

정치와 데이터분석

강원대학교 춘천캠퍼스

박상훈

2025년 2학기 (2025-07-30 부)

E-mail: sh.park.poli@gmail.com

Office Hours: 목. 13:30-16:30; [예약링크](#)

Office: 영상바이오관 405호

Web: [e-루리: 정치와데이터분석](#)

Class Hours: 화. 10:00 - 13:00

Class Room: 사회과학대학1호관 109호(전산실)

강의 개요

본 강의는 2025년 2학기 강원대학교 춘천캠퍼스에서 진행되는 정치와 데이터분석으로, 인문사회 데이터를 분석하고 가설을 검증하는 방법을 익히는 것을 목표로 한다. 이를 위해 기초 통계 이론 강의와 R 실습을 병행하여 진행한다. 수업을 통해 학생들은 정량적 데이터 분석 기법의 이론과 실제에 대한 탄탄한 기초를 다지게 될 것이다. 학생들이 스스로 연구를 설계하고 데이터를 분석하여, 최종적으로 국내 학술지 게재를 목표로 자신의 논문을 완성할 수 있는 능력을 배양하는 데 그 목적이 있다. 강의를 통해 학생들은 변수의 분포를 요약하고, 변수 간의 관계를 추정하며, 통계 그래픽을 통해 분석 결과를 시각적으로 표현하는 방법을 학습할 것이다. 이 수업은 통계나 R 프로그램을 전혀 다뤄본 경험이 없는 학생들을 위해 설계되었다. 정치학연구방법론과 달리 경험적인 자료와 기법들을 중심으로 강의가 이루어지며, 사칙연산과 논리적 사고 능력만 갖추었다면 누구나 쉽게 따라올 수 있도록 매우 기초적인 내용부터 차근차근 강의할 것이다. 다만 수업 내용의 소화를 위해 매주 1~2시간의 복습이 요구된다.

교재 및 참고문헌

Moderndive 에 기초한 슬라이드 및 예제를 활용하여 R을 이용한 실습을 진행한다.

- Chester Ismay, Albert Y. Kim 저. 양승훈 역. 2023. 데이터 과학을 활용한 통계: R Tidyverse 패키지로 손쉽게 익히는 통계학 입문. 자유아카데미.
 - 원서는 오픈북으로 제공됨: <https://moderndive.com/> (이후 **Moderndive**)
- Sean Gailmard. 2014. *Statistical Modeling and Inference for Social Science*. New York: Cambridge University Press (이후 **Gailmard**)
- Paul M. Kellstedt and Guy D. Whitten. 2018. *The Fundamentals of Political Science Research, 3rd ed.* New York: Cambridge University Press (이후 **Kellstedt and Whitten**)

해당 주차가 마무리된 후, 강의 슬라이드가 e-루리 게시판과 강의 홈페이지(https://pherephobia.github.io/TEACHING_POLANAL)을 통해 업로드될 것이며, 필요한 추가 읽기자료 역시 e-루리 게시판에 업로드될 것이다. 강의는

Modernrive 교재를 중심으로 이론과 실습에 대한 내용을 재구성한 슬라이드를 사용한다.

이 강의에서의 예제와 과제, 시험 대부분은 R과 RStudio를 사용하게 될 것이므로 아래의 프로그램을 다음의 온라인 강의를 참고하여 설치하기를 바란다.

- R: <https://www.r-project.org/>
- RStudio: <https://posit.co/download/rstudio-desktop/>
- Quarto: <https://quarto.org/>
- TeXLive: <https://www.tug.org/texlive/> (for MAC OS, MacTeX)
- 데이터 사이언스의 첫걸음, R 4.0.2 버전, Rtools40, 그리고 Rstudio 설치 (Windows 10) | 슬기로운 통계생활
 - MAC OS의 경우에는 Rtools는 별도로 설치할 필요가 없음.
- 한글 깨짐 방지를 위한 윈도우10 한글 설정 총정리 + RStudio 한글 환경설정 | 슬기로운 통계생활
- tinytex을 사용한 texlive 2020 간편 설치와 한글 사용 설정 | 슬기로운 통계생활

성적 평가

실제 수업시간의 1/4 이상 결석하는 경우 성적은 자동적으로 F가 부과되며, 기타 사항은 강원대학교 학사 규정에 따라 처리한다. 최종 성적은 강원대학교 학사 규정에 따라 수강인원이 19명 이하인 경우 절대평가로 이루어진다. 성적이 60점 미만이거나 과제물 및 시험에서의 부정행위가 확인될 경우, 최종 성적은 F로 처리된다. 절대평가 시 성적 평가 기준은 다음과 같다:

- A+: 94-100; A0: 90-93; B+: 84-89; B0: 80-83; C+: 74-79; C0: 70-73; D+: 64-69; D0: 60-63; F: 60점 미만

평가 항목

평가 항목	배점	비고
실습과제	40%	2025년 9월 9일(화)부터 12월 9일(화)까지 총 11회의 실습과제가 있다. R을 활용한 간단한 분석 및 시각화에 대한 문제를 풀고 다음 강의 전까지 제출하는 것으로 한다. 이 중 가장 성적이 낮은 하나의 과제를 최종 점수 반영에서 제외한다. 실습과제에 대한 자세한 내용은 아래 확인.
중간시험	25%	중간시험은 Closed Book In-Class 형식이며, 2025년 10월 28일(화) 강의시간에 진행한다. 자세한 내용은 별도 공지.
기말시험	25%	기말시험은 Closed Book In-Class 형식이며, 2025년 12월 일(화) 강의시간에 진행한다. 자세한 내용은 별도 공지.
출결	10%	학기 중 강의에 성실히 참여하는지에 따라 최종 점수의 10%를 부여.
합계	100%	출결은 담당교원의 호명에 의한 출결 처리를 원칙으로, 강원대학교 출결관리 (https://www.kangwon.ac.kr/www/contents.do?key=2069&)에 따름.
실습과제	40%	

실습과제

R을 활용한 기초 데이터 분석을 마치고 해당 주차에 학습한 내용을 응용하여 풀 수 있는 실습과제를 그 다음 주 강의 시간 전까지 제출한다. 지연제출 시 1일 당 20%의 패널티가 부여된다.

시험(중간시험 + 기말시험)

수강생들은 중간시험과 기말시험을 치르게 된다. 두 시험 모두 강의실에서 필기시험으로 진행하며 각각 10월 28일 화요일, 12월 16일 화요일에 진행한다. 각 시험은 최종 성적의 25%에 해당한다. 객관식, 약술형, 논술형 등으로 구성되며, 자세한 내용은 각 시험 전에 별도 공지한다.

연구 윤리 준수

AI 및 대규모 언어 모델(*Large Language Models, LLMs*) 사용에 대한 주의사항

본 강의에서는 ChatGPT를 비롯한 대규모 언어 모델의 사용을 권장하지 않는다. 물론 LLMs가 유용하게 쓰이는 경우가 있다. 다만 이러한 도구들은 세심한 기술과 충분한 연습이 필요하며, 특정한 상황에서만 효과적으로 기능할 수 있다.

비교의 방법을 통해 정치학적 질문에 답하도록 연구를 설계하는 것이 어려운 것은, 일련의 절차들을 통해 우리의 생각이 정리되고 발전하기 때문이다. AI에게 연구설계를 맡긴다면, 결국 사고 과정도 포기하는 것과 같다. 회의 일정을 잡는 이메일을 쓸 때는 큰 문제가 없을 수 있지만, 사업 계획서나 정책 제안서, 법률 문서를 작성할 때는 심각한 문제가 될 수 있다(Rosenzweig 2023). LLMs나 AI를 사용하여 요구된 분량을 채우고 과제를 제출할 수는 있겠지만, 그 내용은 실질적 의미가 없을 것이다. 철학적으로 이러한 글쓰기는 ‘헛된 것(bullshit)’이라고 정의된다(Hicks, Humphries, and Slater 2024). 철학적 의미의 ‘헛된 것’이란 “진실 여부를 고려하지 않은 채 생산된 발화나 텍스트”를 의미한다(Hicks, Humphries, and Slater 2024, 2). 이는 진실도 거짓도 아닌, 단지 저자가 해당 주제를 이해하고 있는 것처럼 보이게 하는 텍스트이다(Hicks, Humphries, and Slater 2024, 3).

AI 생성 콘텐츠 여부를 판별하는 데 특별한 시간을 할애하지는 않을 것이다. AI가 작성한 과제라 하더라도 자동으로 0점 처리하지는 않으며, 제출된 내용 자체의 가치를 평가할 것이다. 경험상 AI 생성 텍스트는 별도의 확인 없이도 대개 50% 이하의 평가를 받는다. LLMs가 생성하는 텍스트는 본질적으로 진실성이 결여되어 있어 우수한 평가를 받을 수 없기 때문이다.

표절 금지

- 심각한 표절 및 부정행위가 발견되는 경우 최종 성적으로 F가 부여된다.
- 타인의 연구결과를 적절한 인용 없이 자신의 연구에 사용한다면 이는 표절 행위에 해당한다. 출처와 인용을 정확하게 밝히더라도 반드시 필요한 경우가 아니라면 직접적인 인용 대신 자신의 목소리로 정제된 표현을 활용하여야 한다. 특히, 인터넷에 게재된 논문, 기사, 오피니언 등을 적절한 확인 절차와 인용 과정 없이 이용하는 것은 절대로 금한다.
- 일반적으로 통용되는 인용 스타일을 일관적으로 활용하여야 한다. *APSA Style Manual for Political Science*의 사용을 권장하지만, 그 외에 다른 인용 양식(APA, MLA, Chicago Style Manual etc.)을 사용하여도 무관하다. 단, 작성하는 문서 전체에 걸쳐 일관된 하나의 인용 양식을 사용하여야 한다.

강의 일정

9월						
	1	2	3	4	5	6
7	8	9	10	11	12	13
14	15	16	17	18	19	20
21	22	23	24	25	26	27
28	29	30				
일	월	화	수	목	금	토

10월						
			1	2	3	4
5	6	7	8	9	10	11
12	13	14	15	16	17	18
19	20	21	22	23	24	25
26	27	28	29	30	31	
일	월	화	수	목	금	토

11월						
						1
2	3	4	5	6	7	8
9	10	11	12	13	14	15
16	17	18	19	20	21	22
23	24	25	26	27	28	29
30						
일	월	화	수	목	금	토

12월						
	1	2	3	4	5	6
7	8	9	10	11	12	13
14	15	16	17	18	19	20
21	22	23				
일	월	화	수	목	금	토

강의일
 학기
 시험
 추석연휴

비고: 중간시험(10월 28일), 추석연휴(10월 3일-9일), 기말시험 제출(12월 16일)

강의진도 계획

강의 일정에 변동이 있는 경우 사전에 안내될 것이다. 주로 강의자가 수강자에게 학습 내용을 전달하는 강의형으로 진행된다. 📖: 교재, 📄: 학술논문, 🌐: 인터넷 자료, 📺: 영상자료.

제1주차 09/02 화요일: 강의 소개

강의 전반과 강의계획서에 대한 내용을 함께 검토하고, 강의에 필요한 자료를 어떻게 검색하는지 및 R, RStudio, Quarto 등 학기 간에 필요한 설치 프로그램과 실습과제, 중간시험, 기말시험에 대해 안내한다.

제2주차 09/09 화요일: 과학적 연구와 이론, 인과관계와 연구설계

- 📖 **Kellstedt and Whitten**, Chapters 1-4.
- 🌐 **Moderndive**, Chapter 1: <https://moderndive.com/1-getting-started.html>
- 💡 실습과제 1 (9월 8일 월요일 11:59 PM까지)

제3주차 09/16 화요일: 데이터와 정보 I: 개념의 측정

- 📖 **Kellstedt and Whitten**, Chapter 5.
- 📖 **Gailmard**, pp.14-21.

🔗 **Moderndive**, Chapter 3, 4: <https://moderndive.com/3-wrangling.html> <https://moderndive.com/4-tidy.html>

💡 실습과제 2 (9월 15일 월요일 11:59 PM까지)

제4주차 09/23 화요일: 데이터와 정보 II: 단변량 & 양변량 분포

📖 **Kellstedt and Whitten**, Chapter 6.

📖 **Gailmard**, pp. 21-37.

🔗 **Moderndive**, Chapter 2: <https://moderndive.com/2-viz.html>

💡 실습과제 3 (9월 22일 월요일 11:59 PM까지)

제5주차 09/30 화요일: 확률과 통계적 추론 I: 확률, 모집단과 표본

📖 **Kellstedt and Whitten**, pp.143-148.

📖 **Gailmard**, pp.83-115.

🔗 **Moderndive**, Chapter 7: <https://moderndive.com/7-sampling.html>

📺 찰떡 비유로 한 방에 이해 시키는 '조건'이 들어있는 '확률' I 조건부 확률 개념 때려잡기 I 정승제의 확률과 통계

📺 중심극한정리와 표본 평균의 분포의 관계 완벽 정리.zip

💡 실습과제 4 (9월 29일 월요일 11:59 PM까지)

제6주차 10/07 화요일: 추석 연휴로 인한 휴강

💡 실습과제 5 (10월 6일 월요일 11:59 PM까지)

제7주차 10/14 화요일: 확률과 통계적 추론 II: 정규분포와 신뢰구간

📖 **Gailmard**, pp.160-168, 187-203.

📖 **Kellstedt and Whitten**, pp.148-160.

🔗 **Moderndive**, Chapter 8: <https://moderndive.com/8-confidence-intervals.html>

💡 실습과제 6 (10월 13일 월요일 11:59 PM까지)

제8주차 10/21 화요일: 가설 검정 I

📖 **Gailmard**, pp.236-246.

📖 **Kellstedt and Whitten**, pp.161-166.

🔗 **Moderndive** , Chapter 9: <https://moderndive.com/9-hypothesis-testing.html>

💡 실습과제 7 (10월 20일 월요일 11:59 PM까지)

제9주차 10/28 화요일: 중간시험 (Closed Book In-Class)

- 강의실에서 필기시험으로 진행되며 최종 성적의 25%에 해당
- 객관식, 약술형, 논술형 등으로 구성

제10주차 11/04 화요일: 가설 검정 II

❓ 중간시험 리뷰

📖 **Kellstedt and Whitten** , pp.166-188.

제11주차 11/11 화요일: 단순회귀분석

📖 **Kellstedt and Whitten** , pp.188-215.

🔗 **Moderndive** , Chapter 5: <https://moderndive.com/5-regression.html>

🔗 **Moderndive** , Chapter 10.2: <https://moderndive.com/10-inference-for-regression.html#regression-interp>

💡 실습과제 8 (11월 10일 월요일 11:59 PM까지)

제12주차 11/18 화요일: 다중회귀분석 I

📖 **Kellstedt and Whitten** , Chapter 10.

🔗 **Moderndive** , Chapter 6: <https://moderndive.com/6-multiple-regression.html>

💡 실습과제 9 (11월 17일 월요일 11:59 PM까지)

제13주차 11/25 화요일: 다중회귀분석 II

📖 **Kellstedt and Whitten** , pp.246-258.

🔗 **Moderndive** , Chapter 6: <https://moderndive.com/6-multiple-regression.html>

💡 실습과제 10 (11월 24일 월요일 11:59 PM까지)

제14주차 12/02 화요일: 진단(Diagnostics)

🔗 [lindia](#)

- `lindia`는 선형 모델의 진단 플롯(diagnostic plots)을 간편하게 시각화할 수 있도록 `ggplot2`에 추가 기능을 제공하는 확장 패키지

- `lindia`의 모든 함수는 `lm` 객체(`lm()` 및 `glm()` 포함)를 입력받아, 선형 진단 플롯을 `ggplot` 객체 형태로 반환합니다. 다음 코드는 `Cars93` 데이터셋으로 단순 선형 모델을 생성한 후, `lindia::gg_diagnose()` 함수를 호출하여 분포의 전반적인 특징을 시각화하는 방법을 나타냅니다.

💡 실습과제 11 (12월 1일 월요일 11:59 PM까지)

제15주차 12/09 화요일: 비선형 관계, 로지스틱 회귀분석

🔗 [Binary Response Models from Brenton Kenkel's Practical Data Analysis for Political Scientists](#)

제16주차 12/16 화요일: 기말시험 (Closed Book In-Class)

- 강의실에서 필기시험으로 진행되며 최종 성적의 25%에 해당
- 객관식, 약술형, 논술형 등으로 구성