfaction_2	일자리 만족도

## 3) 가공변수 및 무응답 대체 변수

- **가공변수:** 장애인고용패널 데이터에는 기존 변수를 이용하여 산출한 가공변수가 다수 포함되어 있다. 가공변수의 변수명은 기존 변수명 앞에 'a'를 추가하여 구분한다. 예를 들어, 8차조사 기준 패널의 실 거주지(area08)을 공표 기준인 수도권, 광역시도, 기타시도로 범주화한 지역 가공변수의 변수명은 'aarea08'이다.
- **무응답 대체:** 지난조사의 임금근로자 및 자영업자의 월평균 소득 문항에서 항목무응답이 발생한 경우, 통계적 방법을 적용하여 무응답 대체가 이루어졌다. 무응답 대체는 아래 변수에 대해 1차와 2차의 자료에 한해 실시하였으며, 기존 변수명 앞에 'm'을 추가하여 원 변수와 구분하였다.

## 〈무응답 대체 참고〉

무응답 대체를 위해 항목 무응답이 다수 발생한 변수이면서 활용도가 높은 변수를 우선적으로 검토하였으며, 임금근로자의 “월평균 소득” 및 비임금근로자의 “연간 매출액” 변수가 선정되었다.

우선적으로 단위무응답 및 항목무응답에 대한 현황을 분석한 후, 임금근로자의 월평균소득이 결측인 경우, 임금 산정방식(주당임금, 일당임금, 시간당임금)과 근무조건(근무일수, 근무시간)을 고려하여 무응답 대체를 실시하였다. 해당 방법을 적용하기 어려운 응답분류에는 적절한 통계방법론을 적용하여 산출하였다. 비임금근로자도 임금근로자와 동일한 방식으로 응답군을 분류한 후 연간매출액, 연간매출액 범주, 일자리 소재지를 고려하였으며, 적절한 통계방법론을 고려하여 무응답 대체를 실시하였다.

위와 관련하여 자세한 사항이 필요한 경우 공단 패널 담당자에게 문의하면 된다.

〈표 4-11〉 월평균 소득 변수 일람(연속형/범주형/범주형 가공/무응답 대체)

변수명	내용	비고
ca_2603	임금근로자의 월평균 소득	연속형 응답
ca_2604		범주형 응답
aca_2604		가공변수(범주형으로 recode)
<b>mca_2604</b>		<b>범주형 응답 및 완전무응답에 대한 무응답 대체</b>
cb_1701	비임금근로자의 매출액 및 소득 (연간 매출액)	연속형 응답
cb_1702		범주형 응답
acb_1702		가공변수(범주형으로 recode)
<b>mcb_1702</b>		<b>범주형 응답 및 완전무응답에 대한 무응답 대체</b>

\* 1~8차조사 모두 적용

#### 4) 직업코드 및 산업코드

- 장애인고용패널 조사자료의 산업분류와 직업분류는 한국표준산업분류(KSIC) 및 한국표준직업분류(KSCO)를 기준으로 하였으며, 대분류(1자리수)와 중분류(2자리수) 코드를 함께 제공한다.
- 3차조사 이후부터는 제10차 개정 한국표준산업분류와 제7차 개정 한국표준직업분류가 적용되었다.
- 산업과 직업 코딩 작업은 실사 과정에서 산업과 직업에 대한 정확한 정보를 조사원이 1차적으로 수집하고, 수집된 정보를 바탕으로 중앙의 전문가가 코드를 부여하는 방식을 취하고 있다.

〈표 4-12〉 한국표준산업분류 코드(제10차 개정, 대분류 기준)

코드	내용
1	농업, 임업 및 어업
2	광업
3	제조업
4	전기, 가스, 증기 및 공기 조절 공급업
5	수도, 하수 및 폐기물 처리, 원료 재생업
6	건설업
7	도매 및 소매업
8	운수 및 창고업
9	숙박 및 음식점업
10	정보통신업
11	금융 및 보험업
12	부동산업
13	전문, 과학 및 기술 서비스업
14	사업시설 관리, 사업 지원 및 임대 서비스업
15	공공 행정, 국방 및 사회보장 행정
16	교육 서비스업
17	보건업 및 사회복지 서비스업
18	예술, 스포츠 및 여가관련 서비스업
19	협회 및 단체, 수리 및 기타 개인 서비스업
20	가구 내 고용활동 및 달리 분류되지 않은 자가 소비 생산활동
21	국제 및 외국기관

〈표 4-13〉 한국표준직업분류 코드(제7차 개정, 대분류 기준)

코드	내용
1	관리자
2	전문가 및 관련 종사자
3	사무 종사자
4	서비스 종사자
5	판매 종사자
6	농림어업 숙련 종사자
7	기능원 및 관련 기능 종사자
8	장치,기계조작 및 조립종사자
9	단순노무 종사자
10	군인

## 제5장 주제별 가이드

### 1. 개인 자료

- 개인을 추적하여 조사하는 장애인고용패널조사의 특성상 자료의 분석 단위는 패널 개인이며, 개인용 자료에 포함된 주요 내용들을 살펴보면 다음과 같다.
- 본 가이드에는 각 영역별로 주요 내용에 해당하는 변수만을 기재하였다. 장애인고용패널 데이터에는 주요 변수 외의 다양한 세부적인 변수와 다수의 가공변수가 포함되어 있으므로, 변수의 상세한 구성과 가공변수 산출 기준 등을 확인하기 위해서는 별도로 제공되는 코드북을 같이 활용하는 것을 권장한다.

#### [참고사항]

5장 주제별 가이드의 변수명에서 밑줄“\_”로 기재된 부분은 조사차수 두 자리가 들어갈 자리를 비워둔 것이다. 4장에서 살펴본 바와 같이 변수명은 아래와 같은 체계로 구성되어 있으며, MAIN DATA 혹은 WIDE DATA 분석 시 조사차수에 해당하는 부분에 분석하고자 하는 차수를 기재해야 한다.

파트명	조사차수	변수번호
a	08	0310

위에 제시된 변수의 경우, 변수명을 a010310으로 수정하면 동일한 내용의 1차년도 자료를 이용할 수 있다.

LONG TYPE 통합 데이터를 이용하는 경우, 변수명에 조사차수가 포함되지 않은 변수명을 사용하므로 상기 변수명에서 조사차수를 제외한 a0310 변수를 사용하면 자료 분석이 가능하다.

### 1) 패널 기본정보

- 패널 기본정보는 지난 조사 시 응답한 정보를 확인하고 지난 조사 이후 변화사항을 질문하여 조사 시점 현재 최종 정보로 갱신하였다.
- 생년월일은 가족관계등록부나 주민등록상 등재된 생년월일과 상관없이 실제로 출생한 생년월일(양력 기준)을 기준으로 응답하였으며, 음력으로 응답한 경우 음력 생일을 양력으로 환산하여 기입하였다. 혼인상태의 경우 가족관계증명서 또는 주민등록과는 상관없이 사실상의 결혼관계를 기준으로 하였다.
- 연령 및 최종학교 관련 변수는 연령대와 학력 범주를 구분한 가공변수가 제공된다.

〈표 5-1〉 패널 기본정보 변수

항목	변수명	비고
성별	a_0205	
연령	a_0310	
혼인상태	a_0405	
최종학교	a_0505	
졸업여부	a_0510	

### 2) 장애정보

- 장애정보는 지난 조사 시 응답한 정보를 확인하고 지난 조사 이후 변화사항을 질문하여 조사 시점 현재 최종 정보로 갱신하였다.
- 장애인지시기 및 장애발생시기, 장애발생원인, 일자리에서 발생 유무 등은 1차조사 시 모든 패널에게 확인하였으며, 2차조사 이후부터는 주 장애유형이 변경된 경우에 한하여 질문하였다. 분석의 편의를 위해 전체 응답자의 장애인지시기 및 장애발생시기 가공변수가 포함되어 있다.

〈표 5-2〉 패널 장애정보 변수

항목		변수명	비고
주된 장애유형	15개 구분	a_0710	
	4개 구분	type_04	가공변수
	6개 구분	type_06	가공변수
장애등급*		a_0715	1-4차조사 장애등급제 기준 장애등급
장애정도**		a_0716	장애등급제 폐지로 5차조사에서 신규 변수
심한장애 여부(장애인복지법)***		grade_01	장애등급제 폐지로 5차조사에서 신규 가공변수
중증 여부(장애인고용촉진 및 직업재활법)****		grade_02	가공변수
중복장애 유무		a_0805	
중복장애 유형		a_0810	
장애정보 변화시기		a_0901~a_0903	주 장애정보가 하나 이상 변경된 패널
장애정보 변화사유		a_1001	
장애유형별 특성	지체	주된 형태	a_1101
		장애부위	a_110201~a_110203
		활용보장구	a_110301~a_110306t
	뇌병변	진단명	a_1201
		활용보장구	a_120201~a_120205t
	시각	독서가능여부	a_1301
		점자해독가능여부	a_1302
	청각/언어	대화가능여부	a_1401
		의사소통방법	a_1402~a_1402t
	지적/자폐성	의사소통능력	a_1501
		인지능력	a_1502
		문해력(읽기/쓰기)	a_150301~a_150302
		대중교통수단 이용능력	a_1504
	신장	투석여부	a_1601
	심장/신장 호흡기/간	이식여부	a_1602
	장애인지시기 및 발생시기		a_1701~a_1707
현재 장애상태		a_1901~a_1903	

\* 1~4차조사에서는 장애등급 1급, 2급 및 3급 장애인 중 뇌병변·시각·지적·자폐성·정신·심장·호흡기·뇌전증 장애인 및 팔에 장애가 있는 지체장애인은 중증으로, 나머지는 경증으로 분류

\*\* 장애등급제 폐지(2019년 7월)에 따라 5차조사부터 장애등급 변동 시 「장애인복지법」 기준의 장애정도로 응답, 장애의 정도가 심한 장애인이면 중증, 장애의 정도가 심하지 않은 장애인이면 경증으로 분류

\*\*\* 심한 장애 여부는 「장애인복지법」 기준으로, 심한 장애는 기존 1~3급, 심하지 않은 장애는 기존 4~6급에 해당

\*\*\*\* 중경증 여부는 「장애인고용촉진 및 직업재활법」 기준으로, 중증장애인은 1~2급에 해당하는 장애인과 뇌병변, 시각, 지적, 자폐성, 정신, 심장, 호흡기, 뇌전증 및 팔에 장애가 있는 지체장애인 3급이 해당

### [참고사항]

장애정보 변화시기 등 특정 시기를 질문하는 문항에서, 연도에 응답하고 월 문항에 응답하지 않은 경우에는 계절 범주로 재질문하여 응답을 받았다. 이는 이후 특정 날짜에 관련된 문항에서 모름/응답거절이 가능한 경우 동일하게 적용되었다.



### 3) 일자리 정보 확인 및 경제활동상태 판별

- 지난 조사 당시 응답한 일자리의 정보 확인 및 지속 여부, 지난조사 이후 변화 여부를 파악하는 문항은 2차년도 조사부터 신설되었다. 데이터에는 기존 일자리의 현재기준 최종 현황을 수록하였다.
- 어떤 일자리에서 종사상 지위가 임금근로자, 자영업자, 무급가족종사자 간에 바뀌었다면 기존 일자리를 그만두고 새로운 일자리를 시작한 것으로 간주하였다.
- 지난 조사 이후 새로 시작한 일자리가 있는 경우 신규 일자리의 개수 및 내용을 파악하는 문항이 2차년도 조사부터 신설되었다.
- 경제활동상태 판별 문항은 응답자의 지난주 취업여부, 근로시간, 구직활동 여부 등을 파악하여 취업자/구직자/비경제활동인구 구분을 위한 정보를 제공한다.

〈표 5-3〉 일자리 정보 확인 및 경제활동상태 판별 변수

항목	변수명	비고
지난조사 일자리 정보 확인	b_0119~b_0127 b_0219~b_0227 b_0319~b_0327	지난조사 당시 일자리 개수에 따라 최대 3회까지 반복 질문
신규 일자리 정보 확인	b_0501~b_0507 b_0601~b_0607 b_0701~b_0707	지난조사 이후 신규 일자리 개수에 따라 최대 3회까지 반복 질문
경제활동상태 판별	b_2701~b_3201	
월별 경제활동상태	b_3301~b_3312	

## 4-1) 취업자(근로시간)

- 취업자의 근무일수와 근무시간은 일이 불규칙하거나 일시휴직 또는 계절적 요인으로 잠시 일을 쉴 경우 등은 일이 있을 때의 통상적인 근무일수와 근무시간을 기준으로 응답받았다. 초과근로일, 초과근로시간을 포함하며 실제 근무한 시간을 기록하되, 30분 이상은 반올림하여 한 시간으로 간주하였다.
- 초과근로와 관련한 문항은 임금근로자에 한하며, 이는 정규근로시간 이외의 근로시간을 의미한다. 여기서 정규근로시간은 실제 근무시간이 아닌 일자리(직장)에서 일하기로 되어 있는 근무시간으로, 근로기준법에 정해진 근로시간이나 회사와 문서상 또는 암묵적으로 계약한 근로시간, 관습에 의해 정해진 노동시간을 의미한다. 일하는 시간이 월별, 주별 등 다르게 정해져 있으면 평균적인 시간으로 응답받았다.
- 규칙적 근로 여부는 비임금근로자에 한하며, 일거리의 많고 적음과 상관없이 일자리에 정기적으로 출근하는 것은 규칙적으로, 일자리 또는 일거리 자체가 '있다 없다' 하는 경우는 불규칙적인 것으로 간주하였다.
- 결근일수는 근무하기로 되어 있으나 장애, 치료, 치료 등으로 인해 근무하지 못한 일로 지난달을 기준으로 하였다.

〈표 5-4〉 취업자의 근로시간 관련 변수

항목	변수명	비고
규칙적 근로 여부	cb_1501	비임금근로자
주당 근무일수	ca_2501 ca_2502 cb_1502 cb_1503	취업자
일평균 근무시간	ca_2503 ca_2504 ca_2507 cb_1504 cb_1505 cb_1507	취업자
초과근로 빈도	ca_2505	임금근로자
장애/건강문제로 인한 결근여부	ca_2508 ca_2509 cb_1601 cb_1602	취업자

#### 4-2) 취업자(소득)

- 임금근로자의 월평균 소득은 세금, 보험료 등 각종 공제액을 공제하기 전의 총수령액을 기준으로 파악하고, 현물로 받은 경우는 시가로 환산하여 포함하였다. 월급 외에 상여금, 보너스, 초과근로수당 등도 포함하였다.
- 연봉, 주급 단위로 임금이 산정되는 경우 월평균으로 환산하여 포함하였으며, 일용근로자는 지난 3개월간의 평균 소득으로 기재하였다. 취업한지 아직 1개월이 되지 않은 경우에는 향후 받기로 되어 있는 임금을 기준으로 하였다.
- 비임금근로자의 소득 현황은 매출액 현황 및 소득(순수익) 여부, 창업 초기 대비 변화 등을 파악하였다. 매출액(판매액)이란 조사대상자가 운영하는 사업체에서 판매하는 물품(상품)이나 제공하는 서비스를 통해 얻게 되는 수입금액을 의미하며, 소득은 매출액에서 총비용(재료비, 인건비, 세금, 영업비용 등의 전반적인 운영비)을 제외하고 남은 금액을 의미한다.
- 창업한지 아직 1년이 되지 않은 경우에는 현재까지의 매출액(판매액)을 감안한 추정 매출액을 계산하여 응답하였다.

〈표 5-5〉 취업자의 소득 관련 변수

	항목	변수명	비고
임금근로자	임금 산정 방식	ca_2601	
	월평균 소득	ca_2603 ca_2604 aca_2604	
	체불임금 여부	ca_2701	
	체불기간	ca_2702	체불임금이 있는 경우
	체불액	ca_2703	체불임금이 있는 경우
비임금근로자	연간 매출액	cb_1701 cb_1702 acb_1702	
	창업초기 대비 매출액 변동	cb_1703	
	월평균 소득 여부	cb_1801	
	월평균 소득 여부	cb_1802 cb_1803 acb_1803	소득(순수익)이 있는 경우
	창업초기 대비 소득액 변동	cb_1804	

#### [참고사항]

월평균 소득, 매출액 등 연속형 변수 형태로 묻는 문항의 경우 모름/응답거절시 범주형으로 재질문하여 응답을 유도하였다. 이는 모든 연속형 변수 문항에 동일하게 적용되었다.

#### 4-3) 취업자(근무환경)

- 취업자의 근무환경은 일자리의 안전성 및 쾌적성, 장애편의시설 설치현황, 출퇴근 교통수단 및 소요시간 문항 등으로 구성되었다. 출퇴근 관련 문항은 2차조사부터 조사되었다.

〈표 5-6〉 취업자의 근로환경 관련 변수

항목	변수명	비고
근무장소의 안전성	ca_3401 cb_2301	취업자
근무장소의 쾌적한 정도	ca_3402 cb_2302	취업자
업무 중 질병이나 사고를 당한 경험	ca_3501 cb_2401	취업자
근무 시 발생 가능한 안전 관련 위험요소*	ca_3502~ca_3509 cb_2403~cb_2410	취업자
직장 내 재난안전 대응을 위해 필요한 장애 관련 사항**	ca_3510~ca_3518 cb_2411~cb_2419	취업자
직장 내 대피지원인 여부*	ca_3519 cb_2420	취업자
장애편의시설 설치현황	ca_3601~ca_3611 cb_2501~cb_2511	취업자
노동조합 설립여부 및 가입여부	ca_3701, ca_3702	임금근로자
사업체의 산재보험 가입여부	cb_2402	비임금근로자
출퇴근 교통수단*	ca_3801~ca_3808t cb_2601~cb_2608t	취업자
출퇴근 소요시간*	cb_2701 ca_3901	취업자

\* 2차조사 신규 문항

\*\* 7차조사 신규 문항

#### 4-4) 취업자(직무만족도)

- 취업자의 직무만족도(Job Satisfaction) 측정항목은 노동패널조사를 벤치마킹한 장애인고용패널 1차 웨이브 당시 문항에서 '교육훈련 기회' 항목을 제외한 것과 동일하다. 이는 미네소타대학교의 직업심리연구소(Vocational Psychology Research) 미네소타만족지수 (Minnesota Satisfaction Questionnaire, MSQ)와 Smith, Kendall & Hulin(1969) 등이 개발한 직무기술지표(Job Description Index: JDI) 등에 근거하여 사용되고 있는 항목들을 한국의 실정에 맞게 일부 수정하여 축약형으로 작성한 것이다.
- 임금근로자는 10개 항목, 비임금근로자는 8개 항목으로 구분되어 있으며, 현 일자리에 대한 전반적 만족도를 같이 파악한다.

〈표 5-7〉 취업자의 직무만족도 변수

항목	변수명	비고
임금 또는 소득	ca_4501 cb_3101	취업자
취업의 안정성	ca_4502 cb_3102	취업자
하고 있는 일의 내용	ca_4503 cb_3103	취업자
근로환경	ca_4504 cb_3104	취업자
근로시간	ca_4505 cb_3105	취업자
개인의 발전가능성	ca_4506 cb_3106	취업자
의사소통 및 인간관계	ca_4507 cb_3107	취업자
인사고과의 공정성	ca_4508	임금근로자
복리후생	ca_4509	임금근로자
장애에 대한 배려와 편의	ca_4510 cb_3108	취업자
현 일자리에 대한 전반적 만족도	ca_4511 cb_3109	취업자

#### 4-5) 취업자(전반적 직무만족도)

- 취업자의 전반적 직무만족도(job satisfaction) 측정항목은 노동패널조사를 벤치마킹한 장애인 고용패널 1차웨이브 당시 문항과 동일하다. 이는 직무환경을 요인별로 설명하고 있는 Brayfield & Rothe(1951)의 '직무만족지표(Job Satisfaction Index, JSI)'에서 5개 문항을 선별하였다.
- Brayfield, A. H., & Rothe, H. F. (1951). An index of job satisfaction. Journal of applied psychology, 35(5), 307.
- 임금근로자 및 비임금근로자에게 5개의 항목으로 질의하고 있다.

〈표 5-8〉 취업자의 전반적 직무만족도 변수

항목	변수명	비고
나는 현재 하고 있는 일에 만족하고 있다.	ca_3301 cb_2201	취업자
현재 하고 있는 일을 열정적으로 하고 있다.	ca_3302 cb_2202	취업자
현재 하고 있는 일을 즐겁게 하고 있다.	ca_3303 cb_2203	취업자
현재 하고 있는 일을 보람을 느끼면서 하고 있다.	ca_3304 cb_2204	취업자
별다른 일이 없는 한 현재 하고 있는 일을 계속하고 싶다.	ca_3305 cb_2205	취업자

## 5) 임금근로자

- 임금근로자는 지난 조사 이후 일자리 정보에서 파악된 현재 주업의 종사상 지위가 상용근로자, 임시근로자, 일용근로자인 경우에 해당되며, 해당 주업의 일자리 정보, 고용형태 및 업무 정보, 근로조건 등을 1차년도 이후부터 지속적으로 조사하고 있다.
- 현재 주업이 1차조사와 동일한 경우, 취업현황 관련 문항을 질문하지 않고 대신 지난조사 대비 일자리의 전반적인 변화사항을 확인한다. 근무환경 중 출퇴근 환경에 대한 문항은 2차조사에 새로 추가되었다. 부업 문항은 1차년도에만 조사된 후 제외되었다.
- 최저임금과 관련된 문항은 노동이슈 현안 파악 및 반영을 위해 4차조사에서 한시적으로 추가된 후 제외되었고, 6차조사에서는 코로나19 지속으로 인한 일자리(직장) 내 어려움, 고용안정 변화 등 문항 등 5차조사에 신설되었던 문항을 보완하였다.
- 7차조사에서는 시의성에 맞지 않는 마스크 참여 여부나 직장의 코로나 안전 여부 문항을 제외하여 조사하였다.
- 8차조사에서는 한국종사상지위분류 개정 사항 반영을 위한 일부 문항 추가가 이뤄졌으며, 해당 문항을 통해 기존 및 신규 임금근로자 분류를 동시에 조사하였다.

〈표 5-9〉 임금근로자 주요 변수

항목	변수명	비고
일자리 기본정보 (사업체명, 소재지, 사업내용, 업무내용 등)	ca_0101~ca_0302	
취업현황 (근무시작시기, 취업배경 등)	ca_0401~ca_1001t	현재 주업이 변경된 경우에만 응답
일자리 주요 특성 (일자리 분류, 근로자 수 등)	ca_1101~ca_1302	
고용형태 (유급견습, 훈련생 및 인턴 파견/용역 현황, 근로계약기간 등)	ca_1401~ca_2402	8차조사에 고용형태 세분화를 위한 문항 추가
근무일 및 근무시간	ca_2501~aca_2510	
업무특성 (업무시작시기, 능력 및 기술, 전반적 직무만족도)	ca_3001~ca_3306	
근무환경 (안전성, 장애인시설, 출퇴근 현황 등)	ca_3401~ca_3901	
차별경험	ca_4001~ca_420109t	
애로사항 및 이직 직무만족도	ca_430101~ca_4403 ca_4501~ca_4511	
지난조사 대비 일자리 변화	ca_4601~ca_4605	현재 주업이 지난조사와 동일한 경우에만 응답
임금 및 복리후생	ca_2601~ca_2908	
최저임금 (최저임금 수준, 일자리에 대한 영향)	ca_5201, ca_5301~ca_5304	4차조사 한시적 문항
정당한 편의제공 경험	ca_540101, ca_540101t	
직장을 계속 다니기 위해 필요한 사항	ca_550101~ca_550103, ca_5501t	
직장 내 인식개선 교육 실시 여부 및 효과	ca_5601~ca_5604t	5차조사 한시적 문항
직장 내 인식개선 교육 경험 및 연도	ca_540101~ca_5605	6차조사 신규 문항
코로나19 지속으로 인한 일자리(직장) 내 어려움, 고용안정 변화	ca_5701~ca_5719, ca_5801	

- 비정규직 판별 문항은 통계청 경제활동실태조사의 기준을 적용하여 한시적근로, 시간제근로, 비전형 근로(파견/용역, 특수형태, 일일, 가정 내 근로) 여부를 확인하는 문항으로 구성되었으며, 개인의 주관적인 비정규직 여부 자기판단 문항이 포함되었다.
- 3차조사에는 비정규직 응답자를 대상으로 현 일자리 내 동일(유사)업무 수행 정규직 유무, 전환 가능여부 및 방법, 비정규직 처우개선 필요 분야 등의 문항이 추가되었다.

〈표 5-10〉 비정규직 판별 변수

항목	변수명	비고
근로계약 현황	ca_1401~ca_1404	
일일근로 여부 및 유형	ca_1501~ca_1503	
고용지속가능성 판단	ca_1601	
현 일자리의 기대근속기간	ca_1603, ca_1604	
시간제 근로 여부	ca_1701	
파견(용역) 근로 여부	ca_1801	
가정 내 근로 여부	ca_2201	
특수형태근로종사자 여부	ca_2202	
비정규직 여부 자기판단	ca_2301	
현 일자리 내 동일(유사)업무 수행 정규직 유무, 전환가능여부 및 방법, 업무수준 차이 등	ca_4701~ca_4805	
비정규직 처우개선 필요 분야	ca_490101, ca_490102	

## 6) 비임금근로자

- 비임금근로자는 지난 조사 이후 일자리 정보에서 파악된 종사상 지위가 자영업자 혹은 무급가족 종사자인 경우에 해당하며, 매출액, 순수익, 사업체 현황, 업무특성 및 근무환경 등을 변수를 1차년도 이후부터 지속적으로 조사해왔다.
- 임금근로자 문항과 마찬가지로 현재 주업이 1차조사와 동일한 경우, 취(창)업현황 관련 문항을 질문하지 않는다. 근무환경 중 출퇴근 환경에 대한 문항은 2차조사에 새로 추가되었으며, 부업 문항은 1차년도에만 조사된 후 제외되었다.
- 6차조사에서는 임금근로자 문항과 마찬가지로 5차조사에 신규로 도입되었던 코로나19 지속으로 인한 사업체 운영의 어려움, 사업체 운영상황 변화 등 사회변화에 대응하기 위한 문항을 보완 하였다.
- 7차조사에서는 시의성에 맞지 않는 코로나19 지원금의 도움정도, 마스크 참여 여부나 직장의 코로나 안전 여부 문항을 제외하여 조사하였다.

〈표 5-11〉 비임금근로자 주요 변수

항목	변수명	비고
사업체명, 소재지, 사업내용, 업무내용	cb_0101~cb_0301	
근무시작시기, 창업배경 및 초기투자비용	cb_0401~cb_1108t	현재 주업이 변경된 경우에만 응답
사업체 주요 특성(사업체 분류, 종업원 수 등)	cb_1201~cb_1402	
근무일 및 근무시간	cb_1501~cb_1602	
매출액 및 소득	cb_1701~cb_1804	
업무시작시기, 능력 및 기술, 전반적 직무만족도	cb_1901~cb_2206	
안전성, 장애편의시설, 출퇴근 현황 등*	cb_2301~cb_2701	
애로사항 및 이직	cb_280101~cb_3003	
직무만족도	cb_3101~cb_3109	
사업체 운영 및 성과에 대한 평가	cb_3201	
코로나19 지속으로 인한 사업체 운영의 어려움, 사업체 운영상황 변화	cb_3501~cb_3510, cb_3601	

\* 7차조사 시 안전 관련 위험요소 문항이 신규로 추가됨

## 7) 미취업자

- 경제활동상태 판별 결과 미취업자에 해당하는 실업자와 비경제활동인구가 응답하는 문항으로, 향후 취업 희망여부와 희망하는 일자리, 구직활동 경험 등을 파악하였다. 향후 취업 희망여부 및 희망하는 일자리 정보, 구직활동 경험 등은 1차년도부터 지속적으로 조사해오고 있으며, 미취업자의 마지막 취업경험은 1차년도에만 조사하였다.
- 희망하는 일자리(직장)에 대한 정보는 실업자 및 향후 취업을 희망하는 비경제활동인구 즉, 취업 희망 미취업자를 대상으로 질문하였으며, 구직활동 문항은 실업자 전체와 비경제활동인구 중 지난조사 이후 구직활동 경험이 있는 패널에게 파악하였다.
- 취업자와 마찬가지로 7차조사에서는 6차조사에 이어 미취업자에게 코로나19 지속이 구직활동에 미친 영향 파악을 위해 관련 문항을 추가하였다.

〈표 5-12〉 미취업자 주요 변수

항목	변수명	비고
향후 취업의사	cc_0101, cc_0102	
향후 취업 희망이유	cc_0201	
취업 희망 일자리 정보	cc_0301~cc_0311	
구직활동 경험	cc_0401~cc_1001	
코로나19 지속이 구직활동에 미친 영향	cc_1301	



8) 취업을 위한 노력과 지원

- 취업 및 취업유지 활동 경험은 1차조사에는 최근 3년간의 경험을, 2차조사 이후부터는 지난조사 이후 경험을 조사하였다. 해당 활동 경험이 있는 응답자에 한하여 취업 및 취업유지활동의 성과 정도를 항목별로 응답하였다.

〈표 5-13〉 취업 및 취업유지활동 관련 변수

항목	변수명		비고
	경험여부	성과정도	
인터넷을 통해 채용, 창업정보 탐색	d_0101	d_0201	
신문, 취업정보지, 전단지(벽보)를 통해 채용, 창업정보 탐색	d_0102	d_0202	
공공 취업알선기관에 구직등록(인터넷 포함), 취업알선	d_0103	d_0203	지난조사 이후 구직경험이 없는 미취업자 제외
공공근로, 자활근로, 장애인 복지일자리 등 국가/지자체 창출 재정지원 일자리사업 참여 신청 및 근로	d_0104	d_0204	
민간 취업알선업체에 구직등록, 취업알선	d_0105	d_0205	
주변 지인에게 일자리 정보 제공이나 취업 부탁	d_0106	d_0206	
직접 일자리(사업체)에 전화하거나 방문하여 채용 문의	d_0107	d_0207	
다른 일자리(사업체)에 이력서 제출	d_0108	d_0208	취업자
다른 일자리(사업체)의 면접 참여	d_0109	d_0209	취업자
공공/민간/사업체/학교의 직업교육훈련 참여	d_0110	d_0210	
사설학원에서 외국어, 자격 취득, 채용시험 등 준비	d_0111	d_0211	
자격증을 취득하기 위한 준비	d_0112	d_0212	
일자리(사업체)에서의 현장실습이나 인턴	d_0113	d_0213	
채용시험 준비	d_0114	d_0214	
창업을 위한 활동	d_0115	d_0215	지난조사 이후 구직경험이 없는 미취업자 제외
국가 등으로부터 장애인 자립, 창업 자금 등 지원, 융자	d_0116	d_0216	
기타	d_0117	d_0217	

- 공공 취업알선기관 이용 경험이 있거나, 공공, 민간, 회사, 사설 직업교육훈련을 받은 경험이 있는 응답자의 경우, 경험 횟수에 따라 최근순으로 최대 다섯 개까지 관련 내용을 응답받았으며, 이용 횟수만큼 관련 변수가 생성되어 있다.
- 공공 취업알선기관 이용경험은 고용센터, 한국장애인고용공단, 지자체(행정복지센터(주민센터)), 복지관, 장애인 단체/협회 등의 기관에 구직등록(인터넷 포함)을 하거나 취업알선을 받는 경우를 의미하며, 공공근로, 자활근로, 장애인 복지일자리 등은 해당하지 않는다.
- 직업교육훈련 이용경험은 산업인력공단, 장애인고용공단, 직업전문학교, 기능대학 등 공공기관 및 교육기관에서 훈련을 받았거나, 다니고 있는 또는 다녔던 회사의 재직자 교육훈련, 창업교육, 사설학원에서 취업을 위한 외국어, 자격 취득, 채용시험 등을 준비한 경우를 포함한다.

〈표 5-14〉 공공 취업알선기관 및 직업교육훈련 이용경험 변수

항목	변수명	비고
공공 취업알선기관 이용 횟수	d_0401	
공공 취업알선기관 이용 현황 (구직등록 기관, 시기, 알선 횟수 및 만족도 등)	d_0501~d_0511 d_0601~d_0611 d_0701~d_0711 d_0801~d_0811 d_0901~d_0911	이용경험 횟수에 따라 최대 5회까지 반복 질문
직업교육훈련 이용 횟수	d_1001	
직업교육훈련 이용 현황 (교육 종류, 교육기간, 분야, 도움된 정도 등)	d_1101~d_1120 d_1201~d_1220 d_1301~d_1320 d_1401~d_1420 d_1501~d_1520	이용경험 횟수에 따라 최대 5회까지 반복 질문

- 취업을 위해 필요한 사항 및 일자리(근무지)에서 장애로 인해 필요한 지원사항에 대해서는 향후 취업의사가 없는 비경제활동인구를 제외한 모든 응답자에게 조사하였다. 해당 문항은 현재 취업 여부와 무관하게 향후 이직 혹은 취업하게 되면 필요한 내용을 응답하도록 하였다.
- 취업자에게는 직업생활을 유지하는데 장애로 인해 추가로 소요되는 비용을 지출 항목별로 응답받았다. 단, 2, 3차조사는 소요 비용을 개인 지출과 외부 지원(국가, 민간 등) 비용으로 구분하여 응답받았고, 4차조사 이후부터는 이를 단순화 및 범주화하여 받았으므로 1차조사와 이후 조사년도 데이터의 변수명이 동일한 일련번호로 구성되어 있지 않다. 해당 변수 분석 시에는 코드북을 통해 변수명 구조를 이해한 후 분석하는 것을 권장한다.

〈표 5-15〉 취업관련 필요사항 및 지원사항, 추가 소요비용 관련 변수

항목	변수명	비고	
취업을 위해 필요한 사항	필요 정도	d_1601~d_1613t	향후 취업의사가 없는 비경제활동인구 제외
	시급한 정도	d_1701~d_1713	
일자리(근무지)에서 장애로 인해 필요한 지원사항	필요 정도	d_2001~d_2009t	
	지원/활용 여부	d_2101~d_2109	
	도움 정도	d_2201~d_2209	
직업생활 유지에 장애로 인해 추가로 소요되는 비용	출퇴근, 이동을 위한 교통비	d_2701~d_2704, d_3601, d_360101t	취업자만 응답 ※ 1차조사와 변수명 구조가 상이함
	장애인보조기구 구입/유지비	d_2801~d_2804, d_3602, d_360201t	
	직장생활 지원을 위한 인력 활용비	d_2901~d_2904, d_3603, d_360301t	
	직장생활 유지를 위한 장애치료 및 투약비	d_3001~d_3004, d_3604, d_360401t	
	장애치료, 장애로 인한 휴식 등을 위한 휴가 비용	d_3101~d_3104, d_3605, d_360501t	
	기타	d_3201~d_3204, d_3606, d_360601t	
	추가비용의 부담정도	d_3401	
직업생활 유지를 위한 추가 지원 희망항목	d_3501		

### 9) 직업적 능력

- 취업관련 업무능력, 경력, 사회적 지지 정도 등 응답자의 직업적 능력 전반을 1차조사 이후 동일하게 조사하였다.
- 대학교 전공은 4차조사 이후 「한국표준교육분류」 제정 및 시행으로 인해 기준을 재정비하여 1~3차조사와 비교 시 유의해야 한다.

〈표 5-16〉 직업적 능력 주요변수

항목	변수명	비고
취업관련 업무능력	e_0101~e_0115	
주관적 근로능력 평가	e_0201	
일자리에서 일을 한 기간(누적)	e_0301	
가장 오래한 직무를 수행한 기간	e_0302	
(고등학교) 학교 유형	e_1401	
(대학교) 전공 및 대학 소재지,	e_1402~e_1403, e_1704	「한국표준교육분류」 제정에 따른 기준 재정비
(대학교) 장애학생 특별전형 입학여부	e_1404	
취업에 도움을 줄 수 있는 지인의 수	e_1501~e_1508	
주변 사람과 관계를 유지하는 정도	e_1601~e_1607	

- 자격증 관련 정보는 자격증 이름, 유형, 취득시기, 취업 및 업무 관련 도움정도에 대해 1차년도 부터 동일하게 조사되었다.
- 1차조사 시에는 보유하고 있는 자격증 전체 현황을 조사하였고, 2차조사 이후에는 지난조사 당시 응답한 자격증 현황을 TAPI 시스템을 통해 확인한 후 지난조사 이후 새로 취득한 자격증에 대해 응답하였다. 2차조사 이후의 메인 데이터에는 신규 취득 자격증 변수와 조사시점 기준 누적 현황 가공변수가 모두 포함되어 있으나, 지나치게 변수 개수가 많아지는 것을 방지하기 위해 WIDE TYPE/LONG TYPE 통합 데이터에는 누적 현황 가공변수만 수록하였다.

〈표 5-17〉 자격증 유형 분류 기준

자격증 유형	자격증 종류
국가전문자격증	변호사, 법무사, 회계사, 세무사, 의사, 약사 등
국가기술자격증 (기술사/기사/기능장/산업기사 (국가기술자격1급)/기능사(국가 기술자격2급)/기타국가자격증)	① 용접, 유체기계, 산업기계, 건설기계, 공조냉동기계, 발송배전, 건축전기 설비, 자동차 정비, 자동차검사, 보일러, 금속, 정보통신, 토질 및 기초, 토목구조, 도로 및 공항, 수자원개발, 상하수도, 토목시공, 측량 및 지형 공간정보, 건축구조, 건축기계설비, 건축시공, 정보관리, 전자계산조 직응용 등 이공계열 ② 시각디자인, 제품디자인, 조리, 제빵, 미용사, 세탁 등 ③ 직업상담사, 사회조사분석사, 한글속기사, 워드프로세서, 컴퓨터 활용능력, 전산회계사, 전자상거래, 비서 등 전문사무영역의 자격증은 1급일 경우 산업기사, 2급~3급일 경우 기능사에 해당됨
민간자격증	정보검색사, TEPS, 한자능력급수, 번역사 1~3급, 증권분석사, 1종 투자 상담사, 정보검색사 1~3급, 사회보험관리사 1~3급, 웹프로그래머 1~3급, PC 정비사 1~2급, 인터넷 정보검색사 1~2급, 무역영어, 증권분석사, 댄스스포츠 지도사 등
국제(외국)자격증	일본어능력시험(JLPT), TOEIC, TOEFL, 미국공인회계사(AICPA), 국제재무 분석사(CFA), 선물중개인(AP) 등

〈표 5-18〉 자격증 주요 변수

항목	변수명	비고	
지난조사 이후 신규 자격증 현황	신규 자격증 보유 여부 및 개수	e_0401, e_0402	신규 취득 자격증 개수에 따라 최대 9회까지 반복 질문
	자격증 이름	e_0501	
	자격증 취득시기	e_0502~e_0504	
	자격증 형태	e_0505	
	자격증 분야	e_0506	메인 데이터에만 수록
	자격증 취득이유	e_0510	
	현 일자리 취업에 도움된 정도	e_0507	
	현 업무수행에 관련된 정도	e_0508	
향후 (다른)일자리 취업에 도움되는 정도	e_0509		

〈표 5-19〉 자격증 주요 변수(계속)

항목		변수명	비고
누적 자격증 현황	자격증 보유 여부 및 개수(누적)	ae_0401, ae_0402	조사년도 현재 보유 자격증 개수에 따라 최대 9개 자격증 정보 기재.  메인 데이터 및 통합 데이터에 수록
	자격증 ID 및 조사차수	ae_050001, ae_050002	
	자격증 이름	ae_0501	
	자격증 취득시기	ae_0502~ae_0504	
	자격증 형태	ae_0505	
	자격증 분야	ae_0506	
	자격증 취득이유	e_0510	
	현 일자리 취업에 도움된 정도	ae_0507	
	현 업무수행에 관련된 정도	ae_0508	
	향후 (다른)일자리 취업에 도움되는 정도	ae_0509	

#### 10) 자기효능감/자아존중감/장애수용정도

- 취업관련 태도 및 환경 문항에서는 자기효능감/자아존중감/장애수용정도 각각에 대한 4점 혹은 5점 척도 문항을 응답한다. 일부 문항은 분석 시 역코딩이 필요하며, 해당 변수는 응답값을 역코딩한 가공변수가 제공된다.
- 항목별 척도의 출처는 아래와 같다.
  - 자기효능감(General Self-Efficacy): Schwarzer, R., & Jerusalem, M.(1995)이 개발하고, Young-Min Lee., Schwarzer R., and Jerusalem, M.(1994)에 의해 한국판으로 타당화한 일반적 자기효능감 척도 10문항 사용
  - Schwarzer, R., & Jerusalem, M. (1995). Generalized Self-Efficacy scale. In J. Weinman, S. Wright, & M. Johnston, Measures in health psychology: A user's portfolio. Causal and control beliefs (pp. 35-37). Windsor, UK: NFER-NELSON.
  - Young-Min Lee., Schwarzer R., and Jerusalem, M. (1994), Korean Adaptation of the General Self-Efficacy Scale, <http://userpage.fu-berlin.de/~health/korean.htm>
  - <http://userpage.fu-berlin.de/~health/korean.htm> 참고
  - 자아존중감: Rosenberg(1965)에 의해 개발되고 전병재(1974)에 의해 번안된 자아존중감 척도 10문항 사용
  - Rosenberg, M. (1965). Society and the adolescent self-image. Princeton, NJ: Princeton University Press.
  - 전병재(1974), 자아개념 측정가능성에 관한 연구, 연세대학교대학원 학회지, 11(1), 107-130.
  - 장애수용도: Kaiser, Wingate, Freeman & Chandler(1987)의 장애수용척도(Disability

Acceptance Scale; DAS) 9문항 및 백영승, 김성희, 유미, 김춘수(2001)가 개발하여 강용주, 구인순, 박자경(2008)이 타당화 연구를 실시한 자아수용검사의 장애극복요인 3문항 추가로 총 12개 문항 사용

- Kaiser, S. B., Wingate, S. B., Freeman, C. M., & Chandler, J. L. (1987). Acceptance of physical disability and attitudes toward personal appearance. *Rehabilitation Psychology*, 32(1), 51.
- 백영승, 김성희, 유미, 김춘수. (2001). 자아인식검사 개발 연구보고서. 성남: 한국장애인고용공단 고용개발원
- 강용주, 구인순, 박자경. (2008). 자아 인식, 자아 수용 검사 타당화 연구. 성남: 한국 장애인 고용공단 고용개발원.

〈표 5-20〉 자기효능감 변수

	항목	변수명	비고
자기 효능감	어려운 일이라도 열심히 노력한다면 나는 해낼 수 있다	f_0101	
	목표에 집중해서 그것을 이루어내는 것은 나에게 쉬운 일이다	f_0102	
	나는 나의 능력을 믿기 때문에 어려움에 부딪쳐도 당황하지 않는다	f_0103	
	필요한 노력을 기울이기만 한다면 나는 대부분의 문제를 해결할 수 있다	f_0104	
	어려움이 처하더라도 나는 해결책을 찾을 수 있을 것이다	f_0105	
	나는 예상치 못한 일이라도 효율적으로 처리해 낼 것이라고 믿는다	f_0106	
	내 재능 덕분에 나는 예상치 못한 상황에 어떻게 대처해야 할지 알고 있다	f_0107	
	어떤 문제에 부딪칠 때 나는 보통 몇 가지 해결방법을 찾을 수 있다	f_0108	
	누군가가 내 의견에 반대하더라도 나는 내가 원하는 방식으로 할 수 있는 방법을 찾을 수 있을 것이다	f_0109	
	나에게 어떤 일이 생기더라도 나는 해결해 낼 수 있을 것이다	f_0110	

〈표 5-21〉 자아존중감 변수

	항목	변수명	비고
자아 존중감	나는 다른 사람처럼 가치 있는 사람이라고 생각한다	f_0601	
	나는 좋은 성품을 가졌다고 생각한다	f_0602	
	나는 대체적으로 실패한 사람이라고 생각한다	f_0603	
		af_0603	가공변수(역코딩)
	나는 다른 사람들만큼 일을 잘 할 수가 있다	f_0604	
	나는 자랑할 것이 별로 없다	f_0605	
		af_0605	가공변수(역코딩)
	나는 내 자신에 대해 긍정적인 태도를 가지고 있다	f_0606	
	나는 나 자신에 대하여 대체적으로 만족한다	f_0607	
	나는 나 자신을 좀 더 존중할 수 있으면 좋겠다	f_0608	
af_0608		가공변수(역코딩)	
나는 가끔 내 자신이 쓸모 없는 사람이라는 느낌이 든다	f_0609		
	af_0609	가공변수(역코딩)	
나는 때때로 내가 좋지 않은 사람이라고 생각한다	f_0610		
	af_0610	가공변수(역코딩)	

〈표 5-22〉 장애수용정도 변수

	항목	변수명	비고
장애 수용 정도	장애인이기 때문에 사람들을 잘 사귀지 못한다	f_0701	
		af_0701	가공변수(역코딩)
	장애 때문에 세상을 더 넓게 생각하게 되었다	f_0702	
	장애 때문에 할 수 없는 일이 생기면 속상하다	f_0703	
		af_0703	가공변수(역코딩)
	나는 장애 때문에 괴로워하지 않는다	f_0704	
	나는 장애인이지만 내 인생에 만족한다	f_0705	
	장애 자체보다 인생을 어떻게 사느냐가 더 중요하다	f_0706	
	장애는 내 인생에 가장 많은 영향을 미친다	f_0707	
		af_0707	가공변수(역코딩)
	장애 자체보다 정직이 더 중요하다	f_0708	
	인생은 외모보다 훨씬 더 중요한 것들이 많다	f_0709	
장애인이라는 것을 잊고 살만큼 재미있는 일이 많다	f_0710		
장애를 가지고 있지만 내 인생은 부족함이 없다	f_0711		
장애 때문에 불편하지만 마음만 먹으면 무엇이든 할 수 있다	f_0712		

## 11) 일상생활과 삶의 질

- 건강상태 및 생활습관, 일상생활 시 도움 필요여부 및 도움제공자, 일상 및 여가활동, 차별경험 등 일상생활 관련 문항은 대부분의 문항이 1차조사 이후 큰 변동 없이 유지되고 있다.
- 다만, 8차조사의 경우 여가시간 외에도 가사노동 시간, 돌봄 노동시간과 배우자의 분담 수준에 대한 충분도를 확인하는 문항이 일부 추가되었다.

〈표 5-23〉 일상생활 관련 주요 변수

항목	변수명	비고
전반적 건강상태	g_0101~g_0103	
만성질환 여부	g_0201	
만성질환 종류	g_020201~g_020218t	
만성질환 치료활동 여부	g_0203	
운동, 수면, 식사습관	g_0301~g_0502	
흡연 및 음주 여부	g_0601~g_0702	
일상생활 중 스트레스, 우울경험 여부	g_0801~g_0803	
주관적 행복감 정도	g_0804	
일상생활 중 타인의 도움 필요 여부	g_0901	
도움제공자 유무	g_0902	
주요 도움제공자	g_090301, g_090302	
평소 주로 하는 활동	g_100101~g_100103	
여가활동 및 사회참여 활동 경험여부	g_1101~g_1301	8차조사 시 여가 시간 외에도 가사 및 돌봄 노동에 대한 현황 파악
외출 시 이동수단 및 이동시 어려움	g_140101~g_1402t	
일상에서의 차별경험	g_1501	
일상생활 만족도, 주관적 사회경제적 지위	g_1601~g_1701	



## 12) 개인소득

- 패널의 개인소득은 지난 1년간 발생한 세후 총 근로소득(임금소득+사업소득)을 기준으로 조사하였으며 전체 패널 중 근로소득이 발생한 경우에만 응답받도록 하였다.
- 자영업자가 적자일 경우, 임금근로자가 임금이 체불된 경우에는 일했음에도 소득이 '0'일 수 있다. 응답 시에는 주업과 부업을 따로 나누지 않고 임금소득과 사업소득으로만 나눠서 응답하였다.
- 임금소득/사업소득 중 한 항목에서라도 모름/응답거절한 경우는 범주형 문항을 통해 재응답 받도록 하였으며, 응답 항목의 중간값으로 소득의 총계를 계산하였다. 즉, 임금소득에서 모름/응답거절 후 범주형에서 '120만원 이상 240만원 이하'를 응답한 경우 중간값인 180만을 총계를 계산하는데 사용하였다.

※ '60만원 미만'인 경우에는 60만원, '6억원 이상'으로 응답한 경우에는 6억원으로 계산하여 총계 산출

〈표 5-24〉 패널 개인소득 관련 변수

항목	변수명	비고
연간 임금소득*	h_2301	
	h_2302	범주형 응답
	ah_2301	범주형 응답 포함 가공변수
연간 사업소득**	h_2305	
	h_2306	범주형 응답
	ah_2305	범주형 응답 포함 가공변수
연간 근로소득	ah_2309	범주형 응답 포함 가공변수

\* 임금소득: 고용주와의 계약을 통해 노동의 대가로 받게 되는 보수. 조사대상자가 회사나 공장 등에 직원으로 입사를 해서 근로의 대가로 받는 보수를 통해 얻는 소득

\*\* 사업소득: 자영업(노점 포함)이나 사업주의 경우, 자신의 사업체 운영을 통해 얻는 소득

## 13) 노후준비

- 노후준비 관련 문항은 조사시점 기준 만 40세 이상인 패널에 한하여 질문하였다.

〈표 5-25〉 노후준비 관련 변수

항목	변수명	비고
노후생활 중 염려되는 사항	h_400101	만 40세 이상인 경우 응답
	h_400102	
	h_4001t	
노후생활을 위한 경제적 준비 여부	h_4002	
노후생활을 위한 경제적 준비 사항	h_400301	
	h_400302	
	h_4003t	
경제적 준비 정도	h_4004	

## 14) 코로나19

- 코로나19 상황 아래 한시적으로 코로나19의 영향력 등에 대해 파악하고자 6차조사에 최초로 신설되었고, 7차조사 및 8차조사에도 유지되었다.
- 코로나19 상황이 지속됨에 따라 7차조사에서는 예방접종, 확진 경험 등과 같은 시의성에 맞는 문항을 추가하였다.
- 8차조사에서는 예방접종, 확진 경험 등과 같은 시의성에 맞는 문항을 추가하였다.

〈표 5-26〉 코로나19 관련 주요 변수

항목	변수명	비고
패널 본인 및 가구원 코로나19 검사 여부	i_0401~i_0403	6차조사 한시적 문항
패널 본인 및 가구원 코로나19 검사 결과	i_040501, I_040502	6차조사 한시적 문항
패널 본인 및 가구원 자가격리 여부	i_0501~i_0503	6차조사 한시적 문항
코로나19 예방접종 횟수	i_1201	7차조사 신규 문항
코로나19 확진 경험 여부	i_1202	7차조사 신규 문항
코로나19 확진 시기	i_1203~i_1204	7차조사 신규 문항
코로나19 확진 이후 건강 회복 상태	i_1301~i_1303	7차조사 신규 문항
코로나19 지속으로 인한 변화 사항	i_0601~i_0611	7차조사 신규 문항
코로나19 지속 관련 필요 지원 사항	i_140101~i_140102	7차조사 신규 문항
코로나19 확산으로 인해 어려워진 사항	i_070101, i_070102	6차조사 한시적 문항
일상생활에서 코로나19 감염 위험 체감 수준	i_0801	
코로나19 인식	i_0901~i_0904	6차조사 한시적 문항
삶의 질 변화	i_1001	6차조사 한시적 문항
코로나19 백신 접종 여부 및 의사	i_1101, i_1102	6차조사 한시적 문항

## 2. 가구 자료

- 개인을 추적 조사하는 장애인고용패널의 특성상 대부분의 조사 내용이 개인을 기준으로 구성 되어 있으나, 패널이 속한 가구의 일반사항을 같이 파악하고 있다. 가구 관련 내용은 가구원 정보, 가구 소득 및 소비, 자산, 부채, 주거형태 등을 조사하였다.

### 1) 가구원 정보

- 가구란 혈연관계를 떠나 1인 또는 2인 이상이 모여서 취사, 취침 등 함께 생계를 같이 하는 생활단위로, 주민등록 상 세대와 가족의 개념과는 다르다. 가구원은 혈연관계와는 상관없이 생계를 함께 하고 동거하는 구성원을 말하며, 비혈연 동거인도 생계를 같이 하면 가구원에 포함된다. 반면 혈연인 가족이라 하더라도 생계를 달리 하면 가구원에 포함되지 않는다. 단, 타 지역 근무, 학업, 군복무, 치료 및 요양 등의 이유로 일시적으로 떨어져 살고 있는(다시 돌아올 것으로 전제된) 직계가족은 가구원에 해당한다.
- 가구원 정보는 가구원 1인에 대한 일련의 정보가 하나의 세트로 구성되어 있다. 즉 가구원 1명에 대한 성별, 패널과의 관계, 생년월, 장애여부, 혼인상태, 취업상태 등의 정보가 한 묶음으로 제공 되고, 이후에 다른 가구원의 정보가 제시되는 식이다.
- 가구주란 가족관계 또는 세대주와 관계없이 가구를 대표하고 가계를 책임지고 있는 생계책임자를 말한다. 본 조사에서는 가구원 현황을 모두 응답한 후 패널을 포함한 가구원 중 가구주 한 명을 응답하도록 하였다.
- 1차조사에는 가구를 이루어 살고 있는 가구원 전체에 대해 응답받았다. 2차조사 이후에는 기존 가구원의 현재 지속 여부 및 기존 응답의 수정/변화 여부를 확인한 후, 지난조사 이후 새로 추가된 가구원 현황을 응답받았다.
- 2차조사 이후 데이터에는 기존, 신규 가구원 현황을 종합한 최종 가구원 가공변수가 생성되어 있다. 메인 데이터에는 기존 가구원 변수, 신규 가구원 변수, 최종 가구원 변수가 모두 포함되어 있으며, WIDE/LONG TYPE 통합 데이터에는 최종 가구원 변수만 수록하였다.
- 기존 가구원 변수는 이전조사 당시 응답한 가구원 정보를 포괄한다. 당시 응답한 내용의 수정 및 변화여부를 확인하여 최종 정보로 반영한 데이터가 수록되어 있다.

〈표 5-27〉 기존 가구원 관련 변수\*

항목	변수명	비고
기존 가구원 중 지속가구원 수	h_0101	메인 데이터에만 수록
가구원 지속여부	h_0203	
가구원 제외사유	h_0204, h_0204t	
성별	h_0206	
패널과의 관계코드	h_0208	
출생년도	h_0210	
연령	ah_021001	
출생월	h_0211, h_0213	
최종학교	h_0215	
졸업여부	h_0217	
등록장애인 여부	h_0219	
장애유형	h_0221	
장애등급	h_0231	
장애부위	h_0225	
혼인상태	h_0227	
취업상태	h_0229	

\* 가구원 지속여부(h\_0203)부터 취업상태(h\_0229) 변수까지가 가구원 1명에 대한 정보를 포함한 한 세트의 구성되어 있다. 상기 변수명은 첫 번째 가구원에 해당하는 변수이며, 기존 가구원 중 지속가구원 수(h\_0101)만큼의 데이터가 수록되어 있다.

\* 장애등급제 폐지로 인해 5차조사에서부터 장애등급을 기존 1~6급에서 장애의 정도가 심한 장애, 장애의 정도가 심하지 않은 장애로 이원화하여 변수도 수정되었다.(기존 h\_0223)

- 신규 가구원 변수는 지난조사 이후 전입, 결혼, 출생 등으로 추가된 가구원에 대한 정보를 기존 가구원과 동일하게 응답한다.

〈표 5-28〉 신규 가구원 관련 변수\*

항목	변수명	비고
지난조사 이후 신규 편입 가구원 수	h_1101	메인 데이터에만 수록
편입 사유	h_1202, h_1202t	
성별	h_1203	
패널과의 관계코드	h_1204	
출생년도	h_1205	
연령	ah_120501	
출생월	h_1206, h_1207	
최종학교	h_1208	
졸업여부	h_1209	
등록장애인 여부	h_1210	
장애유형	h_1211	
장애등급	h_1216	
장애부위	h_1213	
혼인상태	h_1214	
취업상태	h_1215	

\* 편입 사유(h\_1202)부터 취업상태(h\_1215) 변수까지가 가구원 1명에 대한 정보를 포함한 한 세트에 구성되어 있다. 상기 변수명은 첫 번째 신규 가구원에 해당하는 변수이며, 지난조사 이후 신규 편입 가구원 수(h\_1101)만큼의 데이터가 수록되어 있다.

\* 장애등급제 폐지로 인해 5차조사에서부터 장애등급을 기존 1~6급에서 장애의 정도가 심한 장애, 장애의 정도가 심하지 않은 장애로 이원화하여 변수도 수정되었다.(기존 h\_0212)

- 기존 가구원 변수와 신규 가구원 변수를 통합한 최종 가구원 변수는 기존 가구원 중 현재 지속 중인 가구원과 신규 가구원 정보가 모두 포함되어 있다. 최종 가구원 수는 최대 9명이며, 가구원 ID와 조사차수를 구분하는 변수가 추가되어 있다.

〈표 5-29〉 최종 가구원 현황 가공변수\*

항목	최종 가구원 1	최종 가구원 2	...	최종 가구원 9
가구원 id	ah_020001	ah_030001		ah_100001
조사차수	ah_020002	ah_030002		ah_100002
가구주 여부	ah_020003	ah_030003		ah_100003
성별	ah_0206	ah_0306		ah_1006
패널과의 관계코드	ah_0208	ah_0308		ah_1008
출생년도	ah_0210	ah_0310		ah_1010
연령	ah_021002	ah_031002		ah_101002
출생월	ah_0211	ah_0311		ah_1011
최종학교	ah_0215	ah_0315		ah_1015
졸업여부	ah_0217	ah_0317		ah_1017
학력(6개 구분)	ah_021701	ah_031701		ah_101701
등록장애인 여부	ah_0219	ah_0319		ah_1019
장애유형	ah_0221	ah_0321		ah_1021
장애등급	ah_0231	ah_0331		ah_1031
장애부위	ah_0225	ah_0325		ah_1025
혼인상태	ah_0227	ah_0327		ah_1027
취업상태	ah_0229	ah_0329		ah_1029
가구주 번호	ah_4101			
패널의 가구주 여부	h_4102			
패널을 포함한 가구원 수	h_4103			

\* 편입 사유(h\_1202)부터 취업상태(h\_1215) 변수까지가 가구원 1명에 대한 정보를 포함한 한 세트로 구성되어 있다. 상기 변수명은 첫 번째 신규 가구원에 해당하는 변수이며, 지난조사 이후 신규 편입 가구원 수(h\_1101)만큼의 데이터가 수록되어 있다.

\* 장애등급제 폐지로 인해 5차조사에서부터 장애등급을 기존 1~6급에서 장애의 정도가 심한 장애, 장애의 정도가 심하지 않은 장애로 이원화하여 변수도 수정되었다.(기존 ah\_0223 ~ ah\_1023)

## 2) 가구소득

- 가구 근로소득은 개인 근로소득에 가구원의 근로소득을 더하여 지난 1년간의 해당가구 세후 근로소득 총액(임금소득 및 사업소득)을 조사하였다. 개인소득과 마찬가지로 모름/응답거절인 경우에는 범주형 문항으로 재질문하여 응답받았다(개인소득 참조).
- 기준시점(조사이전년도) 이후에 분가한 가구원이 있는 경우에는 분가하기 전까지의 근로소득을 합산하였다.

〈표 5-30〉 가구 근로소득 관련 변수

항목	변수명	비고
연간 가구주 임금소득*	h_2303	
	h_2304	범주형 응답
	ah_2303	범주형 응답 포함 가공변수
연간 가구주 사업소득**	h_2307	
	h_2308	범주형 응답
	ah_2307	범주형 응답 포함 가공변수
연간 가구주 근로소득	ah_2311	범주형 응답 포함 가공변수
가구 전체 임금소득	h_2401	
	h_2402	범주형 응답
	ah_2401	범주형 응답 포함 가공변수
가구 전체 사업소득	h_2403	
	h_2404	범주형 응답
	ah_2403	범주형 응답 포함 가공변수
가구 전체 근로소득	ah_2405	범주형 응답 포함 가공변수

\* 임금소득: 고용주와의 계약을 통해 노동의 대가로 받게 되는 보수. 조사대상자가 회사나 공장 등에 직원으로 입사를 해서 근로의 대가로 받는 보수를 통해 얻는 소득

\*\* 사업소득: 자영업(노점 포함)이나 사업주의 경우, 자신의 사업체 운영을 통해 얻는 소득

- 가구 근로 외 소득은 패널 가구의 연간 '임금/사업소득 외에서 발생한 소득(근로소득 이외 소득)'의 세후 총소득을 응답받았으며, 기준시점 이후 분가한 가구원이 있는 경우에는 분가하기 전까지의 근로소득 이외 소득을 합산하였다.

〈표 5-31〉 가구 근로 외 소득 관련 변수

항목	변수명		
	연속형(기본)	범주형	범주형 포함
금융소득(개인연금소득 포함)	h_2501	h_2502	ah_2501
부동산소득	h_2503	h_2504	ah_2503
연금 (연금일시금, 기초연금, 국민연금 장애연금 제외)	h_2505	h_2506	ah_2505
산재보험(장해일시금, 유족일시금 제외)	h_2507	h_2508	ah_2507
보훈연금, 고용보험 및 기타	h_2509	h_2510	ah_2509
국민기초생활보장 급여	h_2511	h_2512	ah_2511
장애인연금/경증장애수당	h_2513	h_2514	ah_2513
장애연금(국민연금)	h_2515	h_2516	ah_2515
장애아동수당	h_2517	h_2518	ah_2517
그 외 장애인특별지원금	h_2519	h_2520	ah_2519
코로나19 가구 지원금	h_2533	h_2534	ah_2533
기타 공적이전소득	h_2521	h_2522	ah_2521
가구원 외 따로 살고 있는 가족, 친척으로부터 받은 사적이전소득	h_2523	h_2524	ah_2523
그 이외 사적이전소득 (친구, 이웃, 종교단체, 민간단체 등)	h_2525	h_2526	ah_2525
기타 소득	h_2527	h_2528	ah_2527
가구 근로 외 소득 총계	ah_2531(범주형 응답 포함 가공변수)		
가구 소득 총계(근로소득+근로외소득)	ah_2532(범주형 응답 포함 가공변수)		

### 3) 생활비

- 가구 생활비는 지난 1년간 항목별로 지출된 한 달 평균 금액을 측정하였다. 소득 문항과 마찬가지로 항목별 모름/응답거절 시 범주형 문항으로 응답하도록 하였으며, 범주형 문항에 응답한 경우 중간값으로 소비의 총계를 계산하였다. 기준시점 이후에 분가한 가구원이 있는 경우에는 분가하기 이전까지의 금액을 계산하였다.
- 최대한 개별 가구원이 사용한 소비 금액을 파악하되, 불가피한 경우에는 가구 내 이전지출에 포함시키고 해당 금액을 다른 항목과 중복될 수 없도록 하였다. 혼수장만, 교통사고 보상금 지불 등 특별한 일로 인해 일회성으로 지불한 금액은 제외하였다.



〈표 5-32〉 생활비 관련 변수

항목	변수명		
	연속형(기본)	범주형	범주형 포함
식비	h_2601	h_2602	ah_2601
외식비	h_2603	h_2604	ah_2603
교육비 - ①공교육비	h_2605	h_2606	ah_2605
교육비 - ②사교육비	h_2607	h_2608	ah_2607
차량유지비, 대중교통비	h_2609	h_2610	ah_2609
주거비	h_2611	h_2612	ah_2611
경조사비	h_2613	h_2614	ah_2613
보건의료비	h_2615	h_2616	ah_2615
교양오락비	h_2617	h_2618	ah_2617
내구재 구입비	h_2619	h_2620	ah_2619
통신비	h_2621	h_2622	ah_2621
가구 내 이전지출	h_2623	h_2624	ah_2623
가구 간 이전지출	h_2625	h_2626	ah_2625
피복비	h_2627	h_2628	ah_2627
현금 및 각종 기부금	h_2629	h_2630	ah_2629
국민연금 및 건강보험료	h_2631	h_2632	ah_2631
담배, 술 구입비	h_2633	h_2634	ah_2633
생필품 구입비	h_2635	h_2636	ah_2635
저축	h_2637	h_2638	ah_2637
각종 이자비용(원리금 합계)	h_2639	h_2640	ah_2639
그 외 기타 세금	h_2641	h_2642	ah_2641
코로나19로 인한 추가 지출	h_2647	h_2648	ah_2647
기타 생활비 지출금액	h_2643	h_2644	ah_2643
한 달 평균 생활비 총계	ah_2645(범주형 포함 가공변수)		

### 3. 직업력 자료

- 장애인고용패널조사의 직업력(work history) 자료는 1차 조사 이후로 조사된 개인의 모든 일자리에 관련된 정보가 담겨 있는 자료이다. 직업력 자료를 통해 조사시점 이후 패널이 경험한 모든 일자리의 역사를 종합적으로 볼 수 있다.
- 직업력 자료는 개인의 직업을 자료 구성의 기본 단위로 한다. 데이터는 WORK1, WORK2의 두 가지 형식으로 제공되며, 자료 형식에 따라 변수명 구조가 상이하므로 자료 이용 시 코드북을 활용하는 것이 바람직하다.

#### 1) 직업력 자료(WORK1)의 구조

- 직업력 WORK1 자료는 패널 개인의 직업을 자료 구성의 기본 단위로 한다. 8가지 유형별 조사표를 토대로 구성되어 있고 해당 자료에는 주된 일자리 정보를 포함하여 1개월 이상 지속하였거나 1개월 이상 지속하지 않았지만 새로 시작하여 현재 일하는 모든 일자리 정보가 포함되어 있다.
- 개인이 종사한 일자리 정보를 모두 포함하고 있기 때문에 횡구조를 가지며, 여러 개의 일자리에 종사한 패널의 경우 일자리 수만큼의 행을 가지고 있다.
- 직업력 WORK1은 분석을 위한 KEY 변수와 패널 기본정보 변수, 조사표 내용을 반영한 주제별 변수로 구분된다.
- 직업력 관련 KEY 변수는 해당 일자리의 기본정보를 반영하며, 주요 내용은 다음과 같다.
  - **일자리 ID(jobid):** 일자리 ID는 조사시기와 해당 조사시기에 파악된 일자리 개수, 일자리를 구분할 수 있는 변수로 일정한 패턴을 가지고 구성되어 있다.
  - **일자리 종류(jobtype):** 일자리 종류는 일자리의 종사상 지위 및 지속여부를 1~8가지의 유형으로 나타낸 정보이다.

〈표 5-33〉 일자리 종류(jobtype) 구성 체계

구분	종사상 지위	지속여부	jobtype
지난 조사 당시 하고 있던 일	임금근로자	지속	01
		그만둠	02
	비임금근로자	지속	03
		그만둠	04
새로 조사된 일	임금근로자	지속	05
		그만둠	06
	비임금근로자	지속	07
		그만둠	08

- **일자리 최초 조사 차수(jobfiryear):** 일자리가 최초로 조사된 차수를 의미한다.
- **일자리 고유 번호(jobseq):** 일자리 고유번호는 기존 일자리 ID 변수와 함께 패널별 일자리 매칭에 사용되는 변수로, 일자리 고유번호는 차수 변화에 관계 없이 하나의 일자리에 고유한 일자리 번호를 부여하는 것이다.
- **지난조사 시 일자리 개수(기존 일자리)(jobold):** 지난조사 시 일자리 개수에 해당하며, 1차조사

당시 응답한 주업 및 부업 일자리의 수를 의미한다.

- **지난조사 이후 총 일자리 개수(jobtotal):** 지난조사 이후 총 일자리 개수로, 현재 지속여부와 무관하게 1차조사 이후 패널이 경험한 모든 일자리 수의 합계에 해당한다.
- **현재 지속 일자리 개수(jobctotal):** 현재 지속하고 있는 일자리의 수를 의미한다.
- **조사 차수별 누적 일자리 개수(jobatotal):** 조사되었던(그만두거나 지속하는 경우도 포함) 누적 일자리 개수를 의미한다.
- **직전 조사 일자리ID(jobpreid):** 직전 차수의 일자리 ID를 의미한다.
- **주업여부(jobmain):** 해당 일자리가 패널의 주업에 해당하는지를 구분하는 변수이다.

〈표 5-34〉 직업력 KEY 변수(WORK1)

항목	변수명	비고
일자리 ID	jobid	
일자리 종류	jobtype	
일자리 최초 조사 차수	jobfiryear	
일자리 고유 번호	jobseq	
지난조사 시 일자리 개수(기존 일자리)	jobold_	
지난조사 이후 총 일자리 개수	jobtotal_	
현재 지속 일자리 개수	jobctotal_	
조사 차수별 누적 일자리 개수	jobatotal_	각 차수까지 누적된 일자리 총개수는 joballtotal_로 되어 있음
직전 조사 일자리ID	jobpreid	
주업여부	jobmain	

- 주제별 변수는 일자리(사업체) 기본 정보, 일자리 특성, 근로조건 등의 내용으로 구성되어 있으며, 변수 구성은 직업력 유형에 따라 8가지로 구분된다.

〈표 5-35〉 직업력 주제별 변수(임금근로자, WORK1)

항목	변수명*			
	유형1	유형2	유형5	유형6
사업내용 및 산업코드	job01_0101 ~job01_0102	job02_0101 ~job21_0102	job05_0101 ~job05_0102	job06_0101 ~job06_0102
직업내용 및 직업코드	job01_0103 ~job01_0104	job02_0103 ~job02_0104	job05_0103 ~job05_0104	job06_0103 ~job06_0104
주로 사용하는 장비 또는 기계	job01_0105	job02_0105	job05_0105	job06_0105
종사상 지위	job01_0106	job02_0106	job05_0106	job06_0106
근무시작시기			job05_0107 ~job05_0109	job06_0107 ~job06_0109
근무종료시기		job02_0110 ~job02_0112		job06_0110 ~job06_0112
실제 근무일수 및 근무시간	job01_0125 job01_0126	job02_0125 job02_0126	job05_0125 job05_0126	job06_0125 job06_0126
월평균 소득	job01_0127 job01_0128	job02_0127 job02_0128	job05_0127 job05_0128	job06_0127 job06_0128
장애가 일하는 데 미치는 영향 정도	job01_0131	job02_0131	job05_0131	job06_0131
업무 만족도	job01_0132	job02_0132	job05_0132	job06_0132
일자리 만족도	job01_0133	job02_0133	job05_0133	job06_0133
일자리 지속 근무 의향	job01_0134		job05_0134	
일자리 종류	job01_0120	job02_0120	job05_0120	job06_0120
사업체 규모	job01_0121 job01_0122	job02_0121 job02_0122	job05_0121 job05_0122	job06_0121 job06_0122
정규직 여부	job01_0124	job02_0124	job05_0124	job06_0124
노동조합 유무 및 가입여부	job01_0130	job02_0130	job05_0130	job06_0130
사회보험 가입여부	job01_012901 ~job01_012905	job02_012901 ~job02_012905	job05_012901 ~job05_012905	job06_012901 ~job06_012905

\* 임금근로자에 해당하는 직업력 유형은 다음과 같음.

유형1: 임금근로자-기존 일자리-지속

유형2: 임금근로자-기존 일자리-그만됨

유형5: 임금근로자-신규 일자리-지속

유형6: 임금근로자-신규 일자리-그만됨

〈표 5-36〉 직업력 주제별 변수(비임금근로자, WORK1)

항목	변수명*			
	유형3	유형4	유형7	유형8
사업내용 및 산업코드	job03_0101 ~job03_0102	job04_0101 ~job04_0102	job07_0101 ~job07_0102	job08_0101 ~job08_0102
직업내용 및 직업코드	job03_0103 ~job03_0104	job04_0103 ~job04_0104	job07_0103 ~job07_0104	job08_0103 ~job08_0104
주로 사용하는 장비 또는 기계	job03_0105	job04_0105	job07_0105	job08_0105
종사상 지위	job03_0106	job04_0106	job07_0106	job08_0106
근무시작시기			job07_0107 ~job07_0109	job08_0107 ~job08_0109
근무종료시기		job04_0110 ~job04_0112		job08_0110 ~job08_0112
실제 근무일수 및 근무시간	job03_0128 job03_0129	job04_0128 job04_0129	job07_0128 job07_0129	job08_0128 job08_0129
월평균 소득	job03_0133 job03_0134	job04_0133 job04_0134	job07_0133 job07_0134	job08_0133 job08_0134
장애가 일하는 데 미치는 영향 정도	job03_0135	job04_0135	job07_0135	job08_0135
업무 만족도	job03_0136	job04_0136	job07_0136	job08_0136
일자리 만족도	job03_0137	job04_0137	job07_0137	job08_0137
일자리 지속 근무 의향	job03_0138		job07_0138	
사업체 종류	job03_0122	job04_0122	job07_0122	job08_0122
종업원 수	job03_0125 job03_0126	job04_0125 job04_0126	job07_0125 job07_0126	job08_0125 job08_0126
연평균 매출액	job03_0130 job03_0131	job04_0130 job04_0131	job07_0130 job07_0131	job08_0130 job08_0131

\* 비임금근로자에 해당하는 직업력 유형은 다음과 같음.

유형3: 비임금근로자-기존 일자리-지속

유형4: 비임금근로자-기존 일자리-그만둠

유형7: 비임금근로자-신규 일자리-지속

유형8: 비임금근로자-신규 일자리-그만둠

## 2) 직업력 자료(WORK2)의 구조

- 직업력 WORK2는 연구자의 분석 편의성 제고를 위해 패널별 1개의 행(row)에 직업력 데이터를 축약한 구조로 구성되어 있다.
- 패널별 일자리 정보를 첫 번째 일자리 ~ 마지막 일자리로 구분하고 일자리별 핵심정보(일자리 유형, 지속여부, 일자리 시작/종료 시기, 임금 or 매출액, 일자리 전반적 만족도 등) 만을 포함하고 있다.
- 직업력 WORK2는 패널 기본정보 변수, 일자리별 핵심 정보 변수로 구분된다.

〈표 5-37〉 직업력 일자리별 핵심 정보 변수(WORK2)

항목	변수명	비고
일자리 ID	job__id	1. 일자리 개수별로 세트 구성(첫 번째 일자리의 경우 001, 두 번째는 002로 변수 생성) 2. 8차조사에서 최대 마지막 일자리는 7개로 일곱 번째 일자리 변수까지 구성되어 있음
일자리 종류	job__type	
일자리 최초 조사 차수	job__firyear	
일자리 고유 번호	job__seq	
일자리 이전 차수 ID	job__preid	
일자리 주업여부	job__main	
일자리 지속여부	job__continue	
일자리 사업내용	job__industrial	
일자리 사업내용_산업코드(대)	job__industrial_code1	
일자리 사업내용_산업코드(중)	job__industrial_code2	
일자리 본인이 주로 하는 일	job__job	
일자리 본인이 주로 하는 일_직업코드(대)	job__job_code1	
일자리 본인이 주로 하는 일_직업코드(중)	job__job_code2	
일자리 주로 사용하는 장비 또는 기계	job__equipment	
일자리 종사상 지위	job__status	
일자리 시작한 연도	job__start_year	
일자리 시작한 월	job__start_month	
일자리 시작한 월(계절)	job__start_season	
일자리 그만둔 연도	job__finish_year	
일자리 그만둔 월	job__finish_month	
일자리 그만둔 월(계절)	job__finish_season	
일자리 근속기간(조사월 기준, 개월 수)	job__period	
일자리 근속기간(조사월 기준, 4개 구분)	job__period_t	
일자리 실 근무일수(일주일 평균)	job__working_7days	
일자리 실 근무시간(일 평균)	job__working_1day	
일자리 월 평균 임금_임금근로	job__type1_income	
일자리 월 평균 임금(범주)_임금근로	job__type1_income_t	
일자리 연 평균 매출액_비임금근로	job__type2_income	
일자리 연 평균 매출액(범주)_비임금근로	job__type2_income_t	
일자리 장애가 일하는 데 미치는 영향 정도	job__impact	
일자리 업무 만족도	job__satisfaction_1	
일자리 만족도	job__satisfaction_2	

## 부 록

### 데이터의 활용

#### 데이터 사용 전 기본 숙지사항

- 장애인고용패널조사는 2016년 1월 1일 기준 만 15~64세의 패널을 구축하였으므로 8차년도(2023년) 현재 패널의 연령은 만 22~71세이다.
- 기 공개된 데이터는 추가적인 자료정리 및 확인 과정을 거쳐 일부 수치가 변경될 수 있으므로, 8차년도 데이터 공표 시 제공되는 차수별 데이터를 이용하여야 한다.
- 순위형 문항은 총 4개의 변수(1순위, 1순위-기타, 2순위, 2순위-기타)로 구성되는데 1순위 보기 중 “없음” 선택이 가능한 문항과 불가능한 문항으로 구분되어 있다. 중복응답형 문항은 모두 선택이 가능하도록 구성되어 있으며 보기 개수대로 변수가 생성되고 응답한 보기의 값이 나열되어 있다.
- 연속변수 형태로 묻는 문항의 경우 모름/응답거절 시 범주형 문항으로 재질문하여 응답을 받았기 때문에, 해당 변수는 연속형을 범주형으로, 또는 범주형을 연속형으로 변환하여 분석에 활용하여야 한다. 월평균 소득 등 주요 변수에 대해서는 연속형 혹은 범주형으로 변환한 가공변수가 제공된다.
- 각 차수별로 변수값이 다른 경우가 있을 수 있으므로, 같은 변수라도 코드북 등의 내용을 숙지한 후 분석에 활용해야 한다.
- 조사표 분기 또는 문항 분기로 인해 실제 응답자 수가 맞지 않는 경우가 있을 수 있고 특정 연령과 장애유형 등을 고려하여 대상자를 추출할 경우 표본 수가 적거나 비어 있는 셀이 발생할 수 있으므로 반드시 조사표와 코드북을 확인하여야 한다. 데이터 이용 시 기초분석을 통해 관련 변수들의 빈도값이 코드북의 내용과 일치하는지를 사전에 확인하고 missing 값의 처리를 주의하여야 한다.
- 특히, 7차조사부터는 통합(WIDE) 데이터와 직업력 데이터 간 활용성 강화를 위해 직업력(WORK2) 데이터의 일부 항목을 통합(WIDE)에 추가하였으니, 패널별 구체적인 일자리 분석을 위해선 WORK2 데이터를 활용해야 한다.

## 1. 데이터를 불러오기 위한 경로 지정 방법

- FAQ의 예제 프로그램은 SAS, SPSS, STATA 형태로 제공하며 각 프로그램을 실행하기 위해 데이터의 경로를 지정하는 것은 다음과 같다.

### SAS

SAS의 경우 임시라이브러리(work)가 아닌 PSED 데이터만 저장할 수 있는 라이브러리를 지정하여 사용하면 편리하다. 라이브러리 이름은 psed로 지정하며, 그 방법은 다음과 같다.

```
/*library name을 psed로 지정하는 방법*/
libname psed 'd:\psed\sas\';
위의 명령문에서 d:\psed\sas\ 부분을 이용자가 사용하기 편리한 경로로 변경해주면 된다.
```

### SPSS

```
/*D드라이브에서 8차년도 메인데이터 불러오는 방법*/
get file='d:\psed\spss\main08.sav'.
위의 명령문에서 d:\psed\spss\ 부분을 이용자가 사용하기 편리한 경로로 변경해주면 된다.
```

### STATA

```
/*D드라이브에서 8차년도 메인데이터 불러오는 방법*/
use "d:\psed\stata\main08.dta", clear
위의 명령문에서 d:\psed\stata\ 부분을 이용자가 사용하기 편리한 경로로 변경해주면 된다.
```



## 2. 차수별 데이터를 횡으로 합치는 방법(merge)

- 장애인고용패널조사는 1~8차년도 각각의 메인 데이터와 모든 차수를 포괄하는 통합 데이터를 제공하고 있다. 따라서 데이터 분석 시에는 각 차수의 메인 데이터에서 필요한 변수를 추출하여 분석용 데이터를 통합(MERGE)하거나, 통합 데이터에서 필요한 변수를 선별하여 추출하는 방법을 사용할 수 있다. 두 가지 방법에 대한 절차는 다음과 같다.

### (1) 1~8차년도 메인 데이터를 횡으로 Merge하는 방법

- ① 1~8차년도 메인 데이터에서 각각 필요한 변수를 추출한다.
- ② 필요한 변수를 추출한 1~8차년도 데이터를 기준변수 pid(패널ID)로 정렬한다.
- ③ pid를 기준으로 merge 한다.

### SAS

```
data a*;
set psed.main*;
keep pid emp**02 emp**03 emp**04 emp**05 emp**06 emp**07 emp**08;
proc sort; by pid;
run;
```

위의 명령문에서 \*에 Merge하고자 하는 차수를 대입하여, 저장하고자 하는 변수의 해당 차수 데이터를 생성한다.

```
data b;
merge a1 a2 a3 a4 a5 a6 a7 a8;
by pid;
run;
```

생성된 데이터를 Merge 한다. 위 명령어는 1차부터 8차까지의 데이터 a1~a8을 Merge 한 데이터 b를 생성한다.

## SPSS

```

get file='d:\Wpsed\Wspss\Wmain*.sav'
keep=pid emp**02 emp**03 emp**04 emp**05 emp**06 emp**07 emp**08.
sort cases by pid.
save out='d:\Wpsed\Wspss\Wa*.sav'
compressed.

```

위의 명령문에서 \*에 Merge하고자 하는 차수를 대입하여, 저장하고자 하는 변수의 해당 차수 데이터를 생성한다.

```

match files
/file='d:\Wpsed\Wspss\Wa1.sav'
/file='d:\Wpsed\Wspss\Wa2.sav'
/file='d:\Wpsed\Wspss\Wa3.sav'
/file='d:\Wpsed\Wspss\Wa4.sav'
/file='d:\Wpsed\Wspss\Wa5.sav'
/file='d:\Wpsed\Wspss\Wa6.sav'
/file='d:\Wpsed\Wspss\Wa7.sav'
/file='d:\Wpsed\Wspss\Wa8.sav'
execute.
save out='d:\Wpsed\Wspss\Wb.sav'
compressed.

```

생성된 데이터를 Merge 한다. 위 명령어는 1차부터 8차까지의 데이터 a1~a8을 Merge 한 데이터 b를 생성한다.

## STATA

```

use "d:\Wpsed\Wstata\Wmain*.dta", clear
keep pid emp**02 emp**03 emp**04 emp**05 emp**06 emp**07 emp**08
sort pid
save "d:\Wpsed\Wstata\Wa*.dta", replace

```

위의 명령문에서 \*에 Merge하고자 하는 차수를 대입하여, 저장하고자 하는 변수의 해당 차수 데이터를 생성한다.

```

use "d:\Wpsed\Wstata\Wa1.dta", clear
merge pid using "d:\Wpsed\Wstata\Wa2.dta",
merge pid using "d:\Wpsed\Wstata\Wa3.dta",
merge pid using "d:\Wpsed\Wstata\Wa4.dta",
merge pid using "d:\Wpsed\Wstata\Wa5.dta",
merge pid using "d:\Wpsed\Wstata\Wa6.dta",
merge pid using "d:\Wpsed\Wstata\Wa7.dta",
merge pid using "d:\Wpsed\Wstata\Wa8.dta",
unique sort save "d:\Wpsed\Wstata\Wb.dta", replace

```

생성된 데이터를 Merge 한다. 위 명령어는 1차부터 8차까지의 데이터 a1~a8을 Merge 한 데이터 b를 생성한다.

## (2) 통합데이터(WIDE TYPE)에서 필요한 변수만 추출하기

WIDE TYPE 통합데이터는 이미 1~8차년도의 메인데이터가 횡으로 merge된 파일이므로, 위의 과정을 거치지 않고 바로 필요한 변수만 저장해서 사용할 수 있어 1번 방법에 비해 편리하게 이용할 수 있다.

### SAS

```
data a1;
set psed.wide;
keep pid emp0102 emp0103 emp0104 emp0105 emp0106 emp0202 emp0203
emp0204 emp0205 emp0206 emp0302 emp0303 emp0304 emp0305 emp0306
emp0402 emp0403 emp0404 emp0405 emp0406 emp0502 emp0503 emp0504
emp0505 emp0506 emp0602 emp0603 emp0604 emp0605 emp0606 emp0702
emp0703 emp0704 emp0705 emp0706 emp0801 emp0802 emp0803 emp0804
emp0805 emp0806.
run;
```

### SPSS

```
get file='d:\psed\spss\wide.sav'
/keep=pid emp0102 emp0103 emp0104 emp0105 emp0106 emp0202 emp0203
emp0204 emp0205 emp0206 emp0302 emp0303 emp0304 emp0305 emp0306
emp0402 emp0403 emp0404 emp0405 emp0406 emp0502 emp0503 emp0504
emp0505 emp0506 emp0602 emp0603 emp0604 emp0605 emp0606 emp0702
emp0703 emp0704 emp0705 emp0706 emp0801 emp0802 emp0803 emp0804
emp0805 emp0806.
sort cases by pid.
save out='d:\psed\spss\wa1.sav'
compressed.
```

### STATA

```
use "d:\psed\stata\wide.dta", clear
keep pid emp0102 emp0103 emp0104 emp0105 emp0106 emp0202 emp0203
emp0204 emp0205 emp0206 emp0302 emp0303 emp0304 emp0305 emp0306
emp0402 emp0403 emp0404 emp0405 emp0406 emp0502 emp0503 emp0504
emp0505 emp0506 emp0602 emp0603 emp0604 emp0605 emp0606 emp0702
emp0703 emp0704 emp0705 emp0706 emp0801 emp0802 emp0803 emp0804
emp0805 emp0806.
sort pid
save "d:\psed\stata\wa1.dta", replace
```

### 3. 차수별 데이터를 종으로 합치는 방법(set)

- 차수별 시계열 분석을 위해서 데이터를 종으로 합쳐야 하는 경우가 자주 발생한다. 이러한 경우 각 차수별 공통적으로 포함된 변수를 추출한 뒤 변수명을 동일하게 변경 후 종으로 합치면 된다.
- 통합데이터(LONG TYPE)에는 모든 조사년도의 변수를 종으로 통합하여 제공하기 때문에, 필요한 변수만 따로 저장하여 사용하면 더욱 편리하게 이용할 수 있다. 각각의 경우에 대한 절차는 다음과 같다.

#### (1) 1~8차년도 메인 데이터를 종으로 합치는 방법

- ① 각 차수별 공통변수들의 변수명을 동일하게 변경한다.
- ② 각 차수별로 구분할 수 있는 더미 변수를 만든다.
- ③ 1~8차년도의 자료를 종으로 합친다.

#### SAS

```

data a*;
set psed.main*;
keep pid p0* emp0*2 emp0*3 emp0*5 emp0*6 emp0*7 emp0*8;
if p=* then dummy=*;
rename p0*=p emp0*2=emp2 emp0*3=emp3 emp0*5=emp5 emp0*6=emp6
emp0*7=emp7 emp0*8=emp8;
run;

```

위의 명령문에서 \*에 종으로 합치고자 하는 차수(1~8)를 대입하여, 저장하고자 하는 변수의 해당 차수 데이터를 생성한다.

```

data b;
retain pid p emp2 emp3 emp5 emp6 emp7 emp8 dummy;
set a1 a2 a3 a4 a5 a6 a7 a8;
run;

```

종으로 합친 데이터를 생성한다. 위 명령어는 1차부터 8차까지의 데이터 a1~a8을 종으로 합친 데이터 b를 생성한다.

## SPSS

```
get file='d:\wpsed\wspss\wmain*.sav'  
/keep=pid p** emp**02 emp**03 emp**04 emp**05 emp**06 emp**07  
/rename=(p**=p) (emp**02=emp02) (emp**03=emp03) (emp**04=emp04)  
(emp**05=emp05) (emp**06=emp06) (emp**07=emp07) (emp**08=emp08).  
compute dummy=*
```

```
save out='d:\wpsed\wspss\wa*.sav'.
```

위의 명령문에서 \*에 종으로 합치고자 하는 차수를 대입하여, 저장하고자 하는 변수의 해당 차수 데이터를 생성한다.

```
add files/file='d:\wpsed\wspss\wa1.sav'
```

```
/file='d:\wpsed\wspss\wa2.sav'
```

```
/file='d:\wpsed\wspss\wa3.sav'
```

```
/file='d:\wpsed\wspss\wa4.sav'
```

```
/file='d:\wpsed\wspss\wa5.sav'
```

```
/file='d:\wpsed\wspss\wa6.sav'
```

```
/file='d:\wpsed\wspss\wa7.sav'
```

```
/file='d:\wpsed\wspss\wa8.sav'.
```

```
execute.
```

```
save out=/file='d:\wpsed\wspss\wb.sav'
```

```
compressed.
```

종으로 합친 데이터를 생성한다. 위 명령어는 1차부터 8차까지의 데이터 a1~a8을 종으로 합친 데이터 b를 생성한다.

## STATA

```

use "d:\Wpsed\Wstata\Wmain*.dta", clear
gen dummy=*
keep pid p0* emp**02 emp**03 emp**04 emp**05 emp**06 emp**07 emp**08
dummy
rename p0*=p
rename emp**02=emp02
rename emp**03=emp03
rename emp**04=emp04
rename emp**05=emp05
rename emp**06=emp06
rename emp**07=emp07
rename emp**07=emp08

```

save "d:\Wpsed\Wstata\Wa\*.dta", replace  
위의 명령문에서 \*에 종으로 합치고자 하는 차수를 대입하여, 저장하고자 하는 변수의 해당 차수 데이터를 생성한다.

```

use "d:\Wpsed\Wstata\Wa1.dta", clear
appending use "d:\Wpsed\Wstata\Wa2.dta"
appending use "d:\Wpsed\Wstata\Wa3.dta"
appending use "d:\Wpsed\Wstata\Wa4.dta"
appending use "d:\Wpsed\Wstata\Wa5.dta"
appending use "d:\Wpsed\Wstata\Wa6.dta"
appending use "d:\Wpsed\Wstata\Wa7.dta"
appending use "d:\Wpsed\Wstata\Wa8.dta"
order pid p emp02 emp03 emp04 emp05 emp06 emp07 emp08 dummy
save "d:\Wpsed\Wstata\Wb.dta", replace

```

종으로 합친 데이터를 생성한다. 위 명령어는 a1, a2, a3, a4, a6, a7 a8을 종으로 합친 데이터 b를 생성한다.

## (2) 통합데이터(LONG TYPE)에서 필요한 변수만 추출하기

통합데이터(LONG TYPE)는 메인데이터가 종으로 합쳐진 파일이므로, 위의 과정을 거치지 않고 바로 필요한 변수만 저장해서 사용할 수 있다. 아래 과정에서 변수를 rename 하지 않고 바로 필요한 변수만 추출하여 사용하면 된다. 이미 LONG TYPE 데이터에서는 공통 변수명으로 rename이 된 상태이며 wave 변수를 사용하여 차수를 나타내었다. 모든 차수에서 조사된 변수가 포함되어 있으므로, 조사된 해당차수를 확인하는 것은 데이터 코드북을 이용하면 된다.

### SAS

```
data a1;
set psed.long;
keep pid p emp02 emp03 emp04 emp05 emp06 emp07 emp08;
run;
```

### SPSS

```
get file='d:\psed\spss\long.sav'
/keep=pid p emp02 emp03 emp04 emp05 emp06 emp07 emp08.
sort cases by pid.
save out='d:\psed\spss\wb.sav'
compressed.
```

### STATA

```
use "d:\psed\stata\long.dta", clear
keep pid p emp02 emp03 emp04 emp05 emp06 emp07 emp08
sort pid
save "d:\psed\stata\wb.dta", replace
```

#### 4. STATA 사용 시 한글 라벨 깨지는 문제

- STATA 14ver 이상부터 unicode 등의 호환 문제로 한글 라벨이 깨지는 상황이 다수 발생

<p>Variable List (R)</p> <table border="1"> <thead> <tr> <th>Variable</th> <th>Label</th> </tr> </thead> <tbody> <tr><td>pid</td><td>패널 ID</td></tr> <tr><td>gender</td><td>성별</td></tr> <tr><td>birthy</td><td>출생연도</td></tr> <tr><td>birthm</td><td>출생월</td></tr> <tr><td>birthd</td><td>출생일</td></tr> <tr><td>wave</td><td>조사 피수</td></tr> <tr><td>pid104</td><td>(가공) 연속응답여부</td></tr> <tr><td>p</td><td>조사 응답여부</td></tr> <tr><td>rpreason</td><td>조사 성공여부(사유)</td></tr> <tr><td>year</td><td>조사 연도</td></tr> <tr><td>month</td><td>조사 월</td></tr> <tr><td>day</td><td>조사 일</td></tr> <tr><td>wt01</td><td>환단면 거주지</td></tr> <tr><td>wt02</td><td>중단면 거주지</td></tr> <tr><td>area</td><td>실 거주지역</td></tr> <tr><td>aarea</td><td>(가공) 실 거주지역</td></tr> <tr><td>emp02</td><td>(가공) 경제활동상태(2개 구분)</td></tr> <tr><td>emp03</td><td>(가공) 경제활동상태(4개 구분)</td></tr> <tr><td>emp04</td><td>(가공) 경제활동상태(4개 구분)</td></tr> <tr><td>emp05</td><td>(가공) 경제활동상태(5개 구분)</td></tr> <tr><td>emp06</td><td>(가공) 경제활동상태(6개 구분)</td></tr> </tbody> </table> <p>use C:\Users\VGLIM\Desktop\bae.stunicode\long.dta</p> <p>〈STATA 13ver 이하〉</p>	Variable	Label	pid	패널 ID	gender	성별	birthy	출생연도	birthm	출생월	birthd	출생일	wave	조사 피수	pid104	(가공) 연속응답여부	p	조사 응답여부	rpreason	조사 성공여부(사유)	year	조사 연도	month	조사 월	day	조사 일	wt01	환단면 거주지	wt02	중단면 거주지	area	실 거주지역	aarea	(가공) 실 거주지역	emp02	(가공) 경제활동상태(2개 구분)	emp03	(가공) 경제활동상태(4개 구분)	emp04	(가공) 경제활동상태(4개 구분)	emp05	(가공) 경제활동상태(5개 구분)	emp06	(가공) 경제활동상태(6개 구분)	<p>Variable List (R)</p> <table border="1"> <thead> <tr> <th>Variable</th> <th>Label</th> </tr> </thead> <tbody> <tr><td>pid</td><td>패널 ID</td></tr> <tr><td>gender</td><td>###</td></tr> <tr><td>birthy</td><td>###</td></tr> <tr><td>birthm</td><td>###</td></tr> <tr><td>birthd</td><td>###</td></tr> <tr><td>wave</td><td>###</td></tr> <tr><td>pid104</td><td>###</td></tr> <tr><td>p</td><td>###</td></tr> <tr><td>rpreason</td><td>###</td></tr> <tr><td>year</td><td>###</td></tr> <tr><td>month</td><td>###</td></tr> <tr><td>day</td><td>###</td></tr> </tbody> </table> <p>Notes:</p> <ol style="list-style-type: none"> <li>Unicode is supported; see <a href="#">help unicode advice</a>.</li> <li>New update available; type <code>update all</code>.</li> </ol> <p>〈STATA 14ver 이상〉</p>	Variable	Label	pid	패널 ID	gender	###	birthy	###	birthm	###	birthd	###	wave	###	pid104	###	p	###	rpreason	###	year	###	month	###	day	###
Variable	Label																																																																						
pid	패널 ID																																																																						
gender	성별																																																																						
birthy	출생연도																																																																						
birthm	출생월																																																																						
birthd	출생일																																																																						
wave	조사 피수																																																																						
pid104	(가공) 연속응답여부																																																																						
p	조사 응답여부																																																																						
rpreason	조사 성공여부(사유)																																																																						
year	조사 연도																																																																						
month	조사 월																																																																						
day	조사 일																																																																						
wt01	환단면 거주지																																																																						
wt02	중단면 거주지																																																																						
area	실 거주지역																																																																						
aarea	(가공) 실 거주지역																																																																						
emp02	(가공) 경제활동상태(2개 구분)																																																																						
emp03	(가공) 경제활동상태(4개 구분)																																																																						
emp04	(가공) 경제활동상태(4개 구분)																																																																						
emp05	(가공) 경제활동상태(5개 구분)																																																																						
emp06	(가공) 경제활동상태(6개 구분)																																																																						
Variable	Label																																																																						
pid	패널 ID																																																																						
gender	###																																																																						
birthy	###																																																																						
birthm	###																																																																						
birthd	###																																																																						
wave	###																																																																						
pid104	###																																																																						
p	###																																																																						
rpreason	###																																																																						
year	###																																																																						
month	###																																																																						
day	###																																																																						

- 한글 변수 깨짐 현상 발생 시 해당 명령문 입력

#### STATA

```
Clear
cd "d:\wpsed\wstata"
unicode encoding set euc-kr
unicode translate psed.dta, transutf8
use psed
```

```
. use "C:\Users\VGLIM\Desktop\long.dta"

. clear

. cd "C:\Users\VGLIM\Desktop"
C:\Users\VGLIM\Desktop

. unicode encoding set euc-kr
(default encoding now euc-kr)

. unicode translate long.dta, transutf8
(using euc-kr encoding)

File summary (before starting):
  1 file(s) specified
  1 file(s) to be examined ...

File long.dta (Stata dataset)
  all variable names translated
  all data labels translated
  all variable labels translated
  all value-label names translated
  all value-label contents translated
  all str# variables translated

File successfully translated

File summary:
  all files successfully translated

. use long
```

실행화면

- 위 명령문 실행 시 오류 발생하면 아래 명령문 추가 기입

#### STATA

```
unicode translate psed.dta, transutf8 invalid(ignore)
use psed
```