

2011년도 국가통계 품질개선지원 연구용역

『대중교통현황조사』 품질개선지원 최종결과보고서

- 「대중교통환승실태 및 이용자 만족도 조사」 표본설계 -

2011. 10.

제 출 문

제 출 문

통계청장 귀하

본 보고서를 『대중교통현황조사』 품질개선지원 연구
용역 과제의 최종 연구결과물로 제출합니다.

[개선지원 부문]: 「대중교통환승실태 및 이용자 만족도조사」 표본설계

2011년 10월 30일

청주대학교 바이오정보통계학과 교수 류제복 ①

연구진

책임연구원 류제복 (청주대학교 교수)

연구원 이승주 (청주대학교 교수)

요 약 문

최종결과보고서 요약문

연구과제명	「대중교통현황조사」 대중교통환승실태 및 이용자 만족도 조사 표본설계
주 제 어	모집단 분석, 표본설계, 항목무응답대체
연 구 기 간	2011.06.27 ~ 2011.10.30 (4개월)
연 구 기 관	
연구진구성	류제복, 이승주
<p>“대중교통현황조사”(국토해양부, 교통안전공단)는 정부의 각종 대중교통 지원정책의 합리적인 근거 제공 및 국내 대중교통수준의 해외 비교를 통해 최적의 대중교통체계 발전을 유도하기 위한 것이다.</p> <p>통계청에서는 통계작성기관을 대상으로 한 통계품질 개선지원 수요조사 결과를 토대로 본 통계를 2011년 품질개선을 위한 과제로 선정하게 되었다.</p> <p>이번 연구를 통해 “대중교통 환승실태 및 이용자만족도 조사”를 위한 표본설계와 설문지 등에 대해 미흡한 점을 보완해서 실제적인 통계품질이 개선될 수 있는 방안들을 제시하였다.</p> <ul style="list-style-type: none"> ○ 현행 조사(2010년)에 대한 분석결과 표본배정의 타당성이 부족하였다. 새로운 표본설계에 사용할 연구변수를 선정하였다. ○ 신규 모집단으로 2010년 인구주택총조사 자료에서 15세 이상 인구를 사용하였고 지역별 인구분포를 살펴보았다. ○ 새로운 표본설계를 위해 층화, 표본크기 및 표본배정 그리고 추정을 다루고 항목무응답대체 방법을 소개하였다. <p>본 연구를 통해 향후 연구 검토해야할 사항들도 제시하였다. 즉,</p> <ul style="list-style-type: none"> ○ 현행 조사를 가구단위 조사로 전환하는 문제 ○ 조사지점 및 응답자 선정방법 ○ 현행 설문지의 수정 및 보완 사항 	

목 차

1장. 서론	1
1절. 연구배경	1
2절. 연구목적 및 필요성	1
3절. 연구내용 및 방법	2
4절. 연구흐름도	4
2장. 통계개요	5
1절. 현황	5
2절. 문제점	6
3절. 개선방안 개요	6
3장. 세부 개선방안	7
1절. 연구 협의내용	7
2절. 현행 조사에 대한 분석	13
3절. 모집단 분석	19
4절. 새로운 표본설계	22
(1) 층화	23
(2) 표본크기 결정 및 표본배정	23
(3) 추정	28
(4) 무응답대체	29

4장. 결론 및 제안	34
참고문헌	37
<부록 1>	39
<부록 2>	41

- 표 목 차 -

<표 1> 그룹별/시도별 조사대상 시·군	14
<표 2> 2010년 그룹별/시도별 표본수	15
<표 3> 월 평균 대중교통비용: 평균/표준편차	17
<표 4> 월 평균 대중교통비용: 변동계수/상대표준오차	18
<표 5> 2010년 15세 이상 인구수에 따른 조사대상 시·군	20
<표 6> 2010년 15세 이상 시·도별 인구수 및 가중치	21
<표 7> 목표오차별 표본크기	25
<표 8> 새로운 표본설계에 의한 그룹별/시도별 표본수	27

- 그림 목 차 -

<그림 1> 연구흐름도 4

1장. 서론

1절. 연구배경

“대중교통현황조사”(국토해양부, 교통안전공단)는 2009년 통계청의 정기 품질진단 결과 통계작성 여건과 조직 리더 및 통계작성 담당자의 통계에 대한 인식이 전반적으로 우수한 것으로 평가되었다. 그리고 통계품질을 향상하기 위한 몇 가지 문제점이 지적되고 개선사항들이 제시되었다.

통계청에서는 통계작성기관을 대상으로 한 통계품질 개선지원 수요조사 결과를 토대로 본 통계를 2011년 품질개선을 위한 과제로 선정하게 되었다.

2절. 연구목적 및 필요성

본 연구의 목적 및 필요성을 요약해 보면 다음과 같다.

□ 연구목적

- 2009년도 정기 품질진단에서 개선이 요구된 사항들 중에서 현재까지 개선이 미흡한 부분에 대한 지원을 통해 통계품질 향상을 목적으로 한다.
- 통계품질 향상을 통해 중앙 정부부처 뿐만 아니라 지방자치단체에서 보다 정확한 대중교통관련 정책을 수립 및 집행하는데 도움을 주고자 한다.

□ 연구 필요성

- 개선 요구사항들 중에서 통계생산 담당자와 관계자들이 해결하기 어려운 부분에 대해 외부 전문가들의 도움이 필요하다.
- “대중교통현황조사”의 “환승실태 및 이용자 만족도 조사”에서 표본설계와 설문지 등의 미흡한 부분을 보완해야 실제적인 통계품질 개선이 이루어 질 수 있다.
- 향후 통계 담당자가 원활하게 통계생산을 지속적으로 할 수 있는 실무적 지원이 필요하다.

3절. 연구내용 및 방법

국토해양부와 교통안전공단에서 생산되고 있는 “대중교통현황조사”는 6개 부문에 대해 관측조사, 탑승조사, 그리고 설문조사로 실시되고 있다. 본 연구는 설문조사 부문인 “환승실태 및 이용자 만족도 조사”로 한정한다. 한편 연구 방법은 “대중교통현황조사”와 관련된 조사결과보고서, 품질진단결과보고서, 표본설계 보고서 등에 대한 문헌연구와 2010년도 조사 자료를 분석해서 적절한 표본설계방법을 연구한다.

연구 내용을 요약하면 다음과 같다.

□ 현행 “환승실태 및 이용자 만족도 조사”에 대한 현황과 문제점

- “2010년 대중교통현황조사” 조사결과보고서, “2010년 환승실태 및 이용자 만족도 조사” 보고서, 통계품질진단보고서, 그리고 관련자료 등을 토대로 현황분석 및 문제점 파악

- 통계담당자, 실사 담당자, 그리고 조사 자문교수 등과의 회의를 통해 추가 사항 파악

□ **환승실태 및 이용자 만족도조사를 위한 표본설계의 적합성**

- 추출틀의 적합성 검토
- 표본추출 방법과 조사방법의 타당성 검토
- 표본의 대표성, 추정, 그리고 가중치 문제 등 전반적인 표본설계 내용 검토

□ **설문조사표의 검토**

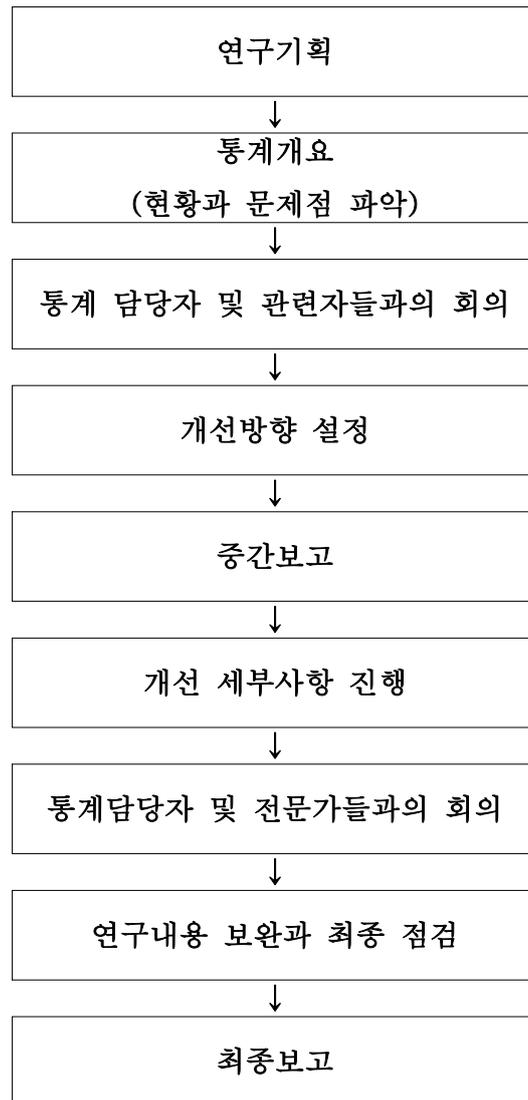
- 설문 문항의 적절성
- 조사방법에 따른 설문 문항의 적합성
- 응답자 응답부담

□ **개선방안 및 기타**

- 조사 자료에 대한 분석의 적절성
- 현행 시·군·구 통계에 대한 타당성 점검
- 향후 연구방향
- 무응답대체: 평균대체
⇒ 응답항목들이 대부분 범주형 자료이므로 평균대체는 적절치 않으므로 범주형 자료의 대체에 적절한 대체방법 등을 제안

4절. 연구흐름도

본 연구용역의 수행을 위한 연구 흐름도는 <그림 1>과 같다.



<그림 1> 연구흐름도

2장. 통계개요

“대중교통현황조사”는 일반·조사통계(승인번호 11669)로 대중교통의 육성 및 이용 촉진에 관한법률 제16조(대중교통현황조사)와 동법 시행규칙 제4조(대중교통현황의 조사 등), 그리고 제5조(대중교통 현황조사의 의뢰)에 의해서 매년 조사가 실시, 공표된다.

1절. 현황

□ 조사목적

- 대중교통의 육성·지원을 위한 정책의 효과적인 수립에 필요한 기초 자료를 조사·제공함으로써 도시별 대중교통 육성을 통한 대중교통수단의 이용 및 촉진을 도모하고 도시규모 및 특성, 장래 도시교통여건 변화 등을 고려한 최적의 대중교통체계 발전을 유도하기 위함
- 정부의 각종 대중교통 지원정책의 합리적인 근거 제공 및 국내 대중교통수준의 해외 비교를 통해 정책 시사점을 도출함

□ 조사대상

- 본 통계는 대중교통관련 사회·경제적 지표, 대중교통 운영자의 경영여건, 대중교통수단 및 시설의 현황, 대중교통의 이용실태, 대중교통 운행현황, 대중교통의 개선을 위해 필요한 사항 등 6개 부분에 대해서 조사가 실시된다.
- 부문별로 조사대상은 대중교통 이용자, 대중교통 회사, 대중교통 운영자, 대중교통 등이 된다.

□ 조사범위

- 조사지역은 전국을 5개 권역으로 나눔
- 행정 구역상 68개 시·군, 7대 특별시/광역시, 40개 시, 21개 군에 대해서 149개 존을 선정

□ 조사내용 및 방법

- 조사항목은 6개 조사부문, 8개 지표, 24개 조사항목, 56개 세부 조사 사항으로 구성
- 조사방법은 관측조사(차종별 교통량조사), 탑승조사(노선별 이용인원, 대중교통 운행속도), 그리고 설문조사(환승유형 및 만족도) 등을 사용

2절. 문제점

- 모집단 정의가 명확치 않고 표본추출틀이 불완전
- 표본크기 계산방법, 표본추출방법, 추정공식 등이 구체적으로 제시되지 않는 등 전반적인 표본설계가 미비
- 현행의 조사지점과 표본의 대표성에 의문

3절. 개선방안 개요

- 모집단의 재정립과 최적의 표본추출틀 설정
- 추출방법, 조사방법 등의 문제점 개선
- 표본규모와 시·군 통계의 타당성 검토

3장. 세부 개선방안

1절. 연구 협의내용

본 연구의 원활한 진행을 위해서 국토해양부, 교통안전공단, 그리고 실사기관의 담당자들과 협의한 내용을 정리하면 다음과 같다.

□ 1차 연구협의회

- 일시 : 2011년 8월 10일(수), 오후 2시 ~ 4시 30분
- 장소 : 청주대학교 류제복교수 연구실
- 참석 :
 - 교통안전공단 김명희 연구원, 리서치 랩 김영남 책임연구원
 - 청주대학교: 류제복(연구책임자), 이승주(연구원)

□ 협의내용

- 모집단
 - 전국 68개 시·군지역 도시철도/ 시내버스 및 농촌버스 이용자
(주 4회 이상 이용하는 15세 이상 남·녀)
 - 대중교통을 주로 이용하는 사람을 대상으로 조사하기 때문
 - 그러나 대중교통을 이용하는 사람들을 대상으로 조사하고,
그중에서 자주 이용하는 사람(주 4회)을 분류해서 분석 가능
- 표본추출틀
 - 조사지점 전체 목록과 지점별 이용자 수에 대한 통계
 - 표본의 대표성을 확보하기 위해서

1) 현행조사지점에 대한 전체 목록과 지점별 노선 수 자료를 확보
(이용자수에 대한 자료가 없기 때문)

- 조사 지점에서 표본 선정을 랜덤하게 실시(예를 들어 일정한 간격으로 계통추출해서 조사대상을 선정. 그러나 이러한 것이 현실적으로 어려움)
- 그리고 조사일이 화~목이므로 그 외의 시간대(주말 등) 이용자들이 누락되므로 이에 대한 대비책이 마련되어야 함

2) 가구 단위조사로 전환

- 인구주택총조사 자료를 이용해서 가구를 선정한 후 대상자들에게 전날의 환승실태와 만족도조사를 실시
- 이 경우 확률추출로의 전환이 가능하고 조사 시 어려움을 해결할 수 있음
- 그러나 해당 가구원들이 대중교통을 이용하지 않는 경우 정해진 규모의 표본조사를 위해서 대체 표본을 사용해야 하므로 대체표본 규모를 예측해야 한다.

○ 조사범위

- 조사시점(2010년 6월 8일~7월 15일, 화-목 오전 7시~오후 9시)
 - 5개 권역, 7대 (특별/직할시/광역시), 6개 광역자치단체(강원, 전북, 제주 제외)
 - 시·군 통계 신뢰성 문제
- ⇒ 2011년도에는 강원, 전북, 제주를 포함한 162개 시·군을 모두 조사

○ 표본규모 및 표본추출방법

- 표본조사 지점 선정
- 표본규모 특히 시·군 통계의 오차
- 현행 추출방법은 비확률추출
⇒ 확률추출로 전환

- 표본의 대표성

⇒ 2011년도 표본은 34,500명으로 확대

시·군의 통계발표 시에 오차수준을 검토해서 발표를 결정

○ 조사내용 및 방법

- 조사내용을 환승실태와 이용자 만족도로 분리하는 방안
- 추출틀을 무엇으로 하는가에 따라 고려
- 현행의 149개 조사지점에서 조사할 것인지 또는 가구 방문으로 할 것 인지 논의

⇒ 가구단위조사로 전환하면 상당부분 문제 해결

○ 기타

- 항목무응답 대체 방법의 적절성
- 현재는 조사결과의 단순 집계
- 보고서 파일, 표본설계보고서, 설문지, 마이크로 자료 등이 필요
- 구조방정식모델을 이용한 가중값 산정?
- 2010년도 설문지

⇒ 현재의 평균대체 방법은 적절치 않음

확률추출로 전환하지 않으면 추정 불가능

□ 2차 연구협의회

- 일시 : 2011년 9월 9일(금), 14:00-16:00
- 장소 : 국토해양부 교통정책실 도시광역교통과
- 참석 :
 - 통계청 김경용사무관, 국토해양부 황현대주무관, 교통안전공단 김명희연구원
 - 청주대학교 류제복교수(연구책임자)

□ 협의내용

- 1차 회의에서 논의된 사항 중 재 검토가 필요한 사항들을 요약

○ 모집단

- 조사대상을 현행과 같이 주 4회 이상 이용하는 15세 이상 남녀로 할지 여부
 - ⇒ 대중교통이용의 과대 추정문제가 발생
- 표본추출틀
 - ⇒ 현행의 표본추출틀은 불명확하고 또한 조사지점을 교통안전공단에서 자체적으로 선정하지 않음
- ※ 표본의 대표성을 갖기 위해서는 표본추출틀의 대표성이 확보되어야 하며 사용 가능한 표본추출틀을 설정해야 함
- ※ 현행의 조사지점은 이용자들이 많은 지역을 선정하고 있는데 명확한 근거가 필요함

○ 조사범위

- 조사시점(화-목; 오전 7시 ~ 오후 9시)
 - ⇒ 화-목 이외의 요일 이용자들은 누락(예, 주말 이용자들)
- 시·군 통계 신뢰성
 - ⇒ 현재는 일괄적으로 표본을 선정. 시·군 통계는 신뢰성을 검토해서 발표

○ 표본규모 및 표본추출방법

- 표본조사지점 선정의 근거가 미흡하며 조사지점 수는 얼마인지?
- 표본규모와 시·군 통계의 오차는?
 - 2010년은 13,080명, 68개 시·군의 149개 존
 - 2011년은 34,500명으로 확대하고 전국에서 162개 시·군을 조사
- 현행 추출방법은 비확률추출
 - ⇒ 확률추출로 전환
- 표본의 대표성

○ 조사내용 및 방법

- 조사방법
 - 조사방법의 보완 또는 변경
 - 그리고 현행과 같이 조사지점에서 대면조사는 조사내용과 시간상 비현실적임
- 조사내용
 - 조사표의 수정, 보완이 요구됨

● 추가사항

- 2010년 조사는 68개 시·군(7개 특별시와 6개 광역시에서 40개 시와 21개 군)에서 149개 존을 시·군별로 어떻게 배분하는가?
1개 군에 1-2개?
 - 조사지역은 존당 5~6개
 - 2011년도에는 1,393개 조사지역 사용
- 1개 조사지점에서 몇 명을 조사하는가?
- 주요변수
 - 1주간 대중교통 이용횟수, 한 달 평균 대중교통비용, 환승률 등
- 조사지점 목록 작성 시 행정구역, 지점별 노선수, 구분(버스정류장, 지하철역, 환승센터 등)
- 조사결과는 전국 통계와 16개 광역자치단체에 대한 통계를 제시
현재와 같은 상황에서 시·군별 통계발표는 검토해야할 것임(신뢰도 점검이 필요함)
- 표본설계를 할 경우 2~3년 주기로 가능
- 표본배정 시 2010년도 인구주택총조사자료의 활용
- 요청 자료 확인
 - 2010년도 대중교통이용실태 및 이용자만족도조사결과 마이크로자료, 조사지점 정보, 조사지점당 조사 인원수

2절. 현행 조사에 대한 분석

“대중교통현황조사”는 조사지역을 행정구역 특성과 도시철도 운행유무에 따라 5개 그룹으로 나누고 있다. 2010년도(밀줄 친 시·군)와 2011년도 조사대상은 <표 1>에 있다. <표 1>에 의하면 2010년도는 7개 특별/광역시와 강원도, 전라북도 및 제주도를 제외한 6개 광역자치단체의 68개 시·군에서 13,080명을 조사하였다. 특히 경기도는 모든 시·군을 조사한 반면에 나머지 5개 자치광역단체는 일부 시·군만을 대상으로 조사가 실시되었다.

한편 2011년도에는 2010년에 제외되었던 제주특별자치도와 강원도, 그리고 전라북도가 포함되고 162개의 모든 시·군을 대상으로 34,500여명을 조사한다. 2010년과 마찬가지로 8개의 특별/광역시(자치도)는 시로 분류하였다.

<표 2>는 2010년도 그룹별/ 시도별 표본수를 나타내고 있다. 표본규모는 지역별 인구수를 고려하여 서울 1,000명, 부산, 대구, 인천은 600명, 광주와 대전은 500명, 그리고 울산은 400명을 할당하였다. 시는 인구규모에 따라 120명~300명을 배정하였고, 군은 일괄적으로 80명씩 배정하였다. 경기도에는 전체 표본수의 41.7%인 총 5,450명이 배정되어 표본이 한곳에 지나치게 편중되었다. 한편 충남과 전남은 조사대상 지역 중에서 인구수가 가장 작은 충북의 570명보다도 적은 520명을 각각 배정하였다. 이러한 결과는 표본배정이 인구규모를 기준으로 하고 있다고 하지만 배정기준이 적절치 못함을 보여주고 있다.

<표 1> 그룹별/시도별 조사대상 시·군

행정구역	행정구역 특성										
	특별시/광역시		시			군					
도시철도 운영유무 해당도시수 (162)	운영	비운영	운영			비운영					
	6	1	24			10	40				
	A그룹		B그룹		C그룹	D그룹	E그룹				
인구수 해당도시수 (162)	100만 이상		30만 이상	30만 미만	30만 이상	30만 미만	군지역*				
	6	1	14	10	11	39	81				
특별시 광역시 자치도 (8)	서울 부산 대구 대전 인천 광주	울산			제주						
경기(31) (27시, 4군)			수원 고양 안산 용인 시흥 의정부 남양주	성남 부천 안양 광명 평택 의정부 화성	군포 의왕 오산 과천시 구리 양주 동두천	파주 광주 하남 포천	이천 김포 안성	양평 가평 연천	여주		
강원(18) (7시, 11군)				춘천	원주	강릉 속초 태백 삼척	홍천 평창 영월 화천	고성 정선 인제 양구	철원 원성 영양		
충북(12) (3시, 9군)					청주	충주	제천	음성 옥천 괴산	청원 영동 보은	진천 증평 단양	
충남(16) (7시, 9군)			천안	아산		서산 보령 공주	논산 계룡	당진 예산 금산	홍성 서천 청양	연기 부여 태안	
전북(14) (6시, 8군)					전주 익산	군산 김제	정읍 남원	완주 실상 무주	부안 순창 장수	고창 진안	
전남(22) (5시, 17군)						여수 목포 나주	광양 순천	해남 영암 완도 신안 진도	고흥 안동 보성 곡성 진성 장흥	화성 영광 담양 함평 구례	순창 광안 평택
경북(23) (10시, 13군)					포항 구미	경산 영주 김천 영천	경주 안동 상주 경북	칠곡 예천 봉화 영양	의성 청도 청송 울릉 성주	울진 영덕 군위 고령	
경남(18) (8시,10군)				양산	창원 진해 마산 김해 진주	거제 사천	통영 밀양	창녕 합천 함양	거창 남해 산청 의령	함안 동성 고성	

* 2010년도 조사지역은 밀줄 친 지역이고, 2011년 조사부터는 창원, 진해, 마산을 창원시로 통합

<표 2> 2010년 그룹별/시도별 표본수

행정구역	행정구역 특성						합계	
	특별시/광역시		시			군		
도시철도 운영유무 해당도시수	운영	비운영	운영		비운영		162	
	6	1	24		10	40		81
	A 그룹		B 그룹		C 그룹	D 그룹		E 그룹
인구수 해당도시수	100만 이상		30만 이상	30만 미만	30만 이상	30만 미만	군지역	162
	6	1	14	10	11	39	81	
서울	1,000							1,000
부산	600							600
대구	600							600
인천	600							600
광주	500							500
대전	500							500
울산		400						400
제주								
경기			3,300	910	200	720	320	5,450
강원								
충북					250		320	570
충남						360	160	520
전북								
전남						120	400	520
경북					200	240	400	840
경남				130	770		80	980
합계	3,800	400	3,300	1,040	1,420	1,440	1,680	13,080

본 조사는 환승실태와 이용자 만족도조사이므로 환승실태와 만족도를 나타내는 변수를 주요변수라 할 수 있다. 따라서 환승과 관련된 변수로는 환승을 위한 이동시간, 대기시간, 이용요금, 그리고 환승만족도 등이다. 이용자 만족도에 대한 항목으로는 대중교통 서비스 품질 요인 평가 관련 만족도 항목과 전반적인 만족도로 교통수단 만족도, 기대대비 만족도, 교통시설 만족도 등이다. 보고서에 의하면 만족도 점수는 구조방정식모형으로 구한 가중값을 이용하여 계산한 가중평균이다. 그러나 2010년 조사 자료를 분석한 결과 대부분의 항목들은 범주형 자료이고 이용자만족도 점수도 단순평균한 것이다. 또한 이들 범주형 변수들을 평균한 값은 통계적으로 의미를 부여하기 어려울 뿐만 아니라 점수화된 것들의 평균과 표준편차가 지역별로 변별력이 떨어져서 새로운 표본설계를 하는 데 한계가 있다. 따라서 1주간 대중교통 이용횟수와 월 평균 대중교통비용 변수를 연구변수로 사용하였다.

<표 3>과 <표 4>는 월 평균 대중교통비용에 대해서 지역별, 인구규모별, 그리고 도시철도 운행 유무에 따라 분류된 그룹별 평균, 표준편차, 변동계수, 그리고 상대표준오차이다. 월 평균 대중교통비용에 대한 전국 표본평균은 45,845원이고 표준편차는 26,646원이다. 지역별로는 서울이 56,810원으로 제일 많고 충남이 30,400원으로 제일 적다. 그룹별로는 도시철도를 운행하는 인구 30만 이상의 B그룹(시)이 52,880원으로 제일 많고 도시철도를 운행하지 않는 E그룹(군)이 38,104원으로 제일 적다.

1주간 대중교통 이용횟수에 대한 지역별, 인구규모별, 그리고 도시철도 운행 유무에 따라 분류된 그룹별 평균, 표준편차, 변동계수, 그리고 상대표준오차는 <부록 1>에 있다.

<표 3> 월 평균 대중교통비용: 평균/표준편차 (단위; 원)

행정구역	행정구역 특성						계	
	특별시/광역시		시			군		
도시철도 운영유무 해당도시수	운영	비운영	운영		비운영		162	
	6	1	24		10	40		81
	A 그룹		B 그룹		C 그룹	D 그룹		E 그룹
인구수 해당도시수	100만 이상		30만 이상	30만 미만	30만 이상	30만 미만	군지역	162
	6	1	14	10	11	39	81	
서울	56,810 28,453							56,810 28,453
부산	44,900 23,259							44,900 23,259
대구	45,215 15,964							45,215 15,964
인천	53,199 31,165							53,199 31,165
광주	43,772 15,867							43,772 15,867
대전	34,971 13,573							34,971 13,573
울산		44,925 17,571						44,925 17,571
제주								
경기			52,880 33,698	44,533 21,778	32,530 12,681	47,308 33,320	48,715 30,815	49,759 31,573
강원								
충북					33,896 12,055		31,969 12,068	32,814 12,089
충남						30,066 12,768	31,150 15,485	30,400 13,656
전북								
전남						45,950 13,978	37,171 16,125	39,197 16,075
경북					35,620 11,734	40,666 20,828	38,657 22,572	38,508 20,046
경남				33,600 16,628	43,959 20,405		36,000 16,695	41,935 20,033
계	47,939 24,514	44,925 17,571	52,880 33,698	43,165 21,502	39,403 17,841	41,777 27,109	38,104 21,459	45,845 26,646

<표 4> 월 평균 대중교통비용: 변동계수/상대표준오차 (단위:%)

행정구역	행정구역 특성						계	
	특별시/광역시		시			군		
도시철도 운영유무 해당도시수	운영	비운영	운영		비운영		162	
	6	1	24		10	40		81
	A 그룹		B 그룹		C 그룹	D 그룹		E 그룹
인구수 해당도시수	100만 이상		30만 이상	30만 미만	30만 이상	30만 미만	군지역	162
	6	1	14	10	11	39	81	
서울	50.085 1.584							50.085 1.584
부산	51.803 2.115							51.803 2.115
대구	35.309 1.441							35.309 1.441
인천	58.582 2.392							58.582 2.392
광주	36.250 1.621							36.250 1.621
대전	38.814 1.736							38.814 1.736
울산		39.113 1.956						39.113 1.956
제주								
경기			63.726 1.109	48.903 1.622	38.984 2.757	70.433 2.625	63.256 3.536	63.452 0.860
강원								
충북					35.566 2.249		37.749 2.110	36.843 1.543
충남						42.467 2.238	49.713 3.930	44.921 1.970
전북								
전남						30.422 2.777	43.382 2.169	41.012 1.798
경북					32.944 2.330	51.218 3.306	58.390 2.920	52.058 1.796
경남				49.489 4.341	46.419 1.673		46.376 5.185	47.771 1.526
계	51.135 0.830	39.114 1.956	63.726 1.109	49.814 1.545	45.278 1.202	64.889 1.710	56.319 1.374	58.121 0.508

3절. 모집단 분석

“환승실태 및 이용자 만족도 조사”는 만 15세 이상인 사람들을 대상으로 한다. 따라서 가장 최근의 모집단 자료로 2010년 인구주택총조사 자료에서 15세 이상 인구를 모집단으로 한다.

2010년도 조사에서는 조사대상 지역을 구분하기 위해서 인구수를 기준으로 하였다. 그러나 조사대상이 15세 이상 인구이므로 15세 이상의 인구수를 기준으로 조사지역을 재분류하였다. 2010년 인구주택총조사에 의하면 총 인구는 48,580,293명인데 15세 이상 인구는 40,203,788명으로 총 인구의 약 82.8%를 차지하고 있다. 따라서 15세 인구수를 기준으로 조정된 조사대상 지역은 <표 5>에 있다. 2010년도 조사에서 인구 30만의 기준이 모집단을 새롭게 정의함으로써 15세 이상 인구 25만 기준으로 변경되었고, 이로 인해 전라북도의 익산이 C그룹에서 D그룹(15세 이상인구 25만 미만)으로 변경되었다.

2011년부터 본 조사는 162개의 시·군을 조사하여 16개 자치광역단체와 시·군별 통계를 생산한다. 전국 추정치를 생산하기 위해서 162개 층(시·군)의 가중치(15세 이상 총 인구수에 대한 해당 층의 15세 이상 인구수의 비)를 계산하였다. 8개 부차모집단(8개 광역자치단체)에 대한 추정을 위해서는 광역자치단체별로 가중치를 별도 계산해서 추정에 반영하면 된다.

<표 6>은 15세 이상 시·도별 인구수와 가중치를 나타내고 있으며, 전국 162개 시·군의 인구수와 가중치는 <부록 2>에 있다.

<표 5> 2010년 15세 이상 인구수에 따른 조사대상 시·군

행정구역	행정구역 특성									
	특별시/광역시		시			군				
도시철도 운영유무 해당도시수 (162)	운영	비운영	운영			비운영				
	6	1	24			10	40	81		
	A그룹		B그룹		C그룹	D그룹	E그룹			
인구수 해당도시수 (162)	80만 이상		25만 이상	25만 미만	25만 이상	25만 미만	군지역			
	6	1	14	10	10	40	81			
특별시 광역시 자치도 (8)	서울 부산 대구 대전 인천 광주	울산			제주					
경기(31) (27시, 4군)			수원 의정부 부천 평택 고양 용인 화성	성남 안양 광명 안산 남양주 시흥	동두천 과천시 구리 오산 군포 의왕 양주 춘천	파주 하남 안성 광주	이천 김포 포천	여주 연천 가평	양평	
강원(18) (7시, 11군)					원주	강릉 태백 삼척	동해 속초	홍천 평창 화천 고성	횡성 정선 양구 영양 강릉	영월 철원 인제 양양
충북(12) (3시, 9군)					청주	충주	제천	청원 영동 음성	보은 진천 단양	옥천 괴산 증평
충남(16) (7시, 9군)			천안	아산		공주 서산 계룡	보령 논산	금산 서천 예산	연기 청양 태안	부여 홍성 당진
전북(14) (6시, 8군)					전주	군산 정읍 김제	익산 남원	완주 장수 고창	진안 임실 부안	무주 순창
전남(22) (5시, 17군)						목포 순천 광양	여수 나주	담양 고흥 장흥 영암 영진	곡성 보성 강진 무안 장성 광안	구례 화순 해남 함평 완도 신안
경북(23) (10시, 13군)					포항 구미	경주 안동 영문	김천 영양 상주 경산	군위 영고 영천	의성 영주 성주 봉화 영양	청송 도곡 칠곡 울진
경남(18) (8시, 10군)				양산	진주 김해 창원	통영 밀양	사천 거제	의령 고성 산청	함안 남해 합성 합천	창녕 동해 거창

* 전북 익산은 15세 인구기준으로 C그룹에서 D그룹(25만 미만의 시)으로 변경

<표 6> 2010년 15세 이상 시·도별 인구수 및 가중치

행정구역	인구수				가중치
	특별시/광역시/ 자치도	시	군	계	
해당 도시수	8	73	81	162	
전국	19,181,024	17,782,726	3,240,038	40,203,788	
서울	8,276,179			8,276,179	0.20586
부산	2,930,640			2,930,640	0.07289
대구	2,043,570			2,043,570	0.05083
인천	2,194,851			2,194,851	0.05459
광주	1,196,956			1,196,956	0.02977
대전	1,233,503			1,233,503	0.03068
울산	877,758			877,758	0.02183
제주	427,567			427,567	0.01063
경기		8,930,597	233,440	9,164,037	0.22794
강원		914,720	319,461	1,234,181	0.03070
충북		826,474	422,608	1,249,082	0.03107
충남		1,128,963	543,684	1,672,647	0.04160
전북		1,221,144	256,158	1,477,302	0.03675
전남		811,152	641,484	1,452,636	0.03613
경북		1,757,286	432,892	2,190,178	0.05448
경남		2,192,390	390,311	2,582,701	0.06424

4절. 새로운 표본설계

2010년도에는 강원도, 전라북도, 제주도를 제외하고 서울을 비롯한 7개 특별/광역시와 경기도를 포함한 6개 광역자치단체의 68개 시·군에서 13,080명을 조사하였다. 그러나 2011년도에는 모든 지역 즉, 8대 특별/광역시(자치도)(서울특별시, 부산광역시, 대구광역시, 인천광역시, 광주광역시, 대전광역시, 울산광역시, 제주특별자치도)와 8개 광역자치단체(경기도, 충청남도, 충청북도, 강원도, 전라남도, 전라북도, 경상남도, 경상북도)에서 총 162개 시·군을 대상으로 2011년 8월 23일부터 34,500명을 표본조사 한다.

2009년도 정기품질진단 결과 제기된 문제점과 개선사항, “2010년 대중교통현황조사” 보고서와 관련 자료, 그리고 교통연구원의 담당자 및 실사팀장과의 회의결과 본 과제의 목적을 달성하기 위해서는 “환승실태 및 이용자 만족도조사”에 대한 새로운 표본설계가 요구된다.

새로운 표본설계의 특징은 다음과 같다.

- 본 조사를 통해 16개 특별/광역자치단체에 대한 통계는 물론 8개 광역자치단체의 시·군별 통계 생산을 목적으로 하므로 8개 광역자치단체의 시·군을 부차모집단(광역자치단체)의 층으로 한다. 그러나 8개 특별/광역시(자치도)는 구별 통계를 생산하지 않는다.
- 새로운 표본설계에서 사용할 주요변수로는 월 평균 대중교통비용을 사용한다. 월 평균 대중교통비용이 1주간 대중교통 이용횟수보다 변동이 크기 때문이다.

- 조사는 각 지역에 할당된 표본에 대해서 조사지점(버스정류장, 지하철역, 환승센터 등)을 선정에서 조사를 실시하고 있다. 그러나 현재 조사지점에 대한 모집단 정보와 추출방법 등에 대한 자료가 전무한 상태이므로 이 부분에 대해서는 다루지 않는다.

(1) 총화

“대중교통현황조사”의 “환승실태 및 이용자 만족도 조사”는 15세 이상의 전 국민을 대상으로 전국단위의 통계는 물론 16개 특별/광역시자치단체의 통계생산을 목적으로 한다. 2011년부터는 전국의 162개 시·군별 통계생산을 계획하고 있으므로 162개의 시·군을 층으로 한다.

16개의 특별/자치광역시자치단체를 각각의 부차모집단으로 간주하지만, 8개의 특별/광역시(자치도)는 구별 통계를 생산하지 않으므로 이들 8개의 특별/광역시(자치도)는 각각 하나의 층이 된다. 그러나 경기도 등 8개 광역자치단체는 시·군 층으로 구성한다.

조사결과 자료는 필요에 따라 사후총화할 수 있다. 예를 들어, 2010년도 조사결과에 따르면 모집단을 수도권, 부산/울산권, 대구권, 광주권, 대전권 등 5개 권역으로 분류하여 통계를 생산하고 있다. 또한 전철/도시철도의 운행지역과 비운행지역, 그리고 인구규모 등으로 분류하여 통계를 생산하기도 하므로 사후총화과정이 필요하다.

(2) 표본크기 및 표본배정

2011년부터는 전국의 162개 시·군별 통계를 생산하고자 하므로 이들을 층으로 하였다. 이들 층을 부차모집단으로 하여 부차모집단별 표본크기를 구하기 위해서 다음 식을 사용한다.

$$n = \frac{\left(\frac{CV}{O}\right)^2}{1 + \frac{1}{N}\left(\frac{CV}{O}\right)^2}$$

여기서, CV 는 부차모집단의 월 평균 대중교통비용에 대한 변동계수, O 는 부차모집단의 목표오차(상대표준오차), N 은 부차모집단의 15세 이상 인구수이다.

본 조사는 인구규모, 도시철도 운행여부 등에 따라 여러 개의 그룹으로 나누고 설문지도 5개 종류로 구분하였다. 따라서 표본을 결정하기 위해서 이러한 그룹에 따른 목표오차를 달리하여 표본을 정하였다.

표본크기를 정하는데 다음과 같은 규칙을 사용하였다.

첫째, 기존의 조사체계에 따라 그룹을 정하며, 15세 이상의 인구규모에 따라 재분류하였다.

둘째, 동일 광역시 내에서 조사가 이루어지지 않은 곳은 같은 그룹에서 조사가 이루어진 시나 군의 평균을 사용하였다. 예를 들면, 충청북도에서 조사가 실시되지 않은 충주시와 제천시 조사가 이루어진 청주시의 자료만을 이용하였다. 물론 청주시의 15세 이상 인구 규모가 다르지만 동일한 시로 간주하였다. 그리고 진천 등 조사가 이루어지지 않은 군 지역은 조사가 이루어진 군들의 평균값을 사용하였다.

셋째, 강원도와 전라북도는 전체적으로 조사가 이루어지지 않았으므로 해당 그룹의 평균값을 사용하였는데, 표본 규모가 유사한 다른 지역과 큰 차이가 발생하였다. 그 원인은 해당 그룹의 전체 평균이 경기도의 영향을 크게 받기 때문이다. 따라서 경기도를 제외

한 충북, 충남, 전남, 경북, 경남 지역의 그룹별 평균값을 사용하였다.

그룹별 목표오차를 달리 한 경우의 표본크기는 <표 7>에 계산되어 있다.

<표 7> 목표오차별 표본크기

행정구역	목표오차				표본크기
	특별시/광역시	시		군	
		25만 이상	25만 미만		
1안	1%	2%	3%	4%	54,630
2안	1.5%	2%	3%	4%	46,740
3안	1.5%	3%	3%	4%	36,816
4안	2%	3%	3%	4%	34,054
5안	2%	3%	4%	5%	25,014

조사지역을 <표 7>에서와 같이 4그룹으로 구분하여 목표오차를 달리 하였다. 목표오차는 그룹별로 추정하고자 하는 정도에 따라 달라지므로 조사주관기관이 조정할 수 있다. 여기서는 2011년도의 표본규모와 비슷한 3안을 기준으로 36,816명의 표본을 제시하였다. 그룹별/ 시도별 표본수는 <표 8>에 있고 162개 시·군의 표본수는 <부록 2>에 있다.

<표 7>과 <부록 2>의 표본배정결과 몇 가지 특이 사항 및 고려 사항은 다음과 같다.

- 표본크기는 모집단 변동계수에 영향을 받는다. 모집단 변동계수의 추정치로 2010년도 자료를 사용하다보니 2010년도의 변동계수가 큰 인천광역시는 표본규모가 크게 되었다.
- 경기도의 안성시와 화성시 등도 변동계수가 큰 것에 기인하여 표본규모가 크게 되었다.

- 경기도 과천시, 충북 청원군 그리고 전남의 곡성군 등의 변동계수는 지나치게 작아 표본규모가 타 군에 비해 작게 되었다.

상기와 같이 일부 지역에서 변동계수가 지나치게 크거나 작은 경우는 해당 지역의 조사결과가 타 지역과 상당히 다르기 때문에 발생하게 된다. 그 원인은 실사과정에서의 품질 차이에 의한 경우도 있다. 따라서 조사과정에 보다 철저한 준비가 필요하다.

3안에 따른 162개 시·군에 배정된 표본수는 다음 사항들을 고려해서 조정할 수 있다.

- 2010년 조사에서는 최소표본으로 80명(군 지역)을 배정하였으므로 새로운 표본설계에서도 조사비용이나 실사 등을 고려해서 최소표본규모를 정한다.
- 표본수가 일부 과다 또는 과소 배정된 지역은 과거 수년간의 조사(또는 2011년)결과를 검토해서 조정할 수 있다.
- 조정된 표본규모는 가중치 조정을 통해 추정과정에 반영해 줄 수 있다.

새로운 표본설계에서는 전국 162개 시·군이 모두 조사대상에 포함되었으나 이용 가능한 자료가 2010년의 68개 시·군 자료이므로 한계가 있었다. 특히 2010년도에 조사가 이루어지지 않은 지역은 활용 가능한 자료가 매우 제한적이었다. 그러나 새로운 표본설계에 의해 조사가 실시된 후 조사 자료는 향후 새로운 표본설계에 매우 유용하게 활용될 것으로 기대된다.

<표 8> 새로운 표본설계에 의한 그룹별/시도별 표본수

행정구역	행정구역 특성						합계	
	특별시/광역시		시			군		
도시철도 운영유무 해당도시수	운영	비운영	운영		비운영		162	
	6	1	24		10	40		
	A 그룹		B 그룹		C 그룹	D 그룹		E 그룹
인구수 해당도시수	80만 이상		25만 이상	25만 미만	25만 이상	25만 미만	군지역	162
	6	1	14	10	10	40	81	
서울	1,115							1,115
부산	1,192							1,192
대구	554							554
인천	1,524							1,524
광주	584							584
대전	669							669
울산		679						679
제주					182			182
경기			5,984	1,556	169	3,230	912	11,851
강원				272	185	975	1,265	2,697
충북					141	280	748	1,169
충남			187	187		939	1,157	2,470
전북					185	975	920	2,080
전남						515	1,646	2,161
경북					242	2,174	2,213	4,629
경남				272	683	965	1,340	3,260
합계	5,638	679	6,171	2,287	1,787	10,053	10,201	36,816

(3) 추정

본 조사를 통해서 16개 특별/광역시자치단체(부차모집단)에 대한 통계는 물론 경기도, 강원도 등의 8개 광역자치단체의 시·군별 통계생산을 목적으로 하므로 표본조사된 자료를 바탕으로 모집단 특성치에 대한 추정이 필요하다. 따라서 전체 모집단에 대한 추정과 부차모집단에 대한 추정을 나누어서 살펴본다.

□ 전체 모집단에 대한 추정

- 추정에 사용될 기호들은 다음과 같다.

N_h : h 층의 15세 이상 인구수, $h=1, 2, \dots, L$

n_h : h 층의 표본수, $h=1, 2, \dots, L$

$N = \sum_{h=1}^L N_h$: 15세 이상 전체 인구수

$n = \sum_{h=1}^L n_h$: 전체 표본수

y_{hi} : h 층의 i 번째 층(시, 군)으로부터 얻은 월 평균 대중교통비용
(또는 1주간 대중교통 이용횟수 등)에 대한 측정값

$f_h = \frac{n_h}{N_h}$: 추출률

$W_h = \frac{N_h}{N}$: h 층의 상대크기

- 전체 모집단에 대한 월 평균 대중교통비용의 추정치인 평균(\bar{y})은 다음과 같다.

$$\bar{y} = \sum_{h=1}^L W_h \bar{y}_h = \sum_{h=1}^L W_h \frac{1}{n_h} \sum_{i=1}^{n_h} y_{hi}$$

그리고 모집단의 월 평균 대중교통비용에 대한 추정치인 \bar{y} 의 분산 추정치는 다음과 같다.

$$var(\bar{y}) = \sum_{h=1}^L W_h^2 (1 - f_h) \frac{s_h^2}{n_h}$$

여기서, $s_h^2 = \frac{1}{n_h - 1} \sum_{i=1}^{n_h} (y_{hi} - \bar{y}_h)^2$ 이다. 한편 전체 모집단에 대한 월 평균 대중교통비용의 추정치인 \bar{y} 의 상대표준오차(CV)는 다음과 같다.

$$\widehat{CV}(\bar{y}) = \frac{\sqrt{var(\bar{y})}}{\bar{y}} \times 100$$

□ 부차모집단(층)에 대한 추정

- 16개 부차모집단(특별/광역자치단체)에서 8개의 부차모집단(경기도 등의 광역자치단체)에 대한 추정은 전체모집단에 대한 추정과 동일하다. 다만, 부차모집단별로 층의 수는 시·군의 수가 된다. 이때 가중치는 해당 부차모집단(광역자치단체)의 총 15세 이상 인구수에 대한 시·군의 15세 이상 인구수가 된다.

(4) 무응답대체

본 조사에서는 모든 설문 문항에 대해서 응답하지 않는 단위무응답(unit nonresponse)과 일부 항목에 대해서만 응답하지 않는 항목무응답(item nonresponse)이 발생한 것으로 예상된다. 그러나 응답거부 등의

이유로 발생하는 단위무응답은 응답자를 대체하는 방식으로 처리하고 있다. 다만 설문지의 일부 항목에 응답하지 않은 항목무응답에 대해서는 평균대체 방법을 사용하여 무응답 항목을 대체해 주고 있다. 그러나 본 조사 항목의 대부분이 범주형이므로 평균대체 방법은 적절치 않다. 따라서 항목무응답이 발생한 경우 대체 방법을 소개한다. 여기서는 류제복·이승주(2011)의 내용을 인용하고, 자세한 사항들은 조사통계연구회(2000)와 최필근(2009a, 2009b, 2010) 등을 참고하면 된다.

□ 연역적 대체(deductive imputation)

연역적 대체방법은 비교적 간단하고 쉽게 항목무응답을 대체할 수 있는 방법이지만 논리적 제약조건이 명확하고 무응답과 관련된 보조자료가 충분해야 가능하다. 추론에 사용된 자료가 확실하지 않을 경우에는 무응답에 대한 대체값에 오류가 생기게 된다. 따라서 연역적 대체를 위해서는 결측값과 관련된 항목들에 응답한 자료들 간의 일치성을 점검해야 한다.

□ 평균대체(mean imputation)

평균대체는 무응답에 대한 대체값으로 대체군내의 응답값들의 평균을 사용하는 방법으로 간편하고 이용하기 쉽다. 항목변수가 양적변수이고 구하고자 하는 통계량이 평균일 때 적절하나 본 조사에서와 같이 범주형 자료의 경우는 적절치 않다. .

크기 N 인 모집단에서 크기 n 인 단순확률표본을 추출하였을 때 응답표본과 무응답표본을 각각 R 과 R^c 로 표기하자. 이때 응답단위는 r 개이다. 그러면 평균대체는 $(n-r)$ 개의 무응답 단위의 대체값으로 r 개의

응답단위들의 평균값을 사용한다. 즉, 대체후의 자료는 다음과 같다.

$$y_k^* = \begin{cases} y_k, & k \in R, \quad (\text{응답값}) \\ \bar{y}_r, & k \in R^c, \quad (\text{대체값}) \end{cases}$$

따라서 모집단 평균은 다음과 같이 관찰된 값들과 대체값들의 평균으로 추정된다.

$$\begin{aligned} \bar{y}_I &= \frac{1}{n} \left[\sum_{k \in R} y_k + \sum_{k \in R^c} \bar{y}_r \right] = \frac{1}{n} [r\bar{y}_r + (n-r)\bar{y}_r] \\ &= \bar{y}_r \end{aligned}$$

평균대체는 무응답들에 모두 동일한 평균값을 대체하게 되므로 평균값의 빈도가 지나치게 높아서 관심변수의 경험적 분포가 왜곡되는 단점이 있다.

□ 비대체(ratio imputation)

무응답에 대한 대체값으로 관심변수와 관심이 있는 보조변수를 이용하여 다음과 같이 대체하는 방법을 비대체라 이 대체 비대체방법을 사용하기 위대한 대관심변수와 보조변수가 원점을 지나 대직선 관계를 갖고 분산이 보조변수에 비례하는 경우 효과적이다.

$$y_k^* = \begin{cases} y_k, & k \in R, \quad (\text{응답값}) \\ \left(\frac{\bar{y}_r}{\bar{x}_r} \right) x_k, & k \in R^c, \quad (\text{대체값}) \end{cases}$$

한편 기존의 대체값에 확률오차인 e_k 를 더해주는 랜덤비대체방법도 사용한다.

□ 회귀대체(regression imputation)

비대체방법과 유사하게 무응답에 대한 대체값으로 관심변수와 관련

이 있는 보조변수를 이용한다. 이때 관심변수와 보조변수가 절편이 있는 직선 관계를 갖고 관심변수의 분산이 동일 할 때 유용하다.

$$y_k^* = \begin{cases} y_k, & k \in R, \quad (\text{응답값}) \\ \bar{y}_r + b(x_k - \bar{x}_r), & k \in R^c, \quad (\text{대체값}) \end{cases}$$

여기서 $b = \sum_{k \in R} (x_k - \bar{x}_r)(y_k - \bar{y}_r) / \sum_{k \in R} (x_k - \bar{x}_r)^2$ 이다.

랜덤비대체와 유사하게 기존의 대체값에 확률오차인 e_k 를 더해주는 랜덤회귀대체방법을 사용하기도 한다.

□ 핫덱대체(hot-deck imputation)

1962년 미국의 CPS(current population survey)에서 연간소득에 대한 결측값을 대체하는 데 처음으로 사용된 핫덱대체는 최근에 실제조사에서 가장 널리 사용되고 있다. 특히 본 조사에서 흔히 사용되는 범주형 자료의 경우 매우 유용하다.

랜덤핫덱은 무응답 대체를 위해서 대체군내의 응답값들 중에서 확률적으로 하나의 값을 선정해서 대체하는 방법이다. 즉,

$$y_k^* = \begin{cases} y_k, & k \in R, \quad (\text{응답값}) \\ y_i, & k \in R^c, \quad (\text{대체값}) \end{cases}$$

이다. 여기서 y_i 는 응답값들 중에서 확률적으로 선정된 값이다. 랜덤핫덱 방법은 표본대체 후에도 표본의 분포가 그대로 유지되는 장점이 있고 통계량의 형태에 관계없이 사용될 수 있다.

한편 응용핫덱은 랜덤핫덱의 응용으로 연속형 항목에서도 적용할 수 있도록 한 방법이다. 대체군 사용 시 범주형 항목은 항목값의 일치 여부를 판단해서 점수를 부여하고 연속형 범주는 신뢰구간을 이용해서 그 구간에 포함되면 비슷한 개체로 간주하여 점수를 부여한다. 모든

대체군의 항목들에 대해서 가정 점수가 높은 개체를 선택하고 그 중에서 하나를 확률적으로 선택해서 대체하는 방법이다.

그밖에도 축차핫덱(sequential hot-deck), 가중핫덱(weighted hot-deck) 등의 방법을 사용하기도 한다.

4장. 결론 및 제안

“대중교통현황조사”(국토해양부, 교통안전공단)는 정부의 각종 대중교통 지원정책의 합리적인 근거 제공 및 국내 대중교통수준의 해외 비교를 통해 최적의 대중교통체계 발전을 유도하기 위한 것이다.

2009년도 통계청의 정기품질진단 결과 통계품질 향상을 위한 몇 가지 문제점이 지적되고 개선사항들이 제시되었다. 이와 관련하여 일부 전문성이 요구되는 부문에 대해 통계청에서는 작성기관의 품질개선지원 수요를 바탕으로 본 통계를 2011년 품질개선지원 대상으로 선정하게 되었다.

이번 연구를 통해 “대중교통 환승실태 및 이용자 만족도 조사”의 표본설계와 설문지 등에 대한 미흡한 점을 보완해서 실제적인 통계품질이 개선될 수 있는 방안들을 제시한다. 그리고 향후 연구 검토해야할 사항들도 제안한다.

□ 결론

- 현행 조사(2010년)에 대한 분석결과 표본배정의 타당성이 부족하였다. 새로운 표본설계에 사용할 연구변수를 선정하였다.
- 신규 모집단으로 2010년 인구주택총조사 자료에서 15세 이상 인구를 사용하였고 지역별 인구분포를 살펴보았다.
- 새로운 표본설계를 위해 층화, 표본크기 및 표본배정 그리고 추정치를 다루고 항목무응답 대체방법을 소개하였다.

□ 제안

- 본 조사를 인구주택총조사 자료를 이용한 가구단위조사로의 전환을 검토한다. 이의 장점은 다음과 같다.
 - 모집단을 대표할 수 있는 표본추출틀의 확보가 가능
 - 표본설계를 바탕으로 표본규모 선정, 표본추출, 추정 등이 가능하여 표본의 대표성을 높일 수 있음
 - 현행 조사는 선정된 조사지점에서 조사가 이루어지므로 시간적으로 조사수행이 원활하지 못함
- 현행 조사지점 선정방법의 보완사항
 - 조사범위에 속하는 모든 지역의 정류소, 환승지역 등의 조사 대상 지점 목록과 그 지점에서의 노선 수 자료를 확보(대중교통 이용객 수에 대한 자료부족)하여 조사지점 선정에 활용
 - 조사지점에서 응답자는 계통추출로 선정하고 조사시간도 조정이 필요함
- “대중교통 환승실태 및 이용자만족도조사” 2011년 설문지(특별/광역시 기준)에 대한 검토사항
 - part-II의 “문5-1”은 Part-I의 “문1”인 기본적인 서비스에 대한 질문을 종합한 만족도와 유사함.
 - part-II의 “문5-3”은 Part-I의 “문3”인 외부 환경요소와 관련된 서비스에 대한 질문을 종합한 만족도와 유사함.
 - 따라서 part-II의 전반적인 만족도는 part-I 으로부터 얻을 수 있으므로 생략 가능
 - “문7-1”에서 3번째 질문인 대중교통 접근시간에 대한 만족도는 개인에 따라 같은 시간이라도 다를 수 있으므로 접근수단

과 접근시간만 질문.

－ “문7-2”에는 다음과 같은 사유로 응답자가 정확하게 응답하기 곤란하므로 보완 필요.

① 환승지점일 경우 다음 수단에 대한 시간이나 비용을 언급하기 어려움

② 환승일 경우 요금계산이 곤란

－ “문8”은 환승만족도에 대한 질문이므로 3)항의 대중교통 요금체계를 대중교통 환승요금체계로 수정

※ 시, 군 등의 타 설문지에도 동일하게 적용

참 고 문 헌

- 교통안전공단(2010), “2010 대중교통현황조사-환승실태 및 이용자 만족도조사”
결과보고서.
- 국토해양부, 교통안전공단(2010), “2010년 대중교통현황조사” 결과보고서
- 김영원 · 류제복 · 박진우 · 홍기학 공역(2009). 표본조사의 이해와 활용,
센게이지러닝코리아(주).
- 류제복 · 이승주(2010), “정보통신공사업경영분석 표본설계방법연구”, 통
계청 연구용역 최종보고서.
- 류제복 · 이승주(2011), “『국민건강 영양조사』 품질개선지원” 통계청
연구용역 최종보고서.
- 박홍래(2000). 통계조사론(개정판), 영지문화사.
- 산업자원부 · 한국전자거래진흥원(2006), ICT 투자규모 산출방법연구.
- 조사통계연구회(2000), *무응답오차*, 자유아카데미.
- 최필근(2009a), “농업총조사 무응답 대체기법 연구(I)”, *연구보고서*, 통계개발
원.
- 최필근(2009b), “농업총조사 무응답 대체기법 연구(II)”, *연구보고서*, 통계개발
원.
- 최필근(2010), “출생전후기 사망통계의 무응답 대체기법”, *연구보고서*, 통계개
발원.
- 통계청(2009), “대중교통 현황조사 2009년 정기통계품질진단”, *연구용역 최종
보고서*.
- 한국조사연구학회(2008), 2008년 산업 · 직업별 고용구조조사 표본설계

연구.

한국통계학회(2006), 사업체근로실태조사 표본설계 최종보고서.

Cochran, W. G.(1977), Sampling Techniques, third edition, John Wiley & Son, Inc.

<http://kosis.kr/> : 통계청 국가통계포털

<부록 1>

1주간 대중교통 이용횟수: 평균/표준편차

(단위; 원)

행정구역	행정구역 특성						계	
	특별시/광역시		시			군		
도시철도 운영유무 해당도시수	운영	비운영	운영		비운영		162	
	6	1	24		10	40		81
	A 그룹		B 그룹		C 그룹	D 그룹		E 그룹
인구수 해당도시수	100만 이상		30만 이상	30만 미만	30만 이상	30만 미만	군지역	162
	6	1	14	10	11	39	81	
서울	10.774 5.613							10.774 5.613
부산	9.912 4.254							9.912 4.254
대구	10.343 4.062							10.343 4.062
인천	11.048 4.781							11.048 4.781
광주	9.400 3.583							9.400 3.583
대전	8.296 3.634							8.296 3.634
울산		9.503 3.658						9.503 3.658
제주								
경기			10.151 4.919	8.490 4.306	9.200 2.196	8.479 3.809	7.663 3.562	9.472 4.620
강원								
충북					7.884 3.262		6.869 3.043	7.314 3.179
충남						6.456 2.933	6.213 2.974	6.381 2.945
전북								
전남						7.042 2.974	7.358 2.934	7.285 2.944
경북					8.080 2.659	7.563 3.886	7.098 3.056	7.464 3.252
경남				7.100 3.223	8.843 3.771		7.238 3.045	8.481 3.710
계	10.106 4.640	9.503 3.658	10.151 4.919	8.316 4.209	8.617 3.384	7.701 3.656	7.146 3.142	9.150 4.417

<부록 1:계속> 1주간 대중교통 이용횟수: 변동계수/상대표준오차 (단위; %)

행정구역	행정구역 특성						계	
	특별시/광역시		시			군		
도시철도 운영유무 해당도시수	운영	비운영	운영		비운영		162	
	6	1	24		10	40		81
	A 그룹		B 그룹		C 그룹	D 그룹		E 그룹
인구수 해당도시수	100만 이상		30만 이상	30만 미만	30만 이상	30만 미만	군지역	162
	6	1	14	10	11	39	81	
서울	52.095 1.647							52.095 1.647
부산	42.916 1.752							42.916 1.752
대구	39.267 1.603							39.267 1.603
인천	43.276 1.767							43.276 1.767
광주	38.117 1.705							38.117 1.705
대전	43.806 1.959							43.806 1.959
울산		38.499 1.925						38.499 1.925
제주								
경기			48.462 0.844	50.713 1.681	23.874 1.688	44.918 1.674	46.487 2.599	48.774 0.661
강원								
충북					41.376 2.617		44.307 2.477	43.459 1.820
충남						45.441 2.395	47.879 3.785	46.161 2.024
전북								
전남						42.240 3.856	39.884 1.994	40.412 1.772
경북					32.905 2.327	51.385 3.317	43.063 2.153	43.573 1.503
경남				45.389 3.981	42.642 1.537		42.072 4.704	43.753 1.398
계	45.915 0.745	38.499 1.925	48.462 0.844	50.617 1.570	39.272 1.042	47.474 1.251	43.967 1.073	48.274 0.422

<부록 2>

2010년 15세 이상 시·군별 인구수, 가중치 및 표본수

행정구역	인구수				가중치	표본수
	특별시/광역시/ 자치도	시	군	계		
해당도시수	8	73	81	162		
전국	19,181,024	17,782,726	3,240,038	40,203,788		36,816
서울	8,276,179			8,276,179	0.205856	1,115
부산	2,930,640			2,930,640	0.072895	1,192
대구	2,043,570			2,043,570	0.050830	554
인천	2,194,851			2,194,851	0.054593	1,524
광주	1,196,956			1,196,956	0.029772	584
대전	1,233,503			1,233,503	0.030681	669
울산	877,758			877,758	0.021833	679
제주	427,567			427,567	0.010635	182
경기		8,930,597	233,440	9,164,037		11,851
수원시		866,531		866,531	0.021553	445
성남시		788,645		788,645	0.019616	326
의정부시		343,628		343,628	0.008547	355
안양시		498,566		498,566	0.012401	345
부천시		706,321		706,321	0.017569	421
광명시		269,439		269,439	0.006702	334
평택시		308,792		308,792	0.007681	678
동두천시		75,007		75,007	0.001866	282
안산시		570,768		570,768	0.014197	571
고양시		736,149		736,149	0.018310	489
과천시		54,394		54,394	0.001353	58
구리시		151,482		151,482	0.003768	254
남양주시		421,317		421,317	0.010480	342
오산시		143,378		143,378	0.003566	233
시흥시		314,649		314,649	0.007826	239
군포시		224,500		224,500	0.005584	246
의왕시		118,685		118,685	0.002952	117
하남시		115,525		115,525	0.002873	276
용인시		673,354		673,354	0.016749	457
파주시		261,768		261,768	0.006511	169
이천시		156,630		156,630	0.003896	381
안성시		146,162		146,162	0.003636	1,416
김포시		176,500		176,500	0.004390	328
화성시		364,404		364,404	0.009064	982
광주시		182,518		182,518	0.004540	476
양주시		147,706		147,706	0.003674	366

<부록 2: 계속> 2010년 15세 이상 시·군별 인구수, 가중치 및 표본수

행정구역	인구수				가중치	표본수
	특별시/광역시/ 자치도	시	군	계		
해당도시수	8	73	81	162		
전국	19,181,024	17,782,726	3,240,038	40,203,788		36,816
포천시		113,779		113,779	0.002830	353
여주군			83,843	83,843	0.002085	272
연천군			35,578	35,578	0.000885	170
가평군			43,141	43,141	0.001073	274
양평군			70,878	70,878	0.001763	196
강원		914,720	319,461	1,234,181		2,697
춘천시		230,894		230,894	0.005743	272
원주시		255,366		255,366	0.006352	185
강릉시		184,823		184,823	0.004597	195
동해시		75,133		75,133	0.001869	195
태백시		43,696		43,696	0.001087	195
속초시		66,799		66,799	0.001662	195
삼척시		58,009		58,009	0.001443	195
홍천군			53,299	53,299	0.001326	115
횡성군			32,662	32,662	0.000812	115
영월군			30,467	30,467	0.000758	115
평창군			32,108	32,108	0.000799	115
정선군			31,044	31,044	0.000772	115
철원군			35,871	35,871	0.000892	115
화천군			18,627	18,627	0.000463	115
양구군			16,061	16,061	0.000399	115
인제군			23,757	23,757	0.000591	115
고성군			23,436	23,436	0.000583	115
양양군			22,129	22,129	0.000550	115
충북		826,474	422,608	1,249,082		1,169
청주시		543,082		543,082	0.013508	141
충주시		169,928		169,928	0.004227	140
제천시		113,464		113,464	0.002822	140
청원군			117,540	117,540	0.002924	62
보은군			26,758	26,758	0.000666	75
옥천군			42,580	42,580	0.001059	56
영동군			40,427	40,427	0.001006	118
진천군			49,082	49,082	0.001221	133
괴산군			28,029	28,029	0.000697	76
음성군			68,133	68,133	0.001695	76

<부록 2: 계속> 2010년 15세 이상 시·군별 인구수, 가중치 및 표본수

행정구역	인구수				가중치	표본수
	특별시/광역시/ 자치도	시	군	계		
해당도시수	8	73	81	162		
전국	19,181,024	17,782,726	3,240,038	40,203,788		36,816
단양군			24,463	24,463	0.000608	76
증평군			25,596	25,596	0.000637	76
충남		1,128,963	543,684	1,672,647		2,470
천안시		460,889		460,889	0.011464	187
공주시		104,584		104,584	0.002601	242
보령시		82,298		82,298	0.002047	187
아산시		221,945		221,945	0.005520	187
서산시		126,601		126,601	0.003149	187
논산시		100,837		100,837	0.002508	149
계룡시		31,809		31,809	0.000791	174
금산군			45,889	45,889	0.001141	96
연기군			68,593	68,593	0.001706	165
부여군			58,676	58,676	0.001459	128
서천군			47,286	47,286	0.001176	128
청양군			26,187	26,187	0.000651	128
홍성군			70,127	70,127	0.001744	128
예산군			67,125	67,125	0.001670	128
태안군			46,945	46,945	0.001168	128
당진군			112,856	112,856	0.002807	128
전북		1,221,144	256,158	1,477,302		2,080
전주시		530,222		530,222	0.013188	185
군산시		214,544		214,544	0.005336	195
익산시		245,884		245,884	0.006116	195
정읍시		92,371		92,371	0.002298	195
남원시		65,858		65,858	0.001638	195
김제시		72,265		72,265	0.001797	195
완주군			68,891	68,891	0.001714	115
진안군			18,009	18,009	0.000448	115
무주군			19,037	19,037	0.000474	115
장수군			16,845	16,845	0.000419	115
임실군			20,901	20,901	0.000520	115
순창군			21,811	21,811	0.000543	115
고창군			46,527	46,527	0.001157	115
부안군			44,137	44,137	0.001098	115
전남		811,152	641,484	1,452,636		2,161

<부록 2: 계속> 2010년 15세 이상 시·군별 인구수, 가중치 및 표본수

행정구역	인구수				가중치	표본수
	특별시/광역시/ 자치도	시	군	계		
해당도시수	8	73	81	162		
전국	19,181,024	17,782,726	3,240,038	40,203,788		36,816
목포시		202,421		202,421	0.005035	103
여주시		222,189		222,189	0.005527	103
순천시		209,509		209,509	0.005211	103
나주시		67,793		67,793	0.001686	103
광양시		109,240		109,240	0.002717	103
담양군			36,382	36,382	0.000905	100
곡성군			23,891	23,891	0.000594	35
구례군			19,498	19,498	0.000485	95
고흥군			56,553	56,553	0.001407	95
보성군			35,263	35,263	0.000877	95
화순군			51,315	51,315	0.001276	114
장흥군			30,953	30,953	0.000770	95
강진군			29,774	29,774	0.000741	95
해남군			56,852	56,852	0.001414	95
영암군			46,927	46,927	0.001167	95
무안군			56,674	56,674	0.001410	95
함평군			27,346	27,346	0.000680	179
영광군			41,590	41,590	0.001034	95
장성군			32,928	32,928	0.000819	78
완도군			40,472	40,472	0.001007	95
진도군			24,830	24,830	0.000618	95
신안군			30,236	30,236	0.000752	95
경북		1,757,286	432,892	2,190,178		4,629
포항시		427,539		427,539	0.010634	121
경주시		215,140		215,140	0.005351	267
김천시		108,863		108,863	0.002708	267
안동시		142,263		142,263	0.003539	267
구미시		320,785		320,785	0.007979	121
영주시		93,056		93,056	0.002315	267
영천시		82,645		82,645	0.002056	431
상주시		85,418		85,418	0.002125	267
문경시		59,827		59,827	0.001488	266

<부록 2: 계속> 2010년 15세 이상 시·군별 인구수, 가중치 및 표본수

행정구역	인구수				가중치	표본수
	특별시/광역시/ 자치도	시	군	계		
해당도시수	8	73	81	162		
전국	19,181,024	17,782,726	3,240,038	40,203,788		36,816
경산시		221,750		221,750	0.005516	142
군위군			18,278	18,278	0.000455	68
의성군			46,934	46,934	0.001167	167
청송군			21,618	21,618	0.000538	166
영양군			14,859	14,859	0.000370	165
영덕군			32,521	32,521	0.000809	166
청도군			34,567	34,567	0.000860	329
고령군			27,173	27,173	0.000676	115
성주군			32,506	32,506	0.000809	167
칠곡군			91,084	91,084	0.002266	209
예천군			38,489	38,489	0.000957	166
봉화군			27,891	27,891	0.000694	166
울진군			40,152	40,152	0.000999	166
울릉군			6,820	6,820	0.000170	163
경남		2,192,390	390,311	2,582,701		3,260
진주시		280,665		280,665	0.006981	227
통영시		106,333		106,333	0.002645	241
사천시		88,564		88,564	0.002203	241
김해시		382,654		382,654	0.009518	252
밀양시		85,583		85,583	0.002129	241
거제시		178,947		178,947	0.004451	242
양산시		202,008		202,008	0.005025	272
창원시		867,636		867,636	0.021581	204
의령군			22,667	22,667	0.000564	134
함안군			49,917	49,917	0.001242	134
창녕군			48,387	48,387	0.001204	134
고성군			44,778	44,778	0.001114	134
남해군			38,944	38,944	0.000969	134
하동군			36,552	36,552	0.000909	134
산청군			28,430	28,430	0.000707	134
함양군			32,887	32,887	0.000818	134
거창군			48,851	48,851	0.001215	134
합천군			38,898	38,898	0.000968	134