

Social Science Research [in Korean]

Principles
Methods and
Practices

Jaeki Song • Miri Kim • Anol Bhattacharjee

SOCIAL SCIENCE RESEARCH:

PRINCIPLES, METHODS, AND PRACTICES [In Korean]

Jaeki Song, Ph.D.

Texas Tech University
Lubbock, Texas, USA
jaeki.song@ttu.edu

Miri Kim, Researcher

Sogang University
Seoul, South Korea
mirikim@sogang.ac.kr

Anol Bhattacharjee, Ph.D.

University of South Florida
Tampa, Florida, USA
abhattach@usf.edu

First Edition

Copyright © 2014 by Jaeki Song·Miri Kim·Anol Bhattacharjee

Published under

Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License



Social Science Research: Principles, Methods, and Practices [In Korean], 1st edition

By Jaeki Song·Miri Kim·Anol Bhattacharjee

First published 2014

ISBN-13: 978-0-615-97043-1

ISBN-10: 0615970435

Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License:

Users are free to use, copy, share, distribute, display, and reference this book under the following conditions:

- **ATTRIBUTION:** Whole or partial use of this book should be attributed (referenced or cited) according to standard academic practices.
- **NON-COMMERCIAL:** This book may not be used for commercial purposes.
- **SHARE ALIKE:** Users may alter, transform, or build upon this book, but must distribute the resulting work under the same or similar license as this one.

For any reuse or distribution, the license terms of this work must be clearly specified. Your fair use and other rights are in no way affected by the above.

Copyright © 2014 by Jaeki Song·Miri Kim·Anol Bhattacharjee

Recommended Citation

J, Song., M, Kim., & Bhattacharjee, A., “Social Science Research: Principles, Methods, and Practices [In Korean]” (2014). KOCW Open Access Textbooks.

<http://www.kocw.net/home/common/contents/document/SocialScienceResearch.pdf>

This Book is brought to you for free and open access by Korea Open Courseware

저자 소개

송재기

Texas Tech University의 Rawls 경영대학의 경영정보전공 교수로 재직 중이며 가장 주된 연구분야는 웹기반 기술 수용 및 혁신과 비즈니스 분석이다. 구체적인 연구 주제로는 IT 서비스 혁신, 정보기술의 확산, 데이터 마이닝과 텍스트 마이닝을 활용한 비즈니스 분석 등이 있다. 미국에 상주하는 한국인 학자 중에서 최고의 학자로 손꼽힐 만큼 왕성한 연구활동을 수행하고 있는 학자 중 한 명이며, 연구 업적은 *Management Science*, *Journal of Management Information Systems*, *Journal of the Association for Information Systems*, *IEEE Transactions on Professional Communication*, *Decision Support Systems*, *Information & Management* 등 다양한 국제 학술지에 40 여편의 논문을 수록하고 있다.

김미리

서강대학교 경영전문대학원의 연구원으로, 인간적 관점에서 혁신을 통한 서비스 시스템 경영에 관한 연구를 관심 있게 수행하고 있다. 주요 연구 주제로는 서비스 혁신 프로세스 개발, 모바일 기술을 이용한 경영혁신분석, 효율성 향상을 위한 혁신적인 의료 및 관광 오픈 시스템 개발, 인간의 가치를 고려한 소셜네트워크 상에서 효과 분석 등이 있다. *Samsung SDS Journal of IT Service* 에 논문을 수록하였고 국내 정부 및 민간 연구과제 프로젝트를 수행하기도 하였다. 2011년 한국 경영정보학회 우수논문상을 수상하였고, 이 후 ICIS, AMCIS, PACIS 등 다양한 국제 학술대회에서 논문을 발표하였다. 또한 최근 연구실적들이 국제 우수 저널들에 채택되는 등 신진학자로서 지속적인 연구활동을 하고 있다.

Anol Bhattacharjee

University of South Florida의 교수이자 Citigroup의 자문위원으로 활동하고 있으며 정보시스템 연구의 세계 10위 안에 손꼽히는 학자이다. 그리고 최근 2001년부터 2010년까지 *MIS Quarterly*와 *Information Systems Research* 의 정보시스템 분야 세계 2대 학술지에 우수 학자 8번째로 연구업적으로 기록하였다. 현재까지 15편동안 연구를 활발히 하고 있으며 50여편의 논문을 수록하고 2권의 책을 편찬하여 구글 스칼라에서 그의 논문과 책이 4,000번 이상 인용되었다. *MIS Quarterly* 의 편집위원으로 4년동안 활용하였으며 2003년에는 연구 공로를 인정받아 총장으로부터 최우수 연구자상을 수상하였다. 지금까지 세계 각국의 대학교들과 연구활동을 하고 있다.

목 차

연구의 소개

제 1 장 과학과 과학적 연구	1
제 2 장 기본 개념 및 사고의 전환	8
제 3 장 과학적 연구과정	16
제 4 장 과학적 연구에 관한 이론	24

경험적 연구의 기초

제 5 장 연구설계	34
제 6 장 구성개념의 측정	41
제 7 장 신뢰성과 타당성	52
제 8 장 표본추출	62

자료수집

제 9 장 조사연구	70
제 10 장 실험연구	82
제 11 장 사례연구	93
제 12 장 해석적 연구	103

자료분석

제 13 장 질적 분석	111
제 14 장 양적 분석-기술적 통계	116
제 15 장 양적 분석-추론적 통계	125

제 1장 과학과 과학적 연구

연구(research)란 무엇인가? 이러한 질문을 누구에게 하느냐에 따라 각기 다른 답변을 얻게 될 것이다. 어떤 사람은 일상에서 자신이 원하는 상품이나 서비스를 구매하기 위해 다양한 방법 가운데 최적의 온라인 웹사이트를 발견하는 것을 연구라 말할 것이다. 텔레비전 뉴스 프로그램에서 다가올 선거 당선자나 정부 지원프로젝트에 대한 국민의 여론을 조사하는 것을 연구라고 부른다. 대학생들은 숙제나 과제를 수행하기 위해 인터넷이나 도서관을 통해 원하는 정보를 찾는 것을 연구라 한다. 대학원생들은 연구를 수행하기 위해 자료를 수집하고 분석하는 것을 연구라 한다. 기업이나 컨설팅 회사의 연구는 기업의 공급망에서 병목점(bottleneck)을 파악하거나 데이터마이닝을 통해 소비자 구매행태를 파악하여 조직의 문제점을 해결하는 것이다. 하지만 위에서 ‘과학적 연구(scientific research)’라고 하기 어렵다. 연구(research)는 과학에 공헌해야 하며 과학적 방법(scientific method)을 따라야 하기 때문에 이번 장에서 ‘과학(science)’과 ‘과학적 방법(scientific research)’의 구성요소 대해 자세히 살펴보겠다.

1.1 과학(Science)

과학이란 무엇인가? 몇몇 사람들은 과학이란 뛰어난 학생들만 관심을 가지는 고등학교나 대학교 수준의 물리학, 화학, 생물학을 지칭한다고 생각한다. 또한, 어떤 사람들은 과학자들이 실험실에서 특수한 장비를 사용한 기술을 과학이라 지칭한다. 근본적으로, ‘과학’이라는 단어는 지식의 의미를 가지고 있는 라틴어 scientia 에서 비롯되었다. **과학(science)**이란 지식의 어떠한 분야이건 ‘과학적 방법(scientific method)’을 이용한 ‘지식체계(body of knowledge)’를 시스템적으로 구축하는 것을 의미한다. 과학은 크게 자연과학과 사회과학의 두 가지 범주로 나뉘어 진다. **자연과학(natural science)**은 자연적으로 형성된 빛, 물체, 물질, 지구, 천체, 인체와 같은 물체나 현상에 관한 지식이다. 뿐만 아니라 자연과학은 물리과학, 지구과학, 생명과학 등 세분화 된다. 물리과학(physical sciences)은 물체에 관한 물리학, 물질에 관한 화학, 천체에 관한 천문학으로 구성되어 있다, 지구과학은 지구에 관한 지리학과 같은 학문으로 구성되어 있다. 생명과학은 인체를 대상으로 하는 생물학과 식물을 대상으로 하는 식물학으로 구성되어 있다. 반면에 **사회과학(social science)**은 사람, 집단, 기업, 사회, 경제와 개별적 집단적 행태에 관한 지식이다. 이와 같이 사회과학은 인간행동을 대상으로 하는 심리학, 집단이나 사회의 행태에 관한 사회학, 기업이나 시장을 대상으로 하는 경제학으로 나눌 수 있다.

자연과학은 사회과학과 다음 몇 가지 점에서 크게 다르다. 자연과학은 정확하고 정밀하며 결정론적 입장을 취하고 있다. 예를 들어, 물리학에서 빛이나 소리, 전자기 등에 대한 과학적 실험은 실험자나 장소, 시간에 상관없이 항상 동일한 결과를 가져온다. 만약 두 학생이 동일한 물리실험을 수행하였는데 서로 다른 결과 값을 도출했다면 그 것을 적어도 한 명의 학생이 실수를 했다는 의미를 가지게 된다. 따라서 자연과학은 실험자의 실수가 없다면 소리의 속도, 빛의 속도나 행성의 궤적은 동일하게 측정될 것이다. 반면에 사회과학은 자연과학에 비해 정확하지 않고 모호하며, 덜 결정론적이다. 예를 들어, 사람의 행복을 가상의 기기를 사용하여 측정한다고 할 때, 그 사람의 행복지수는 날에 따라서, 같은 날이라도 시간에 따라서 다를 수 있다. 사람의 행복은 당일 접하는 뉴스나 당일 오전에 발생한 사건에 따라 다르다. 더욱이 사람의 행복을 정확하게 측정할 수 있는 도구나 계측기준이 없다. 그러므로 동일한 사람이라 하더라도 계측장비에 따라 ‘좀더 행복’ 하거나, ‘덜 행복’ 하게 보일 수 있다. 다시 말해서 사회과학에서는 높은 수준의 측정오류(measurement error)가 존재하며 사회과학 정책결정에 있어서도 상당한 불확실성과 일치된 동의를 적은 편이다. 예를 들어, 자연과학에서는 빛의 속도나 지구의 자전 속도에 대해 과학자들 사이에서는 대부분 동의하고 있다. 하지만 사회과학에서는 많은 사회 과학자들이 테러리즘을 줄이거나 경기침체를 해결하는 방법에 있어 의견을 달리한다. 따라서 사회과학을 연구하기 위해서는 사회과학에 내재되어 있는 모호성, 불확실성, 측정오류를 보다 깊게 연구해야 한다.

1.2 과학적 지식

과학의 목적은 과학적 지식을 생성하는 것이다. **과학적 지식(scientific knowledge)**은 일반화된 법칙이나 이론의 체계를 의미한다. 과학적 방법을 통한 현상이나 흥미로운 행동들을 설명하는 것이다 (과학적 방법은 이후 장에서 자세하게 다룬다). **법칙(law)**은 현상이나 행동들의 관찰된 패턴이라면 **이론(theory)**은 근본적인 현상이나 행동에 대한 체계적 설명이다. 예를 들면, 물리학에서 뉴턴의 운동법칙은 첫째, 제 1 법칙은 관성의 법칙으로 모든 물체의 질량 중심은 그 상태를 바꿀만한 힘이 강제로 주어지지 않는 한, 정지상태를 유지하거나 일정한 운동을 하여 진행방향으로 계속 움직이는 상태를 유지하려는 성질이고, 둘째, 제 2 법칙은 가속도의 법칙으로 일정한 질량을 가진 물체의 가속도는 힘에 비례하려는 성질이고, 셋째, 제 3 법칙은 작용과 반작용의 법칙으로 두 물체가 충돌할 때 어떤 현상이 발생하는지에 관한 것이다. 뉴턴의 세 가지 법칙은 고전역학의 운동물체에 관한 이론을 기초로 구성되어 있다. 마찬가지로 광학은 빛의 특성을 설명하며 빛이 다른 매개체에서 어떻게 반응하는지 설명한다. 전자기 이론은 전기의 특성과 전기가 어떻게 발생하는지를 설명한다. 양자역학은 이원자입자(subatomic particle)의 특성을 설명하고 천문학은 별과 다른 천체의 특성을 설명한다. 열역학은 에너지의 특성과 기계적 작동을 설명한다. 물리학 대부분은 고등학교 교과서나 대학교재에서 대체적으로 이러한 이론들을 각 장으로 나누어서 설명하고 있다. 비슷한 이론은 사회과학에서도 이용 가능하다. 예를 들어, 심리학의 인지부조화이론(cognitive dissonance theory)은 사건에 대한 관찰이 그 사건에 대한 초기 인식과 다를 경우 어떻게 반응하는지를 설명한다. 일반억제이론(general deterrence theory)은 사람들이 부적절하거나 불법적으로 웹사이트에서 음악을 다운로드 하거나 불법적으로 소프트웨어를 복제하는 등의 범죄적 행동을 보이는데 대해 설명하고 있다.

과학은 연구의 목적에 따라 분류될 수 있다. 기초과학 혹은 순수과학은 가장 기본적인 물체, 힘, 그들간의 관계나 그들을 지배하는 법칙에 관한 것으로 물리학, 수학, 생물학이 이에 해당된다. 응용과학 또는 실용과학은 물리현상을 바탕으로 한 기초과학에서 얻은 과학적 지식을 적용하는 과학이다. 공학(engineering)은 주변 환경에서 생선된 문제에 대해 일반성 있는 해결 방법을 제시하는 학문으로 자연 과학적 지식이나 기법을 이용한다. 예를 들어, 물리학이나 화학과 같은 기초과학을 바탕으로 한 응용과학이며 의학은 생물학과 같은 기초과학을 바탕으로 한 응용과학이다. 기초과학이나 응용과학 모두 인간의 삶을 위해 필요하다. 그렇지만 응용과학은 스스로 토대를 세울 수 없고 기초과학의 발전에 따라 함께 발전한다. 예를 들어, 물리학의 전기(electricity)에 대한 이해 없이 공학(engineering)에서 전구를 만들 수 없다. 대체적으로 대학에서는 주로 기초과학의 발전에 역량을 쏟는 반면 산업이나 기업에서는 응용과학에 초점을 맞추는 경향이 있다.

과학적 연구의 목적은 자연이나 사회현상을 설명하는 법칙을 발견하거나 이론을 공준(公準)하는 것이다. 다시 말해서 과학적 지식을 형성하는 것이다. 하지만 사회과학에서의 이러한 지식은 완벽하지 않거나 진실과 다를 수 있음을 이해하는 것은 중요하다. 하나의 보편적인 사실이 존재하지 않을지 모르나 다수의 진실이 평등하게 존재 할지도 모르기 때문이다. 따라서 과학적 지식을 바탕으로 한 이론은 반드시 이해해야 하며 그것은 어떤 특정한 현상을 한 명의 과학자가 제안한 하나의 설명이라는 것이다. 실제 사회현상을 얼마나 적절하게 설명하는지에 따라 이론으로 정립된다. 과학의 과정은 좀 더 정확한 도구와 논리적인 근거를 활용한 관찰을 통해 빈약한 이론으로부터 튼튼한 이론으로 발전시키는 시간의 과정이다.

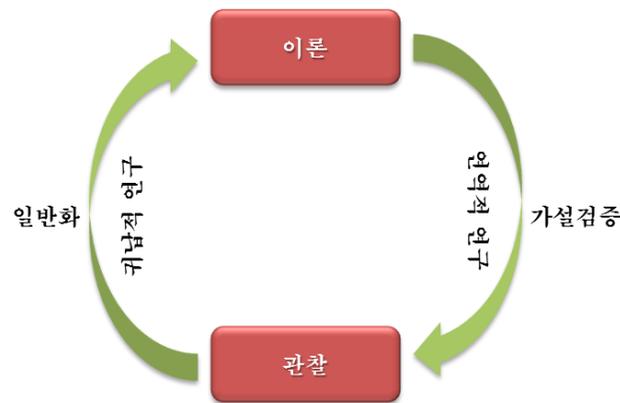
우리는 현상을 설명하는 논리와 관찰의 과정을 통하여 법칙과 이론에 대해서 배웠다. 이론과 관찰은 과학적 지식의 두 가지 근원이다. 과학에서 이론(theory)과 관찰(observation)은 서로 밀접한 관계를 가지고 있으며, 이론 없는 관찰이 있을 수 없으면 관찰 없는 이론이 존재할 수 없다. 이론은 우리가 관찰한 것이 의미 있는지 제공하고 관찰은 기존의 이론을 입증하고 강화시켜주거나 새로운 이론을 정립하는데 도움을 준다.

1.3 과학적 연구

과학의 발전을 위해서 우리는 항상 이론적 측면(theoretical level)과 실증적 측면(empirical level) 사이를 반복해서 발전시켜 나아가야 한다. 이론적인 측면은 자연현상이나 사회현상에 대하여 추상적인 개념(concept)을 발전시키고 그 개념들 사이의 관계를 설명한 것이다(예, ‘이론’ 개발). 반면에 실증적 측면은 이론적 개념과 개념간의 관계를 검증하여 이론들이 얼마나 실제와 부합하고 관찰된 데이터를 바탕으로 현상을 잘 설명해주는지에 검증하는 것이다. 이론(theory)은 사실(truth)이 아니며 과학자들이 주장하는 단지 특정현상에 대한 설명이라는 점을 주지해야 한다. 이론의 적합도(goodness)는 이론이 실제 관찰치와 얼마나 적합한지에 따라 판단된다. 과학적 연구(scientific research)는 이론을 제기하고 이론적 개념들을 검증하고 이론과 실증적 관찰치의 적합성을 조사하여 실제 관찰치를 좀더 잘 설명할 수 있는 수정된 이론을 제시하고 수정된 이론을 재검증하고 다시 관찰을 통해 이론을 향상시키는 순환과정을 거치게 된다. 이와 같이 끊임없이 이론과 관찰을 연구하는 과정을 통해 이론은 보다 세련되고 현실을 잘 설명할 수 있게 된다.

연구자의 과학적 탐구는 귀납법과 연역법의 두 가지 형태를 취할 수 있다. **귀납적 연구(inductive research)**에서는 연구자의 목적이 관찰된 데이터로부터 이론적 개념이나 개념간의 관계 패턴을 추론하는데 있다. **연역적 연구(deductive research)**에서는 연구자의 목적이 이미 알려진 이론적 개념이나 개념간의 관계를 새로운 데이터를 통해 검증하는데 있다. 그러므로 귀납적 연구는 이론생성(theory-building)을 위한 접근방법이라고 볼 수 있고 연역적 연구는 이론검증(theory-testing)을 위한 접근방법이라고 볼 수 있다. 이론검증은 단순히 이론을 검증하는데 목적이 있는 것이 아니라 주어진 이론을 개선하고 향상시키며 확장하는데 있다. 그림 1.1 은 귀납적 연구와 연역적 연구를 연구의 순환과정을 통해 설명하고 있다. 연구의 순환과정에서 귀납적 연구와 연역적 연구는 이론과 데이터를 연결하는 두 개의 중심점에 있다는 것에 주목할 필요가 있다. 만일 연구자가 이론과 관찰된 데이터 모두를 잘 알지 못하면 연구자는 귀납적이든 연역적이든 연구하기 어렵다. 따라서 연구자는 귀납적 연구와 연역적 연구 모두를 할 수 있어야 한다. 연구자는 귀납적으로 이론을 세우고 연역적 방법으로 이론을 검증하여 현상의 관측치를 좀더 잘 설명할 수 있도록 초기 이론을 수정할 수 있어야 한다.

이론형성(귀납법)과 이론검증(연역법)은 과학의 발전에 있어서 매우 중요하다. 아무리 좋게 보이는 이론이라 할지라도 이론이 실제와 부합하지 않다면 이론은 가치가 적다. 이와 마찬가지로 아무리 많은 데이터가 있다할지라도 관찰된 데이터가 이론적 지식을 형성하는데 도움이 되지 않는다면 그러한 데이터는 가치가 적다. 그러므로 그림 1.1 을 순환으로 보기보다 이론과 관찰된 데이터 사이에서 지속적인 반복을 통한 좀더 나은 이론을 도출하는 나선형 관점으로 바라보아야 한다.



<그림 1.1> 연구의 순환과정

특히 사회과학에서 이론형성과 이론검증은 이론적 개념의 부정확성, 개념을 측정하는 부정확한 도구들, 연구자가 관심을 가지는 현상에 영향을 미치는 기타 많은 설명되지 않는 요인들로 인해 특히나 어렵다. 예를 들어, 칼 막스의 공산주의는 효과적인 경제 운용방안으로 수십년간 지속되었다가 결국 자본주의와 견주어 경제성장과 사회복지를 형성시키는데 비효율적인 방안으로 판명되었다. 따라서 중국 같은 사회주의 경제체제도 개별기업의 이윤극대화를 목표로 하는 자본주의 경제체제로 전환되고 있다. 그러나 최근의 미국에서 주택담보대출과 금융산업의 붕괴는 자본주의

또한 문제를 내포하고 있고, 앞서 가정하고 있는 것만큼 성장이나 사회복지를 향상시키는 효과적인 수단이 아님이 판명되었다. 자연과학의 이론과 달리 사회과학에서의 이론은 완전하지 못하며 연구자로 하여금 이론을 발전하고 새로운 대체 이론을 생성할 수 있는 무한한 가능성을 열어놓고 있는 것이다. 그렇지만 일반적으로 귀납적(이론생성) 연구는 기존에 밝혀진 이론이나 설명이 적을 때, 유용하고 연역적(이론검증) 연구는 동일 현상에 대해 많은 경쟁이론이 존재할 경우 유용하다고 알려져 있다. 반면, 많은 경쟁 이론이 존재할 경우, 연구자는 어떤 이론이 어떤 조건하에 가장 적합한지 생각해보아야 한다.

그러므로 앞서 과학적 연구로서 언급한 이론적 측면과 실증적 측면을 각각 수행하기 위해서는 이론적·방법론적 스킬 셋(skill set)이 요구된다. 방법론적 기술(know-how)은 상대적으로 표준화되었고 학문에 걸쳐 크게 다르지 않아 교육을 통해 쉽게 체득할 수 있다. 그러나 이론적 기술(know-what)은 영역(domain)에 따라 다르고 수십 년 간의 관찰과 심사숙고가 필요하며 누가 가르쳐서 배울 수 없고, 경험을 통해 터득하는 수 밖에 없다. 방법론적 기술은 일반적 연구자가 되기 위해 필요한 기술이고 이론적 기술은 우수한 연구자가 되기 위해 필요한 기술이라고 할 수 있다. 관찰된 사건에 대한 설명은 그리 쉽지 않으며 관찰된 데이터를 연결적이고 전반적으로 조감하여 해석할 수 있는 이론적 기술(skill)이 필요하다. 그러기 위해서는 직관, 통찰력)과 개인적 경험이 필요하고 매우 어려운 작업이다. 이러한 것을 수행할 수 있는 연구자들은 해당분야에서 최고의 연구자들과 하겠다. 갈릴레오(Galileo), 뉴턴(Newton), 아인슈타인(Einstein), 닐슨 브르(Niels Bohr), 아담스미스(Adam Smith), 찰스 다윈(Charles Darwin), Herbert Simon(허버트 사이먼)과 같은 위대한 과학자들은 대 이론가(master theoreticians)들이었으며 과학의 발전방향을 전환하게 한 이론들로 인해 오랫동안 기억된다. 그들은 이론은 상정(想定)하고 과학으로 정립하였다.

1.4 과학적 방법

앞에서 우리는 과학이 과학적 방법을 통해 얻은 지식이라고 하였다. 그렇다면 ‘과학적 방법’이란 무엇인가? **과학적 방법(scientific method)**은 현상에 대한 관찰을 정확하게 검증하고, 결과를 해석하며, 도출된 결과를 일반화 할 것 인지와 같은 과학적 지식을 설립하기 위한 기술적인 표준화된 집합체이다. 과학적 방법은 연구자가 독립적이고 편견 없이 이론을 검증하는 것이다. 만일 이론이 관찰된 현실과 동떨어지면 기각하고, 현존하는 이론들을 향상시키거나 수정할 수 있도록 한다. 과학적 방법은 다음 4 가지 특성이 있다.

- **재현가능성(replicable):** 다른 연구자가 독립적으로 과학적 연구를 실행하여 유사하거나 동일한 결과를 얻을 수 있어야 한다.
- **정확성(precise):** 측정하기 어려운 이론적인 개념은 반드시 다른 연구자가 동일한 개념을 측정하기 위하여 그 이론을 정의하고, 검증할 수 있도록 정확하게 사용해야 한다.
- **반증가능성(falsifiable):** 이론은 반드시 반증 가능한 방법으로 제시되어야 한다 검증할 수 없거나 위조된 이론들은 과학적 이론이 아니며 그와 관련된 어떠한 지식도 과학적 지식이 아니다. 그리고 부정확한 용어를 사용하거나 정확하게 측정되지 않은 개념을 사용한 이론은 검증할 수 없기 때문에 과학적이지 않다.

- **간결성(parsimonious):** 하나의 현상을 설명하는 의견이 있을 경우, 과학자들은 가장 심플하며 논리적인 이론을 추구한다. 간결성은 ‘오컴의 면도날(Occam’s razor)’ 이라고 부르는데 이는 불필요한 복잡성을 피하고 가장 간결한 이론을 취해야 한다는 원칙에서 비롯된 것이다. 따라서 과학자들에게 지나치게 많은 개념과 개념간의 관계를 설명하는 이론을 추구하는 것을 방지한다.

과학적 방법을 통해 법칙이나 이론을 검증할 수 없는 탐구들은 ‘과학’이라 할 수 없다. 예를 들어, 신학(종교에 관한 연구)은 신학적 개념(신, 천국, 지옥의 존재)을 독립적인 연구자들이 공식적 방법을 통해 편견없이 검증할 수 없으므로 과학이 아니다. 유사하게 예술, 음악, 문학, 인문학, 법률은 비록 그러한 분야가 창조적이고 추구할 만한 가치가 있다 할지라도 과학으로 간주되지 않는다.

사회과학에 적용된 과학적 방법으로는 질적·양적 자료를 다루는 기법, 통계분석, 실험방법, 설문조사, 사례연구와 같은 전통적으로 계속되어오는 다양한 연구전통(research tradition)이 있다. 이 책의 대부분은 이러한 방법들을 이해하기 위해 저술되었다. 그렇지만 과학적 방법은 연구자가 현상을 어떻게 관찰하고, 어떻게 관찰결과를 해석할 것인가와 같은 실증적 측면의 중요성도 인식해야 한다.

1.5 과학적 연구의 종류

과학자들은 왜 과학적 연구를 수행하는가? 여기에는 탐구(exploration), 기술(description), 설명(explanation)의 세가지 이유가 있다. **탐구적 연구(exploratory research)**는 새로운 분야를 탐구하는 과정에서부터 시작한다. 이 연구는 (1) 특정한 문제의 규모나 범위를 살피거나 (2) 그러한 문제에 대한 시발점(initial idea) 또는 감(hunch)을 얻기 위해서거나 (3) 그러한 문제에 대해 연구의 확장 가능성을 검증하기 위해서이다. 예를 들어, 시민들이 경기침체기간 동안 정부정책에 만족하지 않는지를 검증하려면 탐구적 연구는 시민들의 불만족 정도를 측정하거나 그러한 불만족을 어떻게 설명할 수 있는지(공공시위의 빈도, 뉴스나 출간물에 나타난 의견 등), 불만족에 대한 가상적 원인들(인플레이션, 이자율, 실업, 높은 세율에 대한 비효율적 정부 정책)을 중심으로 이루어져야 한다. 탐구적 연구는 여러 출간물로부터 나오는 국내총생산, 실업률, 소비자지수나 경제학자, 정부관료 등의 전문가 인터뷰 자료, 유사한 문제를 해결하기 위한 기존의 정책수행 자료를 조사하기도 한다. 이러한 연구는 해당문제를 정확하게 이해하는데 큰 도움이 될 수는 없지만 그 자체로 유용하고 가치 있는 활동이다.

두 번째 연구는 **기술적 연구(descriptive research)**이다. 연구자들은 관심 있는 현상을 관찰하고 관찰한 것을 기록하려 한다. 이러한 관찰은 과학적 방법에 의거해 수행되어야 하며 관찰된 결과는 훈련되지 않은 사람이 수행한 것 보다 정확하고 신뢰할 수 있다. 예를 들어, 미국의 인구조사나 노동통계는 미국 인구의 구성이나 가치변화나 미국노동력(장단기 실업, 산업별 고용률, 기업크기 별 고용 등)을 정확하게 기술하는데 있다. 이러한 측정에 있어 유사한 설문항목이 인구조사나 고용설문에 계속해서 사용되고 있다. 만일 설문항목이 다르게 되면 인구구성이나 고용의 이전과 이후를 비교하기 어렵다. 다른 기술적 연구로는 특정주민에 대한 민족지학적 연구가 있다. 이 연구에서는 어떻게 특정주민들이 청소년 범죄율, 도시에서 집단폭력, 남아공화국과 같은 분리주의 국가에서 종교적, 문화적, 윤리적 실천의 변화와 전개와 같은 문제에 당면하여 해결해가게 하는가를 기술하고 중동에서 트위터와 인스턴스 메시지와 같은 기술의 역할을 기술하는 연구들이 있다.

세 번째 연구는 **설명적 연구(explanatory research)**이다. 설명적 연구는 현상이나 문제점, 행동들을 관찰하고 설명하는 것이다. 기술적 연구가 무엇, 어디서, 언제에 관한 질문에 답을 주지만 설명적 연구는 ‘왜’와 ‘어떻게’에 관한 질문의 답을 찾는 것이다. 연구에서 인과관계의 요인을 파악하고 현상으로부터 관찰된 결과를 확인함으로써 이들을 ‘점 연결(connect the dots)’을 통해서 왜 그리고 어떻게 그러한 요인이 관심 있는 현상에 영향을 미치는지 설명한다. 예를 들어, 이러한 연구는 청소년 범죄나 집단폭력의 원인을 분석하고 그러한 사회적 문제를 극복할 수 있는 전략을 제시한다. 박사나 학교 연구자들은 일부 탐색적·기술적 연구를 연구초반에 수행하고 있지만 대부분은 설명연구와 관련되어 있다. 명백한 사건에 대한 관찰을 통한 설명적 연구는 이론적이고 직관적, 통찰력, 연구자의 경험과 같은 해석적인 스킬(skill)이 요구된다.

제 2장 기본 개념 및 사고의 전환

좋은 연구를 위해 선행되어야 하는 조건에는 연구자처럼 생각하고 행동할 수 있는 훈련을 하는 것이다. 이는 연구자의 추상적이며 관심있는 현상을 시각화하는 것을 의미한다. 다시 말해서 눈에 보이지 않는 무형의 개념이나 패턴을 유형화 시키고 이러한 관찰들을 서로 연결시킴으로써 일반적인 법칙이나 이론으로 생성하고 발전시켜 다른 문맥(context)에 접목시키는 것이다. 연구자처럼 생각하는 것은 분석단위(unit of analysis), 구성개념(construct), 가설(hypothesis), 조작화(operationalization), 이론(theory), 모델(model), 귀납적 유추(induction), 연역적 추리(deduction) 등과 같은 과정을 포함하며 본 장에서는 이러한 과정에서 반드시 이해해야 할 연구의 기본개념 및 용어들에 대해 알아본다.

2.1 분석단위(Unit of Analysis)

사회과학에서 연구를 시작하기 전에 첫 번째로 결정되어야 할 사항은 **분석단위(unit of analysis)**이다. 분석단위란 연구조사의 연구 대상(target)을 지칭하는 것으로 전형적으로 사용되는 단위에는 개인, 집단, 조직, 국가, 기술, 물체 등이 있다. 예를 들어, 당신이 사람들의 쇼핑 습관이나 새로운 서비스에 대한 만족도에 대한 관심이 있다면 연구단위는 개인이 되는 것이다. 만약 기업 내에서 프로젝트 팀의 특징이나 프로젝트 성공 요인 등에 관심이 있다면 분석단위는 집단이 되는 것이고, 연구의 목적이 기업이 어떻게 하면 이윤을 더 향상시킬 수 있는가 혹은 좀 더 나은 의사결정을 하는지에 관심이 있다면 분석단위는 기업이 되는 것이다. 여기서 의사결정자들은 개인이지만 회사의 의사결정을 대표하는 사람이기 때문에 개인적인 결정보다는 회사내의 개인이 되는 것이다. 그리고 연구자가 서로 다른 국가들의 문화에 대한 관심을 가지고 있다면 분석단위는 국가가 되는 것이다. 이와 같이 어떠한 객체도 분석의 단위가 될 수가 있다. 웹 페이지(web page)가 분석의 단위가 될 수 있는 경우도 있다. 예를 들어, 연구자의 관심이 어떻게 하면 웹 페이지를 사용자에게 매력적으로 만들 수 있을지 연구한다면 분석단위는 사용자가 아닌 웹 페이지가 되는 것이다.

분석단위를 정확히 이해하는 것은 힘들다. 예를 들어 특정 지역에서 높은 범죄율이 왜 나타나는지 연구하고 싶다면 우리의 분석단위는 범죄나 범인이 아닌 지역이 될 것이다. 왜냐하면 우리의 관심사는 지역이지 범죄가 아니기 때문이다. 반면에 서로 다른 지역에서 살인, 강도와 같은 서로 다른 범죄의 형태를 비교하고자 한다면, 분석단위는 범죄가 된다. 만약 불법적인 행동에 관여하게 된 범죄자들을 연구하고자 한다면 분석단위는 개인-범죄자가 되는 것이다. 그러므로 동일한 연구 분야 내에서의 두 가지 이상 관련된 관심주제가 있다면 서로 다른 분석단위가 도출 될 것이다.

분석단위를 설정하고 이해하는 것은 데이터(data)를 어디에서 누구를 대상으로 수집하는가를 결정하기 때문에 매우 중요한 요소이다. 만약 분석단위가 웹 페이지상에서 무언가를 분석하는 것이라면 데이터는 실제 웹 페이지에서 수집되어야만 한다. 만약 사람들이 웹 페이지를 어떻게 이해하는 등의 설문조사(survey)를 통해서 데이터를 수집해서는 연구 목적과 부합하지 않는다. 기업을 대상으로 하는 연구를 진행하기 위해서는 기업의 크기, 수익, 체계, 기업의 수용 능력과 같은 기업 단위(organizational-level)의 변수들을 측정해야 하는 것이다. 이와 같은 데이터는 기업의 회계 기록이나 최고 경영자의 설문조사를 통해서 수집할 수 있다. 앞서 언급한 최고 경영자는 개인 단위(individual level)로 보이기도 하지만 조직 단위(organizational level)로도 사용 가능하다. 각 조직에는 한 명의 최고 경영자가 있기 때문이다. 이와 같이 몇몇 변수(variable)는 하위 단위(lower level), 혹은 상위 단위(high level)로 사용 가능한 분석단위이다. 예를 들면, 당신이 조직 내의 팀워크에 대한 연구를 한다면, 당신은 서로 다른 팀에 속해있는 개별 팀원들을 상대로 설문조사를 하고 팀 단위의 변수를 사용하기 위하여 그들의 개인적인 수치를 평균화하여 하나의 팀 수치로 변형시킬 것이다. 변수(variable)의 사용에 대하여 살펴보도록 한다.

2.2 개념(Concept), 구성개념(Construct), 그리고 변수(Variable)

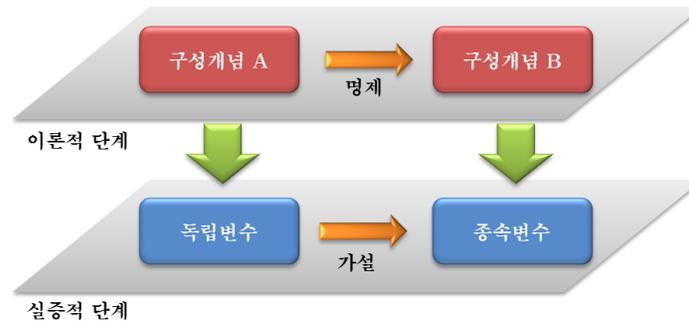
대부분의 연구는 탐구적(exploratory), 기술적(descriptive), 설명적(explanatory)으로 구분할 수 있다고 제 1 장에서 배웠다. 대부분의 연구가 그렇지만, 특히 설명적 연구는 관찰된 자연현상이나 사회현상을 설명하는데 그 목적이 있다. 이를 위해서는 설명적 연구는 **개념(concept)**을 구체화시키고 그것의 특징을 결정 짓고 관계를 형성시킨다. 개념이란 철학적이며 논리적인 관점에서 우리가 경험할 수 있는 사물이나 객체(object)가 가지고 있는 공통적인 성질을 통해 만들어낸 관념으로 우리를 둘러싸고 있는 것들에 대한 사람들 사이에 동의된 생각이다. 좀 더 기술적인 의미로 어떤 객체에 대한 의미를 포함하는 것으로부터 한 단계 높은 추상적인 존재까지를 의미한다. 사람, 회사, 차와 같은 객체들은 개념이 아니지만 그것들과 포괄적이며 공통적인 특성이나 행동과 같은 것을 개념으로 볼 수 있다. 예를 들어 이민에 대한 개인의 태도, 혁신을 위한 기업의 능력, 자동차의 무게와 같은 것이다.

이러한 관점에서 보았을 때 개념을 이해하는 것은 조금 복잡할 수 있다. 우리는 일상에서 우리 자신 스스로가 인지하거나 못한 다양한 종류의 개념들을 사용하고 있다. 어떤 개념은 시간에 걸쳐 오랫동안 발전되고 구체화 되어 언어로 다른 사람들과 공통으로 사용한다. 현상을 설명하는 원리(discipline)나 언어에서 빌려와 새로운 형태로 개념을 구체화 하기도 한다. 예를 들어, 테크노스트레스(technostress)는 새로운 기술을 배워야 하는 사람들이 느끼는 정신적인 스트레스로 새로운 개념인 것이다.

개념은 구체적일 수도 있고 때로는 더 추상적일 수도 있다. 우리가 흔히 말하는 사람의 몸무게, 오늘의 날씨 등은 상대적으로 객관적이며 구체적이다. 하지만 개인의 성향은 상대적으로 추상적인 개념으로 이를 구체화하기 쉽지가 않다. 이와 같이 개념들은 각각 가지고 있는 추상적인 레벨이 존재하게 된다. 이와 같이 추상적인 개념을 설명하기 위해 구성체 또는 **구성개념(construct)**을 사용한다. 다시 말해서, 구성개념은 주어진 현상을 설명하기 위해 선택되거나 창조된 추상적 개념이다. 사람의 몸무게와 같이 간단한 구성개념으로 표현되기도 하고, 어휘력, 문법력, 대화능력 등을 포함한 복합적인 구성개념으로 표현 될 수 도 있다. 전자의 경우를 **단차원 구성개념(uni-dimensional construct)**이라 표현하고 후자의 경우를 **다차원 구성개념(multi-dimensional construct)**라고

표현한다. 다차원 구성개념의 경우에 개념(concept)과 구성개념(construct)의 차이가 더 명확해지는데 이는 상위레벨의 추상화가 구성개념(예, 정보의 질)이고 하위레벨의 추상화는 개념(예, 정확성, 연계성, 이해성)이다. 그러나 이러한 구분은 단차원 구성개념에서는 명확하지 않다.

과학적인 연구에서 구성개념을 이용하기 위해서는 매우 신중한 절차를 걸쳐야 한다. 위에서 살펴본 바와 같이 구성개념은 추상적인 개념이기 때문에 이 개념을 사용하는 사용자들 간에 서로 다른 이해와 개념의 혼돈을 최소화 시켜야 하며 이를 위해서는 구성개념을 정확히 정의하는 것이 중요하다. 예를 들어, 각 개인의 소득을 지칭할 때는 월 소득인지 년 소득을 의미하는지, 세금부과 전의 명목소득인지 세금 부과 후의 실질소득을 의미하는지, 개인 소득인지 가족 소득을 의미하는지를 정확하게 명시해야 한다. 이와 같이 구성개념을 정의하기 위해서는 두 가지 정의가 존재한다. 하나는 사전적 정의(dictionary definitions)이고 또 다른 하나는 **조작적 정의(operational definitions)**이다. 과학적 연구에서는 사전적 정의 보다는 조작적 정의를 더 많이 사용하는데 이는 구성개념이 실제로 어떤 계량적 방법을 통해 측정되는지를 의미하기 때문이다. 예를 들어, 기온이라는 구성개념의 조작적 정의에는 반드시 화씨, 섭씨, 혹은 켈빈 척도를 사용하여 측정할 것인지 포함해야 하며 앞서 언급한 소득 또한 그렇다.

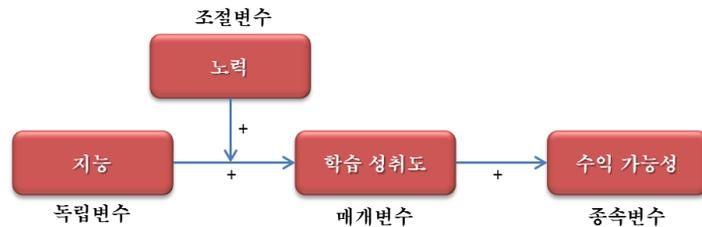


<그림 2.1> 연구에서 이론적·실증적 단계의 관계

이와 같이 구성개념과 조작적 정의를 응용하여 변수(variable)라는 개념과 혼용하여 사용하고 있는 경우가 종종 있다. 어원적으로 변수는 부정적-긍정적, 낮음-높음과 같이 양적으로 다양하지만, 구성개념은 그렇지 않다. 정확히 말해서, **변수(variable)**란 주어진 현상을 설명하기 위해 구체화된 구성개념이 조작적 정의를 통해 측정이 가능한 상태로 된 것을 의미한다. 즉 구성개념은 직접적으로 측정할 수 없기 때문에 대리적인 방법(proxy)을 통해 변수를 통해 측정을 한다. 예를 들어, 사람의 지적능력(intelligence)을 측정하기 위해 IQ 점수를 사용하였다면 지적능력이 구성개념이 되는 것이고 IQ 점수가 변수가 되는 것이다. IQ 점수는 사람의 지적 능력을 측정하는 하나의 방법으로서 사용되기도 하지만 항상 지적 능력을 측정하는 방법이라고는 할 수는 없다. 따라서 변수는 구성개념을 올바르게 설정하기 위한 다양한 측정 방법 중 하나이며 연구 목적과 구성개념에 맞는 측정방법이 사용되어야 하는 것이다. 또 다른 예시는 소비자 만족도를 측정하기 위해 여러 가지 문항을 사용하였다면 소비자 만족도가 구성개념이 되고 그 문항들이 변수가 되는 것이다. 중요한 것은 구성개념을 측정하기 위해 사용된 변수들의 측정이 좋을 수도 또는 상대적으로 좋지 않을 수도 있는 것이지 절대적인 측정 방법은 없다. 따라서 연구자는 좋은 측정방법의 개발을 위해 노력을 해야

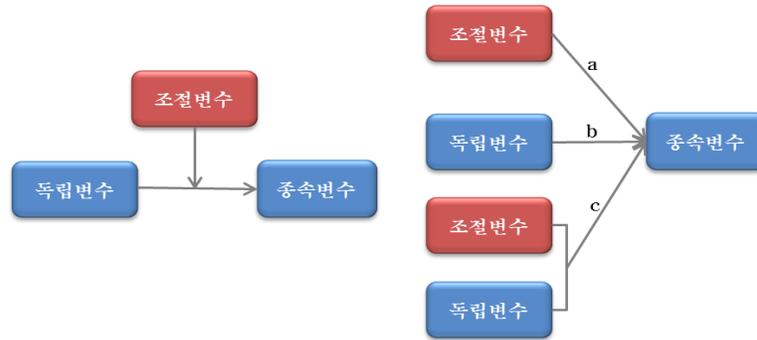
한다. 그림 2.1 에서 보듯이 과학적 연구는 이론적인 단계(theoretical plane)와 실증적인 단계(empirical plane)가 맞물려 진행되어야 한다. 구성개념(construct)은 이론적 단계를 거쳐 개념화(conceptualize) 되고 변수(variable)는 실증적인 단계에서 조작화(operationalize) 과정을 거친다. 이를 통하여 연구자는 본인의 생각을 이론적·실증적 단계를 통해 구체화를 시키는 것이다.

변수를 어떻게 사용하는지에 따라 변수는 독립변수(independent variable), 종속변수(dependent variable), 매개변수(mediating variable), 조절변수(moderating variable), 그리고 통제변수(control variable)로 나눌 수 있다. **독립변수(independent variable)**란 다른 변수를 설명하는 변수를 의미하고, 그 다른 변수는 종속 변수라고 한다. **종속변수(dependent variable)**는 독립 변수들에 의해 설명 되는 변수를 의미한다. **매개변수(mediating variable or intermediate variables)**는 독립변수가 종속변수에 미치는 중간에 개입한 변수이다. **조절변수(moderating variable)**란 독립변수와 종속변수간에 관계의 방향과 강도에 영향을 미치는 변수이다. **통제변수(control variable)**란 연구자가 연구설계에 고정시킴으로써 통제하고자 하는 개입변수를 의미한다. 예를 들어, 높은 지능이 학생들의 학습 성취도를 향상시킨다는 것을 알아보고자 한다면, 지능은 독립변수이고, 학습성취도는 종속변수이다. 하지만 이외에도 종속변수를 설명하지 못하는 관련 없는 변수들이 존재하기도 한다. 그러나 이러한 변수들이 종종 종속변수에 영향을 미치기도 한다. 이러한 변수들은 과학적 연구에서 반드시 통제되어야 하는데 이런 것들을 통제 변수가 되는 것이다.



<그림 2.2> 구성개념간의 관계구조 네트워크

다양한 변수들의 역할의 차이점을 이해하기 위해서 그림 2.2 의 예제를 살펴보자. 만약 지능이 학습성취도에 영향을 미친다고 설정하면 지능은 IQ 점수에 의해 측정된 독립변수, 학습 성취도는 GPA 로 측정된 종속변수가 된다. 또한, 학습성취도가 노력(정의 관계 또는 부의관계)에 따라 다르게 나타날 것이라면 노력은 조절변수가 된다. 학습 성취도가 더 높은 수익 가능성의 중간 단계로서 본다면 수익 가능성은 학습 성취도의 종속 변수가 될 것이며, 학습 성취도는 지능과 수익 가능성 관계에서 매개 변수가 된다. 그러므로 변수는 변수들 사이의 관계에 따라서 독립, 종속, 조절, 매개변수로 정의된다. 특정개념과 관련되는 개념상호간 체계적인 이론적 모델 구축을 **노모로지컬 네트워크(nomological network)**라고 한다. 이처럼 연구자처럼 생각하고 관찰한 것은 현상을 구성개념으로 정립하는 것뿐만 아니라 그들 사이의 네트워크를 시각화하는 것도 일컫는다.



<그림 2.3> 조절변수의 이해

조절변수는 과학적 연구에서 매우 유용하게 이용될 수 있는 변수의 형태이나 매우 해석하기 어렵다. 그림 2.3 에서 볼 수 있듯이 조절변수는 종속변수에 두 번 영향을 미치기 때문이다. 종속변수는 다음 아래와 같이 나타낼 수 있다.

$$\text{종속변수} = a(\text{독립변수}) + b(\text{조절변수}) + c(\text{독립변수} \times \text{조절변수})$$

따라서 연구자는 연구의 목적에 따라 조절변수의 사용은 매우 신중하게 결정하여야 한다.

2.3 명제(Proposition)와 가설(Hypothesis)

그림 2.2 는 지능, 노력, 학업성취도, 그리고 수의 가능성과 같은 이론적 구성개념들이 어떻게 관계구조 네트워크 안에서 서로 연관되어 있는지 보여준다. 과학적 연구에서 각각의 관계들을 명제(proposition)라고 한다. 주어진 현상이나 행동들에 대하여 설명하고자 할 때, 단순히 주요 개념이나 관계들을 나열하는 것이 아니라 구성개념들 간의 관계에 대한 패턴을 구별하고 설명할 수 있어야 한다. 따라서 이러한 관계들의 패턴을 명제라고 하는 것이다. **명제(propotion)**란, 이론에 근거하여 도출되고 이를 사실로 받아들여 질 수 있는 개념들 간의 관계를 임시적으로 추측한 것을 서술한 것이다. 예를 들어, “학생들의 지능이 그들의 학습성취도에 영향을 미친다” 가 있다. 이 서술된 문장을 수집된 데이터를 사용하여 증명할 수 있다. 결과에 따라 명제가 사실인지 아닌지 판단할 수 있는 것이다. 일반적으로 명제는 논리적으로 설명하는 연역적 방법(deduction)이나 실증적 관찰을 통한 귀납적인 방법(induction)을 바탕으로 하고 있다.

명제는 추상적 구성개념으로 이루어져있기 때문에 직접적으로 테스트 할 수 없다. 대신에 그것들을 측정 가능한 변수로 변환시켜 간접적으로 측정한다. 명제의 실증적 작성법은 변수들간의 관계로 설명되는데 이것을 **가설(hypothesis)**이라고 부른다. 앞서 언급한 것처럼 IQ 와 GPA 점수는 지능과 학습성취도를 각각 측정한 변수들인데 위에 제시된 명제는 가설 형태인 “학생의 IQ 점수는 GPA 에 영향을 미친다” 로 변환 가능하다. 명제가 이론적 단계(theoretical plane)이라면 가설을 실증적

단계(empirical plane)이다. 그러므로 가설은 실증적으로 분석이 가능해야 하며 실증분석 결과가 검증되거나 검증되지 않는 결론이 도출되어야 한다.

가설은 강하거나 약할 수 있다. “학생들의 IQ 점수는 학습 성취도와 관련이 있다” 라는 가설은 두 가지의 약점을 가지고 있다. 첫째, 위의 가설에 두 변수간의 관계가 정(+) 의 관계인지 부(-)의 관계인지가 나타나 있지 않다. 두 번째로는 어떤 변수가 독립변수인지 종속변수인지의 인과성이 결여되어 있다. 따라서 좀 더 명확한 가설을 만든다면 “학생들의 IQ 점수는 학업성취도에 긍정적인 영향을 미친다” 라고 설정되어야 한다. 이 가설은 그림 2.2 를 나타낸다.

과학적인 가설에는 독립변수와 종속변수가 명확하게 설정되어야 한다. 위의 예에서 IQ 점수는 독립변수로서 ‘원인(cause)’ 의 역할을 하고 학습 성취도는 종속변수로서 ‘효과(effect)’ 의 역할을 가지고 있다. 나아가 가설은 검증결과가 옳고 그름(true or false)으로 판명되어야 한다. 예를 들면, 높은 지식수준이 높은 학습 성취도를 결정짓는다(true) 또는 높은 지식수준이 학업성취도의 결정에 아무런 역할을 하지 못한다(false)와 같은 결론을 도출하여야 한다. “학생들은 일반적으로 지식이 높다” 라던가 “모든 학생들은 성공적으로 학업을 마친다” 라던지 하는 가설은 어떤 것이 독립변수이고 어떤 것이 종속변수인지가 명확하지 않기 때문이다. 즉 원인과 결과가 명확히 나타나 있지 않으므로 이와 같은 것들은 가설이라고 하기가 어렵다.

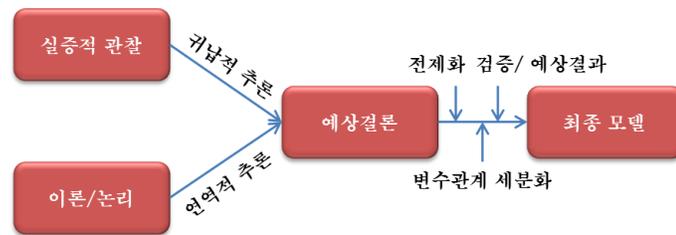
2.4 이론(Theory)과 모델(Model)

이론(theory)은 주어진 가정과 조건하에서 어떤 현상이나 행동을 설명하는 서로 연관된 구성개념과 명제들의 체계적인 집합체이다. 명제는 두·세 가지의 연결된 구성개념들이고 이러한 명제들의 연결된 집합을 이론이라고 한다. 따라서 이론은 많은 구성개념 및 명제들의 체계(system)를 나타내기 때문에 명제나 가설에 비해 보다 넓고 복잡하며 추상적이라고 할 수 있다.

과학적 연구에 익숙하지 않은 사람들은 종종 이론을 추측(speculation)의 관점으로 본다. 그러나 실용적이고 사실적인 것은 과학적으로 이론을 입증하는데 매우 중요한 요인이다. 훌륭한 과학적 이론은 명백한 사실을 사용하여 입증된 것이고 그것은 실제적으로 가치가 있다. 유명한 독일의 심리학자 쾰트 레빈(Kurt Lewin)은 “실제 없는 이론은 없으며 이론 없는 실체는 세울 수 없다” 라고 하여 이론과 실체는 분리될 수 없음을 설명하였다. 그러므로 과학적 연구에서는 이론과 실제 모두 중요한 요소이다.

이론은 사회적 현상이나 자연적 현상을 설명한다. 제 1 장에서도 강조했듯이 좋은 설명과 빈약한 설명이 존재 하듯 좋은 이론과 그렇지 않은 이론도 존재할 수 있다. 따라서 제 3 장에서는 좋은 이론을 어떻게 실제에서 평가할 수 있는지 살펴본다. 다시 한번 언급하자면 연구자는 이론이 절대 ‘진실(truth)’ 이 아니라는 것을 명심해야 한다. 어떠한 이론도 절대적인 것은 없으며 이론은 어떤 한 사람이 제안했다고 반드시 진실로 받아들여지는 것은 아니기 때문이다. 그러므로 과학적 연구는 빈약한 이론을 좀더 설명력을 가지고 있는 이론으로 발전시키는 것이다. 연구자들의 이러한 도전은 현상을 기존보다 좀 더 풍부한 이론으로 정립시키는데 있다.

이론과 함께 자주 쓰이는 용어 중 하나가 모델이다. **모델(model)**은 연구를 위해 구성된 전체 혹은 부분적 시스템이라고 표현된다. 이론의 기본적인 목적이 현상을 설명(explanation)이라면 모델은 현상을 표상화(representation) 하는 것이다. 모델은 의사결정자들이 주어진 투입을 바탕으로 중요한 결정을 하는데 사용하기도 한다. 예를 들어, 마케팅 관리자가 올해 광고 비용을 작년 광고 비용, 판매수익, 시장 성장, 경쟁상품과 관련된 수치를 사용하여 어떤 상품에 어떻게 사용하였는지 알아보려고 할 때 모델을 사용한다. 혹은 일기예보를 할 때 바람의 강도, 방향, 기온, 습도의 수치를 바탕으로 미래 기상 패턴을 예측하는 모델을 사용하기도 한다. 이러한 모델들은 매우 유용하며 모델은 매우 다양한 형태를 취하기도 한다. 대표적으로 수학적 모델, 네트워크 모델, 경로 모델(path model)등이 있다. 또한, 모델은 기술적(descriptive), 예측적(predictive), 규범적(normative)인 형태를 갖는다. 기술모델(descriptive model)은 변수들을 시각화하고 그 변수들의 관계를 시스템 안에서 표현한다. 앞선 광고 지출 모델이 그러하다. 예측모델(predictive model)은 미래의 사건에 대한 예견을 표현하고자 할 때 사용하며, 기상예보가 대표적인 예이다. 규범모델(normative model)은 일반적으로 규범 활동을 표현하는 가이드의 역할로 사용된다.



<그림 2.4> 모델을 만들기 위한 과정

이론과 모델의 발전 과정에는 귀납적(inductive)과 연역적(deductive) 추론을 포함하고 있다. **연역적 추론(deduction)**은 이론적이거나 논리적인 이유를 바탕으로 현상이나 행동에 대해 결론을 이끌어냄으로써 전제(premise)화 하는 것이다. 예를 들면, 어떤 한 은행에서 직원에게 강력한 윤리성을 강요한다(전제 1). ABC는 그 은행의 종업원이다(전제 2). 그렇다면 ABC는 그 윤리성을 따를 것이라고 믿는다(결론). 따라서 연역적 추론은 기본 전제와 근거가 올바르다면 그 결과는 반드시 옳다는 것이다.

귀납적 추론(induction)은 사실과 관찰된 증거를 바탕으로 결론을 그려나가는 과정이다. 예를 들어, 한 회사가 프로모션 캠페인에 많은 비용을 지출했다(관찰 1). 그러나 판매량을 증가하지 않았다(관찰 2). 따라서 프로모션 캠페인은 잘못되었을 가능성이 있다(결론). 그러나 여기서 낮은 판매량을 경제침체나 경쟁상품의 등장, 자사 브랜드 파워, 공급망 문제와 같은 다른 변수들이 존재한다. 따라서 귀납적 추론을 통한 결론은 오직 가설이며, 반증될 수 있다. 일반적으로 연역적 방법에서의 결론이 귀납적 방법의 결론보다 강하지만 연역적 방법의 기본명제가 잘못되었다면 그 결과는 반드시 잘못된 결론을 갖는다는 의미를 포함하는 점을 가지고 있다는 것을 명심해야 한다.

이론과 모델 설정의 과정에서 귀납적 추론과 연역적 추론에 대하여 보여주고 있다(그림 2.4). 귀납적 추론은 우리가 어떤 현상을 관찰한 후 “왜 이런 일이 벌어졌는가?” 라는 질문으로 시작한다. 이에

대한 답을 도출하기 위하여 우리는 가설을 설립하게 된다. 그리고 나서 우리는 결론을 도출하기 위해 기존 이론으로부터 현상을 잘 설명할 수 있는 로직과 근거 있는 전제를 바탕으로 연역적 추론을 통해 설명한다. 이와 같이 연구자는 귀납적 추론과 연역적 추론을 적절히 사용하여 현상을 가장 잘 설명할 수 있는 방법을 모색하여야 한다.

제 3 장 과학적 연구과정

과학적 연구는 과학적 방법을 통해 과학적 지식을 얻는 과정이라고 제 1 장에서 살펴 보았다. 그렇다면 어떻게 연구를 수행할 것인가? 이 장에서 우리는 과학적 연구과정, 연구과정에 담긴 가정(assumption)과 각 과정에서 기대되는 결과(outcomes)를 다룰 것이다.

3.1 사회연구의 패러다임

연구를 계획하고 수행하는 것은 연구자의 멘탈 모델이나 사고하고 관찰하는 틀에 의해 형성된다. 이러한 멘탈 모델이나 프레임을 **패러다임(paradigm)**이라 부른다. 패러다임에 대한 개념은 토마스 쿤(Tomas Kuhn, 1970)이 과학적 혁명의 구조(The structure of scientific revolution) 책에서 주장한 것으로 그는 자연과학의 역사를 통해 패러다임의 변화를 살펴보았다. 패러다임 개념은 사회과학에서도 동일하게 적용할 수 있다. 특히 동일한 사회현상이라 할지라도 사람들마다 각기 다르게 인식할 수 있어서 생각과 추리과정 또한 달라질 수 있기 때문이다. 예를 들어 보수주의자와 자유주의자는 국민의 삶에 대해서 다른 인식을 하는 경향이 있으며 따라서 사회문제를 어떻게 해결할 것인가에 대해서도 다른 의견을 가지고 있다. 보수주의자들은 세금감면이 국민들의 소비능력을 확대시켜 침체된 경기를 촉진시키는데 적당한 방안이라고 보고 있는 반면 자유주의자들은 정부가 보다 많은 가처분 소득 늘리기 위해 고용창출이나 공공사업에 직접적으로 투자해야 한다고 보고 있다. 이와 마찬가지로, 서구 사회는 개인의 권리, 즉 사생활 권리, 언론의 자유에 대한 권리, 총기소유의 권리와 같은 권리를 중시하고 있다. 반면 아시아 사회는 개인의 권리가 가족, 조직, 국가의 권리와 균형을 갖는 경향이 있어 정부정책들은 좀더 공공적인 성격을 갖는다. 이러한 관점의 차이로 인해 서구 국가들은 아시아 국가들의 정책이 인간의 존엄성을 많이 해친다고 보고 있으며 반대로 아시아 국가는 서구 사회가 개인 탐욕적이고 높은 범죄율을 가지고 있으며 개인주의를 지나치게 추종한다고 보고 있다. 따라서 우리가 가지고 있는 패러다임은 '색안경'같은 것이어서 어떻게 세상을 바라 보는지, 어떻게 세상을 연구해야 하는지, 그리고 어떻게 생각과 관찰을 구상해야 하는지에 따라 결정된다.

패러다임은 암묵적이고 내재적이며 실증적으로 검증된 것이 아니기 때문에 인식하기 매우 어렵다. 그러나 패러다임을 인식하는 것은 동일한 사회현상에 대해 연구자들간의 인식의 차이를 조정할 수 있는 열쇠가 된다. 예를 들어, 자유주의자들은 중등교육을 향상시키는 최적의 방법이 많은 교사를 고용하는 것이라고 보고 있으나, 보수주의자들은 학교 선택권과 같은 수단으로 교육을 민영화하는 것이 더 효과적인 방법이라고 보고 있다. 그렇게 보는 시각에는 보수주의자들은 학교들 간의 자유경쟁이 현재 약화된 학교시스템을 향상시키기 위한 해결책으로 경쟁시장이 더 큰 역할을 할 것

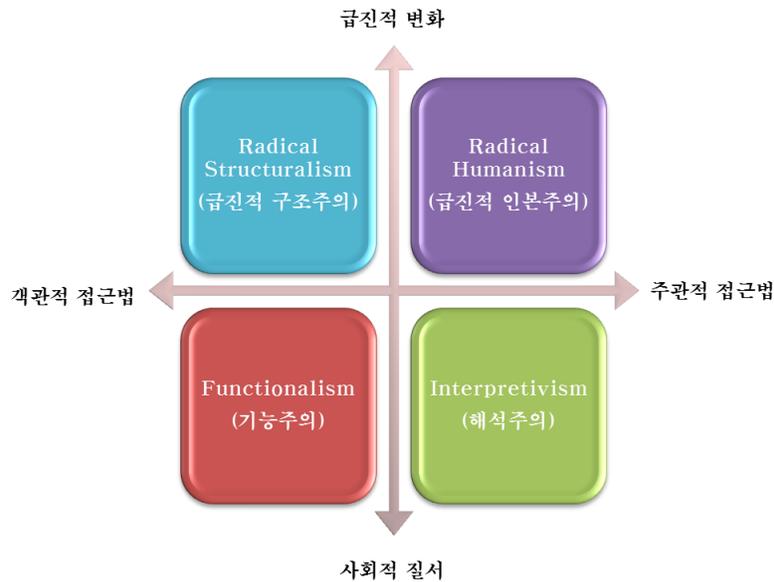
이라고 믿고 있으며, 이에 반해 자유주의자들은 좀더 많은 훈련된 교사들을 확보하는 것이 문제를 해결하는 것이라고 믿고 있는 것이다. 사회과학 연구에 있어서 만일 연구자가 어떤 기술이 어떤 조직에서는 성공적으로 시행되고 어떤 조직에서는 실패하는지 살펴보고자 할 때, 세상을 ‘합리적 시각(rational lens)’으로 바라보는 연구자는 조직 내의 부적절한 기술의 도입이나 기술과 작업환경과의 부조화와 같은 문제 등의 합리적인 설명을 찾으려고 한다. 반면 동일한 문제에 대해 ‘사회적 시각(social lens)’으로 보는 연구자는 사용자의 훈련 부적절함이나 경영지원의 부재와 같은 사회적 결핍에서 원인을 찾으려고 한다. 그리고 ‘정치적 시각(political lens)’으로 보는 연구자는 기술 실행과정의 근간이 되는 조직적 정치(organizational politics)의 이유를 찾으려고 한다. 그러므로 연구자들의 사회현상을 바라보는 패러다임은 연구자들이 측정하고자 하는 개념이나 관찰, 문제에 대한 사후설명을 제한하기도 한다. 그러나 복잡한 사회현상을 고려해 볼 때, 위와 같은 각각의 패러다임은 현상에 대한 부분적인 설명은 가능하지만 완전히 문제를 이해하기 위해서는 다양한 패러다임의 이해와 적용이 필요하다.

사회과학자들은 실증주의와 포스트실증주의 두 가지 대중적인 패러다임을 통하여 이론을 설명하고 있다. **실증주의(positivism)**는 프랑스 철학자 콩트(Auguste Comte, 1798-1857)의 연구를 바탕으로 한 것으로, 자연과학의 방법론을 사회과학에 확장시키는 것을 기반으로 20 세기 중반까지 지배적인 과학 패러다임이었다. 실증주의는 과학이나 지식의 창출이 관찰하고 측정할 수 있는 것으로 한정해야 하고 직접적으로 검증할 수 있는 이론을 바탕으로 해야 한다고 보고 있다. 즉 지식은 경험과 실증적 검증에 기반을 둔 것이라고 보는 인식론 관점의 과학 철학이다. 실증주의 초기에는 과학적 탐구를 종교로부터 분리하고 경험주의(empiricism)를 맹목적으로 신뢰하게 하게 됨으로써, 관찰과 측정이 과학적 탐구의 핵심이 되었다. 따라서 관찰된 사실이 아닌 확대와 추론에 대한 어떤 시도도 거부되었다. 그러나 이와 같은 실증주의에 대한 회의가 20 세기 중반부터 시작되어 그 이후에는 **후기실증주의(post-positivism)**라는 새로운 패러다임이 등장 하였다. 후기실증주의는 보편적인 법칙을 먼저 구성하고 논리적 추론을 통해 실증적 현상을 설명하는데 노력했다. 후기 실증주의자들은 과학을 확실한 것이 아니라 확률적인 관점에서 바라보고 사회적 현상을 좀더 잘 이해하기 위해 탐구하였다. 후기실증주의자는 우리가 보고 있는 세상은 우리의 마음에 의해 주관적으로 생성되었다고 보는 관점에서 주관주의자(subjectivists)와 외부적 실체가 개인의 사고와 독립적으로 존재하지만 확실하게 우리가 그러한 실체를 알 수 없다는 비판적 실재론자(critical realist)로 구분하고 있다.

브루넬과 모건(Burrell and Morgan, 1979)은 사회학적 패러다임과 조직 분석(Sociological Paradigms and Organizational Analysis)이란 책에서 사회과학자들이 사회현상을 인식하고 연구하는 방법은 연구자의 두 가지 근본적인 철학적 가정, 즉 존재론적 가정과 인식론적 가정에 의해 형성된다고 보았다. **존재론적 가정(ontological assumption)**은 우리가 어떻게 세상을 보는가에 대한 가정이다. 예를 들어, 사회적 질서나 부단한 변화와 같은 것이다. **인식론적 가정(epistemological assumption)**은 세상을 연구하는 가장 좋은 방법에 대한 가정으로 예를 들어, 객관적 접근 또는 주관적 접근 같은 것이다. 이와 같은 존재론적, 인식론적 두 가지 가정을 활용하여 우리는 사회과학 연구를 그림 3.1 에 나타난 것처럼 네 가지로 분류하였다.

만일 연구자가 세상을 사회적 질서로 구성되어 있다고 보고(존재론적 접근), 질서화된 사건이나 행동의 패턴을 찾고자 할 때, 연구의 가장 적합한 방법이 객관적 접근(인식론적 접근)이라고 생각한다. 이러한 연구자들은 설문조사와 같이 표준화된 데이터를 수집하여 관찰하고 해석하는

기능주의적(functionalism) 패러다임을 취한 것이다. 그러나, 만일 연구자가 사회적 질서를 연구하는 최적의 방법이 사회적 실체와 관련되어 있는 여러 행위자들의 주관적 해석을 통해서 이루어져야 한다고 보다면 이는 해석주의적(interpretivism) 패러다임을 바탕으로 한 것이다. 반대로 연구자가 세상은 급진적 변화로 구성되어 있고 객관적 접근법을 이용하여 변화를 이해하거나 규정한다면, 급진적 구조주의(radical structuralism) 패러다임을 취하는 것이다. 연구자가 행위자들의 주관적 관점을 이용하여 사회변화를 이해하려고 한다면, 급진적 인본주의(radical humanism) 패러다임을 취하는 것이다.



〈그림 3.1〉 사회과학 연구의 네 가지 패러다임
(출처: Burrell and Morgan, 1979)

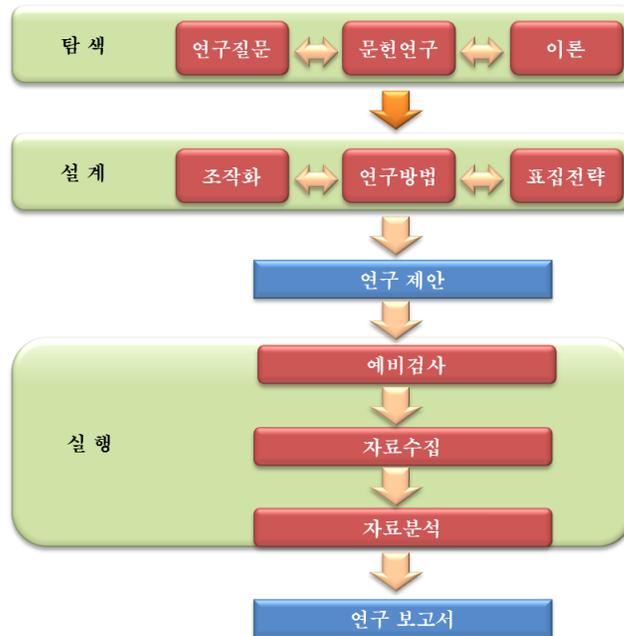
사회과학 연구의 대부분은 자연과학과 기능적 패러다임을 따르고 있다. 기능주의자들은 사회적 질서나 패턴은 기능적 구성요소를 바탕으로 이해할 수 있다고 보았다. 따라서 사회문제를 작은 부분으로 나누고 그러한 작은 문제를 하나 혹은 그 이상의 구성요소로 구분 짓고 설문조사나 실험연구와 같은 객관적 방법을 사용하여 연구할 수 있다고 믿었다. 그러나 후기 실증주의의 출현으로 사회과학 연구자들은 점점 사회적 질서를 인터뷰나 인류학적(ethnographic) 연구를 통해 이해하였다. 이것은 사회연구에 있어서 급진적이거나 특이하거나 변화무쌍한 사건 대신에 일반화할 수 있는 행동, 사건, 현상에 대한 패턴을 이해하는데 주로 관심이 있기 때문에 급진적 인본주의나 급진적 구조주의는 극히 소수의 연구자들이 속해있었다. 그럼에도 불구하고 당신이 왜 자유주의가 중동지역에서 급격하게 일어나고 있는지에 대한 연구, 혹은 그러한 변화가 튀니시아(Tunisia)에서는 성공하고 시리아(Syria)에서는 아직 성공하지 못했는지 연구하고자 한다면, 아마 급진적 인본주의의 관점이 적합할 것이다. 사회현상과 조직현상은 질서와 변화의 두 가지 요소로 구성되어 있다. 예를 들어, 조직의 성공은 기업 환경에 있어 통제되지 않고 항시 변화는 경쟁자, 경쟁 상품, 공급자, 소비자뿐만 아니라 명확하게 정의되고 공식화된 비즈니스 프로세스, 작업 프로세스와 작업책임에

의해서도 결정된다. 그렇기 때문에, 어떤 조직이 다른 조직보다 왜 성공적인가와 같은 연구는 사회현상에 대해 포괄적이고 좀더 완벽한 이해를 위해서 연구의 다양한 패러다임을 인정하고 적용해야 한다.

3.2 연구과정

과학적 연구는 일반적으로 관찰, 이론적 설명, 검증의 사이클을 통한 과정을 거친다. **관찰(observation)** 단계에서는 자연이나 사회 현상, 사건, 연구자의 관심 있는 행동들을 관찰한다. 이론적 **설명(rationalization)** 단계에서는 연구자가 관찰한 서로 다른 조각들을 합리적으로 연결시키는 과정이며 이것들의 연결은 이론으로 만들어 질 수 있다. 마지막으로 **검증(validation)** 단계에서는 자료 수집과 분석의 과정을 과학적 방법을 사용하여 검사한다. 연구자가 관찰한 것들을 이론적으로 설명하는지(귀납적 연구)와 이론을 검증하는지(연역적 연구)에 따라 연구 설계(research design)를 구분하고 관찰-이론적 설명-검증의 사이클은 귀납적-연역적 연구의 사이클과 매우 비슷하다.

대부분의 전통적인 연구방법은 연역적이고 기능적인 경향이 있다. 연구과정을 그림 3.2가 도식화하여 보여주고 있다. 연구를 수행하는 과정은 탐색, 설계, 실행의 세 범주로 나누어 수행하게 되는데 일반적인 연구 설계가 모두 이와 같은 지침이나 과정을 따르지는 않는다. 연구의 과정은 연구 주제에 대한 목적과 필요에 따라 수정되어야 한다.



<그림 3.2> 기능적 연구방법

연구의 첫 번째 단계는 연구자가 관심 있는 특별한 행동, 사건, 현상에 대해 하나 이상의 **연구질문(research question)**을 설정하는 것이다. 연구질문은 연구자가 연구를 통해 알고 알고 싶은 현상이나, 사건, 행동과 같은 세부 주제를 의미한다. 예를 들면, 온라인 상에서 고객이 상품이나

서비스를 직접 보거나 경험하지 않은 상태에서 구매하게 하는 요인은 무엇인가? 또는 어떠한 교육 프로그램이 고등학교 학생들의 창의성을 증진시키는데 도움이 되는가? 또는 어떤 사람들은 왜 폭력적인 행동을 저지르는가? 등이 있다. 연구질문은 무엇, 왜, 어떻게, 언제, 등의 문제를 탐구할 수 있도록 한다. 연구질문의 범위를 넓게 설정하는 것은 연구의 흥미를 이끌어내는데 충분하다. 예를 들어, 회사의 혁신은 어떻게 이루어지는가? 는 중국 회사는 서비스 분야에서 혁신을 어떻게 하는가? 라는 연구질문보다 더 흥미롭다. 반면, 제한적이고 협소한 연구질문(예/아니오와 같은 대답)은 비교적 덜 유용하고 흥미를 끌지 못한다. 따라서 폭넓은 주제는 사회현상의 미묘한 뉘앙스(nuances)을 찾아내는데 유용하다. 흥미롭던 흥미롭지 않던 모든 연구질문은 미리 선정된 연구주제를 바탕으로 시작되어야 한다. 상대적으로 흥미로운 질문들의 조건으로는 많은 사람들로부터 관심을 가질만하고, 가상적 문제가 아닌 실질적 문제를 다루고 있고, 연구질문에 대한 결과가 당연한 것이 아닌 여러 요소에 따라 달라지는 연구주제들이다. 연구질문의 예로는 어떤 사회적 현상의 (a) 결정요인과 (b) 결과는 무엇인가? 이러한 연구질문은 실질적 관찰이나 이전 연구문헌에서 규명되지 않은 것을 바탕으로 하며 충분히 관심 가질 만한 주제이어야 한다.

두 번째 단계는 관심연구 분야에 대한 **문헌연구(literature review)**를 수행하는 것이다. 문헌연구의 주목적은 세 가지로 구분된다. (1) 관심분야에 대한 지식의 현재 상태를 조사하기 위해서, (2) 연구의 주요 저자, 논문, 이론, 결과를 알기 위하여, (3) 연구자의 지식과 문헌 연구의 차이를 인식하기 위하여 수행한다. 오늘날 문헌연구는 온라인 데이터베이스의 키워드 검색으로 간편하게 이루어진다. 연구자라면 반드시 키워드를 통한 관련된 논문들을 살펴보아야 한다. 그리고 최소한 초록 및 주요 내용을 숙지해 두어야 하고 철저하게 이루어져야 하며 두·세 개의 학술지나 특정 방법론, 최근의 자료에만 의존해서는 안된다. 문헌연구는 표를 활용하여 정리할 수 있고, 개념도(concept matrix)와 같은 프레임워크(framework)를 사용하여 정리할 수 있다. 제대로 수행된 문헌연구는 초기의 연구자의 연구질문이 문헌 연구를 통해 이미 수행되어 검증되었거나, 새롭고 흥미로운 연구주제일 수 있다. 따라서 이러한 과정을 통해 연구자는 자신이 가지고 있는 연구질문에 대한 통찰력이나 결과를 예측할 수 있다.

기능주의 연구(연역적 연구)는 이론-검증의 과정을 세 번째 단계에서 포함하고 있다. 연구질문을 설명하기 위해 하나 이상의 이론이 설정되어야 한다. 문헌조사는 관심 있는 현상과 잠재적으로 관련된 여러 개념들을 파악할 수 있게 하는 반면, **이론(theory)**은 파악된 개념이 연구하고자 하는 현상과 논리적으로 관련이 있는지 파악할 수 있도록 한다. 이론이 현상과 관련이 없거나, 일부만 관련이 있거나, 관련 없는 개념들을 측정하게 하기도 하고 의미 있는 결과를 우연히 얻을 가능성을 낮게 하기도 한다. 모든 이론이 모든 사회적 상황을 연구하는데 적합한 것은 아니기 때문에 선정한 연구질문을 설명할 수 있어야 하고 가정이 충족되도록 신중하게 선택해야 한다.

네 번째 단계는 **연구설계(research design)**이다. 연구질문에 대한 만족할 만한 결과를 얻기 위하여 수행해야 하는 활동들을 청사진화(blueprint)하는 것이다. 본 단계에서는 연구방법, 구성개념의 조작화, 데이터의 표집전략을 세운다.

다음 단계는 이론으로부터 검증된 연구질문에 대한 예측 가능한 결과와 관련된 추상적인 개념을 개념화(conceptualize)하고 **조작화(operationalize)**하는 것이다. 개념화는 개념들과 변수들의 의미를 정의하고 규정하는 것을 의미한다. 조작화는 이론적으로 추상화 된 개념들을 어떻게 정확하게 측정할

지 규정하는 것이다. 예를 들어, 편견(prejudice), 소외(alienation), 자유주의(liberalism)과 같은 사회과학의 개념은 정의하기 어려울 뿐만 아니라 측정은 더욱 어렵다. 조작화는 구성개념의 ‘조작적 정의(operational definition)’ 부터 시작한다. 그리고 연구자는 문헌연구를 통해 기존의 측정방법을 살펴보고 자신의 연구에 적합하게 수정하여 사용한다. 만약 문헌을 통한 측정이 존재하지 않거나 빈약하다면 연구자는 구성개념을 최대한 설명할 수 있는 측정도구를 개발할 수 있다. 이것은 몇 개의 설문항목을 사용할 것인지, 어떤 설문항목을 사용할 것인지를 정확하게 명시해야 한다.

연구자는 연구질문에 적합한 자료수집을 위해 어떤 **연구방법(research method)**을 사용할지 고민해야 한다. 실험(experiment)이나 설문조사(survey)와 같은 정량적 방법(quantitative method)과 사례연구(case research), 현장연구(action research)와 같은 정성적 방법(qualitative method)을 사용하거나, 두 가지 방법 모두 사용할 수 있다. 만약 연구자가 실험연구방법을 선택하였다면 실험 설계는 어떻게 할 것인가? 설문조사를 수행한다면 우편설문, 전화설문, 웹 설문을 할 것인가? 와 같은 구체적인 계획을 세워야 한다. 복잡하고 불확실한 사회현상을 파악하기 위해서는 복수방법 접근법(multi-method approaches)이 적합할 것이다. 복수방법 접근법은 각 방법의 장점을 취할 수 있고 단일 방법으로는 얻을 수 없는 통찰력을 제공한다.

연구자는 연구하고자 하는 모집단의 대상을 선정하고 모집단에서 **표집(sampling)**을 어떻게 추출할지에 대한 전략을 세워야 한다. 연구자의 편의에 의한 임의적인 표집이 되지 않도록 주의해야 한다. 예를 들어, 연구자가 회사의 개인, 팀, 부서 단위로 설문조사를 해야 한다면, 연구자는 관심 대상을 설정 한 뒤 자료를 수집해야 한다. 뿐만 아니라 표집은 연구질문에 대한 분석단위(unit of analysis)와 매우 관련이 깊다.

어떤 대상을 조사할 것인가(대상), 무엇을 조사할 것인가(개념), 어떻게 데이터를 수집할 것인가(연구방법)를 결정한 이후에 연구자는 **연구수행(research execution)** 단계를 진행한다. 예비 검사, 자료수집, 그리고 분석이 본 연구 수행 단계에 포함된다.

예비검사(pilot testing)는 연구 진행과정에서 매우 중요하다. 그것은 당신의 연구설계에 대한 잠재적인 문제들을 감소시키는데 도움을 줄 뿐만 아니라 구성개념의 신뢰성(reliability)과 타당성(validity)을 갖춘 측정도구의 사용을 확인시켜 준다. 예비검사의 표집은 일반적으로 모집단의 작은 부분이 사용된다.

예비검사가 성공적으로 완료되면, 연구자는 다음 단계로 **자료수집(data collection)**을 진행한다. 수집된 데이터는 연구자가 설계한 연구방법에 따라 양적 혹은 질적 자료가 된다. 자료수집은 연구수행에 있어서 가장 지엽적인 부분이다. 따라서 지엽적인 데이터를 바탕으로 일반화 하는 것은 부분적이고 불완전할 수 있기 때문에 실증적 일반화(empirical generalization) 하는데 주의해야 한다. 만약 잘못된 데이터를 수집하게 되면 그릇된 결론에 다다를 수 있기 때문에 자료를 수집하기 이전에는 어떤 개념(concept)을 측정하는지, 그러한 개념을 어떻게 조작화 하는지에 대해 상당한 주의를 기울여야 한다. 만일 연구자가 정교한 계획 없이 바로 자료를 수집하면 연구질문에 대한 해답과 관련이 없거나 불완전하거나 또는 쓸모 없게 될 수 있으며 자료수집에 소요되었던 노력을 물거품이 되고 말 것이다.

연구자는 처음에 제시하였던 연구질문에 대한 결론을 도출하기 위해 수집된 데이터를 활용하여 분석 또는 해석을 한다. 양적 또는 질적 자료에 따른 **자료분석(data analysis)**은 회귀분석 또는 구조방정식과 같은 양적 분석(quantitative analysis)이나 코딩 또는 콘텐츠 분석과 같은 질적 분석(qualitative analysis)으로 구분 지어 수행하게 된다.

최종단계에서 연구자는 연구의 전체과정을 연구논문, 연구서(monograph), 또는 보고서 형태로 **연구 보고서(research report)**를 작성한다. 이러한 과정에는 연구질문, 연구동기, 문헌연구, 이론, 개념, 조작화, 연구방법, 표집, 자료수집, 분석, 결론 등에 있어서 왜 그러한 선택을 하였는지 적절하게 피력해야 한다. 연구과정은 다른 연구자가 반복하여 수행할 수 있어야 하며 연구결과를 검증하거나 연구의 공헌을 평가할 수 있도록 상세하게 기술하여야 한다. 다른 연구자들을 위한 문서화가 되지 않는다면 그러한 연구는 향후 가치가 없게 될 것이다. 따라서 문서작성은 과학의 점진적 진화를 위해 반드시 필요한 요소이다.

3.3 일반적인 연구논문 형태

I. 서론

- 연구주제의 현황이 무엇이며 이 논문에서 연구주제가 왜 중요한가?
- 특별히 검증하려는 목적이 무엇인가?
- 이 논문만의 특별한 장점(unique selling point)은 무엇인가?
- 논문의 구조 (roadmap for the paper)는 어떻게 이루어지는가?

II. 문헌연구, 이론적 배경, 연구 모형

- 연구주제에 대해 이미 선행된 논문들과 결과는 무엇인가?
- 선행논문들의 단점은 무엇인가 또는 선행논문에서 검증되지 않은 내용들은 무엇이고 따라서 본 논문에서 무엇을 검증하려 하는가?

III. 가설 설정

- 가설 설정의 이론적 근거
- 가설설정

IV. 연구조사 방법

- 표집
- 변수
- 측정방법
- 연구설계

V. 결론

- 통계적 검증 결과

VI. 토의 및 결론

- 중요한 검증결과 및 활용방안
 - 이론적 측면에서의 활용방안
 - 실무적 측면에서의 활용방안
 - 방법론 측면에서의 활용방안
- 한계
- 미래 연구 방안
- 결론

부록 <if needed>

참고문헌

제 4 장 과학적 연구에 관한 이론

앞 장에서 살펴 보았듯, 과학이란 일련의 과학적 방법론을 통해 실증적으로 검증된 ‘이론(theory)’의 집합이 지식의 형태로 나타난 것을 말한다. 이번 장에서 우리는 이론이 무엇이며, 연구를 함에 있어 이론이 왜 필요한지, 이론을 이루는 바탕은 무엇인지, 또 이론을 어떻게 평가하고 연구에 적용시킬 것인지를 살펴보고자 한다. 끝으로 사회과학에서 자주 사용되는 다섯 가지 이론에 대해서 살펴볼 것이다.

4.1 이론(Theory)

이론은 자연적 혹은 사회적 사건 및 현상을 설명한 것이다. 학술적으로, 과학적 이론은 구성개념들 간의 집합체이며, 그들의 관계를 규명 짓는 명제들로 이루어져 있다. 이는 현상을 논리정연하고 시스템적이며 일관성 있게 설명할 수 있도록 하는 것을 말한다(Bacharach 1989)¹.

사건의 논리적 개연성에 대한 설명 없이 몇몇의 단순한 예측치를 근거로 현상에 대한 결과의 예측이 가능하기도 하다. 그러나 이론은 일정 사건에 대하여 단순히 설명하거나 예측을 할 뿐 아니라, 왜 해당 사건이 일어났는지 논리적으로 설명할 수 있다. 예를 들어, 자본시장 분석가들은 주요 기업들의 자본 현황 보고서나 정부기관에서 발표한 새로운 데이터 등을 활용하여 ‘관련성 있는(correlated)’ 관찰을 통해 주식시장 추세를 예측 할 수 있다. 예측은 이와 같이 단순히 연관성(correlations)만을 요구한다. 하지만 사건에 대한 설명은 원인과 결과 관계(cause-effect relationship)에 대한 이해가 바탕이 되어야 한다. 원인과 결과를 이해하기 위해서 세 가지 조건을 충족시켜야 한다. (1) 구성개념 간의 상관관계 (2) 사건의 진행(원인은 반드시 결과 이전에 일어나야 한다.) (3) 실증적인 검증을 통한 대체 가설의 기각이 있다. 과학적 이론은 방법론을 통한 실증적인 검증으로 입증 될 수 있다는 데에 이론적, 철학적, 또는 다른 형태의 그것과 구분 된다.

설명(explanations)은 개별적(idiographic) 설명과 보편적(nomothetic) 설명으로 나뉜다. **개별적 설명(idiographic explanation)**이란 개별적 상황이나 사건에 대하여 자세하게 기술 하는 것을 말한다. 예를 들어, 여러분이 시험을 잘 치르지 못했다고 한다면 그 이유는 다음과 같이 여러 가지로 나타날 것이다. (1) 시험날짜를 잊어버린 경우 (2) 교통체증으로 시험장에 늦게 도착한 경우 (3) 시험 중간에 패닉 상태가 된 경우 (4) 전날 일을 늦게 까지 하느라 시험공부를 못한 경우 (5) 여러분의 애완견이 교과서를 손상시킨 경우 등 해당 설명이 자세하고 명확하게 보일지라도, 다른 비슷한 상황까지 모두

¹Bacharach, S. B. (1989). “Organizational Theories: Some Criteria for Evaluation,” *Academy of Management Review* (14:4), 496-515.

다섯 가지의 상황에 적용될 수는 없다. 다시 말해서, 보편적일 수 없다는 것이다. 반면, **보편적인 설명(nomothetic explanation)**은 특정한 상황에 대한 기술보다는 해당 상황·사건과 유사한 다른 일련의 사건까지도 일괄적으로 설명하는데 중점을 둔다. 예를 들어, 학생들이 시험을 잘 치르지 못한 이유에 대해 학생들이 시험에 대한 걱정, 집중력 부족, 및 기타 장애로 인해 공부할 여건이 되지 못하였다고 설명할 수 있는 것이다. 보편적 설명이 사건 및 대상에 대한 보다 보편적인 설명을 하는 것을 목표로 하기 때문에, 세부적이고 완전한 설명을 갖기에는 불충분하다. 그러나, 보편적인 설명은 몇 가지의 변수로만 사건을 설명하므로 경제적인 이점을 가지고 있다. 그렇기 때문에 이론은 일련의 현상에 대한 보편적인 설명을 가지고 있어, 이론적인 설명은 보편적인 성격을 띤다고 할 수 있다.

이론을 살펴볼 때, 우리는 어떤 것이 이론이 될 수 없는 지를 주의하여 살펴보아야 한다. 이론은 단순한 자료나 사실, 경험적 발견의 나열 혹은 분류가 아니다. 돌맹이를 쌓는다고 집이 완성될 수 없듯이, 사실의 집합이 이론이 될 수 없는 것이다. 마찬가지로, 구성개념의 집합이 이론이 되지 않는다. 이론은 개별적인 구성개념 이외에도 구성개념 간의 연결 고리가 되는 명제, 설명, 조건들이 함께 있어야 하기 때문이다. 자료, 사실, 결과가 경험적·관찰적인 측면의 접근 방법이라면, 이론은 보다 개념적이고 논리적인 접근 방법이다.

우리가 연구를 할 때 이론을 활용하는 데에는 여러 가지 장점이 있다. 첫째, 이론은 자연적·사회적 현상을 설명하기 위해 필요한 주요 인과관계를 규명하는 주요 요인과 결과에 대한 논리를 제공한다. 둘째, 기존 연구에서 규명한 구성개념 간의 인과관계들을 살펴보고, 구성개념들 사이에 영향을 주는 요인들을 발견할 수 있다. 셋째, 기존 연구에서 언급한 구성개념이 연구자가 자신의 연구에서 적용 가능한지에 대해 일종의 가이드 라인을 확인할 수 있다. 넷째, 이론은 다른 이론과의 지식의 차이를 발견하고 축적하여, 기존의 이론이 재조명 받을 수 있도록 도움을 준다.

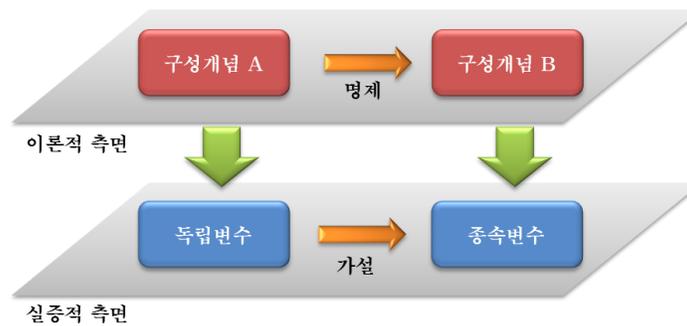
그러나, 이론은 현상을 완벽히 설명하는데 한계성을 가지고 있다. 복잡한 현실을 이론적으로 단순화하여 설명하는 과정에서 이론이 가지고 있는 구성개념과 그들의 관계에 한계성으로 충분한 설명이 이루어지지 않기 때문이다. 단순하고 보편화된 이론은 복잡한 현실을 설명하기 하기에 충분치 않다. 나아가, 이론은 연구자들의 한계적인 시각에서 바라본 결과이기 때문에, 이론에서 정립되지 않은 중요한 개념을 연구에 포함시키지 못했을 가능성도 염두에 두어야 한다.

4.2 이론의 구성요소

데이비드 위튼(David Whetten, 1989)은 이론을 구성하는데 구성개념(construct), 명제(proposition), 논리(logic), 한계적 조건/가정(boundary condition/assumption)과 같이 네 개의 구성요소가 있다고 설명했다. 구성개념은 현상을 설명하는 주요 개념으로서 이론의 ‘무엇(what)’에 해당되며, 명제는 구성개념이 어떻게 연계되어 있는지에 대한 ‘어떻게(how)’, 논리는 왜 일련의 개념들이 연관성이 있는 지에 해당되는 ‘왜(why)’, 한계적 조건/가정은 어떠한 상황에서 개념들 간의 인과관계를 적용시킬 수 있는지의 ‘누구, 언제, 그리고 어디서(who, when, and where)’을 기술한다. 비록 구성개념과 명제가 앞선 제 2장에서 설명되었지만, 여기서 보다 구체적인 설명을 하고자 한다.

구성개념(construct)는 연구하고자 하는 특정한 현상을 설명하기 위해 정의한 추상적인 개념을 말한다. 제 2장에서 살펴보았듯, 구성개념은 정해진 키, 몸무게와 같이 일차원이거나, 인성, 문화와 같이

다차원적일 수 있다. 나이, 교육수준, 회사크기와 같은 구성개념은 쉽게 이해하고 개념화 할 수 있으나, 창의력, 편견, 조직의 민첩성과 같은 구성개념은 복잡하고 이해하기 어렵다. 신뢰, 태도, 학습수준과 같은 개념 또한 일시적인 경향에 대한 정의일 뿐 지속적인 상태를 나타낼 수 없어 개념을 정의하는데 난해하다. 그럼에도, 모든 구성개념은 명확하고 애매모호하지 않는 정의를 가지고 있어야 하며 구성개념이 개별적·집단적 수준에서 어떻게 측정 될 것인가를 정확히 명시해야 한다. 연구에서 측정 가능한 추상적 구성개념을 **변수(variable)**라고 한다. 예를 들어, IQ 지수(Intelligence Quotient Score: IQ score)는 지적능력(intelligence)이라는 추상적인 개념을 측정 가능하게 하는 변수이다. 앞에서 언급하였듯이, 과학적 연구는 이론적인 측면과 실증적인 측면으로 이루어 진다고 하였다. 구성개념이 이론적인 측면에서 현상을 개념화 시킨 것이라면, 변수는 실증적인 측면에서 현실세계를 관찰하고 측정하는데 사용된 것이다. 이러한 변수들은 앞서 2 장에서 설명하였듯이 독립변수(independent variable), 종속변수(dependent variable), 매개변수(mediating variable), 조절변수(moderating variable) 등으로 분류 될 수 있다. 이론측면의 개념을 정립하는 구성개념과 현실측면을 관찰하고 측정하는 변수의 구분은 아래와 같이 도식화 할 수 있다(그림 4.1).



<그림 4.1> 이론적·실증적 개념의 구분

명제(proposition)는 연역적 추론을 바탕으로 구성개념 간의 관계를 정의한 것이다. 명제는 ‘-다’와 같은 평서법으로 기술되며, 인과관계를 나타내어야 한다. 예를 들면, 원인 X 로 인해 결과 Y 가 뒤따른다. 명제는 단순히 예측으로만 끝날 수 있으나, 검증 가능한 형태여야 하고 실증적 증거가 없을 경우 부정할 수 있어야 한다. 구성개념과 마찬가지로, 명제는 이론적 측면으로 나타나면서 각 구성개념을 측정하는 변수들의 관계들을 설명할 수 있어야 한다. 이렇듯, 변수들 간의 관계를 검증 가능한 형태로 서술한 명제를 **가설(hypothesis)**이라고 한다. 이론적 형태로 나타나는 명제와 검증 가능한 형태의 가설의 구분은 그림 4.1 에서도 쉽게 알 수 있다.

이론을 구성하는 세 번째 요소인 **논리(logic)**는 해당 이론의 바탕이 되는 명제들에 대한 기본 배경을 제시한다. 즉, 논리는 이론적 구성개념과 그 관계에 대한 의미를 구성하기 때문에 ‘접착제(glue)’와 같은 역할을 한다. 이것은 이론의 가장 중요한 핵심이 되는 ‘설명(explanation)’을 가능하게 만든다. 만약, 논리가 충분하지 않으면, 명제는 설득력을 잃고 모호하게 되어 이론의 구성력을 해치게 된다.

마지막으로, 이론은 가치관, 시간, 공간 등의 한정적인 요소들에 대한 **가정(assumption)**으로 **경계조건(boundary conditions)**을 가지며, 동일한 통제상황에서 똑같이 적용 될 수 있어야 한다. 예를

들어, 경제학 이론에서 인간이 이성적이며 비용과 이득을 바탕으로 경제적이고 효율적인 의사결정을 한다고 가정한다. 반면, 정치학 이론에서 인간은 합리적이고 경제적이기 보다 본인의 가질 수 있는 권력을 위해 정치적으로 행동한다는 가정을 바탕으로 한다. 이러듯 각기 다른 가정을 바탕으로 하는 이론들은 직접적인 비교 대상이 될 수 없기 때문에 정치적인 힘의 구조와 조직에 대한 연구 없이 경제 이론을 적용할 수 없는 것이다. 이와 마찬가지로, 이론은 문화적 가정(예: 개인적 - 집단적 수준), 시간적 가정(예: 시간대 별 인간의 행동), 공간적 가정(예: 분야별 이론의 적용)을 포함하고 있다. 이론을 보다 정확하게 검증하기 위해서는, 이러한 가정에 대한 이해가 충분히 되어야 한다. 불행하게도 이론학자들 내포된 가정을 명확하게 이론에서 제시하고 있지 않기 때문에 연구자들은 이론에서 검증하고자 하는 현실에 대해 명확한 관찰을 바탕으로 한 가정을 설정해 두지 않는다면, 이론을 적용하는데 많은 문제가 발생하게 될 것이다.

4.3 이론의 평가기준

이론은 복잡한 현실을 단순하게 하거나 일부를 설명한다. 따라서, 이론에 대한 설명이 얼마나 잘 되어있느냐에 따라 이론의 적합성이 결정되는 것이다. 그렇다면 이론의 ‘적합성(goodness)’은 어떻게 평가 할까? 연구의 성질과 연구자의 목적에 따라 이론의 적합성에 대한 평가에는 다양한 기준이 적용되나, 다음 주요 항목들은 공통적으로 눈 여겨 보아야 한다.

- **논리의 일관성(logical consistency):** 이론의 구성개념(construct), 명제(proposition), 경계조건(boundary condition)들 간에 논리적인 일관성을 갖추고 있어야 한다. 만약 이론을 구성하는 위의 세 요소들이 일관적이지 못하다면 이론은 설득력을 잃게 된다. 예를 들면, 합리성을 가정으로 하는 이론에서 구성개념이 비합리성이라는 개념을 포함한 경우이다.
- **설명 능력(explanatory power):** 연구에 적용되는 이론이 현실을 얼마나 잘 설명하고 있는지 확인해야 한다. 연구를 통해 설명하고자 하는 현상을 기존의 이론보다 정확하게 설명할 수 있어야 한다. 여기서 이론의 설명력은 회귀 방정식에서 분산 값인 R-square 로 측정한다.
- **반증 가능성(falsifiability):** 영국의 철학자 칼 포퍼(Karl Popper)는 1940 년쯤, 연구의 배경이 되는 이론은 타당성을 갖추면서도 오류에 대한 검증 능력도 함께 포함되어야 한다고 주장하였다. 반증 가능성이란 만약 실증적 데이터가 이론적 명제와 일치하지 않을 경우, 이론이 틀렸음을 입증하는 것이다. 다시 말하면, 실증적으로 검증 할 수 없는 형태는 이론이 될 수 없다. 가령 “높은 온도의 날씨는 더운 날씨이다.” 와 같은 문장은 날씨의 ‘더운’ 정도와, 온도의 ‘높음’ 의 정도가 정확하게 정의 되지 않은 모호한 문장으로 구성되어 있기 때문에 이론으로 성립되기가 어렵다. 반증 가능성은 적절한 구성개념의 측정으로 증명 될 수 있는 반대적 설명을 인정하는 것이다. 하지만 이론의 반증 가능성이 이론은 반증성을 가지고 있어야 한다는 의미와 구별된다. 만약 이론이 실증적 검증을 통하여 반증성을 가지고 있을지라도 연구에 따라 그 이론이 적합하지 않을 수 있기 때문이다.
- **간결성(parsimony):** 이론의 간결성은 이론을 설명하는데 사용 된 변수의 수에 근거한다. 13 세기 영국의 논리학자 윌리엄 오컴(William of Ockham)은 이론이 사회현상을 보다 단순하고 간결한 변수와의 가정을 통해 설명되어야 한다고 주장하였다. 이를 오컴의

면도날(Ockham's razor)이라고 한다. 복잡한 사회현상을 설명하기 위하여 종종 구성개념의 개수를 늘리기 마련이다. 그러나 이러한 접근은 현상을 간결하고 보편화 하고자 하는 이론의 연구목적에 불분명하게 한다. 간결성은 이론의 자유도(degrees of freedom)와 관련이 있으며, 자유도가 높을 수록 이론의 맥락, 배경에 적용할 수 있는 보편성이 높다고 본다.

4.4 이론화에 대한 접근

연구자들은 이론을 정립하기 위해 어떻게 접근하는가? 스테인필드와 풀크(Steinfeld & Fulk, 1990)² 의 연구에서 이론화에 대해 네 가지 접근법을 제시하고 있다. 첫 번째로는 현상의 패턴을 관찰하고 그를 바탕으로 귀납적인 추론을 통해 이론을 구성한다. 경험적 관찰을 통해 배경이 되는 이론을 형성하는 과정을 ‘근거이론형성(grounded theory building)’ 이라고 한다. 이것은 연구자의 관찰력과 해석능력이 주로 요구되며, 이 과정으로부터 추출되는 이론은 주관적이고 확인하기 어려운 형태일 수도 있다. 또한, 연구자가 관찰한 패턴에 대해 일관성을 주장하지 못한다면, 관찰 결과가 이론으로 연결되기 어렵다. 근거이론(grounded theory)에 대한 설명은 정성연구(qualitative research)에서 더 자세히 설명하기로 한다.

두 번째 접근방법은 연구하고자 하는 현상에 대해 기존의 프레임워크(framework) 안에서 적절한 요인들을 적용하여 상방향식 접근방법(bottom-up)의 개념 분석을 통해 이론을 정립한다. 기존의 틀이 투입-진행-결과와 같은 단순한 체계를 가지고 있다면 연구자는 다양한 카테고리의 투입변수를 사용하여 결과에 대한 연결고리를 찾아 낼 수 있다. 예를 들어 투입변수는 개인, 조직, 기술적 요인이 결과와 관련이 있거나 투입과 결과 사이에 존재하는 진행부분을 세분화하여 접근할 수 있다. 이 절차는 귀납적 접근을 통해 이루어지고 연구자의 논리력과 해석능력, 그리고 현상에 대한 사전지식의 정도에 따라 연구를 진행한다.

세 번째 접근방법으로는 기존 이론의 범위를 넓히거나 수정하여, 새로운 맥락에서 현상을 설명하는 것이다. 예를 들어, 개인의 학습 이론을 확장시켜 조직의 학습능력을 설명하는 것이다. 이 과정에서 기존의 이론에서 조작된 개념의 정의, 명제, 한계조건들은 새로운 맥락에서 연구 할 경우, 수정되거나 재정의 될 수 있다. 그러므로 연역적인 접근을 통해 개발된 이론은 사회과학 이론을 더욱 풍부하게 하였으며 기존의 것보다 효과적인 방법으로 설명할 수 있게 만들었다.

마지막으로 기존의 이론에서 설명된 현상과 해당 연구에서 설명하고자 하는 현상간의 유사점을 바탕으로 하여, 기존의 이론을 새로운 맥락과 구조에서 적용시키고 이해하는 것이다. 이 접근 방법은 연역적인 유추를 사용하여 이론을 형성하는데 아날로그 방식이며 가장 창의적인 방법이다. 예를 들어, 마커스(Markcus, 1987)³ 는 핵폭발과 통제할 수 없는 네트워크의 확장, 네트워크 기반의 사업의 유사점을 찾고, 네트워크 성장에 관한 결정적 다수 이론(critical mass theory)을 적용하여 연구를 하였다. Markus 는 핵 폭발 과정에서 폭발 에너지를 지속시키기 위한 방사능의 임계질량이 필요한

² Steinfeld, C.W. and Fulk, J. (1990). "The Theory Imperative," in Organizations and Communications Technology, J. Fulk and C. W. Steinfeld (eds.), Newbury Park, CA: Sage Publications

³ Markus, M. L. (1987). "Toward a 'Critical Mass' Theory of Interactive Media: Universal Access, Interdependence, and Diffusion," *Communication Research* (14:5), 491-511.

것처럼, 인간의 네트워크 또한 성장을 지속시키기 위하여 충분한 사람이 필요하다고 설명하였으며 화학적인 이론의 논리를 전혀 새로운 맥락에서 적용하고 상황을 설명하였다.

4.5 사회과학 이론의 예시

이번에는 다양한 사회과학 원리에 대한 이론을 간략히 살펴보고자 한다. 이론들은 각기 다른 구성개념, 명제, 한계조건, 가정, 배경 논리를 통해 다양한 사회적 행동을 설명하고 있다. 여기서 설명되는 이론들은 이론의 형성과 구조에 대한 이해를 돕기 위함이며, 이론의 세부적인 설명은 생략하기로 한다.

대리이론(Agency Theory or Principal-Agent Theory)은 조직경제학을 기초로 하여 이익과 목적이 상충하는 두 집단(예: 고용주-고용인, 기업 경영진-주주들, 판매자-구매자) 간의 관계를 설명한 이론으로, 로스(Ross, 1973)⁴의 연구에서 처음 소개 되었다. 대리 이론에서는 두 집단 간의 서로 상충하는 목표를 최적화 시키는 조건을 찾는 것이 그 목적이 된다. 이 이론의 핵심적인 가정은 인간은 자신의 개별적 이익을 추구하며, 합리적이며, 위험기피적인 특성을 개인과 조직 수준에 적용할 수 있다는 것이다.

대리이론의 대상이 되는 두 집단은 주인(principal)과 대리인(agent)로 설명 되는데, 주인은 대리인을 고용하여 일정한 업무를 지정하게 된다. 여기서 주인은 집단의 목표를 보다 효율적이고 경제적으로 업무를 마치는 것이 되고, 대리인은 큰 위험을 감수하지 않으며 집단의 이익보다는 본인의 개인적 이익(주로 임금)을 추구하게 된다는 것이다. 따라서, 개인과 집단의 이익은 상충할 수 밖에 없다. 특히 주인 집단이 대리인 집단의 행동이나 능력에 대해 정확히 관찰하지 못할 경우 정보의 부족으로 대리인 집단이 도덕적 해이(moral hazard) 등의 이유로 업무능력이 반감하거나, 자신의 능력을 충분히 발휘 하지 못하는 환경을 조성하는 등의 문제들이 야기된다. 월급제도와 같이 행동 기반의 전형적인 관계는 이러한 문제를 해결하지 못한다. 따라서, 대리이론은 결과에 초점을 맞추라고 제안하고 있다. 가령, 업무의 난이도 및 업무 수행 정도에 따라 임금을 지급하는 계약, 기존의 업무 수행에 대한 월급에 성과급과 커미션을 더하는 복합적 형태의 계약 형태이다. 또한 직원에게 주식을 배분하는 것 또한 결과 중심의 계약이라고 할 수 있다. 대리이론은 행동 기반 계약의 효율성을 높이는 방법론을 제시한다. 가령, 조직 내 감시자를 두어 모니터링을 하며, 직원과 경영진 간의 정보의 차이를 줄이거나, 업무 배분 구조를 더욱 체계화 시키는 것이 한 예가 된다고 할 수 있다.

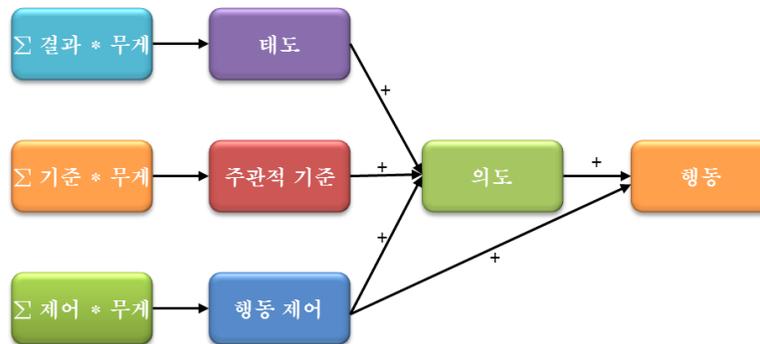
계획적행동이론(Theory of Planned Behavior: TPB)은 아이젠(Ajzen, 1991)⁵이 주장한 이론이다. 이것은 사회심리학적 맥락에서 인간의 행동을 일반화한 것으로 개인 행동으로 범위에 적용시킨 이론이라고 할 수 있다. 여기서 개인 행동은 이성적인 판단과 더불어 논리적 사고와 사회적 책임감에 대한 고려를 바탕으로 결정 된다. 이 이론은 인간의 행동(behavior)에 영향을 주는 인간의 의도(intention)에 근거한다. 이는 인간의 행동에 대한 개인적 태도(attitude), 주관적인 규범 (subjective norm), 행동조절(behavioral control)을 바탕으로 형성된다(그림 4.2). 태도(attitude)는 특정 행동에 대한

⁴ Ross, S. A. (1973). "The Economic Theory of Agency: The Principal's Problem," *American Economic Review* (63:2), 134-139.

⁵ Ajzen, I. (1991). "The Theory of Planned Behavior," *Organizational Behavior and Human Decision Processes* (50), 179-211.

개인의 신념과 희망하는 결과에 대하여 느끼는 긍정적·부정적 감정으로 정의할 수 있다. 주관적 규범(subjective norm)은 사람이 자신의 주위에 있는 친구, 동료, 상사 등에 대해 갖는 기대치와 중요한 기대역할에 대한 개인적 관점을 말한다. 행동조절 (behavioral control)은 개인의 행동에 대한 내부적·외부적 통제 요건을 말한다. 내부적 제어는 어떠한 목적을 가진 행동을 하고자 하는 개인의 능력(자기 효능감: self-efficacy)을 말하며, 외부적 제어는 행동을 가능하게 하는 외부적 요소(가능 조건: facilitating conditions)를 말한다. 계획적행동이론은 인간이 주어진 행동을 하고자 하나 뒷받침이 될만한 요소가 없는 경우, 행동 제어는 행동에 직접적으로 영향을 주는 요인으로 작용하며, 의도는 간접적으로 영향을 미친다고 설명하고 있다.

계획적행동이론은 합리적행동이론(Theory of Reasoned Action: TRA)으로부터 확정된 이론이다. 합리적행동이론에서 태도, 주관적 규범은 주요 요소로 꼽히나, 행동제어는 포함되지 않는다. 행동 제어라는 개념은 아이젠(Ajzen)이 인간의 개인행동에 대해 완벽하게 제어할 수 없을 경우에 대한 고민으로 추가한 것이다. 예를 들어, 웹 서핑을 하는데 빠른 인터넷 연결이 없는 경우와 같은 때와 같다.

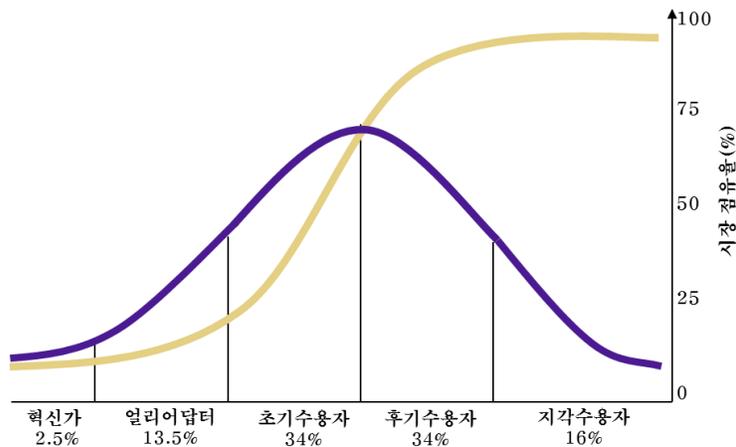


<그림 4.2> 계획적 행동 이론

혁신확산이론(Innovation Diffusion Theory: IDT)은 커뮤니케이션 이론에서 중요한 이론 중 하나로, 혁신적인 기술의 확산이 잠재된 수용자에게 어떻게 습득되는 가를 설명하는 것이다. 프랑스 사회학자 가브리엘 타르드(Gabriel Tarde)에 의해 소개된 이 이론은 1962년 에버렛 로저(Everett Roger)의 확산 연구를 바탕으로 구체적으로 발전되었다. 이 이론에서 핵심이 되는 네 가지 개념은 혁신(innovation), 소통 채널(communication channel), 시간(time), 사회체계(social system)이다. 혁신은 새로운 기술, 아이디어, 수용자를 포함하며 개인 혹은 조직 단위에서 설명된다. 거시적 관점에서, 혁신 확산은 하나의 사회체계 내에 속한 사람들 사이에서 새로운 혁신을 접하고 그에 대한 가능성을 소통채널(예: 대중매체, 얼리어답터)을 통해 확산시키는 커뮤니케이션의 한 과정으로 본다. 더불어 확산은 시간의 제약을 받는다. 즉, 확산은 소수의 얼리어답터에 의해 천천히 수용되며, 빠르게 많은 인구들에게 퍼지며, 확산의 속도를 늦추며 혁신 수용인구 또한 줄게 되는 형태를 갖고 있다. 확산의 형태는 그림 4.3에서 보듯이 S자 형태를 띠며 정규분포를 갖는다. 수용시기에 초점을 맞추어 혁신 수용자를 혁신가(innovator), 얼리어답터(early adaptor), 초기수용자(early majority), 후기수용자(late majority),

지각수용자(laggard)로 구분한다. 확산의 정도는 오피니언리더(opinion leader)와 그렇지 않은 사람들의 사회적 체계에 따라 다르게 나타난다.

미시적 관점에서 로저(Rogers, 1995)⁶는 혁신의 수용을 다섯 단계로 구성하였다 (1) 인지단계: 수용자가 미디어나 개인 채널을 통해 혁신을 처음 접하는 단계, (2) 설득단계: 먼저 혁신을 접한 이들로부터 혁신을 시도해 볼 것을 설득 받는 단계, (3) 결정단계: 혁신을 수용할 것인지를 판단하는 단계, (4) 수용단계: 처음 혁신을 받아들이는 단계, (5) 확산단계: 결정에 대한 지속적으로 사용할 것인가에 대한 결론을 내리는 단계가 혁신확산의 과정이다(그림 4.4). 다음의 다섯 가지 혁신에 대한 특성은 수용자들을 혁신 수용의 단계로 이끈다. (1) 상대적 이점(relative advantage): 이전의 혁신에 비교하여 상대적 장점에 대해 기대, (2) 적합성(comparability): 혁신이 개인의 업무형태, 가치, 신념과 적합한가에 대한 정도, (3) 복잡성(complexity): 혁신이 수용하는데 얼마나 어려운지 여겨지는 정도, (4) 시도가능성(trialability): 혁신에 대해 테스트 할 수 있는 트라이얼(trial)의 존재여부, (5) 관찰가능성(observability): 혁신의 결과가 얼마나 정확하게 관찰 가능한가의 정도이다. 마지막 두 특성은 혁신에 관한 연구에서 많이 다루어져 왔다. 다른 네 가지 특성과 달리, 복잡성만 혁신의 수용과 반상관관계를 가진다. 혁신의 수용은 수용자 개인이 위험을 부담하는 정도, 교육적 수준, 도시화 수준, 소통력에 따라 차이를 보이게 된다. 초기수용자는 보다 모험적이고, 교육적 수준이 높은 경향이 있으며, 혁신에 대한 정보를 매스미디어에서 주로 접하게 된다. 후기수용자일수록 가족이나 친구와 같이 주변인으로부터 주요 정보를 얻어 혁신을 접하게 되는 경우가 많다. 혁신 확산 이론은 대표적으로 ‘친혁신편향(pro-innovation bias)’으로 많은 비판을 받기도 하는데 그 이유는 혁신이 개인의 혁신 수용단계에서 중간에 중단되거나 개인적으로 거부되는 경우, 기술의 확산 기간 동안 다른 혁신으로 전환되는 경우 등 간과하고 있다는 것이다.



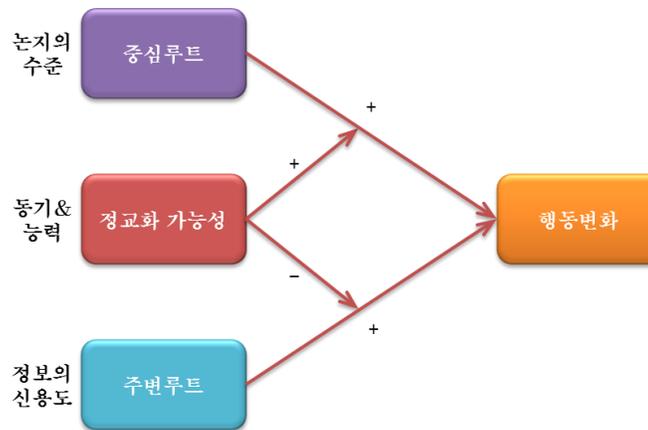
<그림 4.3> 혁신 확산의 S 곡선

⁶ Rogers, E. (1962). *Diffusion of Innovations*. New York: The Free Press. Other editions 1983, 1996, 2005.



<그림 4.4> 혁신 확산의 과정

정교화 가능성 모델(Elaboration Likelihood Model: ELM)은 페티와 카시오포(Petty & Cacioppo, 1986)⁷의 연구에 의해 발전된 이론으로, 태도의 형성과 심리학적 관점에서 보는 이중과정이론(dual-process theory)이다. 개인의 태도 변화가 특정 대상, 사건, 행동 등에 대해 어떻게 영향을 받아 변화하는지 그 상관관계로 보이는 요소를 설명하였다. 정교화 가능성 모델에서 개인의 태도는 두 개의 ‘루트(root)’를 통해 영향을 받는다고 가정한다. 두 가지 루트는 중심경로(central route)와 주변경로(peripheral route)로 나뉘는데, 이는 정보의 처리 능력과 정교화의 요구 수준에 따라 다르게 나타난다(그림 4.5). 중심경로는 주제와 관련되어 공식적인 판단을 내리기 전, 주제와 관련된 비공식 메시지를 고민하고, 그 가운데서 개인적 관련성 및 중요성을 처리하는 과정이다. 주변경로를 통해서 하는 의사결정은 주요 시사점을 메시지의 핵심이나 의도가 아닌 이전의 사용자, 전문가 의견 등 주변 ‘큐(cue)’를 통해 하는 것이다. 그리고 의식적인 것을 요구하지 않으며 이러한 중심경로와 주변경로는 정교화 가능성 모델에서 각각 주장의 질(argument quality)와 주변정보(peripheral cues)와 같은 구성개념을 사용하여 태도가 바뀔 수 있다고 보았다.



<그림 4.5> 정교화 가능성 모델

중심경로와 주변경로의 영향력은 개개인이 중심메시지를 얼마나 수용하고 자극을 받는지 그 수준에 의해 결정되고 이것은 정교화 가능성, 혹은 관여도(elaboration likelihood)라고 불리는 수용에 대한 정교화 과정을 거치게 된다. 높은 수준의 정교화 가능성(관여도)을 가지고 있는 사람들은 주어진 정보에 대하여 좀 더 깊이 있는 생각하는 과정을 거치게 됨으로써 주장의 질에 영향을 미친다.

⁷ Petty, R. E., and Cacioppo, J. T. (1986). *Communication and Persuasion: Central and Peripheral Routes to Attitude Change*. New York: Springer-Verlag.

반면에 낮은 수준의 정교화 가능성(관여도)을 가지고 있는 사람들은 좀 더 주변 정보의 영향을 받는다. 정교화 가능성(관여도)은 개인의 특성보다는 상황의 특성에 많은 영향을 받는다. 예를 들어, 의사는 어떤 질병에 대해 진단 할 때, 그들의 가지고 있는 경험에 중심경로로 두게 된다. 그러나, 그들이 자동차를 고를 때는 그 외 주변경로를 통한 결정을 하게 될 것이다. 이렇듯, 정교화 가능성 모델에서는 새로운 제품이나 사회적 변화에 대해 대상의 태도 변화가 어떠한 과정을 보이는지에 보여준다.

일반억제이론(General Deterrence Theory: GDT)은 18 세기 두 실용주의학자 쎬사레 베카리아와 제레미 베타(Cesare Beccaria & Jeremy Betham)에 의해 형성된 이론이다. 일반 억제 이론은 범죄와 범죄대책 두 가지 모두를 설명하는데, 개인이 왜 사회에서 분리되어 반사회적, 범죄 행동을 하게 되는지를 설명하고 있다. 인간은 기본적으로 합리적인 결정을 한다는 가정 하에, 인간은 자신이 비용과 이점을 고려하여 일탈적 행위를 자유 의지에 따라 결정한다는 것이다. 인간은 본연 이익의 최대화를 추구하고, 일탈을 통해 얻어지는 이익은 범죄에 대한 처벌과 그 고통에 대한 예측으로 제어가 된다는 것이다. 처벌의 신속성, 엄격성, 확실성이 억제이론을 구성하는 주요 구성개념이다.

범죄와 관련된 정통 해석적 연구에서는 일반적으로 범죄행동이 가난, 교육의 부재, 심리적 요소에 의해 발생된다고 하였으며, 그들에게 훈련과 의학적 치료를 제공함으로써 재활과 선도가 가능하다고 주장한다. 그러므로, 일반억제이론에서는 범죄자의 개인적 요소(개인의 가치관 이나 부에 대한 요구 등)와 환경적인 맥락(피해자의 안전도, 지역 경찰의 효율성 등)에 의해 범죄가 결정되는 것이라고 보았다. 일반억제이론의 중점은 사회적응을 위한 범죄자의 재활과 미래범죄의 예방, 범죄자들의 범죄행동에 대한 흥미와 관심을 억제시켜 원천적으로 범죄행위를 제한한다는 것이다. 가령, 경찰의 순찰 횟수를 늘리거나, 법원에서 가석방제도를 없애고 사형제도를 권장하는 등 예비범죄자들에게 가시적인 경고를 하는 것이다. 일반 억제 이론은 정통적인 범죄뿐 아니라, 현대의 지적재산권 침해와 같은 화이트칼라 범죄도 포함한다.

제 5 장 연구설계

연구설계는 실증적 연구에서의 자료를 수집하기 위한 포괄적인 연구 계획을 말한다. 즉, 특정한 연구질문에 대한 해답을 얻거나 가설을 검증하기 위한 연구 과정을 의미하며, 실증적 연구의 ‘청사진(blueprint)’ 과 같은 것이다. 연구설계는 다음과 같이 세 단계를 거치며 구체화 시킬 수 있다. (1) 자료수집(data collection), (2) 측정도구 개발(instrument development), (3) 표본추출(sampling)이 그것이다. 측정도구와 표본추출에 대한 설명은 다음 장에서 각각 자세히 설명하기로 하고, 이번 장에서는 연구설계와 맥락을 같이 하기도 하는 데이터 수집 과정에 대해 알아보기로 한다.

자료수집(data collection) 방법은 보편적으로는 실증적 방법과 해석적 방법 두 가지로 크게 구분된다. **실증적 방법(positivist method)**은 통상 연구실의 실험이나 설문조사 등 이론과 가설을 검증하는 데 사용되고, **해석적 방법(interpretive method)**은 행동연구나 사회문화 연구에서 이론을 수립하는데 적용된다. 실증적 방법에서는 연역적 접근을 바탕으로 한 실증적 데이터를 수집하여 이론을 실증적으로 검증한다. 반면, 해석적 방법은 귀납적 추론을 바탕으로 수집된 데이터에서 현상을 이론으로 이끌어 정립한다. 실증적·해석적 방법은 주로 양적 조사(quantitative research)와 질적 조사(qualitative research)가 종종 혼용 되어 사용하는데, 양적 조사와 질적 조사는 수집·분석된 자료의 종류를 의미한다. 다시 말해, 점수나 단위와 같이 수치상으로 나타나는 데이터들로 회귀분석 등을 자료분석을 하는 것을 양적 연구, 심층면접이나 인터뷰 및 관찰로 얻어진 자료를 기반으로 코딩(coding) 하여 그 의미를 해석하는 것을 질적 연구라고 한다. 실증적 연구는 양적 자료를 주로 활용하지만, 질적 자료 또한 연구 대상에 포함된다. 해석적 연구에서는 주로 질적 자료를 기반으로 하지만 양적 자료를 참고 하기도 한다. 이렇듯, 질적 자료와 양적 자료 중 어느 것에 치우치지 않고, 연구 목적에 적절히 함께 활용하는 혼합조사설계(mixed-mode design)가 복잡한 사회 현상에 대하여 보다 참신한 통찰력을 도출하는데 많은 도움이 된다.

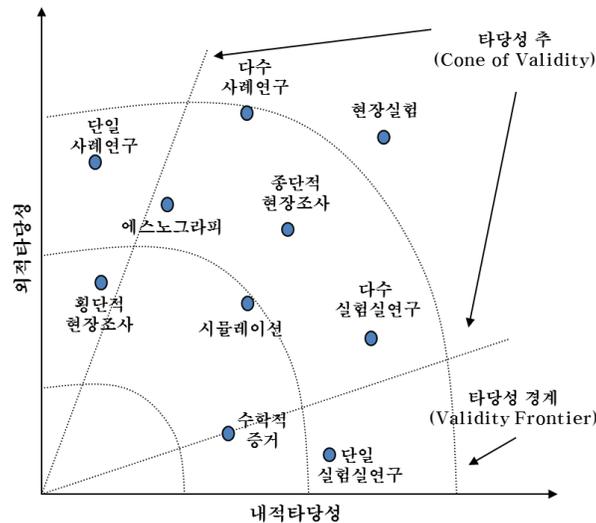
5.1 연구설계의 기본 요건

연구설계가 얼마나 해당 가설과 연구주제를 정확하게 검증할 수 있는지는 다음의 주요 네 가지 요건을 충족하는지 살펴봐야 한다. (1) 내적 타당성(internal validity), (2) 외적 타당성 (external validity), (3) 개념 타당성(construct validity), (4) 통계적 결론 타당성(statistical conclusion validity)

내적 타당성(internal validity)은 통상 인과관계(causality)라고도 하는데, 종속변수의 변화가 다른 연구 범위 밖의 외생 변수들이 아닌 실제 이론상에서 정의된 독립변수의 변화 때문에 일어난 것인가에 관한 것이다. 인과관계를 성립하기 위해 필요한 세 가지 조건은 다음과 같다. (1) 원인과 결과 사이의

공변량(covariance): 어떠한 결과가 일어나기 위해서는 원인이 존재 하여야 하며, 원인이 없으면 결과 또한 발생하지 않는다. (2) 사건의 일관성: 원인은 결과보다 먼저 앞서서 발생해야 한다. (3) 대체 설명의 부재 혹은 허위상관(spurious correlation)이 존재하지 않아야 한다. 실험실 연구와 같은 연구설계는 외생변수가 철저하게 통제되는 동안에 조작을 통한 독립변수에 따른 종속변수의 변화를 관찰함으로써 내적 타당성은 타 연구에 비해 높다. 그러나, 현장조사(field surveys)와 같은 연구에서는 현실적으로 독립변수에 영향을 주는 외생변수의 통제가 어렵다. 왜냐하면 원인과 결과가 일시적으로 동시에 측정되고 원인과 결과를 제대로 설명할 수 있는 사건의 일관성 또한 보장하기가 어렵기 때문이다. 실험실 연구가 비록 타 연구에 비해 내적 타당성이 상대적으로 높다 하더라도 연구 내에서 사전 배경지식, 사전연구, 분석내용 등에서 오차 등 다양한 변수가 생길 수 있기 때문에 그 타당성을 늘 보장할 수 없다. 이것에 대한 구체적인 설명은 이후 실험설계 부분에서 논의될 것이다. 무엇보다도 연구자는 연구의 목적 및 방법에 따라 충족되어야 할 내적 타당성의 수준이 각각 다르다는 것을 상기 할 필요가 있다.

외적 타당성(external validity)은 보편성이라고 하는데 이는 실험결과의 일반화에 관한 문제이며, 표본을 모집단으로 일반화(population validity) 시키거나, 혹은 시점 및 상황 등에 대한 확장(ecological validity)에 대한 것이다. 예를 들어, 어떠한 연구가 미국 내 금융기업을 대상으로 얻은 결과를 전세계 금융기업 전체에 대해(population validity), 또는 미국 내 전체 기업을 대상으로(ecological validity) 적용할 수 있을까? 개인, 집단, 회사와 같은 폭넓은 연구대상을 바탕으로 설문조사하여 데이터를 수집하는 것이 여러 외생변수에 대해 제한 및 통제가 가능한 실험실 연구에 비해 현실세계의 현상을 일반화 하기가 수월하고 보편성을 상대적으로 높게 한다. 연구설계에 있어 나타날 수 있는 내적·외적 타당성의 범위는 다음 그림 5.1 과 같이 나타날 수 있다.



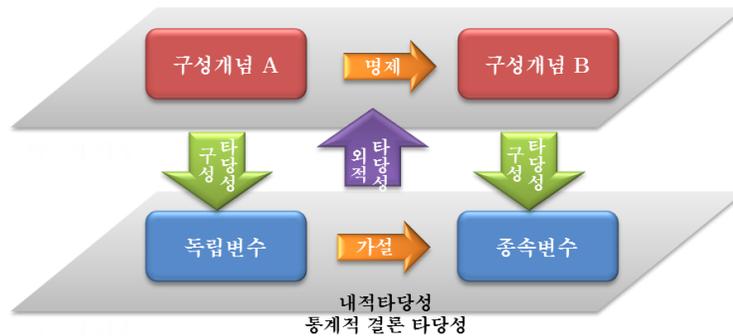
<그림 5.1> 내적 타당성과 외적 타당성

일부 학자들은 내적 타당성과 외적 타당성 간에 상충관계(tradeoff)가 형성된다고 한다. 외적 타당성이 높을수록 내적 타당성이 낮거나, 그 반대의 경우가 된다는 주장이다. 그러나 이러한 주장은 일부에 해당한다. 현장실험(field experiment), 추적조사연구(longitudinal field survey),

다중사례연구(multiple case study)는 내적·외적 타당성이 모두 높게 나타나는 연구들이다. 저자는 그림 5.1 에서 보듯이, 타당성의 추(cone of validity) 내에 위치하는 내적·외적 타당성이 적절한 수준의 연구를 선호한다. 그렇다고 해서, 추 외부의 연구들이 효용성이 떨어진다면, 연구로서의 가치가 떨어지는 것은 아니다. 연구자들은 각자의 연구의 목적과 분야의 특성에 따라 부합하는 내적·외적 타당성의 정도를 결정하고 그에 따라 결정하는 것이다.

개념 타당성(construct validity)은 연구에서 측정하려고 하는 이론적 구성개념을 얼마나 정확하게 측정하는 가를 검사하는 것이다. 가령, 공감(empathy), 변화에 대한 기피(resistance to change), 체제 학습(organizational learning)과 같은 사회과학 구성개념들은 정의하기 어려우며 측정 하기도 쉽지 않다. 적합한 개념 타당성을 위해서는 공감(empathy)이라는 개념이 다른 유사한 개념에 빗대어 측정되지 않아야 한다. 실증적 연구에서 개념 타당성 측정은 예비검사의 상관관계 측정이나, 요인분석(factor analysis)을 통하여 알아볼 수 있다.

통계적 결론 타당성(statistical conclusion validity)은 유효한 통계적 수치를 통해 결론을 검증하는 것이다. 가설을 검증하는데 적절한 통계 기법을 사용하였는지, 혹은 특정 통계적 기법들의 가정이 변수를 잘 포함하고 있는지 등을 보는 것이다. 특히, 해석적 연구 설계는 특별한 통계적 기법을 차용하지 않아 통계적 의미가 유의하지 않다. 서로 다른 종류의 타당성이 이론적 세계와 경험적 세계에서 어떻게 작용하는지 그림 5.2 에서 설명하고 있다.



<그림 5.2> 과학적 연구에서의 타당성 종류

5.2 내적 타당성과 외적 타당성을 향상시키는 요건

연구자들이 기대하는 좋은 연구설계는 내적 타당성과 외적 타당성이 높은 연구이다. 외적 및 내적으로 타당성이 높은 연구는 가설을 검증하고 표본의 적용 범위를 넓혀 일반화 시키는 가장 기본적인 전제가 되기 때문이다. 연구의 내적 타당성(causality)을 높이기 위해서는 다음과 같은 통제방법이 있다. (1) 조작(manipulation), (2) 제거(elimination), (3) 상쇄(inclusion), (4) 통계의 통제(statistical control), (5) 무작위화(randomization)

연구설계 과정 중 독립 변수의 **조작(manipulation)** 혹은 실험자극(treatment)을 통하여 타당성을 높일 수 있다. 즉, 여러 단계에 걸쳐 실험자극을 주고, 자극을 준 집단과 그렇지 않은 집단 간의 효과를

비교하여 결과를 도출한다. 가령, 의학분야에서 실험자극은 새로 개발한 시약품, 교육 연구에서는 강의법 등이 있다. 이러한 실험자극을 활용한 통제는 실험연구(experimental) 또는 준실험연구(quasi-experimental) 설계에서 주로 이루어 지며, 설문조사와 같은 비실험연구에서는 쓰이지 않는다. 실험자극을 주는 연구과정에서 그 대상의 구분이 제대로 이루어 지지 않는다면, 실험자극이 미치는 영향을 정확하기 파악할 수 없다.

실험자극(treatment)를 통제된 상황에서 연구에 영향을 미치는 외생변수를 **제거(elimination)**하는 기법은 연구의 타당성을 높이기도 한다. 또한, 연구설계에서는 외생변수를 포함하고 요인설계에 따른 종속변수의 영향을 따로 분석 하는 방법을 **인클루션기법(inclusion technique)**이라고 한다. 특히, 일반화를 가장 높일 수 있는 방법 중 하나이며 많은 표집자료가 필수적이다. **통계적통제(statistical control)**는 외생변수를 발굴하고 통계적 검증(statistical testing) 과정에서 공변량(covariates)을 측정함으로써 가능하다.

마지막으로 **무작위기법(randomization technique)**은 무작위 표본추출을 통한 외생변수에 대한 종속변수에 미치는 영향을 제거한다. 무작위화에는 두 가지 방법이 있다. (1) 모집단에서 표본을 무작위로 추출해내는 **무작위선택(random selection)**, (2)확률적 방법으로 선택된 표본이 무작위로 처리집단에 배정하는 **무작위배정(random assignment)**이 있다. 무작위화는 실제 연구의 대상이 되는 모집단에서 표본집단을 추출하여 이론을 일반화시키는 외적 타당성에 적합하다. 표본의 자원과 제한적 요인들 때문에 무작위선택(random selection)이 어려울 경우 무작위배치(random assignment)를 해야 한다. 그러나 이 경우, 모집단이 여러 관점으로부터 구성되어 존재하여 연구자가 이를 통제할 수 없다면 일반화하기는 어렵다.

5.3 자주 쓰이는 연구설계 형태

앞서 언급하였듯이, 연구설계는 과학적 연구에서 연구 목적에 따라 실증적, 해석적 연구 크게 두 가지로 나뉜다. 실증적 연구는 이론을 검증하고 해석적 연구는 이론을 설정한다. 실증적 연구에서는 현실에서 관찰된 사실을 바탕으로 일반화 할 수 있는 일련의 법칙과 패턴을 찾는다면, 해석적 연구에서는 사회 현상을 보다 주관적인 관점에서 이론적 해석을 하는 것을 목적으로 한다. 실증적 연구는 주로 실험실 연구, 현장조사, 설문조사, 이차자료를 통한 분석, 또는 사례연구 등이 있으며, 해석적 연구에는 사례연구, 현상 분석학, 민족지학 등이 있다. 사례연구는 실증적, 해석적 연구를 수행할 수 있는 연구 설계에 해당된다. 즉, 어떠한 특정 방법이 모든 사회과학 연구 적절하다고 할 수 없기 때문에 연구의 목적에 따라서 설계해야 한다. 포커스그룹(focus group)이나 실험실 연구는 탐색적 연구(exploratory research)에서 주로 쓰이고, 민족지학(ethnography)은 기술적 연구(descriptive research)에 적합하여 주로 쓰이는 방법이다. 더욱 자세한 설명은 9 장부터 12 장에서 하겠다.

실험연구(experimental study)는 원인과 결과를 서로 다른 시간에 측정함으로써 인과관계를 실증적으로 검증하는 방법이다. 연구에서 종속변수에 영향을 미치는 독립변수를 구분하여 실험집단에 자극(treatment)을 주어 피실험집단(control group)과의 비교 연구를 하는 것이 주된 과정이다. 가령, 새로 개발한 약품에 대한 실험실 연구를 하고자 할 때, 우리는 우선 연구 대상을 무작위로 뽑아 자극집단과 통제집단의 두 집단으로 나눈다. 그리고 첫 번째 실험집단에는 약에 대한 자극에 노출 시키고, 피실험집단에는 플라시보 효과를 이용하여 가짜 약을 제공한다. 조금 더 복잡한 연구에서는

약의 사용량, 식사습관 과 같은 다양한 요인을 함께 연구하게 된다. **진실험실연구(true experimental design)**에서도, 연구 대상은 무작위로 서로 다른 집단으로 나뉘게 된다. 만약, 연구대상을 분별하는 과정이 무작위화 되어있지 않은 경우를 **준실험설계(quasi experimental design)**라 한다. 실험연구는 대학교 내 연구실과 같이 인위적으로 통제 된 실험실 조건에서 진행되거나(laboratory experiment), 실제 사건이 자주 발생하는 현장(field experiment)에서 실시된다. 실험실 연구에서는 연구자들의 연구하고자 하는 변수와 외생변수에 대한 통제가 가능하나, 현장실험에서는 그렇지 않다. 따라서, 실험실 연구에서는 내적 타당성이 높을 가능성이 큰 반면, 현장실험은 외적 타당성이 높을 수 밖에 없다. 실험을 통한 수집된 데이터는 양적통계기법을 활용하여 분석 된다. 즉, 변수들에 대한 통제가 전제조건이 되는 실험실 연구에서는 각 변수들의 영향에 대한 민감도를 측정할 수 있기에 높은 내적 타당성을 갖는다는 장점을 가지고 있지만, 실험실 밖 현장에서는 다양한 외생변수들에 노출 되어있는 복잡한 사회 현상을 검정하기 때문에 낮은 외적 타당성을 갖는 다는 약점을 가지고 있다. 아울러, 연구자는 외생변수에 대한 구분과 제대로 통제되지 못한 상황에서의 실험연구는 내적 타당성에 큰 문제가 된다.

현장조사(field survey)는 비실험형태의 연구설계로 독립변수나 실험 자극에 대한 조작을 거치지 않고, 통계적 수단을 통해 변수들의 효과를 측정한다. 현장조사에는 설문 조사와 심층 면접 등을 통해 연구대상의 실제 행동, 신념, 처해진 상황 등에 대해 파악한다. **횡단면적 현장조사(cross-sectional field survey)**는 단순 설문문항(single questionnaire)을 통해 독립변수와 종속변수를 동시에 측정하는 방법인 반면, **추적연구(longitudinal study)**에서는 일정한 시간을 두고 독립변수를 측정 한 뒤에 종속변수를 측정하여 측정 시간을 달리한다. 현장조사는 자료의 수집이 현장에서 이루어지기 때문에 많은 변수를 측정하고 다양한 시각과 이론을 적용할 수 있어 외적 타당성을 보장 할 수 있다는 것이 큰 강점이 있다. 그러나 이러한 조사환경이 연구대상의 반응을 측정하는데 다양한 편견이 개입될 수 있다. 그 결과 인과관계를 측정하는데 큰 영향을 미치게 되어 연구의 내적 타당성이 상대적으로 결여 될 수 있다는 문제가 존재한다.

이차자료분석(secondary data analysis)은 연구자에 의해서가 수집된 자료가 아닌 자료에 대해 재해석 하는 것을 말한다. 예를 들어, 정부기관 위탁 자료처, UN 개발계획프로그램(United Nations development program)과 같은 정부자료조사기관이나, 다른 연구에서 조사된 자료, 혹은 주식시장 및 e-bay 와 같이 실제 매출이 집계되는 시장 자료 등이 주로 사용 된다. 이차 자료분석 방법은 기존의 연구자가 직접 일차 자료를 수집하였던 통상의 연구들과는 다르다. 일차자료는 일차자료와 비교하여 자료를 수집하기에 비용과 시간이 소모되었지만, 이차자료분석은 연구 주제의 해답을 찾기 위한 단계적 분석방법을 활용하기 때문에 효율적이다. 단, 연구자가 수집하고자 하는 이차자료가 연구의 진행방향과 일치하기가 어렵고, 연구에 적합한 체계를 갖추지 못한 자료일 확률이 높다. 초기의 자료가 수집되었을 당시의 목적이 연구자가 연구하고자 하는 목적과 다를 수 있기 때문이다. 만약 인과관계의 명확하지 않는 과정이 수행된다면, 기간의 타당성(interval validity)에 치명적일 수 있다.

사례연구(case research)는 보다 특정한 상황이나 문제에 대해 심층적인 조사를 하는 것을 말하며, 주로 인터뷰, 심층관찰, 내적·외적 자료를 활용하여 데이터를 수집한다. 사례연구는 특성상 특정 가설을 검증하는 실증적(positivist) 연구나, 이론을 설정하는 해석적(interpretive) 연구방법 모두 가능하다. 이 연구의 장점은 특정 사회현상에 대하여 이전에 발견하지 못하였던 사회, 문화, 정치 등 다양한 분야에 걸친 변수를 찾아낼 수 있다. 분석을 통해 질적 연구로 수행되는 경향이 있다. 결과에

대한 연구자의 관찰능력과 해석능력에 따라 해석의 의미와 정도에 큰 차이를 보일 수 있지만, 한 가지의 사례연구를 통해 일반화하기는 어렵다. 따라서 여러 유사 사례들을 살펴보는 다수사례연구(multiple case study)가 요구된다.

포커스그룹연구(focus group research)는 연구자가 6-10 명으로 구성된 작은 집단을 한 장소에 모여두고 한 시간 반에서 두 시간 정도 특정 현상에 대해 토론하는 것을 지켜보는 연구이다. 토론은 충분한 사전교육을 거친 진행자(facilitator)가 참여자들에게 연구의 목적과 주제를 제시함으로써 시작되는데, 연구자는 참여자들의 생각과 경험을 충분히 반영한 응답을 토론 내에서 이끌어내야 한다. 포커스 집단 연구는 연구대상이 되는 샘플 집단이 크지 않으며 다양한 변수들에 대한 통제가 이루어지기 어렵기 때문에 내적 타당성이 결여될 수 있다. 따라서, 해석적(explanatory), 기술적(descriptive) 연구에 자주 쓰이지만 특히 해석적 연구에 적합한 방법이다.

행동연구(action research)는 복잡한 사회 현상은 가장 잘 이해할 수 있는 방법으로 특정한 ‘행동’을 취함으로써, 현상과 관련된 행동의 영향을 관찰하는 것이다. 연구자는 통상 컨설턴트 나 조직의 한 구성원으로서 참여하고, 실제적으로 새로운 조직 체계 및 기술에 대해 조직 구성원이 어떠한 반응을 보이는지를 관찰한다. 연구자의 특정 행동은 반드시 이론을 바탕으로 행해져야 하며 왜, 어떠한 방식으로 행동을 하였으며 변화된 결과에 대하여 충분한 설명해야 한다. 연구자는 결과가 나타나는 과정을 관찰하고 지속적으로 행동과 결과에 대한 이론적 통찰을 정형화 시키면서 연구를 진행해 나아가야 한다. 초기 이론은 문제되는 상황을 연구자가 특정행동을 취함으로써 성공적으로 해결하였을 때 그 타당성을 갖는다. 또한 지속적인 문제 해결과 통찰력이 다른 행동연구와 구별된 특징을 가지고 있다면 행동 연구는 학문과 실제의 중간다리의 역할을 하는 탁월한 연구방법이 된다. 더불어, 행동연구는 다른 곳에서 찾아볼 수 없는 특수한 사회적 현상을 파악하는데 좋은 방법론이다. 다만, 연구자의 관찰력과 해석력을 요구하기 때문에 연구자에 따라 결과의 일반화가 제한되기도 한다.

에스노그라피(ethnography)은 문화적 맥락에서의 사회현상 연구해야 하는 해석적 연구 설계 방법이다. 연구자가 어떠한 사회 현상을 연구하기 위해서는 해당 사회에 소속이 되어 오랜 시간(약 8 개월~2 년) 동안 그 사회의 가치관 등에 대해 파악하고 현상을 그 맥락에서 이해하여야 한다. 자료는 주로 현상에 대한 관찰 된 공식적·비공식적인 사회의 구성원과의 상호작용을 통해 다양한 반응을 수집해야 하며, 그 해석은 ‘이해(sense-making)’를 포함한다. 연구자는 독자들이 해당 사회에 대해 간접경험을 할 수 있도록 자신의 경험을 아주 상세하게 기술해야 한다. 이 방법론은 사회 현상의 전체적 맥락을 이해할 수 있다는 장점이 있으나, 응답자의 편견 등이 개입 될 여지가 남아있다. 무엇보다, 시간과 비용이 많이 들며, 결과가 해당 문화가 아닌 다른 문화권으로 일반화 하기에는 상대적으로 어렵다는 단점이 있다.

5.4 연구설계의 선택

앞서 설명한 많은 연구 방법 중 어떤 것이 좋은 방법론이라고 말하기는 어렵다. 연구자는 보통 자신이 가장 편하거나 수행하는데 어려움이 없는 연구방법론을 선택하기 마련이나, 무엇보다 연구자 자신이 하고자 하는 목적과 제기된 문제를 해결하는데 가장 적합한 방법을 선택해야 한다. 연구주제가 개념들 간의 관계를 규명하기 어렵고 연구문제의 본질을 파악하고자 한다면, 개별적 단위의 분석이 가능한 포커스 집단(focus group)이나 집단 단위의 분석이 가능한 사례연구(case

study)인 탐구적(exploratory) 연구가 이상적일 것이다. 연구하고자 하는 현상에 대해 충분한 뒷받침이 될 적합한 이론을 찾지 못했거나 새로운 이론을 성립하고자 하는 경우에는 사례연구(case study) 혹은 에스노그래피연구(ethnography)인 해석적(interpretive) 연구가 적합하다. 또는, 어떠한 이론을 활용하여 현상을 설명하기에 적합한지 실증적인 검증이 필요한 경우, 실험설계(experimental design), 설문조사(survey research), 이차자료분석(secondary data analysis)과 같은 실증적(positivist) 연구를 수행한다.

특정한 연구설계 관계없이, 연구자는 설문지(questionnaire), 인터뷰(interview), 관찰(observation), 문서(document), 이차자료(secondary data) 등의 질적(qualitative) 및 양적(quantitative) 자료를 적절히 조합할 필요가 있다. 예를 들어, 양적 자료를 수집하기 위해 설문 항목을 훌륭하게 구성했다 할지라도, 응답자들의 의견을 수집할 수 있는 주관식의 문항을 포함하여 설문지 항목에서 미처 깊고 넘어가지 못한 부분까지 확인할 수 있어야 한다. 마찬가지로, 사례연구(case research)에서 면대면 인터뷰(face-to face interviews)를 동반하여 질적 자료를 수집하는 동안에 양적 자료의 가치 또한 고려되어야 한다. 예를 들어, 조직의사결정에 관한 연구에서 조사자가 조직이 의사결정 하는데 걸리는 시간, 의사결정 참여자 수, 의사결정이 필요한 의제 수 등을 수치상으로 표기 함으로서, 응답자들의 응답으로부터 수집하기 어려운 통계적 수치 데이터를 함께 얻을 수 있다. 따라서 연구자는 연구 목적에 따라 가능한 다양한 방법론을 통해 유의미한 데이터를 수집하고, 결과에 대하여 다양한 해석으로의 가능성을 열어두어야 하는 것이다.

제 6 장 구성개념의 측정

이론적 명제는 추상적인 구성개념 간의 관계로 구성된다. 이론적 명제와 같은 이론을 검증하기 위해서는 각 구성개념들을 정확하고 올바른 과학적인 방법으로 측정 한 후에 구성개념들 간의 관계를 파악해야 한다. 측정(measurement)은 현실세계에 대하여 보다 신중하고 세심한 관찰이 필요하며 경험적 연구방법론의 핵심이 된다. 사람의 나이, 몸무게, 회사규모 등과 같이 사회과학에서 사용하는 구성개념들의 일부는 측정하는데 큰 어려움이 없는 반면 창의력, 편견, 소외감 등은 실제로 측정하기가 어려운 개념구성들에 해당된다. 이번 장에서는 이러한 구성개념을 측정하기 위한 개념과 조작과정을 살펴본다.

6.1 개념화 (Conceptualization)

개념화(conceptualization)란 불분명한 구성개념과 그것의 구성요소들을 보다 간결하고 명확한 용어로 정의하는 정신적 과정이다. 예를 들어, 우리는 자주 ‘편견’이라는 단어를 쓰며 그 단어에 대해 특정 이미지를 마음 속에 가지고 있다. 그러나, 우리는 이러한 일상적인 단어를 정확한 정의와 함께 대답하는 것은 어렵다. 그렇다면, 누군가 특정 인종에 대해 부정적인 발언을 한 것을 인종 차별이라고 정의할 것인가? 동일한 직종 안에서 여성이 남성보다 상대적으로 급여가 적다면 성적 차별이라고 정의할 것인가? 교회를 다니는 사람이 믿지 않는 사람들과 지옥의 관계에 관한 것들을 종교적 편견이라고 할 것인가? 편견에는 여러 가지의 종류가 존재하는지, 또 존재한다면 과연 무엇인지, 이러한 편견의 정도에도 차이가 있는지 이러한 문제들을 해결하는데 가장 중요한 이슈는 바로 편견이라는 구성개념에 대해 정확하게 측정하는 것이다. 따라서 구성개념의 범위를 분명히 설명하는 것 또한 개념화 과정의 일부이다.

사회과학과 관련된 구성개념들의 모호하고 불분명한 특성 때문에 개념화 과정(conceptualization process)은 특히 중요하다. 예를 들어, 사회과학 연구에서 ‘연민(compassion)’이라는 개념을 ‘공감(empathy)’ 혹은 ‘감성(sentimentality)’ 과 같은 단어들과 동일하다고 보아야 하는가? 만약 ‘연민은 공감과 양의 관계에 있다.’ 라는 가정명제가 있다고 하자. 연구자는 여기서 연민(compassion)과 공감(empathy)의 유사한 개념을 분리시켜 이 명제를 실증적으로 검증해야 한다. 만약, 종교적 신념이 깊은 사람들에 비해 믿지 않는 사람들은 사후에 그들의 죄로 인해 지옥으로 가게 된다고 믿기 때문에 만약 그들이 말하는 믿지 않는 사람들을 믿게끔 만들어야 한다고 주장한다면 그들의 행동은 편견에서 비롯된 것이라고 할 것인가 아니면 다른 이들에 대한 연민에서 출발한다고 할 것인가? 이러한 구성개념에 대한 정의는 편견 혹은 상상을 동반한 추론에서 시작되기도 하며 혹은 그것도 할 수 없는 경우가 있다. 하지만 현실에서 이러한 개념은 실제로 존재한다. 가령, 특정 개념에 대해 심리 혹은 정신적인 것과 관련된 구성개념을 다루는 과정을

‘구체화(reification)’라고 한다. 구체화란 구성개념에 대한 정의와 이들을 측정 가능한 측정변수로 조작하는 것에 중점을 둔다. 따라서 연구자들은 구성개념을 어떠한 객관적인 기준보다는 연구 범위 내에서 공유된 합의와 정신적 이미지(개념)에 기반한다고 보아야 한다.

구성개념을 구체화 하는데 있어서 중요한 한 가지는 이것을 일차원 혹은 다차원으로 명시할 것인지 결정해야 한다. **일차원(unidimensional)**의 구성개념은 보통 하나의 관점을 가지고 측정을 하기 때문에 한 가지의 측정 방법이나 검증을 수행한다. 간단한 구성개념의 예로 사람의 몸무게, 풍속 등이 있으며 자부심(self-esteem)과 같은 조금 복잡한 구성개념도 일차원으로 가정 할 수는 있으나, 그 가정 자체가 비현실적이라는 비판을 피할 수 없다. 따라서 **다차원(multidimensional)**의 구성개념은 두 개 이상의 관점을 포함한 복잡한 형태를 가지게 된다. 가령, 학습능력을 정의하기 위하여 수학적, 언어적 능력의 두 가지 관점을 제시한다면 학습능력의 구성개념은 다차원적이라고 말한다. 여기서 수학적, 언어적 능력은 반드시 분리하여 서로 다른 검증방법을 통하여 측정해야 하며, 그 결과 값을 평균적으로 합산하여 학습능력에 대한 전반적인 수치를 나타낸다.

6.2 조작적 정의 (Operationalization)

이론적 구성개념이 정의되면 연구자는 이것을 어떠한 방법으로 측정할 것인가? **조작적 정의(operationalization)**는 구성개념을 측정하기 위한 **지표(indicator)** 또는 항목(item)들을 개발하는 과정을 말한다. 예를 들어, 연구자가 ‘사회 경제적인 지위’와 같은 관찰로 표현하기 힘든 구성개념을 ‘가정의 경제적 수준’에 의해 정의 한다면 응답자들에게 질문 할 “당신의 가정 연간 수입 액은 얼마입니까?”는 이를 측정할 수 있는 하나의 지표가 될 것이다. 사회과학 연구에서 구성개념은 다분히 주관적이며 불분명한 내재적 의미를 포함하고 있기 때문에, 이에 대하여 연구자들은 다수의 지표들을 활용하여 측정한다. 반면, 나이, 성별, 교육수준, 수입 등 인구 통계학적 구성개념은 다수의 지표들이 불필요하기도 한다. 일련의 과정으로부터 연구자는 이 지표들이 얼마나 구성개념을 가장 잘 나타내고 있는지 검증하기 위하여 정확성을 평가해야 한다. 이것을 신뢰성(reliability)이라고 한다.

구성개념이 이론적 측면에서 개념화되었다면 지표(indicator)는 구성개념과 달리 실증적 측면에서 적용되는 사항이라고 생각하면 된다. 하나의 구성개념을 정의하는 실증적 수준에서의 통합된 지표를 **변수(variable)**라고 한다. 변수는 연구에서 어떻게 적용되는가에 따라 독립, 종속, 매개, 조절 변수로 구분된다. 또한 각각의 지표들은 다양한 **속성(attributes or levels)**을 가지고 있으며, 이들 각각은 하나의 **가치(value)**로 표현된다. 예를 들어, ‘성별’ 변수는 여성, 남성의 두 가지 속성을 가지고 있다. 마찬가지로, 고객의 만족도를 조사하고자 할 경우 ‘아주 불만족’, ‘조금 불만족’, ‘보통’, ‘조금 만족’, ‘아주 만족’과 같이 다섯 단계의 척도로 표현하여 측정할 것이다. 이러한 속성들의 가치는 **양적(quantitative)** 수치로 나타내거나 **질적(qualitative)** 비수치 형태가 될 수도 있다. 양적 자료는 회귀분석과 구조방정식 모델인 양적 자료 분석기법을 활용하고, 질적 자료는 코딩과 같은 기법을 통하여 분석된다. 사회과학 연구에서 수 많은 변수들은 양적인 의미를 포함할지라도 주로 질적인 경우가 많다. 예를 들어, 소비자 만족도를 조사할 때, ‘아주 불만족’, ‘조금 불만족’, ‘보통’, ‘조금 만족’, ‘아주 만족’으로 구분하고 다섯 속성을 1부터 5까지의 숫자를 부여한다. 이것은 질적 자료분석하기 위한 정교한 통계적인 방법으로 사용할 수 있다. 이러한 수치는

응답자가 자신의 만족에 대한 평가를 표현하는 것이지만, 비록 연구자가 만족이라는 변수를 양적인 형태로 나타냈을지라도 여전히 그것은 질적인 형태라는 것을 상기해야 한다.

지표는 반영적(reflective) 또는 형성적(formative)으로 구분 된다. **반영적 지표(reflective indicator)**란 구성개념을 ‘반영’ 하고 있는 측정방법을 말한다. 가령 ‘독실함(religiosity)’을 구성개념으로 측정하고자 할 경우, 응답자에게 얼마나 신앙심을 가지고 있는지 그리고 종교 집회에 참여하는가의 여부가 독실함의 반영적 지표이다. **형성적 지표(formative indicator)**는 구성개념을 결정짓고 ‘형성’ 하는 측정 방법을 말한다. 이러한 지표들은 구성개념의 각기 다른 특성과 차원을 나타낸다. 예를 들어, ‘독실함’이 신념, 헌신, 의식적인 차원들로 구성되어 정의되었다면 각 차원들을 측정하는 지표들이 바로 형성적 지표인 것이다. 따라서 일차원적 구성개념은 반영적 지표를 사용하여 측정하며 다차원적 구성개념은 다수의 차원에서 형성된 조합으로 측정한다. 하지만 자부심과 같은 추상적인 구성개념을 측정할 경우 다수의 반영적 지표를 사용하며, 다수의 측면으로 구성된 다차원적 구성개념은 하나 또는 그 이상의 반영적 지표로 측정될 수 있다.

6.3 측정 수준

구성개념에 대한 조작적 정의를 할 때 가장 먼저 해야 할 것은 측정 수준을 어떻게 할 것인가이다. **측정수준(level of measurement)**은 **척도(rating scale)**라고도 한다. 이것은 하나의 지표에 내재된 가치를 의미한다. 예를 들어, 남성과 여성(M and F or 1 and 2)은 ‘성별’이라는 지표에 내재된 두 가지 측정 수준이다. Science 학술지에 실린 심리학자 스텔리 스미스 스티븐(Stanley Smith Stevens 1946)의 ‘측정 척도의 이론(On the theory of Scales of Measurement)’ 연구에서는 과학적 측정 방법의 척도를 명목척도(nominal scale), 서열척도(ordinal scale), 등간척도(interval scale), 비례척도(ratio scales)로 분류하였으며 각 척도의 통계적인 특징은 다음 표 6.1에 기술하였다.

척도	집중경향	통계기법	척도변경
명목 척도	최빈값	chi-square	1 대 1 (동일시)
서열 척도	중앙값	백분위, 비모수통계	단순증가형태 (서열)
등간 척도	산술적평균, 치역, 분산	상관관계, 회귀분석, 변인분석	양의 선형 (유사형)
비례 척도	기하적평균, 조화평균	변동계수	양의 유사관계 (승법적, 곡선형)

주의: 분류가 높은 척도는 아래 분류 척도의 통계적 기법 모두 활용 가능함

<표 6.1> 척도의 특성 및 활용 통계기법

명목척도(nominal scale)는 범주척도(categorical scale)라고도 불리며, 범주화된 데이터를 측정한다. 이 척도는 공통적인 특성을 지닌 지표 혹은 변수들을 측정할 때 주로 쓰이게 된다. 성별(두 가지 특성; 남성, 여성), 산업 분류(제조업, 금융업, 농업 등) 그리고 종교 소속(기독교, 이슬람교, 유대교 등) 등이 명목 척도가 활용되는 대표적인 예이다. 앞선 성별을 측정할 때, 남성은 1, 여성은 2 와 같은 형식으로 임의의 수치를 적용하는 경우 숫자에는 특별한 의미가 부여되지 않게 된다. 1 과 2 사이에 어떠한 수학적 공식이 성립이 되지 않으며, 단지 남성을 M 과 여성을 F 로 표현하는 방식이 되는 것이다. 명목 척도는 단순히 구성개념의 각기 다른 특성에 **이름(name)** 또는 **수준(label)**을 부여하는 것이다. 따라서, 명목 척도의 적절한 중심 값을 나타내는 것은 중앙값이나 평균값이 정의될 수 없기

때문에 최빈값(mode)라고 한다. 주로 사용되는 통계적 기법은 카이스퀘어(chi-square) 혹은 도수분포(frequency distribution)가 있으며, 1 대 1 척도 변경만 가능하다(예: 1 = 남성, 2 = 여성).

서열척도(ordinal scale)는 학생의 GPA 와 시험성적 등을 바탕으로 반에서 등수를 매기는 것처럼 *서열이 정해진(rank-ordered)* 데이터를 측정하는 척도이다. 그러나, 서열 척도에서는 속성의 실제 값이나 그들 간의 유사점 혹은 차이점을 평가 할 수가 없다. 예를 들어, 학생들의 시험성적 순위가 실제 GPA 나 시험 성적과 아무런 의미가 없고 학생 간에 얼마나 잘했는지 상대적인 차이를 측정하는 것이다. 전형적인 예시로는 자연 과학 분야에서 모스경도계(Moh's scale of mineral hardness)가 있다. 모스경도계는 광물의 경도를 측정하는 것으로서, 한 광물이 다른 광물에 흠집을 낼 수 있는가가 기준의 척도가 된다. 즉, 다이아몬드는 지구상에서 모든 광물에 흠집을 낼 수 있어 가장 단단한 광물이라고 결론을 내리는 것이다. 그러나, 광물의 단단한 정도를 수치상으로 나타내거나 하지는 않고, 다른 광물과 비교하여 단단한 정도를 상대적으로 측정하는 것이다. 서열 척도도 명목 척도와 같이 '나쁨', '중간', '좋음' 혹은 '매우 불만족', '조금 불만족', '보통', '조금 만족', '매우 만족' 으로 표현한다. 하지만 응답자의 '조금 불만족' 은 '매우 만족' 보다 만족도가 상대적으로 떨어진다는 의미를 나타낸 것이며 이를 만족의 수준으로 수치화 할 수 없다. 명목 척도의 집중 경향도는 최빈값(mode)이 되며, 평균값(mean)으로 해석 될 수 없다. 따라서, 백분위나 비모수적 분석과 같은 통계 분석은 가능할 수 있으나, 상관관계분석, 회귀분석, 분산분석법과 같은 정교한 분석에는 적합하지 않다.

등간척도(interval scale)는 서열대로 정리(rank-ordered)되어 있어야 할 뿐만 아니라, 서열들 간에 동일한 간격이 존재해야 한다. 가령, 온도(화씨, 섭씨)를 살펴보면 화씨 30 도와 40 도 사이의 간극과 화씨 80 도와 90 도 사이의 간극이 일치한다. 이와 마찬가지로, 설문지의 연간소득에 대한 응답 척도를 살펴 보면 \$0 - \$10,000, \$10,000 - \$20,000, \$20,000 - \$30,000 과 같이 항목 간 간극의 차가 일치하고 각 항목의 중간 값은 각기 \$5,000, \$15,000, \$25,000 과 같이 일정한 간격으로 존재한다. 이것을 등간 척도라고 부른다. IQ 지수도 등간 척도에 해당한다고 볼 수 있다. IQ 지수 또한 100 에서 110 까지 사이의 간극과 110 에서 120 까지 사이의 간극이 같다고 보기 때문이다. 따라서 등간 척도에서는 하나가 다른 하나와 비교될 때 '얼만큼 더' 값을 가지게 되는지를 측정 할 수 있다는 점이 명목 척도와 서열 척도를 구분 가능하게 한다. 중앙 집중화 경향의 측정은 평균(mean), 중앙값(median), 최빈값(mode) 모두를 포함하며, 또한 표준 편차와 같은 분산의 분포도의 측정으로 가능하다. 활용 가능한 통계적 분석방법으로는 명목척도와 서열척도에서 활용되는 일련의 통계적 기법을 포함하며 상관관계분석, 회귀분석, 분산분석법 등이 있다. 척도 변경은 양의 선형함수가 가능하다. 앞서 설명한 것처럼 만족도 측정은 등간 척도가 아니다. 그 이유는 만족도의 수준이 동일한 간극을 갖고 있지 않기 때문이다. 다시 말해서 매우 만족함과 만족의 차이가 보통과 조금 만족의 차이의 수준이 동일하다고 볼 수 없기 때문이다. 그러나, 사회과학 연구에서 이를 잘못 인식하여 항목간의 만족도 간극이 일정하다고 여기는 경향이 있어 서열척도의 분석기법을 사용하는 오류를 범하는 경우가 종종 발생한다.

비례척도(ratio scale)는 명목척도, 서열척도, 등간척도의 특성을 모두 포함하면서, '절대적 영점(true zero point)' 을 가지고 있다. 또한 0 의 값은 실질적으로 해당 구성개념의 부재를 말한다. 자연과학이나 공학 등에서 주로 쓰이는 질량, 비행고도, 충전 정도와 같은 척도들이나, 사회과학에서 주로 활용되는 나이, 제임기간, 회사규모와 같은 변수들이 이러한 비례척도에 해당된다. 예를 들어,

회사 규모의 값이 0 이라고 한다면, 회사 내 사원이나 이익창출 활동이 전혀 없다는 것을 의미하는 것이다. 화씨와 섭씨의 온도척도와 달리, 켈빈(Kelvin)의 온도 척도는 비율척도이다. 켈빈 온도에서 0 의 값 (섭씨 -273.15 도)은 모호한 값이 아닌, 운동 에너지의 값이 0 이 되는 상태를 나타내는 값이기 때문이다. 이러한 척도를 우리는 ‘비례척도’ 라고 하고 이는 두 개의 지점이 비례 상의 유의미한 값을 가지고 있으며 해석 가능하다. 즉, 10 명의 사원을 고용하고 있는 어떠한 회사의 규모는 5 명의 사원을 고용하는 회사의 두 배에 해당한다고 보는 것이다. 동일하게 10,000 명의 사원을 고용한 회사는 상대적으로 5,000 명을 고용한 회사와 다르다는 것이다. 중앙 집중화 경향을 측정하는데 기하평균 및 조화평균을 포함하며 스튜던트화 범위분포(studentized range)와 변동 계수(coefficient variation)와 같은 비율 측정 방법이 가능하다. 비율 척도에서는 모든 통계기법의 활용이 가능하며, 승법적(multiplicative) 또는 대수(logarithmic)와 같은 양의 관계에서 정교화된 척도변경이 가능하다.

위에서 살펴 본 명목, 서열, 등간, 비율의 네 가지 척도를 바탕으로 연구자는 사회과학에서 쓰이는 특정 비례척도를 개발 할 수 있다. 통상적인 비례 척도에는 이진척도(binary), 리커트(Likert), 의미차별(semantic differential), 거트만(Guttman) 척도가 있으며 이에 대한 설명을 살펴보자.

이진척도(binary scale). 이진척도는 두 가지 값으로 구성되는 명목척도이다. 즉, 예-아니오, 참-거짓과 같은 경우를 말한다. 전형적인 이진법에는 ‘정치적 활동’ 의 구성개념을 예로 들 수 있는데, 여섯 항목에 각각 이진법 척도로 나타낸다(표 6.2). 이진 척도는 각각의 항목에서 두 가지의 유의미한 값을 가지고 있으며, ‘예’ 로 응답한 횟수를 세어 0 부터 6 까지의 값으로 변환한 것이 바로 개인의 정치적 활동에 대한 척도가 되는 것이다. 어떻게 이러한 항목을 이끌어내는지에 대한 메커니즘을 이해하기 위해서는 ‘척도(scaling)’ 를 이해해야 한다. 이에 대한 설명은 이후에 좀더 자세히 살펴볼 것이다. 또한 이진 척도는 성별을 남-여, 고용형태를 정규직-계약직과 같은 서로 다른 값들을 가지고 있다. 고용형태의 경우는 실업, 정규직, 계약직, 은퇴의 두 가지 이상의 값으로 표현 가능하다면 이진 척도 측정이 불가능 하지만, 이것은 여전히 명목척도화 된 항목들이라 말할 수 있다.

당신은 국회의원에게 청원서를 써본 적이 있습니까?	예	아니오
당신은 정당에 가입한 적이 있습니까?	예	아니오
당신은 정치적 운동을 위해 기금을 기부한 적이 있습니까?	예	아니오
당신은 국회의원 후보자의 정치활동비를 기부해 본 적이 있습니까?	예	아니오
당신은 신문이나 잡지에 정치성이 있는 글을 기고해 본적이 있습니까?	예	아니오
당신은 다른 사람의 투표의사결정에 영향을 끼친 일이 있습니까?	예	아니오

<표 6.2> 정치적 활동에 대한 이진 척도의 예시

리커트 척도(Likert scale). 이것은 사회과학 연구에서 서열이 정해진 데이터를 분석하는데 주로 쓰이며 렌시스 리커트(Rensis Likert)가 고안한 척도이다. 리커트 척도에서 쓰이는 항목은 응답자들이 특정 항목에 대하여 동의하는 정도를 ‘매우 찬성’ 부터 ‘매우 반대’ 까지 5 단계 혹은 7 단계의 범위를 설정하여 측정한다. ‘직원의 자부심’ 의 구성개념을 측정하기 위해서 여섯 항목과 각각의 항목은 리커트 척도로 구성되었다(표 6.3). 리커트 척도는 총화평정척도(summated scales)라고 하는데 이는 응답자의 응답을 평가하는 단일 평가지표로서 각 문항에 따른 척도 값들을 합산하여 도출하기 때문이다.

	매우 반대	조금 반대	보통	조금 찬성	매우 찬성
나는 직장에 만족하고 있다.	1	2	3	4	5
나는 다른 직장동료와 잘 지내는 편이다.	1	2	3	4	5
나는 직장상사와의 관계에 만족하는 편이다.	1	2	3	4	5
나는 현재 직장에 만족한다고 다른 사람에게 이야기 할 수 있다.	1	2	3	4	5
나의 직장동료는 나를 존중해주는 편이다.	1	2	3	4	5
내가 하는 일은 내 회사의 성과에 기여한다고 생각한다.	1	2	3	4	5

<표 6.3> 직원의 자부심에 대한 항목과 리커트 척도

리커트 항목들은 이진 척도의 항목보다 응답자의 중립적인 응답에 대해 더욱 정교하게 측정을 가능케 한다. ‘앵커(anchors)’ 라고 주로 불리는 3-9 개의 값들이 사용되고 응답자의 ‘중립성’ 을 측정하기 위해 척도의 개수는 홀수로 구성한다. 몇몇 연구에서는 척도의 중간 값을 제외시켜 짝수 개의 척도를 이용하여 응답자로 하여금 찬성 혹은 반대의 의견 표명을 할 수 밖에 없게끔 강요하는 방법을 사용하기도 하는데, 이는 응답자의 의견을 정확하게 전달 할 수 없기 때문에 바람직하지 못한 방법이다. 리커트 척도에서 중요한 것은 항목(item)들과 지표(indicator)들이 달라진다 하더라도 ‘매우 반대’, ‘매우 만족’ 과 같은 앵커(anchors)는 그대로 여야 한다는 점이다. 앵커(anchors)들 사이의 동일한 간격을 가져야 할 필요가 없기 때문에 이것이 서열척도에 해당한다고 볼 수 있으나 연구자들은 등간척도로 여긴다.

의미차별척도(semantic differential scale). 응답자들이 하나의 항목에 대해 그들의 의견이나 감정을 서로 다른 한 쌍의 형용사 형태로 응답하는 복합척도이다. 예를 들어, ‘공중의료보험 제도에 대한 태도’ 라는 구성개념은 표 6.4 와 같이 네 가지 항목을 사용하여 측정 할 수 있다. 리커트 척도와 같이 전반적인 척도 수치는 각 항목들의 수치의 합으로 나타낼 수 있다. 여기서 주의해야 할 점은 리커트 척도에서 서술(statement)의 변화가 있더라도, 앵커(anchor)는 항목(item)과 관계없이 동일하게 구성했다면, 의미차별척도에서 서술(statement)은 구성개념이며 앵커(anchor)는 항목(item)에 따라서 한 쌍의 형용사 형태로 변형해야 한다. 따라서 의미차별척도는 인간 태도나 행동, 사물에 대한 감정 등을 측정하는데 훌륭한 연구 방법이라고 알려져 있다.

공중의료보험제에 대해 어떻게 생각하십니까?						
	아주	조금	보통	조금	아주	
좋은	<input type="checkbox"/>	좋지 않은				
유용한	<input type="checkbox"/>	유용하지 않은				
배려적인	<input type="checkbox"/>	배려가 없는				
흥미로운	<input type="checkbox"/>	지루한				

<표 6.4> 공중의료보험제도에 대한 태도에 대한 의미차별척도

거트만척도(Guttman scale). 이 척도는 루이스 거트만(Louis Guttman)에 의하여 개발되었으며, 구성개념에 대한 약한 강도부터 강한 강도까지 강도에 따라 서열적으로 배열한 일련의 항목들을 활용하는 척도이다. 예를 들어, ‘이민자에 대한 태도’ 에 대한 구성개념은 다섯 개의 항목으로 측정

가능하다(표 6.5). 거트만 척도에서 각 항목들은 강도에 따라 가중치가 부여되기 때문에, 다양한 가중치를 포함하는 항목들이 있으며 각각의 응답은 합쳐져 종합적인 관찰 값을 나타낸다.

다음 이민자들에 대한 의견에 당신은 어떻게 생각합니까?		
당신은 이민자들이 당신의 나라의 국적을 얻는 것에 거리낌을 느낍니까?	예	아니오
당신은 이민자들이 이웃에 사는 것이 거리껴 지십니까?	예	아니오
당신은 이민자들의 이웃이 되는 것을 꺼립니까?	예	아니오
당신은 이민자들을 친구로 사귀는 것이 꺼려질 것 같습니까?	예	아니오
당신의 가족 중 누군가가 이민자와 결혼을 한다면 꺼려질 것 같습니까?	예	아니오

<표 6.5> ‘이민자들에 대한 태도’에 대한 거트만 척도

6.4 척도화 (Scaling)

앞서 우리는 구성개념의 설명하기 위해 고안된 항목(item)과 지표(indicator)에 따라 응답자의 반응을 측정하는 방법을 살펴보았다. 그렇다면 각각의 지표들을 설정하는 과정인 척도화(scaling)를 자세히 살펴보려고 한다. 엄밀히 말해, **척도화(scaling)**는 양적으로 혹은 측정 가능한 단위를 관측하기 어려운 구성개념에 대하여 질적 판단을 연관시켜 측정하는 과정의 일부분이다. 척도화에 대하여 스티븐(Stevens 1946)은 ‘일정한 규칙에 따라 수의 의미를 부여하는 것’이라고 설명하였다. 명확한 용어를 추상적인 개념을 사용하여 측정하는 이러한 과정은 실증적인 사회과학 연구에서 가장 어려운 작업 중 하나로 남아 있다.

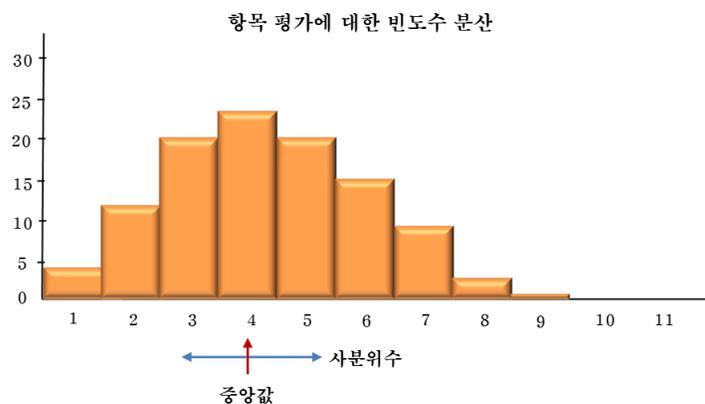
척도화 과정의 결과물은 바로 **척도(scale)**이며 이것은 구성개념에 대한 항목과 지표를 측정하기 위한 실증적인 구조이다. 위에서 설명 하였듯이 척도(scales)는 비례척도(rating scales)와 조금 다른 맥락에서 이해해야 한다. 비례척도는 주어진 항목 내에서 명목척도인 예-아니오, 등간 척도인 ‘매우 만족’, ‘매우 불만족’을 응답자의 반응으로 측정하고 분류를 한다. 비례척도를 어떤 하나의 문장에 대해 적용하는 것을 척도화라고 하지는 않는다. 다시 말해서 척도화란 비례척도를 항목에 적용시키기 전에 척도 항목들을 개발하는 과정이다.

척도는 구성개념의 일차원(unidimensional) 또는 다차원(multidimensional)에 따라 결정 된다. 일차원 척도는 단편적으로 높고 낮음의 정도로 구성된 단일 척도로 구성개념을 측정한다. 물론, 어떠한 척도는 다수의 항목을 포함하고 있으면서도, 동일한 관점을 사용한 항목들로 구성되어 있다. 특히 사회과학의 구성개념들을 측정할 경우 이러한 현상이 종종 보여지는데, 예를 들면 ‘자부심’은 높고 낮음과 같이 일차원적 척도를 갖는 경우가 있다. 한편, 다차원 척도는 구성개념의 서로 다른 관점을 가지고 측정하기 위한 검증과 다양한 항목들을 적용하여 다차원적 구성개념을 종합적으로 측정한다. 예를 들어, 학문적 소질(academic aptitude)이란 개념을 측정하는데 있어 학생의 수리적, 언어적 능력의 두 가지 측면으로 검증을 하고자 한다면, 두 검증의 결과로 합산한 점수가 학문적 소질의 최종 측정 값이 될 것이다. 사회과학 연구에서의 대부분의 척도는 일차원적이며 이는 세 가지 접근법을 통해 형성된다.

일차원 척도 방법론은 20 세기 초반에 널리 알려지고 정립되었다. 일차원 척도에서 가장 많이 알려진 세가지 방법론은 (1) 서스톤의 등현간격기법(Thurstone’s equal-appearing scaling method) (2)

리커트의 총화척도(Likert's summative scaling), (3) 거트만의 누적척도(Guttman's cumulative scaling)이다. 세 가지 접근법은 비슷한 점들이 가지고 있지만, 가장 큰 차이점은 각 척도의 비율을 결정짓는 기준과 통계적 방법을 사용한 최종 항목들을 선택하는 것이다.

서스톤의 등현간격기법(Thurstone's equal-appearing scaling method). 초기 척도 이론을 정립한 학자 가운데 1925년 루이스 서스톤(Louis Thurstone)은 등현간격기법(method of equal-appearing intervals)을 소개하였다. 이 방법론은 구성개념에 대한 정확한 정의를 바탕으로 측정 가능한 척도를 개발하는 것이다. 여기서 항목들은 측정될 구성개념에 대해 잘 알고 있는 전문가에 의해 결정된다. 하나의 구성개념과 관련된 약 80개에서 100개 정도의 후보 항목을 추려내고, 이를 응답자들이 동의-반대의 형태로 응답할 수 있도록 문장으로 구성한다. 그 다음, 구성개념을 잘 표현하고 있는 항목들에 대한 판단을 할 수 있는 패널들을 구성한다. 패널들은 학문적으로 구성개념에 대한 훈련이 선행되어 있거나, 해당 연구 현상에 대해 어느 정도 인지하고 있는 응답자를 무작위로 추출한다. 그리고 패널들은 각 항목들이 구성개념을 얼마나 잘 표현하고자 하는지를 자신의 의견과 비교하여 1부터 11까지의 척도로 나타낸다 (1: 연결성이 없음, 11: 연결성이 높음). 그리고 그 결과를 바탕으로 각 항목들의 중앙값과 사분위수(inter-quartile range)를 계산하고 히스토그램의 형태로 나타낸다(그림 6.1). 결국 척도 항목은 중앙값의 범위를 중심으로 간격이 일정한 형태로 선택되고, 그 항목들을 빈도수가 높은 중앙값과 동일한 집단에 포함되어 가장 작은 사분위수에 존재하는 항목을 선택한다. 여기서 연구자는 항목 선별을 위해 통계적 분석에 완전히 의존하기 보다는 항목에 따른 각각의 수준(level)과 적절한 문장으로 구성되어 있는지 확인하는 분석방법이 응답자들의 명확한 이해를 돕는다. 척도 항목의 중앙값은 구성개념을 표현하는 종합적 척도 값에 사용되는 가중치를 의미한다. 1부터 11까지의 일정한 간격으로 구성된 자(ruler)와 비슷하기 때문에 이 기법을 등현간격기법이라고 한다.



<그림 6.1> 서스톤 척도 항목의 히스토그램

서스톤(Thurstone)은 일차원 척도법으로 두 가지 방법을 더 제시하였는데, 순차적등간기법(method of successive intervals)과 쌍대비교기법(method of paired comparisons)이 그것이다. 두 방법은 패널에게 항목에 대해 점수를 매기는 과정을 요구하는 것을 제외하고 등현간격기법과 유사하다. 예를 들어, 쌍대비교기법의 경우, 패널이 구성개념을 설명하고 있는 각각의 문장에 대해 짝을 지어 평가를

하게끔 하는 것이다. 하지만 이 접근법에서 평가해야 할 항목들이 많다면 등현간격기법에 비해 다루기 불편하며 시간이 많이 걸린다는 단점이 있다.

리커트의 총화척도기법(Likert's summative scaling method). 리커트의 방법론은 머피와 리커트(Murphy and Likert 1938)의 연구에서 발전된 일차원 척도로, 가장 널리 사용하고 있는 세가지 척도 관점을 소개했다. 서스톤(Thustone)의 척도 기법과 동일하게 리커트(Likert)의 척도 기법도 구성개념에 대한 정확한 정의를 바탕으로 80 개에서 100 개 정도의 후보 항목을 설정하는 데서 출발한다. 각 항목들은 패널들에 의해 1 부터 7 까지의 척도로 평가 되며, 1 은 '매우 반대', 7 은 '매우 찬성' 을 나타낸다. 일련의 항목들은 다음과 같은 단계들을 거치며 최종 척도로 선택된다. (1) 개별 항목의 점수를 합한 값과 각 항목의 점수에 대한 이변상관계수를 계산한 뒤, 낮은 점수(예, 0.60 이하)의 항목들은 제외시킨다. (2) 높은 사분위수와 낮은 사분위수에 해당하는 항목들의 평균을 내고 그 평균 값을 비교하는 t 검정을 수행한 후에 높은 t 값을 갖는 항목을 채택한다. (3) 마지막으로 연구자는 앞선 단계의 결과를 바탕으로 상대적으로 적은 10 개에서 15 개의 항목을 추출한다. 리커트(Likert)의 방법론은 모든 항목에 동일한 가중치를 적용하기 때문에 항목들에 대한 응답자들의 반응은 종합적으로 합산이 가능하다. 따라서 이 방법론은 총화척도기법으로 명명되었다. 여기서 반드시 주의해야 할 점은 구성개념을 측정하려는 항목들 가운데 반대방향의 의미를 가지는 항목들이 존재한다면, 코딩도 1 은 5 로, 2 는 4 로 변형한 후에 종합해야 한다.

거트만(Guttman)의 누적척도기법(Guttman's cumulative scaling method). 거트만(Guttman 1950)이 고안한 누적척도기법은 에모리 보가더스(Emory Bogadus)의 사회적거리이론을 기반으로 두고 있다. 즉, 인간이 사회적 관계에 참여하려는 의지는 다양한 강도의 정도로 나타나는데, 이를 최소한의 강도-최대한의 강도까지 배열한 일련의 항목들을 사용하여 측정한다. 기본적인 개념은 하나의 항목에 동의한 사람은 나머지 항목에 대해서도 동의를 하게 된다는 것이나, 실질적으로 완벽히 누적된 형태로 나타나는 항목들의 군집은 찾기가 어렵다. 척도도식법(scalogram analysis)을 활용하여 집단 지어진 항목들이 누적도와 연관성이 얼마나 있는지를 검증한다.

이전에서 설명한 척도 기법들과 마찬가지로 거트만의 척도 기법도 구성개념에 대한 명확한 정의를 바탕으로 한 후보 항목들의 집단을 설정하는 것으로부터 시작한다. 이를 판단하는 패널 집단이 후보 항목의 선호도에 따라 예-아니오로 응답을 하게 된다. 그 다음, 모든 후보 항목에 대한 패널들의 응답을 매트릭스나 표로 정리한다. 매트릭스는 '예' 의 응답률을 내림차순으로 정렬하고 동일한 숫자의 '예' 는 최소한의 동의한 숫자를 바탕으로 왼쪽부터 오른쪽으로 정렬한다(표 6.6). 따라서 척도는 왼쪽부터 시작하면서 오른쪽 방향으로 누적된다. 그러나, 누적이 되지 않는 예외의 경우도 생기기 때문에 척도를 전체적으로 누적하지 못하기도 한다. 항목들을 정렬하기 위해서는 척도도식법(scalogram analysis)이라는 데이터 분석기법을 사용한다. 또한, 응답자들이 개별 항목에 대한 점수와 종합적인 점수를 도출하기 위해서 통계적인 분석기법을 사용한다.

응답자	항목 12	항목 5	항목 3	항목 22	항목 8	항목 7
29	Y	Y	Y	Y	Y	Y	
7	Y	Y	Y	-	Y	-	
15	Y	Y	Y	Y	-	-	
3	Y	Y	Y	Y	-	-	

32	Y	Y	Y	-	-	-
4	Y	Y	-	<u>Y</u>	-	-
5	Y	Y	-	-	-	-
23	Y	Y	-	-	-	-
11	Y	-	-	<u>Y</u>	-	-

Y 는 행렬 안에서 완전 누적이 되지 않는 예외를 표시한 것

〈표 6.6〉 거트만 척도를 이용한 분류평가 매트릭스

6.5 지표 (Index)

지표(Index)는 일련의 규칙과 공식들을 통한 **복합적 구성개념(multiple constructs)** 혹은 **구성요소(component)**를 종합적으로 합산한 점수를 말한다. 하나의 구성개념을 동일한 관점으로 혹은 서로 다른 관점으로 측정하는 지에 따라 지표는 척도와 다른 특성을 가지고 있다. 가장 잘 알려진 지표의 예로는 미국 노동부의 노동통계청에서 매달 산정하는 소비자물가지표(consumer price index: CPI)가 있다. 소비자물가지표(CPI)는 소비자들이 통상적으로 재화와 서비스를 얼마나 지불하는지를 여덟 가지의 대분류(음식, 주거, 의복, 교통, 보건, 여가, 교육, 소통 및 그 외)와 200 개의 소분류로 세분화하여 조사하는 것이다. 미국 정부에서는 매달 80,000 개 항목의 물가를 조사하고, 각 지역의 물가지수와 구매비율을 고려하여 가중치를 계산하게 된다. 그리고 전문 분석가들이 이를 합산하여 전반적인 물가지수에 대한 공식과 패턴을 도출하게 된다.

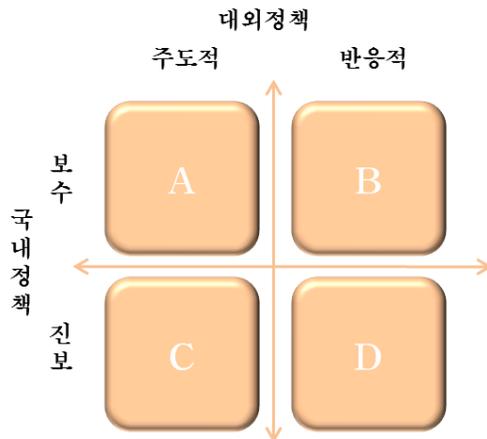
또 다른 예시로는 사회-경제적 지위(socio-economic status: SES)가 있으며, 이를 던칸(Duncan)의 사회-경제적 지표(socioeconomic index: SEI)라고도 한다. 이 지표는 수입, 교육, 직업의 구성개념을 종합한 지표이다. 수입은 달러로 측정되고, 교육은 교육기간 또는 학위수여 여부, 그리고 직업은 범주화된 상태로 세분화한다. 이렇게 다양한 각기 다른 측정치들이 전체적인 사회-경제적 지표를 구성하게 되며, 직업교육(직업을 가진 사람들의 1 년 이상 고등 교육을 받은 백분율)과 직업수입(직업을 가진 사람들의 연간 소득 이외의 특정 수당을 받는 백분율)에 가중치를 복합적으로 활용한다. 그러나, 사회-경제적 지위의 지표 측정에 대해서는 학계에서 많은 논란이 제기되고 있다.

지표를 생성하는 과정은 척도의 그것과 유사하다. 첫째, 지표와 구성요소들에 대해서 개념화 혹은 명확한 정의가 있어야 한다. 지표를 정의하는 것이 가능하더라도 실제 판단하는 이들이 어떻게 생각하느냐에 따라 어떤 구성개념이 지표에 포함해야 하는지 말아야 하는지 다르게 나타날 수 있기 때문이다. 예를 들어, 사회-경제적 지위 지표에서 수입이 교육이나 직업과 관련이 없다고 할 경우, 지표를 구성하기 위해 이 세 가지 구성요소 중 하나만을 포함시킬 것인가, 아니면 모두를 포함시켜야 하는가? 이것은 이론과 선행연구, 전문가와의 충분한 의견을 수집하여 해결할 수 있다. 둘째, 각 구성요소를 조작적 정의를 내리고 측정한다. 예를 들어, 특히 일정 시간에 따라 직업군이 변화를 하기 때문에 직업에 대해서 어떻게 분류를 할 것인지 고려해야 한다(예, 웹 개발자는 인터넷의 발전 이전에는 직업으로 존재하지 않았다.) 셋째, 지표의 점수를 산출하기 위해서는 명확한 규칙과 공식이 존재해야 한다. 이 과정은 주관적인 입장이 많이 개입될 가능성이 높다. 마지막으로, 지표로부터 산출된 점수는 기존에 존재하거나 새로운 데이터를 기반으로 검증해야 한다.

지표와 척도는 단순한 수적인 표현 혹은 구성개념을 표현하는 값으로 구성되지만 실상 많은 차이가 존재한다. 첫째, 지표는 서로 다른 특성을 가진 구성요소들의 집합체이다. 예를 들면, 사회-경제적 지위 지표가 소득, 교육수준, 직업으로 구성되어 있다. 그러나 척도는 일련의 비슷한 항목들을 리커트의 5 점 척도와 같이 동일한 비율 척도를 사용하여 구성한 것이다. 둘째, 지표는 가격, 소득과 같이 객관적으로 측정 가능한 값으로 구성된 반면 척도는 태도, 편견, 자부심 등과 같이 주관적이고 판단적인 측정을 바탕으로 설계된다. 어떤 학자들은 척도를 산출하는 방법론의 정교함이 척도와 지표를 구분 짓는 기준이라 주장하고, 한편에서는 지표를 측정하는 방법론 또한 정교함에 있어서 우열을 가리기 힘들다고도 주장한다. 그럼에도 불구하고 지표와 척도 모두 사회과학 연구에 있어 중요한 방법론 수단임에는 이견이 없다.

6.4 유형분류체계 (Typology)

척도와 지표는 일차원 구성개념에 대한 서열측정 수단이 된다. 그러나 연구자들이 두 개 이상의 구성개념을 포괄하는 새로운 범주를 개발하기도 하는데 이를 **유형분류체계(typology)**라고 한다. 척도 및 지표와 달리, 유형분류체계는 명목적 변수만을 포함한 다차원 측면들로 구성된다. 예를 들어, 신문에서 국내·외 정책에 대한 논조를 분석하여 신문사의 정치적 이데올로기에 관한 유형분류를 새로이 창출 할 수 있다(그림 6.2). 이러한 과정을 통해 신문사의 ‘정치적 성향’을 A 부터 D 까지로 네 가지 유형으로 구별 가능하고 심지어 다른 속성들에 따라 분류 모델을 개발할 수도 있다.

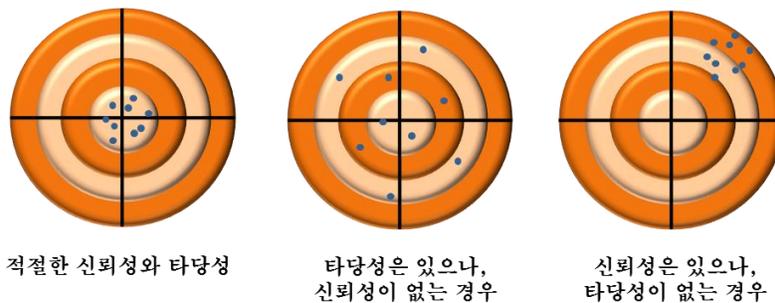


<그림 6.2> 신문의 다차원 유형 분류 체계

제 7 장 신뢰성과 타당성

사회과학 연구에서 구성개념을 측정하는 과정의 어려움을 제 6 장을 통해 살펴보았다. 실생활에서 존재하지 않는 모호하고 추상적인 개념에 대해 측정하는 것은 쉽지 않다. 더욱이 다차원의 구성개념이라면 측정은 더욱 어려워질 뿐만 아니라 연구자들이 지향하는 의미의 정확도도 떨어지게 마련이다. 따라서, 연구자들은 이러한 척도를 검증할 때 다음 사항들을 꼭 유의해야 한다. (1) 실제로 관측할 수 없는 구성개념에 대한 척도들이 실제 측정이 가능한 것인가이다. 다시 말하면 척도가 ‘유효(valid)’ 해야 한다. (2) 구성개념에 대해 일관성 있고 정확하게 측정하였는지, 다시 말해 척도가 ‘신뢰할 만한(reliable)’ 해야 한다. 척도의 측정에 있어서 신뢰성과 타당성은 ‘심리 측정적 속성(psychometric properties)’ 이라고 불리며 사회과학 연구에서 측정 과정을 평가하는 적절성 및 정확성을 뒷받침하는 중요한 개념이다.

측정의 신뢰성과 타당성은 늘 함께 공존하는 것은 아니다. 즉, 구성개념을 일관성 있게 측정하는 과정이라도 잘못된 구성개념을 측정한다면, 신뢰성을 존재하지만 타당하지는 않다. 이와 마찬가지로 올바른 구성개념을 측정하더라도 일관성이 없다면 타당성을 존재하지만 신뢰성이 없다. 신뢰성과 타당성을 가지고 다수의 항목을 통한 구성개념의 측정은 과녁의 중심 근처에 모여 군집화 된다(그림 7.2). 타당성과 신뢰성을 가지고 있는 구성개념은 점들이 과녁의 중심으로 몰려 있다. 하지만 타당성만 있고 신뢰성이 없는 구성개념은 점들이 과녁의 중심을 향해 있지만 몰려 있지 않고 흩어져 있다. 반면, 신뢰성만 있고 타당성이 없는 구성개념은 점들이 과녁의 중심이 아닌 다른 지점에 몰려있는 것을 볼 수 있다. 이와 같이 구성개념의 적절한 측정을 위해서는 신뢰성과 타당성은 필수 조건이다.



<그림 7.1> 신뢰성과 타당성의 비교

7.1 신뢰도(Reliability)

신뢰도(reliability)는 구성개념의 측정이 일관적이거나 의존적인 정도를 나타내는 것이다. 다시 말해서, 연구 환경이 변하지 않는 가정 하에서 특정한 구성개념을 여러 번에 걸쳐 측정하였을 때, 매번 동일한 결과를 도출할 수 있는 정도를 말한다. 신뢰 할만 하지 않는 측정도구를 사용하여 사람들이 당신의 몸무게를 추측한다고 하자. 여기서 사람들은 서로 다른 측정방법으로 측정하기 때문에 일관성이 존재하지 않는다. 따라서 ‘추측(guessing)’으로서의 측정 기법은 신뢰할만하지 못하다. 몸무게를 측정하는데 있어서 가장 신뢰할 만한 측정 방법은 저울을 사용하는 것이다. 저울은 실제 몸무게의 변화가 있지 않는 이상 수 차례 측정하더라도 동일한 결과를 나타내기 때문에 일관성을 가지고 있으며 이것은 곧 신뢰성이 있다고 할 수 있다. 인간의 판단은 시간, 장소, 상황에 따라 달라지기에 측정 자체가 일관성을 갖고 있지 못하다. 따라서 인간의 주관적인 판단은 신뢰도를 떨어뜨린다고 할 수 있다.

신뢰성은 일관성이라는 개념을 포함하고 있지만 정확성까지 보장하는 것은 아니다. 앞선 몸무게의 예시에서 저울의 눈금이 잘못 표기 되었다면, 당신의 몸무게를 정확하게 측정하지는 못하게 됨으로써 그 결과, 타당성을 가지고 있는 측정이라고 할 수 없다. 그럼에도 불구하고, 눈금이 잘못 표기된 저울은 당신의 몸무게에 대해 일관성 있는 결과를 보여주기 때문에 이것은 신뢰성을 가지고 있다고 한다.

사회과학에서 결여된 관찰로부터의 신뢰성 측정 문제에는 여러 가지 이유가 있다. 궁극적인 요인 중 하나는 관찰자 혹은 연구자의 주관적 개입이다. 예를 들어, 회사에서 직원의 도덕성을 직원의 웃는 횟수, 농담을 하는지 등으로 측정한다고 하자. 연구자가 직원들끼리 소통을 할 시간이 없을 때와 여유로운 시기 이렇게 두 번에 걸쳐 진행 되었다면, 두 개의 시나리오로부터 상이한 결과가 나타날 것이다. ‘관찰(observation)’은 질적 측정 기법이다. 반면, 연구에서 양적 측정 기법을 통한 신뢰도를 향상시키는 경우가 종종 있다. 가령 한 달 동안 도덕성의 측정 항목을 횟수로 세어 판단하는 것이다. 물론, 도덕성을 측정하는데 있어서 불만사항이라는 항목이 타당할 수도 있고 아닐 수도 있다. 그러나, 인간의 주관적인 관점의 개입이 적을수록 신뢰성을 더 가지고 있다. 신뢰성을 해치는 또 다른 주요 요인은 모호하고 부정확한 질문을 던지는 것이다. 예를 들어, 연구 대상에게 봉급이 얼마인지를 물어봤을 때, 응답자는 봉급이 일년을 기준으로 대답해야 하는지, 월별을 기준으로 하는 것인지, 아니면 시간당 임금을 말하는 것인지 혼동함으로써 그 결과 서로 다른 기준이 적용된 응답으로 측정기준을 신뢰할 수 없게 된다. 셋째로 꼽을 수 있는 요인은 연구 대상이 현상에 대한 연구 주제가 접해 보지 못하였거나 관심이 없는 경우이다. 예를 들어, 미국 대학 학생들에게 캐나다와 슬로베니아의 관계에 대해 만족하는지에 대한 질문이나, 기업 CEO 에게 사내 결혼이 기업의 성과에 미치는 영향 등에 대한 질문은 응답자로 하여금 주제와 관련해서 연관성 높지 않은 응답을 수집함으로써 연구 결과에 부정적인 영향을 미칠 수 있다.

그렇다면 신뢰성 있는 측정을 개발하기 위해서 어떻게 해야 할까? 사회과학 연구의 측정 과정에서 개입 될 수 있는 연구자의 주관적 개입을 최대한 배제하고 최대한 정확한 배경 지식을 가지고 질문에 응할 수 있는 연구대상으로부터 반응을 수집해야 한다. 여기서 질문 항목을 정확한 단어의 선택과 어렵지 않고 간결한 표현으로 구성하여 응답자가 잘못 해석하는 오류가 없도록 해야 한다.

이러한 방법들은 신뢰성을 보다는 높일 수 있게 된다. 측정도구는 반드시 신뢰성 검증을 해야 하기 때문에 신뢰성을 측정 할 수 있는 다양한 방법들을 소개한다.

평정자간 신뢰도(inter-rater reliability). 관찰자 간의 신뢰도(inter-observer reliability)라고도 하는데, 이는 동일한 구성개념에 대해 측정하고자 할 때, 두 명 이상의 독립적인 평가자(관찰자) 사이에서 일관성 있는 결과를 도출했는지 평가하는 것이다. 통상, 예비조사(pilot test)를 통해서 평가를 하게 되는데 구성개념의 측정 수준에 따라 두 가지 방법이 있다. 측정 해야 할 값이 범주화되어 있다면, 평가자들은 범주를 정의하고 각각의 관찰을 어떤 범주에 포함시킬지, 평가자들 사이의 의견 일치성이 어느 정도인지를 백분위로 나타내는 것이 바로 평정자간 신뢰도이다. 예를 들어, 두 명의 평가자가 100 개의 관찰 결과를 세 범주로 분류하고 그들의 일치성이 75%가 된다고 하면, 평정자간 신뢰도는 0.75 가 된다. 만약 측정이 등간척도 이거나 비례척도로 구성되어 있다면 두 평가자의 측정치를 단순 상관관계(simple correlation)로 도출하여 평정자간 신뢰도를 측정 할 수 있다. 예를 들면, 교내 활동이 매 5분 마다 두 명의 관찰자들로부터 1-7 사이의 응답 척도로 표현하는 경우이다.

검사-재검사 신뢰도(test-retest reliability). 동일한 구성개념에 대하여 두 번의 검사 결과 값들 사이의 일관성을 측정하는 것이다. 검사는 동일한 표본이 사용되지만 서로 다른 시간대에 이루어 진다. 두 번의 측정을 통한 검사 결과가 일정하게 도출되었다면 신뢰도가 있다고 말할 수 있다. 이 때, 각기 두 번의 검증 결과의 상관관계가 검사-재검사 신뢰도 계수가 된다. 여기서 주목해야 할 점은 두 번의 검증 시간 간격이다. 일반적으로, 시간 간격이 길 수록 두 관측 결과가 바뀔 가능성이 높다. 즉, 시간 간격이 짧을수록 검사-재검사 신뢰도가 높아진다.

반분 신뢰도(split-half reliability). 하나의 구성개념의 절반과 나머지 절반의 일관성을 측정하는 것이다. 예를 들면, 주어진 구성개념이 열 개의 항목으로 이루어져 있다면 무작위로 다섯 개의 항목으로 묶은 두 집단을 응답자들에게 배포한 뒤, 반응을 수집하고 점수화한다. 그리고 전체 점수와 각 집단의 점수를 활용하여 반분 신뢰도를 측정한다. 많은 항목을 포함한 구성개념의 경우 두 개의 결과 값이 비슷하게 도출되기 때문에 이 기법은 시스템적으로 많은 항목을 측정해야 하는 점이 있으며 구성개념의 신뢰도는 과대 평가되는 경향이 있다.

내적 일관성 신뢰도 (internal consistency reliability). 동일한 구성개념 하에서 서로 다른 항목들 사이의 일관성을 측정하는 것이다. 가령, 다수의 항목으로 이루어진 구성개념에서 항목들 간에 유사한 정도를 반영하고 있는지를 측정함으로써 내적 일관성을 검증한다. 내적 일관성 신뢰도는 구성개념을 이루고 있는 항목들 간의 평균을 측정하고, 항목-전체 상관관계의 평균이나 크롬바크 알파(Cronbach's alpha) 계수로 산출한다. 예를 들어, 여섯 항목의 척도로부터 15 쌍이 도출되고 그들간의 상관관계를 측정하는 것이다. 항목-전체 상관관계의 평균을 도출하기 위해서 우선, 각 여섯 항목의 값을 더해 '전체(total)항목' 을 만들고 개별 여섯 개의 항목 값과 전체 항목 값의 상관관계를 산출한 뒤 평균을 구한다. 위의 두 신뢰도 측정 방법들은 구성개념을 설명하고 있는 항목의 개수와는 관계가 없다. 크롬바크 알파(Cronbach's alpha)는 신뢰도 평가를 다음과 같은 공식을 통해 산출한다.

$$\alpha = \frac{K}{K-1} \left(1 - \frac{\sum_{i=1}^K \sigma_{Y_i}^2}{\sigma_X^2} \right)$$

여기서 K 는 측정하려는 항목의 개수, σ_X^2 은 관찰된 전체 점수의 분산도, $\sigma_{Y_i}^2$ 은 항목 i 에 대한 관찰된 분산을 나타낸다. 표준화한 크롬바크 알파(standardized Cronbach's alpha)는 다음 공식을 통하여 계산한다.

$$\alpha_{standardized} = \frac{K\bar{r}}{(1 + (K - 1)\bar{r})}$$

여기서 K 는 항목의 개수, \bar{r} 은 항목 간의 상관관계 평균을 의미한다. 이것은 상관관계 매트릭스의 위쪽 삼각형 혹은 아래쪽 삼각형에서의 K 계수의 평균 $(K-1)/2$ 을 의미한다.

7.2 타당성 (Validity)

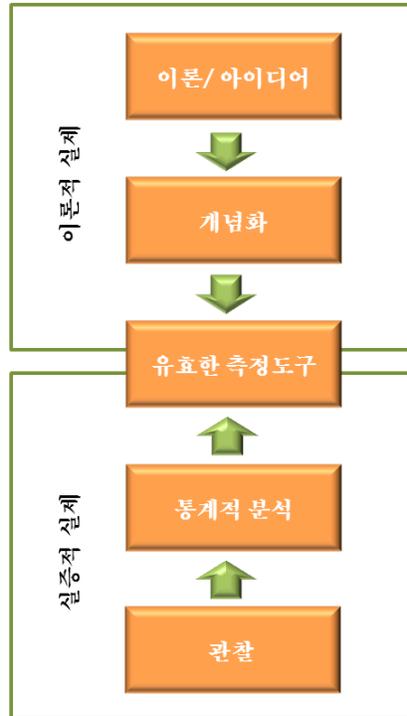
타당성(Validity), 혹은 구성 타당성(construct validity)이라고도 불리는 이 개념은 실제 측정하고자 하는 구성개념을 정확하게 측정하고 있는가에 대한 것이다. 예를 들어, ‘연민’이라는 개념을 측정하고자 할 경우, 실제로 ‘연민’에 대하여 측정을 하고 있는지, 아니면 ‘동정심’과 같이 다른 개념이 측정이 되는지에 대한 것이다. 타당성은 이론적 및 실증적 접근을 통한 평가가 가능하고 가능한 두 가지 접근을 모두 활용하는 것이 바람직하다. 이론적 측면에서의 타당성 평가는 이론적으로 구성개념을 얼마나 잘 조작적 정의로 표현하였는지를 측정한다. 이러한 유형의 타당성을 ‘**번역 타당도(translational validity)**’ 또는 표현 타당도(representational validity)라고 한다. 이 타당성은 안면 타당성 (face validity)과 내용 타당성(content validity)로 구성된다. 번역적 타당도는 주로 전문가로 구성된 패널에 의해 평가된다. 이들은 각각의 항목들이 구성개념의 정의와 잘 이루어졌는지 평가를 하게 되는데 여기서 Q-솔트(Q-sort)라는 질적 기법을 사용한다.

실증적 관찰을 바탕으로 한 하나 이상의 외부 요소들과 측정하고자 하는 구성개념이 얼마나 잘 연결되어 있는지가 타당성의 실증적 검증과 직결되는 문제인데 이를 **기준 타당도(criterion-related validity)**라고 한다. 이것은 수렴 타당도(convergent validity), 판별 타당도(discriminant validity), 동시 타당도(concurrent validity), 예측 타당도(predictive validity)로 구성되어 있다. 반면 번역 타당도(translational validity)는 측정하려는 구성개념이 이론의 개념을 잘 반영하였는지를 평가한다면, 기준 타당성은 이론적 구성개념의 측정이 실제 연구 설계에서 설정한 것처럼 잘 수행되고 있는가를 평가하는 것이다. 또한 관찰된 데이터는 상관관계 분석이나 요인 분석과 같은 통계적 기법을 활용한 양적 분석에 중점을 둔다. 타당성의 이론과 실증적 세계에서 평가는 사회과학 연구에서 충분히 타당성을 검증하기 위한 필요한 방법이다.

다양한 종류의 타당도는 **측정 과정(measurement procedures)**과 관련이 있으며 **가설 검증 과정(hypotheses testing procedures)**과는 구별된다. 즉, 내적 타당도(internal validity: causality), 외적 타당도(external validity: generalizability), 통계적 결과 타당도(statistical conclusion validity)가 있으며 마지막의 통계적 결과 타당도는 다음 장에서 자세히 설명한다.

안면 타당도(face validity). 구성개념을 측정하는 표면 상의 합리적인 측정의 정도를 나타내는 지표(indicator)이다. 예를 들어, 충분한 설명 없이 한 사람의 종교의례 참가 횟수가 그 사람의 종교에 대한 신념을 나타내는 지표가 될 수 있다. 따라서, 해당 지표는 안면 타당도를 가지고 있다. 반면,

도서관에서 책을 대출하는 횟수를 직원의 도덕성의 지표로 측정한다면 이해 가능한 지표로 대표하기는 어렵기 때문에 안면 타당도가 부족하다. 특히, 조직행동 연구에서 사용되는 측정이 안면 타당도를 충분하게 만족하지 못하는 경우가 많이 있었다. 조직이 새로운 지식 및 기술을 받아들이는 능력을 연구 및 개발 투자 집중도라는 하나의 한 지표로 보는 경우가 대표적인 예이다. 만약 우리가 연구하고자 하는 구성개념의 정의가 모호하고 다른 연관된 개념들과 쉽게 구분을 할 수 없을 경우에는 전문가 패널을 고용하여 각 개념의 안면 타당도를 검증 받을 수 있다.

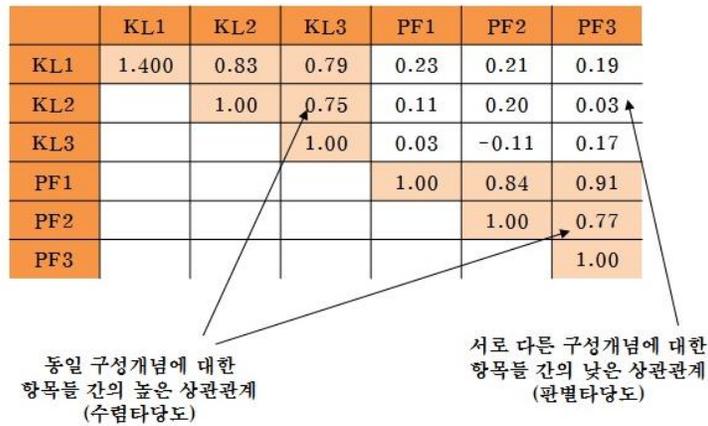


〈그림 7.2〉 타당도 평가에 대한 두 가지 접근법

내용 타당도(content validity). 측정하려는 구성개념의 관련된 내용 도메인과 척도 항목들이 얼마나 잘 일치하는지 평가하는 것이다. 예를 들어, 당신이 ‘레스토랑 서비스의 만족도’ 을 측정하고자 한다면, 음식의 질, 직원의 친절도, 대기시간, 레스토랑의 전반적인 분위기를 포괄하는 구성개념을 내용 도메인으로 정의해야 한다. 그리고 구성개념의 내용 타당도를 만족시키기 위해서 레스토랑의 고객이 위에서 언급한 사항들에 대해 만족하는지를 측정하는 지표를 사용하여 측정해야 한다. 이 접근 방법은 구성개념의 전반적인 내용 도메인에 대한 자세한 설명이 요구되며 자부심이나 지능과 같은 복잡한 구성개념은 더욱 그러하다. 전문가 패널을 활용하여 안면 타당도와 동일하게 구성개념의 내용 타당도 검증이 가능하다.

수렴 타당도(convergent validity). 실제 측정하고자 하는 방법이 측정 방법이 잘 알려진 구성개념을 얼마나 근접하게 측정하는지를 나타내는 것이다. **판별 타당도(discriminant validity)**는 측정에 포함되지 않는 구성개념을 측정하지 않고 있는 정도를 나타낸다. 일반적으로 수렴 타당도와 판별 타당도는 일련의 구성개념들을 종합적으로 측정한다. 예를 들어, 조직 내 지식의 수준이 조직의 성과와 관련이

있다고 할 경우, 연구자가 측정하고 있는 조직 내 지식을 실제 조직 내 지식의 수준으로 측정하고 있는지(수렴 타당도), 조직의 성과를 측정하고 있지는 않는지(판별 타당도)를 동시에 살펴 볼 수 있다. 수렴 타당도는 구성개념의 특정 지표와 다른 지표의 비교를 통해 유사성으로 나타나지만 판별 타당성은 하나의 구성개념과 관련된 지표들이 다른 구성개념의 지표들과의 상이성으로부터 나타난다. 위의 예시로 돌아가서 조직 내 지식의 수준과 조직의 성과를 각각 세 개의 항목으로 측정한다고 하면, 각 항목간의 이진상관관계를 도출할 수 있다. 아래의 상관관계 매트릭스를 살펴보면 조직 내 지식의 수준과 조직의 성과의 개별 항목들 사이에는 높은 상관관계를 가지고 있음을 보여주지만 두 구성개념들의 항목들간의 상관관계는 낮게 나타났다. 따라서 항목 간의 수렴 타당도는 판별 타당도와 함께 동시에 확인할 수 있다(표 7.1).



<표 7.1> 수렴 타당도 및 판별 타당도의 이진 상관관계 분석

수렴 타당도와 판별 타당도를 판단하는 통계적 방법론에는 탐색적 요인분석(*exploratory factor analysis*)이 있다. 이것은 데이터축소기법(*data reduction*)으로써 이진 상관관계 분석을 바탕으로 구성개념의 주어진 항목들을 더 적은 수의 요인들의 집합체로 축소시켜 상관관계를 분석하는 분석방법이다. 여기서 주요인분석(*principal components analysis*)이라는 통계적 기법을 사용하는데 각 요인들은 이론에서 검증하고자 하는 구성개념과 밀접한 관계가 있어야 한다. 추출된 요인들은 고유치(*eigenvalue*)가 1.0 이상이 되어야 하고 요인들은 직각(*orthogonal*) 또는 사각(*oblique*)회전 기법을 사용하여 추출된다. 적절한 수렴 타당도를 측정하기 위해서는 구성개념에 해당하는 각각의 요인들은 0.6 (또는 0.7) 이상의 요인 적재 값(*factor loading*)이 나와야 하고 이를 동일성-요인 적재 값(*same-factor loading*)라고 한다. 반면에 판별 타당도는 0.3(또는 0.4)이하의 요인 적재 값(*factor loading*)이 나타날 경우 크로스-요인 적재 값(*cross-factor loading*)이라고 한다. 수렴 타당도와 판별 타당도를 더욱 정교하게 측정하기 위해서 사용하는 방법은 다속성 다측정 기법 매트릭스(*multi-trait multi-method matrix: MTMM*) 접근법이다. 이 접근 기법은 개별의 구성개념을 두 개 이상의 서로 다른 방법을 사용하여 측정해야 한다. 예를 들어, 설문조사와 개별적 관찰을 병행하는 방법, 교육의 질을 평가하는데 교사와 학부모라는 두 집단의 응답자를 대상으로 하는 설문조사가 있다. 이 방법은 다른 타당도 검증 방법과 비교하여 상대적으로 까다로워 널리 쓰이는 기법이 아니기 때문에 설명을 생략하도록 한다.

	요인 1	요인 2
KL1	0.88	0.13
KL2	0.93	0.11
KL3	0.87	0.03
PF1	0.17	0.93
PF2	-0.03	0.85
PF3	0.07	0.78

↑ 높은 동일성-요인적재량 (수렴타당도)
← 낮은 크로스-요인적재량 (판별타당도)

〈표 7.2〉 수렴 타당도 및 판별 타당도의 탐구적 요인 분석

기준 타당도(criterion-related validity)는 구성개념을 기존에 있던 측정 기준이나 미래에 개발 될 기준들과 얼마나 잘 관련이 되어 있는가를 측정을 하게 되는데, 이것을 각각 동시 타당도와 예측 타당도라고 부른다. **예측 타당도(predictive validity)**는 이론적으로 기대되는 결과와 측정을 통한 미래의 결과가 얼마나 정확히 예측할 수 있는가를 검증하는 정도이다. 학생의 대학 수학능력 시험과 같은 공인된 시험 점수와 높은 GPA 성적을 측정하여 성공적인 대학생활을 예측할 수 있을까? 이것은 이론적으로 구성개념들 간의 네트워크가 얼마나 잘 형성되어 있는지를 살펴보면 알 수 있다. **동시 타당도(concurrent validity)**는 하나의 측정 방법이 동시에 다른 측정 방법과 얼마나 잘 연계성이 있는가를 검증한 것이다. 가령, 학생들의 미분학 성적이 선형 대수학 성적과 높은 상관 관계가 있다고 할 수 있을 것인가? 여기서 두 성적은 모두 수학적 문제를 검증하는 것이기 때문에 동시 발생적으로 관계가 존재한다. 수렴 및 판별 타당성과는 다르게, 동시 및 예측 타당도는 사회과학 연구에서 그것들과 다르게 상대적으로 중요하게 다뤄지지 않는 편이다.

7.3 측정 이론

지금까지 살펴본 신뢰도와 타당도를 **고전적 검증이론(classical test theory)** 혹은 **참값이론(true score theory)**이라고 불리는 수학적 검증으로 알아보도록 한다. 이것은 측정이 실제 어떻게 작동하고, 무엇을 측정하고, 무엇을 측정하지 않는지를 조사하는 심리측정 이론이다. 이 이론에서 측정오류가 없다면 모든 관찰은 정확하게 관찰된 참 값 T(true score T)를 가진다는 것을 가정한다. 하지만 측정오류 E(measurement errors E)는 참 값 T로부터 관찰 값 X(observed score X)의 편차가 생긴다.

$$X (\text{Observed score}) = T (\text{True score}) + E (\text{Error})$$

관찰 값과 참 값의 분산으로 다음과 같은 공식도 성립이 가능하다. 오차의 분산인 $\text{var}(E)$ 를 최소화함으로써 관찰 값 X는 참 값 T를 측정할 수 있다.

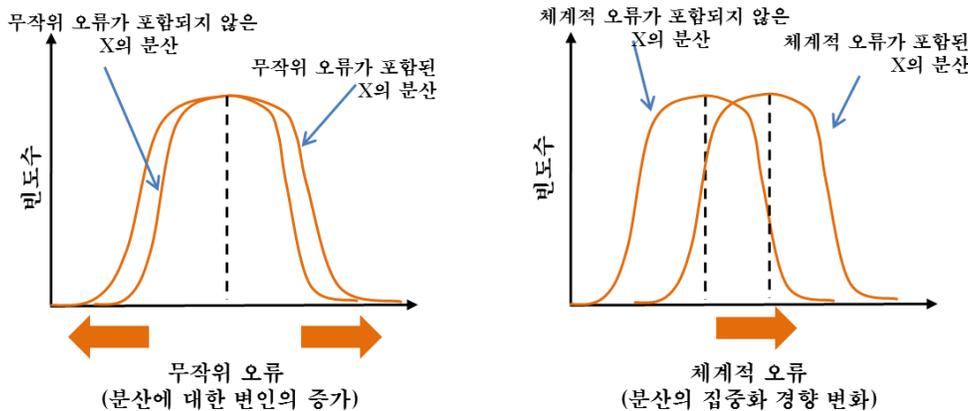
$$\text{var}(X) = \text{var}(T) + \text{var}(E)$$

측정오류는 무작위 오류(random error)와 체계적 오류(systematic error)로 나뉜다. 무작위 오류(random error)는 잘 알려지지 않아 발견하기 어렵고 통제 불가능한 외생요인들이 무작위로 관찰 값에 영향을 미쳐 발생하는 오류이다. 예를 들면, 측정시간 동안 응답자의 기분 좋은 감정상태는 항목을 측정하는 반응에 영향을 미치게 된다. 응답자가 기분이 좋다면, ‘자부심’, ‘만족’, ‘행복’ 과 같은 구성개념에 대한 항목을 측정한 결과 값은 대체적으로 그렇지 않은 응답자에 비하여 긍정적인 반응으로 나타날 것이다. 그러나 연구에서 응답자의 감정상태와 같은 외생요인에 대한 예측과 통제는 굉장히 어렵다. 사회과학 연구뿐만 아니라 조직연구 또한 예외는 아니다. 기업의 성과를 측정하기 위해 정부의 규제나 환경의 변화와 같은 외생 요인이 기업의 성과에 영향을 미치기도 한다. 그러므로 무작위 오류는 측정을 하는 과정에서의 ‘잡음(noise)’ 으로 여겨지며, 위와 같은 특징때문에 연구의 값으로 측정할 수 없어 무시된다.

체계적 오류(systematic error)는 실제 측정과정에서 어떤 체계나 방법의 문제로 인해 개입되는 오류이다. 앞선 기업의 성과 예제에서 제조나 서비스기업보다 금융 위기의 영향을 많이 받은 금융기업을 대상으로 그 영향을 알아보려고 할 때, 표집 자료의 대상을 금융기업만 조사하게 된다면 연구자는 수집된 데이터로부터 금융기업의 체계적인 성과 감소를 관찰하게 될 것이다. 무작위 오류는 표본에서 관찰 값의 긍정 및 부정적으로 영향을 미치거나 혹은 어떠한 영향을 미치지 못할 수도 있지만 체계적 오류는 전체 표본에 긍정적 혹은 부정적 영향을 미친다. 따라서 체계적 오류는 ‘편견(bias)’ 과 혼용되어 쓰이며 반드시 수정해서 연구를 진행해야 한다. 관찰 값이 무작위 오류와 체계적 오류를 모두 포함한다면 참값 방정식은 다음과 같이 표현할 수 있다.

$$X = T + E + E_r + E_s$$

E_r 과 E_s 는 각각 무작위 오류와 체계적 오류를 나타낸다. 통계학적 측면에서 보았을 때, 무작위 오류는 표준편차와 같은 가변성이 관찰된 값의 분포에 변화를 주지만 중앙값에는 영향을 미치지 않는다. 반면, 체계적 오류는 중앙집중경향에 영향을 미치지만 가변성으로 인한 분포에는 영향을 미치지 않는다(그림 7.3).



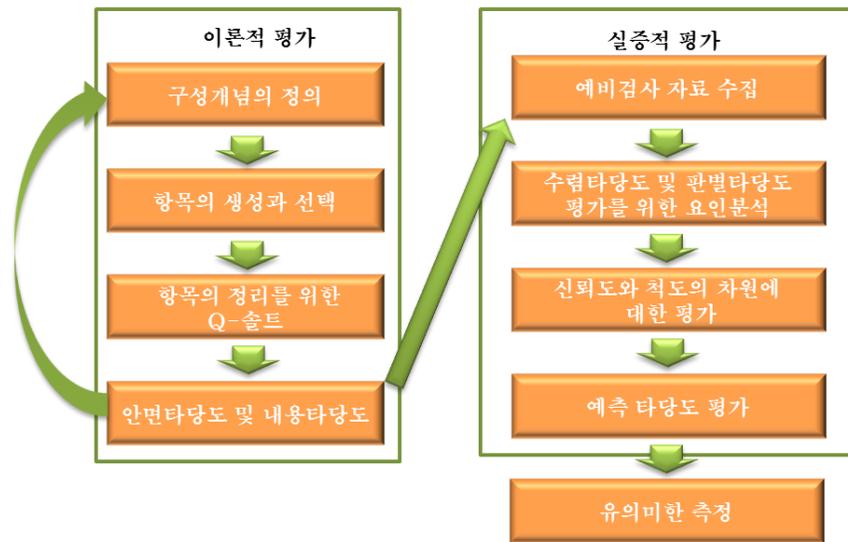
<그림 7.3> 무작위 오류와 체계적 오류의 영향

무작위 오류와 체계적 오류는 측정과정에서 어떠한 의미를 가지고 있을까? 가변성이 증가하면, 무작위 오류는 구성개념의 신뢰도를 감소시킨다. 한편, 중앙집중화 경향의 이동으로 체계적 오류의 구성개념의 타당도를 해치는 결과를 초래하게 된다. 타당도의 문제는 신뢰도의 그것보다 더 심각한 문제를 야기한다. 그 이유는 연구자가 측정하고자 하는 구성개념이 아닌 다른 구성개념의 측정으로 인한 유의미하지 않는 측정이기 때문이다. 그러므로 타당도의 문제점은 통계적 분석에도 영향을 미쳐 문제를 야기하기도 한다. 다음의 공식은 참 값과 관찰 값의 유사성을 비율로 나타낸 신뢰도를 나타낸다. 만일 $\text{var}(T)$ 와 $\text{var}(X)$ 가 동일할 경우, 참 값의 가변성은 관찰 값의 가변성과 같음을 의미하고 여기서 신뢰도는 1.0 이 된다.

$$\text{var}(T)/\text{var}(X) = \text{var}(T)/[\text{var}(T) + \text{var}(E)]$$

7.4 측정 타당성에 대한 종합적 접근

타당성에 대한 평가는 이론적·실증적으로 검증해야 한다. 그림 7.4 는 타당성을 평가하는 순차적 과정을 설명하고 있다.



<그림 7.4> 타당도 측정에 대한 종합적 접근

이론적·실증적 평가를 통한 종합적 접근은 이론적인 실제에서부터 출발한다. 첫째, 구성개념에 대한 개념화이다. 여기서 구성개념을 정의하고 그것을 구성하는 도메인(domain)과 측면(dimension)을 정의해야 한다. 다음, 각 구성개념의 정의 바탕으로 실질적으로 측정할 수 있는 지표(indicator)와 항목(item)을 개발한다. 선행연구를 통해 항목을 선별하는데 많은 정보를 얻을 수 있다. 항목은 단순하며 이해하기 쉬운 형태로 더 명확하고 적절하게 다시 쓰여져야 한다. 이 과정에서 전문가 패널로 구성된 평가단을 통해 각각의 항목을 평가하고 Q-솔트 분석을 한다. 이 분석 과정에서 평가자들은 개념적 정의를 포함한 전체 구성개념의 목록과 각각의 구성개념들을 측정할 지표들로 구성된 인덱스 카드를 받아보게 된다. 인덱스 카드에는 하나의 지표가 표기되어 있다. 그리고 평가자들은 개별적으로 인덱스 카드를 읽고 주어진 구성개념의 정의를 바탕으로 지표들의 가독성,

명확성, 의미를 판별하고 그것들을 군집화한다. 평가자 간 신뢰도(Inter-rater reliability)는 그들이 군집화한 것들 사이의 동의하는 정도를 평가하는 것이다. 만약 불분명한 항목 또는 지표는 평가자들 사이에서 계속 문제가 제기될 것이며 이것은 재조사가 되거나 정확한 의미를 포함한 단어로 표현되거나 연구의 항목에서 제외될 것이다. 개별 구성개념은 10-15 개 정도의 항목으로 선별되어 추후 분석에 활용된다. 분석에 활용할 항목들은 안면 타당도와 내용 타당도를 검증하기 위하여 다시 평가자들에 의해 조사된다. 항목들의 검증이 유의미하지 않다면, 연구자는 개념적 정의를 바탕으로 새로운 항목을 개발해야 한다. 두 세번에 걸친 Q-솔트 과정을 통해 구성개념을 가장 잘 표현하는 항목들 얻을 수 있다.

다음으로 타당도 과정은 실증적인 실재로 이동한다. 연구의 측정도구가 구성이 되었다면, 연구대상을 대표하는 집단을 추려 예비조사를 실시한다. 여기서 수집된 데이터는 수렴 타당도와 판별 타당도를 측정하기 위해 SAS 또는 SPSS 등과 같은 통계 분석 프로그램을 활용한다. 상관관계 분석과 해석적 요인 분석을 하고 요인 적재 값에 적합하지 않는 항목은 이 단계에서 제외한다. 그리고 남아 있는 척도들은 크롬바트 알파(cronbach alpha)와 같은 내적 타당도 검증을 통한 신뢰도를 측정한다. 설명하고자 하는 구성개념이 일차원인지, 다차원인지에 따라 척도의 측면이 달라지게 된다. 다음, 이론상으로 정의한 구성개념의 예측 능력을 평가하기 위해 회귀 분석이나 구조 방적식 모델을 활용한다. 구성개념이 신뢰도 및 타당도의 요건을 갖추었다고 판단이 되면, 연구 결과의 적절성과 정확성을 기대해 볼 수 있다.

위에서 설명한 타당성을 측정하기 위한 종합적인 접근법은 연구자의 시간과 노력을 많이 요구한다. 그럼에도 불구하고 이 단계적 과정으로부터 과학적 연구의 규범에 부흥하는 측정척도를 개발 할 수 있는 것이다. 결점을 가지고 있는 척도를 사용한 연구 결과는 의미가 없기 때문에 척도 타당도는 매우 중요한 과정이며 실증적 연구에서 반드시 포함되어야 한다.

제 8 장 표본추출

모집단에 대한 관찰과 통계적 추론을 위해 관심 모집단의 부분 집합(표본)을 선택하는 통계학적인 과정을 **표본추출(sampling)**이라고 한다. 사회과학 연구는 일반적으로 특정한 모집단의 행동패턴을 추론하는 것과 관련이 있다. 우리는 실행가능성, 비용제한 등의 이유로 관찰하고자 하는 전체 모집단에 대해 연구할 수 없기 때문에, 관심 모집단으로부터 대표성을 지니는 표본을 선택하여 관찰과 분석을 한다. 표본을 통해 얻어진 결론을 일반화 하기 위해서는 모집단을 진정성 있게 대표할 수 있는 표본을 선택하는 것이 매우 중요하다. 대통령 선거 전 CNN / Gallup poll, ABC 그리고 CBS 와 같이 다른 투표 집단에 의해 실시되는 여론조사 혹은 출구조사와 같은 예는 일치성이 없거나 오류가 있는 추론의 근본적인 원인이 된다.

8.1 표본추출 과정



<그림 8.1> 표본추출 과정

표본추출 과정은 몇 가지 과정으로 구성된다. 첫 번째 단계는 목표 모집단을 정의하는 것이다. **모집단(population)**은 연구자가 연구하고 싶은 특성을 지니고 있는 모든 사람들 혹은 항목들(분석 단위)이라고 정의될 수 있다. 분석 단위는 당신이 도출하고자 하는 과학적 추론을 위해 활용되는 개인, 집단, 조직, 나라, 물건이나 다른 독립체 등이 될 수 있다. 모집단은 종종 분석단위가 분명할 때가 있다. 예를 들어, 제조자는 완성품이 특정한 품질을 생산하는데 있어서 필요조건을 충족하는 생산라인에서 제조되어야만 하는지 혹은 재작업이 되어야 하는지를 결정하고자 할 때, 모집단은 생산시설에서 제조된 전체 완성품들의 집합으로 구성된다. 하지만 연구자가 정의하고자 하는 목표

모집단을 설정하기 어려울 때도 있다. 만약 고등학생들이 학교에서 학습의 근본적인 동기를 알고 싶어 한다면, 무엇이 목표 모집단이라 할 수 있는가? 고등학교 학생들, 선생님, 교장선생님, 부모가 목표 모집단일까? 이 문제에 대한 올바른 답은 고등학생들이다. 왜냐하면, 당신은 고등학생들의 성적에 관심이 있는 것이지, 선생님, 부모, 학교의 성과에 관심이 있는 것이 아니기 때문이다. 똑같은 맥락으로, 어떤 바퀴가 결합을 가지고 있는 무엇인지 검증하기 위해서 룰렛바퀴(roulette wheel)에 대한 품질을 분석한다면, 관심 모집단은 한 개의 룰렛 바퀴가 아닌, 여러 개의 서로 다른 룰렛바퀴가 된다.

표본추출 과정의 두 번째 단계는 **표집틀(sampling frame)**을 선택하는 것이다. 이것은 표본이 도출될 수 있는 목표 모집단의 접근 가능성을 의미한다. 예를 들어 당신의 목표 모집단이 전문직 종사자라고 한다면, 전 세계의 모든 직원을 접근하는데 불가능하다. 그렇기 때문에 좀 더 현실적인 표집틀은 연구에 참여하고자 하는 하나 혹은 두 개의 지역회사에 속해있는 개별 직원 명단이 될 것이다. 만약 목표 모집단의 분석단위가 조직이라면, 포춘 500 대 기업의 명단이나 뉴욕 증시에 등록된 S&P 기업의 명단이 적절한 표집틀이 될 것이다.

표집틀이 모집단을 전적으로 대표하지 않는다면, 표본에서 얻어진 결론은 일반화할 수 없다. 예를 들어, 모집단 내의 직원들의 자부심을 연구하고자 할 경우, 목표 모집단이 조직의 전체직원이고, 표집틀이 미국의 중서부 자동차 회사의 직원들이라고 한다면, 그러한 부분 집단으로부터의 도출된 결과는 미국의 전반적인 노동자들의 결과로 일반화 할 수 없다. 왜냐하면 미국의 자동차 산업은 50 여년간 심한 경쟁 압력을 겪어 왔으며, 많은 개편과 감축의 사건이 있었기 때문에 직원들의 사기와 자부심이 낮게 나타나는 결과가 초래했기 때문이다. 뿐만 아니라, 대다수의 미국 노동자들은 자동차 회사가 아닌 서비스 산업이나 소규모 비즈니스 산업에 종사하고 있다. 그렇기 때문에, 미국의 자동차 회사 직원들의 표본은 미국의 노동자들을 대표해 주지 않는다. 이와 마찬가지로, 포춘 500 대 기업은 미국 대기업 500 개를 포함하고 있는데, 이것은 대부분이 중소기업으로 이루어진 미국 기업을 대표하지 못한다. 따라서, 이것 또한 편향된 표집틀이라 할 수 있다. 이와 반대로, S&P 리스트는 S&P 대기업주, 중견기업주, 소기업주 등을 활용하여 중소기업 및 대기업을 선택 가능하게 하지만, 공개적으로 거래되는 회사들만 포함(개인 회사는 포함하지 않음)한다는 점에서 여전히 편향적이다. 하지만 표본이 연구자가 정말 얻고자 하는 정보를 가지고 있는 모집단과 똑같은 필요는 없다. 예를 들어, 연구자가 새로운 금연 프로그램의 성공률을 알고자 한다면, 목표 모집단은 그 프로그램에 참여할 수 있는 세계 각국의 흡연가들이 될 수 있지만, 그것은 알려지지 않은 모집단일 수도 있다. 그렇기 때문에, 연구자는 지역의료시설의 금연센터에 도착한 환자들 가운데 금연 프로그램에 노출되지 않은 환자를 표본으로 할 수 있고, 이러한 경우에는 표본추출은 관심 모집단과 일치되지 않는다.

표본추출 과정에서의 마지막 단계는 잘 정의된 표본추출기법을 활용하여 표집틀로부터 표본을 선택하는 것이다. 표집기법(sampling technique)은 두 가지 큰 범주로 확률(임의)표본추출과 비확률 표본추출로 집단화된다. 확률표본추출은 결과의 일반화가 연구에서 중요할 경우에 이상적이다. 하지만, 비확률표본추출은 특별한 환경에서의 연구일 경우 정당성을 가지고 있다. 위의 두 가지 기법들은 다음에서 논의될 것이다.

8.2 확률표본추출(Probability Sampling)

확률표본추출(probability sampling)은 모집단 내에 있는 모든 단위가 표본으로 선택 받을 기회를 가지고 있는 기법이다. 기회는 표본의 평균이나 표준편차들과 같은 표본통계량이 도출되면 선택확률에 따라, 모집단 매개변수를 편향되지 않도록 추정 할 수 있다. 모든 확률표본추출은 대개 두 가지 속성을 가지고 있다. (1) 모집단 내에 모든 단위는 표본으로 추출될 확률이 0 이 아니다. (2) 표본추출과정은 무작위 선택 과정을 포함한다. 다음 확률표본추출법의 유형들을 소개하도록 하겠다.

단순임의표본추출법(simple random sampling). 모집단의 모든 부분집합(표집들)은 표본으로 선택될 수 있는 동일한 확률을 지니고 있다. 표집들에 존재하는 총 N 개의 단위에서 어떠한 n 단위에 대한 집합을 선택하는 확률은 NC_n 이다. 그렇기 때문에, 가중치 없는 표본통계량은 모집단의 매개변수에 대한 편향 없는 추정 값이다. 단순임의표본추출은 표집들로부터 응답자를 임의적으로 선택하지만, 표집들이 매우 클 경우에는 대개 표에 나타난 임의적인 수나 컴퓨터화된 임의 숫자 생성기를 활용하게 된다. 예를 들어, 엑셀을 활용하여 스프레드시트상에 존재하는 1,000 개의 기업목록 가운데 200 개 기업을 조사하고자 선택한다면, 'RANDO'함수를 활용하여 임의적인 숫자를 생성할 수 있다. 그리고 각각에 대응되는 수를 오름차순으로 정렬하고 그 결과값 가운데 순번 1 부터 200 개까지의 기업을 선택한다. 이것이 확률표본추출법 중 가장 간단한 방법으로 이 기법의 강점이기도 하다. 그 이유는 표집들이 세분화되고 부분화되지 않았기 때문에, 표본은 편향되지 않고 표본으로부터의 결론은 확률표본추출 기법 중 가장 일반화가 가능하기 때문이다.

계통표본추출법(systematic sampling). 표집들이 어떠한 기준에 의해 정렬되고, 요소는 정렬된 목록에서 규칙적인 간격으로 선택된다. 계통표본추출은 임의적인 선택과 함께 시작되어 모든 K 번째 요소를 계속적으로 선택하는 과정이다. $k = N/n$, k 는 표집단위 크기 N 을 희망표본 크기 n 으로 나눈 비율이며, 표본추출비율(sampling ratio)이라고 불린다. 시작점이 자동적으로 목록의 맨 처음이 아니라, 임의적으로 선택된 첫 번째 k 요소라는 것이 중요하다. 1,000 개 기업에서 200 개 기업을 선택하는 앞선 예제에서와 같이, 1,000 개의 기업을 그들의 크기 순서(직원 수, 세입)등의 오름차순(혹은 내림차순)으로 정렬하고, 정렬된 목록에서 임의의 첫 번째 기업을 선택하고 매번 다섯 번째 기업을 계속해서 선택하면 된다. 이러한 과정은 대기업 혹은 소기업에 대한 지나친 대표성(overrepresentation)을 가지고 있지 않으며, 모든 크기의 회사들이 표집들 내에서 균등하게 선택 가능하게 한다. 즉, 표본은 적어도 한 가지 기준 내에서 모집단을 대표한다고 할 수 있다.

층화표본추출법(stratified sampling). 층화표본추출에서 표집들은 모집단을 계층에 따라 범주화하고 이를 하위집단으로 구성한다. 반면에, 단순임의추출은 전체 모집단에서의 임의적 표본추출을 통해 행해진다. 1,000 개의 기업에서 200 개의 기업을 선택하는 앞선 예제에서, 기업의 규모에 기초하여 대기업(500 명 이상의 직원 보유), 중견기업(50~500 명의 직원), 소기업(50 명)으로 범주화 할 수 있다. 그 뒤에 각각의 하부 집단에서 임의적으로 67 개의 기업을 선택하여 총 200 개의 기업 표본을 선택할 수 있다. 하지만, 표집들 내에는 대기업보다 훨씬 많은 수의 소기업이 존재하기 때문에, 같은 수의 대기업, 중견기업, 소기업의 표본을 갖는 것은 모집단을 대표하지 못한다. 이는 개별 집단 내에서 표본의 비율이 표집들 혹은 관심 모집단에서의 비율을 반영하지 못하는 것이다, 적은 비율의 집단(대기업)이 과도하게 표본화 되었기 때문에, 이것을 비확률적층화표본추출법(non-proportional stratified sampling)이라고 한다. 모집단에서의 비율로 하부 집단을 선택하는 것이 하나의 대안

기법이라고 할 수 있다. 예를 들어, 100 개 대기업과, 300 개 중견기업과, 600 개 소기업이 존재한다면, 20 개의 대기업 표본, 60 개의 중견기업 표본, 120 개의 소기업 표본을 추출할 수 있다. 이 경우, 모집단의 확률 분포가 표본에서도 유지되기 때문에 확률층화표본추출법(proportional stratified sampling)이라고 부른다. 비확률적표본추출 접근은 대기업 표본과 같은 작은 하부 집단을 표현하는데 특히 효과적이지만, 확률적표본추출과 비교하여 비확률적 접근의 결과물이 전체 모집단에 대한 하부 집단의 확률에 따라 가중되어 있다면 그 대표성은 약하다.

군집표본추출법(cluster sampling). 지리적으로 넓은 지역에 존재하는 모집단이 있다면, 전체 모집단을 단순임의표집이 불가능할 수 있다. 이러한 경우, 모집단을 지리적인 경계를 따라 군집화 하는 것이 효과적일 수 있는데, 임의로 몇개의 군집을 표본조사하고, 그 군집의 모든 개체(unit)를 측정하는 것이다. 예를 들어, 뉴욕주의 시정부를 표본조사 하고 싶다면, 주의 모든 곳을 다니며 모든 시공무원들을 인터뷰하는 것(단순임의표집을 하는 것)보다, 자치주를 바탕으로 이 정부들을 군집화할 수 있다. 임의로 세 개의 자치주를 묶고, 이 자치주들의 각 공무원들을 인터뷰하면 된다. 그러나 군집 간 차이에 따라서 군집 표본의 표본 추정 값의 가변성은 일반적으로 단순임의 표본 표집의 가변성보다 높기 때문에 단순임의표본표집 방법의 결과보다 표본에 있어 일반화되기 어렵다.

대응쌍표본추출법(matched-pairs sampling). 연구자들은 종종 한 가지 구체적인 기준을 바탕으로 하나의 모집단 내에서 두 하위 집단을 비교하고 싶어 한다. 예를 들어, 어떤 기업들은 지속적으로 다른 기업보다 왜 수익성이 좋은가? 이러한 연구를 수행하기 위해서 표집틀은 기업의 총이익, 주당 순익, 혹은 다른 수익성의 척도를 바탕으로 ‘높은 수익성’ 과 ‘낮은 수익성’ 으로 구분한다. 그리고 하위 집단에 속한 기업 가운데 단순임의표본을 고르고, 기업의 크기, 업종, 그리고 다른 대응하는 기준들을 바탕으로 두 번째 하위 집단에 속한 기업과 짝을 지어본다. 이제 더 세밀하게 조사할 수 있도록 높은 수익성 회사와 낮은 수익성 회사로 짝지어진 두 개의 쌍이 된 표본이 존재하게 된다. 이렇게 쌍으로 표집을 하는 방법이 다른 하위 집단들 사이에서 조울증의 차이를 이해하는 적절한 방법으로 쓰인다.

다단계표본추출법(multi-stage sampling). 앞서 기술된 확률표집은 모두 단일 표본추출법의 예들이다. 표집이 필요로 하는 방법에 따라, 단일표본추출 방법을 조합하여 다단계표집 방법으로 사용할 수 있다. 예를 들어, 기업 규모에 따라 계층으로 나눌 수 있고, 각 계층에 맞게 계통표집을 할 수 있다. 이것은 층화표집과 계통표집의 이단계 조합이다. 마찬가지로, 학군에 따른 학생들의 성적을 살펴보고자 한다면, 먼저 학군으로 군집화하고, 각 군집을 단순임의표본추출 기법을 활용하여 표본을 설정한다. 그리고 군집 내 학교들을 대상으로 성적에 따라 단순임의 표본을 추출하고 표본 내의에서 성적에 따라 학생들을 단순임의표본으로 선택한다. 이 경우, 군집화와 단순임의표본으로 이루어진 네 단계의 표집과정을 거치게 된다.

8.3 비확률적표본추출(Non-Probability Sampling)

비확률적표본추출(nonprobability sampling). 모집단의 단위들이 표본으로 선택의 기회가 없거나 선택 확률이 정확히 결정되지 않은 상황의 표집기법이다. 일반적으로, 단위는 쿼터(quota)나 편의에 따른 비확률 기준에 따라 선택된다. 선택이 임의적으로 결정되지 않기 때문에, 비확률적 표집은 표본오차의 값이 아닌, 표집편향(sampling bias)에 영향을 받을 수 있다. 따라서, 표본으로부터의

정보는 모집단으로 일반화 할 수 없다는 특징을 가지고 있는데 비확률적표집 방법의 종류는 다음과 같다.

편의표본추출법(convenience sampling). 우연적 표집법 혹은 기회 표집법이라고도 하는데, 이것은 모집단으로부터 접근성이 쉽고, 이용 가능한, 혹은 편리한 상황으로부터 추출한 표본이다. 예를 들어, 연구자는 쇼핑몰 밖에 서 있고 사람들이 쇼핑몰 안으로 들어갈 때 설문지를 나눠주고 데이터를 수집한다면, 대상자 표본은 편리한 표본추출 방법을 사용한 것이다. 이것은 계획적으로 타 쇼핑몰에서 쇼핑하는 사람들을 제외시키기 때문에 비확률적 표본추출이다. 당신이 선택한 표본으로부터 얻을 수 있는 의견은 쇼핑몰의 특징(예를 들어, 고급품 가게는 더 부유한 사람들을 끌어들일 것이다), 고객의 인구학적 개요, 혹은 그 위치(예를 들어, 대학에 가까운 쇼핑몰은 주로 독특한 구매 성향의 대학생들의 마음을 끌 것이다)와 같은 특정 쇼핑몰의 특별한 성질을 반영할 것이고, 따라서 전체적인 구매자들의 견해를 대표하지는 못할 것이다. 따라서 이런 관찰의 과학적 일반화 가능성은 매우 제한적이다. 편의표본추출의 다른 예로는, 특정 수업에 등록된 학생들을 표본조사 하는 것이나, 특정 병원에 가는 환자들을 표본조사 하는 것 등이 있다. 이런 방법의 표본조사는 일반화 가능한 추론을 얻는 것보다는 측정도구의 타당성을 목적으로 하는 예비검사에서 가장 유용하다.

할당표본추출법(quota sampling). 층화 추출법에서와 같이 모집단이 상호배타적인 하위집단으로 나뉘어져 있고, 각 하위 집단에서 선택된 비확률 집단은 미리 정해진 할당량(quota)으로 정해져 있는 기법이다. **비례할당추출법(proportional quota sampling)**에서는, 각 하위 집단의 대상자 비율이 모집단의 비율과 대응 해야한다. 예를 들어, 미국 인구가 70% 코카시안, 15% 히스패닉, 그리고 13% 아프리카인으로 구성되어 있고, 98 명의 표본을 가지고 투표 선호도에 대해서 조사하려면, 쇼핑몰 밖에 서서 사람들에게 투표 선호도에 대해서 물어볼 수 있다. 15 개의 응답을 히스패닉 하위집단으로부터 얻은 후로는(혹은, 아프리카인 집단으로부터 13 개의 응답을 얻게 되고 나서는) 다른 집단에서는 계속 표본조사를 해야한다. 그렇기 때문에 히스패닉으로 보이는 사람들에게는 더 이상 질문하여 반응을 확인하지 않아야 표본의 인종구성이 전체적인 미국인구의 인종구성과 일치하게 된다. **비비례할당표본추출법(non-proportional quota sampling)**은 비례적으로 대표할 표본이 없어도 된다는 점에서 덜 제한적이지만, 각 하위집단의 최소한의 사람들은 만나야 한다. 이 경우, 코카시안, 히스패닉계 미국인, 아프리카인 미국인의 하위집단에서 각각 50 명의 응답자를 뽑기로 결정하고, 각 하위집단의 할당량에 충족되면 그만하게 된다. 어느 쪽의 할당추출법도 미국 인구를 대표하지는 못할 것인데, 수집하는 장소의 변수가 존재하기 때문이다. 실험이 뉴욕에서의 쇼핑몰에서 이루어졌는지 캔자스의 쇼핑몰에서 이뤄졌는지에 따라서 결과가 완전히 달라질 수 있다. 비비례방법은 모집단에 대해서도 잘 대표하지 못하지만, 적은 수의 불충분한 집단의 의견을 담아내거나 과다 표집을 할 때 유용하다.

전문가표본추출법(expert sampling). 연구되고 있는 현상에 대해 전문지식을 바탕으로 비확률적 방법을 활용하여 응답자를 선택하는 방법이다. 예를 들어, 사베인즈옥슬리법과 같은 새로운 정부의 정책의 효과를 알아보기 위해, 이 법에 익숙한 법인 회계사 집단을 표본 조사할 수 있다. 이러한 접근의 이점은, 전문가들이 비전문가들에 비해 이러한 주제에 대해 더 익숙한 경향있다. 비록 결과가 전체의 모집단으로 일반화할 수 있는 것은 아니지만, 전문가 표본으로부터의 의견이 전문가와 비전문가를 함께 포함하는 표본보다는 더 신뢰할 수 있는 특징을 가지고 있다.

스노우볼표본추출법(snowball sampling). 연구자는 연구에서 기준에 부합하는 몇몇 응답자들을 확인하고, 그들이 생각하는 연구기준에 부합되는 다른 사람을 추천해달라고 부탁함으로써 표집하는 방법이다. 만약 당신이 컴퓨터 네트워크 관리자에 대해서 조사를 대상으로 조사하고 싶고 그런 일을 하는 한, 두명만 알고 있다면, 그들을 대상으로 시작해서 그들에게 네트워크 관리 일을 하는 다른 사람들을 추천해달라고 부탁하는 것이다. 비록 스노우볼표본추출법이 모집단을 대표할 수 있는 표본을 이끌어내는 것은 거의 어렵지만, 접근하기 어려운 모집단에 대해서나 표본추출의 틀이 이용 불가능한 경우에는 유일한 방법이 될 수도 있다.

8.4 표본추출의 통계

앞부분에서는 모집단, 표본통계량, 표집편향 등과 같은 용어들에 대해 알아보았다. 이 단원에서는 이런 용어들이 무엇을 의미하는지, 그리고 서로 어떤 연관 있는지에 대해 이해하고자 한다.

리커트 척도 항목에 대한 응답자의 반응과 같이 특정하게 관찰된 상황을 주어진 단위를 사용하여 측정할 때, 이렇게 관찰된 값을 응답(response)이라고 한다(그림 8.2). 다시 말하면, **응답(response)**은 표본 단위로부터 추출된 측정값이다. 동일한 측정도구를 사용하더라도 서로 다른 응답자로부터의 응답은 각각의 항목에 대해 개별적으로 서로 다르게 나타날 것이다. 동일한 항목이나 관찰에 대해 응답자들로부터 도출된 반응들을 발생 빈도를 바탕으로 **도수분포(frequency distribution)**로 나타낼 수 있다. 다수의 응답자로부터 응답한 데이터를 수집하여 도수분포화 한다면 벨 모양 곡선과 닮은 **정규분포(normal distribution)**를 따르게 된다. 이것은 표본평균이나 표준 편차와 같은 표본 전체의 전반적인 특징을 평가하는 데 사용될 수 있다. 이런 표본 평가들을 **표본통계량(sample statistic)**이라고 하며 관찰된 데이터를 통해 평가된 값을 ‘통계’라고 한다. 모집단도 전체를 표본 조사했을 때 얻을 수 있는 평균과 표준편차를 가지고 있다. 하지만 전체 모집단은 절대 표본 조사될 수 없다. 모집단의 특징은 항상 알려져 있지 않기 때문이며, 이것을 **모수(population parameter)**라고 한다. 데이터를 통하여 통계적으로 추정되지 않으면 ‘통계’가 아니다. 표본이 모집단을 완벽하게 대표하고 있지 않다면 표본통계량은 모수와 다를 수 있고, 이 둘의 차이를 **표본오차(sampling error)**라고 한다. 이론적으로 표본의 크기를 점점 증가시키므로써, 다시 말해서 표본의 수가 모집단의 수에 점점 가까워지면, 표본오차는 작아지고 표본통계량은 점점 모수의 근사치에 대응하게 된다.

표본이 모집단을 완전히 대표하게 되면, 추정된 표본통계량은 대응하는 이론상의 모수와 동일해야 한다. 표본통계량이 타당할 정도로 모수에 근접해 있는지를 어떻게 알 수 있을까? 여기서 **표본분포(sampling distribution)**의 개념에 대해 이해해야 할 필요가 있다. 그림 8.3 에 나타나 있는 것과 같이 주어진 모집단에서 서로 다른 세 종류의 임의표본을 고르고 각 표본의 표본평균과 표준편차와 같은 표본통계량을 도출했다고 생각해 보자. 각각의 임의표본이 모집단을 완전히 대표하고 있다면, 세 임의표본의 표본평균은 동일할 것이며(모수와도 동일하며), 표본평균의 가변성은 0 이 될 것이다. 그러나 이것은 임의표본이 모집단에서 서로 다른 부분 집합을 구성하고 있기 때문에 극단적으로 있음직하지 않으며, 따라서 표본들 사이에서 평균들은 서로 미세하게 다를 수 있다. 세 표본평균을 가지고 빈도 히스토그램(frequency histogram)을 나타내보면 표본들의 숫자가 3 에서 10 으로, 3 에서 100 으로 늘어가면, 빈도 히스토그램은 표본분포가 될 것이다. 따라서 표본분포는 표본들의 집단들로부터 나온 표본통계량의 도수분포이다. 도수분포는 단일표본의 응답(관찰)분포이다. 도수분포와 같이, 표본분포도 평균을 중심으로 멀리 흩어진 값과 평균을 중심으로 많은 모인

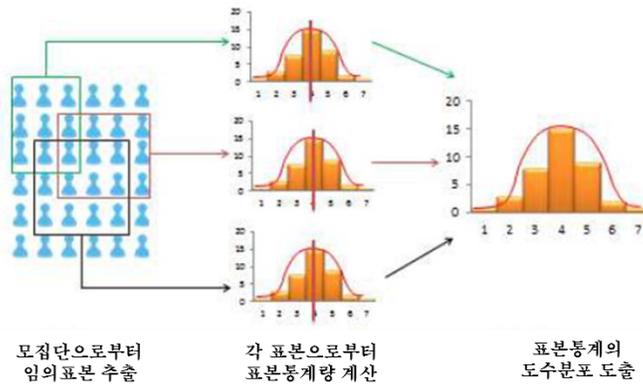
표본통계량을 가지는 경향이 있다. 무한대로 큰 표본의 그 분포는 정규분포로 나타낼 것이다. 예를 들어, 표집통계의 표준 편차와 같이 표본분포에서 표본통계량의 가변성이나 산포도는 **표준오차(standard error)**라고 한다. 반대로, 표준편차(standard deviation)라는 용어는 단일표본에서 관찰된 응답의 가변성이라는 뜻을 갖는다.

항목명	ATT1	ATT2	ATT3	INT1	INT2	INT3	PROO	MP3
6	6	6	6	6	6	6	2	2
1	1	1	1	1	1	1	5	10
6	5	4	2	5	5	1	5	5
6	6	6	5	5	5	2	4	
6	6	6	6	5	4	2	2.5	
7	7	7	7	7	7	6	2	3
5	4	5	5	4	5	2	1	
6	6	6	6	7	7	3	10	
7	6	6	5	5	4	2	1	
4	4	4	4	4	4	1	0	
5	5	5	4	4	4	2	1	
4	4	4	4	4	4	2	5	
6	6	5	5	5	5	2	4	
-1	-1	-1	-1	-1	-1	-1	-1	1
-1	-1	-1	-1	-1	-1	-1	-1	-1
-1	-1	-1	-1	-1	-1	-1	-1	-1
6	6	6	6	6	6	2	15	
6	6	5	5	5	4	2	1	
1	1	1	1	1	1	1	0	
6	6	6	6	6	6	2	6	
6	6	5	5	6	5	2	5	
5	5	5	2	1	2	2	7	
4	5	6	1	5	2	2	3	
6	7	7	6	7	6	2	5	

<그림 8.2> 표본통계

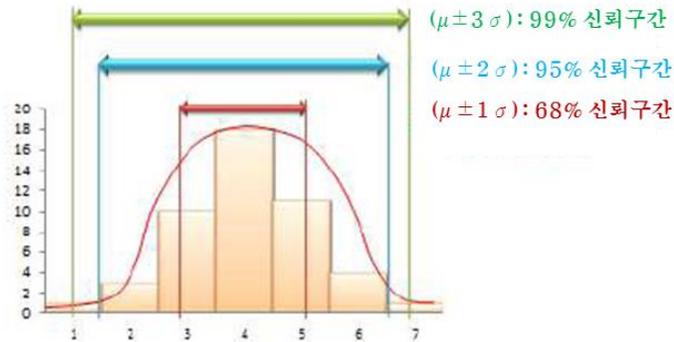
표본분포에서 표본통계량의 평균값은 알려지지 않은 모수의 추정으로 추정된다. 표본분포를 기반으로 모수 예측을 위한 신뢰구간을 측정하는 것이 가능하다. **신뢰구간(confidence interval)**은 표본통계량의 특정한 구간 안에 모수가 있을 가능성을 측정한 것으로 모든 정규분포는 68-95-99 퍼센트 규칙을 따르는 경향이 있다(그림 8.4). 68-95-99 퍼센트 규칙은 평균 값에서 한 개의 표준편차가 ($\pm 1\sigma$) 구간 안에 있을 경우가 68%, 평균 값에서 두 개의 표준편차가 ($\pm 2\sigma$) 구간 안에 있을 경우가 95%, 평균 값에서 세 개의 표준편차가 ($\pm 3\sigma$) 구간 안에 있을 경우가 99%를 넘는 것을 말한다. 무한한 표본 안의 표집분포가 정규분포에 가까워지면 68-95-99 규칙은 적용될 수 있고 다음과 같이 말할 수 있다.

- (표본통계량 \pm 한 개의 표준오차) 는 모수의 신뢰구간이 68%라는 것을 나타낸다
- (표본통계량 \pm 두 개의 표준오차) 는 모수의 신뢰구간이 95%라는 것을 나타낸다
- (표본통계량 \pm 세 개의 표준오차) 는 모수의 신뢰구간이 99%라는 것을 나타낸다



<그림 8.3> 표본 분포

표본은 표본분포로 측정될 수 없거나 표본분포가 68-95-99 규칙을 따르지 않을 경우에는 ‘편향적(biased)’ 이 된다. 즉, 모집단을 대표하지 않는 것이다. 우리는 회귀분석을 사용하여 유의성 검사를 할 때, 대체로 회귀계수를 $p < 0.05$ 로 설정한다. 이것은 표본통계량이 95%의 신뢰구간과 일치하는 모수를 예측할 수 있는 것과 같다. 예를 들어, ‘6 시그마’의 기준은 99%의 신뢰구간을 벗어난 결함있는 제품을 찾아내는데 있거나 그 목적이 $p < 0.01$ 에서 유의성 검정을 나타낸다.



<그림 8.4> 68%-95%-99% 신뢰구간

제 9 장 조사연구

조사연구(survey research)는 사람의 선호도, 생각, 행동에 대한 데이터를 체계적으로 수집하기 위해 표준화된 설문이나 면접을 활용하는 연구방법이다. 일찍이 인구조사가 고대 이집트 시대에 만들어졌을지라도 전형화 된 연구방법의 조사는 사회학자 파울 라자펠드(Paul Lazarsfeld)에 의해 시작되었다. 그는 1930-1940 년대 라디오가 미국인들의 정치적 성향 구성에 미치는 영향을 알아보기 위한 연구를 그 후 이 방법론은 사회과학의 양적 연구에 있어 아주 중요하고 일반적인 연구방법론이 되었다.

조사방법으로는 기술조사, 탐색조사, 설명조사가 있다. 조사연구는 개인을 분석의 구성단위로서 할 경우 가장 적합하다. 예를 들어 분석의 구성단위가 집단이거나 조직, 또는 한 쌍(구입자와 판매자와 같은 조직의 한 쌍)인 것들도 조사연구의 대상이 될 수 있지만 이러한 연구에서 종종 ‘핵심 정보제공자’ 또는 ‘대리인’ 과 같이 구성단위는 특정한 개인을 대상으로 한다. 선택된 정보 제공자가 현상에 대해 적합한 지식을 갖고 있지 않거나 편향된 의견을 가지고 있다면 응답자는 편견을 가진 대상이 될 수 있다. 예를 들어, 최고경영자가 직원들의 자사에 대한 의견이나 팀워크에 대해 충분히 알지 못할 가능성이 높기 때문에 직원들의 자부심이나 팀워크에 대한 연구에 있어 잘못된 정보제공자를 선택하였다고 할 수 있다.

다른 연구방법과 비교하여 조사연구는 몇 가지 강점을 가지고 있다. 첫째, 관찰하기 힘든 사람들의 선호도, 특성, 태도, 믿음, 행동, 또는 사실에 근거한 정보와 같은 아주 광범위한 자료를 측정하기 위한 훌륭한 수단이다. 둘째, 모집단의 범위가 너무 커서 직접적으로 관찰하기 어려운 자료를 간접적으로 수집하는데 가장 이상적인 방법이다. 국가와 같이 큰 범위는 작은 표본에서 모집단을 적절하게 대표한다는 것을 보장하기 위해서 꼼꼼한 표집방법을 사용하여 우편이나 전자메일, 또는 전화조사로 활용될 수 있다. 셋째, 응답자는 질문에 대하여 민감하지 않으며, 편의성 때문에 설문조사는 대부분의 응답자에게 선호된다. 넷째, 면접(interview)은 표본이 존재하지 않는 노숙자나 불법체류자들과 같은 특정 집단으로부터 접근 가능한 유일한 방법이라고 할 수 있다. 다섯째, 표본이 큰 설문조사는 다수의 변수를 분석함으로써 작은 영향도 알아내는 것이 가능하고 설문조사의 설계 구성에 따라 모집단의 하위집단의 비교분석도 가능하다(예, 집단 안에서의 분석과 하위집단 간 분석). 여섯째, 조사연구는 실험연구나 사례연구보다 연구자의 시간, 노력과 비용의 측면에서 경제적이지만 특이한 단점을 갖고 있다. 이 장의 마지막에서 논의될 무응답자, 편향적인 표집법, 사회적인 바람직함에 의한 편향, 그리고 회상에 의한 편향이 그것이다.

데이터가 어떻게 수집되는지에 따라 조사방법은 두 개의 넓은 범주로 나누어 진다. 설문조사(우편, 집단 관리 또는 온라인 조사)와 면접조사(개인, 전화, 포커스그룹 인터뷰)이다. 설문조사는 온전히

응답자에 의해 이루어지는 방법인 반면에 면접은 응답자의 응답을 기반으로 한 면접관에 의해 이루어진다. 각각의 설문조사와 면접조사는 비용, 목표 모집단의 범위, 질문할 때 연구자의 융통성에 따라 강점과 약점을 가진다.

9.1 설문조사(Questionnaire Surveys)

프란시스 갈톤(Francis Galton)에 의해 고안된 **설문(questionnaire)**은 표준화된 방식으로 응답자의 반응을 알아보기 위해 질문들로 구성된 연구도구이다. 질문은 비구조적일 수도, 구조적일 수도 있다. 비구조적 질문은 응답자에게 그들의 생각·의견을 글로써 반응을 살펴보는 구성이라면 구조적 질문은 응답자가 질문을 읽고 주어진 선택문항으로부터 답을 고르는 구성이다. 구조적 설문에서 질문에 대한 개인적인 응답은 통계적 분석을 위해 복합척도 또는 복합지수로 측정된다. 질문은 응답자가 읽을 수 있고, 이해할 수 있고, 의미를 갖는 형태로 설계되어야만 한다. 그렇기 때문에 설문조사는 아이들이나 문맹자와 같은 특정 집단에 적합하지 않거나 유용하지 않을 수 있다.

대부분의 설문조사는 동일한 질문을 많은 사람들에게 우편으로 보내고 응답자들은 자발적으로 편한 시간에 응답을 마치면 봉투에 넣어 연구자에게 다시 회신하는 자기기입식(또는 자계식)**우편조사(self-administered mail survey)**인 경우가 많다. 대부분 국가의 우편요금은 비용측면에서 저렴하기 때문에 우편조사의 장점을 가지고 있다. 하지만 대부분의 사람들이 조사요청을 무시하는 경향이 있기 때문에 메일조사의 응답률은 꽤 낮은 편이다. 또한 응답자들이 설문을 완성하고 다시 회신하기까지(어쩌면 설문지를 잃어버릴 수도 있다) 몇 달 정도의 시간적인 지체가 있다. 이러한 이유로 연구자가 반드시 응답자가 설문을 회신하고 있는지 관찰해야 하며 무응답자에게 설문을 상기시키는 내용을 보냄으로써 지속적으로 관리해야 한다(한 달에서 한달 반 사이에 두·세 개의 상기 메시지가 적당하다). 설문조사는 응답자가 작성한 설명이나 응답자에 의해 자세히 쓰인 응답이 필요한 경우에는 적합하지 않다. 경도설계(longitudinal design)는 서로 다른 시간에 동일한 응답자들의 반응을 조사하는데 사용될 수 있지만 응답률은 다음 조사까지 급격하게 떨어질 것이다.

조사의 두 번째 종류는 **집단조사설문(group-administered questionnaire)**이다. 응답자는 공통된 장소와 시간에 있어야 하며 그리고 각 응답자들은 설문조사를 그 공간 안에서 완성해야 한다. 응답자들은 다른 사람들과 상호의 영향 없이 그들의 응답을 독립적으로 작성해야 한다. 이러한 설문방식은 연구자에게 간편하고 높은 응답률을 보장한다. 만약 응답자가 특정한 질문을 이해하지 못하면 설명을 요청하는데 응할 수 있다. 만약, 회사 중역으로부터 설문조사를 할 수 있도록 승인을 받았다면 구내식당이나 회의실에 직원들을 모아 조사하는 것은 상대적으로 쉽게 응답을 얻을 수 있다.

최근의 설문조사는 인터넷이 바탕으로 된 온라인이나 웹 조사를 통해 이루어진다. 응답자는 설문조사 참여에 대한 메일을 받게 되며, 이것은 설문조사가 완성되어야 하는 온라인 웹사이트와 연결된다. 또는 설문이 이메일에 포함되어 있어서 설문을 완성하고 메일로 회신할 수 있다. 이러한 조사들은 관리하는데 매우 저렴하고 결과도 즉각적으로 온라인 데이터베이스에 기록이 가능하다. 그리고 필요하다면 설문은 쉽게 수정될 수도 있다. 하지만 설문 웹사이트가 패스워드가 없거나 복수제출을 방지하도록 설계되지 않았다면 응답자들은 쉽게 응답을 변경 할 수도 있다. 게다가 어린아이, 가난한 사람, 연장자, 소수집단과 같이 컴퓨터가 없거나 인터넷 접속이 어려운 사람에게는 실시될 수 없기 때문에 표집편향(sampling bias)이 발생하므로 중요한 문제가 될 수 있다. 응답자 표본대상이 많은

시간을 온라인에서 보내고 설문을 완성할 능력과 시간이 있는 젊은 세대들에게 치우칠 가능성이 있다. 만약 설문링크가 연구대상에게 이메일로 전송 되는 것이 아니라 불특정 다수가 사용하는 게시판을 통해 전해진다면 컴퓨터를 사용한 응답률은 많은 문제를 야기한다. 이러한 이유들 때문에 많은 연구자들은 응답자가 선호하는 응답방식을 선택할 수 있는 듀얼미디어조사(dual-media survey)를 선택한다(예, 우편 조사와 온라인 조사).

설문문항을 만들고 구성하는 것은 기술이다. 반드시 질문의 내용, 표현이나 단어선택, 서식, 그리고 순서와 관련하여 많은 결정프로세스와 함께 이루어져야 한다. 이러한 모든 것들이 설문조사를 통한 결과에 중요하게 영향을 미칠 수 있다.

응답서식(response format). 설문 문항은 구조적이거나 비구조적 형태이다. 구조적인 문항에 대한 응답은 다음에 제시되는 한 가지 응답 서식을 따라야 한다.

- **이분화된응답(dichotomous response)**은 ‘참/거짓, 예/아니오, 동의/비동의’ 와 같이 응답자가 두 가지 가능한 대안 중 한 가지를 선택하는 것을 말한다. “특정한 환경에서 사형제도가 정당화 될 수 있다고 생각하는가?” (동그라미 하시오)라는 질문은 예/아니오 로 나타낼 수 있다.
- **명목응답(nominal response)**은 두 개의 이상의 배열되지 않고 서열화 되지 않은 선택대안을 말한다. 가령 “당신이 종사하고 있는 직종이 어느 산업에 속해 있습니까?” 라는 질문에 제조업 / 고객 서비스 / 소매업 / 교육 / 건강관리(의료) / 관광&서비스 / 외 라는 대안을 제공하는 것이 그 예라 할 수 있다.
- **서열응답(ordinal response)**은 서열화된 두 개 이상의 선택대안을 응답자에게 제공하는 것이다. 가령, “최종 학력은 무엇입니까?” 라는 질문에 대안으로 고등교육과정 / 학사과정 / 대학원 과정 이라고 대안을 제시하는 것이 그 예라 할 수 있다.
- **구간수준응답(interval-level response)**은 리커트 척도나 어의차이 척도 혹은 거트만 척도 등에서 제시되는 5 점 혹은 7 점 범위의 문항을 응답자에게 제시하는 것을 말한다. 이 척도의 유형은 지난 장에서 논의된바 있다.
- **연속적응답(continuous response)**은 응답자의 나이나 회사에서의 경력과 같이 의미 있는 영점을 지닌 비율척도화 된 연속적 값을 응답자에게 선택하도록 제시하는 것을 말한다. 이러한 응답은 일반적으로 빈칸으로 두어 응답자가 직접 작성하도록 제시하는 경향이 있다.

문항 내용과 단어 선택(question content and wording). 설문조사에서 질문 문항에 따라 도출된 응답은 매우 민감하게 나타난다. 설문 문항의 틀이 응답자가 보기에 불완전하거나 애매모호한 문항으로부터 얻은 무의미한 응답은 가치가 없는 결과를 초래하게 될지도 모른다. 딜먼(Dillman, 1978)은 좋은 설문 문항을 만들기 위한 몇 가지 규칙을 제시해 주고 있다. 설문조사에서 모든 단일 문항은 다음에 제시되는 사안에 대한 면밀한 검증을 거쳐야만 한다고 설명하고 있다.

- *문항이 명료하고 이해 가능한 것인가:* 설문 문항은 단순한 언어로 기술되어야 하고 능동태가 선호되며, 응답자가 이해하지 못할 수 있는 복잡한 단어나 은어(전문용어)가 없어야만 한다. 설문지에 포함되어 있는 모든 문항은 비슷한 방식으로 기술되어 응답자로 하여금 읽기 쉽고 이해할 수 있게 해야 한다. 만약 설문 대상이 의사, 변호사, 연구자와 같이 전문용어를 매일 사용하는 환경에 있는 특정 집단을 대상으로 하는 경우는 예외적으로 전문용어나 복잡한 단어를 활용할 수 있다.
- *문항이 부정문으로 기술되었는가:* 예를 들어, “만약 지방자치단체는 세금을 인상하지 말아야 한다면……” 과 같이 부정적으로 기술된 문항은 응답자를 혼란하게 하고, 결국 부정확한 응답을 피할 수 없게 된다. 따라서 이러한 문항들을 피해야 하며, 모든 경우에 이중 부정문 또한 피해야 한다.
- *문항이 애매모호한가:* 설문문항은 응답자에 따라 다르게 해석될 수 있는 단어나 표현들(예, ‘어떠한’ 혹은 ‘단지’)와 같은 단어를 피해 작성해야 한다. 응답자에게 연 수입에 대해 물어본다면, 임금을 의미하는지, 배당금, 임대수입을 의미하는지 혹은 다른 종류의 수입을 의미하는지 와 같이 명확하게 제시해야 한다. 이와 마찬가지로, 동일한 질문이 개인적 수입을 의미하는지, 가계 수입(배우자의 수입도 포함)을 의미하는지, 개인 수입과 기업의 수익을 동시에 의미하는지에 대해서도 명확하지 않다. 응답자에 따른 서로 다른 해석에 대한 응답데이터는 연구자가 정확하게 해석할 수 없는 결과를 초래하게 될 것이다.
- *문항이 편향되어 있거나 가치 판단적인 단어를 포함하고 있는가?:* 편향은 응답자에게 특정한 방식으로 응답을 유도하는 문항의 특성을 말한다. 케네스 라신키(Kenneth Rasinky 1989)는 정부 지출에 대한 사람들의 태도와 관련된 여러 가지 연구를 진행하였고, 그 속에서 응답자들이 ‘복지’ 라는 것과 ‘빈곤층에 대한 지원’ 이 같은 의미를 지니고 있는 용어임에도 불구하고 후자 쪽으로 더 많이 응답하는 경향을 보이는 것을 발견하였다. 이와 비슷하게 ‘범죄율을 낮추는 것’ 을 더 지지하였고(‘법 집행’ 을 적게 지지함), ‘대도시 문제에 대한 해결’ 을 더 지지하였고(‘대도시에 대한 지원’ 을 덜 지지함), ‘약물 중독을 해결하는 것’ (‘약물 재활’ 을 덜 지지함)을 응답자들이 더 지지하고 있다는 것을 발견할 수 있다. 편향된 언어나 어조로 구성된 질문은 응답을 왜곡시키는 경향을 가지고 있다. 편향된 단어를 사전에 예측하는 것은 어렵지만, 설문 문항은 이러한 편향된 언어를 피할 수 있도록 면밀히 검토되어야 한다.
- *문항이 이중적으로 해석되는가:* 이중적으로 해석되는 문항은 다중의 답을 지닐 수 있는 문항이라 할 수 있다. “당신은 작업을 할 때 제공되는 하드웨어와 소프트웨어에 대해 만족하는가?” 라는 문항을 예로 들어보자. 이 예에서 만약 응답자가 하드웨어에는 만족하고 소프트웨어에는 만족하지 못하는 경우나, 혹은 그 반대의 경우에는 어떻게 응답해야 할까? 이렇게 이중적으로 해석되는 문항은 (1) 당신은 작업을 할 때 제공되는 하드웨어에 대해 만족하는가? (2) 당신은 작업을 할 때 제공되는 소프트웨어에 대해 만족하는가? 로 서로 다른 문항으로 분리해서 질문해야 한다. 다른 예로, 당신의 가족들이 공영 TV 를 선호하는가? 라는 질문에 몇몇의 응답자는 공영 TV 자체에 대해 만족한다고 응답할 수 있고, 아동용 Sesame Street 케이블 TV 프로그램을 선호한다고 응답할 수도 있다.

- *문항이 너무 일반적인가:* 때때로, 너무 일반적인 문항은 응답자의 인식을 정확히 전달하지 못할 수 있다. 만약 설문 대상자에게 “특정한 책에 대해 얼마나 좋아하는지”에 대해 묻고 ‘전혀 좋아하지 않는다’ 부터 ‘매우 좋아한다’ 까지 응답 범위를 제공하였다고 하자. 만약 그 사람이 ‘매우 좋아한다’ 를 선택한다면, 그것은 무엇을 의미하겠는가? 이 질문 대신에 “이 책을 다른 사람에게 추천해줄 것인가?”, “같은 저자가 지은 다른 책을 읽을 계획을 가지고 있는가” 와 같이 좀 더 특정하게 행동에 관한 질문을 하는 것이 좋다. 이와 마찬가지로, “당신의 회사가 얼마나 큰가?” (응답자 별로 다르게 해석될 수 있다.)라고 묻는 대신에, “회사에 종업원이 얼마나 많이 있습니까?” 혹은 “회사의 연 수입이 얼마나 되는가?” 와 같이 회사의 규모를 측정할 수 있는 구체적인 질문을 하는 것이 좋다.
- *문항이 너무 상세한가:* 연구 목적을 포괄하지 않는 너무 상세한 질문을 피하는 것이 좋다. 예를 들면, “가족 구성원에 연년생의 자녀가 필요한가?” 나 “가족 구성원으로 몇 명의 아이가 적당하다고 생각하는가?” 와 같은 질문이 그 예라 할 수 있다. 확실하지 않다면, 보편적인 측면보다는 세부 사항 측면에 실수를 범하는 것이 낫다.
- *문항이 추정하고 있지는 않는가:* 만약 “세금 할인에 대한 이익이 무엇인지?” 물어볼 경우, 응답자가 세금 할인을 이익으로 인식할 것이라고 추정하고 있는 것이다. 하지만 대부분의 사람들은 세금 할인을 이익으로 인식하지 않을 수 있다. 왜냐하면, 세금 할인은 아이들을 위한 공공 교육에 적은 양의 재원 투입을 의미한다고 인식할 수 있으며, 시민들에게 더 적은 공공 서비스(경찰, 구급차, 소방 서비스)를 제공할 것이라고 인식할 수 있기 때문이다. 따라서 추정에 기반한 문항은 피하는 것이 좋다.
- *문항이 가상적인 개념인가:* TV 게임 쇼에서 사회자가 “만약 당신이 백만 달러를 쇼에서 획득할 경우, 어느 곳에 지출하겠는가?” 와 같은 질문하는 것을 많이 보았을 것이다. 대부분의 응답자들은 그 정도의 돈을 가지고 있던 경험이 없거나 혹은 그러한 일에 대해 상상해본 적이 없다. 그렇기 때문에, 그들의 대답은 세계일주를 한다거나, 식당이나 바를 산다거나, 교육에 사용하겠다거나, 은퇴 후 노후자금으로 활용한다거나, 부모나 자녀들을 돕는다거나, 또는 호화로운 결혼식을 치를 것이라고 응답하는 등 매우 임의적인 경향을 지닐 수 있다. 가상적인 질문은 가상적인 답변을 이끌게 되고, 그것은 과학적인 추론을 하는데 활용되지 못한다.
- *응답자가 질문에 정확하게 대답하기 위한 정보를 가지고 있는가:* 우리는 종종 어떠한 문제에 대답하기 위해서는 필요한 정보가 있다고 가정을 하게 된다. 하지만 현실에서는 그렇지 못한 경우가 있게 마련이다. 연구자가 연구대상으로부터 응답을 얻어질지라도, 그 응답은 질문문항에 대한 응답자의 정보 부족으로 인해 정확하지 않을 수 있다. 예를 들어, 회사의 CEO 에게 그들이 알지 못하는 “회사 내 매일의 공정 세부사항을 알고 있는가?”, 선생님들에게 “그들의 학생이 얼마나 많이 배웠는가”에 대해서나 혹은 고등학교 학생들에게 “미국정부가 ‘피그만 위기’ 때 정확하게 대응하였는가?” 와 같은 질문을 해서는 안 된다는 것이다.

문항의 순서배열(question sequencing). 일반적으로 문항은 논리적인 흐름을 가지고 있어야 한다. 설문 대상자들로부터 높은 응답률을 달성하기 위해서는 문항들은 가장 덜 민감한 것부터 가장 민감한 것으로, 사실과 행동에 기반을 둔 것으로부터 태도에 기반을 둔 것으로, 일반적인 것에서 특정한 것으로의 흐름을 가지고 배열 되어야 한다. 문항의 순서배열에 대한 몇 가지 규칙은 다음과 같다:

- 응답자들이 쉽게 기억할 수 있고 당혹스럽지 않은 쉬운 문항으로 시작해라. 좋은 대안은 개인차원의 연구를 위한 인구학적 요소(나이, 성별, 교육 수준)나 회사 차원의 연구를 위한(종업원 수, 연간 수입, 산업)등이 될 수 있다.
- 절대 개방형 질문으로 시작하지 마라.
- 사건의 역사적인 순서를 따르고자 한다면, 최초의 것부터 최후의 것까지를 연대순에 따라 배열한다.
- 한 번에 한 가지 주제에 대해서만 질문해라. 주제를 바꾸고자 할 때, “다음 섹션은 xxx 에 대한 당신의 생각을 알아 보고자……” 과 같이 전환어(transition)를 제시해 주어야 한다.
- 필요하다면, 필터나 인지질문을 활용한다. 예를 들어, “5 번 문항에 대해 ‘예’ 라고 답할 경우, 2 번째 섹션으로 이동하고, ‘아니오’ 라고 답할 경우, 세 번째 섹션으로 이동하십시오” 이 있다.

다른 중요한 규칙들. 연구자는 응답자들을 자신과 같이 여길 줄 알아야 한다. 응답자의 시간, 참여, 개인정보에 대한 믿음과 신뢰성에 대해 주의 깊게 살피고 감사할 줄 알아야 한다. 다음에서 제시되는 모든 연구를 위한 전략을 연습하도록 해야 한다.

- 사람들의 시간은 가치 있는 것이다. 그들의 시간에 대해 존중해 줄 수 있어야 한다. 설문을 가능한 한 짧게 만들고 절대적으로 필요한 만큼으로 제한해야 한다. 응답자들은 10~15 분 이상의 시간을 설문에 사용하는 것을 원치 않기 때문에 긴 설문은 응답자로부터 극히 낮은 응답률의 결과를 초래한다.
- 응답자들에게 그들의 응답에 대한 신뢰성을 확인시켜주어야 한다. 예를 들어, 그들의 응답데이터가 어떻게 활용되고(예: 학술적 연구) 도출되는 지를 알려주어야 한다.
- 조직적 차원의 설문을 위해 응답자에게 최종 결과에 대한 보고서를 전달해 줄 것을 확실하게 하며, 이를 정확히 이행할 것임을 밝혀 주어야 한다.
- 응답자에게 연구에 참여해 준 점에 대해 감사의 말을 전해야 한다.
- 마지막으로, 실제 현장에서 응답자에게 설문을 부여하기 전, 설문문항에 대해 예비검사를 진행해야 한다. 이러한 예비검사는 설문 대상자에게 설문문항의 애매모호함이나 명확성의 결여, 문항의 편향된 단어들과 같이 제거되어야 할 오류들에 대해 확인하고 수정할 수 있게 해주기 때문이다.

9.2 면접 조사

면접조사는 설문조사보다 좀 더 개인이 원하는 대로 할 수 있는 자료수집방법이다. 면접 조사는 숙련된 면접관이 연구 원안(설문지와 비슷한)을 활용하여 진행하는 방식을 말한다(예: 표준화된 문항세트). 하지만, 설문과 다르게 면접조사 대본은 응답자에게 보여지지 않으며 면접관을 위한 특별한 지침을 포함할 수 있고, 면접관이 개인 관찰결과나 언급을 기록할 수 있는 공간을 포함할 수 있다. 더욱이, 우편설문과 다르게, 면접관은 응답자로부터 연구주제와 관련된 이슈에 대하여 확인 할 수 있는 기회가 있으며, 그에 대한 추가 문항에서 질문할 수 있는 기회를 가지게 된다. 하지만, 면접 조사는 많은 시간과 자원이 소비되는 기법이다. 또한 특별한 면접 기술이 면접관에게 요구되기도 한다. 면접조사에서 면접관은 측정도구의 한 부분으로 인식되기 때문에, 면접관은 응답자의 반응을 관찰하는 과정에서 편향된 결과를 제공하지 않도록 주의의 기울여야 한다.

가장 전형적인 면접조사의 형태는 개인 혹은 **면대면 면접조사(face-to-face interview)**이다. 면대면 면접조사는 면접관이 응답자에게 직접적으로 질문을 하고 그들의 응답을 기록하는 방식이다. 개인 면접조사는 응답자의 집이나 사무실에서 진행될 수 있다. 이 접근법은 어떤 응답자에게는 선호될 수 있는 반면에, 그들의 집에 낯선 이를 방문하는 것을 불편하게 느끼는 응답자에게는 선호되지 않을 수도 있다. 하지만, 숙련된 면접관들은 응답자들이 협조할 수 있도록 설득할 수 있고, 이는 응답률을 극적으로 향상시키는 결과를 발생시키기도 한다.

개인 면접조사를 변형한 것이 집단 면접조사이며, 또한 집단 면접조사는 **포커스그룹(focus group)**이라고 불리기도 한다. 약 6 명-10 명으로 구성된 응답자 집단을 동일한 장소에서 면접조사를 하는 것이다. 면접관은 기본적으로 토론을 이끌고, 모든 사람이 응답할 수 있는 기회를 부여하는 조력자의 역할을 수행하게 된다. 포커스그룹은 다른 형태의 조사연구보다 복잡한 문제에 대한 깊은 설명을 가능하게 한다. 왜냐하면, 다른 응답자가 이야기하는 것을 듣게 될 때, 사람들은 그들이 생각지 못한 부분에 대한 아이디어나 응답을 도출할 수 있기 때문이다. 하지만 포커스그룹의 토론은 몇몇의 개인에 의해 의견이 주도될 수 있고, 몇몇의 참여자들은 그들의 동료나 상관 앞에서 자신들의 의견을 내기를 꺼려하는 경향을 가지고 있다(특히 실적부진 직원 혹은 회사의 정책과 같은 민감한 주제에 대해 토론할 경우). 적은 수의 표본크기 때문에, 포커스그룹은 기술적 혹은 설명적 연구보다는 탐험적 연구에 활용된다.

면접조사의 세 번째 유형은 **전화 면접조사(telephone interview)**이다. 이 기법에서 면접관들은 표준화된 설문문항에 대한 답변을 도출하기 위하여, 전화번호부상에서 임의적인 사람을 선택하고 설문 대상자에게 전화를 통하여 연락을 하게 된다. 최근 기술적 발전으로 인해 컴퓨터 이용 전화조사(CATI)가 가능해 지게 되었다. CATI 는 학술연구기관, 정부, 상업 설문에 많이 활용되고 있는데 면접관이 전화 오퍼레이터로서의 역할을 수행하게 된다. 면접관들은 컴퓨터 프로그램을 통해서 지침과 질문문항을 스크린으로 확인하고 면접조사를 진행하게 된다. 이 시스템은 자동다이얼 시스템을 통해 설문 대상자들을 임의적으로 선택할 수 있게 하고, 응답 결과가 음성녹음 시스템을 통해 자동으로 기록되게 한다. 만약 응답자와 전화상으로 연결되게 되면, 높은 응답률이 얻어질 수 있다. 하지만 이 기법은 통화품질이 좋지 못한 교외지역에서는 이상적이지 못하며, 그래픽이나 상품의 실증과 같은 비음성 정보에 대한 결과를 추출하기 위해 활용되지 못하는 단점을 지니고 있다.

면접관의 역할. 면접관은 면접조사 상에서 복잡하고 다각적인 역할을 수행하게 되는데, 이는 다음에 제시되는 일들을 수반하게 된다.

- **면접조사 준비하기:** 면접관은 자료수집 과정의 일선에 위치하기 때문에, 자료의 질은 면접관이 잘 숙련되었는가에 기인하게 된다. 면접관은 면접조사 과정과 방법에 대해 교육을 철저히 받아야 하며, 연구의 목적과 더불어, 응답들이 어떻게 저장되고 활용되는지, 면접관 편향의 요인은 무엇인지 정확히 인지하고 있어야 한다. 또한, 공식적인 조사가 진행되기 전에 면접조사의 시간이 얼마나 걸리는지 사전에 리허설을 해보아야 한다.
- **응답자의 협조 요청하기:** 특히 집에서 이루어지는 개인 면접 조사의 경우, 면접관은 응답자의 개인 스케줄에 맞춰 조사를 진행할 수 있도록 준비해야 한다(주말과 같이 예상치 못한 시간이라 할지라도 이다). 그들은 응답자가 연구에 참여할 수 있는 판매원과 같은 역할을 수행할 수 있어야 한다.
- **응답자에게 동기부여하기:** 응답자는 자주 면접관으로부터 면접에 대한 동기를 얻게 된다. 만약 면접관이 재미있지 않거나 주의를 기울이지 않는다면, 응답자는 유용하고 유익한 응답을 제공하려는 동기부여를 갖기 못하게 된다. 면접관은 연구에 대한 열정을 보여주어야 하고, 응답자에게 이 연구가 얼마나 중요한지를 설명해 주어야 하며, 면접 조사 동안에 응답자의 요구에 주의를 기울여야만 한다.
- **어떠한 혼란이나 걱정에 대해 명확하게 해주기:** 면접관은 응답자의 걱정에 대해 그들의 입장에서 생각해 줄 수 있어야 한다. 더욱이, 질문 문항이 대본상 포함되어 있지 않더라도 필요하다면, 응답자를 배려하는 질문을 면밀히 살펴서 질문해야 한다.
- **응답의 질에 대해 살펴보기:** 면접관은 응답자로부터 획득한 데이터의 질을 판단할 수 있는 가장 좋은 위치에 있기 때문에 관찰과정에서 응답자의 몸짓이나 바디랭귀지를 통해 응답에 대해 정확하게 보충할 수 있다.

면접조사 시행하기. 면접관은 면접조사를 시행하기 전, 이를 위한 연구자로부터의 커버레터, 정확한 설문지 사본, 응답자의 사진 ID, 응답자에 연락할 수 있는 전화번호 등의 키트(kit)을 준비해야 한다. 면접관은 약속시간 이전에 미리 응답자에게 연락을 취해야 할 것을 권한다. 그리고 면접조사를 시작함에 있어 “면접 조사를 위해 들어가도 되겠습니까?” 라는 말보다 “당신과 중요한 연구주제에 대하여 이야기하고자 몇 분 동안 면접조사를 진행하고자 합니다” 와 같이 위엄 있고 신뢰성 있는 톤으로 말을 해야 한다. 그리고 면접관은 개인적인 소속을 밝히는 동시에 자신을 설명하고, 연구의 목적에 대해 1~2 문장으로 설명을 한다. 또한, 1 분이내에 응답자의 코멘트나 자발적인 참여에 대한 기밀성에 대해 보증해 준다고 언급해야 한다. 면접관은 말을 많이 하거나 전문용어를 사용하지 않아야 하고, 응답자가 특별히 요청하지 않는 경우를 제외하고는 세부사항을 제공해 주지 않아야 한다. 만약 면접관이 진행하고자 하는 면접을 녹음하고자 한다면, 시작 전에 반드시 응답자의 동의를 얻어야만 한다. 면접조사가 녹음되는 경우라도, 면접관은 응답자의 주요 단어나, 조사, 축어구 등을 주의 깊게 살펴야 한다.

면접 조사 동안, 면접관은 설문지 스크립트에 따라서 쓰여진 순서대로 질문하여야 하며, 문항을 좀 더 친숙하게 할 수 있도록 단어를 변경해서는 안 된다. 또한 질문의 순서를 변경해서는 안 되고, 응답되었을 수도 있는 어떠한 문항에 대해서 건너뛰어서는 안 된다. 이러한 문제가 발생하지 않도록 실제 면접 조사에 앞서, 리허설을 통해 수정되어야 한다. 면접관은 응답자의 말을 끊어서도 안 된다. 만약 응답자가 피상적인 답변을 했을 경우, 면접관은 응답자에게 더욱더 배려심 있게 질문하여 정보를 획득할 수 있도록 해야 한다. 몇 가지 질문 조사 기법은 다음과 같다:

- 조용하게 묻기: 잠시 정지하고 기다리는 것(다른 질문으로 넘어가지 않고)은 응답자로 하여금 면접관이 좀 더 상세한 응답을 기다리고 있다는 것을 깨닫게 할 수 있다.
- 명시적인 격려: 간헐적으로 “어-허” 나 “그렇군요” 와 같은 감탄사를 사용하는 것은 응답자의 상세한 응답을 이끌어내기도 한다. 하지만, 면접관은 응답자가 한 말에 대해 ‘동의’ 나 ‘비 동의’ 의 표현해서는 안 된다.
- ‘상세한 말 요청하기: 면접관은 “그것에 대해 상세하게 말해주실 수 있나요?” 혹은 “1 분 전에 고등학교에서의 경험을 이야기 해주셨습니다. 혹시 좀 더 말해주실 수 있나요?” 와 같은 질문을 한다.
- 반영: 면접관은 정신과 의사처럼 응답자가 했던 말을 반복할 수 있다. “내가 들은 당신의 경험은 대단히 충격적인군요” 라고 말한 후, 잠시 정지하고 응답자가 상세하게 말할 수 있도록 기다리는 것이 그 예라 할 수 있다.

면접조사가 끝난 후, 면접관은 응답자에게 감사를 표현하고 결과가 언제 나오는지 말해주어야 한다. 또한 서둘러서 조사를 마무리해서는 안 된다. 면접조사 동안에 이루어진 면접관이 작성한 어떠한 말이나, 주요 관찰사항은 연구자가 응답자의 코멘트를 해석하는데 도움을 줄 수 있다.

9.3 조사연구에서의 편향

조사연구의 많은 강점과 이점에도 불구하고 조사연구를 통해 도출된 추론들을 무효화 시킬 수 있는 체계적인 편향들을 가지고 있다. 비응답편향(non-response bias), 표집편향(sampling bias), 사회적으로망성(또는 사회규범 부응성)편향(social desirability bias), 기억편향(recall bias)의 다섯 개의 편향들이 있다.

비응답편향(non-response bias). 조사연구는 일반적으로 응답률이 낮은 것으로 잘 알려져 있다. 우편조사에서 조사 대상자에게 여러 번 상기시켜줌에도 불구하고 일반적으로 15-20%의 응답률을 가지고 있다. 만약 조사 대상자의 대다수가 조사에 참여하지 못한다면 연구결과의 타당성에 대한 문제를 제기될 수 있다. 다시 말해서 시스템적인 이유로 인해 비응답자들이 조사에 참여를 하는지 안 하는지가 중요한 이유가 되는 것이다. 예를 들어, 불만족한 고객은 만족한 고객들보다 그들의 경험에 대해 더 큰 목소리를 내기 때문에 만족한 고객들에 비해 인터뷰 요청이나 설문조사에 응답하고자 하는 성향이 더 높다. 그래서 응답자 표본은 원래 생각했던 기본적인 표본에서보다 불만족한 고객들에 대해 더 높은 비율을 갖게 될 것이다. 이러한 경우, 결과는 일반화하기가 어려울 뿐만

아니라 관찰된 결과들도 역시 편향된 표본으로부터 나와 인위적인 값이 되기 쉽다. 다음과 같은 몇 가지 전략들이 응답률을 높이는데 이용될 수 있다.

- *사전통보(advance notification)*: 조사 대상자들에게 미리 간단한 편지를 보내는 것은 다가올 조사에 대해 응답자들이 미리 준비할 수 있는 시간을 제공함으로써 응답의 성향을 개선하는데 도움을 준다. 보내는 편지에는 연구의 목적과 중요성, 자료수집의 방식(예, 전화설문, 메일설문 등), 협조에 대한 감사인사 등을 언급해야 한다. 또한 응답자가 이 연구에 참여할 의지가 있는지 없는지를 나타내는 요금별납 된 엽서를 회신해주도록 요청하는 것도 포함될 수 있다.
- *내용의 적절성(relevance of content)*: 만약 연구가 응답자에게 관심이 있는 사항이거나, 관심분야와 관련성이 있는 내용에 대한 질문이라면 응답자는 자신과 관련 있지 않는 연구일 때보다 응답할 가능성이 높다.
- *우호적인 응답자 설문문항(respondent-friendly questionnaire)*: 짧은 설문 문항일수록 긴 문항에 비해 더 높은 응답률을 끌어내기 쉽다. 게다가 분명하고, 불쾌하지 않으며, 답하기 쉬운 질문일수록 응답률이 더 높은 경향이 있다.
- *보증(endorsement)*: 조직을 다루는 연구는 조직의 회사중역으로부터 연구의 중요성을 입증받는 것이 도움이 된다. 이러한 보증은 연구자의 신뢰성을 높일 수 있도록 추천 양식이나 커버레터의 형태로 추가하여 구성한다.
- *추가요청문(follow-up request)*: 복합적인 추가 요청문은 비응답자로부터 답변을 유도하게끔 만든다.
- *인터뷰 진행자 교육(interviewer training)*: 인터뷰 응답률은 훈련된 인터뷰 진행자에 의해 결정되기 때문에 그들에게 인터뷰 요청방법과 잠재적인 응답자들을 찾기 위해 전산화된 다이얼링 기술을 사용하고, 연락이 닿지 않는 응답자들에게 일정을 재통보 하는지에 관하여 교육 받는다.
- *인센티브(incentives)*: 적어도 어떤 모집단에서의 응답률은 현금이나 기프트 카드, 또는 펜, 복권 응모권, 추첨권, 할인 쿠폰, 자선단체 기부와 같은 인센티브 형태를 활용하여 높일 수 있다.
- *비금전적 인센티브(non-monetary incentives)*: 특히 비즈니스는 금전적인 인센티브보다 비금전적인 인센티브에 응답하기 더 쉽다. 비금전적 인센티브의 예는 조사의 모든 응답들을 종합한 것에 대해 기업의 개인적인 의견들을 비교한 벤치마킹 보고서가 될 수 있다.
- *기밀성과 사적권리(confidentiality and privacy)*: 마지막으로 응답자들의 사적인 정보와 답변은 제 3자에게 전달되지 않는다는 보장은 응답률을 높이는데 도움이 될 것이다.

표집편향(sampling bias). 전화조사는 공식적으로 전화번호부에 있는 임의적 표본의 방식으로 조사가 수행될 때 전화번호 목록에 없거나 휴대전화를 가진 사람들, 그리고 전화로 응답하는 것이 불가능한 사람들은(예, 일하는 사람들) 제외하게 된다. 또한 하루 중 많은 시간을 집에서 머무르는 무직자, 장애인, 연장자 같이 일반전화 서비스가 가능한 응답자는 연구 조사대상을 균형적이지 못하게 한다. 온라인 조사는 즉각적으로 인터넷을 사용할 수 있는 학생과 젊은 사람들로부터 응답을 수집하게 되므로 불균형적이며 형편이 어렵거나 연장자와 같이 인터넷 또는 컴퓨터에 접근하는데 어려움이 있는 사람들을 배제하는 경향이 있다. 설문조사는 글을 읽고, 이해하고, 질문에 의미있는 응답을 할 수 없다고 판단되는 아이들과 문맹자를 배제하게 된다. 표집편향의 다른 유형은 잘못된 모집단의 표집과 관련이 있다. 이러한 예로 학생들(아이들)의 학습성취에 대해 교사(부모)에게 물어보거나 CEO에게 회사 내의 운영상의 세부사항들에 대해 물어보는 것을 들 수 있다. 이러한 편향들은 모집단을 대표하지 못한 표본이 되거나 편향된 표본으로부터 나온 추론에 대해 일반화 가능성을 어렵게 한다.

사회적바람직성편향(social desirability bias). 대부분의 사람들이 사회적으로 바람직하게 보이려는 편의를 의미한다. 다시 말해 응답자의 대다수는 자신이나, 고용인, 가족, 또는 친구들에 대하여 부정적인 의견이나 난처한 언급을 피하려고 한다. 당신의 프로젝트 팀이 제대로 역할을 하지 못하고 있다고 생각하는지, 당신의 직장에는 사내정치가 심한지, 인터넷에서 불법으로 음악파일을 다운로드 받은 적이 있는지와 같은 부정적인 질문들에 대해서 연구자는 아마도 응답자로부터 진실한 답변을 받을 수 없을 것이다. 자신들을 사회적으로 바람직한 태도를 가진 것으로 보이기 위해 ‘진실을 맴도는’ 응답자의 이러한 성향은 조사연구로부터 얻어진 응답들의 타당성을 약하게 하는 ‘사회적바람직성편향’이라고 일컫는다. 사실상 설문조사에서 사회적바람직성편향을 극복하기 위한 방법은 없다. 하지만 인터뷰 조사 시, 인터뷰 진행자는 일관성 없는 답변들을 알아채고 추가질문을 하거나 응답자의 답변을 보강하기 위해 개인적인 관찰력을 사용할 수 있다.

기억편향(recall bias). 설문항목에 대한 응답은 종종 주제의 동기부여, 기억, 응답 능력에 따라 다르게 나타난다. 특히 아주 오래 전에 일어난 사건에 대해 다룰 경우 응답자는 그 사건에 대해서 시간이 흐르면서 자신의 동기부여나 행동에 대해 정확히 기억하지 못할 수도 있다. 예를 들어, 응답자에게 어린 시절의 기억할 만한 자신의 생일과 같은 사건이나 일년 전 자신의 컴퓨터 기술의 활용 정도를 말해보라고 한다면 그들의 응답은 아마도 회상하는데 어려움이 있기 때문에 정확하지 않을 것이다. 기억편향을 극복하는 하나의 가능한 방법은 기억으로부터의 동기부여나 인지를 끌어내도록 하기보다는 그 특정한 사건이 일어났을 때와 같이 그 사건 안에 응답자의 기억에 머물게 하는 것이다.

동일방법편의(common method bias). 동일방법편의는 횡단면 조사나 동일한 도구를 사용하는 설문조사처럼 같은 시점에서 독립변수와 종속변수를 측정함으로써 그 사이에 공유된 허위 공분산의 양을 나타낸다. 이러한 경우에 조사하고자 하는 현상을 측정도구와 충분히 분리하지 않았을 것이다. 표준화된 통계적 검사는 하몬드(Harmond)의 단일요소 검사(Podsakoff et al. 2003), 린델과 휘트니(Lindell & Whitney 2001)의 표시변수기법(market variable technique) 등이 있다. 만약 독립변수와 종속변수를 각 다른 시점에서 측정하는 추적조사(longitudinal survey)를 하거나 독립변수의 측정을 설문지를 바탕으로 하는 자기점수화(self-rating) 하거나 종속변수를 컴퓨터를

이용한 기록과 같은 변수를 서로 다른 수단을 통해 측정한다면 이런 편향은 잠재적으로 피할 수 있다.

제 10 장 실험연구

실험연구는 연구 설계상에서 ‘국제 표준(gold standard)’로 여겨지며 연구설계 가운데 가장 엄격하다. 하나 혹은 그 이상의 독립 변수들은 연구자에 의해 처리(treatment)되고, 처리될 변수는 연구에 적합한 수준으로 할당된다. 그리고 종속변수에 영향을 미치는 처리에 따른 결과를 관찰한다. 실험연구의 강점은 연구결과에 영향을 미치는 외생변수의 효과를 사전에 통제할 수 있고 처리(treatment)의 조작을 통한 인과관계를 분석하기 때문에 실험 자체로 **내부 타당성(internal validity)**을 만족하게 된다.

실험연구는 기술적 연구와 탐험적 연구보다 인과관계를 검증하는 설명적 연구에 가장 적합하다. 연구에서 상대적으로 제한되거나 잘 정의된 독립변수는 조작되거나 통제되어 실험실 혹은 실제 현장에서 수행된다. **실험실 실험(laboratory experiment)**은, 실험실과 같은 인공적인 공간에서 이루어지기 때문에 높은 내부 타당성을 갖게 되는 경향이 있지만, 외부 타당성은 낮게 나타나는 결과를 가져온다. 왜냐하면, 실험실이라는 공간은 실제 세계를 전부 반영 할 수 없기 때문이다. **현장실험(field experiment)**은 실제와 같은 현장 공간에서 실험이 진행되고 내부 타당성 및 외부 타당성이 모두 높게 나타난다. 그러나 현장실험연구는 처리(treatments)를 조작하고 외생변수의 효과를 통제하는 것이 어렵기 때문에 다른 연구방법과 비교하여 상대적으로 흔히 사용하는 기법은 아니다.

실험연구는 두 가지 큰 범주로 집단화될 수 있다. 진실실험설계(true experimental design)와 준실험설계(quasi-experimental design)가 그것이다. 두 설계 모두 처리(treatment)에 대한 조작이 필요하지만, 진실실험설계는 임의추출이 요구되고 준실험설계는 그렇지 않다는 차이점이 있다. 비실험연구(non-experimental research)는 실질적으로 연구설계가 아니라, 설문연구, 관찰연구, 상관관계연구와 같이 처리에 대한 조작을 실행하지 않는 모든 종류의 연구를 포함하는 용어이다.

10.1 기본 개념

처리집단과 통제집단. 실험연구에서 몇몇 대상은 처리(treatment)라 불리는 한 개 이상의 실험 자극을 부여 받게 되는데 이를 **처리집단(treatment group)**이라고 한다. 반면에 자극을 부여 받지 못한 다른 대상들은 **통제집단(control group)**이라고 한다. 처리집단 내의 대상이 통제집단 내의 대상보다 결과 변수에 우호적으로 나타날 경우, 처리는 성공적이라고 한다. 실험자극은 다중적인 단계로 부여될 수 있는데 이 경우, 하나 이상의 처리집단을 가지게 된다. 예를 들어, 치매와 같은 특정한 의학 상태를 다루기 위한 신약의 효과를 검증하는 실험연구에서 치매환자 집단을 임의적으로 세 집단으로 분류한다고 가정해 보자. 첫 번째 집단은 약을 많이 복용하게 하고, 두 번째 집단은 약을 적게

복용하게 하고, 세 번째 집단은 설탕약과 같은 약리학적으로는 전혀 효과가 없는 약을 제공해 주게 될 경우, 첫 번째와 두 번째 집단은 실험집단이고 세 번째 집단은 통제집단이다. 약을 투여한 후, 실험 집단의 대상이 통제집단보다 환자의 상태가 상당히 호전되었다면 우리는 약이 효과적이라고 말할 수 있다. 또한 약의 복용 양으로부터 두 실험 집단의 차이를 비교하고자 한다면 많은 약을 복용한 집단이 적은 약을 복용한 집단에 비해 더 효과적인지 검증 할 수 있다.

처리에 대한 조작. 실험연구에서 처리는 타 연구 방법들과 다른 독특한 특성을 가지고 있다. 처리에 대한 조작은 원인-결과 관계의 원인(cause)을 통제하는 데 도움이 된다. 실험연구의 타당성은 처리를 얼마나 잘 조작하였는가에 달려 있다. 처리에 대한 조작은 실제로 실험연구에 앞서 예비검사를 통해 살펴보아야 한다. 측정은 처리를 부여하기 전에 행해지는 **사전측정(pretest measure)**과 처리를 부여한 후에 행해지는 **사후측정(posttest measure)**이 있다.

임의선택과 임의추출. 임의선택(random selection)은 모집단이나 표본들에서 표본을 임의적으로 얻는 과정이다. 이 접근법은 조사연구에 전형적으로 사용되고 있는 방법이고 모집단 내에 있는 개별 단위는 표본으로 선택될 수 있는 기회를 가지고 있다. 반면에 임의추출(random assignment)은 표본을 실험집단 혹은 통제집단에 임의로 추출하는 과정을 말한다. 임의선택은 표본추출(sampling)과 관련이 있기 때문에 연구결과와 외적 타당성(일반화 가능성)과 긴밀한 관련을 가지고 있다. 반면, 임의추출은 설계와 관련이 있기 때문에 내적 타당성과 관련성이 크다. 잘 설계 된 실험연구에서는 임의선택과 임의추출을 동시에 가지고 있을 수 있다. 하지만, 준 실험연구는 임의선택과 임의추출 모두를 포함하지 않는다.

내적 타당성의 위험요소. 실험설계가 다른 연구방법들과 비교하여 추론에 대한 내적 타당성이 더 엄격하다. 처리조작을 통한 원인을 통제할 수 있는 장점 때문에 내적 타당성 대한 위험영향이 없는 것이다. 하지만 내적 타당성에 대한 몇 가지 위험요소들은 존재한다. 예를 들어, 고등학생들의 수리 능력을 개선하기 위한 보충학습프로그램의 효과에 대한 연구이다.

- **과거사건위험요소.** 관찰된 결과(종속 변수)가 실험처리보다 외생변수나 과거의 사건에 의해 영향을 받을 가능성이 존재하는 경우이다. 예를 들어, 보충학습프로그램으로 인한 점수 향상은 프로그램의 수행 결과라기보다는 학교에서 수학 시험을 위한 준비하는 단계에서 영향을 받을 가능성이 있다는 것이다.
- **성숙위험요소.** 관찰된 결과가 실험에 대한 처리보다 그 대상의 자연스러운 성숙에서 비롯될 가능성이 있다는 것이다. 예를 들면, 복잡한 개념을 이해하는 지적 능력의 일반적인 향상이 있다.
- **검증위험요소.** 응답자의 사후반응은 그들의 선행반응에 의해 결정되는 예비-사후 설계에서 나타나는 요소이다. 예를 들어, 학생들이 사전검사에서의 그들이 응답한 내용을 기억한다면, 그들은 사후검사에서 그 답안을 반복적으로 작성하려는 경향을 가지게 된다. 예비검사를 수행하지 않는 것이 이 위험요소를 제거할 수 있는 방법이다.

- **수단위협요소.** 사전-사후에서 발생하며 사전검사와 사후검사에서의 점수의 차이는 보충 수학 프로그램 때문이 아니라, 사후검사의 난이도가 사전검사보다 높고 낮음 등의 측정 수준의 변화로 인해서 발생할 수 있는 가능성이 있다.
- **피험자탈락 위험요소.** 처리집단과 통제집단 내에서 학생들의 수학 공부에 대한 중도 포기 비율이 다를 가능성이 있다. 이러한 중도포기는 사전검사에서 낮은 점수를 받은 학생에게 나타나기 쉬운 시스템적 이유 때문에 발생한다. 만약 수학점수가 낮은 학생이 중도에 포기한다면, 사후검사의 결과는 높은 점수를 받은 학생들의 우세로 인해 인위적으로 부풀려 질 수 있다.
- **회귀위협요소.** ‘평균으로의 회귀’ 라고도 불리며, 사후검사에서 측정된 집단의 전체 점수가 기대하는 방향보다 평균을 향해 회귀하는 통계학적 경향을 의미한다. 예를 들어, 대상이 사전검사에서 높은 점수를 받았다면, 평균에 가까워지기 위해 사후검사에서 낮은 점수를 받는 경향이 있다. 사전검사에서의 높은 점수는 평균에서 떨어져 통계학적으로 일탈(aberration)현상 일 수 있기 때문이다. 이 문제는 두 측정 수단이 불완전한 상관관계를 지니고 있을 때 비-임의 표본에서 더 지배적으로 나타난다.

10.2 이집단 실험설계

가장 간단한 진실실험설계는 하나의 처리집단과 하나의 통제집단으로 구성하는 두 집단의 실험 설계이다. 이것은 처리(treatment)로 조작 가능한 독립변수의 효과를 검증하는데 가장 적합하다. 이집단 실험설계는 사전·사후검사 통제집단설계와 사후검사통제집단설계로 구분되며 변동은 공분산설계를 포함하기도 한다. 이 설계들은 표준화된 표기법을 활용하여 표현되는데, R 은 집단의 대상에 대한 임의적인 추출을 나타내고, X 는 처리집단에 처리가 부여된 것을 나타내며, O 는 종속변수의 사전검사 혹은 사후검사의 관찰을 나타낸다. 이러한 표식은 처리집단과 통제집단에 대한 사전검사와 사후검사의 관찰결과를 구별하기 위함이다.

사전·사후검사통제집단설계(pretest-posttest control group design). 이 설계에서 대상은 처리집단과 통제집단을 임의적으로 추출하고 사전검사에서 종속변수에 대해 초기 값을 측정한 후 처리집단을 독립변수로 표현되는 처리(treatment)로 부여하고 사후검사에서 종속변수를 다시 측정한다. 다시 말해서 처리를 통한 종속변수의 사전·사후검사의 차이를 비교하여 실험의 효과를 측정하는 것이다. 이 설계의 표기법은 그림 10.1 에서 보여진다.

R	O ₁	X	O ₂	(처리 집단)
R	O ₃		O ₄	(통제 집단)

<그림 10.1> 사전·사후검사통제집단설계

사전·사후검사설계에서 실험적 처리의 효과 E 는 처리집단 및 통제집단의 사전-사후검사 결과간의 차이를 통해 측정된다. 이 설계의 통계학적 분석은 처리집단과 통제집단 사이의 단순 분산분석(ANOVA)을 통해 이루어진다. 사전·사후검사설계는 피험자 탈락(mortality), 검증, 회귀와 같은

내부 타당성에 영향을 미치는 위협요소를 다루는데 이러한 위협요소는 처리집단과 통제집단에 비슷한(임의적인) 방식으로 영향을 미칠 수 있다. 선택 위협요소는 임의추출을 통해 통제될 수 있지만 내적 타당성에 영향을 미치는 부수적인 위협요소는 여전히 존재할 수도 있다. 예를 들어, 피험자 탈락(mortality)은 두 집단 사이에 중도 포기 비율로 나타나는 문제점이 발생할 수 있으며 사후검사가 평범하지 않은 주제나 내용을 다루고 있을 경우, 사전검사는 사후검사를 편향되게 만들기도 한다.

$$E = (O_2 - O_1) - (O_4 - O_3)$$

사후검사통제집단설계(posttest-only control group design). 이 설계는 사전검사를 생략하는 사전 사후검사 통제집단 설계의 단순한 형태이다. 설계의 표기법은 그림 10.2 에서와 같이 나타난다.

R	X	O ₁	(처리 집단)
R		O ₂	(통제 집단)

<그림 10.2> 사후검사통제집단설계

처리효과는 두 집단간의 사후검사 결과를 비교 함으로써 간단하게 계산할 수 있다. 이 설계의 적절한 통계학 분석은 두 집단의 분산분석(ANOVA)을 통해 알 수 있다. 이 설계는 내적 타당성 측면에서 사전·사후검사통제집단설계보다 더 매력적이라 할 수 있다. 그리고 성숙, 검증, 회귀, 선택, 사전 사후검사의 상호작용에 대한 위협요소는 통제할 수 있지만, 피험자 탈락(mortality) 위협요소는 여전히 존재한다.

$$E = (O_1 - O_2)$$

공분산설계(covariance design). 종속변수의 측정은 공분산(covariance)이라고 불리는 외생변수에 의해 영향을 받는다. 실험연구에서 공분산이 주요 관심 변수는 아니지만, 종속변수에 영향을 미치는 잠재 효과를 제거하기 위해 통제되어야 한다. 그러므로 독립변수의 효과를 좀 더 정확하게 측정할 수 있다. 지금까지 논의된 실험설계는 이러한 공분산을 통제하지 않았다. 공분산설계는 상반변수설계라고도 불리며 사전 사후검사 통제집단설계의 특별한 형태이며 사전검사에서 종속변수보다 공분산을 측정하는데 중점을 둔다. 설계 표기법은 그림 10.3 에서와 같이 나타난다. 아래 그림에서 C 는 공분산을 나타낸다.

R	C	X	O ₁	(처리 집단)
R	C		O ₂	(통제 집단)

<그림 10.3> 공분산설계

사전검사가 종속변수가 아닌 공변량에 비중을 두어 측정하기 때문에 사후검사에서 처리효과는 처리집단과 통제집단 사이의 차이로 측정한다.

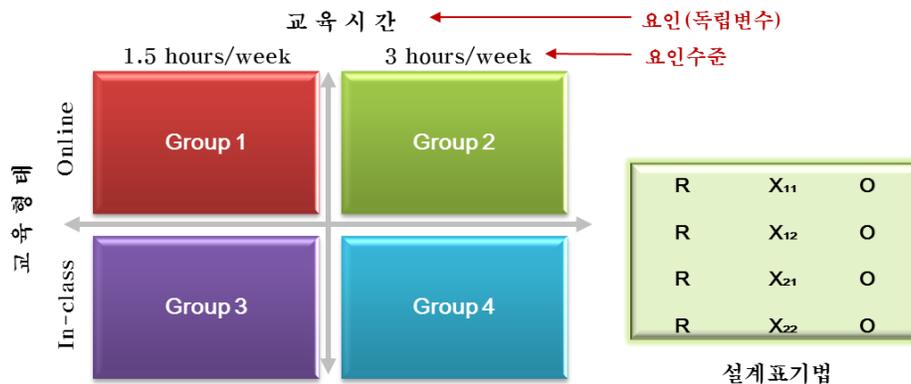
$$E = (O_1 - O_2)$$

공분산의 존재로 올바른 통계학적 분석은 두 집단의 공분산분석(ANCOVA)이다. 이 설계는 사후검사통제집단설계의 모든 장점을 가지고 있으면서 공변량을 통제하기 때문에 내부 타당성을 확보하게 된다. 공분산설계는 사전·사후검사통제집단설계로 확장할 수 있다.

10.3 요인 설계

만약 두 개 이상의 독립변수(처리)를 조작할 경우, 이 집단 설계는 적절하지 않다. 이 경우 우리는 그 이상의 집단설계(higher-group design)가 필요하며, 이를 실험연구에서 흔히 사용되는 것으로 요인설계라고 부른다. 각각의 독립변수를 요인(factor)이라 하고 개별요인의 하부구조를 단계(level)라고 한다. 요인설계는 연구자가 종속변수에 대한 각각의 처리된 개별 주요효과(main effect)를 검증할 뿐만 아니라, 독립변수들 간의 상호작용효과(interaction effect)도 검증한다.

가장 기본적인 요인 설계는 2×2 요인 설계이며, 두 개의 처리로 구성되며 각각 두 단계(높고 낮음/존재와 존재하지 않음)의 하부구조를 갖는다. 예를 들어, 두 가지 형태의 교육형태(강의실 수업과 온라인 지도)의 학습결과를 비교하고, 지도시간(주당 1.5 시간 혹은 3 시간)에 따라 효과가 다를음을 증명하고 싶다고 가정해 보자. 이 경우, 교육형태와 지도시간이라는 두 가지 요인이 있다. 두 요인 모두 두 단계를 갖게 된다(그림 10.4). 만약 지도 시간의 세 번째 단계(주당 6 시간)를 추가하고자 한다면, 두 번째 요인은 세 가지 단계를 가지게 되며, 전체적으로 2×3 요인 설계로 변한다. 반면에, 집단작업의 존재 여부에 대한 세 번째 요인으로 추가하고자 한다면, 2×2×2 요인 설계가 된다. 이 표기법에서 각각의 숫자는 요인을 나타내 주고, 각 요인에 대한 값은 요인 내의 단계를 나타낸다.



<그림 10.4> 2 x 2 요인 설계

그림 10.4의 오른쪽 박스에서 보듯이 요인 설계는 설계 표기법을 활용하여 표현할 수 있다. R은 처리집단에 대한 임의 추출한 것을 나타낸다. X는 처리집단 자체를 나타내며, X의 아래 기입된 숫자는 요인의 단계를 나타낸다. O는 종속 변수에 대한 관찰을 나타낸다. 2×2 요인 설계는 네 가지 처리집단을 가지게 되며, 각각의 요인에 대해 두 가지 단계를 갖게 되어 총 네개의 결합물을 가지게 되는 것이다. 더불어 2×3 요인 설계는 여섯 개의 처리집단을 가지게 되고, 2×2×2 설계는 총 8개의

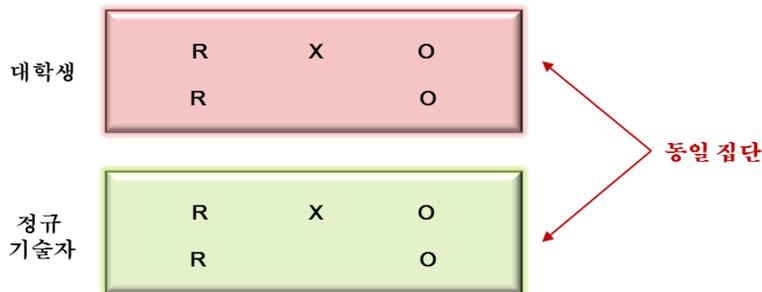
처리집단을 가지게 된다. 코헨(Cohen)의 통계적 파워(statistical power) 계산방식에 의거하여 기본적으로 요인 설계에서는 개별 셀의 표본크기가 최소 20 개 이상이 되어야 한다. 2×2×2 요인 설계는 각 셀에 20 개의 표본이 있어야 하므로 최소 160 개 이상의 표본이 필요하게 된다. 이처럼 단계나 요인을 늘리는 것은 자료수집 비용을 상당히 증가하게 만든다. 제한된 자원으로 인하여 때때로 요인설계에서 어떤 셀은 아무런 처리를 하지 않을 수도 있는데 이를 불완전요인설계(incomplete factorial design)이라고 한다. 이러한 불완전요인설계는 불완전한 요인에 대한 추론으로 도출하는 능력을 손상시키기도 한다.

요인설계에서 **주효과(main effect)**는 종속변수가 한 가지 요인에 의해 여러 단계에서 의미있는 차이를 보이거나, 다른 요인이 모든 단계에 영향을 미치게 될 경우 존재한다. 주요효과의 평가는 요인 단계에 따른 종속변수의 변화가 없는 경우, 가치가 없는 것으로 간주한다. 위의 예에서 학생들을 가르치는 교육형태와 지도시간 모두 학습 결과에 있어 주요효과를 가지고 있음을 발견할 수 있을 것이다. **상호작용효과(interaction effect)**는 하나의 요인이 두 번째 요인의 단계에 따라 종속변수의 효과에 차이를 발생 시킬 경우 존재하게 된다. 예제에서 보듯이, 교육형태에 따른 학습효과가 교육시간이 주당 3 시간보다 주당 1.5 시간에서 높게 나타날 경우, 학습효과에 대한 교육형태와 지도시간 사이의 상호작용효과가 있다고 말할 수 있는 것이다. 상호작용효과의 존재는 주요효과와 관련성을 없게 하거나 주요효과를 해석하는 것이 무의미하게 만들 수 있다.

10.4 하이브리드 실험설계(Hybrid Experimental Design)

하이브리드 실험설계는 다수의 설계들을 결합한 형태로 무선화구획설계, 솔로몬 4 집단설계, 전환반복설계 세 가지가 있다.

무선화구획설계(randomized block design). 실험설계가 반복되는 동안에 개체 모집단이 동종의 하위집단으로(블록이라고 칭함) 분류되는 사전 사후검사 통제집단설계와 사후검사통제집단의 변동을 말한다. 예를 들어, 두 개의 동일 집단으로 이루어진 대학생들과 정규기술자들 간의 동일한 사후검사설계를 반복하려면 두 블록의 개체는 임의적으로 처리집단(동일한 처리를 받음) 또는 통제집단으로 나누어지게 된다(그림 10.5). 무선화구획설계의 목적은 ‘잡음(noise)’ 또는 블록간의 차이에 기인하는 데이터의 변동을 줄임으로써 실제효과가 더 정확히 발견될 수 있도록 하는 것이다.



<그림 10.5> 무선화블록설계

솔로몬 4 집단설계(solomon four-group design). 이 설계에서 표본은 두 개의 처리집단과 두 개의 통제집단으로 나누어진다. 하나의 처리집단과 하나의 통제집단은 사전검사를 하고 나머지 처리집단과 통제집단은 검사를 하지 않는다. 이 설계는 사전·사후검사통제집단설계와 사후검사통제집단설계의 결합을 나타내며 사후검사설계에서 아닌 사전·사후검사설계에서의 사후검사측정에서 사전검사측정의 잠재적 편향효과를 검사하기 위함이다. 솔로몬 4 집단 설계의 표기는 그림 10.6 과 같다.

R	O	X	O
R	O		O
R		X	O
R			O

<그림 10.6> 솔로몬 4 집단 설계

전환반복설계(switched replication design). 측정의 세 파동(waves)과 두 위상(phases)에서 시행되는 이집단 설계이다. 첫 번째 위상에서의 처리집단은 두 번째 위상에서 통제집단이 되며, 첫 번째 위상의 통제집단은 두 번째 위상에서 처리집단이 된다(그림 10.7). 다시 말해서 본래의 설계는 두 집단간의 처리/통제 역할을 전환하여 반복하는 것이다. 그리고 이 연구의 마지막에서 모든 참가자들은 첫 번째 또는 두 번째 위상에서 처리를 받게 될 것이다. 이 설계는 단계적으로 이루어지거나 계획된 일정대로 반복적으로 일어나는 조직 상황을 검증하는데 가장 유효하다.

R	O	X	O	O
R	O		O	X

<그림 10.7> 전환반복설계

10.5 준실험설계(Quasi-Experimental Design)

준실험설계(quasi-experimental design)는 진실실험설계와 거의 동일하다. 두 설계의 차이점 가운데 준실험설계의 핵심요소는 무작위 추출이다. 예를 들어, 전체집단 또는 하나의 조직이 처리집단으로 사용되면 동일한 산업 내에서 다른 집단 혹은 타 조직이 통제집단이 된다. 무작위 추출의 약점은 하나의 집단이 다른 집단에 비해 특정 내용을 많이 가짐으로써 모집단에서 표본추출의 집단 간 동일하지 않는 선택편향(selection bias)의 문제를 가진다. 따라서 준실험설계는 위협에 연관된 선택의 다양성 때문에 내부 타당성 면에서 진실실험설계의 수준에 미치지 못한다. 위협에는 선택-성숙 위협(비율에서 성숙되는 처리집단과 통제집단), 선택-과거사건의위협(외부 요인이나 역사적인 사건에 따라 영향을 받는 처리집단과 통제집단), 선택-회귀위협(비율에 따라 사전검사와 사후검사 간의 평균으로 회귀하는 처리집단과 통제집단), 선택-수단위협(측정법에 따른 처리집단과 통제집단), 선택-검사(사전검사에 따른 처리집단과 통제집단), 선택-피험자탈락(차등이 있는 중퇴율에 따른 처리집단과 통제집단)이 존재한다. 일반적으로 이러한 선택위협 가능성 존재할 경우 준실험설계를 피하는 것을 따른다.

진실험설계가 무작위 추출을 생략함으로써 준실험설계로 변환될 수 있다. 예를 들어, 사전·사후검사통제집단설계의 준 동질화 상태를 비동등집단설계(nonequivalent groups design: NEGD)라고 하며 무작위 추출 R 은 비임의적 추출 N 으로 교체된다(그림 10.8). 이처럼 전환반복설계의 준실험설계를 비동등전환반복설계(non-equivalent switched replication design)라고 한다(그림 10.9).

N	O	X	O
N	O		O

<그림 10.8> NEGD 설계

N	O	X	O	O
N	O		O	X

<그림 10.9> 비동등전환반복설계

이 밖에 진 실험 설계에는 해당하지 않는 특유의 비동등설계가 몇 가지 있다. 이러한 설계들에 해 알아보자.

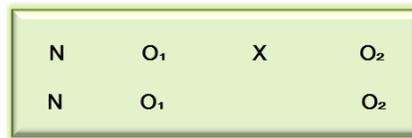
불연속회귀분석 설계(regression-discontinuity design). 이것은 비동등한 사전·사후검사설계이다. 대상은 사전 측정을 통한 최종허용 점수를 바탕으로 처리집단과 통제집단으로 구분한다. 예를 들어 심각하게 아픈 환자는 신약의 효과를 검사하기 위해 처리집단에 할당되고 증상이 가벼운 환자는 통제집단에 할당 될 것이다. 다른 예를 들자면, 시험점수가 평균 이하인 학생들은 그들의 성적 향상을 위한 보충학습프로그램에 참가하는 처리집단으로 선택되고 평균 이상의 높은 점수를 받은 학생들은 보충학습프로그램에 선택되지 않는다. 이 설계의 표기는 다음과 같으며 C 는 최종허용점수를 나타낸다.

C	O	X	O
C	O		O

<그림 10.10> 불연속 회귀분석 설계

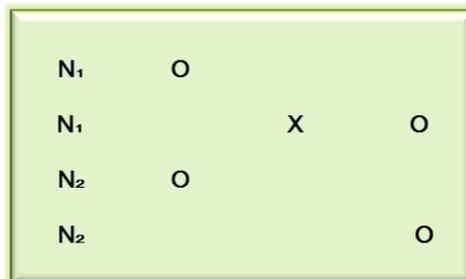
최종허용점수로 인하여 관찰된 결과는 내부 타당성에 영향을 미치는 새로운 위협함수가 될 수 있다. 하지만 최종허용점수를 사용하는 것은 제한적이기 때문에 비용이 높은 자원들을 모집단 내에서 임의추출보다 그것들을 가장 필요로 하는 집단에게 배정하는 것으로 준실험처리와 동시에 이루어진다. 불연속회귀분석설계에서의 통제집단 점수는 처리집단 점수를 비교한 기준이 아니다. 만약 통제집단에서 사전검사와 사후검사 점수 사이에 불연속성이 없고 처리집단에서 불연속성이 존재한다면, 불연속성은 처리효과의 증거로 본다.

대리사전검사(proxy pretest design). 그림 10.11 에서 보듯이 대리 사전검사는 사전검사 점수의 처리가 적용된 후에 측정된다는 한 가지 중요한 차이점을 제외하고는 비동등집단설계(NEGD)와 매우 유사하다. 연구자가 사전검사 데이터를 측정할 수 없는 경우 혹은 어떤 프로그램의 효율성을 측정하기 위한 연구에서 그것이 이미 시작한 후에는 일반적으로 적용하여 사용된다. 이런 상황에서 연구자가 대처할 수 있는 가장 최선의 선택은 서로 다르게 측정된 사전기록을 이용하는 것이다. 예를 들어, 교육 프로그램의 효율성을 측정하고자 할 경우, 사전에 기록 된 학생들의 평점이 사전검사 데이터가 될 수 있다. 하지만 대리사전검사의 변동은 대상의 사전검사 데이터의 사후검사의 재수집된 데이터를 사용할 지도 모른다. 이것은 사전검사 데이터를 기억하는 것이 소환경향(recall bias)을 가져올 수 있지만 그림에도 불구하고 대리사전검사의 변화는 종속변수의 변화를 측정할 수 있도록 도와준다.



<그림 10.11> 대리사전검사설계

독립표본사전·사후검사설계(separate pretest-posttest samples design). 동일한 대상으로부터 사전검사와 사후검사를 측정한 데이터를 수집하기 어려울 때 유용하다. 그림 10.12 에서 보듯이, 이 설계의 네 집단 가운데 두 집단은 하나의 단일 비동등 집단이며, 다른 두 집단은 또 다른 하나의 비동등집단으로 구성된다. 예를 들어, 한 도시에서만 시행되는 새로운 온라인 서비스에 대한 소비자 만족을 검사하고자 할 경우, 온라인 서비스를 시행하는 도시의 소비자는 처리집단이 되고 그렇지 않은 도시의 소비자는 통제집단이 된다. 만약, 동일 소비자들로부터 사전검사와 사후검사의 측정이 어렵다면, 소비자 만족을 서로 다른 소비자 집단으로 나눈다. 그리고 새로운 프로그램을 시행한 후, 동일한 시점에서 소비자 만족을 측정한다. 소비자 만족은 처리집단이기도 하며 통제집단으로도 측정된다. 이 설계는 어떤 특정한 소비자 만족의 변화를 시행 이전과 이후의 측정을 할 수 없고 소비자 만족의 평균 값만을 알 수 있기 때문에 강력한 실험설계 방법은 아니다. 하지만 이 설계는 내부 타당성이 낮음에도 불구하고 여전히 동일한 대상으로부터 사전 사후검사로부터 측정된 데이터를 수집하는 것이 불가능 할 경우 유용한 방법이다.

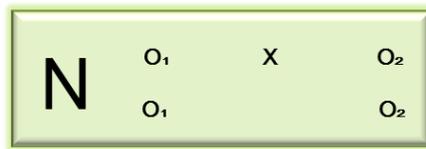


<그림 10.12> 독립표본 사전사후검사 설계

비동등종속변수설계(nonequivalent dependent variable design; NEDV). 종속변수에 대해 두 개의 측정방법을 가지고 있는 단일-집단 사전·사후준실험설계이다. 이론적으로 하나의 측정은 처리에 의해

영향을 받고 다른 측정은 그렇지 않는 것이다. 예를 들어, 고등학생들을 위한 새로운 미적분학 교육과정을 만들고자 한다면 이 교육과정의 사후검사 결과는 대수학 점수가 아닌 미적분학 점수와 관계가 있을 것이다. 하지만 대수학 점수의 사후검사 역시 과거 사건과 성숙 위험요소와 같은 외부 요소들 때문에 경과에 영향을 미칠 가능성이 있다. 그러므로 대수학의 사전사후검사 결과는 통제 측정 값으로 사용하고 미적분학의 사전·사후검사는 처리 값으로 사용한다. 그림 10.13 에서 보듯이 단일집단은 N 으로 나타내고 동일집단의 미적분학과 대수학의 사전검사는 O_1 , 사후검사는 O_2 로 나타낸다. 이 설계는 내부 타당성 측면에서 약하지만 독립된 통제집단을 사용하지 않는다는 것이 장점이다.

비동등중속변수설계의 변동성은 다수의 결과변수와 각각의 변수가 처리에 의해 얼마나 영향을 받는지 설명하는 이론을 적용한 패턴을 결합한 설계이다. 이것은 연구자가 이론적 예측 값이 실제 관측 값과 일치되는지 여부를 검사할 수 있다. 따라서 패턴매칭기술(pattern-matching technique)은 이론과 실제 관찰된 패턴 사이의 일치하는 정도를 바탕으로 본래 비동등 중속변수 설계의 낮은 내부 타당성에 대한 염려를 완화시키는 강력한 방법이다.



<그림 10.13> NEDV 설계

10.6 실험연구의 위험

실험연구는 가장 어려운 연구설계 중 하나로 다른 방법론들과 충돌이 잦아 그 선택에 있어서 신중하게 고려되어야 한다. 실험연구가 가지고 있는 위험성들로는 첫째, 실험연구가 가설설정과 검증에 의해 이론이 요구됨에도 불구하고 대부분의 실험연구들이 비이론적이다. 이론 없이 검증된 가설은 일시적인 경향이 있으며 의미가 없거나 비논리적일 수가 있다. 둘째, 실험연구에서 사용되는 다수의 측정도구들은 신뢰성과 타당성을 검증하지 않으며 연구들 사이에서 어떤 측정도구가 유효한지 견줄 수 없다. 따라서 이러한 측정수단을 이용하여 만들어진 결과 역시 비교할 수 있는 부분이 사라지게 된다. 셋째, 많은 실험연구들이 관련 없는 종속변수나, 상호작용이 존재하지 않거나, 통제 요소가 없고 처리집단에 동일하지 않는 처리가 존재하지 않는 부적합한 연구설계를 사용한다. 이러한 연구로부터의 결론들은 내부 타당성이 약할뿐만 아니라 결과에 대한 의구심을 유발하기 쉽다. 넷째, 실험연구에서 사용되는 처리(treatment)는 매우 다양하고, 비교할 수 없으며, 연구마다 동일하게 나타나지 않고, 개체 모집단이 적절하지 않게 사용되기도 한다. 예를 들어, 전문적인 경험이 전무한 대학생들에게 마케팅 매니저인 척하면서 복잡한 예산할당업무를 요청하는 경우가 있다. 잘못된 조사대상으로부터 수집된 결과는 그들로부터 적합한 업무의 수행결과를 이끌어 내지 못하여 내부 타당성에 대한 새로운 위협을 발생시킨다. 이것은 결과적으로 해석이 불가능하고 의미가 없으며 타 연구들과 통합하는 것 역시 불가능 하게 된다.

실험연구 설계에서 적절한 실험처리의 설계는 매우 중요한 작업이다. 처리는 실험방법의 이유이자 목적이기 때문에 반드시 고려되어야 한다. 적합한 실험처리 설계 작업을 하기 위해서 연구자는

사전에 검증된 작업으로 수행하고 처리조작 점검을 수행하고, 필요하다면 반복적으로 예비검사를 실시한다. 또한 응답자들에게 단순하고 친숙한 작업을 할 수 있도록 설계한다.

요약하자면, 이번 장에서는 실험연구 설계방법의 핵심 개념을 소개하고 다양한 진실험설계와 준실험설계에 대해서 소개했다. 실험연구가 대부분 내부 타당성이 있을 지라도 내부 타당성이 낮은 설계는 간과되어서 안 된다.

제 11 장 사례연구

사례연구는 자연적인 배경 안에서 하나 혹은 그 이상의 측면에 대해 오랜 시간을 거쳐 집중적으로 연구하는 방법이다. 인터뷰, 관찰, 사전기록문서, 이차자료분석과 같은 다양한 자료수집 방법들은 연구자가 관심 있는 현상에 대하여 풍성하고, 상세하고 연구 현상과 관련된 추론을 얻는데 활용될 수 있다. 사례연구는 실증주의적 방식을 활용한 이론 검증과 해석적 방식을 이용한 이론 구축이 가능하다. 이 방법은 다른 사회과학 분야보다 경영학 연구 분야에서 널리 활용되고 있다.

사례연구는 실험연구와 조사연구와 같은 연구방법들과 비교하면 독특한 강점들을 지니고 있다. 첫째, 사례연구는 이론 구축 및 검증에 모두 활용되며, 특히 실증분석방법은 이론검증을 통해 연구한다. 해석적 사례연구의 경우, 알려지지 않은 구성개념이 연구과정 중에 드러나기도 한다. 둘째, 연구 초기의 연구과정에서 그것과 관련성이 적거나 핵심적인 부분을 다루지 않을 경우에는 연구질문이 수정될 가능성이 있다. 하지만, 이것은 자료수집 단계가 완료된 후의 실증 방법에서는 불가능하다. 셋째, 사례연구는 연구맥락과 관련된 자료를 풍부한 집합체들로부터 탐색할 수 있는 장점을 가지고 있기 때문에 다른 연구방법과 비교하여 풍부하고, 현상을 설명하는 맥락에 적합하고, 정확한 해석을 가능하게 도와주는 역할을 한다. 넷째, 연구자의 관심현상은 다양한 참여자들의 관점으로부터 연구되며, 단위분석(예: 개인과 조직)이 가능하다.

사례연구는 다음과 같은 몇 가지 약점을 가지고 있다. 첫째, 실험처리나 통제 혹은 추론의 내적 타당성을 포함하지 않는다. 물론 이러한 문제는 실험연구를 제외한 모든 연구방법의 문제이기도 하다. 하지만, 사례연구에서 활용하는 ‘자연적 통제(natural control)’가 통제의 문제점으로 중요하게 다뤄진다. 둘째, 사례연구를 통해 도출된 추론의 질은 연구자의 통찰력에 의하여 좌우된다. 숙련된 연구자는 초보 연구자가 놓칠 수 있는 개념이나 패턴 등을 더 잘 찾아낼 수 있다. 그렇기 때문에 결과물은 때때로 주관적이라고 비평 받기도 한다. 마지막으로, 추론은 연구맥락과 관련이 깊기 때문에 다른 맥락이나 다른 조직에 적용한 추론을 일반화하기 어렵다.

사례연구는 경영 수업에서 배우는 하버드 사례연구와 같은 사례서술(case description)과는 다르다는 점을 명심해야 한다. 사례서술은 전형적으로 학생들 사이에서 토론 및 비판적 사고를 촉진하기 위한 목적을 가지고 조직의 문제와 조직이 어떻게 문제를 해결했는지 상세히 묘사하는 것이라면, 사례연구는 과학적인 방법을 활용하여 조직의 현상에 대한 설명을 도출하기 위한 연구 방법이라 할 수 있다.

사례연구방법은 연구자의 연구 능력을 요구하는 어려운 연구기법이기 때문에 오류를 범하기 쉽다. 벤바사와 그의 동료들은 (Benbasat et al., 1987)⁸ 사례연구에서 자주 발생하는 다섯 가지 문제에 대해 서술하였다. 첫째, 많은 사례연구들이 특정한 연구질문 없이 시작되었기 때문에 그에 따른 특정한 해답을 얻거나 통찰력 있는 추론으로 끝나지 못하는 경향이 있다. 둘째, 사례의 범위는 연구질문에 들어 맞는 것이 선택되기 보다는 종종 접근성과 편리성에 의해 선택이 되기 때문에 연구질문에 적절한 답을 얻지 못할 수 있다. 셋째, 연구자들은 다중평균을 활용하여 수집된 데이터에 대해 측정하지 않아 편향된 인터뷰 대상으로부터의 반응에 대한 편향된 해석이 되는 것이다. 넷째, 많은 연구들은 연구 당시, 어떤 인터뷰 질문이 사용하였는지, 어떤 문서형태로 분석하였는지, 인터뷰 대상의 조직차원의 포지션은 각각 어떻게 구성되었는지 등에 대한 상세한 자료를 제시해주지 않는다. 이것은 수집된 데이터의 신뢰성에 대해 물음을 던지게 된다. 마지막으로, 사례연구가 종단연구로서의 강점에도 불구하고 많은 사례연구들은 현상에 대해 종단적인 방식을 따르지 않고 현재 과정 및 현상에 대한 횡단적이고 제한적인 의견을 나타내기만 한다.

11.1 사례연구에서의 주요 의사결정

사례연구를 고려할 때, 연구자에 의한 몇 가지 의사결정이 필요하다. 첫째, 사례연구가 연구질문에 대한 해답을 얻기 위한 정확한 방법인가? 특히, 사례연구 방법은 구성개념들간의 관계를 발견해 나가는 이론구축이 형성단계에서 적절한 탐구적 연구이다. 그리고 경험과 행동요인에 대한 비판적 연구 또는 원인과 결과보다 복잡하고 일시적인 과정을 이해하기 위한 목적이 있는 연구이다. 따라서 사례연구는 조직의 변화나 큰 규모의 기술 실행 프로젝트와 같은 다수의 참여자와 상호작용하는 일련의 사건을 포함하고 복잡한 조직적인 연구에 잘 들어 맞는다.

둘째, 무엇이 사례연구를 위한 적절한 분석단위인가? 사례연구는 다양한 분석단위로 조사가 가능하기 때문에 연구자는 현상을 관찰하기에 앞서 개인, 집단, 조직 혹은 다중 레벨 가운데 분석단위를 결정해야 한다. 예를 들어, 집단 의사결정이나 집단 작업에 대한 연구는 집단 활동 안에서 개인의 참여와 같은 개인 단위의 구성개념과 집단의 응집력 혹은 집단 리더쉽과 같은 집단 단계의 구성개념을 활용한다. 이것은 개인 단계에서 집단 단계까지의 구성개념을 통합함으로써 연구를 보다 더 풍부하게 이해하기 위함이다.

셋째, 연구자는 단일사례 혹은 다중사례 설계를 연구에 적용해야 하는가? 단일사례 설계는 상황이 매우 독특하거나 일반적으로 과학적인 조사를 통한 접근이 불가능할 경우, 잘 형성된 이론을 검증하기 위한 비판적인 사례가 있는 경우에 이론 형성을 위한 단계에서 매우 적합하다. 반면, 다중 사례 설계는 이론검증을 위한 단계에서 적합하다. 예를 들면, 결과를 일반화하는 경우, 현상에 대한 좀 더 풍성한 해석을 위한 목적에 더 잘 부합한다. 인(Yin, 1984)⁹ 은 복제논리를 가진 다중사례연구의 활용을 추천하였다. 그 이유는 각각의 사례는 하나의 실험연구와 비슷하고, 과학적인 엄격성의 비슷한 규칙을 따라 귀납적으로 이론구축을 하고 연역적으로 이론을 검증하였기 때문이다.

⁸ Benbasat, I., Goldstein, D. K., and Mead, M. (1987). "The Case Research Strategy in Studies of Information Systems," *MIS Quarterly* (11:3), 369-386.

⁹ Yin, R. K. (2002), *Case Study Research: Design and Methods*. Thousand Oaks, CA: Sage Publications.

넷째, 사례연구를 위해 어떤 현장을 선택할 것인가? 사례연구를 진행할 수 있는 추론이 주어진 맥락 차원에서 설명되어야 하기 때문에 현장 선택의 과정은 매우 중요하다. 이것은 또한 잘못된 현장 선택으로 잘못된 추론을 이끌 수 있기 때문이다. 만약, 연구의 목표가 이론을 검증 혹은 추론의 일반화를 검증하는 것이라면, 비슷하지 않은 현장선택의 과정은 관찰 값의 분산을 증가시키기 위해서 수행되어야만 한다. 예를 들어, 연구의 목적이 회사의 기술 실행 과정을 이해하기 위함이라면, 대기업, 중견기업, 소기업 등의 회사규모에 따라 기술 실행의 과정이 같은지, 다른 지 검증해야 한다. 현장의 선택은 기회주의적이거나 편리함 등에 기초해서는 안되며, ‘이론적 표집방법’을 따르고 연구질문에 적합해야 한다.

다섯 번째, 사례연구에서 어떤 자료수집기법이 활용되어야 하는가? 인터뷰는 사례연구에서 가장 널리 사용되지만, 인터뷰 자료는 다른 연구기법으로 수집한 데이터의 부족한 부분을 보충하거나 정보를 제공하는 역할을 한다. 앞에서 언급한 다른 연구기법으로는 직접적 관찰(예: 연구자가 경영자 미팅, 브리핑, 계획 단계에 참석하는 것), 참고자료(예: 내부 보고서, 발표자료, 각서, 뉴스기사 등), 기록물(예: 조직 차트, 재무 기록 등), 인공물(예: 기기, 결과물, 도구) 등이 있다. 또한, 연구자는 인터뷰 대상들의 반응을 비교함으로써 관찰된 데이터의 타당성을 검증해야 한다.

11.2 사례연구 진행하기

대부분 사례연구는 해석적인 경향이 있다. 해석적 사례연구는 귀납적인 기법으로써 하나 이상의 현장에서 관찰된 데이터는 새로운 이론을 구축하거나 존재하는 이론확장을 목적으로 개념과 패턴을 체계적으로 분석하고 합성한다. 이센하르트(Eisenhardt 1989)¹⁰는 사례연구로부터 이론구축을 위한 ‘로드맵’을 제안했다. 이 로드맵은 해석적인 사례연구를 수행하기 위해서 이어질 단계들을 재배치하거나 수정하였다. 그러나 표집방법, 자료수집, 분석기법은 기존과 동일하다.

연구질문을 정의하라. 모든 과학적 연구가 그렇듯 사례연구에서도 연구질문은 이론적으로나 실증적으로 흥미로운 것을 바탕으로 정의한다. 연구자는 사례연구의 기반을 세우기 위해서 연구질문에 대한 가능한 대답나 기초 구성개념을 직감을 통해 설정한다. 예상 구성개념은 실증적인 사례연구에서 이론에 기초하여 개발하지만 해석적 연구에서 이론이나 가설이 없이 어렵짐작으로 구성된다. 이러한 연구질문이나 구성개념은 해석적인 사례연구에서 필요하다면 변경될 수 있지만 실증적인 연구에서는 그렇지 않다.

사례현장을 선택하라. 연구자는 사례현장을 구성하기 위해 임의적 표집방법이 아닌 ‘이론적 표집방법’의 과정을 활용해야만 한다. 사례현장의 선택은 통계적 혹은 연구자의 관심보다는 이론적인 특성에 적합하도록 설정해야 한다. 이론적 특징으로는 기초이론을 확장, 이전의 사례를 대체할 만하거나, 이론적 범주를 발전시키기 위함이 있다. 선택된 사례현장은 연구질문에 대한 현상을 잘 설명할 수 있어야 하고 외생변수의 영향을 최소화함과 동시에 종속변수의 측정값을 최대로 할 수 있어야 한다. 연구의 목적이 어떤 기업이 다른 기업들과 비교하여 어떻게 혁신을 잘 하는지를 조사한다면, 연구자는 동일 산업군 내에서 비슷한 규모의 기업들을 통한 자료를 수집해야

¹⁰ Eisenhardt, K. M. (1989). “Building Theories from Case Research,” *Academy of Management Review* (14:4), 532-550.

한다. 이것은 산업과 규모의 영향을 최소화하고 더 혁신적이고, 덜 혁신적인 기업을 선택하여 기업의 혁신 변동성을 잘 설명할 수 있도록 도와주기 때문이다. 연구자는 각 기업 내의 프로젝트 관리자에게 연락한 뒤, 인터뷰하는 동안에 연구에 참여한 연구자, 연구 보조원들의 소개, 연구의 목적과 현상을 조사 대상자에게 설명한다. 그리고 연구를 통한 긍정적인 잠재효과는 물론 수집된 데이터가 어떻게 사용될지, 인터뷰 예상 시간 등을 설명한다. 또한, 연구자는 인터뷰 응답자와 회사에 대한 익명성을 보장해 주어야 한다.

인터뷰 프로토콜을 만들어라. 사례연구에서 주요한 자료수집 형태는 인터뷰이기 때문에 인터뷰 프로토콜은 인터뷰 과정을 안내하기 위해서 반드시 설계되어야 한다. 이것은 일반적으로 질의 응답되어야 할 항목들로 구성되어 있다. 질문은 개방형(비구조화), 폐쇄적(구조화), 또는 두 개의 혼합 형태로 이루어 진다. 인터뷰 프로토콜은 설계한 방법대로 엄격히 수행되어야 하고 인터뷰 과정 동안, 인터뷰 진행자는 질문의 순서를 바꾸거나 어떠한 질문도 생략할 수 없다. 하지만 응답자의 대답이 애매하거나 혹은 흥미롭다면 이와 관련해서 질문하는 것은 허용된다. 인터뷰 진행자는 중립적인 톤을 유지하면서 질문해야 하고, 응답자의 대답을 특정한 방향으로 이끌어서는 안되며, 어떠한 응답에도 동의 혹은 부정의 반응을 보이면 안 된다. 인터뷰 기법에 대한 자세한 설명은 ‘설문조사’ 장에서 논의되었다. 가능하다면, 인터뷰 자료를 보충 및 검증하기 위해 내부 문서, 각서, 연차 보고서, 재무 보고서, 뉴스 기사와 같은 직접적인 관찰을 활용한 부차적인 자료를 수집해야 한다.

응답자를 선택해라. 연구자가 관심 있는 현상을 다양한 관점에서 분석하기 위해 인터뷰 응답자는 조직적, 부서별, 직급별에 따라 선택해야 한다. 응답자들의 임의적인 표본 수집이 가장 선호되는 편이지만, 눈덩이 표본(snowball sample)을 통한 수집방법도 가능하다. 응답자들은 연구자가 연구하고 있는 현상과 관련있는 인물이어야 하며, 인터뷰 진행자의 질문에 정확하고 적절하게 대답하기 위한 그들의 능력과 응답하고자 하는 의지를 가지고 있어야 한다.

자료수집을 시작해라. 미래의 참고자료로 활용하기 위해 인터뷰 자료를 전자화시켜 기록이 가능하다. 하지만 연구자는 이러한 기록을 수행하기에 앞서 반드시 응답자들의 동의가 있어야 함을 명심해야 한다. 인터뷰가 진행되는 동안, 인터뷰 진행자는 응답자로부터의 중요한 내용이나 결정적인 관찰, 행동적인 반응(예: 응답자의 보디 랭귀지), 응답자들에 대한 인상에 대해 기록한다. 인터뷰가 종료된 후에는 분석을 위해 글자화된 문서로 기록되어야 한다.

사례 간 분석을 시행하라. 자료분석의 단계는 자료수집 이후에 이루어지거나 동시에 일어난다. 자료수집과 분석이 동시에 이루어지는 것은 자료분석으로부터 통합된 주제를 바탕으로 자료수집의 방향을 수정할 수 있는 장점을 가지고 있으며 주제에 대해 더 깊이 있게 증명할 수도 있다. 자료 분석은 두 단계를 통해 이루어지는데 첫 번째는 사례 간 분석이다. 연구자는 다양한 사례현상으로부터 각각 도출된 개념들을 통합하고 연구질문에 대한 이론과 도출된 개념들에 대한 패턴을 조사한다. 연구자는 주관적인 관점에서 사례현상에 대한 경험이나 개인적 관찰을 활용하여 연구질문에 대한 인터뷰 자료를 수집한다. 글라세와 스트라우스(Glasser and Strauss 1967)의 근거이론(grounded theory)은 코딩 전략을 사용하는데 개방형 코딩, 축 코딩, 선택적 코딩과 같은 기법을 활용하여 추론에 대한 관찰된 데이터들과의 연결을 이끌어준다. 홈그라운기술(homegrown technique)은 데이터의 네트워크 다이어그램과 같은 도해적 표현이나 종단자료(longitudinal data)를 위한 서열분석(sequence analysis)를 사용한다.

교차사례분석을 시행하라. 다중-현장 사례연구는 자료분석의 두 번째 단계로서 교차사례분석이 필요하다. 이 분석에서, 연구자는 서로 다른 사례현장에서 수집한 자료를 바탕으로 비슷한 개념과 패턴을 찾고, 결론에 영향을 미치지 않는 맥락의 차이는 무시한다. 도출된 패턴은 근거이론을 검증하는데 활용되거나 이론을 일반화하기 위해 개념을 추가하거나, 관계를 설정함으로써 이론을 발전시키기도 한다. 예를 들어, 연구자는 범주(예: 회사 규모, 산업 등)를 선택하고, 집단 간의 비슷한 점이나 차이점(예: 성과가 높은 사람들의 집단 vs 성과가 낮은 사람들의 집단, 혁신가 집단 vs 혁신에 도태된 집단)을 찾을 수도 있다. 또한, 유사점과 차이점을 공통으로 가지고 있는 형태의 짝들을 목록화하여 비교할 수 있다.

가설을 구축하고 검증하라. 일반화가 가능한 사례현장에서 도출한 개념과 주제를 기초하여, 잠정적인 가설을 구축한다. 이러한 가설들은 관찰한 데이터나 수집한 데이터들을 비교하여 그것이 올바르게 정립되었다는 증거를 통해 구성개념과 그 관계를 반복적으로 살펴보아야 한다. 만약 그렇지 않다면, 구성개념과 그들 간의 관계는 재정립되어야 한다. 또한 연구자는 문헌연구를 통해 구성개념과 가설을 비교함으로써, 내적 타당성과 일반화 가능성을 확보해야 한다. 연구에서 반대되는 결과물은 이론의 통찰력을 이끌어 내기 위해서 창의적인 사고를 활용하여 조화시켜야 한다. 이론과 데이터 간의 지속적인 반복 작업은 기존의 이론을 변화하거나 새로운 통찰력을 이끌어내지 못하고, 이미 ‘이론적인 포화’ 상태에 이르러 이론구축단계가 끝났다고 본다.

사례연구 보고서를 써라. 연구자가 보고서를 쓰는 과정은 표본수집, 자료수집, 가설개발, 자료분석을 위해 명백하고 상세하게 그 과정을 설명해야 한다. 이것은 독자들이 보고된 추론에 대한 합리성, 강점, 일관성 등을 평가할 수 있도록 한다. 명확한 연구방법의 과정은 연구의 결과가 연구자의 의식을 통한 편향되지 않음을 증명 시켜준다.

11.3 해석적 사례연구 표본

해석적인 사례연구를 학습하는 가장 좋은 방법은 실례를 조사하는 것이다. 이센하들트(Eisenhardt 1989)¹¹의 연구는 초고속 환경(HVE)에서 경영자들이 어떻게 의사결정을 하는지 연구하였다. 이 장에 수록된 시놉시스를 읽기 전에 “*Academy of Management Journal*”에 수록된 원본 문서를 읽어 보는 것을 공부하는 독자들에게 추천한다. 이센하들트는 몇몇 HVE 회사의 경영진이 다른 회사보다 어떻게 빠른 의사결정을 하는지, 빠른 의사결정이 회사성과를 향상시키는지 혹은 악화시키는지에 대하여 알아보려 하였다. 여기서 HVE는 수요, 경쟁, 기술적인 변화가 너무 빠르고, 정보의 활용이 부정확하거나 불가능하거나 혹은 쓸모없는 형태의 비연속적인 것으로 정의하였다. 이 연구에서 내재되어 있는 가정은 (1) HVE에서 부적절한 정보를 가지고 빠른 의사결정을 하는 것은 어렵고 (2) 빠른 의사결정은 효율적이지 못해 이는 회사성과를 저해할 가능성이 있다는 것이다.

이센하들트는 경영진들의 의사결정에 대한 선행 연구들을 재검토하면서 몇 가지 패턴을 발견하였다. 특별하게 어떠한 패턴들도 초고속 환경에 특화된 것은 아니었다. 회사는 편리한 방법으로 빠른 의사결정을 하고 있었으며, 이러한 결정에는 매우 적은 자원, 최소한의 대안들, 제한된 분석방법,

¹¹ Eisenhardt, K. M. (1989). “Making Fast Strategic Decisions in High-Velocity Environments,” *Academy of Management Journal* (32:3), 543-576.

의사결정에 관여하는 제한된 참여자, 의사결정의 중앙화, 내부적 갈등이 제한된 상황에서 이루어지고 있었다. 하지만 이센하들트는 이와 같은 관점이 그가 가지고 있는 연구질문을 설명하기 위해 필요하지는 않다고 생각하였다. 급격한 환경과 불완전한 정보의 활용을 바탕으로 한 의사결정일지라도 높은 수준의 의사결정은 반드시 필요하다고 언급하였다.

이러한 현상을 조사하기 위해서 이센하들트는 개인용 컴퓨터 산업에 속해있는 여덟 개의 기업을 대상으로 귀납적 연구를 수행하였다. 1980 년대의 개인용 컴퓨터 산업은 UNIX 운영체제, RISC 아키텍처, 64KB 메모리의 개발과 함께 매우 빠른 속도로 발전하고 있었다. 뿐만 아니라 IBM 이 개인용 컴퓨터 시장에 진입함으로써 경쟁은 가중되었으며, 두 자리 수의 수요 증가는 초고속 환경을 만들어 냈다. 이것은 반향 논리를 가진 다중사례설계이며, 이것은 각각의 사례에서 얻은 추론과 다른 사례로부터 얻은 추론을 서로 비교하여 확인이 가능하거나 그렇지 않게 나타난다. 사례현장은 연구자의 접근 가능성과 거리 시간상의 근접성으로 선택된다. 하지만, 캘리포니아 지역의 실리콘 벨리에 위치해 있는 모든 기업들은 초고속 개인 컴퓨터 산업에 속해 있으며 이러한 동일 산업과 동일 지역에 있는 기업들의 결합 현상은 산업과 지역적 위치에 따른 종속변수의 편차를 배제시켰다.

이 연구는 의사결정(기업 별 다중 전략 의사결정 비교), 경영진(전략적인 의사결정에 책임을 가진 서로 다른 팀들 간의 비교), 기업(전반적인 기업의 성과) 분석의 다중-단계를 포함한다. 위와 관련해서 다음과 같이 수집되었다.

- 최고 경영자들로부터 최초 인터뷰: 최고 경영자에게 자사의 경쟁력 있는 전략, 차별화되는 역량, 주요 경쟁사, 성과, 현재 가장 중요한 전략적 의사결정에 대해 질문했다. 이 인터뷰를 기초로 몇몇 기업들로부터의 전략적인 의사결정은 회사의 투자를 위해 선택되었다. 이를 위하여 (1) 의사결정이 회사의 전략적 포지셔닝을 포함하는가, (2) 의사결정이 높은 이해관계로 이루어져 있는가, (3) 의사결정이 복합적인 수단을 포함하고 있는가, (4) 의사결정이 기업의 전략적 의사결정 과정을 대표해 주는 것인가의 네 가지 기준이 활용되었다.
- 부서장들과의 인터뷰: 각각의 부서장들에게 자사의 경쟁력 있는 전략, 기능적인 전략, 최고 경영자들로 이루어진 팀의 구성원, 팀과의 자연적인 상호작용과 그 빈도, 전형적인 의사결정 과정, 과거의 의사결정이 어떻게 이루어졌는지, 그러한 의사결정을 하는데 시간이 얼마나 걸리는지에 대한 16 개의 개방형 질문을 했다. 인터뷰는 1.5 ~ 2 시간 가량 진행되었고 길게는 4 시간동안 이루어졌다. 응답자의 질문에 대답한 내용은 그들이 가지고 있는 인식이나 해석보다는 ‘언제 사건이 일어났는지’, ‘사건 당시 응답자는 무엇을 하였는지’ 와 같은 사실에 초점을 두었다. 인터뷰는 두 사람에게 의해 진행되었으며 데이터는 인터뷰 진행자와 보조자가 받은 응답자에 대한 인상과 사실에 대해 교차검증을 수행하였다. 모든 인터뷰 자료는 전자화 되어 기록되었고, 인터뷰 과정 중 중요한 내용에 대해 노트에 기록하는 작업도 시행되었다. ‘24 시간 룰’ 을 활용하여, 인터뷰에 대한 상세한 현장 노트는 24 시간 이내에 작성이 완료되었는데 이것은 응답자가 대답한 내용 가운데 놓친 부분이나 그들의 인상에 대한 느낌을 가능한 소실되지 않게 하기 위해서였다.
- 질문지: 각 기업의 경영진들의 양적인 자료를 수집하기 위하여 설문 조사지를 작성하였다.

- 이차 자료: 산업 보고서와 전략적 의사결정에 책임이 있는 경영진들의 인구통계학적 자료를 바탕으로 하는 기업의 내부 자료와 재무적 성과 등을 활용하였다.
- 개인 관찰: 연구자는 두 기업이 하루 동안 진행되는 전략 세션 및 주간 경영진 회의에 참석하였다.

자료분석은 양적 기술과 질적 기술의 결합을 포함한다. 양적 자료는 기업 혹은 의사결정에 따른 패턴을 분석한 것이다. 질적 인터뷰 자료는 한 명 이상의 경영진에 의해 언급된 프로파일의 특성을 활용하여 의사결정의 분위기 프로파일을 결합한 것이다. 사례 간 분석을 수행하기 위해서 의사결정 과정을 타임라인에 따라 경영진들에 의해 결정된 중요한 전략적 의사결정을 구성함으로써 수행된다. 크로스 사례분석은 의사결정의 속도나 기업의 성과와 같은 변수를 범주화하여 이들 간의 차이점과 유사성을 쌍으로 구성하여 비교한다. 이러한 분석들을 바탕으로 잠정적인 구성개념들과 명제들은 기업범주 안에서 각각의 의사결정 과정을 통해 귀납적으로 유도된다. 각 의사결정 사례는 연구에서 제안할 관계를 성립하기 위한 검증자료로 활용하기 위해 다시 한번 검토한다. 추론된 명제는 과거에 연구된 문헌들의 결과와 비교하고 그것들이 다른 문헌과 차이점은 무엇인지, 그리고 연구자의 사례연구를 통한 결과가 어떤 새로운 통찰력은 끌어내는지 살펴보아야 한다. 마지막으로 문헌을 통한 유효한 명제는 빠른 HVE 에서 기업의 전략적 의사결정의 귀납적 이론과 통합된다.

다중사례연구로부터 도출된 추론은 기존 문헌연구의 몇 가지 의사결정 패턴과 모순된다. 첫째, 빠른 속도 환경(HVE)에서 의사결정자는 기존 문헌에서 적은 정보를 활용할 것이라는 결과와 다르게 더 많은 정보를 이용했다. 게다가 의사결정자는 실시간 정보를 활용하는 것으로 나타났는데 이것은 기업이 가지고 있는 문제점, 시장의 기회, 그리고 기업 환경의 급격한 변화에 정확하고 빠르게 대응할 수 있게 한다. 둘째, 의사결정이 빠른 경영자는 더 많은 대안들을 고려하였다. 그들은 자극적이고 복합적인 대안들을 가지고 있었던 반면에 의사결정이 느린 경영자는 단계적이며 상대적으로 적은 대안들을 가지고 있었다. 셋째, 의사결정이 빠른 경영자는 기존 문헌연구에서 도출한 결과와 같이 중앙에 집중된 의사결정이나 조언을 통한 의사결정을 하지 않았다. 오히려 이런 기업들은 두 단계의 의사결정 과정을 가지고 있었는데 첫 번째 단계에서는 경험을 가지고 있는 카운슬러에게 조언을 받고 두 번째 단계에서 대안들의 비교를 빠르게 하여 의사결정을 하는 것이다. 넷째, 의사결정이 빠른 경영자는 과거 문헌연구를 통해 기대했던 것과 마찬가지로 충돌이 적게 일어났지만 의사결정의 속도를 개선하기 위하여 더 나은 충돌해결 기술을 가지고 있었다. 마지막으로, 의사결정이 빠른 경영자는 그들이 가지고 있는 직관력, 감정, 중요한 의사결정의 과정을 빠르게 활용하기 때문에 월등한 기업성과를 보여주었다.

11.4 실증적 사례연구 표본

사례연구는 가설이나 이론을 검증하기 위한 실증적인 방식으로 사용된다. 이러한 연구들이 많이 존재하지 않지만 마커스(Markus, 1983)¹² 의 Golden Triangle Company(GTC) 사례는 기술 실행에 대해 실증적으로 연구하였다. 이 연구의 목적은 기업의 생산성을 향상시키기 위해 새로운

¹² Markus, M. L. (1983). "Power, Politics, and MIS Implementation," *Communications of the ACM* (26:6), 430-444.

회계정보시스템(Financial Information System: FIS)을 도입해야 하는 이유에 대한 것이었다. GTC의 회계사들의 성과는 GTC 본사의 회계사들로부터 도움을 받지만 GTC 지점의 부서 회계사들은 이러한 도움에 저항을 가지고 있었던 것이다. 이러한 흥미로운 현상은 GTC 기업이 가지고 있는 특징을 나타내기 때문에 마커스의 연구는 단일-사례연구로 수행되었다.

회계정보시스템의 사용자들의 저항을 가져오는 이유를 이해하기 위해서 마커스는 세 가지의 이유를 상정했다. (1) 시스템-결정이론: 저항은 시스템의 기술적인 결함, 나쁜 인체 공학적인 디자인, 불편한 사용감과 같은 요인들로부터 유발된다. (2) 사용자-결정이론: 저항은 시스템을 사용하는데 있어 회계사의 인지 스타일이나 개인적 특성과 같은 사용자의 내적 요소에 의해 유발된다. (3) 상호작용이론: 저항은 사용자나 시스템의 본질적인 요소로 유발되는 것이 아니라 두 요소 간의 상호작용에 의해 유발된다. 구체적으로 말해서 상호작용이론은 회계정보시스템을 실행함으로써 조직 내 파워를 재분배하는 상황을 불러 일으키는 것이다. 시스템의 도입으로 조직적 지위, 관계, 권력을 잃어버리는 회계사들은 시스템을 거부하게 되며 그 반대의 회계사들은 찬성할 것이다.

이 세가지 이론을 검증하기 위해서 마커스는 각각의 이론적 해석으로부터 기대되는 결과를 예측하고 GTC에서 관찰한 결과와 비교하여 결과에 대한 예측을 분석했다. 예를 들어 시스템-결정이론은 사용자의 저항이 부적합한 시스템 때문에 유발되기 때문에 시스템의 기술적인 문제를 해결함으로써 그들의 저항을 감소시킬 수 있다고 보았다. 회계정보시스템을 사용하는 컴퓨터는 강력한 운영 시스템, 회계 정보의 즉시처리를 지원하는 온라인 프로세스, 관리자로부터 개발된 단순하고 새로운 회계 소프트웨어가 더해져 지속적으로 향상되었다. 이런 변화들이 만들어지고 일년 후에도 사용자들은 여전히 시스템을 거부되었으며 시스템은 반드시 교체되어야만 한다고 느꼈다. 따라서 시스템-결정이론은 기각되었다.

사용자-결정이론은 회계사들의 개인적인 저항 요소를 교체하거나 또는 시스템 저항이 낮은 사용자들을 선임하는 것이 회계정보시스템에 대한 저항을 감소시킬 수 있을 것이라고 보았다. 그 후에 GTC는 업무순환과 이동정책을 활용하여 회계사들을 저항이 있는 부서들에 안팎으로 이동시켰다. 하지만 저항은 여전히 존재했을 뿐만 아니라 증폭시켰다. 특별한 예로, 회계부서에서 일하면서 시스템의 설계자였던 회계사는 기업의 부서 관리자의 사무실로 옮긴 이후, 시스템에 대한 저항이 시작했다. 사용자-결정이론의 실패로 이 이론 역시 기각되었다.

마지막으로 상호작용이론은 사전 실행의 단계에서 설명되지 못한 권력의 불균형과 재분배가 설명되지 않으면 저항을 사용자나 시스템의 요소가 감소시킬 수 없다고 보았다. 회계정보시스템의 도입 전, GTC의 회계부서는 회계운영과 관련된 모든 회계자료를 가지고 있다고 생각했다. 그렇기 때문에 그들의 자료들은 두꺼운 매뉴얼 원장에 따라 보관되었고, 자료에 대한 접근을 관리 및 공식적인 보고서로 기록되기 이전에 비공식적인 회계 사건들을 조정했다. 기업의 회계사들은 부서별 회계사들에 크게 의존하고 있었는데 그들이 기업보고나 통합적인 자료정리를 위해 부서별 자료들에 접근할 수 있었기 때문이다. 회계정보시스템은 기업의 모든 자료를 수집하고 통합하여 하나의 기업 데이터베이스에 자동으로 관리됨으로써 부서별 회계사들이 필요 없게 되었으며 그들의 업무에 포함되었다는 인식이 줄어들었다. 기업의 회계사들은 더 이상 부서별 회계사를 통하지 않고도 데이터베이스에 내장된 자료를 직접 요청할 수 있게 되었다. 게다가 부서들 간의 성과를 비교하고 분석 가능해졌으며 경영진들의 이상한 패턴이나 활동들 찾아내 결국 부서의 권력을

무력화시켰다. 비록 마커스가 실증적으로 이론을 검증하지 않았을지라도 시스템-결정이론과 사용자-결정이론과 결부하여 조직의 권력에 대한 재분배를 관찰함으로써 상호작용 이론의 정당성을 이끌어냈다.

11.5 전통적 연구와의 비교

가설검증을 하는 실증적 사례연구는 자연과학 연구자들로부터 비판을 받기도 한다. 그 이유는 통제된 관찰 값이라고 보기 부족하고 통제된 추론이라고 보기 어렵거나 반복 가능성이 없으며 일반화하기 어렵기 때문이다. 하지만 이러한 비판들은 적합한 사례연구 설계를 통해 극복될 수 있다. 예를 들어 사례연구에서 통제된 관찰 값은 실험적인 혹은 통계적인 통제로부터 획득하기는 어렵다. 반면 사례연구자들은 ‘자연적인 통제(natural control)’를 사용함으로써 통제범위의 결핍을 보상할 수 있다고 보았다. 마커스(Markus, 1993)의 연구에서 의미하는 자연적인 통제는 회계정보시스템의 도입 초기에 지지하던 회계사가 관리 부서로 옮긴 후에 저항을 보이기 시작한 그의 행동 변화이다. 이것은 그의 새로운 부서에서의 역할이다. 따라서 사례연구는 전통적 연구 방법과 비교하여 이러한 자연적인 통제들은 사전에 예측할 수 없다. 사례연구를 수행하는 연구자는 통제 가능한 변수들을 찾는 노력을 함으로써 이러한 통제들에 대해 간과하지 않아야 한다. 놀랍게도 자연적인 통제는 천문학, 지리학, 인체생물학과 같은 자연과학 분야에서도 사용된다.

통제된 추론의 문제점은 사례현장에서 주어진 대부분의 질적 자료의 추론을 뒷받침 할 수 있는 적절한 양적 자료가 부족한 것과 관련이 있다. 가설검증을 위한 양적 자료의 부족이 사례연구에서 통제된 추론의 행동을 예측함으로써 획득 가능하다. 이러한 예측은 이론적 배경을 바탕으로 지속적으로 예측을 검증해왔던 것이다. 마커스는 자신의 연구에서 이를 적용하여 저항에 대한 세 가지의 이론적 가설을 각각 설정하였다. 그리고 실제로 관측된 행동과 설정한 가설이 일치하지 않아 두 개의 가설을 기각하였다. 이 경우에서 가설은 통계학적 검증보다 논리적 명제를 사용하여 분석하였다. 그 이유는 통계검증은 논리의 한 부분이기 때문이다.

셋째, 반복 가능성의 문제점은 주어진 사례현장의 특이성과 독특성을 가진 현상을 반복하여 관측하기 어렵다는 것이다. 하지만 다른 연구자가 마커스의 세 가지 이론을 사용하여 앞선 현장의 특징을 가지고 있는 새로운 사례현장에서 동일한 이론을 검증한다면 앞선 세 가지의 결과에 대한 예측은 검증될 것이다. 다시 말하자면, 사례연구의 현상이 반복되지 않을지라도 사례연구의 추론은 반복될 가능성이 있다.

넷째, 사례연구는 일반화되지 않을 수도 있는 독특하고 반복되지 않는 현상들을 살펴보는 경향이 있다. 특히, 자연과학에서의 일반화는 추가적인 연구에 의해 성립된다. 서로 다른 현장에서 도출된 예측들이 다른 연구들과 통합되고 일관성 있는 결과들로 검증된다면 사례연구는 일반화할 수 있다.

마지막으로 영국 철학자 칼 포퍼(Karl Popper)는 과학적 이론의 네 가지 필요 조건에 대해 설명했다. (1) 이론은 반증 가능해야 하며, (2) 논리적으로 일관성이 있어야 하고, (3) 충분히 예측할 수 있어야 하며, (4) 경쟁 이론보다 더 나은 해석을 제공해야 한다. 사례연구에서 앞의 세 개의 필요조건은 사례연구의 현장, 대체 가능한 예측의 수, 분석의 단계를 늘려 관측된 결과 값의 자유도를 높임으로써 만족시킬 수 있다, 이것은 마커스의 연구에서 부서별 회계사와 기업 회계사 간의 다중

집단의 행동을 조사하고 세 가지 경쟁 이론에 따른 충분한 설명을 제공함으로써 검증되었다. 포퍼의 네 번째 조건은 하나의 가설이 두 개의 경쟁 가설보다 더 나은 관측된 관찰을 함으로써 검증하였다.

제 12 장 해석적 연구

앞장에서 해석적 연구(interpretive research)의 소개와 여러 사례들을 살펴보았다면, 이번 장에서는 해석적 연구의 다른 면을 살펴보고자 한다. 앞서 소개하였듯이, 실험실 연구와 설문조사와 같이 실증적(positivist)이고 연역적(deductive) 방법론은 가설과 이론을 검증하기 위한 방법론이고, 행동연구나 민족지학연구와 같이 해석적(interpretive)이고 귀납적(inductive) 방법론은 이론을 성립하기 위한 방법론이다. 실증적 방법론과 달리, 해석적 연구는 특정 현상에 대한 관찰데이터의 해석과 그를 바탕으로 한 이론의 구성에서 출발한다.

‘해석적 연구(interpretive research)’는 종종 ‘질적 연구(qualitative research)’와 혼용되기도 하는데, 실제 두 연구의 기본개념에는 차이가 있다. 해석적 연구의 기본 연구 패러다임(3장에서 설명)은 사회적 현상이 객관적이기 보다 인간의 경험과 사회적 맥락에서 이해되어야 한다는 것을 전제로 하고 있으며(존재론), 주로 특정인의 주관적인 해석이 요구되는 연구(인식론)에 적합하다. 해석적 연구는 사회현상 안에 내재되어 있는 추상적 개념을 찾아내기 어렵다는 가정에서 이것을 이해(sense-making)하고 해석(interpret)을 목적으로 하는 것이다. 반면, 실증적 연구는 현실세계를 관찰 가능한 형태로 일련의 맥락에서 추상적인 개념으로 정의하고 객관적이고 체계적인 수단을 통해 실증적인 검증을 할 수 있다고 가정하고 있어, 해석적 연구와 대조적인 면을 보인다. 따라서, 연구자는 연구하고자 하는 현상의 특성에 대해 충분히 고려하여, 두 방법론 중 적합한 것을 선택할 필요가 있다.

한편, 질적 연구(qualitative research)와 양적 연구(quantitative research)은 자료의 형태, 자료수집과 분석방법에 따라 실증적 연구이거나 자료 기반의 연구를 수행한다. 질적 연구는 통상 인터뷰나 관찰과 같이 수치로 표현되지 않는 자료를 바탕으로 수행하고, 양적 연구는 점수, 통계치 등 수적인 데이터를 바탕으로 수행하는 연구이다. 따라서, 질적 연구에서는 회귀분석(regression analysis)과 같이 통계적인 해석이 아닌 내용분석(content analysis)과 같은 간접적인 코딩을 통한 해석이 필요하다. 코드화 된 질적 자료는 코드가 나타난 개수에 따라 양적 연구처럼 보일 수는 있지만 통계적인 분석을 포함하지 않는 경우가 대부분이다. 정통 해석적 연구자들은 코딩 접근법(coding approach)이 객관적이고 통계적인 시각을 포함하였다하여 질적 연구에 포함시키지 말아야 한다는 주장을 하기도 한다.

해석적 연구가 주로 질적 자료를 기반으로 하지만, 종종 보다 정확하고 객관적인 이해를 돕기 위해서 양적 자료를 활용하곤 한다. 예를 들어, 이센하르트(Eisenhardt 1989)의 사례연구에서 대기업이 보다 전략적이고 효율적인 의사결정을 거치는 과정을 살펴보기 위해 전략적 의사결정이 소요되는 기간(대략 1 개월 반에서 18 개월), 의사결정에서 고려되는 의제의 수 등 수치적인 데이터를 수집하는

동시에 응답자들이 조직내부에서 생기는 의사충돌에 대해 어떻게 반응하는지에 대해 설문 조사를 하기도 했다. 응답자들의 주관적인 편견 없이, 수적 데이터를 이용하여 빠른 의사결정을 하는 집단과 의사결정이 상대적으로 느린 집단을 객관적으로 나눌 수 있게 해주었고, 연구자는 각각 두 집단에서 생기는 의사 충돌에 관해서도 살펴 볼 수 있었다. 이렇듯, 해석적 연구뿐 아니라 실증적 연구에서도 보다 정확한 연구결과를 얻기 위해 질적 자료와 양적 자료 모두를 살펴 보아야 한다. 질적·양적 자료를 함께 사용하는 것을 ‘혼합설계(mixed-mode design)’라고 하며, 이를 이용하여 보다 새로운 시각에서 과학적이고 통찰력을 가진 연구로 접근 할 수 있다.

해석적 연구는 인류학, 사회학, 심리학, 언어학, 기호학에서 출발하여 실증적 연구보다 훨씬 이전인 19세기 초반부터 널리 사용되어왔다. 실증적 연구자들이 종종 해석적 연구가 질적 자료를 수집하고 해석하는 과정에서 주관적이고 오류가 발생할 수 있음을 지적한다. 그러나, 실증적 연구 또한 이론을 일반화시키고 적용하는데 있어 많은 오류를 보여, 해석적 연구의 신뢰성과 타당성에 대한 문제제기로 인하여 1970년대 이후부터 해석적 연구가 재조명하기 시작했다.

12.1 해석적 연구와 실증적 연구 비교

앞서 방법론의 기본 패러다임인 존재론(ontological), 인식론(epistemological)을 바탕으로 한 해석적 연구와 실증적 연구를 살펴보았듯이, 두 연구에는 여러 차이점이 존재한다. 첫째, 해석적 연구는 이론적 표본추출(theoretical sampling)을 바탕으로 하고 있으며 연구 배경, 응답자, 그리고 연구하고자 하는 현상이 목적에 맞게 제대로 조사될 수 있는지, 타 연구와 구분 지을 수 있는 특성이 무엇인지 등 이론적인 차원에서 고려된 표집을 의미한다. 이와 반대로, 실증적 연구에서 표집은 모집단을 무작위로 추출하는 등 무작위표본추출(random sampling)을 바탕으로 일반화 할 수 있는 수준이어야 한다. 따라서, 해석적 연구에서는 표본의 크기와 표집방법에 관계없이 표집이 해당연구의 목적과 특성을 잘 반영하여야 하지만 실증적 연구는 그렇지 않다.

두 번째로, 해석적 연구에서 연구자는 아주 중요한 역할을 한다. 민족지학(ethnography), 행동연구(action research), 참여관찰(participant observation)과 같은 연구에서 연구자는 사회 현상의 일부로 간주되며, 연구자의 역할과 참여과정은 자료분석 동안에 명확히 구분 지어져야 한다. 사례연구와 같은 방법론에서의 연구자는 ‘중립(neutral)’의 입장을 지키고 자료수집 과정과 분석 과정을 객관적인 관점에서 수행해야 한다. 해석적 연구에서 연구자의 시각과 편견은 해석의 주관성이라는 특성에 큰 영향을 끼치는 요인은 아니다. 그러나 실증적 연구에서 연구자는 독립적이고 외적인 변수로 작용하기 때문에, 자료수집 및 분석 과정에 어떠한 편견도 허용하지 않는다.

셋째, 해석적 분석은 현실로부터 분리되고 통제된 상황이라기 보다 현실 그 자체의 맥락으로 이해한다. 실증적 연구에서 통계적 기법이 주로 쓰이는 반면, 해석적 분석은 사회 현상에 포함되어 있는 언어, 기호, 사회의 구성원의 관점에 대해 중점을 둔다. 또한, 구성개념의 타당성이나 통계적인 검증력보다 자료수집과 분석에 대한 접근에 초점을 맞추어 연구를 하는 것이다.

마지막으로, 해석적 연구에서 자료수집과 분석은 유기적이고 반복적으로 진행되어야 한다. 예를 들어, 연구자는 다음 인터뷰의 원활한 진행을 위해 인터뷰를 실시하고 그것에 대해 코딩을 한 후, 다음 인터뷰를 진행해야 한다. 반복적인 분석은 연구자에게 연관성있는 인터뷰 질문지를 완성하는데

있어서 결점을 보완하게 하고, 현상에 대한 이해를 높이는데 도움이 된다. 또한, 기존의 연구질문이 현상을 이해하는데 새롭거나 유용하지 못한 통찰력을 가지고 있다면 연구질문을 새로운 시각으로 재조명 할 수 있는 것이다. 실증적 연구에서는 이렇듯 연구의 재구성과 자료수집의 반복이 실질적으로 어려우나, 해석적 연구에서는 하나의 연구과정으로 인정된다는 것이 큰 장점이라고 할 수 있다.

12.2 해석적 연구의 장단점

해석적 연구만이 가지는 여러 가지 장점이 있다. 첫째, 사회적 현상에 숨겨진 복합적이고 다각화된 프로세스들을 이해하는데 적합하다는 것이다. 특히, 회사 내부의 관계, 사무실 내부의 권력구조와 같은 질적 자료들은 양적 자료로 쉽게 관찰 되지 않는다. 둘째, 현상을 설명하는 배경이론이 될 만한 이론이 현재 존재하지 않거나 충분치 않을 경우, 새로운 이론을 성립하는데도 많은 도움이 된다. 셋째, 특정한 시각과 맥락의 차원에서 해석이 필요한 현상이나 사건을 연구하는데 적합하다. 마지막으로, 해석적 연구는 연구주제와 관련된 흥미로운 점들에 대해 발견할 수 있으며 차후 연구할 수 있는 부분에 대해 제시할 수 있다.

해석적 연구에는 장점만 있는 것은 아니다. 첫째, 해석적 연구는 자료수집과 분석 능력만이 요구되는 실증적 연구와 달리 많은 시간과 노력, 집중이 요구된다. 분석하고자 하는 주제에 적합한 데이터가 충분하지 않은 경우, 설정한 가정을 제대로 검증하지 못할 수 있고, 반대로 너무 많게 되면 연구자가 데이터를 활용할 수 있는 효용치를 벗어나게 된다. 둘째, 해석적 연구에서 개인적인 편견 없이 사회 현상과 그 배경이 되는 개념을 구성해 낼 수 있는 정확한 관찰력과 분석능력을 갖춘 훈련된 연구자가 수행해야 한다. 셋째, 연구대상이나 자료는 현상에 대해 신뢰성을 가지지 못하거나 충분한 지식을 반영하지 못한 채 편향될 수 있다. 또한, 연구대상과 조사자 간의 신뢰관계가 형성되지 못한 경우 연구대상의 답변이 왜곡될 가능성이 있는데, 이를 피하기 위해 관계를 형성하는데 시간이 오래 걸리게 된다. 해석적 관점에서 연구하는 연구자는 잘 보이지 않은 변인들 까지도 관찰할 수 있어야 하고 그 문제에 대해서 충분히 이해해야 한다. 넷째, 자료에 따른 해석은 현상을 바탕으로 한 조사자의 관점에 따라 크게 좌우되기 때문에 결과를 일반화하여 다른 상황에도 적용하기가 어렵다. 마지막으로, 해석적 연구는 종종 실제 제기된 연구주제를 제대로 해결하지 못하거나 미래행동을 예측하기 어렵다는 단점을 가지고 있다.

12.3 해석적 연구의 특징

모든 해석적 연구는 반드시 다음과 같은 원칙을 따라야 한다.

연구주제는 자연적인 현상을 관찰해야 한다(*naturalistic inquiry*). 연구하고자 하는 현상이 일반적이고 일상적으로 일어나야 한다. 해석적 연구는 사회현상에 대해 특정 맥락을 구분 짓지 않고, 있는 그대로 사회적 배경을 바탕으로 한 해석을 특징으로 수행해야 한다. 해당 사회 맥락에서 민감한 부분을 일반화시키는데 제한을 두기는 하나, 현상을 설명하는 가운데 맥락적 변수는 포함되고 고민되어야 할 부분이다.

연구자는 수단로 인식한다(researcher as instrument). 연구자는 그들의 연구 안에서 사회적 맥락의 한 부분으로 작용하며 자료수집의 수단으로 인식되는 것이 대표적인 예이다. 그렇기 때문에 그들이 사용한 기술이나 연구 대상자들과의 신뢰를 바탕으로 한 관계, 정확한 정보를 분별해 낼 수 있는 능력이 요구된다. 더불어, 사회현상에 대한 개인적 통찰력, 사전지식 및 경험 등은 현상을 제대로 이해하는데 중요하다. 연구자는 연구 과정에서 발생할 수 있는 개인의 편견을 정확히 인지하여야 하고, 이러한 편견이 현상을 설명하는데 있어 절대 문제가 생기지 않도록 주의해야 한다.

해석적 분석이 요구된다(interpretive analysis). 관찰은 해당 사회에 속한 연구대상의 시각에서 해석되어야 한다. 이때 해석은 두 단계를 거쳐야 한다. 첫 번째 단계는 연구대상이 현상에 대한 관찰과 경험을 바탕으로 주관적인 관점에서 이루어져야 한다. 두 번째 단계는 연구대상으로부터 전달받은 경험의 의미를 파악하는 것으로, 연구대상이 왜 그러한 행동을 하게 되었는지에 대한 보다 깊은 해석을 하는 것이다.

표현적 언어를 사용한다(use of expressive language). 응답자의 언어나 비언어적 행동을 서류화하고 해석적 분석을 위해 구성요소로서 살펴볼 필요가 있다. 해석적 연구는 기계가 아닌 인간의 눈에 비쳐진 것을 표현한 것이기 때문에 응답자의 감정이나 표현력을 독자로 하여금 해당 대상에 대해 이해 할 수 있도록 도와야 한다. 또한 상상, 은유, 풍자 등 다양한 화법이 해석적 분석에서 나타날 수 있음을 기억해야 한다.

일시적인 자연현상에 대한 연구이다(temporal nature). 해석적 연구는 연구질문에 대한 특정한 답을 찾아가는 과정뿐 아니라 장기간에 걸친 사회 프로세스에 대해 감각적으로 이해하고자 하는 것이다. 따라서, 해석적 연구는 연구자를 일정기간 동안에 실제 현상이 발생하는 현장에 있게 하여, 관심 현상에 대한 전체적인 그림을 그리도록 하는 것이다.

해석의 순환과정이다(hermeneutic circle). 해석적 분석은 관찰된 부분과 전체적인 사회 맥락을 연결짓는 반복적인 프로세스이다. 참여자들의 다양한 주관적인 관점과 경험들이 유기적으로 연관지어 이론을 구성하는 것이다. 현상에 대한 이해, 의미, 그리고 관찰이 ‘이론적인 포화(theoretical saturation)’ 상태와 같이 충분해지도록 반복하고 여기서 반복적인 확인절차는 관심 현상에 대한 특정한 시점을 부가하지 않는다.

12.4 해석적 자료수집

해석적 연구에서 자료수집은 다양한 기법을 활용하여 얻는다. 가장 많이 쓰이는 방법은 인터뷰(interview)로, 면대면 조사, 전화, 또는 포커스그룹 조사 등이 여기에 해당한다. 인터뷰의 형태와 전략은 앞선 조사연구에서 자세히 설명되었다. 다음으로는 관찰(observation)이 있다. 관찰 기법에는 사례연구와 같이 연구자가 중립적·수동적 외부 관찰자가 되는 직접적 관찰(direct observation)과, 행동연구(action research)와 같이 현상에 능동적 참여자로서 연구대상에게 일정한 자극을 주는 참여관찰(participation observation)이 있다. 마지막으로 문서(documentation) 조사이다. 문서는 내부적·외부적 문서 모두를 포함하며, 간단한 메모, 이메일, 연간보고서, 재무제표, 신문기사, 웹사이트 등 현상에 대해 관찰 가능한 것이어야 한다.

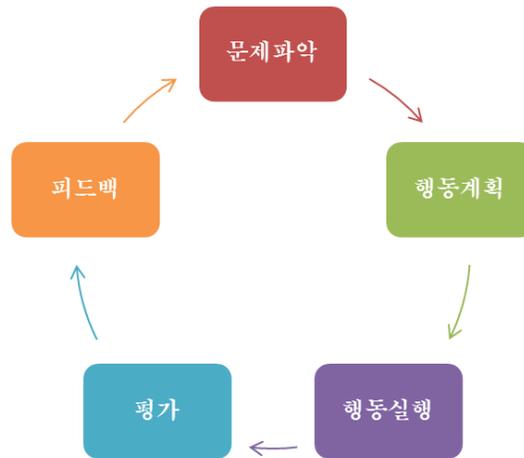
12.5 해석적 연구 설계

사례연구(case research). 앞 장에서 논의되었듯이 사례연구는 현상에 대해 보다 심화된 종단적인 연구방법으로, 표면으로 보이는 사실뿐만 아니라 감춰져 있는 문맥의 이해를 이끌어 내는 것이다. 사례연구는 이론을 설정하는 해석적 관점에서의 연구 설계뿐 아니라 이론을 검증하는 실증적 관점에서도 활용되는 독특한 연구방법 중 하나이다. 사례연구의 기법과 자세한 예시는 이미 앞장에서 소개 된 바 있다. 사례연구를 수행하는 연구자는 현상에 대하여 능동적 참여자라기 보다는 중립적 관찰자에 더 가까우며 직접적인 관찰을 하게 된다. 여타 다른 해석적 관점에서와 같이, 사례연구의 의미 있는 결과는 연구자의 관찰력과 해석력에 따라 크게 좌우 된다.

행동연구(action research). 행동연구는 질적 연구인 동시에 실증적 연구설계이다. 즉, 이론을 설정한다기 보다는 실증적으로 검증하는 차원의 연구에 더 가깝다. 행동연구는 복잡한 사회상이 가지고 있는 변화, 방해 요인과 같은 어떠한 ‘행동’ 이 현상에 어떤 결과로의 영향을 미치는지 살펴봄으로써 그 현상을 충분히 이해하기 위한 가장 좋은 방법이다. 이 방법론에서 연구자는 일반적으로 컨설턴트나 조직의 구성원의 한 사람으로써 그들의 행동이나 행동자체가 현상에 어떻게 영향을 미치는지 관찰하고 그 행동과 현상 사이의 관계에 대해 분석하는 것이다. 여기서 행동의 예로는 기업이 조직의 업무력과 생산성을 높이기 위해 적용시키는 새로운 조직체계, 구성원, 기술 등과 같은 조직적 변화 등이 있다. 연구자의 행동은 늘 충분한 이론적 바탕을 근거로 그 이유와 방법에 대한 기술이 이루어 져야 한다. 또한 연구에서 제기 된 현상 문제를 해결하기 위해 연구자가 선택한 행동이 이론과 적절에 부합하는지도 살펴보아야 한다. 끊임없는 문제해결과 통찰력의 발달은 행동연구를 다른 연구와 구분 짓는 가장 핵심적인 특징이다. 따라서, 행동연구는 연구와 실재를 연결하는 훌륭한 방법론 중 하나이다.

이 밖에도 다양한 행동연구 방법론이 있다. 가장 보편적으로 알려진 것이 참가활동연구(participatory action research)로 수스만과 에버드(Susman&Evered 1978)¹³의 연구에서 소개되었다. 참가활동연구는 다섯 단계를 따른다. (1) 문제 파악(diagnosing): 사회적 맥락 안에서 문제를 제기하고 정의한다. (2) 행동계획(action planning): 대체될 수 있는 해결책을 찾아 평가하여 다음 단계의 행동을 이론적 논리에 맞게 대안들 가운데 결정한다. (3) 행동실시(action taking): 결정된 행동을 직접 하여 실행에 옮기는 것 말한다. (4) 평가(evaluating): 결정된 행동의 영향력이 실제 어떠한지를 평가한다 (실제에서 이론적 영향이 정말로 영향을 미쳤는지 알아보기 위하여 실시함). (5) 피드백(learning): 경험과 행동의 영향력에 대한 피드백을 활용하여 제기되었던 문제에 대한 시사점과 앞으로 발전시켜야 할 부분을 기술한다(그림 12.1 참조). 행동을 평가하고 피드백하는 과정은 문제를 수정하여 보다 적합한 해결방법을 이끌어 낼 수 있고 행동연구의 순환과정은 이러한 수정된 행동결과들이 반복되면서 형성된다. 전체적인 행동연구의 순환과정은 적어도 두 번 이상 실시하도록 제안하고 있는데 첫 번째 순환과정으로부터의 피드백이 두 번째 순환과정에서 실시되기 때문이다. 인터뷰와 문서조사와 같은 방법론은 연구자의 관찰을 통해 데이터가 수집되었다면, 행동연구의 일차적 자료는 참가자의 관찰로 이루어 지기 때문이다.

¹³ Susman, G.I. and Evered, R.D. (1978). "An Assessment of the Scientific Merits of Action Research," Administrative Science Quarterly, (23), 582-603.



〈그림 12.1〉 행동연구 순환과정

민족지학(ethnography). 민족지학은 크게는 인류학에 근거를 두고 있으며, 현상을 해당 문화적 맥락에서 이해해야 한다는 것을 강조하고 있다. 연구자는 해당 사회문화에 충분한 기간 동안 속해(통상 8 개월에서 2 년까지), 그들과 동일한 사회 구성원으로서 참여하여 관찰하고, 기록한 일기를 바탕으로 연구를 진행하게 된다. 일차 자료수집은 참여자의 관찰이 되고 자료분석은 수집된 자료를 이해(sense-making)하는데 초점을 맞추어 진행한다. 또한, 연구자는 현장에서 기록한 충분한 자료를 제공해야 한다. 또한, 현장 자료에는 연구자가 경험한 것들에 대해 상세히 기술하여, 독자가 간접 경험 할 수 있도록 한다. 이 방법론은 연구자에게 두 가지 요소를 요구한다. 첫째, 연구자의 사전지식과 경험에서 특정 관점 혹은 이론을 성립할 수 있는 능력이고 둘째, 학계에서 충분히 받아들일 수 있도록 연구한 현상에 대한 이해를 돕게 하는 설득력이다.

민족지학의 가장 대표적인 사례로는 영국의 동물 행동학 연구자인 제인 구달(Jane Goodall)의 침팬지 연구이다. 구달(Goodall)은 탄자니아의 곰비국립공원 내 침팬지 서식지에서 기거하고, 침팬지와 직접 소통하고 삶을 공유하며 영장류 행동에 대한 연구를 했다. 연구기간 동안 그녀는 침팬지가 어떻게 먹이와 거처를 찾는지, 다른 침팬지와 소통은 어떻게 하고 그 형태는 어떻게 되는지 등을 연구하였다. 최근 민족지학 연구의 예로 미라 블루본드-란저(Myra Bluebond-Langer, 1996)¹⁴의 연구가 있다. 병세가 심한 아이가 있는 가족들 내에서 의사결정이 어떻게 이루어지는지 물리적, 심리적, 환경적, 인종적, 법적, 문화적 측면에서의 영향을 살펴본 연구이다. 연구자는 2 년이 넘는 시간 동안 투병중인 아이들 80 명과 그 가족들과 함께 하였다. 자료수집은 참가자의 관찰과 그들의 가족, 친구, 친척, 병원 관계자들 간의 실제 공식적·비공식적 대화를 기록한 것을 중심으로 하였다.

현상학(phenomenology). 현상학은 주변 현실을 의식적 경험을 통한 연구로서 이해하는 방법론이다. 20 세기 초, 독일 철학자 에드먼 허슬러(Edmund Husserl)는 인간의 경험은 모든 지식의 원천이 된다고 주장하였는데, 현상학은 바로 이 개념을 기반으로 하고 있다. 현상학은 인간의 판단, 지각, 행동과 같은 의식적 경험을 바탕으로 하여, 현상에 대해 체계적으로 반영하고 분석한다. 이 연구의

¹⁴ Bluebond-Langer, M. (1996). In the Shadow of Illness: Parents and Siblings of the Chronically Ill Child. Princeton, NJ: Princeton University Press.

목적은 (1) 참여자의 다양한 주관적인 관점에서 사회현상을 바라보며 기술하는 것, (2) 참여자의 주관적인 경험에서 상징적인 의미와 그 내면의 구조를 이해하는 것이다. 현상학에서는 연구자가 어떠한 가정이나 개인적인 편견을 배제할 것을 강조한다. 참여자가 처해진 환경과 현상을 바라보는 다양한 측면을 조절하여, 그들의 의식적인 고민, 느낌, 행동을 내면적 구조를 이해하고 연구할 수 있어야 하기 때문이다.

어떤 학자들은 현상학을 연구방법론이 아닌 철학으로 분류하기도 한다. 이러한 관점은 조지(Giorgi & Giorgi, 2003)¹⁵의 연구에서 실존적현상학연구방법론(existential phenomenological research method)을 하나의 가이드라인으로서 제시함으로써 등장하였다(그림 12.2). 이 방법론은 크게 자료수집과 자료분석의 단계로 나뉜다. 자료수집의 단계에서 사회 현상에 속해있는 참여자들의 주관적인 경험과 관점을 인터뷰를 통하여 조사한다. 예를 들어, “평소 일상을 설명해주실 수 있습니까?” 나, “어떠한 상황에 대해서 자세히 설명해주실 수 있습니까?” 와 같은 질문을 던지는 것이다. 인터뷰는 모두 녹음 및 기록이 되어 분석의 근거로 쓰인다. 자료분석 단계에서 연구자는 인터뷰 기록을 읽으면서 (1) 전반적인 흐름을 찾고, (2) 참여자들의 주관적인 경험을 바탕으로 한 ‘중요도 단위(units of significance)’ 를 형성한다. 여기서 중요한 단위란 공간 인지(felt-space), 시간 인지(felt-time) 등과 같이 응답자의 심리적 경험을 문서화 한 것을 말한다. 예를 들어, 참여자가 현상을 경험할 때 안정감, 자유감, 속박감이나 즐거움을 느꼈는가?(공간 인지) 또는 참여자가 경험에서 압박을 받거나, 느리거나, 끊기는 느낌을 받았는가?(시간 인지) 와 같은 질문인 것이다. 현상학 분석은 참여자의 과거, 현재, 미래를 포함한 시간적인 관점을 중요시 하며, 연구자는 참여자의 특수한 상황에서 그들의 상상력을 동원하여 참여자를 이해해야 한다. 참여자들의 실제 경험은 서술적으로 기술하거나, 새로운 테마를 활용하여 기술한다. 이후의 분석은 쪼개지고 애매모호한 경험들에 대한 테마를 다양한 의미로 분석해야 한다.



<그림 12.2> 실존적현상학연구방법론

¹⁵ Giorgi, A and Giorgi, B (2003) Phenomenology. In J A Smith (ed.) Qualitative Psychology: A Practical Guide to Research Methods. London: Sage Publications.

12.6 해석적 연구의 검증 기준

실증적 연구가 ‘환원주의적(reductionist)’ 접근법을 활용하여 현실세계를 이론이나 법칙의 형태로 단순화한 것이라면, 해석적 연구는 현실세계를 해당 사회에 속한 구성원의 관점으로 경험을 해석하고 이해하는 것이다. 해석적 연구가 사회의 맥락과 흐름을 중시하는 특성 때문에 그 결과를 다른 맥락으로 보편화하는 힘들다. 그러나 해석적 연구가 상대적으로 주관적이고 경험을 중시하기 때문에 실증주의나 기능주의 학자들은 해석적 연구를 통한 분석이 보다 덜 엄격한 기준의 잣대를 가진다고 생각한다. 해석적 연구는 실증적 연구와 다르게 사회 현상에 대한 존재론적과 인식론적 가정에서부터 그 기준을 가진다. 실증적 연구에서 신뢰도(reliability), 내적 타당성(internal validity), 그리고 보편성(generalizability)이 연구의 기준을 뜻하지만, 해석적 연구에서는 이러한 요소들이 큰 변별력을 갖지 못한다. 다만, 린콜린과 구바(Lincoln & Guba 1985)¹⁶의 연구에서 해석적 연구에 대한 기준을 제시하였다.

의존성(dependability). 해석적 연구는 두 연구자가 동일한 현상에 대해 동일한 자료를 바탕으로 독립적으로 현상을 분석하여 동일한 결과가 도출되었거나, 한 명의 연구자가 시간을 두고 동일한 자료를 사용하여 비슷한 결과를 얻는 것은 연구의 의존성과 진실성이 있다고 판단한다. 이는 실증적 연구에서 신뢰성(reliability)과 같은 개념이라고 보면 된다. 이것은 서로 다른 두 연구자들이 동일한 결과를 얻었을 때 평가자 간의 신뢰도(inter-rater reliability), 혹은 한 명의 연구자가 동일한 결과에 대한 관찰을 반복적으로 한 경우인 조사-재조사 신뢰도(test-retest reliability)와 같은 개념인 것이다. 해석적 연구 결과의 의존성을 높이기 위해서 현상과 사회적 맥락에 대해 정확한 기술을 함으로써 독자로 하여금 해석적 분석을 신뢰할 수 있도록 해야한다.

신빙성(credibility). 만약 독자가 해석적 연구의 추론이 믿을만하다고 여길 때, 해석적 연구가 신빙성이 있다고 할 수 있다. 이 개념은 기능적 연구에서 내적 타당도(internal validity)와 유사한 개념이다. 해석적 연구의 신빙성을 높이기 위해서는 연구자가 충분히 현장에 개입했다는 증거를 제시하여야 한다. 즉, 자료수집 방식, 자료 처리, 관리, 분석하는 과정을 세세하게 기술하는데, 인터뷰의 녹취 기록, 응답자들의 응답에 대한 정확한 기록과 노트 등이 해당된다.

확증성(conformability). 연구를 통해 얻는 시사점이 연구하고자 한 현상을 경험한 사람들(참여자)에 의해 확증을 얻음으로써, 단순히 가설만이 아님을 보증하는 것이다. 이는 기능적 연구에서 객관성과 유사한 개념이다. 해석적 연구는 객관적인 현실의 관점으로 현실을 바라보는 것을 부정하기 때문에 확증성은 ‘내적 주관성(intersubjectivity)’ 라고 표현 할 수 있다. 예를 들어, 연구자가 이론이나 조사 보고서를 바탕으로 현상에 대하여 얻은 추론이 연구의 참여자로부터 확증을 받아 신뢰성을 증명한다.

전이가능성(transferability). 해석적 연구에서 전이 가능성이란 연구의 결과가 다른 상황에서도 적용될 수 있는지에 관한 것이다. 기능적 연구에서는 외적 타당도(external validity)와 비슷한 개념이다. 연구자는 연구하고자 하는 현상에 대한 충분한 설명과 연구 구조에 대한 묘사, 가정, 자료수집 과정 및 결과를 기술함으로써 독자들이 그들 스스로 다른 상황에서도 적용할 수 있는지 판단할 수 있도록 해야 한다.

¹⁶ Lincoln, Y. S., and Guba, E. G. (1985). *Naturalistic Inquiry*. Beverly Hills, CA: Sage Publications.

제 13 장 질적 분석

질적 분석은 인터뷰 기록에서 얻은 문서화 된 자료가 질적 분석의 자료가 된다. 양적 분석이 통계학 주도와 연구자의 독립적인 특성을 가지고 있다면, 질적 분석은 수집된 자료가 사회적인 맥락(context)에서 연구자의 통합적 분석능력에 따라 다르게 나타나기 때문에 연구자 ‘개인’에게 매우 의존하게 된다. 질적 분석에서 강조되는 것은 예상(predicting) 및 설명(explaining)이 아닌 현상에 대한 감각(sense making)이다. 질적 연구에서는 윤리적, 참여적 태도, 분석 전략에 기초한 창의적이고 연구적인 마음가짐이 필요하다. 이 장에서는 몇 가지 질적 연구에 대해 간략하게 설명하겠다. 질적 분석에 흥미 있는 독자들은 마일즈와 후버맨(Miles & Huberman 1984)¹⁷의 책을 읽어 보는 것이 좋다.

13.1 근거이론(Grounded Theory)

참여, 관찰, 심층 인터뷰, 이야기, 묘사, 소리/비디오 기록물 혹은 보조 문서를 통해 얻은 방대한 질적 자료를 어떻게 분석할 것인가? 이러한 텍스트화된 자료를 분석하는 기법 중 하나는 ‘**근거이론(grounded theory)**’ 방법이다. 이 기법은 현상에 대하여 이론을 성립하고자 할 경우, 사회 현상과 관련한 기록된 자료를 해석하는 귀납적 기술이다. 근거이론은 글라셔와 스트라우스(Glaser & Strauss 1967)¹⁸의 상수비교분석방법(method of constant comparative analysis)에서 제시되었으며 이후에 스트라우스와 코빈(Strauss and Corbin 1990)¹⁹이 코딩기법을 소개하면서 정립되었다. 자료를 분류하고 범주화하는 과정을 거쳐 코드형태(개념), 범주화(구성개념), 관계를 설정한다. ‘**근거(grounded in)**’는 실증적으로 관찰된 데이터를 사용하여 해석하기 때문에 붙여진 이름이다. 그렇기 때문에 단적으로 관찰된 증거를 바탕으로 만들어졌으며, 근거이론의 관점은 기존에 연구되어 있지 않은 이론적 기대나 편향이 존재하게 된다.

스트라우스와 코빈(Strauss & Corbin 1998)은 문서화된 자료를 분석하는 세 가지 코딩 기법을 소개하였다. 개방(open), 축(axial), 선택(selective)이 그것이다. **개방코딩(open coding)**이란, 현상으로부터 관찰된 자료의 잠재적인 주요 아이디어나 개념)들을 설정하는 것이다. 연구자는 개념과 관련된 불연속적인 사건이나 아이디어, 행동, 인식이나 관계를 원자료(raw data)로 분석한다. 각각의 개념은 특정한 범주 안에서 연결된다(분석 단위: coding unit). 몇몇 개념은 단순하고, 명료하고,

¹⁷ Miles M. B., Huberman A. M. (1984). *Qualitative Data Analysis: A Sourcebook of New Methods*. Newbury Park, CA: Sage Publications.

¹⁸ Glaser, B. and Strauss, A. (1967). *The Discovery of Grounded Theory: Strategies for Qualitative Research*, Chicago: Aldine.

¹⁹ Strauss, A. and Corbin, J. (1990). *Basics of Qualitative Research: Grounded Theory Procedures and Techniques*, Beverly Hills, CA: Sage Publications.

모호성을 가지고 있지 않은 반면에 몇몇 다른 개념은 복잡하고 애매모호함을 가지면서 그 개념을 보는 사람에 따라 다르게 표현될 수 있다. 따라서 분석단위는 개념에 따라 매우 추상적일 수 있다. ‘조직의 크기’와 같은 단순한 개념은 자료의 몇 가지 단어들로 확인이 가능하지만 ‘조직의 미션’과 같은 복잡한 개념은 몇 페이지에 걸쳐서 분석될 수 있다. 코딩은 표준화된 분석단위이다. 예를 들어, 미국과 프랑스의 소설을 비교하는 연구에서 문장, 단락, 장 혹은 책 전체에 코드를 부여할 수 있다. 코딩은 코드의 질적 분석에서 다중 단계를 통해 가능하지만, 코드가 같은 분석 단위를 가질 때만 비교될 수 있다. 예를 들어, 소설에서 ‘장’이 코딩 단위이면, 프랑스 소설의 경우 23%의 ‘장’이 메타포를 포함하고 있는 반면에, 미국 소설에서는 15%의 ‘장’이 메타포를 포함하고 있다고 말할 수 있다. 개념은 연구자의 고유 이름이나 연구 문헌에서 얻어진 표준화된 라벨을 활용하여 이름 지어질 수 있다. 개념의 기본 집합이 설정되면, 이 개념은 나머지에 자료에 코드로 부여하기 위해 활용될 수 있고, 새로운 개념을 찾거나 오래된 개념을 확인하는 것에 활용될 수 있다. 규모, 색상, 단계의 높고 낮음 등과 같은 각각의 개념이 인식될 수 있는 특성을 확인하는 것은 매우 중요하기 때문에 추후, 비슷한 개념이 광의의 범주로 집단화될 수 있다. 연구자는 관심 있는 현상과 관련 있는 새로운 개념을 개발하는데 있어서 개방적이고 능동적으로 행동하기 때문에 이 코딩 기법을 ‘개방(open)’이라고 한다.

비슷한 개념들은 상위범주로 **범주화(category)**된다. 개념들이 특정한 내용일 수 있지도 모르지만 범주는 광범위 하거나 일반화 할 수 있어서 근거이론 상에서 구성개념으로 발전된다. 범주는 연구자가 반드시 연구해야 하는 개념들로서 그 수를 줄이도록 노력하고 사회 현상을 이해하기 위한 ‘큰 그림’을 그리는데 집중해야 한다. 범주화는 개념을 하위범주에 의해서 구분 짓고 하위 범주를 상위 범주로 포함되어야 한다. 특히 연구의 목표가 현재 있는 이론을 확장하는 경우라면, 기존에 연구된 개념은 범주의 이름을 작성하는데 활용될 수 있다. 하지만, 신념(belief)이나 편향(bias)과 같이 기존의 구성개념을 이용할 때는 주의를 기울여야 한다. 구성개념이 가지고 있는 특성이나 차원은 범주 내에서 반드시 확인해야 한다. 차원(dimension)은 구성개념이 가지고 있는 특성들의 연속체를 나타낸다. 예를 들어, ‘커뮤니케이션 미디어’의 범주는 ‘속도’라는 특성을 가진 범주를 가질 수도 있고, 그것은 빠름, 중간 혹은 느림으로 차원화 될 수 있다. 이러한 범주화는 자료로부터 다른 종류의 커뮤니케이션 미디어 사이의 차이를 나타내게 해주고, 어떠한 커뮤니케이션 미디어가 어떠한 일을 하는지와 같은 패턴(pattern)을 연구자로 하여금 확인할 수 있게 해준다.

근거이론의 두 번째 과정은 **축코딩(axial coding)**이다. 상위범주와 하위범주가 연구자가 연구하고자 하는 현상에 대해 잠정적으로 설명할 수 있는 인과관계나 가설을 형성시켜 준다. 이것은 개방코딩과는 구별되지만, 축코딩은 개방코딩과 연계해 즉각적으로 작동할 수 있다. 범주들 사이의 관계는 자료 안에서 명백하게 발견되거나 미묘하거나 암시적인 것일 수도 있다. 후자의 예에서, 연구자는 현상에 내재되어 있는 환경과 같은 조건(condition), 이러한 조건에서 발생한 사건에 대한 개인의 반응과 같은 행동 혹은 상호작용(action/interaction), 행동과 상호작용의 결과(consequence) 등은 범주를 이해하기 위해 코딩계획(코딩 패러다임이라 불리지만, 3장에서 논의되었던 패러다임과는 다르다)으로 활용할 수도 있다. 조건, 행동/상호작용, 그리고 결과가 무엇인지 확인되면, 이론적 명제가 나타난다. 연구자는 이러한 조건에서 현상이 왜 일어났는지, 어떠한 결과가 있었는지 설명할 수 있다.

근거이론의 세 번째 과정은 **선택적코딩(selective coding)**이다. 가장 중심이 되는 범주 또는 핵심 변수가 포함되고 체계적이고 논리적으로 중심 범주와 이 외의 범주들의 관계를 정의하는 작업이다. 중심 범주는 기존의 범주에서 발전되었거나 코딩 된 범주들을 포함하는 상위 범주가 될 수도 있다. 새로운 데이터들은 중심범주와 다른 범주들 간의 타당성을 검증하기 위하여 선택적으로 표집된다(예, 잠정적 가설). 선택적 코딩은 분석의 범위를 한정하기도 하며 진행을 빠르게 한다. 동시에 코더(코딩을 작업하는 사람)는 현상과 관련된 새로운 자료로부터 관찰된 범주를 조심하여 초기 이론을 좀 더 강화하는 방향으로 작업을 진행해야 한다. 이러한 과정은 근거이론의 귀납적 요소인 이론적 표본추출(theoretical sampling)이라고 한다. 따라서 개방, 축, 선택적 코딩은 연속적인 과정을 거치게 된다. 새로운 자료의 코딩과 이론의 강화는 이론적 **포화점(theoretical saturation)**이 만족할 때까지 진행한다. 이론적 개념은 불완전한 개념들과 흥미로운 현상에 대한 전체론의 이론(holistic theory)을 견고하게 연결시켜 준다.

범주, 관계, 해석을 위한 지속적인 재배치, 통합과 강화의 과정은 ‘상수비교(constant comparison)’라고 하며, 네 단계의 이해와 반복되는 상호작용의 과정을 포함한다. (1) 범주를 검증하기 위한 각 범주들과의 연관된 사건이나 문서를 비교한다. (2) 범주들과 그 범주들의 범위를 통합한다. (3) 중요 개념에 집중하기 위한 이론적 범위를 정한다. (4) 메모, 스토리라인, 다이어그램을 사용한 이론을 작성한다. 주요 범위를 가지고 있는 것은 다른 범주와 통합 할 수 있다는 의미에서 매우 중요하다. 조건, 행동/상호작용, 결과는 주요 범주를 구분 짓는 요소이다. 스트라우스와 코빈(Strauss & Corbin 1990)은 메모, 스토리라인, 개념도와 같은 통합 기술을 제시하고 있는데 **스토리라인(storylining)**은 관찰한 현상에 대한 내용을 재정의하거나 해석하기 위해서 범주들과 그 관계들을 사용한다. **메모(memoing)**는 근거이론 분석을 통하여 실질적인 개념과 이론적으로 코딩 된 관계들에 대한 생각들을 적어 놓은 것이다. 메모는 분석을 하면서 개발된 생각들을 재정의하거나 파악하는데 중요한 도구이며, 2×2 표나 다이어그램, 수치, 또는 그림을 사용하여 패턴이나 관계를 보여준다. **개념도(concept mapping)**는 개념들 사이의 관계를 박스모양과 화살표를 이용하여 도식적으로 표현 한 것이다. 일반적으로 한 장 이상의 문서나 칠판, 그래픽 소프트웨어를 사용하여 주요 개념들이 그려지고 관찰된 자료에 가장 적합한 방향으로 화살표를 사용하여 서로 연결시켜 보여준다.

근거이론을 수행한 후에는 반드시 내적 일치도와 논리에 대해 다시 살펴보아야 한다. 연구자는 중심 구성개념을 특성과 차원을 가지고 있도록 하는데 그렇지 않다면, 자료는 다시 분석되어야 한다. 모든 범주들의 특성과 차원들은 서로 차이를 보여주어야 한다. 예를 들어, 행동의 발생 빈도가 어떤 한 범주에 속해 있다면 그 자료는 반드시 행동의 범주나 빈도의 범주에 속해 있어야 한다는 것이다. 마지막으로 이론은 반드시 원 자료(raw data)에서 분석되어야 타당성을 가질 수 있다. 만약 이론이 관찰된 증거들과 모순된다면 코딩 프로세스는 모순적인 부분들이나 설명되지 않은 점들에 대해서 다시 조정하는 작업을 거쳐야만 한다.

13.2 내용분석(Content Analysis)

내용분석(content analysis)는 누가 무엇을, 누구에게, 왜, 어느 정도까지, 어떠한 효과로 말했는지와 같은 텍스트(text)에서 질적 방법이나 양적 방법을 통해 체계적으로 분석하는 것이다. 내용 분석은 일반적으로 다음과 같은 순서로 수행된다. 첫째, 분석할 많은 분량의 텍스트들(예, 신문기사, 리포트

등) 가운데 연구자는 텍스트의 모집단으로부터 표집작업을 수행한다. 이 과정은 임의적으로 행해져서는 안되며 연구 주제와 적절히 관련이 있는 내용들이 선택되어야 한다. 둘째, 연구자는 표집작업을 통해 선택된 텍스트들을 분류하고 ‘작은 단위(chunk)’의 분석의 단위로 일정한 규칙을 적용시켜 세분화 한다. 이 과정을 단위화(unitizing)라고 한다. 예를 들어, 텍스트 내에서 추정, 영향, 조력자, 장벽과 같은 것들은 단위로 구성 될 수 있다. 셋째, 연구자는 코딩(coding)으로 불리는 과정에서 단위화 된 텍스트들에게 하나 또는 그 이상의 개념을 적용 시키고 구성화한다. 코딩은 연구자가 텍스트를 분류하면서 새로 찾거나 알아내는 것이라는 전제하에 이루어져야 한다. 마지막으로, 코딩 된 자료는 가장 빈번하게 나오는 내용이 무엇인지, 그것들이 다른 것들과 어떻게 연관되어 있는지 알아보기 위해 질적 연구와 양적 연구 모두 분석된다.

내용분석의 단순한 형태는 **평판분석(sentiment analysis)**이다. 이 기법은 물체, 사람, 현상에 대한 사람들의 의견이나 태도분석에 사용된다. 온라인 상에서 정치가에 대한 온라인 메시지를 읽고 그 내용이 긍정적, 부정적, 중립을 가지고 있는지 분류하는 것이다. 여기서 각각의 메시지는 하나의 분석단위로 표현된다. 이 분석방법은 전체 표집군(메시지)이 정치가에 대한 긍정적인지, 부정적인지, 중립을 가지고 있는지 구별하는데 도움을 준다. 비록 이 방법이 수작업을 통해 이루어진다고 하지만, 수백만의 큰 자료에 대해서는 코딩 과정을 자동화하면서 사람의 감각적 기능을 유지해 주는 텍스트 분석 소프트웨어를 사용할 수 있다.

내용분석이 연구자들로부터 자주 비판되어 온 이유 중 하나는 체계적인 절차가 약해, 분석이 다른 연구자들로 하여금 복제가 된다는 것이다. 실링(Schilling 2006)²⁰ 은 다섯 레벨로 구성된 텍스트를 분석하는 과정을 나선형 모델로 제시함으로써 내용 분석의 단점을 반증하였다. (1) 내용 분석을 위해 녹음기록을 글로 옮겨 기록하거나 원자료(raw data)형태로 변환한다. (2) 원자료를 간결한 프로토콜의 형태로 변환시킨다. (3) 프로토콜을 예비범주시스템으로 변환시킨다. (4) 예비범주시스템을 활용하여 코드화 된 프로토콜 형태로 기록한다. (5) 코드화 된 프로토콜을 분석하고 현상에 대해 해석한다.

하지만 내용분석은 몇 가지 한계점을 갖고 있다. 먼저, 코딩 과정에서 문서화 된 형태는 정보의 이용 가능성을 제한한다. 예를 들어, 만약 당신이 사형제도에 대한 사람들의 견해를 연구하고자 할 때, 연구와 관련된 적절한 문서 없이는 내용분석을 하기 어려울 것이다. 둘째, 표집은 반드시 표집편향을 피하기 위해 조심스럽게 행해져야만 한다. 만약 연구의 모집단이 주어진 주제에 대해 공식적으로 발표된 연구 문헌이라고 한다면, 당신은 공식적으로 발표되지 않은 연구나 아직 공식화되지 않은 최근 연구들을 제외해야만 할 것이다.

13.3 결론

질적 연구에서는 양적 연구방법을 이용하여 자료를 분석하거나 수치를 사용하여 현상에 대해 해석하는 것이 받아들여지는데 이러한 분석을 혼합방법분석(mixed-method analysis)이라고 한다. 예를 들어, 인터뷰 기록의 질적 자료는 내용분석을 통해 양적 연구방법으로 분석되거나, 양적 자료는 중심성향(central tendency), 상관관계(correlation)의 측정을 통해 질적 연구방법으로 분석될 수 있다.

²⁰ Schilling, J. (2006). "On the Pragmatics of Qualitative Assessment: Designing the Process for Content Analysis," *European Journal of Psychological Assessment* (22:1), 28-37.

질적 연구와 양적 연구 결과를 결합하는 것은 연구자에게 때때로 하나의 연구방법만 사용하는 것보다 흥미로운 현상에 대해 더욱 풍부한 이해를 돕는다.

최종적으로, 질적 연구방법에서 ATLAS.ti.5, NVivo, 그리고 QDA Miner 와 같은 표준 소프트웨어 프로그램은 코딩과정을 자동화하는데 사용된다. 이런 프로그램들은 빠르고, 효율적으로 사용자가 정의한 규칙에 따라 아주 많은 양의 텍스트를 정리하고, 찾고, 분류하고, 처리할 수 있다. 이렇게 자동화된 분석을 사용하기 위해서는 코딩 스키마(schema)가 만들어져야 하고, 텍스트 내에서 찾기 위한 키워드와 코드를 정리해야 한다. 이것은 초기 매뉴얼 검사를 바탕으로 먼저 이루어져야 한다. 스키마는 코드를 상위 코드의 집단이나 구성개념으로의 계층화된 체계를 바탕으로 구성될 것이다. 코딩 스키마는 정확성과 적절성을 위해 서로 다른 표본을 사용하여 검증해야 한다. 하지만 만약 코딩 스키마가 편향적이거나 정확하지 않다면 전체 모집단으로의 분석결과는 결함을 가지고 있거나 해석하기 불가능할 것이다. 또한 소프트웨어 프로그램을 사용하여 텍스트를 분석할 경우, 특정 단어나 구, 문장 속에 숨겨져 있는 의미(예, 은유, 풍자)까지 해독할 수 없기 때문에 큰 규모의 질적 연구에서 잘못된 해석으로 초래 될 수 있다.

제 14 장 양적 분석-기술적 통계

연구에서 수적 자료는 통계 도구를 활용하여 양적 분석이 가능하다. **기술분석(descriptive analysis)**은 통계학적으로 구성개념이나 그것들 간의 관계를 서술하고, 종합하여 나타내는 것이다. **추론분석(inferential analysis)**은 가설을 통계학적으로 검증(이론검증)하는 것이다. 이 장에서는 기술분석을 위한 통계학적 기법에 대해 알아보고 다음 장에선 추론분석을 위한 통계학적 기법에 대해 알아보겠다. 독자들은 이 장에서 서술되는 개념들을 이해하기 위해 SPSS 나 SAS 같은 통계 프로그램과 친숙해지는 것을 추천한다.

14.1 자료 준비(Data Preparation)

연구에서 자료는 우편조사, 인터뷰, 실험, 관찰 등 다양한 방법으로부터 수집되어야 한다. 그리고 이러한 데이터는 스프레드시트나 텍스트 파일과 같이 기계가 해석할 수 있도록 수적 형태로 변환시켜야 한다. 이러한 데이터의 준비과정은 SPSS 혹은 SAS 와 같은 컴퓨터 프로그램으로 분석될 수 있게 하기 위함이다. 데이터 준비는 다음의 과정을 거치게 된다.

데이터 코딩(data coding). 코딩은 자료를 수적인 형태로 변환하는 과정이다. 코딩하는 과정을 기록해 놓은 책을 코드책이라고 하는데 **코드책(codebook)**은 연구에서 변수에 대한 상세한 내용을 포함하여 이해를 돕는 문서이다. 변수의 항목, 측정방법, 각 항목의 형태(수, 텍스트 등), 각 항목의 응답 척도(명목, 서열, 구간, 비율 척도 중 어느 것으로 측정되었는지, 이러한 척도가 5 점, 7 점, 아니면 다른 타입의 척도인지), 각 항목에 대하여 어떻게 수의 형태로 코드를 부여하였는지가 그것이다. 예를 들어, 측정 항목이 ‘매우 동의하지 않음’에서 ‘매우 동의함’까지 7 점 리커트 척도를 가지고 있다면, 1 을 ‘매우 동의하지 않음’, 4 를 ‘중립’, 7 을 ‘매우 동의함’으로 코드를 부여해야 한다. 산업 형태와 같은 명목자료(nominal data)는 다음과 같은 코딩계획에 의해 숫자의 형태로 코드가 부여되어야 한다. 1 을 제조업, 2 를 소매업, 3 을 금융업, 4 를 의료업 등이다(물론, 명목 자료는 통계학적으로 분석할 수 없다). 나이, 수입, 혹은 시험 점수와 같은 비율척도자료(ratio scale data)는 응답자들에 의해 코드가 부여된다. 때때로 데이터는 자료수집을 위한 형태보다는 다른 형태로 종합될 필요가 있다. 예를 들어, ‘컴퓨터의 이익’과 같은 구성개념을 측정하기 위해 응답자에게 이익에 대한 설문조사를 하고 그것들 중 응답자가 선택할 수 있게 한다면(응답자는 수많은 컴퓨터의 이익들 중에서 그들이 원하는 것을 선택할 수 있음), 선택된 항목은 이익의 종합적인 측정 단위로 활용될 수 있는 것이다. 인터뷰 기록지와 같은 자료는 통계학적인 분석을 위한 수의 형태로 변환되지 못한다. 코딩은 많은 변수와 다수의 측정항목을 포함하는 복잡한 연구에서 특히 중요하다. 코딩과정은 코드의 일관성을 갖기 위해서 코드 팀의 도움을 받게 되는데 코드책은 그들이 코드화된 데이터를 이해하고 해석하는데 도움을 준다.

데이터 입력(data entry). 코드가 부여된 데이터는 스프레드시트나 데이터베이스, 텍스트 파일이나 SPSS 와 같은 통계 프로그램에 직접적으로 입력할 수 있다. 대부분의 통계 프로그램은 데이터를 입력시키기 위한 자료 편집기능을 제공한다. 하지만, 이러한 프로그램은 자신들만의 확장자를 가지고 데이터를 저장한다(예: SPSS 는 .sav 파일형태). 따라서 특정 통계프로그램에 저장된 데이터는 다른 통계프로그램으로 공유게 만드는 것을 어렵게 한다. 그렇기 때문에, 데이터를 스프레드시트나 데이터베이스에 입력하는 것이 활용도가 좋다. 연구자는 데이터를 스프레드시트에 입력시켜 놓으면, 필요에 의해 재 정리 할 수 있고, 통계 프로그램에 따라 공유할 수 있고, 분석을 위해 데이터를 하부집합으로 추출할 수 있다. 65,000 개 이하의 관찰 수와 256 개 이하의 항목은 마이크로소프트 엑셀과 같은 스프레드 시트에 저장할 수 있지만 수 백만개로 관찰된 큰 사이즈의 데이터 집합은 데이터베이스를 필요로 하게 된다. 관찰된 값은 스프레드시트에서 하나의 열로 입력될 수 있고, 측정 항목은 하나의 행에 나타날 수 있다. 입력된 데이터의 정확성을 위해 입력하는 동안에 혹은 입력한 후에 자주 검토해서 정확하게 입력하였는지 확인해야 한다. 특히, 데이터를 입력하는 동안, 코드를 부여하는 사람은 잘못된 데이터를 제대로 다루어야 한다. 응답자가 설문 문항의 내용을 무시하고 모두 ‘매우 동의함’ 이라고 응답하였을 경우, 부정문의 문항을 가지고 있던 것에 대한 응답을 반대코드로 부여하지 못한 예가 있다. 이러한 데이터는 당시에 입력은 가능 할지 모르지만, 차후 분석에서 배제된다.

결측 값(missing value). 결측 데이터는 어떠한 실증 데이터에서 필수 불가결한 부분이다. 응답자는 특정한 질문이 애매모호하게 표현되었다거나 혹은 너무 민감할 경우, 대답을 하지 않을 수 있기 때문이다. 이러한 문제점들은 주요 자료수집 과정이 시작되기 전에 예비검사를 통해 발견하여 수정되어야 한다. 데이터를 입력하는 동안에 어떤 통계 프로그램은 자동적으로 공란 부분에 대하여 결측 값으로 간주하여 처리한다. 또 다른 통계프로그램은 결측 값을 나타내기 위해 -1 이나 999 같은 특정한 수치를 입력해야 한다. 자료분석 동안, 대부분의 소프트웨어 프로그램에서 결측 값을 설정하는 모드에서 하나의 항목에서 관찰된 다수의 관측 값 가운데 결측 값이 한 개라도 있을 경우, 그 항목을 버리는 것을 일률적삭제(listwise deletion)기법이라고 한다. 이것은 표본크기를 급격하게 줄어들게 함으로써 측정하고자 하는 미세한 효과를 찾아내기 매우 어렵게 만든다. 그렇기 때문에 몇몇 소프트웨어 프로그램은 결측 값을 추정 값으로 대체하는 옵션을 허용하는데, 이러한 과정을 대치(imputation)라고 부른다. 예를 들어, 다중항목척도 상에 한가지 항목에 결측 값이 있다면, 대치되는 값은 척도상 남아있는 항목에 대한 응답자의 평균 값이 된다. 만약 결측 값이 단일항목 척도 상에 존재한다면, 많은 연구자들은 항목에 대한 다른 응답자들의 평균 값을 대치 값으로 활용한다. 만약 결측 값이 임의적인 특성이라기보다는 체계적인 특성을 가지고 있다면, 대치 값은 편향될 수 있다. 상대적으로 편향되지 않은 추정 값을 생성하기 위한 두 가지 방법은 최우법(maximum likelihood procedure)과 다중대치법(multiple imputation method)이 있는데, 두 방법 모두 SPSS 와 SAS 와 같은 소프트웨어 프로그램에서 지원된다.

데이터 변환(data transformation). 자료가 의미 있게 해석되기 전, 필요에 의해 데이터 값을 변환한다. 예를 들어, 거꾸로 코드가 부여된 항목들은 근본적인 구성개념에 대해 반대의 의미를 전달하기 때문에 반드시 반전되어야 한다(항목이 다른 항목과 비교되거나 혹은 반전이 되지 않은 다른 항목과 합쳐지기 전에 반전이 되어야 한다). 개별적인 척도 항목을 더하여 척도 측정 값을 생성하고, 관찰된

측정 값 집합에서 가중된 인덱스를 생성하고, 다중 값을 작은 범주로 폭락시키는 것(예, 수입을 수입 범위로부터 폭락시킴) 다른 종류의 데이터 변환이 존재한다.

14.2 단변량분석(Univariate Analysis)

단변량분석 혹은 단일변수에 대한 분석은 하나의 변수를 특징적으로 서술할 수 있는 통계학적 기법을 말한다. 단변량 분석은 (1) 도수분포, (2) 중심화 경향, (3) 분산을 포함한다. **도수분포(frequency distribution)**는 변수의 개별 값에 대한 빈도(혹은 퍼센티지)나 변수 값의 범위를 종합하여 나타낸다. 예를 들어, ‘독실함’의 측정도구로서 응답자들이 종교활동에 얼마나 참여하는지를 범주척도(categorical scale)로 나타내면 ‘전혀’, ‘일년에 한번’, ‘일년에 몇 번’, ‘한 달에 한번’, ‘한 달에 몇 번’, ‘주당 몇 번’, ‘응답하지 않음’과 같다. 각 범주 내에서 관찰의 빈도(퍼센티지)를 세고 그림 14.1 과 같이 표의 형태로 보여준다면, 이것이 도수분포인 것이다. 그림 14.1 의 오른쪽 부분에서 보여지는 것과 같은 분포는 막대 차트의 형태로 묘사된 것이다. 수평축은 각 범주의 변수를 나타내고 수직축은 각 범주에서의 관측 빈도 혹은 퍼센티지를 나타낸다.



<그림 14.1> 독실함의 도수분포

관찰이 독립적이고 임의표본의 도수분포는 종형곡선(bell-shaped curve)을 따르는 경향이 있다. 대부분 관찰 값의 범위가 중앙에 무리를 이루고 있으며 양쪽으로 갈수록 좀 더 적은 수의 관찰 값이 나타나 극단점을 향한다. 이러한 곡선형태를 나타내는 분포를 **정규분포(normal distribution)**라 부른다.

중심화경향(central tendency)은 관찰 값들의 분포가 중앙에 모여있음에 대한 추정치이다. 중심화경향의 세 가지 주요한 추정치가 있다. 평균값(mean), 중앙값(median), 최빈값(mode)이 그것이다. 간단하게 평균이라 불리는 **산술평균(arithmetic mean)**은 주어진 분포에서 모든 값들의 단순 평균값이다. 예를 들어, 여덟 명의 시험점수가 15, 22, 21, 18, 36, 15, 25, 15 이다. 이 값들의 산술평균은 $(15+22+21+18+36+15+25+15) / 8 = 20.875$ 이다. 기하평균(geometric mean)은 n 개의 양수가 있을 때, 이들의 곱에 n 제곱근의 값을 나타내며, 조화평균(harmonic mean)은 n 개의 양수에 대하여 그 역수들을 산술평균 한 것의 역수로 나타낸다. 이러한 평균 값들은 사회연구 데이터를 분석하는데 널리 활용되지 않는다.

중심화경향의 두 번째 측정값은 **중앙값(median)**이며, 중앙값은 분포 내에 있는 값들 가운데 가장 중앙에 위치한 값을 의미한다. 이것은 분포 내의 모든 값을 오름차순으로 정리한 후 중앙 값을

고르는 것으로 계산된다. 분포 내에 짝수개의 값들이 있는 경우에는 두 개의 중앙 값이 존재하게 되는데, 이 때 두 값의 평균을 중앙 값으로 나타내면 된다. 위의 예제에서, 정렬된 값들은 15, 15, 15, 18, 21, 22, 25, 36 이다. 두 중앙 값은 18 과 21 이기 때문에, 중앙값은 $(18+21)/2=20$ 이다.

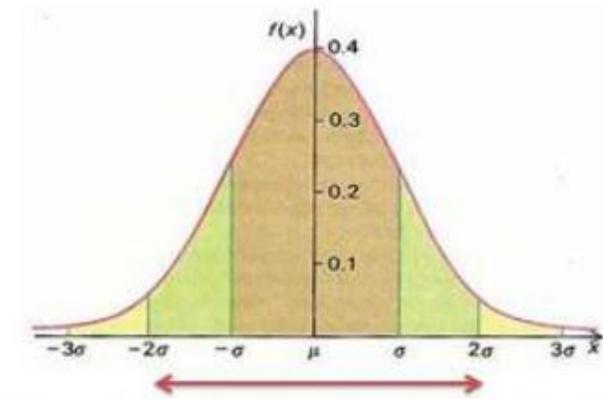
마지막으로 **최빈값(mode)**은 분포 내에 관찰 값들 가운데 가장 많이 나타나는 값이다. 앞선 예제에서 가장 많이 나타나는 값은 15 이며, 이는 시험점수 집합의 최빈값이다. 표본에서 추정되는 평균, 중앙값, 최빈값 혹은 후에 추정되는 모든 값은 **통계량(statistic)**이라고 불리는 것을 잊지 말아야 한다.

분산(dispersion)은 관찰 값이 중앙화 경향 근처에서 얼마나 퍼져있는지를 의미한다. 예를 들어, 평균 주위에 얼마나 타이트하게 혹은 넓게 관찰 값들이 무리를 지어있는 지가 분산이라 할 수 있다. 분산은 범위와 표준편차로 측정할 수 있다. **범위(range)**는 분포에서 가장 높은 값과 가장 낮은 값과의 차이이다. 앞선 예제에서의 범위는 $36-15=21$ 이다.

범위는 이상 값이 존재할 경우 특히 민감하게 반응한다. 예를 들어, 분포 내에서 가장 높은 값이 85 이고 다른 값들은 그대로라고 가정할 경우, 범위는 $85-15=70$ 이 될 것이다. **표준편차(standard deviation)**는 이상값은 각각의 관찰 값의 평균으로부터 얼마나 가깝고 먼지를 나타내는데 그것은 아래의 공식과 같다.

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \mu)^2}{n - 1}}$$

σ 는 표준편차, x_i 는 i 번째 관찰 값, μ 는 산술 평균, n 은 관찰 값의 총 개수이고 Σ 는 모든 값들의 총합을 의미한다. 표준편차의 제곱은 분포의 **분산(variance)**이라고 불린다. 일반적인 도수분포에서 관찰 된 68%의 관측 값은 평균의 하나의 표준편차($\mu + 1\sigma$) 안에 포함되고, 관찰 된 95%의 관측 값은 두 개의 표준편차($\mu + 2\sigma$)안에 위치하고, 관찰된 99.7%의 관측값은 세 개의 표준편차($\mu + 3\sigma$)안에 위치함을 그림 14.2 에서 보여주고 있다.



<그림 14.2> 정규분포

14.3 이변량분석(Bivariate Analysis)

이변량분석은 두 개의 변수가 서로 얼마나 연관되어 있는지를 검증하는 것이다. 보통의 이변량 통계량은 간단히 ‘상관관계’라고 불리는 **이변량상관관계(bivariate correlation)**로 측정한다. 두 변수의 관계의 강도는 -1 부터 1 사이의 숫자로 나타낸다. 나이와 자존감이 얼마나 연관되어 있는지를 20 명의 응답자를 통해 알아본다고 가정해보자(예: 나이가 먹을수록, 자존감은 증가한다, 감소한다, 변하지 않는다). 만약 자존감이 증가한다면, 두 변수간에는 양의 상관관계가 존재하는 것이고, 자존감이 감소한다면, 두 변수간에는 음의 상관관계가 존재하는 것이고, 자존감의 변화가 없다면, 영의 상관관계가 있는 것이다. 상관관계의 값을 계산하기 위해서, 표 14.1 에서 보여주는 데이터 집합체를 살펴보자

관찰수	나이(x)	자존감(y)	xy	x^2	y^2
1	39	4.1	159.9	1521	16.81
2	45	4.6	207	2025	21.16
3	29	3.8	110.2	841	14.44
4	42	4.4	184.8	1764	19.36
5	19	3.2	60.8	361	10.24
6	22	3.1	68.2	484	9.61
7	39	3.8	148.2	1521	14.44
8	30	4.1	123	900	16.81
9	33	4.3	141.9	1089	18.49
10	23	3.7	85.1	529	13.69
11	20	3.5	70	400	12.25
12	18	3.2	57.6	324	10.24
13	24	3.7	88.8	576	13.69
14	22	3.3	72.6	484	10.89
15	29	3.4	98.6	841	11.56
16	35	4.0	140	1225	16.00
17	36	4.1	147.6	1296	16.81
18	37	3.8	140.6	1369	14.44
19	35	3.4	119	1225	11.56
20	32	3.6	115.2	1024	12.96
합계	609	75.1	2339.1	19779.0	185.45

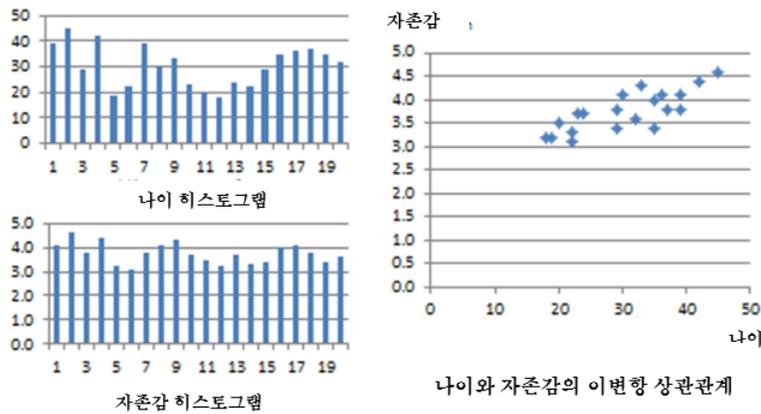
<표 14.1> 나이와 자존감에 대한 데이터

위의 데이터 집합에서 두 변수는 나이(x)와 자존감(y)이다. 나이는 비율척도이고 자존감은 ‘강하게 동의하지 않음’에서 ‘강하게 동의함’의 7 점 리커트 척도를 활용하여 자존감에 대한 다중 항목 척도들을 평균한 값이다. 각 변수의 히스토그램은 그림 14.3 의 왼편에 있으며 이변량 상관관계를 계산하기 위한 공식은 다음과 같다.

$$r_{xy} = \frac{\sum(x_i y_i) - n \bar{x} \bar{y}}{(n-1)S_x S_y} = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n \sum x_i^2 - (\sum x_i)^2} \sqrt{n \sum y_i^2 - (\sum y_i)^2}}$$

r_{xy} 는 상관관계, x와 y는 x와 y의 표본평균 값, S_x 와 S_y 는 x와 y의 표준편차 값이다. 표 14.1에

있는 값을 위의 공식을 활용하여 수동적으로 나이와 자존감 사이의 상관관계 값을 계산하면 0.79 이다. 따라서 나이와 자존감이 강한 양의 상관관계를 지니고 있음을 나타내 준다(예: 자존감은 나이가 들면서 높아지는 경향을 보이며, 나이가 적어질수록 낮아지는 경향을 보인다). 이러한 패턴은 그림 14.3 에서 보이는 나이와 자존감을 비교한 히스토그램을 통해 시각적으로 보여질 수 있다. 그림 14.3 에서의 수직 축은 실제 관찰값을 나타내는 것이지 관찰의 빈도를 나타내는 것은 아님을 주의해야 한다. 그렇기 때문에, 이것은 도수분포가 아닌 히스토그램이라 할 수 있다. 그림 14.3 에서 오른쪽의 이변량 산점도는 기본적으로 가로축을 나이, 세로축을 자부심으로 한 것이다. 이 도표는 양의 상관관계를 나타내는 상승하는 경사면(positive slope)과 닮아있다. 만약 두 변수가 음의 상관관계라면 산점도는 반대로 나이가 증가할수록 자부심이 감소하는 것을 보여주는 내려가는 경사면(negative slope)으로 나타났을 것이다. 만약 두 변수가 상관관계가 없다면, 산점도는 나이의 증가와 자부심은 관계가 없다는 것을 의미하는 수평선(zero slope)에 가까운 모양으로 나타날 것이다.



<그림 14.3> 나이와 자존감의 상관관계와 히스토그램

이변량 상관분석 후에, 연구자는 흔히 상관관계가 유의미한지 아니면 우연적으로 발생된 것인지에 대해 알고 싶어한다. 이러한 문제에 답하는 것은 다음의 가설을 검정함으로써 알 수 있다.

$$H_0: r = 0, H_1: r \neq 0$$

H_0 은 귀무가설(null hypothesis)이고 H_1 은 대립가설(alternative hypothesis)이라고 한다(H_a 로도 나타낸다). 대립가설은 상관관계가 있는지 확인하고자 하는 우리가 실제로 검정하고자 하는 가설이다. H_0 과 H_1 은 단순히 두 개의 가설로 보일 수 있지만 두 가설은 서로 반대되기 때문에 공동으로 하나의 가설을 나타낸다고 할 수 있다. 또한 H_1 은 r 이 0보다 큰지 작은지에 대해 구체적으로 명시하고 있지 않기 때문에 비특질가설(non-directional hypothesis)이라는 것도 알아두어야 한다. 특질 가설(directional hypothesis)은 $H_0: r \leq 0; H_1: r > 0$ 으로 명시될 것이다 (양의 상관관계를 검정할 경우). 비특질가설은 양측 T 검정(two-tailed t-test)을 통해 검증되는 반면에 특질 가설의 유의미한 검정은 단측 T 검정(one tailed t-test)을 이용하여 검증된다.

통계적 검정에서 대립가설은 직접적 검증할 수 없다. 오히려 특정한 확률수준에서 귀무가설을 기각함으로써 간접적으로 증명된다. 통계적 검정은 우리가 가진 표본이 모집단과 똑같지 않고,

표본을 바탕으로 한 추론이 모집단에도 적용될 수 있는지 확신 할 수 없기 때문에 항상 확률론적인 방법을 사용할 수밖에 없다. 통계적인 추론이 순수하게 일어날 확률을 **p 값(p-value)**이라고 한다. p 값은 추론이 맞지 않다고 결론 낼 수 있는 최대수준을 **유의수준(significance level; α)**과 비교한다. 대부분의 통계분석에서 α 값이 0.05 로 설정된다. $\alpha=0.05$ 보다 작은 p 값은 귀무가설을 기각하는데 충분한 근거가 있다는 것을 가리킴으로써 대립가설을 간접적으로 채택하게 된다. 만약 $p>0.05$ 이라면 귀무가설을 기각하거나 대립가설을 채택하는데 통계적인 증거를 충분히 갖고 있지 못하다는 것을 뜻한다.

위의 가설을 검정하는 가장 쉬운 방법은 인터넷이나 통계분야의 표준 교과서에서 찾아 볼 수 있는 통계표(statistical table)로부터 r 의 기각값(critical value)를 찾아보는 것이다(대부분의 소프트웨어 프로그램들도 유의성검정을 수행한다). r 의 기각값은 우리가 원하는 유의수준($\alpha=0.05$), 자유도, 단측검정인지 양측검정인지에 따라 다르게 나타난다. **자유도(degree of freedom)**는 통계적 계산으로 자유롭게 바뀔 수 있는 값이다. 상관관계의 경우 자유도가 $n-2$ 라고 한다면, 표 14.1 의 데이터에서 자유도는 $20-2=18$ 이 된다. 여기에는 단측검정과 양측검정을 위한 두 가지 다른 통계표가 있다. 양측검정표에서 $\alpha=0.05$, $df=18$ 에서 r 의 기각값은 0.44 이다. 컴퓨터로 계산한 상관관계 값은 0.79 이며, 상관관계가 유의미하기 위해서는 반드시 기각값 0.44 보다 크거나 -0.44 보다 작아야 한다. 0.79 는 0.44 보다 크기 때문에 자료에서 나이와 자부심간에 유의미한 상관관계가 있다. 다시 말하면, 이 상관관계가 우연히 일어났을 확률인 오즈비(odds rate)가 5%보다 적다고 결론지을 수 있다. 따라서 우리는 귀무가설($r \leq 0$)을 기각할 수 있으며 또한 대립가설($r > 0$)이 옳은 것이라고 간접적으로 말할 수 있다.

대부분의 연구들은 둘 이상의 변수를 포함한다. 만약 n 개의 변수가 있다면 변수들 사이에 총 $n*(n-1)/2$ 의 상관관계가 가능 할 것이다. 이러한 상관관계는 상관관계 공식을 사용하여 손으로 계산하는 것 보다는 SPSS 와 같은 소프트웨어프로그램을 사용하면 쉽게 계산 된다. 이것은 표 14.2 에서 보여지는 것처럼 상관행렬(correlation matrix)로 나타낼 수 있다. 상관행렬은 행과 열을 따라 변수의 이름을 목록화한 행렬로서 셀 안에 있는 값은 두 변수들 간 이변량 상관관계를 나타낸다. 어떤 변수이던지 자기 자신과는 항상 완벽히 상관관계에 있으므로 행렬의 주 대각선을 따르는 값은 1 이다. 또한 상관관계는 비방향성의 특징을 가지고 있기 때문에 변수 V1 과 V2 의 상관관계는 V2 와 V1 의 관계와 같다. 그래서 아래쪽 삼각형태의 행렬(주로 대각선의 아래쪽)과 위쪽의 삼각형태의 행렬(주로 대각선의 위쪽)은 서로 거울반사가 된 것이라고 할 수 있다. 그러므로 보통은 간단하게 아래쪽 삼각형태의 행렬만을 나타낸다. 만약 상관관계가 구간척도를 사용하여 측정된 변수들을 포함할 경우 이 특수한 형태의 상관관계를 **피어슨적률상관관계(pearson product moment correlations)**라고 한다.

이변량 자료를 나타내는 또 다른 유용한 방법은 교차표(cross-tabulation)이다(보통 축약해서 cross-tab 또는 분할표(contingency table)라고 불림). **교차표(cross-tabulation)**는 둘 이상의 명목형 범주로 나타내는 변수들의 모든 결합빈도(또는 비율)를 나타내는 표이다. 예를 들어, 그림 14.3 에서 보여지는 것과 같이 20 명의 학생들을 표본으로 학점과 성별을 다음과 같이 관찰했다고 가정하자. 성별은 명목형 변수(남자/여자 또는 남/여)이고 학점은 세 레벨로 나누어지는 범주형 변수이다(A,B,C). 데이터의 간단한 교차표는 2×3 행렬에 성별과 학점의 결합분포(joint distribution)로 나타낸다. 즉, 각 성별 범주에서 얼마나 많은 학생들이 각각의 학점에 있는지 빈도수 또는 비율로 나타낸다. 또한, 이 행렬은 A, B, C 의 학점에 남학생과 여학생 전체에 걸쳐 동등하게 분포되어있는지 볼 수 있도록

도와준다. 표 14.3 의 교차표는 A 학점 분포는 전체 학생들 가운데 여학생이 아주 많이 치우쳐 분포되어있는 것을 보여준다. 10 명의 남학생과 10 명의 여학생 표본에서 오직 한 명의 남학생이 A 학점을 받은 것에 비해 여학생은 다섯 명이 A 학점을 받았다. 반대로 C 학점 분포는 전체 학생들 가운데 남학생 쪽으로 편향되어 있다. 단지 한 명의 여학생이 C 학점을 받은 것에 비해 남학생은 세 명이 C 학점을 받았다. 하지만 B 학점의 분포는 여섯 명의 남학생과 다섯 명의 여학생으로 다소 균일하다. 이 표의 마지막 열과 행은 각 범주에 걸친 합계를 나타내고 있으며, 표의 주변을 따라 보여주기 때문에 주변합계(marginal total)라고 한다.

	v1	v2	v3	v4	v5	v6	v7	v8
v1	1.000							
v2	0.274	1.000						
v3	-0.134	-0.269	1.000					
v4	0.201	-0.153	0.075	1.000				
v5	-0.095	-0.166	0.278	-0.011	1.000			
v6	-0.129	0.280	-0.348	-0.378	-0.009	1.000		
v7	0.171	-0.122	0.296	0.086	0.193	0.233	1.000	
v8	0.518	0.238	0.238	-0.227	-0.551	0.082	-0.102	1.000

<표 14.2> 8 개 변수의 가설 상관행렬

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
성별	여	남	여	남	여	남	남	남	여	여	남	남	남	여	여	남	여	여	여	남
성적	A	B	B	B	C	A	C	B	B	A	B	C	C	B	B	B	A	A	B	B

데이터 셋

		성적			총합	
			A	B	C	
성별	남	관찰값	1	6	3	10
		기대값	2.5	5.5	2.0	
	여	관찰값	4	5	1	10
		기대값	2.5	5.5	2.0	
총합			5	11	4	20

성별과 성적의 교차표

<표 14.3> 교차표 분석의 예

표 14.3 에서 남학생과 여학생 사이의 뚜렷한 성적 분포의 패턴을 볼 수 있을지라도 이 패턴이 정말 ‘통계적’으로 유의미하다고 할 수 있을까? 다시 말해 위의 빈도수치는 우연으로부터 기대된 것들과 차이가 있지 않을까? 이 문제에 답하기 위해서는 우리는 2×3 교차표의 각각 셀 안에서 관측치의 기대 값을 계산해야만 한다. 이 계산은 셀의 주변합계와 행의 주변합계를 곱하고 관측치의 총 수로 나누면 된다. 예를 들어 남학생/A 학점 셀에서의 기대 값은 $5 * 10 / 20 = 2.5$ 가 된다. 다시

말하면, 우리는 2.5 명의 남학생이 A 학점을 받을 것이라고 기대하지만 실제로 오직 한 명의 남학생만이 A 학점을 받았다. 기대 값과 실제 값의 차이가 유의한지 아닌지는 카이제곱검정(chi-square test)을 통해 검증될 수 있다. 카이제곱통계량은 전체 셀에 걸쳐진 관측 값과 기대 값 사이의 평균차이로 계산된다. 그리고 카이제곱검정으로 계산된 값과 유의수준($p < 0.05$)을 비교한다. m 과 n 이 행과 열의 수라고 할 때 자유도는 $(m-1)*(n-1)$ 이므로 이 예시에서는 $df = (2-1) * (3-1) = 2$ 이 된다. 어느 통계 책에나 있는 표준카이제곱 표에서 $p=0.05$ 와 $df=2$ 에서의 카이제곱 기각 값은 5.99 이다. 관측된 데이터를 바탕으로 계산된 카이제곱 값은 기각 값보다 작은 1.00 이다. 따라서 관측된 학점패턴이 우연에 의해 기대된 패턴과 통계적으로 다르지 않다고 결론지을 수 있다.

제 15 장 양적 분석-추론적 통계

추론적 통계는 변수간 관계에 대한 결론을 이끌어 내기 위해 활용되는 통계적인 절차이다. 가설을 검정하기 위해 명료하게 설계되었다는 점에서 기술적 통계와 다른 점을 가지고 있다. 수량적인 통계분석은 대부분 SPSS 와 SAS 와 같은 통계 소프트웨어에서 지원된다. 이번 장은 가장 기초적이고 다소 짧은 지침서가 될 것이다. 깊이 있는 공부를 위하여 통계학에 대한 교과서를 찾아 보거나 통계학 수업을 수강하는 것을 추천한다.

15.1 기본 개념

영국의 철학자 칼 포퍼(Karl Popper)는 이론은 증명할 수 없으며 오직 반증할 수 있다고 했다. 예를 들어, 내일 해가 뜰 것이라는 것을 어떻게 증명할 것인가? 포퍼(Popper)는 태양이 매일 뜨는 것을 우리가 기억하기 때문에 내일도 해가 반드시 뜰 것이라는 것을 의미할 수는 없다고 했다. 왜냐하면 귀납적으로 얻어진 이론은 미래의 현상을 예언할 수 있을지 없을지를 추측만 할 수 있기 때문이다. 실제로 우리는 아무런 증명 없이 매일 태양이 뜰 것이라는 이론을 가정한다. 만약, 특정한 날 태양이 뜨지 않는다면, 이론에 오류가 있거나 반증될 수 있다고 제안했다. 이와 같이, 반대되는 증거에 기초하여 가설을 반증할 수 있지만, 진정으로 받아들이지는 못할 것이다. 왜냐하면, 증거가 존재한다는 것은 반대되는 증거가 추후에 관찰되지 않을 것이라는 것을 의미하지는 않기 때문이다. 대립 가설을 진정으로 수용하지 못하는 것은 우리가 귀무가설을 대립가설의 반대로 만들고, 실증분석을 통해 귀무가설을 기각하여 대립가설을 받아들일 수 있도록 만들었기 때문이다. .

사회과학 연구에서 가설 관계를 검정할 때의 두 번째 문제는 종속변수가 무한 개의 외생변수에 의해 영향을 받을 수 있고 이러한 모든 외생변수를 통제하는 것은 쉽지 않다는 것이다. 그렇기 때문에, 관찰된 표본에서 두 변수가 서로 관계가 있어 보일지라도, 모집단 안에서 그들의 관계가 존재하지 않을 수 있다는 것이다. 따라서, 추론적인 통계는 확실하거나 결정론적인 것이 아니라 확률적인 것이다.

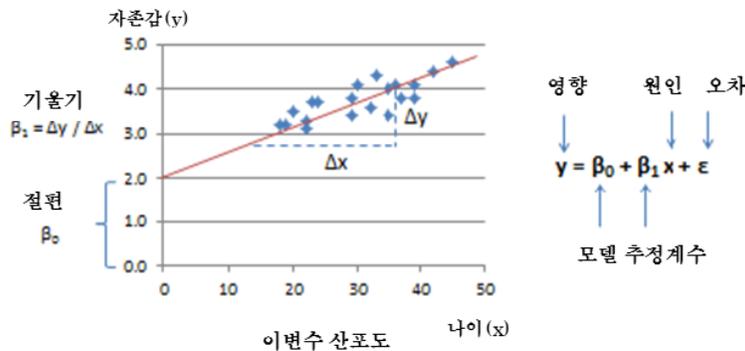
관찰된 표본에서 두 변수의 관계가 의미있다는 것을 어떻게 알 수 있을까? 그리고 우연히 일어난 것이 아니라는 것을 어떻게 알 수 있을까? 레놀드 피셔(Ronald A Fisher)는 역사적으로 매우 유명한 통계학자로, 유의성 검정을 위한 기초 가이드라인을 확립했다. 그는 통계적 결과가 기각될 확률이 5%이거나 그 이하일 때 유의미하게 여겨진다고 말하였다. 추론적 통계에서 이 확률을 p 값(p-value)라고 유의수준(α)라고 한다. P 값(p-value)과 유의수준(α) 사이의 기대되는 관계는 $p \leq 0.05$ 로 표시된다. **유의수준(significance level)**은 표본에서 얻어낸 추론을 일반화시키기 위해 지볼해야 하는 최대위험수준이라고 정의한다. P 값(p-value)이 0.05 혹은 5% 이하일 경우, 귀무가설을 기각하고 Type

I 에러의 범위 안에서 5% 정도의 정확하지 않은 확률을 의미한다. $p > 0.05$ 일 경우, 귀무가설을 기각하거나 대립가설을 승인하는 충분한 근거를 가지고 있지 못하다는 것을 나타낸다.

우리는 세 가지 관련된 통계 개념인 표본분포, 표준오차, 신뢰구간에 대해 이해해야 한다. **표본분포(sampling distribution)**는 모집단으로부터 무한의 표본들로 이루어진 이론적 분포를 의미한다. 하지만 표본은 모집단과 동일하지 않기 때문에, 모든 표본은 내재된 수준오차를 항상 가지고 있다. 이를 **표준오차(standard error)**라고 부른다. 만약 표준오차가 작다면, 표본으로 얻어진 통계 추정치(표본 평균값)는 모집단의 좋은 추정치라고 할 수 있다. 표본 추정치의 정확성은 **신뢰구간(confidence interval: CI)**이라고 정의된다. 95% CI 는 평균값 추정치에서 플러스 마이너스 두 개의 표준편차 안에 있는 것으로 정의된다. 그렇기 때문에, 관찰된 표본 추정치가 95% 수준의 CI 를 갖고 있다면, 모수가 관찰된 표본 추정치의 두 개의 표준편차 안에 95%의 확률로 존재한다고 할 수 있다. P 값(p-value)과 신뢰구간(CI)은 모두 표본이 얼마나 모집단에 근접한지 알 수 있게 해주는 지표이다.

15.2 일반적인 선형 모델

사회과학 연구에서 대부분의 추론적 통계 절차는 **일반선형모델(general linear model; GLM)**이라 불리는 일반통계모델로부터 파생된다. 모델(model)은 데이터의 집합을 활용해서 나타낼 수 있는 추정된 수학적 등식이다. 선형(linear)은 일직선을 의미한다. 따라서 GLM 은 관찰된 데이터의 관계를 선형 패턴으로 나타내 주는 등식이다.



<그림 15.1> 이변수 선형모델

GLM 의 가장 간단한 형태는 두 변수의 선형모델이다. 이것은 하나의 독립변수(원인 혹은 예측 변수)와 하나의 종속변수(효과 혹은 결과변수) 사이의 관계를 검증하는 선형모델이다. 두 변수를 각각 나이와 자존감이라고 가정해보자. 이 관계에 대한 이변수 산포도는 그림 15.1 과 같이 보여지며, 나이(예측 변수)는 수평축(X 축)을 따라 있고 자존감(결과)는 수직축(Y 축)을 따라 존재한다. 산포도에서 개별적인 관찰 값은 나이와 자존감의 결합물로서 선형의 모습을 띤다. GLM 의 기울기와 절편으로 이 지표에 대한 선을 추정할 수 있다. 대수학에서 일직선은 수학적 등식 $y = mx + c$ 를 활용하여 나타내는 것을 보았을 것이다. 이 등식에서 m 은 직선의 기울기(slope: x 단위가 변화할 때

y 가 얼마나 변화하는가)이고, c 는 절편(intercept: x 가 0 일 때, y 의 값)이다. 이 등식을 나타내면 다음과 같다.

$$y = \beta_0 + \beta_1x + \varepsilon$$

β_0 는 기울기, β_1 는 절편, ε 는 오류 값(error term)이다. ε 는 추정치와 실제 관찰 값의 편차이다. 대부분의 관찰이 선에 근접해 있지만 선과 일치하지는 않는다. 그 이유는 GLM 은 완벽하지 않기 때문이다. 두 개 이상의 독립(예측) 변수를 가지고 있는 선형 모델의 경우 시각적으로 삼차원을 생각해야 한다. 결과(y)가 수직축을 따라 형성되고, 두 예측 변수(x1, x2)가 두 수평축을 따르고 있기 때문이다. 둘 이상의 변수들 간의 관계를 나타낸 선을 회귀선(regression line)이라고 하고, β_0 과 β_1 는 회귀계수(regression coefficient)라고 불리고, 회귀계수를 추정하는 과정을 **회귀분석(regression analysis)**이라 한다. 회귀분석을 위한 GLM 은 n 개의 예측변수를 가지고 있다.

$$y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_3 + \dots + \beta_nx_n + \varepsilon$$

위의 등식에서, 예측변수 x_1 는 독립변수 혹은 공변량(covariates)/통제변수(control variables)를 나타낸다. 공변량은 이론적으로 관 심있는 변수가 아니라 종속변수 y 에 영향을 미칠 수 있는 변수로, 독립변수의 잔여효과를 더 정확하게 발견하기 위해 통제되어야 하는 것이다. 회귀 등식에서 공변량은 시스템적인 오류를 찾아내고 오차항(ε)은 임의적인 오차를 찾아 낸다. GLM 의 대부분의 변수들이 구간 혹은 비율척도이지만, 이는 일반적인 경우는 아니다. 어떤 예측변수는 명목변수(예: 성별: 남자 혹은 여자)일 수 있는데, 가변수(dummy variables)로서 코드가 부여될 수 있다. 가변수는 오직 0 혹은 1 로 두 변수 중 한 개로 나타날 수 있다고 가정한 다(성별의 예에서, 남자: 0, 여자: 1). 명목변수의 집합은 n-1 의 가변수를 활용하여 나타낸다. 예를 들어, 농업, 제조업, 서비스 부문을 포함하는 산업부문을 두 가변수(x1, x2)의 조합을 활용하여 나타낼 수 있는데, 농업은 (0,0)으로 제조업은 (0,1)로 서비스는 (1,1)로 표현할 수 있다. 어떤 명목변수가 0 또는 1 로 코드가 부여되는 것은 중요하지 않다. 0 과 1 의 값은 단지 두 집단으로 구별하기 위해 처리(실험연구에서 처리집단과 통제집단을 구별하는 것과 같이)되는 것이지 수적인 양을 의미하는 것이 아니며, 추정된 두 집단을 통계학적 지표를 나타내는 것일 뿐이다.

GLM 은 매우 강력한 통계 도구이다. 단일통계방법이 아닌 많은 예측변수 및 결과변수를 가진 복잡한 분석을 할 때도 활용될 수 있는 방법들의 집합체이기 때문이다. 만약 가예측 변수가 있고, 결과 변수에 대한 가변수의 두 레벨(0 과 1)의 효과를 비교한다고 한다면, 우리는 분산분석(analysis of variance; ANOVA)을 하게 된다. 하나 혹은 그 이상의 공변량의 효과를 통제하는 동안 ANOVA 분석을 하게 된다면, 우리는 공분산분석(analysis of covariance; ANCOVA)을 하게 되는 것이다. 우리는 다수의 결과 변수(예: y_1, y_2, \dots, y_n)를 가질 수 있는데, 이 것은 각각의 결과변수(각각은 그것들만의 상관 계수를 가짐)에 대한 등식을 ‘연립 방정식’을 활용하여 나타낼 수 있다. 만약 다수의 결과 변수가 동일한 예측변수의 집합으로부터 모델링 되었다면, 결과를 도출하기 위하여 다변량 회귀(multivariate regression)분석을 사용한다. 만약 다수의 결과변수에 대해 ANOVA 혹은 ANCOVA 분석을 실시하게 된다면, 결과를 도출하기 위한 분석은 *다변량 ANOVA(multivariate analysis of variance; MANOVA)* 혹은 *다변량 ANCOVA(multivariate analysis of covariance; MANCOVA)*라고 한다. 어떤 회귀등식에서의 결과를 다른 밀접한 관계가 있는 연립 회귀 등식에서의 예측변수로

모델링 하면 구조방정식모델링(structural equation modeling)이라 불리는 복잡한 형태의 분석방법을 가지게 된다. GLM 의 가장 중요한 문제는 모델의 구체화(model specification)이다. 예를 들어, 관심 있는 현상을 가장 잘 나타내기 위해 어떻게 회기 등식을 구체화 할 것인가(혹은 연립 등식)이다. 모델의 구체화는 관찰된 데이터가 모델에 가장 잘 들어 맞게 하는 것이 무엇인가 보다는 현상에 대해 선행 연구된 이론적인 고찰에 기초해야 한다. 데이터의 역할은 모델을 검정하는 것이지 구체화 하는 것에 있지 않기 때문이다.

15.3 두 집단의 비교

가장 간단한 추론적 분석은 통제집단을 설계한 임의화된 사후검증에서 처리집단의 결과와 통제집단의 결과를 비교하는 것이다. 전통적인 수학 교과과정을 선택한 수강생들의 성적과 특별한 수학 교육과정의 수강생들을 비교함으로써 더 좋은 성적을 기대할 수 있는 집단이 어디인지 분석한다고 가정해 보자. 위의 예에서 예측변수는 가변수(1=처리집단, 0=통제집단)로 설정하고, 결과변수인 성적은 비율 척도화 된다(예: 수학 시험 성적은 특별한 프로그램을 따라 결정된다). 이 간단한 설계를 위한 분석기법은 오직 하나의 예측 변수를 포함하기 때문에 *일원변량분석(one-way ANOVA)*이라고 하며 스튜던트 t 검정(Student's t-test)으로 통계적 검정을 수행한다.

t 검정은 1908 년에 더블린(Dublin)에 있는 기네스(Guinness)에서 일하는 윌리엄 셸리 고셋(William Sealy Gosset)이란 화학자가 아일랜드에서의 19 세기 런던에서 유행한 흑맥주인 스타우트(Stout)의 품질을 모니터 하기 위해 활용되면서 소개 되었다. 그의 직원들은 품질을 통제하기 위해 통계학을 활용한다는 사실을 드러내고 싶지 않았기 때문에 고셋(Gosset)은 '스튜던트(Student)' 를 이용하여 t 값을 계산하는 것을 포함하였다. 여기에서 t 는 fisher 가 두 집단 사이에 차이를 보여주기 위해 자주 활용했던 방법이었다. 이러한 이유로 스튜던트 t 검정이라는 이름을 가지게 되었으며 이후 동료 통계학자들에게 알려졌다.

t 검정은 두 집단의 평균 값이 통계학적으로 서로 다른지(양측 검정) 혹은 한 집단이 다른 집단보다 통계학적으로 더 큰(혹은 작은) 평균값을 가졌는지(단측 검정) 알아보기 위한 검정이다. 위의 예제에서, 특별 수학 교과과정을 듣는 수강생이 전통적인 교과과정의 수업을 듣는 수강생보다 성적이 좋은지 검정하기 위해 단측 검정을 실시한다. 이 가설은 다음과 같다

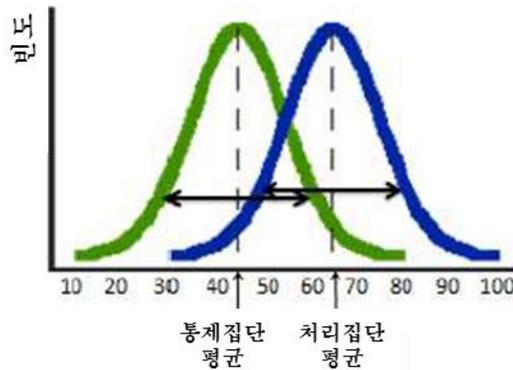
$$H_0: u_1 \leq u_2 \quad (\text{귀무가설})$$

$$H_1: u_1 > u_2 \quad (\text{대안가설})$$

u_1 은 특별 교과과정을 수강하는 학생들(처리집단)의 모평균, u_2 는 전통적 교과과정을 수강하는 학생들(통제집단)의 모평균을 나타낸다. 귀무가설은 항상 '등호' 표시가 있어야 함을 주의해야 하며, 통계적인 유의성 검정의 목표는 귀무가설을 기각하는데 있음을 명심해야 한다.

각 모집단에서 추출된 표본으로부터 어떻게 하면 데이터를 이용하여 모집단의 평균 차이를 추론할 수 있을까? 그림 15.2 의 처리집단과 통제집단의 빈도분석에서 통제집단의 평균값이 45 (0-100 척도에서), 처리집단의 평균값이 65 로 종형(정규)분포를 나타내는 것을 알 수 있다. 이 평균들은 서로 다르게 보이지만 모집단 평균과는(μ) 다른 표본평균(\bar{X})이다. 이것은 표본오류 때문에 발생하게

된다. 표본평균은 특정 신뢰구간에서 모집단의 평균이 존재한다는 것을 확률적으로 보여준 것이다(95%의 신뢰구간은 표본평균의 \pm 두 개의 표준오차이며, 표준오차는 모집단의 수 많은 표본로부터 추출된 표준편차의 분포를 나타낸다). 그러므로 모집단의 평균이 통계적으로 유의성을 갖기 위해서는 표본평균 값뿐만 아니라 표준오차, 표본평균의 빈도분포에서 분산 정도에 의해 결정된다. 만약 펼쳐진 분산 정도가 크다면(즉, 두 종형곡선이 아주 많이 겹친다면), 두 평균의 95%의 신뢰구간 역시 겹치게 된다. 따라서 두 모집단의 평균은 다르다는 것을 ($p < 0.05$)의 범위에서 결론을 내릴 수 없게 된다. 하지만 만약 곡선이 좁게 펼쳐져 있다면(즉, 덜 겹쳐져 있다면) 두 모집단의 평균값이 가지고 있는 신뢰구간은 겹치지 않을 것이기 때문에 통계적으로 귀무가설을 기각하고 두 집단의 모집단 평균은 $p < 0.05$ 에서 유의미하게 다르다고 말할 수 있다.



<그림 15.2> 수강생들의 t 검정

t-검정을 하기 위해서는 반드시 두 표본 집단간의 표본평균이 차이가 있다는 것을 t-통계량으로 먼저 계산해야 한다. 이 통계량은 두 개의 표본평균 차이 값과 두 개의 표준오차 차이 값의 상대적 비율이다. 이것을 수식으로 나타내면 분자는 처리집단(집단 1)과 통제집단(집단 2)간에 표본평균의 차이를, 분모는 두 집단 간의 표준오차의 차이를 가리키며 이를 나타내면 다음과 같다.

$$t = \frac{\bar{X}_1 - \bar{X}_2}{S_{\bar{x}_1 - \bar{x}_2}}, \quad S_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

s^2 는 분산이고 n 은 각 집단의 표본크기이다. 만약 처리평균이 통제평균보다 크다면 t-통계량은 양의 값으로 나올 것이다. t-통계량이 우연히 나올 수 있는 값보다 충분히 크지 아닌지 시험하기 위해서 반드시 SAS 나 SPSS 와 같은 통계 프로그램을 활용하여 계산하거나 인터넷이나 통계교과서에서 찾아 볼 수 있는 통계표의 계산된 t-통계량과 연관하여 확률 또는 p 값을 살펴봐야 한다. t 통계량의 값은 단측검정을 따르는지 양측검정을 따르는지, 확인하고 자유도(df) 계산을 통해 알 수 있다. t-통계량의 자유도는 다음과 같이 계산한다.

$$df = \frac{(s_1^2 / n_1 + s_2^2 / n_2)^2}{(s_1^2 / n_1)^2 / (n_1 - 1) + (s_2^2 / n_2)^2 / (n_2 - 1)}$$

보통 자유도(df)는 $\langle n_1 + n_2 - 2 \rangle$ 과 거의 가깝다. 만약 p 값이 원하는 유의수준($\alpha = 0.05$) 이나 최고수준보다 작다면, 실제로 아무것도 없을 때 처리효과가 있다고 결론 지을 수 있으며(1 종 오류) 귀무 가설을 기각할 수 있다.

처리집단이 통제집단보다 큰 평균값을 가지는지를 확인한 후, 다음은 처리집단이 통제집단과 비교하여 얼마나 효과(effect size; ES)를 갖는지 알아보는 것이다. GLM 에서 결과변수(y), 더미변수로 코드화한 처리변수를 독립변수(x)로 설정하고 회귀분석을 통하여 효과크기를 측정할 수 있다. 처리변수의 회귀계수(β_1)는 회귀직선의 기울기($\beta_1 = \Delta y / \Delta x$) 이며 효과크기의 추정치이기도 하다. 위의 예처럼, 두 변수(0 과 1)로 설정한 가변수 x 는 임의로 $\Delta x = 1 - 0 = 1$ 이기 때문에 효과크기 혹은 β_1 은 간단하게 처리평균과 통제평균의 차이($\Delta y = y_1 - y_2$)가 된다.

15.4 요인설계

앞선 예시를 확장해 교육시간(3 또는 6 시간/주)에 따른 전통적 교육과정(통제)과 특별한 교육과정(처리)의 효과에 대해 알아보자. 교육과정의 종류(전통적/특별한)와 교육시간(3 또는 6 시간/주)의 두 요인으로 2x2 요인설계를 한다. 여기에서 주효과(main effect)라고 불리는 각 요소의 독립효과뿐만 아니라 상호작용효과(interaction effect)라고 하는 두 요소간의 결합확률을 측정한다. 이원요인분석설계(two-way factorial design)를 위한 선형모델 등식은 다음과 같이 설계된다.

$$y = \beta_0 + \beta_1 + \beta_2 x_2 + \beta_3 x_1 x_2 + \varepsilon$$

y 는 수강생들의 후처리 수행점수이고 x_1 은 처리(특별한 교육과정 vs 전통적 교육과정), x_2 은 교육시간(3 또는 6 시간/주)을 나타낸다. x_1 과 x_2 둘 모두 가변수이며 x_2 가 비울척도(3 또는 6)로보이지만 요인설계에서는 두 개의 범주로써 나타내진다. 회귀계수 β_1 와 β_2 는 주효과의 크기를 나타내고 β_3 와 상호작용 효과의 크기를 각각 나타낸다. 또한, 요인설계 모델은 이원분산분석(two-way ANOVA analysis)을 이용하여 분석된다. 다수의 예측변수를 회귀분석은 다중회귀(multiple regression)이지만, 다수의 결과변수를 사용하는 회귀분석은 다변량회귀(multivariate regression)라고 한다.

상호작용 효과를 설명하고자 할 때, 만약 β_3 가 유의하다면 수강생의 점수는 교육시간에 따른 교육과정의 종류에 따라 처리효과가 있다는 것으로 해석한다. 이 경우에 교육과정(β_1)과 교육시간(β_2)의 처리효과에 대해 독립적으로 해석하면 안 된다. 그 이유는 두 처리효과가 각 개별 요인으로부터 분리될 수가 없기 때문이다. 만약, 독립적으로 주효과를 해석하고자 한다면 상호작용효과가 무의미할 때에만 가능하다.

새로운 변수로서 공변량(covariate)도 새로운 회귀 계수(예, β_4)로 요인설계에 포함될 수 있다. 공변량은 구간척도나 비울척도를 통해 측정될 수 있으며 예측변수가 가변수로 설계되었을 때도 가능하다. 공변량에 대한 설명은 다른 예측변수들의 규칙과 동일한 규칙을 따른다.

15.5 기타 양적 분석

이 밖에도 GLM의 변화량을 기반으로 한 유용한 추리적 통계 기술들이 있다. 아래의 통계 기술에 흥미가 있는 독자들은 고급통계 교과서나 고급통계 강의를 참고하는 것을 추천한다.

- *요인분석(factor analysis)*은 관측된 측정항목(item)을 이변량 상관패턴을 바탕으로 요인(factor)이라고 불리는 잠재변수(latent)의 작은 집합으로 통계기법을 활용하여 종합하는데 그 기법이 데이터 감소기술이다. 요인분석은 사회과학 연구의 다중항목 측정에서 수렴 타당도와 판별 타당도를 평가하기 위해 널리 사용된다.
- *판별분석(discriminant analysis)*은 예측변수의 선형결합을 기반으로 하는 명목범주들 가운데 하나로 주어진 관측치를 분류하는 기술이다. 이 기술은 종속변수가 명목변수라는 것만을 제외하고는 다중회귀와 유사하다. 특히 소비자나 제품을 핵심속성을 바탕으로 분류하는 마케팅 분야에서 많이 적용 된다.
- *로지스틱 회귀(logistic regression)*는 종속변수가 범주화로 표현되고 로지스틱 분포를 따른다고 가정하는 GLM이다. 그리고 로지스틱 회귀분석의 목표는 데이터를 로지스틱 곡선에 맞춤으로서 성공적인 결과에 대한 확률을 예측하는 것이다. 하나의 예로 나이, 신체용적지수, 운동식이요법 등의 예측변수를 기반으로 특정기간 동안에 심장마비에 걸릴 확률을 예측하는 것이다. 로지스틱 회귀는 거의 의학분야에서 사용된다. 효과크기 측정은 한 집단 대 다른 집단에서 일어나는 ‘오즈비(odds ratio)’를 바탕으로 한다.
- *프로빗 회귀(probit regression)*은 종속변수가 0과 1(이산값으로 가정할 수 있는 0과 1)사이에서 달라질 수 있는 GLM이며, 표준정규분포를 따른다고 가정한다. 이 회귀분석의 목표는 각 결과에 대한 확률을 예측하는 것이다. 프로빗 회귀는 보험계리과학, 금융서비스, 보험, 개인의 신용등급, 급여, 채무와 대출과 관련된 정보들을 바탕으로 한 신용을 다루는 산업들의 예측분석을 하는데 이용된다. 프로빗 회귀와 로지스틱 회귀는 비슷한 이진결과로서 유사하게 회귀계수를 설명하는 경향이 있다. 하지만 로지스틱 회귀가 상대적으로 계산하고 설명하기가 좀 더 용이하다.
- *경로분석(path analysis)*은 변수집합들 사이의 방향성 관계를 분석하기 위한 다변량 GLM 기술이다. 하나의 등식에서 종속변수가 다른 등식의 독립변수가 되는 복잡한 모델을 설명하는 것이 가능하다. 또한 현대 사회과학 연구에 널리 사용되고 있다.
- *시계열 분석(time series analysis)*은 시간에 따라 지속적으로 변하는 시계열 데이터 또는 변수들을 분석하는 기술이다. 적용되는 예로써 주식 시가변동이나 도시 범죄율을 예측하는 것을 들 수 있다. 이 기술을 경제학이나 금융, 디지털신호처리와 같은 분야에서 널리 사용된다. 특히, 이 기술은 시간에 관계없이 동일한 변수를 측정된 값들 안에서 자기상관관계 또는 상관관계를 찾아내는데 사용된다.