

Quiz 250414 (Wason Selection Task)

coop711

2025-04-14

Contents

1 소개	5
2 국민문해력조사 집계 결과	7
2.1 응답 집계	7
2.2 막대그래프	7
2.3 문해력 점수 계산	8
2.4 Red and Black 비교	9
2.5 문해력 등급 판정	14
2.6 유형별 정답률	15
2.7 어려운 문제?	15
2.8 정답률이 낮은 문제들	16
3 1주차 데이터 실험 집계	21
3.1 실험의 목적	21
3.2 Q1. Dewey as good as elected, statistics convince Roper . . .	23
3.3 Q2. Statistics is the science of learning from data,	24
3.4 Q3. How to lie with statistics	25
3.5 Q4. 종부세	26
3.6 마감 시간으로부터 제출 시간의 분포	29
4 2주차 데이터 실험 집계	33
4.1 실험의 목적	33
4.2 Q1. 춘추전국시대에 국가통계관리의 중요성 강조	36
4.3 Q2. 국가정책을 수립하는 데 통계의 역할	37

4.4 Q3. 우리나라 생산가능인구 감소 시기	38
4.5 Q4. 우리나라 총인구 최대 시기	39
4.6 Q5. 소멸위험 단계 개선 지역	40
4.7 Q6. 조출생률과 합계출산율	41
4.8 Q7. 눈속임 그래프(Cheating Charts)	42
4.9 마감 시간으로부터 제출 시간의 분포	46
5 3주차 데이터 실험 집계	51
5.1 실험의 목적	51
5.2 Q1. 국세와 지방세 비중	53
5.3 Q2. 조세부담률	54
5.4 Q3. OECD 국민부담률	55
5.5 Q4. 과세대상 근로소득 1,200만 원	56
5.6 Q5. 소득세 실효세율	57
5.7 Q6. 기업규모별 과세 현황	58
5.8 Q7. 국민부담률 적정 수준 : 아일랜드와 OECD 평균	60
5.9 마감 시간으로부터 제출 시간의 분포	63
6 4주차 데이터 실험 집계	67
6.1 실험의 목적	67
6.2 Q1. 세종대왕 시대 조세제도	69
6.3 Q2. 공법도입에 대한 대신들의 찬성을	70
6.4 Q3. 공법도입과 품관촌민들의 찬반	71
6.5 Q4. 공법	72
6.6 Q5. 1423년 조선시대 호구와 인구	73
6.7 Q6. 지방관료와 품관촌민	74
6.8 Q7. 부연설명의 효과 : 주당 근로 69시간제 도입 찬반	75
6.9 마감 시간으로부터 제출 시간의 분포	78

CONTENTS	5
7 5주차 데이터 실험 집계	83
7.1 실험의 목적	83
7.2 Q1. 한글의 문자 유형	85
7.3 Q2. 정보혁명과 문자 체계	86
7.4 Q3. 알기 힘든 전문 용어	87
7.5 Q4. 해방직후 비문해율	88
7.6 Q5. 세대간 문해력 격차	89
7.7 Q6. 문해력 격차의 파급효과	90
7.8 Q7. 프레임을 설정하는 단어의 힘	91
7.9 마감 시간으로부터 제출 시간의 분포	94
8 6주차 데이터 실험 집계	99
8.1 실험의 목적	99
8.2 Q1. 월간 독서율	101
8.3 Q2. 지역 및 지역크기별 가구수 비례 무작위추출법	102
8.4 Q3. 한달 독서량의 분포	103
8.5 Q4. 최근 1개월간 독서량	104
8.6 Q5. 20대의 연간독서율	105
8.7 Q6. 50대의 연간독서율	106
8.8 Q7. The more, the better? : 내가 남보다, 혹은 남이 나보다	107
8.9 마감 시간으로부터 제출 시간의 분포	110
9 7주차 데이터 실험 집계	115
9.1 실험의 목적	115
9.2 Q1. 통계학의 기본원리	117
9.3 Q2. 리터러리 다이제스트의 실패	118
9.4 Q3. 1948년, 여론조사가 듀이를 당선시킨 해	119
9.5 Q4. 1948 미 대선 이후	120
9.6 Q5. 표본오차를 반으로 줄이려면?	121
9.7 Q6. 대선 여론조사의 목표모집단?	122
9.8 Wason Selection Task	123
9.9 Q7. Red에 추상적 질문, Black에 구체적 질문	124

9.10 Q8. Red에 구체적 질문, Black에 추상적 질문	127
9.11 Q9. 인지적 편향과 오류	128
9.12 학습 순서의 영향	131
9.13 마감 시간으로부터 제출 시간의 분포	134

Chapter 1

소개

이 문서는 2025년 봄학기 동안 수행된 주요 퀴즈 결과 및 분석을 정리한 보고서입니다.

각 장에서는 개별 퀴즈의 응답 분포, 정답률, 인지편향 관찰 결과 등을 시각화하고 요약합니다.

분석 대상은 다음과 같습니다:

- Wason Selection Task 정답률 변화
- 프레이밍 효과에 따른 반응 차이
- Oxford Happiness Questionnaire 응답 패턴
- 언어별 응답지 차이 (한글 vs 영어)

R과 R Markdown을 기반으로 자동화된 시각화 및 분석을 포함합니다.

Chapter 2

국민문해력조사 집계 결과

2.1 응답 집계

제작자: 김민수

제작일: 2023-09-15

제작자: 김민수

제작일: 2023-09-15

제작자: 김민수

제작일: 2023-09-15

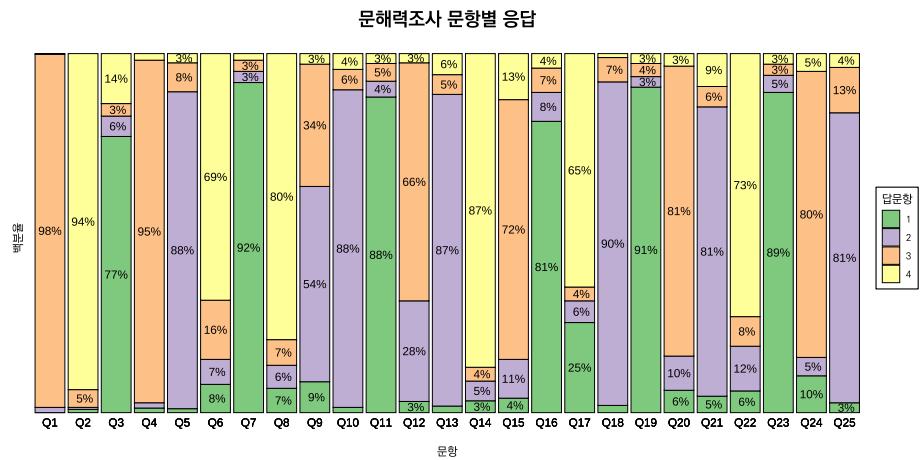
제작자: 김민수

2.2 막대그래프

막대그래프로 답안 분포를 시각적으로 살핀다. 차후 나오는 정답률과 함께 어느 문항에서 어느 답안을 많이 고르는지 파악하는 데 활용한다.

Table 2.1: Counts

Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14	Q15	Q16	Q17	Q18
0	1	77	1	1	8	92	7	9	1	88	3	2	3	4	81	25	0
1	1	6	1	88	7	3	6	54	88	4	28	87	5	11	8	6	0
98	5	3	95	8	16	3	7	34	6	5	66	5	4	72	7	4	0
0	94	14	2	3	69	2	80	3	4	3	3	6	87	13	4	65	0



2.3 문해력 점수 계산

2.3.1 정답과 대조하여 R(Right)/W(Wrong) 표시

Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14	Q15	Q16	Q17	Q18	Q19	Q20	Q21	Q22	Q23	Q24	Q25
R	R	R	R	R	R	W	W	W	W	W	R	R	R	R	R	R	R	R	R	R	R	R	R	
R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	W	R	R	R	R	R	R	
R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	W	R	R	R	R	R	R	
R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	
R	R	R	R	R	R	R	W	W	R	R	W	R	R	W	R	W	R	R	W	R	R	R	R	
R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	

2.3.2 학생별 점수 산출

80, 96, 96, 100, 76, 100, 88, 20, 92, 92, 88, 80, 100, 100, 80, 100, 92, 84, 100, 92, 84, 100, 92, 100, 96, 92, 84, 88, 48, 92, 96, 100, 80, 96, 92, 84, 92, 72, 96, 96, 96, 60, 100, 76, 76, 72, 80, 32, 64, 88, 92, 92, 92, 88, 96, 96, 48, 84, 72, 100, 88, 88, 92, 76, 96, 84, 92, 88, 96, 88, 100, 68, 80, 76, 84, 92, 92, 92, 80, 88, 84, 56, 72, 80, 84, 28, 100, 96, 92, 92, 88, 84, 88, 96, 84, 88, 88, 92, 88, 84, 76, 88, 88, 92, 44, 80, 84, 20, 80, 72, 92, 100, 96, 92, 88, 72, 88, 64, 56, 96, 80, 92, 84, 88, 76, 80, 84, 76, 88, 80, 96, 96, 84, 96, 88, 68, 92, 92, 88, 92, 28, 80, 44, 100, 84, 84, 96, 68, 100, 88, 84, 92, 92, 100, 96, 92, 96, 12, 92, 72, 92, 88, 100, 20, 80, 92, 100, 100, 100, 92, 100, 92, 92, 80, 84, 92, 88, 88, 100, 88, 88, 96, 92, 80, 100, 92, 88, 96, 84,

96, 96, 96, 80, 96, 96, 52, 80, 84, 68, 92, 84, 80, 48, 84, 80, 88, 84, 96,
 36, 88, 92, 88, 88, 60, 80, 84, 88, 92, 84, 64, 92, 92, 96, 88, 88, 88, 88,
 88, 36, 64, 84, 76, 80, 80, 24, 100, 88, 80, 88, 92, 64, 76, 80, 24, 80, 96,
 56, 64, 84, 88, 64, 84, 84, 96, 84, 92, 88, 80, 84, 24, 76, 48, 96, 88, 40,
 80, 84, 80, 88, 92, 92, 80, 96, 64, 96, 40, 92, 92, 88, 84, 100, 36, 80, 92,
 84, 84, 32, 96, 96, 96, 96, 96, 92, 92, 96, 84, 88, 92, 92, 88, 84, 80, 100,
 80, 64, 28, 92, 88, 80, 92, 100, 92, 88, 100, 84, 80, 56, 84, 16, 76, 96,
 80, 84, 68, 92, 88, 84, 88, 96, 88, 56, 36, 88, 88, 72, 76, 92, 96, 80, 92,
 100, 80, 48, 28, 36, 92, 88, 84, 68, 88, 36, 92, 88, 92, 72, 84, 96, 60, 88,
 96, 84, 92, 72, 100, 84, 88, 96, 80, 96, 72, 96, 80, 24, 88, 96, 36, 64, 80,
 88, 92, 92, 80, 96, 92, 88, 80, 96, 88, 80, 52, 92, 92, 44, 96, 44, 88, 100,
 88, 92, 96, 100, 92, 96, 88, 92, 96, 36, 92, 84, 88, 92, 56, 84, 72, 84, 88,
 96, 100, 96, 76, 76, 88, 88, 88, 96, 92, 28, 96, 92, 88, 84, 76, 100, 92,
 92, 88, 100, 84, 60, 84, 84, 84, 76, 100, 100, 92, 16, 84, 52, 84, 88, 88,
 96, 96, 88, 88, 80, 56, 92, 28, 76, 72, 96, 80, 84, 40, 88, 88, 80, 96, 76,
 96, 100, 32, 76, 88, 100, 100, 36, 92, 16, 84, 96, 92, 88, 52, 80, 84, 92,
 40, 96, 96, 32, 48, 60, 92, 96, 68, 48, 48, 88, 52, 92, 52, 88, 80, 24, 76,
 100, 88, 92, 80, 92, 88, 20, 92, 84, 52, 96, 96, 88, 52, 84, 40, 88, 48, 60,
 72, 48, 64, 60, 40, 92, 76 and 28

2.4 Red and Black 비교

2.4.1 Summary

- Red:

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
16	77	88	79.9	92	100

- Black:

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
12	80	88	82.2	92	100

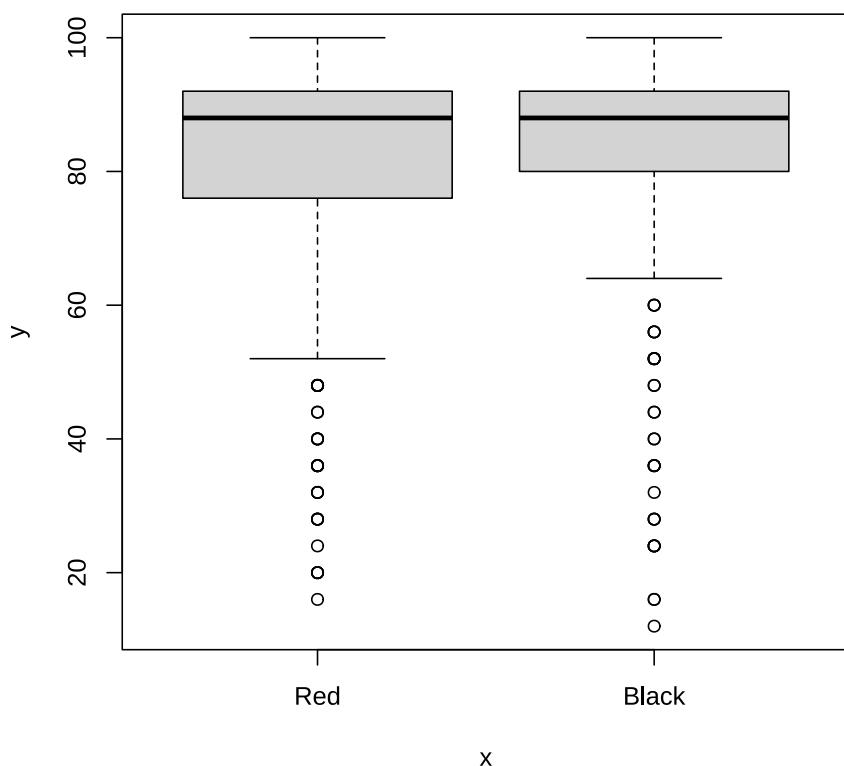
2.4.2 출기-잎 그림

- Red

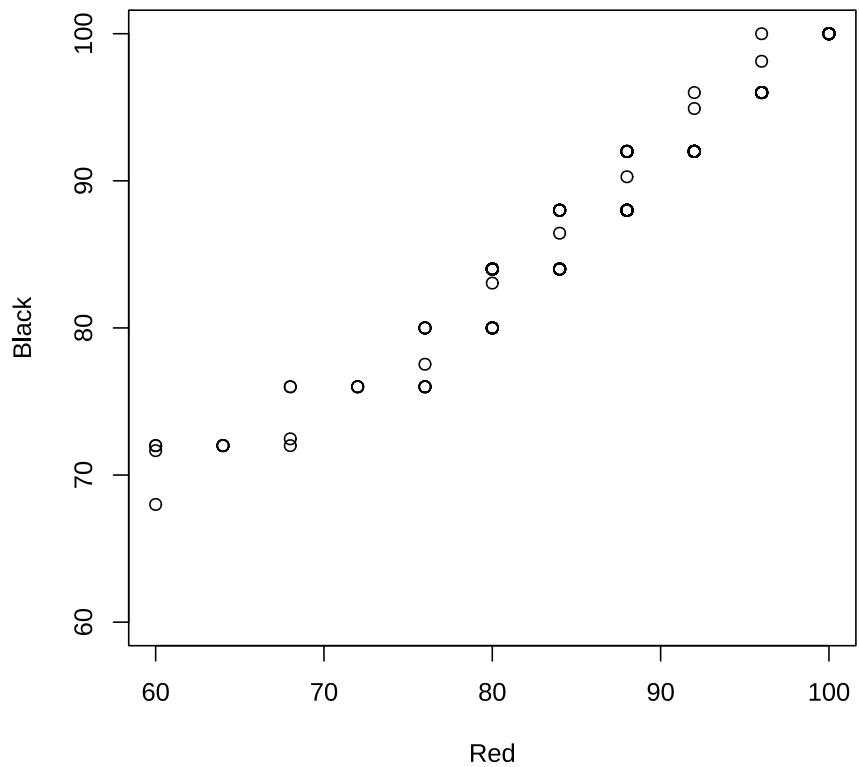
```
##  
## The decimal point is 1 digit(s) to the right of the |
```

- Black

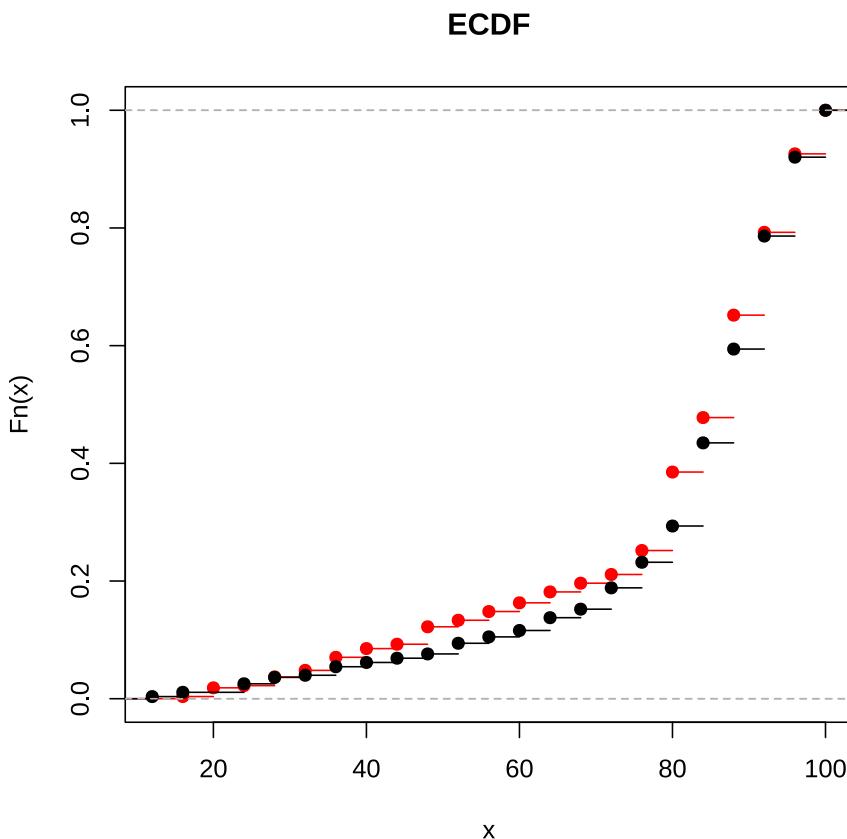
2.4.3 Box Plots



2.4.4 QQ plot



2.4.5 ECDF plot



2.4.6 t-test

Red 와 Black으로부터 관찰된 점수들의 평균에 대하여 t test를 적용하였더니
기초통계량에서 살핀 것을 뒷받침해 주듯이 p-value 가 0.05보다 큰 값이
나왔습니다. 평균은 닮았다는 뜻입니다.

Table 2.5: Welch Two Sample t-test: score by .\$group

Test statistic	df	P value	Alternative hypothesis	mean in group Red	mean in group Black
-1.433	538.2	0.1524	two.sided	79.93	82.23

Table 2.6: (#tab:literacy grade) 문해력 등급 분포

	I	II	III	IV	계
	13	41	55	437	546

Table 2.7: (#tab:literacy grade RnB) 그룹별 문해력 등급 분포

	I	II	III	IV	계
Red	6	27	24	213	270
Black	7	14	31	224	276
계	13	41	55	437	546

2.5 문해력 등급 판정

2.5.1 분포표

- I수준(24점 이하), II수준(28 ~ 48점), III수준(52 ~ 72점), IV수준(76점 이상)

```

\begin{table}
\caption{(#tab:literacy grade) 문해력 등급 분포(%)}

\begin{array}{|c|c|c|c|c|c|} \hline
 & I & II & III & IV & 계 \\ \hline
 2.38 & 7.51 & 10.07 & 80.04 & 100 \\ \hline
\end{array}
\end{table}

```

2.5.2 Red and Black

카이제곱 테스트로 Red와 Black에 들어간 등급별 인원수가 얼마나 닮았는지를 살펴보았지만 t-test와 마찬가지로 통계적으로 유의한 차이가 관찰되지 않았습니다.

Table 2.8: Pearson's Chi-squared test: .

Test statistic	df	P value
5.301	3	0.151

2.6 유형별 정답률

	유형	정답률(%)
문1	사실적	98.4
문2	사실적	93.6
문3	비판적	76.9
문4	사실적	95.4
문5	추론적	88.3
문6	추론적	68.7
문7	추론적	91.9
문8	사실적	79.7
문9	추론적	34.1
문10	추론적	88.5
문11	사실적	87.9
문12	사실적	66.3
문13	사실적	86.8
문14	추론적	87.4
문15	사실적	72.3
문16	사실적	81.1
문17	추론적	65.0
문18	비판적	90.1
문19	사실적	90.7
문20	사실적	80.8
문21	사실적	80.6
문22	비판적	73.3
문23	추론적	89.2
문24	사실적	79.7
문25	추론적	80.8

2.7 어려운 문제?

2.7.1 정답률 80% 이하

	문3	문6	문8	문9	문12	문15	문17	문22	문24
정답률	76.9	68.7	79.7	34.1	66.3	72.3	65	73.3	79.7

2.7.2 정답률 70% 이하

	문6	문9	문12	문17
정답률	68.7	34.1	66.3	65

2.7.3 정답률 60% 이하

	문9
정답률	34.1

2.7.4 정답률 50% 이하

	문9
정답률	34.1

2.8 정답률이 낮은 문제들

2.8.1 문6.

문 6) 다음 도로교통법의 내용을 바르게 이해한 사람은 누구입니까?

제44조 (술에 취한 상태에서의 운전 금지)
 ① 누구든지 술에 취한 상태에서 자동차 등을 운전하여서는 아니 된다.
 ② 경찰공무원은 교통의 안전과 위험 방지를 위하여 필요하다고 인정하거나 제1항의 규정을 위반하여 술에 취한 상태에서 자동차 등을 운전하였다고 인정할 만한 상당한 이유가 있는 때에는 운전자가 술에 취하였는지의 여부를 호흡 조사에 의하여 측정할 수 있다. 이 경우 운전자는 경찰공무원의 측정에 응하여야 한다.(개정 2006.7.19, 시행일 2006.10.20.)
 ③ 제2항의 규정에 의하여 술에 취하였는지의 여부를 측정한 결과에 불복하는 운전자에 대하여는 그 운전자의 동의를 얻어 혈액·채취 등의 방법으로 다시 측정할 수 있다.
 ④ 제1항의 규정에 따라 운전이 금지되는 술에 취한 상태의 기준은 혈중 알코올 농도가 0.05퍼센트 이상으로 한다.



- ① 술을 마셨더라도 운전자가 원하지 않을 경우 음주 측정을 거부할 수 있군.
- ② 운전자의 음주 운전 여부에 대한 최종 판단은 혈액·채취 결과만 인정이 되는군.
- ③ 음주 운전이 의심될 경우 경찰관은 바로 운전자에게 혈액·채취를 명할 수 있군.
- ④ 혈중 알코올 농도가 0.05퍼센트 미만이면 운전을 못할 정도로 취한 상태는 아니라고 보는군.

2.8.2 문9.

문 9) 서울에 사는 철수는 울릉도에 사는 영희에게 선물을 보내려고 합니다. 선물의 무게는 5kg이며, 가로·세로·높이가 각각 30cm입니다. 철수가 지불해야 할 요금은 얼마입니까?

<우체국 택배 요금표>

무게 (kg) (단위) 2kg까지 (60㎠까지)	5kg까지 (80㎠까지)	8kg까지 (100㎠까지)	10kg까지 (120㎠까지)
동양자석	3,000원	4,000원	5,000원
티켓	4,000원	5,000원	6,000원
도색자석	5,000원	6,000원	7,000원

☞ 주의사항 :

1. 크기는 가로, 세로, 높이를 더한 값을 말함.
2. 무게와 크기 중에 하나라도 기준을 초과하면 그에 해당하는 요금이 적용됨.



- ① 5,000원 ② 6,000원
③ 7,000원 ④ 8,000원

2.8.3 문12.

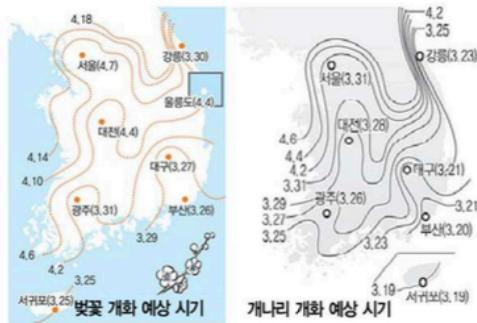
문 12) 다음은 2008년 4월 9일의 선거 때 배포된 '투표 확인증'에 대한 안내문입니다. 바르게 이해한 사람은 누구입니까?

투표 확인증 이용 안내	
<u>사용 가능 시설</u>	• 국립중앙박물관 및 소속 지방박물관(선거 당일만 이용 가능), 국립민속박물관 국립현대미술관, 국가지정문화재
<u>유효 기간</u>	2008년 4월 9일 ~ 4월 30일 (사설별 사용 기한이 다를 수 있으니 꼭 확인하시기 바랍니다.)
<u>이용 방법</u>	• '투표 확인증'을 소지한 분은 국립립 유료시설의 이용·관람시 무료 또는 할인 혜택을 받을 수 있습니다. • 면제·할인은 1인 1회에 한하여, 할인 금액은 2,000원 이내입니다. • '투표 확인증'은 타인에게 양도할 수 없으며, 현금으로 교환되지 않습니다. • '투표 확인증'은 도난·분실·멸실·훼손 등에 대하여 추가 배부하지 않습니다.

- ① 나는 확인증이 두 장 있으니 4,000원을 할인받을 수 있겠네요.
- ② 오늘은 4월 30일이니까, 국립중앙박물관에서 사용할 수 있겠죠?
- ③ 이 주차장은 할인이 안 될지 모르니 미리 물어봐야 해요.
- ④ 확인증을 잃어버렸는데, 다시 발급해 주겠지.

2.8.4 문15.

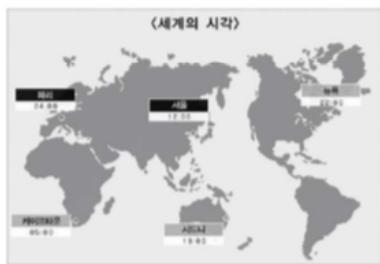
문 15) 다음 그림에 대한 설명으로 잘못된 것은 무엇입니까?



- ① 주요 도시의 꽃이 피는 날짜를 표시한 지도다.
- ② 동남쪽에서부터 꽃이 피기 시작한다.
- ③ 벚꽃이 개나리보다 일주일 정도 일찍 핀다.
- ④ 강릉이 대전보다 꽃이 더 일찍 핀다.

2.8.5 문17.

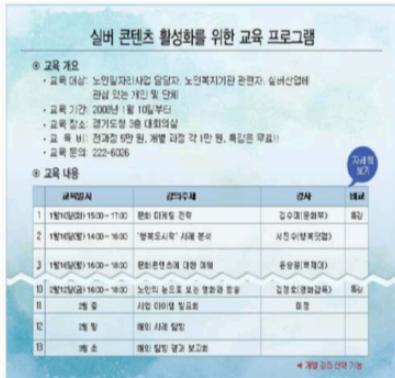
문 17) 프랑스 파리에서 국제 축구 경기가 열립니다. 한국 시간으로 밤 11시에 경기가 열린다면 현지 시간은 몇 시인니까?



- | | |
|---------|---------|
| ① 오전 3시 | ② 오전 6시 |
| ③ 오후 6시 | ④ 오후 3시 |

2.8.6 문22.

문 22) 다음 문화센터 교육 프로그램 내용을 바르게 이해한 것은 무엇입니까?



- ① 이 프로그램은 노인들을 위한 것이다.
- ② 교육은 특정 요일에 실시된다.
- ③ '노인의 눈으로 보는 영화와 방송' 특강은 서진수 씨가 강의한다.
- ④ 5만원을 내면 모든 강의를 들을 수 있다.

Chapter 3

1주차 데이터 실험 집계

3.1 실험의 목적

1주차 구글 예습 설문지 집계결과를 분석합니다.

Q1~Q3에서는 랜덤화의 효과로 Red, Black 이 얼마나 닮았는지 알아봅니다.

Q4에서는 같은 내용의 질문지인데 “바람직한 논의이다”라는 선택지에 부연설명을 붙이거나(Red), “부적절한 논의이다”라는 선택지에 부연설명을 붙였을 때(Black), 부연설명의 여부에 따라 응답이 달라지는지 살펴봅니다.

끝으로 제출시간의 분포가 날마다 고른지, Red, Black 간에는 닮았는지 알아봅니다.

3.1.1 Red, Black을 잘못 표시한 사람들

랜덤화출석부(2월 25일 기준)에 있는 Red, Black 과 실제 구글예습설문지에 올린 Red, Black 이 다른 사람들의 분포를 파악해 보았습니다.

	Red(구글예습퀴즈)	Black(구글예습퀴즈)
Red(랜덤화출석부)	286	8
Black(랜덤화출석부)	6	276
계	292	284

랜덤화출석부에 있는 Red, Black 과 실제 구글 설문지에서 선택한 Red, Black 이 다른 사람들의 수효는 14명입니다.

Red를 Black 이라고 한 사람이 8명, Black 을 Red 라고 한 사람이 6명입니다.

두 가지 방법으로 분석합니다.

우선 Red, Black 을 잘못 선택한 14명을 랜덤하게 둘로 나누면 어느 한 쪽 집단에 들어갈 기대인원은 14명을 둘로 나눈 7(명)이고, 표준오차는 14의 제곱근에 1/2을 곱해 준 1.9명이 됩니다.

실제로 Red를 Black 이라고 한 사람수, 8명이나 Black 을 Red 라고 한 사람수, 6명은 기대인원으로부터 표준오차 범위 안에 아주 잘 들어갑니다.

두 번째 분석 방법은 확률을 계산해 보는 것입니다.

Red, Black 을 잘못 선택한 14명을 랜덤하게 둘로 나눌 때, 실제로 관찰된 8명 이상이나 6명이하로 잘못 선택한 사람수가 나올 가능성은 얼마나 되는가 입니다.

이 경우 공평한 동전던지기를 확률 법칙으로 표현한 이항분포로부터 계산할 수 있습니다.

시행횟수가 14이고 한 번 시행에서 성공확률이 1/2 인 이항분포에서 성공갯수가 6이하이거나 8이상을 관찰할 확률은 0.791입니다.

공평한 동전 던지기에서 앞면이 6개 이하 나오는 확률은 8개 이상 나오는 확률과 같기 때문에 사실상 한쪽만 계산해서 2배 해 주면 됩니다.

이 값을 p-value 라고 하는데, p-value가 0.05보다 작을 때 통계적으로 유의한 차이를 관찰하였다고 말합니다.

즉, 공평한 동전을 던지는 것과 같은 과정이라고 가정하였을 때 실제로 관찰된 값들이 가정으로부터 얼마나 떨어져 있는지를 표현한 것입니다.

0.05, 즉 1/20은 이런 실험을 스무 번 정도 반복하면 1번 나올 정도로 드문 사건을 의미합니다.

즉 가정이 타당하다면 나오기 힘든 결과라는 것입니다.

그런데 Red, Black 을 잘못 표시한 사람들의 분포에서 관찰된 p-value 는 0.05와는 비교도 안될 정도로 큰 값입니다.

따라서 두 집단이 랜덤화 효과가 작동하여 통계적으로 유의한 차이를 보이지 않는다고 할 수 있습니다.

3.1.2 응답인원의 Red, Black

Red 로 응답한 인원은 292명, Black 에 응답한 인원은 284명입니다.

전체 응답인원 576 명을 랜덤하게 둘로 나눌 때 어느 한 쪽의 기대인원은 전체 응답인원의 절반인 288명이고, 표준오차는 전체 응답인원의 제곱근에 1/2을 곱해 준 12 명입니다.

따라서 Red, Black 각 그룹에 관찰된 인원은 기대인원으로부터 표준오차 범위 안에 들어갑니다.

간혹 이 범위를 살짝 벗어나는 경우들이 가끔 나오지만 두배의 표준오차 범위 안에는 거의 다 들어갑니다.

3.2. Q1. DEWEY AS GOOD AS ELECTED, STATISTICS CONVINCE ROPER25

3.2 Q1. Dewey as good as elected, statistics convince Roper

1. 다음 문장에서 statistics의 의미는? *

Dewey as good as elected, statistics convince Roper.

- 통계학
- 통계
- 자료
- 통계청
- 정보

3.2.1 Roper(Counts)

	통계학	통계	자료	통계청	정보	계
Red	26	230	25	6	5	292
Black	13	244	14	8	5	284
계	39	474	39	14	10	576

Table 3.3: Pearson's Chi-squared test: .

Test statistic	df	P value
8.026	4	0.09065

Q1의 집계 결과가 Red, Black 간에 통계적으로 유의한 차이가 있는지 알아보기 위하여 카이제곱 테스트를 수행하였습니다.

그 결과 카이제곱 통계량은 8.03, 자유도는 4, p-value는 0.09065이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

3.2.2 Roper(%)

통계학	통계	자료	통계청	정보	계
6.8	82.3	6.8	2.4	1.7	100.0

정답률은 Red, Black 을 합하여 계산하는데, 82.3(%) 입니다.

3.3 Q2. Statistics is the science of learning from data, ...

2. 다음 문장에서 statistics 의 의미는? *

Statistics is the science of learning from data, ...

- 통계학
 - 통계
 - 자료
 - 통계청
 - 정보
-

3.3.1 ASA(Counts)

	통계학	통계	자료	통계청	정보	계
Red	257	27	4	3	1	292
Black	262	14	3	2	3	284
계	519	41	7	5	4	576

Table 3.6: Pearson's Chi-squared test: .

Test statistic	df	P value
5.403	4	0.2484

Q2의 집계 결과가 Red, Black 간에 통계적으로 유의한 차이가 있는지 알아보기 위하여 카이제곱 테스트를 수행하였습니다.

그 결과 카이제곱 통계량은 5.403, 자유도는 4, p-value 는 0.2484이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

3.3.2 ASA(%)

통계학	통계	자료	통계청	정보	계
90.10	7.12	1.22	0.87	0.69	100.00

정답률은 Red, Black 을 합하여 계산하는데, 90.1(%) 입니다.

3.4 Q3. How to lie with statistics

3. 다음 문장에서 statistics 의 의미는? *

How to lie with statistics

- 통계학
- 통계
- 자료
- 통계청
- 정보

3.4.1 D.Huff(Counts)

	통계학	통계	자료	통계청	정보	계
Red	14	204	41	14	19	292
Black	16	213	28	6	21	284
계	30	417	69	20	40	576

Table 3.9: Pearson's Chi-squared test: .

Test statistic	df	P value
5.967	4	0.2016

Q3의 집계 결과가 Red, Black 간에 통계적으로 유의한 차이가 있는지 알아보기 위하여 카이제곱 테스트를 수행하였습니다.

그 결과 카이제곱 통계량은 5.967, 자유도는 4, p-value 는 0.2016이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

3.4.2 D.Huff(%)

통계학	통계	자료	통계청	정보	계
5.2	72.4	12.0	3.5	6.9	100.0

정답률은 Red, Black 을 합하여 계산하는데, 72.4(%) 입니다.

3.5 Q4. 종부세

“바람직한 논의이다”라는 선택지에 부연설명을 붙이거나(Red), “부적절한 논의이다”라는 선택지에 부연설명을 붙였을 때(Black), 부연설명의 여부에 따라 응답이 달라지는지 살펴본 결과 기대와는 달리 통계적으로 유의한 수준의 차이를 관찰하지 못하였습니다.

앞에서 본 바와 같이 Red, Black 두 집단은 출석부의 다섯 번수에 대하여 랜덤화 과정을 거쳐서 가장 닮은 구성을 찾은 것이기에 Q1, Q2, Q3의 응답 결과도 매우 닮게 나오는데 만약 부연설명이 효과가 없다면 Q4에서의 응답도 닮게 나왔을 것입니다.

지난 학기들과 달리 통계적으로 유의한 차이를 관찰하지 못한 이유를 따져 볼 필요가 있겠습니다.

3.5.1 질문지 선택지에 부연설명

Red

4. (NBS, 24/07/25 질문지 수정) 최근 정치권에서는 고가 1주택 보유자 및 다주택자에 대한 * 종합부동산세 부담을 완화해야 한다는 주장이 대두되고 있는데요. 이에 대한 다음의 의견 중 선생님께서는 어디에 더 동의하십니까?

- 납세자들의 과도한 세금 부담을 낮추기 위한 바람직한 논의이다
- 부적절한 논의이다
- 모름/무응답

Black

4. (NBS, 24/07/25 질문지 수정) 최근 정치권에서는 고가 1주택 보유자 및 다주택자에 대한 * 종합부동산세 부담을 완화해야 한다는 주장이 대두되고 있는데요. 이에 대한 다음의 의견 중 선생님께서는 어디에 더 동의하십니까?

- 바람직한 논의이다
- 상위 계층에 감세 혜택을 집중시킬 수 있는 부적절한 논의이다
- 모름/무응답

3.5.2 집계

	바람직한 논의이다	부적절한 논의이다	모름/무응답	계
Red(바람직한 논의에 부연설명)	95	113	84	292
Black(부적절한 논의에 부연설명)	72	130	82	284
계	167	243	166	576

Table 3.12: Pearson's Chi-squared test: .

Test statistic	df	P value
4.271	2	0.1182

Q4의 Red에는 종합부동산세 부담을 완화해야 한다는 주장에 대하여 바람직한 논의라는 쪽에 긍정적인 부연설명을 붙였는데, 292명이 응답한 가운데 95명이

“바람직한 논의이다”라는 반응을 보이고, 113명이 “부적절한 논의이다”라는 반응을 보입니다.

Black에는 같은 주장에 대하여 부적절한 논의라는 쪽에 부정적인 부연설명을 붙였는데, 284명이 응답한 가운데 72명이 “바람직한 논의이다”라는 반응을 보이고, 130명이 “부적절한 논의이다”라는 반응을 보입니다.

그리고 “모름/무응답”에 답한 인원은 Red에 84명, Black 에 82명이 응답하였습니다.

지난 학기 자료들에서 볼 수 있다시피 카이제곱 테스트는 이와 같은 상황에서 부연설명의 유무가 응답에 미치는 영향이 대부분 통계적으로 유의하다는 것을 보여 줍니다.

그런데, 이번 학기는 매우 예외적으로 그렇지 않은 경우가 관찰되었습니다.

카이제곱 통계량은 4.271, 자유도는 2, p-value 는 0.1182으로 부연설명을 어떻게 붙이느냐에 따라 반응이 다르게 나온다는 것을 보여주고 싶었지만 실제로 관찰된 차이는 Q1 ~ Q3와 마찬가지로 통계적으로 유의한 수준은 아닙니다.

여기서 부연설명이 응답에 영향을 끼치지 않는다고 가정해 봅시다.

그렇다면 Red, Black 의 응답은 Q1~Q3 에서와 같이 랜덤화 효과에 의하여 통계적으로 유의한 차이를 보이지 않을 것입니다.

그런데 실제로 관찰된 카이제곱 통계값과 P-value 는 통계적으로 유의한 차이를 보여 주지 못하는 수준입니다.

따라서 부연설명이 영향을 끼치지 않는다는 가정을 받아들일 수밖에 없게 되었습니다.

지난 학기 자료들이 모두 통계적으로 유의한 차이를 보여 주었던 것과는 달리 이번 학기에 유독 통계적으로 유의한 차이를 보이지 않는 이유는 무엇일까요?

3.5.3 % 비교.

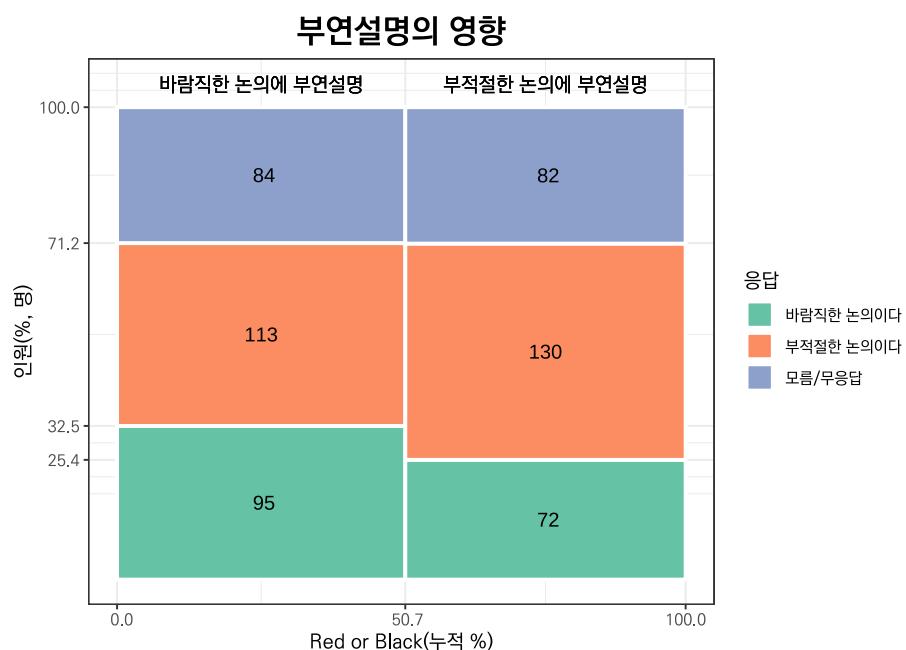
	바람직한 논의이다	부적절한 논의이다	모름/무응답	계
Red(바람직한 논의에 부연설명)	32.5	38.7	28.8	100.0
Black(부적절한 논의에 부연설명)	25.4	45.8	28.9	100.0

“바람직한 논의이다”에 부연설명을 붙인 Red에서 “바람직한 논의이다”라고 응답하는 사람들의 백분율, 32.5(%)은 “부적절한 논의이다”에 부연설명을 붙인 Black 에서 “바람직한 논의이다”라고 응답하는 사람들의 백분율, 25.4(%) 보다 높습니다.

반면 “부적절한 논의이다”에 부연설명을 붙인 Black에서 “부적절한 논의이다”라고 응답하는 사람들의 백분율, 45.8(%)은 Red에서 “부적절한 논의이다”라고 응답하는 사람들의 백분율, 38.7(%) 보다 높습니다.

부연설명을 어디에 붙이느냐에 따라 반응이 치우치는 것을 알 수 있지만 p-value 가 보여주듯이 그 차이가 통계적으로 유의한 수준은 아닙니다.

3.5.4 Mosaic Plot



Mosaic Plot은 이 집계결과를 시각적으로 잘 보여줍니다.

“바람직한 논의이다”에 부연설명을 붙인 Red에서 “바람직한 논의이다”라고 응답한 백분율이 “부적절한 논의이다”에 부연설명을 붙인 Black에서 “바람직한 논의이다”라고 응답한 백분율보다 높고, Black에서 “부적절한 논의이다”라고 응답한 백분율은 Red에서 “부적절한 논의이다”라고 응답한 백분율보다 월등히 높습니다.

3.6 마감 시간으로부터 제출 시간의 분포

3.6.1 분포표

Table 3.14: 일 단위

	[0,1]	(1,2]	(2,3]	(3,4]	(4,5]	(5,6]	(6,7]	(7,8]	(8,9]	(9,10]	(10,11]	(11,12]	(12,13]	(13,14]
Red	46	11	6	11	6	14	10	42	21	19	15	30	35	26
Black	35	10	6	7	5	12	12	34	22	17	26	37	29	32
계	81	21	12	18	11	26	22	76	43	36	41	67	64	58

분포표로부터 두 가지 문제를 살펴보겠습니다.

첫째, 날마다 고르게 제출하는가?

둘째, Red, Black 간에 통계적으로 유의한 차이가 있는가?

각 문제를 살펴보기 위해서는 분포표의 일부분을 대상으로 카이제곱 테스트를 수행합니다.

3.6.2 날마다 고르게 제출하는가?

[0,1]	(1,2]	(2,3]	(3,4]	(4,5]	(5,6]	(6,7]	(7,8]	(8,9]	(9,10]	(10,11]	(11,12]	(12,13]	(13,14]
81	21	12	18	11	26	22	76	43	36	41	67	64	58

Table 3.16: Chi-squared test for given probabilities: .

Test statistic	df	P value
184.8	13	1.766e-32 * * *

날마다 고르게 제출하는지 알아 보았습니다.

분포표의 “계”행에서 ’계’열을 제외하고 카이제곱테스트를 수행합니다.

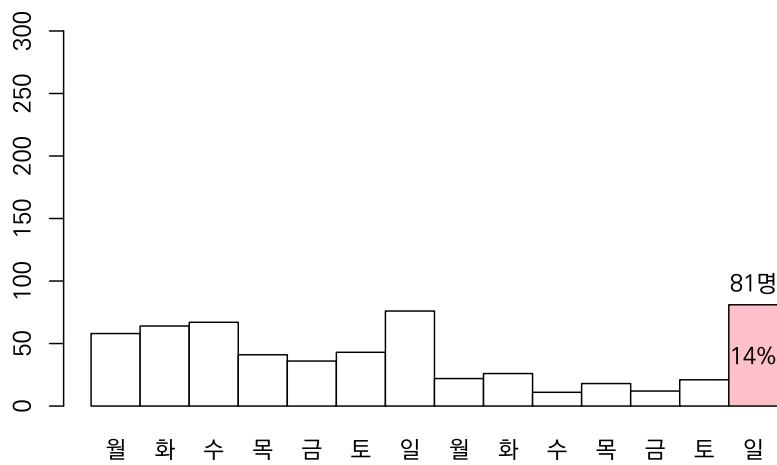
분포표 만으로도 쉽게 파악할 수 있지만 카이제곱테스트가 명확히 해 줍니다.

카이제곱 통계량은 184.812, 자유도는 13, p-value 는 1.8e-32 이므로 결코 고르게 제출한다고 말할 수 없겠습니다.

막대그래프로 살펴 보겠습니다.

3.6.3 막대그래프

Quiz250303 (576명 제출)



막대그래프는 총 제출인원 576(명) 중에 81(명), 14(%)가 마감일에 몰리는 것을 보여주고 있습니다.

3.6.4 Red, Black 간에 닮았는가?

	[0,1](1,2](2,3](3,4](4,5](5,6](6,7](7,8](8,9](9,10](10,11](11,12](12,13](13,14]													
Red	46	11	6	11	6	14	10	42	21	19	15	30	35	26
Black	35	10	6	7	5	12	12	34	22	17	26	37	29	32

Table 3.18: Pearson's Chi-squared test: .

Test statistic	df	P value
8.59	13	0.8032

제출시간의 분포가 Red, Black 간에 닮았는지 알아 보았습니다.

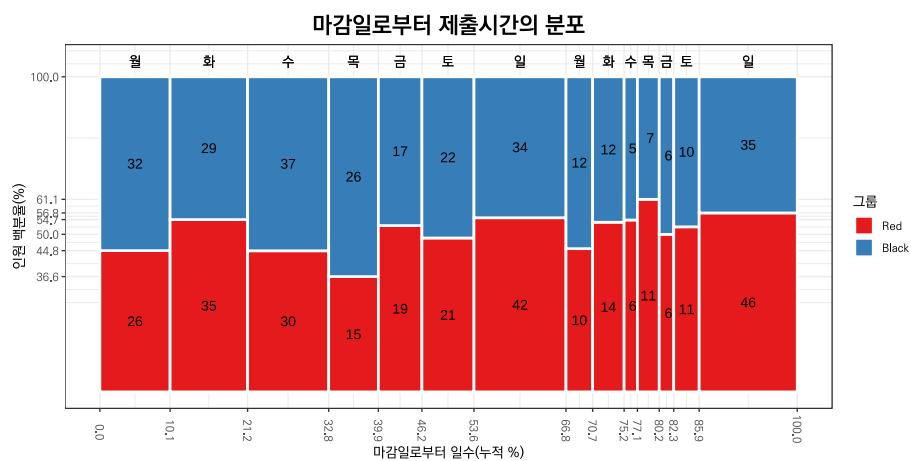
이번에는 분포표의 첫번째와 두번째 행, '계'열을 제외한 나머지 열에 대해서 카이제곱테스트를 수행합니다.

카이제곱 통계량은 8.590, 자유도는 13, p-value 는 0.8032 이므로 제출 시간의 분포는 Red, Black 간에 통계적으로 유의한 차이가 관찰되지 않습니다.

이 사실을 Mosaic Plot 을 이용하여 시각적으로 살펴보겠습니다.

답았다고 느껴지나요?

3.6.5 Mosaic Plot



Chapter 4

2주차 데이터 실험 집계

4.1 실험의 목적

2주차 구글 예습 설문지 집계결과를 분석합니다.

Q1~Q6에서는 랜덤화의 효과로 Red, Black 이 얼마나 닮았는지 알아봅니다.

Q7에서는 같은 눈속임 그래프인데 원형그래프의 각도를 속일 때(Red)와 막대그래프의 높이를 속일 때(Black) 오류를 지각하는데 차이가 있는지 알아봅니다.

끝으로 제출시간의 분포가 날마다 고른지, Red, Black 간에는 닮았는지 알아봅니다.

4.1.1 Red, Black을 잘못 표시한 사람들

제출시간	학번	랜덤화출석부	구글예습퀴즈
2025-03-11 21:44:18	20223501	Red	Black
2025-03-12 01:12:52	20246792	Black	Red
2025-03-12 02:27:14	20242601	Red	Black
2025-03-13 02:06:41	20246737	Red	Black
2025-03-13 19:03:57	20241216	Red	Black
2025-03-14 09:32:05	20241022	Black	Red

제출시간	학번	랜덤화출석부	구글예습퀴즈
2025-03-14 10:22:29	20231725	Red	Black
2025-03-14 15:07:21	20243233	Black	Red
2025-03-15 15:21:53	20213033	Red	Black
2025-03-15 22:20:27	20241201	Red	Black
2025-03-15 22:46:42	20243720	Red	Black
2025-03-16 15:31:37	20223329	Black	Red
2025-03-16 17:32:57	20211021	Black	Red
2025-03-16 19:50:32	20241011	Red	Black
2025-03-16 20:56:30	20213040	Red	Black
2025-03-16 23:10:55	20241028	Black	Red
2025-03-17 21:59:54	20242960	Black	Red
2025-03-20 11:14:21	20246272	Black	Red
2025-03-20 14:54:36	20246923	Black	Red
2025-03-20 20:52:14	20242949	Black	Red
2025-03-22 00:55:57	20246606	Black	Red
2025-03-23 11:51:35	20242585	Red	Black
2025-03-23 14:27:48	20243513	Black	Red
2025-03-23 19:53:42	20227107	Red	Black

	Red(구글예습퀴즈)	Black(구글예습퀴즈)
Red(랜덤화출석부)	268	12
Black(랜덤화출석부)	12	269
계	280	281

랜덤화출석부에 있는 Red, Black 과 실제 구글설문에 올린 Red, Black 이 다른 사람들의 수효는 24명입니다.

Red를 Black 이라고 한 사람이 12명, Black 을 Red 라고 한 사람이 12명입니다.
두 가지 방법으로 분석합니다.

우선 Red, Black 을 잘못 선택한 24명을 랜덤하게 둘로 나누면 어느 한 쪽 집단에 들어갈 기대인원은 24명을 둘로 나눈 12(명)이고, 표준오차는 24의 제곱근에 $1/2$ 을 곱해 준 2.4명이 됩니다.

실제로 Red를 Black 이라고 한 사람수, 12명이나 Black 을 Red 라고 한 사람수, 12명은 기대인원으로부터 표준오차 범위 안에 아주 잘 들어갑니다.

두 번째 분석 방법은 확률을 계산해 보는 것입니다.

Red, Black 을 잘못 선택한 24명을 랜덤하게 둘로 나눌 때, 실제로 관찰된 12명 이상이나 12명이하로 잘못 선택한 사람수가 나올 가능성은 얼마나 되는가 입니다.

이 경우 공평한 동전던지기를 확률 법칙으로 표현한 이항분포로부터 계산할 수 있습니다.

시행횟수가 24이고 한 번 시행에서 성공확률이 $1/2$ 인 이항분포에서 성공횟수가 12이하이거나 12이상을 관찰할 확률은 1.161입니다.

공평한 동전 던지기에서 앞면이 12개 이하 나오는 확률은 12개 이상 나오는 확률과 같기 때문에 사실상 한쪽만 계산해서 2배 해 주면 됩니다.

이 값을 p-value 라고 하는데, p-value가 0.05보다 작을 때 통계적으로 유의한 차이를 관찰하였다고 말합니다.

즉, 공평한 동전을 던지는 것과 같은 과정이라고 가정하였을 때 실제로 관찰된 값들이 가정으로부터 얼마나 떨어져 있는지를 표현한 것입니다.

0.05, 즉 $1/20$ 은 이런 실험을 스무 번 정도 반복하면 1번 나올 정도로 드문 사건을 의미합니다.

즉 가정이 타당하다면 나오기 힘든 결과라는 것입니다.

그런데 Red, Black 을 잘못 표시한 사람들의 분포에서 관찰된 p-value 는 0.05와는 비교도 안될 정도로 큰 값입니다.

따라서 두 집단이 랜덤화 효과가 작동하여 통계적으로 유의한 차이를 보이지 않는다고 할 수 있습니다.

4.1.2 응답인원의 Red, Black

Red 로 응답한 인원은 280명, Black 에 응답한 인원은 281명입니다.

전체 응답인원 561 명을 랜덤하게 둘로 나눌 때 어느 한 쪽의 기대인원은 전체 응답인원의 절반인 280.5명이고, 표준오차는 전체 응답인원의 제곱근에 $1/2$ 을 곱해 준 11.8 명입니다.

따라서 Red, Black 각 그룹에 관찰된 인원은 기대인원으로부터 표준오차 범위, 혹은 두배의 표준오차 범위 안에 들어갑니다.

4.2 Q1. 춘추전국시대에 국가통계관리의 중요성 강조

춘추전국시대에 국가통계관리의 중요성을 역설한 사람은 누구인가? *

- 공자
- 맹자
- 관자
- 노자
- 장자

4.2.1 관자(집계표)

	공자	맹자	관자	노자	장자	계
Red	40	9	224	8	1	282
Black	25	17	223	11	5	281
계	65	26	447	19	6	563

Table 4.4: Pearson's Chi-squared test with simulated p-value (based on 2000 replicates): .

Test statistic	df	P value
9.064	NA	0.05197

Q1의 집계 결과가 Red, Black 간에 통계적으로 유의한 차이가 있는지 알아보기 위하여 카이제곱 테스트를 수행하였습니다.

그 결과 카이제곱 통계량은 9.06, 자유도는 NA, p-value 는 0.05200이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

4.2.2 관자(%)

공자	맹자	관자	노자	장자	계
11.5	4.6	79.4	3.4	1.1	100.0

정답률은 Red, Black 을 합하여 계산하는데, 79.4(%) 입니다.

4.3 Q2. 국가정책을 수립하는 데 통계의 역할

국가 정책을 수립하는 데 있어서 통계의 역할은 어느 정도 중요하다고 생각하는가? *

- 절대로 중요하지 않다
- 거의 중요하지 않다
- 보통이다
- 상당히 중요하다
- 절대적으로 중요하다

4.3.1 통계의 중요성(집계표)

	절대로 중요하지 않다	거의 중요하지 않다	보통이다	상당히 중요하다	절대적으로 중요하다	계
Red	0	2	9	94	177	282
Black	2	4	9	94	172	281
계	2	6	18	188	349	563

Table 4.7: Pearson's Chi-squared test with simulated p-value
(based on 2000 replicates): .

Test statistic	df	P value
2.056	NA	0.6687

Q2의 집계 결과가 Red, Black 간에 통계적으로 유의한 차이가 있는지 알아보기 위하여 카이제곱 테스트를 수행하였습니다.

그 결과 카이제곱 통계량은 2.056, 자유도는 NA, p-value 는 0.6687이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

4.3.2 통계의 중요성(%)

절대로 중요하지 않다	거의 중요하지 않다	보통이다	상당히 중요하다	절대적으로 중요하다	계
0.36	1.07	3.20	33.39	61.99	100.00

정답률은 Red, Black 을 합하여 계산하는데, 62.0(%) 입니다.

4.4 Q3. 우리나라 생산가능인구 감소 시기

우리나라의 생산가능인구는 다음 중 어느 연도부터 감소하기 시작하였는가? *

- 2012
- 2017
- 2022
- 2027

4.4.1 생산가능인구 감소 시기(집계표)

	2012	2017	2022	2027	계
Red	31	224	23	4	282
Black	23	227	25	6	281
계	54	451	48	10	563

Table 4.10: Pearson's Chi-squared test: .

Test statistic	df	P value
1.687	3	0.6399

Q3의 집계 결과가 Red, Black 간에 통계적으로 유의한 차이가 있는지 알아보기 위하여 카이제곱 테스트를 수행하였습니다.

그 결과 카이제곱 통계량은 1.687, 자유도는 3, p-value 는 0.6399이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

4.4.2 생산가능인구 감소 시기(%)

	2012	2017	2022	2027	계
	9.6	80.1	8.5	1.8	100.0

정답률은 Red, Black 을 합하여 계산하는데, 80.1(%) 입니다.

4.5 Q4. 우리나라 총인구 최대 시기

우리나라의 총인구는 어느 연도에 가장 많았는가? *

- 2018
- 2019
- 2020
- 2021

4.5.1 총인구 최대 시기(집계표)

	2018	2019	2020	2021	계
Red	47	27	205	3	282
Black	29	27	215	10	281
계	76	54	420	13	563

Table 4.13: Pearson's Chi-squared test with simulated p-value
(based on 2000 replicates): .

Test statistic	df	P value
8.269	NA	0.03248 *

Q4의 집계 결과가 Red, Black 간에 통계적으로 유의한 차이가 있는지 알아보기 위하여 카이제곱 테스트를 수행하였습니다.

그 결과 카이제곱 통계량은 8.269, 자유도는 NA, p-value 는 0.0325이므로 Red, Black 간에 통계적으로 유의한 차이를 보이고 있습니다.

4.5.2 총인구 최대 시기(%)

	2018	2019	2020	2021	계
	13.5	9.6	74.6	2.3	100.0

정답률은 Red, Black 을 합하여 계산하는데, 74.6(%) 입니다.

4.6 Q5. 소멸위험 단계 개선 지역

다음 중 2013년에는 소멸위험 주의단계였는데, 2018년에는 소멸위험 매우 낮은 단계로 바뀐 곳은? *

- 서울
- 경기
- 세종
- 제주

4.6.1 소멸위험 단계 개선 지역(집계표)

	서울	경기	세종	제주	계
Red	13	14	241	14	282
Black	9	14	245	13	281
계	22	28	486	27	563

Table 4.16: Pearson's Chi-squared test with simulated p-value
(based on 2000 replicates): .

Test statistic	df	P value
0.7955	NA	0.8691

Q5의 집계 결과가 Red, Black 간에 통계적으로 유의한 차이가 있는지 알아보기 위하여 카이제곱 테스트를 수행하였습니다.

그 결과 카이제곱 통계량은 0.795, 자유도는 NA, p-value 는 0.8691이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

4.6.2 소멸위험 단계 개선 지역(%)

서울	경기	세종	제주	계
3.9	5.0	86.3	4.8	100.0

정답률은 Red, Black 을 합하여 계산하는데, 86.3(%) 입니다.

4.7 Q6. 조출생률과 합계출산율

다음 A, B, C 세 동네의 조출생률과 합계출산율을 제대로 구한 것을 고르시오. *



- 조출생률 95.2로 동일, 합계출산율 A : 0.5, B : 0.125, C : 0.25
- 조출생률 95.2로 동일, 합계출산율 A : 0.25, B : 0.125, C : 0.5
- 조출생률 95.2로 동일, 합계출산율 A : 0.125, B : 0.5, C : 0.25
- 조출생률 95.2로 동일, 합계출산율 A : 0.5, B : 0.25, C : 0.125

4.7.1 조출생률과 합계출산율(집계표)

	합계출산율			합계출산율 A : 0.5, B : 0.25, C : 0.125 계
	합계출산율 A : 0.25, B : 0.125, C : 0.25	합계출산율 A : 0.125, B : 0.5, C : 0.25	합계출산율 A : 0.5, B : 0.25, C : 0.125	
Red	163	34	56	29 282
Black	140	56	48	37 281
계	303	90	104	66 563

Table 4.19: Pearson's Chi-squared test: .

Test statistic	df	P value
8.707	3	0.03345 *

Q6의 집계 결과가 Red, Black 간에 통계적으로 유의한 차이가 있는지 알아보기 위하여 카이제곱 테스트를 수행하였습니다.

그 결과 카이제곱 통계량은 8.707, 자유도는 3, p-value 는 0.0335이므로 Red, Black 간에 통계적으로 유의한 차이를 보이고 있습니다.

4.7.2 조출생률과 합계출산율(%)

합계출산율 A		합계출산율 A :		합계출산율 A :		계
합계출산율 A :	: 0.25, B :	0.125, C :	0.125, B : 0.5,	0.5, B : 0.25, C	: 0.125	
0.5, B : 0.125, C : 0.25	0.125	0.5	C : 0.25			53.8 16.0 18.5 11.7 100.0

정답률은 Red, Black 을 합하여 계산하는데, 53.8(%) 입니다.

4.8 Q7. 눈속임 그래프(Cheating Charts)

지난 학기까지 앞에 나오는 선지를 고르기 쉽다는 1번효과에 대한 질문을 만들어서 테스트해 왔지만 효과를 검증하기 어려워 문제를 바꿔 보았습니다.

언론방송에서 가끔 원형그래프나 막대그래프를 제시하면서 숫자와 그림이 맞지 않는 경우를 볼 수 있습니다.

여러분들은 그런 경우에 어떻게 인식하는지 언론기관에서 발표한 눈속임 그래프를 보여줍니다.

Red에는 원형그래프의 각도를 속이고, Black 에는 막대그래프의 높이를 속여 어떤 응답이 나오는지 살펴보았습니다.

여러분들은 대부분 눈속임 그래프에 속지 않고 있습니다.

언론기관들이 왜 이런 짓들을 하는지 궁금해집니다.

Red

(JTBC 2015년 3월3일) 김영란법 통과 관련 여론조사를 원형그래프로 요약하였다. 원형그래프가 여론조사 결과를 제대로 보여 주고 있나요?

Category	Percentage (%)
잘했다 (잘)	64.0
모르겠다 (모르겠다)	28.7
잘못했다 (잘못)	7.3

○ 제대로 보여주고 있다
○ 제대로 보여주고 있지 않다
○ 모름/응답거절

Black

(JTBC 뉴스룸 2017년 2월) 차기 대선주자 지지율 여론조사 결과를 막대그래프로 요약하였다. 막대그래프가 여론조사 결과를 제대로 보여주고 있나요?

후보자	지지율 (%)
문재인	33
안희정	22
황교안	9
안철수	9
이재명	5
유승민	2

○ 제대로 보여주고 있다
○ 제대로 보여주고 있지 않다
○ 모름/응답거절

4.8.1 집계표

	제대로 보여주고 있다	제대로 보여주고 있지 않다	모름/응답거절 계	
Red(김영란법 국회통과)	65	180	37	282
Black(고위공직자 범죄수사처 설립)	100	119	62	281
계	165	299	99	563

Table 4.22: Pearson's Chi-squared test: .

Test statistic	df	P value
26.18	2	2.065e-06 * * *

Q7의 Red에는 김영란법 국회통과에 대한 여론조사 결과를 원형그래프로 나타내었는데 잘했다(64%), 잘못했다(7.3%), 모르겠다(28.7%)의 각도를 데이터와 전혀 맞지 않게 왜곡하여 마치 잘했다와 잘못했다의 비율이 거의 대등한 것처럼 각도를 조정하였습니다.

282명이 응답한 가운데 65명이 결과를 “제대로 보여주고 있다”는 반응을 보이고, 180명이 결과를 “제대로 보여주고 있지 않다”는 반응을 보입니다.

Black은 2017년 대선의 대선주자 여론조사에서 33%의 지지율을 기록한 문재인 예비후보와 22%의 지지율을 기록한 안희정 예비후보의 지지율 막대가 거의 비슷한 것처럼 왜곡하였습니다.

281명이 응답한 가운데 100명이 여론조사 결과를 ”제대로 보여주고 있다”는 반응을 보이고, 119명이 여론조사 결과를 “제대로 보여주고 있지 않다”는 반응을 보입니다.

그리고 “모름/무응답”에 답한 인원은 Red에 37명, Black에 62명이었습니다.

카이제곱 테스트는 이와 같은 상황에서 원형그래프를 왜곡할 때와 막대그래프를 왜곡할 때 인식의 차이가 통계적으로 유의하다는 것을 보여 줍니다.

카이제곱 통계량은 26.180, 자유도는 2, p-value 는 2.1e-06(으)로 그래프의 유형에 따라 눈속임의 인식에 통계적으로 유의한 차이가 관찰된다는 것을 보여줍니다.

여기서 그래프의 유형이 눈속임의 인식에 차이를 주지 않는다고 가정합니다.

랜덤화의 효과로 Red, Black 의 응답은 닮게 마련입니다.

즉, 통계적으로 유의한 차이를 보이지 않게 됩니다.

그러나 실제로 관찰된 카이제곱 통계값의 P-value 는 0.05보다 매우 작은 값입니다.

따라서, 그래프의 유형이 눈속임의 인식에 영향을 끼치지 않는다는 가정은 잘못된 것이죠.

이러한 논증 방식을 귀류법이라고 합니다.

4.8.2 % 비교

	제대로 보여주고 있다	제대로 보여주고 있지 않다	모름/응답거절	계
Red(김영란법 국회통과)	23.0	63.8	13.1	100.0
Black(고위공직자 범죄수사처 설립)	35.6	42.3	22.1	100.0

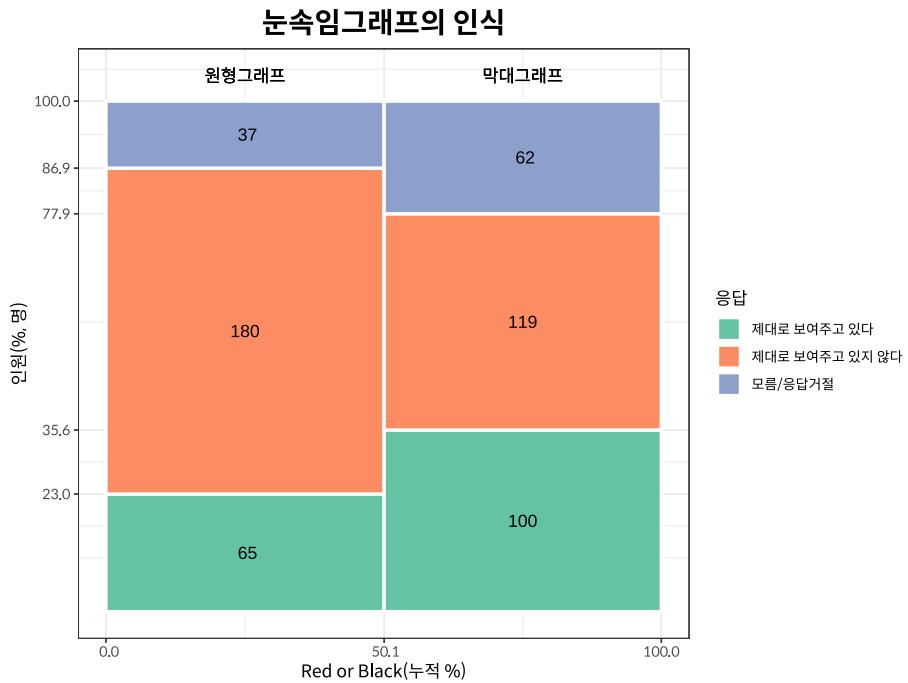
원형그래프의 각도를 왜곡한 Red에서 여론조사 결과를 “제대로 보여주고 있다”고 응답하는 사람들의 백분율, 23.0(%)은 “제대로 보여주고 있지 않다”고 응답하는 사람들의 백분율, 63.8(%) 보다 매우 낮습니다.

반면 막대그래프의 높이를 왜곡한 Black에서 여론조사 결과를 “제대로 보여주고 있다”고 응답하는 사람들의 백분율, 35.6(%)은 “제대로 보여주고 있지 않다”고 응답하는 사람들의 백분율, 42.3(%) 보다 적습니다.

원형그래프에서 눈속임을 지각하는 백분율이 막대그래프에서 눈속임을 지각하는 백분율보다 훨씬 높게 나타나고 있습니다.

원형그래프의 각도를 속이느냐, 막대그래프의 높이를 속이느냐에 따라 반응이 달라진다는 것을 잘 알 수 있습니다.

4.8.3 Mosaic Plot



Mosaic Plot 은 이 집계결과를 시각적으로 잘 보여줍니다.

원형그래프의 각도를 왜곡한 Red 에서 여론조사 결과를 “제대로 보여주고 있다”고 응답한 백분율이 매우 낮고, 막대그래프의 높이를 왜곡한 Black 에서 여론조사 결과를 “제대로 보여주고 있다”고 응답한 백분율은 상대적으로 덜 낮은 것을 시각적으로 알 수 있습니다.

4.9 마감 시간으로부터 제출 시간의 분포

4.9.1 분포표

Table 4.24: 일 단위

	[0,1]	(1,2]	(2,3]	(3,4]	(4,5]	(5,6]	(6,7]	(7,8]	(8,9]	(9,10]	(10,11]	(11,12]	(12,13]	(13,14]	(14,15]
Red	52	17	9	6	12	8	12	51	20	13	21	13	21	27	282
Black	59	11	11	3	6	7	9	43	19	18	20	21	26	28	281
계	111	28	20	9	18	15	21	94	39	31	41	34	47	55	563

분포표로부터 두 가지 문제를 살펴보겠습니다.

첫째, 날마다 고르게 제출하는가?

둘째, Red, Black 간에 통계적으로 유의한 차이가 있는가?

각 문제를 살펴보기 위해서는 분포표의 일부분을 대상으로 카이제곱 테스트를 수행합니다.

4.9.2 날마다 고르게 제출하는가?

(1,2]	(2,3]	(3,4]	(4,5]	(5,6]	(6,7]	(7,8]	(8,9]	(9,10]	(10,11]	(11,12]	(12,13]	(13,14]
28	20	9	18	15	21	94	39	31	41	34	47	55

Table 4.26: Chi-squared test for given probabilities: .

Test statistic	df	P value
170.5	12	3.764e-30 * * *

날마다 고르게 제출하는지 알아 보았습니다.

분포표의 “계”행에서 ’계’열을 제외하고 카이제곱테스트를 수행합니다.

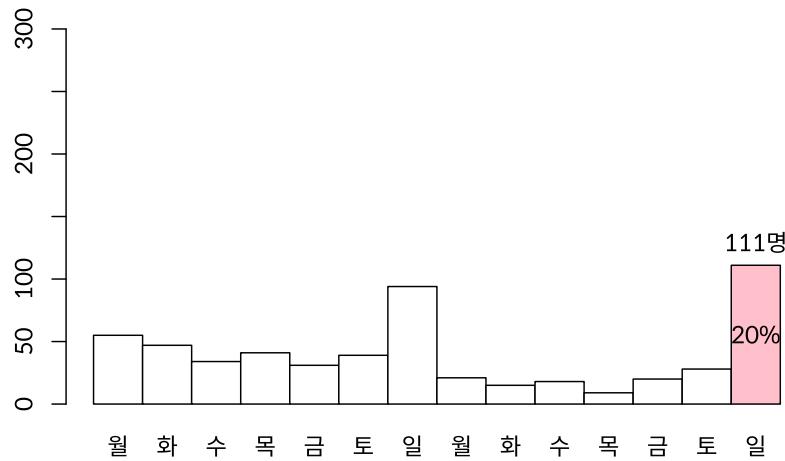
분포표 만으로도 쉽게 파악할 수 있지만 카이제곱테스트가 명확히 해 줍니다.

카이제곱 통계량은 170.50, 자유도는 12.00, p-value 는 3.8e-30 이므로 결코 고르게 제출한다고 말할 수 없겠습니다.

막대그래프로 살펴 보겠습니다.

4.9.3 막대그래프

Quiz250310 (563명 제출)



막대그래프는 총 제출인원 563(명) 중에 111(명), 20(%)가 마감일에 몰리는 것을 명확히 보여주고 있습니다.

4.9.4 Red, Black 간에 닮았는가?

	(1,2]	(2,3]	(3,4]	(4,5]	(5,6]	(6,7]	(7,8]	(8,9]	(9,10]	[10,11]	[11,12]	[12,13]	[13,14]
Red	17	9	6	12	8	12	51	20	13	21	13	21	27
Black	11	11	3	6	7	9	43	19	18	20	21	26	28

Table 4.28: Pearson's Chi-squared test: .

Test statistic	df	P value
8.812	12	0.7189

제출시간의 분포가 Red, Black 간에 닮았는지 알아 보았습니다.

이번에는 분포표의 첫번째와 두번째 행, '계'열을 제외한 나머지 열에 대해서 카이제곱테스트를 수행합니다.

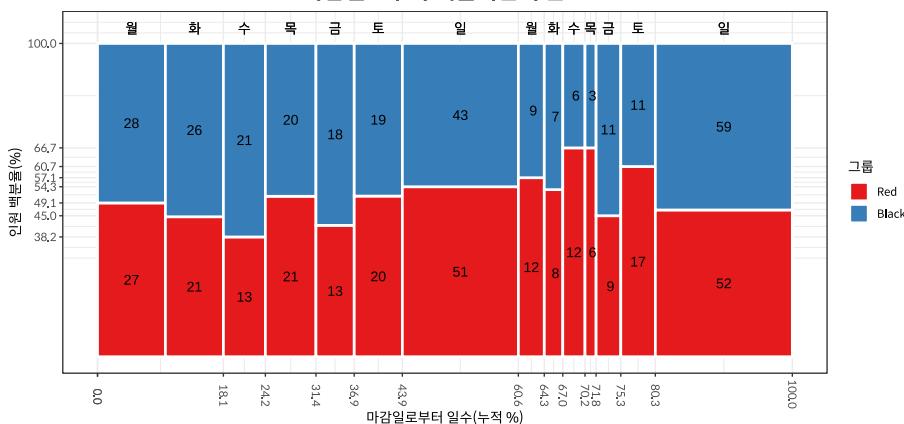
카이제곱 통계량은 8.812, 자유도는 12, p-value 는 0.7189 이므로 제출 시간의 분포는 Red, Black 간에 통계적으로 유의한 차이가 관찰되지 않습니다.

이 사실을 Mosaic Plot 을 이용하여 시각적으로 살펴보겠습니다.

답았다고 느껴지나요?

4.9.5 Mosaic Plot

마감일로부터 제출시간의 분포



Chapter 5

3주차 데이터 실험 집계

5.1 실험의 목적

3주차 구글 예습 설문지 집계결과를 분석합니다.

Q1~Q6에서는 랜덤화의 효과로 Red, Black 이 얼마나 닮았는지 알아봅니다.

Q7에서는 같은 사안에 대해서 질문 안에 편향된 정보를 담아 넣었을 때 Red, Black 의 응답이 어떻게 달라지는지 알아봅니다.

끝으로 제출시간의 분포가 날마다 고른지, Red, Black 간에는 닮았는지 알아봅니다.

5.1.1 Red, Black을 잘못 표시한 사람들

	Red(구글예습퀴즈)	Black(구글예습퀴즈)
Red(랜덤화출석부)	273	5
Black(랜덤화출석부)	4	271
계	277	276

랜덤화출석부에 있는 Red, Black 과 실제 구글설문에 올린 Red, Black 이 다른 사람들의 수효는 9명입니다.

Red를 Black 이라고 한 사람이 5명, Black 을 Red 라고 한 사람이 4명입니다.

두 가지 방법으로 분석합니다.

우선 Red, Black 을 잘못 선택한 9명을 랜덤하게 둘로 나누면 어느 한 쪽 집단에 들어갈 기대인원은 9명을 둘로 나눈 4.5(명)이고, 표준오차는 9의 제곱근에 1/2을 곱해 준 1.5명이 됩니다.

실제로 Red를 Black 이라고 한 사람수, 5명이나 Black 을 Red 라고 한 사람수, 4명은 기대인원으로부터 표준오차 범위 안에 아주 잘 들어갑니다.

두 번째 분석 방법은 확률을 계산해 보는 것입니다.

Red, Black 을 잘못 선택한 9명을 랜덤하게 둘로 나눌 때, 실제로 관찰된 5명 이상이나 4명이하로 잘못 선택한 사람수가 나올 가능성은 얼마나 되는가 입니다.

이 경우 공평한 동전던지기를 확률 법칙으로 표현한 이항분포로부터 계산할 수 있습니다.

시행횟수가 9이고 한 번 시행에서 성공확률이 $1/2$ 인 이항분포에서 성공횟수가 4이하이거나 5이상을 관찰할 확률은 1입니다.

공평한 동전 던지기에서 앞면이 4개 이하 나오는 확률은 5개 이상 나오는 확률과 같기 때문에 사실상 한쪽만 계산해서 2배 해 주면 됩니다.

이 값을 p-value 라고 하는데, p-value가 0.05보다 작을 때 통계적으로 유의한 차이를 관찰하였다고 말합니다.

즉, 공평한 동전을 던지는 것과 같은 과정이라고 가정하였을 때 실제로 관찰된 값들이 가정으로부터 얼마나 떨어져 있는지를 표현한 것입니다.

0.05, 즉 $1/20$ 은 이런 실험을 스무 번 정도 반복하면 1번 나올 정도로 드문 사건을 의미합니다.

즉 가정이 타당하다면 나오기 힘든 결과라는 것입니다.

그런데 Red, Black 을 잘못 표시한 사람들의 분포에서 관찰된 p-value 는 0.05와는 비교도 안될 정도로 큰 값입니다.

따라서 두 집단이 랜덤화 효과가 작동하여 통계적으로 유의한 차이를 보이지 않는다고 할 수 있습니다.

5.1.2 응답인원의 Red, Black

Red 로 응답한 인원은 277명, Black 에 응답한 인원은 276명입니다.

전체 응답인원 553 명을 랜덤하게 둘로 나눌 때 어느 한 쪽의 기대인원은 전체 응답인원의 절반인 276.5명이고, 표준오차는 전체 응답인원의 제곱근에 $1/2$ 을 곱해 준 11.8 명입니다.

따라서 Red, Black 각 그룹에 관찰된 인원은 기대인원으로부터 표준오차 범위 안에 들어갑니다.

5.2 Q1. 국세와 지방세 비중

1. 2020년 우리나라 국세와 지방세의 세수 비중은? *
- 78:22
 - 77:23
 - 76:24
 - 75:25
 - 74:26

5.2.1 국세와 지방세 비중(집계표)

	78:22	77:23	76:24	75:25	74:26	계
Red	16	35	45	29	152	277
Black	16	33	55	23	149	276
계	32	68	100	52	301	553

Table 5.3: Pearson's Chi-squared test: .

Test statistic	df	P value
1.779	4	0.7763

Q1의 집계 결과가 Red, Black 간에 통계적으로 유의한 차이가 있는지 알아보기 위하여 카이제곱 테스트를 수행하였습니다.

그 결과 카이제곱 통계량은 1.78, 자유도는 4, p-value 는 0.7763이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

5.2.2 국세와 지방세 비중(%)

78:22	77:23	76:24	75:25	74:26	계
5.8	12.3	18.1	9.4	54.4	100.0

정답률은 Red, Black 을 합하여 계산하는데, 54.4(%) 입니다.

5.3 Q2. 조세부담률

2. 2020년, 우리나라의 조세부담률은 얼마인가? *
- 10%
 - 15%
 - 20%
 - 25%
 - 30%

5.3.1 조세부담률(집계표)

	10%	15%	20%	25%	30%	계
Red	6	28	223	17	3	277
Black	7	24	221	19	5	276
계	13	52	444	36	8	553

Table 5.6: Pearson's Chi-squared test: .

Test statistic	df	P value
1.003	4	0.9094

Q2의 집계 결과가 Red, Black 간에 통계적으로 유의한 차이가 있는지 알아보기 위하여 카이제곱 테스트를 수행하였습니다.

그 결과 카이제곱 통계량은 1.003, 자유도는 4, p-value 는 0.9094이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

5.3.2 조세부담률(%)

10%	15%	20%	25%	30%	계
2.4	9.4	80.3	6.5	1.4	100.0

정답률은 Red, Black 을 합하여 계산하는데, 80.3(%) 입니다.

5.4 Q3. OECD 국민부담률

3. 2013년부터 2020년까지 국민부담률에 대하여 설명한 것 중 옳지 않은 것은? *

- 조세 뿐 아니라 사회보장기여금을 포함하는 모든 강제적 납부액을 명목 GDP 대비 비율로 표시한 것이다
- 우리나라의 국민부담률은 2010년 22.4%에서 꾸준히 상승하여 2020년 27.9%에 달하였다.
- OECD평균은 2010년 31.6%에서 2020년 33.5%까지 상승하였다.
- 프랑스와 스웨덴의 국민부담률은 꾸준히 40%를 넘고 있다.
- 미국의 국민부담률은 우리나라의 국민부담률보다 항상 높은 수준을 유지하고 있다.

5.4.1 OECD 국민부담률(집계표)

조세 뿐 아니라 사회보장기여금을 포함하는 모든 강제적 납부액을 대비 비율로 표시한 것이다	2010년	OECD평균은 2010년 꾸준히 상승하여 2020년까지 상승하였다.	미국의 국민부담률은 우리나라의 국민부담률보다 항상 높은 수준을 유지하고 있다.	계
Red	11	40	16	198
Black	8	32	18	199
계	19	72	34	553

Table 5.9: Pearson's Chi-squared test: .

Test statistic	df	P value
3.062	4	0.5476

Q3의 집계 결과가 Red, Black 간에 통계적으로 유의한 차이가 있는지 알아보기 위하여 카이제곱 테스트를 수행하였습니다.

그 결과 카이제곱 통계량은 3.062, 자유도는 4, p-value 는 0.5476이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

5.4.2 OECD 국민부담률(%)

조세 뿐 아니라 사회보장기여금을 포함하는 모든 강제적 납부액을 명목 GDP 대비 비율로 표시한 것이다	우리나라의 2010년 꾸준히 상승하여 2020년 27.9%에 달하였다.	국민부담률은 OECD평균은 31.6%에서 2020년 33.5%까지 상승하였다.	프랑스와 스웨덴의 국민부담률은 꾸준히 40%를 넘고 있다.	미국의 국민부담률은 우리나라의 국민부담률보다 항상 높은 수준을 유지하고 있다.	계
3.4	13.0	6.1	5.6	71.8	100.0

정답률은 Red, Black 을 합하여 계산하는데, 71.8(%) 입니다.

5.5 Q4. 과세대상 근로소득 1,200만 원

4. 2020년 과세대상 근로소득이 1,200만원 미만인 신고인원은 어느 정도 규모인가? *

- 952만 명
- 779만 명
- 166만 명
- 38만 명

5.5.1 과세대상 근로소득 1,200만 원(집계표)

	952만 명	779만 명	166만 명	38만 명	계
Red	158	71	39	9	277
Black	150	74	47	5	276
계	308	145	86	14	553

Table 5.12: Pearson's Chi-squared test: .

Test statistic	df	P value
2.155	3	0.5408

Q4의 집계 결과가 Red, Black 간에 통계적으로 유의한 차이가 있는지 알아보기 위하여 카이제곱 테스트를 수행하였습니다.

그 결과 카이제곱 통계량은 2.155, 자유도는 3, p-value 는 0.5408이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

5.5.2 과세대상 근로소득 1,200만 원(%)

952만 명	779만 명	166만 명	38만 명	계
55.7	26.2	15.6	2.5	100.0

정답률은 Red, Black 을 합하여 계산하는데, 55.7(%) 입니다.

5.6 Q5. 소득세 실효세율

5. 2020년 우리나라 소득세의 실효세율은 얼마인가? *

- 0.2%
- 15.1%
- 37.4%
- 5.9%

5.6.1 소득세 실효세율(집계표)

	0.2%	15.1%	37.4%	5.9%	계
Red	2	58	19	198	277
Black	8	48	16	204	276
계	10	106	35	402	553

Table 5.15: Pearson's Chi-squared test: .

Test statistic	df	P value
4.888	3	0.1802

Q5의 집계 결과가 Red, Black 간에 통계적으로 유의한 차이가 있는지 알아보기 위하여 카이제곱 테스트를 수행하였습니다.

그 결과 카이제곱 통계량은 4.888, 자유도는 3, p-value 는 0.1802이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

5.6.2 소득세 실효세율(%)

0.2%	15.1%	37.4%	5.9%	계
1.8	19.2	6.3	72.7	100.0

정답률은 Red, Black 을 합하여 계산하는데, 72.7(%) 입니다.

5.7 Q6. 기업규모별 과세 현황

6. 2020년 우리나라 기업규모별 과세 현황에 대한 설명으로 옳지 않은 것은? *
- 중소기업이 신고법인수의 91%를 차지하는 데 부담하는 세액은 24.6%이다
 - 중견기업은 신고법인수의 0.5%를 차지하는 데 부담하는 세액은 8.7%이다
 - 상호출자제한기업은 신고법인수의 0.1%를 차지하는 데 부담하는 세액은 25.5%이다
 - 그 외 대기업은 신고법인수의 8.3%를 차지하는 데 부담하는 세액은 41.2%이다
 - 법인당 평균세액은 상호출자제한기업 > 그외 대기업 > 중견기업 > 중소기업의 순서로 많이 부담하였다.

5.7.1 기업규모별 과세 현황(집계표)

						그 외	법인당
중소기업이 신고법인수의 91%를 차지하는 데 부담하는 세액은 24.6%이다	중견기업은 신고법인수의 0.5%를 차지하는 데 부담하는 세액은 8.7%이다	상호출자제한기업은 신고법인수의 0.1%를 차지하는 데 부담하는 세액은 25.5%이다	대기업은 신고법인수의 8.3%를 차지하는 데 부담하는 세액은 41.2%이다	평균세액은 > 그외 대기업 > 중견기업 > 중소기업의 순서로 많이 부담하였다.			
Red	14	37	38	50	138	277	
Black	18	33	21	61	143	276	
계	32	70	59	111	281	553	

Table 5.18: Pearson's Chi-squared test: .

Test statistic	df	P value
6.804	4	0.1466

Q6의 집계 결과가 Red, Black 간에 통계적으로 유의한 차이가 있는지 알아보기 위하여 카이제곱 테스트를 수행하였습니다.

그 결과 카이제곱 통계량은 6.804, 자유도는 4, p-value 는 0.1466이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

5.7.2 기업규모별 과세 현황(%)

						법인당	
						그 외	평균세액은
중소기업이 신고법인수의 91%를 차지하는 데 부담하는 세액은 24.6%이다	중견기업은 신고법인수의 0.5%를 차지하는 데 부담하는 세액은 8.7%이다	상호출자제한기업은 신고법인수의 0.1%를 차지하는 데 부담하는 세액은 25.5%이다	대기업은 신고법인수의 8.3%를 차지하는 데 부담하는 세액은 41.2%이다	평균세액은 > 그외 대기업 > 중견기업 > 중소기업의 순서로 많이 부담하였다.			
5.8	12.7	10.7	20.1	50.8	100.0		

정답률은 Red, Black 을 합하여 계산하는데, 50.8(%) 입니다.

5.8 Q7. 국민부담률 적정 수준 : 아일랜드와 OECD 평균

질문 내용에 의도하는 바를 담으면 어떨까요?

OECD 국가 중 국민부담률이 매우 낮은 편인 아일랜드의 사례를 들어서
감세정책이 가져온 긍정적인 효과에 대해서 설명하고 우리나라의 바람직한 조정
방향은 무엇이냐고 묻는 것을 Red, 감세 정책이 가져온 부정적인 효과에 대해서
설명하고 우리나라의 바람직한 조정 방향은 무엇이냐고 묻는 것을 Black에
배치했을 때, 설명이 응답에 영향을 미치지 않으면 Red와 Black에 차이가 없어야
할텐데 집계결과는 어떻게 나오고 있나요?

분명히 영향을 미치고 있는 것으로 보입니다.

통계적으로 매우 유의한 차이가 관찰되고 있습니다.

감세정책의 효과가 긍정적이라고 설명한 Red에서는 낮춰야 한다는 응답이,
감세정책의 효과가 부정적이라고 설명한 Black에서는 높여야 한다는 응답이 높게
나온 것을 볼 수 있고, 따라서 p-value 가 엄청나게 작은 값을 보여주고 있습니다.

Red
<p>7. (chatGPT 의 답변을 토대로 작성) 2019년 기준 아일랜드의 국민부담률은 23.7%로 OECD 회원국 평균 33.6%보다 매우 낮은 수준입니다. 아일랜드의 감세 정책은 대기업들을 유치하여 경제 성장을 촉진하는 데에 도움을 준 것으로 평가합니다. 2022년 아일랜드의 1인당 명목상 GDP는 10만 1,509달러로 세계 2위 수준입니다. 우리나라의 2020년 국민부담률이 27.9%인 점에 비추어 어떤 정책 방향을 추구해야 하겠습니까?</p> <ul style="list-style-type: none"> <input type="radio"/> 국민부담률을 아일랜드 수준으로 낮춰야 한다 <input type="radio"/> 국민부담률을 OECD 평균 수준으로 높여야 한다 <input type="radio"/> 모름/무응답
Black
<p>7. (chatGPT 의 답변을 토대로 작성) 2019년 기준 아일랜드의 국민부담률은 23.7%로 OECD 회원국 평균 33.6%보다 매우 낮은 수준입니다. 아일랜드의 감세 정책은 빈부 격차를 확대하고 불평등을 줄이는 데 도움이 되는 사회 복지 프로그램 및 기타 정책에 자금을 지원하는 정부의 능력을 제한하였습니다. 그 결과 상위 10%의 가구가 전체 부의 54%를 점유할 정도로 불평등이 심화되었습니다. 우리나라의 2020년 국민부담률이 27.9%인 점에 비추어 어떤 정책 방향을 추구해야 하겠습니까?</p> <ul style="list-style-type: none"> <input type="radio"/> 국민부담률을 OECD 평균 수준으로 높여야 한다 <input type="radio"/> 국민부담률을 아일랜드 수준으로 낮춰야 한다 <input type="radio"/> 모름/무응답

5.8.1 집계표

	낮춰야 한다	높여야 한다	모름/무응답	계
Red(감세의 긍정적효과 설명)	90	120	67	277
Black(감세의 부정적 효과 설명)	46	168	62	276
계	136	288	129	553

Table 5.21: Pearson's Chi-squared test: .

Test statistic	df	P value
22.43	2	1.349e-05 * * *

Q7의 Red에는 아일랜드의 사례에서 감세 정책의 긍정적 측면을 설명한 후 우리나라 조세 정책의 방향에 대하여 물었을 때, 277명이 응답한 가운데 90명이 우리나라의 국민부담률을 아일랜드 수준으로 “낮춰야 한다”는 반응을 보이고, 120명이 OECD평균 수준으로 “높여야 한다”는 반응을 보입니다.

Black은 같은 아일랜드의 사례에서 감세 정책의 부정적 측면을 설명한 후 우리나라 조세 정책의 방향에 대하여 물었을 때, 276명이 응답한 가운데 46명이 우리나라의 국민부담률을 아일랜드 수준으로 “낮춰야 한다”는 반응을 보이고, 168명이 OECD 평균 수준으로 “높여야 한다”는 반응을 보입니다.

그리고 “모름/무응답”에 답한 인원은 Red에 67명, Black에 62명이 응답하였습니다.

우연일까요?

모름/무응답에 있어서는 Red, Black이 몹시 닮았습니다.

카이제곱 테스트는 이와 같은 상황에서 감세정책의 긍정적 측면을 부각시킨 경우와 부정적 측면을 부각시킨 경우에 그 차이가 통계적으로 매우, 매우, … 유의하다는 것을 보여 줍니다.

카이제곱 통계량은 22.427, 자유도는 2, p-value 는 1.3e-05으로 감세정책의 어떤 측면을 설명하느냐에 따라 반응이 다르게 나온다는 것을 보여줍니다.

여기서 질문 내용에 의도하는 바를 담더라도 응답에 영향을 끼치지 않는다고 가정합니다.

랜덤화의 효과로 Red, Black 의 응답은 닮게 마련입니다.

즉, 통계적으로 유의한 차이를 보이지 않게 됩니다.

그러나 실제로 관찰된 카이제곱 통계값의 P-value 는 0.05보다 매우 작은 값입니다.

따라서, 질문 내용에 의도하는 바를 담더라도 영향을 끼치지 않는다는 가정은 잘못된 것이죠.

이러한 논증 방식을 귀류법이라고 합니다.

5.8.2 % 비교

	낮춰야 한다	높여야 한다	모름/무응답	계
Red(감세의 긍정적 효과 설명)	32.5	43.3	24.2	100.0
Black(감세의 부정적 효과 설명)	16.7	60.9	22.5	100.0

감세정책의 긍정적 측면을 설명한 Red에서 우리나라의 국민부담률을 “낮춰야 한다”고 응답하는 사람들의 백분율, 32.5(%)은 “높여야 한다”고 응답하는 사람들의 백분율, 43.3(%) 보다 높습니다.

반면 감세정책의 부정적 측면을 설명한 Black에서 우리나라의 국민 부담률을 “낮춰야 한다”고 응답하는 사람들의 백분율, 16.7(%)은 “높여야 한다”고 응답하는 사람들의 백분율, 60.9(%) 보다 훨씬 적습니다.

어느 정책의 긍정적 측면을 설명하느냐, 부정적 측면을 설명하느냐에 따라 반응이 달라진다는 것을 잘 알 수 있습니다.

Red 와 Black 이 워낙 차이가 나지만 전체적으로 어느 정도가 우리나라의 국민부담률을 “낮춰야 한다”하고 어느 정도가 “높여야 한다”고 응답하였는지 합쳐 보겠습니다.

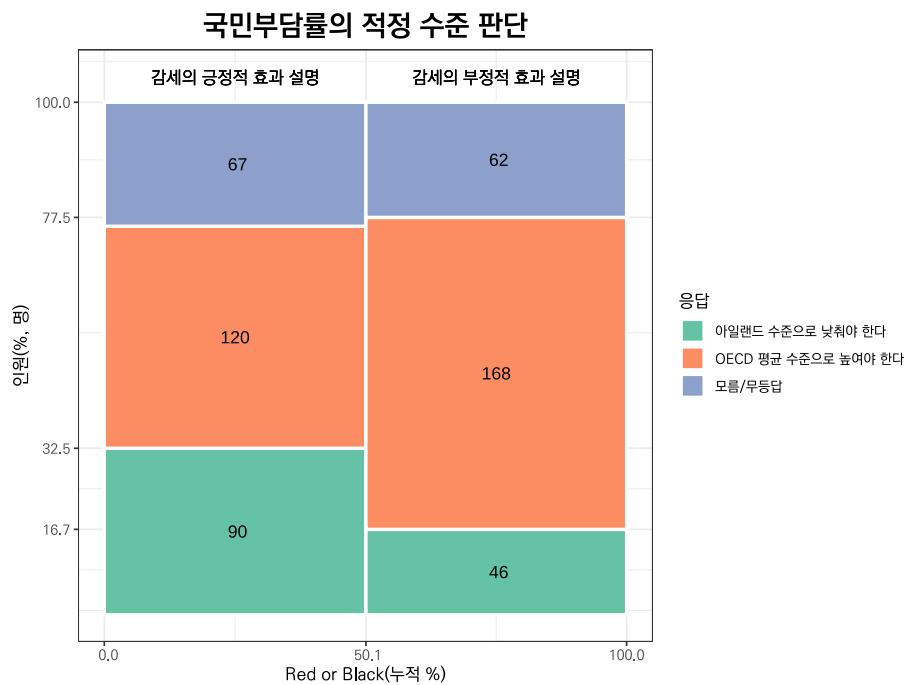
5.8.3 % 합계

낮춰야 한다	높여야 한다	모름/무응답	계
24.6	52.1	23.3	100.0

우리나라의 국민부담률을 “낮춰야 한다”고 응답한 백분율은 Red, Black 합쳐서 24.6%(으)로 우리나라의 국민부담률을 ’높여야한다”고 응답한 백분율, 52.1(%) 보다 상당히 적습니다.

다만, 모름/무응답이 23.3(%)로 적지 않습니다.

5.8.4 Mosaic Plot



Mosaic Plot 은 이 집계결과를 시각적으로 잘 보여줍니다.

감세정책의 긍정적 측면을 설명한 Red 에서 우리나라의 국민부담률을 “낮춰야 한다”고 응답한 백분율이 높고, 감세정책의 부정적 측면을 설명한 Black 에서 우리나라의 국민부담률을 “높여야 한다”고 응답한 백분율이 월등히 높은 것을 시각적으로 알 수 있습니다.

5.9 마감 시간으로부터 제출 시간의 분포

5.9.1 분포표

Table 5.24: 일 단위

	[0,1]	(1,2]	(2,3]	(3,4]	(4,5]	(5,6]	(6,7]	(7,8]	(8,9]	(9,10]	(10,11]	(11,12]	(12,13]	(13,14]	
Red	71	12	6	8	4	6	16	48	20	14	18	17	16	21	277
Black	64	14	3	6	10	4	4	47	30	18	12	17	22	25	276
계	135	26	9	14	14	10	20	95	50	32	30	34	38	46	553

분포표로부터 두 가지 문제를 살펴보겠습니다.

첫째, 날마다 고르게 제출하는가?

둘째, Red, Black 간에 통계적으로 유의한 차이가 있는가?

각 문제를 살펴보기 위해서는 분포표의 일부분을 대상으로 카이제곱 테스트를 수행합니다.

5.9.2 날마다 고르게 제출하는가?

[0,1]	(1,2]	(2,3]	(3,4]	(4,5]	(5,6]	(6,7]	(7,8]	(8,9]	(9,10]	(10,11]	(11,12]	(12,13]	(13,14]
135	26	9	14	14	10	20	95	50	32	30	34	38	46

Table 5.26: Chi-squared test for given probabilities: .

Test statistic	df	P value
410	13	1.714e-79 ***

날마다 고르게 제출하는지 알아 보았습니다.

분포표의 “계”행에서 ’계’열을 제외하고 카이제곱테스트를 수행합니다.

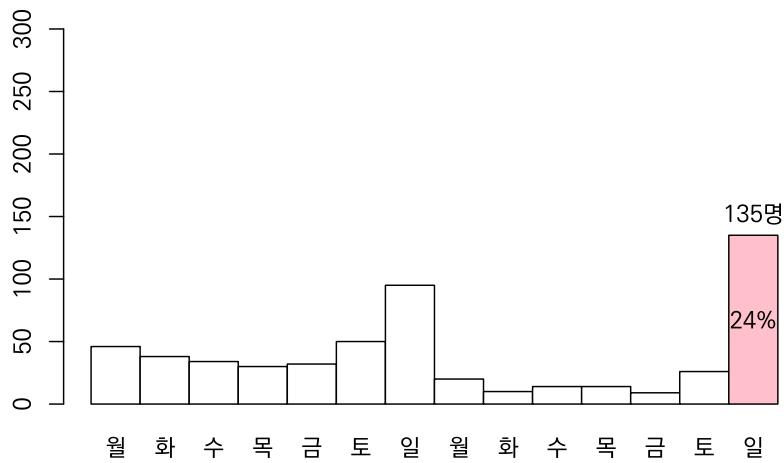
분포표 만으로도 쉽게 파악할 수 있지만 카이제곱테스트가 명확히 해 줍니다.

카이제곱 통계량은 410.01, 자유도는 13.00, p-value 는 1.7e-79 이므로 결코 고르게 제출한다고 말할 수 없겠습니다.

막대그래프로 살펴 보겠습니다.

5.9.3 막대그래프

Quiz250317 (553명 제출)



막대그래프는 총 제출인원 553(명) 중에 135(명), 24(%)가 마감일에 몰리는 것을 명확히 보여주고 있습니다.

5.9.4 Red, Black 간에 닮았는가?

	[0,1](1,2](2,3](3,4](4,5](5,6](6,7](7,8](8,9](9,10](10,11](11,12](12,13](13,14]													
Red	71	12	6	8	4	6	16	48	20	14	18	17	16	21
Black	64	14	3	6	10	4	4	47	30	18	12	17	22	25

Table 5.28: Pearson's Chi-squared test: .

Test statistic	df	P value
16.98	13	0.2003

제출시간의 분포가 Red, Black 간에 닮았는지 알아 보았습니다.

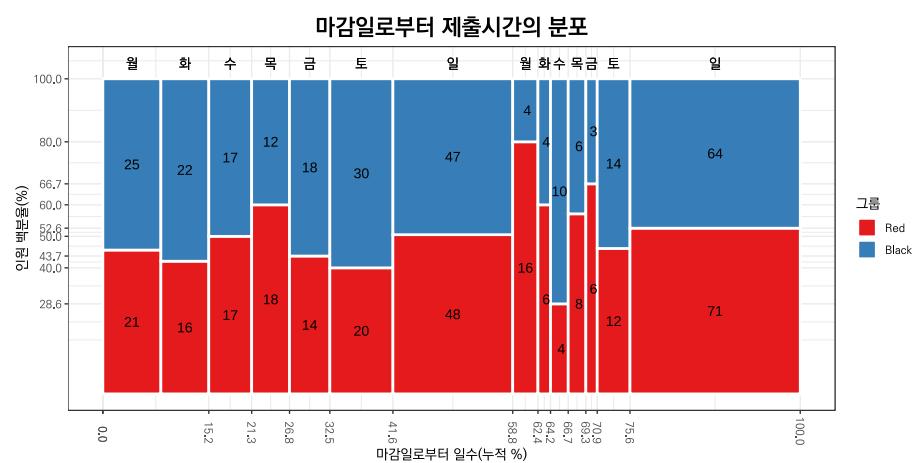
이번에는 분포표의 첫번째와 두번째 행, '계'열을 제외한 나머지 열에 대해서 카이제곱테스트를 수행합니다. 카이제곱 통계량은 16.978, 자유도는 13, p-value

는 0.2003 이므로 제출 시간의 분포는 Red, Black 간에 통계적으로 유의한 차이가 관찰되지 않습니다.

이 사실을 Mosaic Plot을 이용하여 시각적으로 살펴보겠습니다.

답았다고 느껴지나요?

5.9.5 Mosaic Plot



Chapter 6

4주차 데이터 실험 집계

6.1 실험의 목적

4주차 구글 예습 설문지 집계결과를 분석합니다.

Q1~Q6에서는 랜덤화의 효과로 Red, Black 이 얼마나 닮았는지 알아봅니다.

Q7에서는 부연설명을 어느 쪽에 붙이느냐에 따라서 Red 와 Black 의 응답이 달라지는 것을 알아봅니다.

끝으로 제출시간의 분포가 날마다 고른지, Red, Black 간에는 닮았는지 알아봅니다.

6.1.1 Red, Black을 잘못 표시한 사람들

	Red(구글예습퀴즈)	Black(구글예습퀴즈)
Red(랜덤화출석부)	275	1
Black(랜덤화출석부)	3	285
계	278	286

랜덤화출석부에 있는 Red, Black 과 실제 구글설문에 올린 Red, Black 이 다른 사람들의 수효는 4명입니다.

Red를 Black 이라고 한 사람이 1명, Black 을 Red 라고 한 사람이 3명입니다.

두 가지 방법으로 분석합니다.

우선 Red, Black 을 잘못 선택한 4명을 랜덤하게 둘로 나누면 어느 한 쪽 집단에 들어갈 기대인원은 4명을 둘로 나눈 2(명)이고, 표준오차는 4의 제곱근에 1/2을 곱해 준 1명이 됩니다.

실제로 Red를 Black 이라고 한 사람수, 1명이나 Black 을 Red 라고 한 사람수, 3명은 기대인원으로부터 표준오차 범위, 혹은 표준오차 두 배 범위에는 잘 들어갑니다.

두 번째 분석 방법은 확률을 계산해 보는 것입니다.

Red, Black 을 잘못 선택한 4명을 랜덤하게 둘로 나눌 때, 실제로 관찰된 3명 이상이나 1명이하로 잘못 선택한 사람수가 나올 가능성은 얼마나 되는가 입니다.

이 경우 공평한 동전던지기를 확률 법칙으로 표현한 이항분포로부터 계산할 수 있습니다.

시행횟수가 4이고 한 번 시행에서 성공확률이 $1/2$ 인 이항분포에서 성공횟수가 1이하이거나 3이상을 관찰할 확률은 0.625입니다.

공평한 동전 던지기에서 앞면이 1개 이하 나오는 확률은 3개 이상 나오는 확률과 같기 때문에 사실상 한쪽만 계산해서 2배 해 주면 됩니다.

이 값을 p-value 라고 하는데, p-value가 0.05보다 작을 때 통계적으로 유의한 차이를 관찰하였다고 말합니다.

즉, 공평한 동전을 던지는 것과 같은 과정이라고 가정하였을 때 실제로 관찰된 값들이 가정으로부터 얼마나 떨어져 있는지를 표현한 것입니다.

0.05는 이런 실험을 스무 번 정도 반복하면 1번 나올 정도로 드문 사건을 의미합니다.

즉 가정이 잘못되었다는 것입니다.

그런데 Red, Black 을 잘못 표시한 사람들의 분포에서 관찰된 p-value 는 0.05와는 비교도 안될 정도로 큰 값입니다.

따라서 두 집단이 랜덤화 효과가 작동하여 통계적으로 유의한 차이를 보이지 않는다고 할 수 있습니다.

6.1.2 응답인원의 Red, Black

Red 로 응답한 인원은 278명, Black 에 응답한 인원은 286명입니다.

전체 응답인원 564 명을 랜덤하게 둘로 나눌 때 어느 한 쪽의 기대인원은 전체 응답인원의 절반인 282명이고, 표준오차는 전체 응답인원의 제곱근에 $1/2$ 을 곱해 준 11.9 명입니다.

따라서 Red, Black 각 그룹에 관찰된 인원은 기대인원으로부터 표준오차 범위, 혹은 2배의 표준오차 범위 안에 들어갑니다.

6.2 Q1. 세종대왕 시대 조세제도

1. 세종대왕 즉위 당시의 조세제도에 대한 설명으로 적절치 않은 것은? *
- 과전법 체제에서 전국 토지를 세 등급으로 나누고 실제 수확량을 확인하여 징수하였다.
 - '조(租)'는 공전(公田)의 경작자가 국고에 상납하는 지대 또는 사전(私田)의 경작자가 전주에게 바치는 지대를 뜻한다.
 - '세(稅)'는 사전의 소유자가 국가에 상납하는 지대를 뜻한다.
 - 실제로 현장에 나가서 수확량을 파악하고 등급을 매기는 담험(踏驗)을 하였다.
 - 등급에 따라 일정한 비율로 세금을 감면해 주는 대동(大同)법을 실시하였다.

6.2.1 조선초기 조세제도

'조(租)'는 공전(公田)의 경작자가 국고에 상납하는 지대 또는 사전(私田)의 경작자가 전주에게 바치는 지대를 뜻한다.					
과전법 체제에서 전국 토지를 세 등급으로 나누고 실제 수확량을 확인하여 징수하였다.	실제로 현장에 나가서 수확량을 파악하고 등급을 매기는 답험(踏驗)을 하였다.	등급에 따라 일정한 비율로 세금을 감면해 주는 대동(大同)법을 실시하였다. 계			
Red	22	28	26	22	180
Black	17	30	17	19	203
계	39	58	43	41	564

Table 6.3: Pearson's Chi-squared test: .

Test statistic	df	P value
4.082	4	0.3951

Q1의 집계 결과가 Red, Black 간에 통계적으로 유의한 차이가 있는지 알아보기 위하여 카이제곱 테스트를 수행하였습니다.

그 결과 카이제곱 통계량은 4.08, 자유도는 4, p-value는 0.3951이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

6.2.2 조선초기 조세제도(%)

'조(租)'는 공전(公田)의 경작자가 국고에 상납하는 지대 또는 사전(私田)의 경작자가 전주에게 바치는 지대를 뜻한다.	'세(稅)'는 사전의 소유자가 국가에 상납하는 지세를 뜻한다.	실제로 현장에 나가서 수확량을 파악하고 등급을 매기는 답험(踏驗)을 하였다.	등급에 따라 일정한 비율로 세금을 감면해 주는 대동(大同)법을 실시하였다.	계
6.9	10.3	7.6	7.3	67.9 100.0

정답률은 Red, Black 을 합하여 계산하는데, 67.9(%) 입니다.

6.3 Q2. 공법도입에 대한 대신들의 찬성율

2. 공법 도입의 가부를 묻는 국민투표에서 대신들은 몇% 정도 찬성하였는가? *

- 10%
- 30%
- 50%
- 70%
- 90%

6.3.1 공법도입과 대신들(집계표)

	10%	30%	50%	70%	90%	계
Red	153	54	28	27	16	278
Black	160	46	26	32	22	286
계	313	100	54	59	38	564

Table 6.6: Pearson's Chi-squared test: .

Test statistic	df	P value
2.129	4	0.7121

Q2의 집계 결과가 Red, Black 간에 통계적으로 유의한 차이가 있는지 알아보기 위하여 카이제곱 테스트를 수행하였습니다.

그 결과 카이제곱 통계량은 2.129, 자유도는 4, p-value 는 0.7121이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

6.3.2 공법도입과 대신들(%)

10%	30%	50%	70%	90%	계
55.5	17.7	9.6	10.5	6.7	100.0

정답률은 Red, Black 을 합하여 계산하는데, 55.5(%) 입니다.

6.4 Q3. 공법도입과 품관총민들의 찬반

3. 공법 도입의 가부를 묻는 국민투표에서 품관총민의 찬반은 대략 몇 대 몇인가? *

- 7:3
- 6:4
- 5:5
- 4:6
- 3:7

6.4.1 품관총민들의 찬반(집계표)

	7:3	6:4	5:5	4:6	3:7	계
Red	48	159	25	22	24	278
Black	62	162	25	18	19	286

	7:3	6:4	5:5	4:6	3:7	계
계	110	321	50	40	43	564

Table 6.9: Pearson's Chi-squared test: .

Test statistic	df	P value
2.678	4	0.613

Q3의 집계 결과가 Red, Black 간에 통계적으로 유의한 차이가 있는지 알아보기 위하여 카이제곱 테스트를 수행하였습니다.

그 결과 카이제곱 통계량은 2.678, 자유도는 4, p-value 는 0.6130이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

6.4.2 품관총민들의 찬반(%)

7:3	6:4	5:5	4:6	3:7	계
19.5	56.9	8.9	7.1	7.6	100.0

정답률은 Red, Black 을 합하여 계산하는데, 56.9(%) 입니다.

6.5 Q4. 공법

4. 공법의 기본세율은? *
<input type="radio"/> 10분의 1
<input type="radio"/> 15분의 1
<input type="radio"/> 20분의 1
<input type="radio"/> 30분의 1

6.5.1 기본세율

	1/10	1/15	1/20	1/30	계
Red	84	47	132	15	278
Black	88	29	154	15	286
계	172	76	286	30	564

Table 6.12: Pearson's Chi-squared test: .

Test statistic	df	P value
5.936	3	0.1148

Q4의 집계 결과가 Red, Black 간에 통계적으로 유의한 차이가 있는지 알아보기 위하여 카이제곱 테스트를 수행하였습니다.

그 결과 카이제곱 통계량은 5.936, 자유도는 3, p-value 는 0.1148이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

6.5.2 기본세율(%)

1/10	1/15	1/20	1/30	계
30.5	13.5	50.7	5.3	100.0

정답률은 Red, Black 을 합하여 계산하는데, 30.5(%) 입니다.

6.6 Q5. 1423년 조선시대 호구와 인구

5. 1422년 세종실록지리지에 실린 조선시대의 호수와 인구는? *

- 15만호, 32만명
- 20만호, 69만명
- 44만호, 153만명
- 130만호, 5백만명

6.6.1 호구와 인구

	15만호	20만호	44만호	130만호	계
Red	11	164	87	16	278
Black	22	151	105	8	286
계	33	315	192	24	564

Table 6.15: Pearson's Chi-squared test: .

Test statistic	df	P value
8.446	3	0.03765 *

Q5의 집계 결과가 Red, Black 간에 통계적으로 유의한 차이가 있는지 알아보기 위하여 카이제곱 테스트를 수행하였습니다.

그 결과 카이제곱 통계량은 8.446, 자유도는 3, p-value 는 0.03760이므로 Red, Black 간에 통계적으로 유의한 차이를 보이고 있습니다.

6.6.2 호구와 인구(%)

15만호	20만호	44만호	130만호	계
5.9	55.9	34.0	4.3	100.0

정답률은 Red, Black 을 합하여 계산하는데, 55.9(%) 입니다.

6.7 Q6. 지방관료와 품관총민

6. 공법 도입의 가부를 묻는 국민투표에서 지방 관료들의 찬반 동향과 품관총민의 찬반 동향이 서로 반대인 곳은 *
다음 중 어디인가?

- 경기
- 평안
- 전라
- 경상
- 충청

6.7.1 찬반이 반대인 곳(집계표)

	경기	평안	전라	경상	충청	계
Red	25	39	44	40	130	278
Black	25	37	51	43	130	286
계	50	76	95	83	260	564

Table 6.18: Pearson's Chi-squared test: .

Test statistic	df	P value
0.5635	4	0.967

Q6의 집계 결과가 Red, Black 간에 통계적으로 유의한 차이가 있는지 알아보기 위하여 카이제곱 테스트를 수행하였습니다.

그 결과 카이제곱 통계량은 0.563, 자유도는 4, p-value 는 0.9670이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

6.7.2 찬반이 반대인 곳(%)

경기	평안	전라	경상	충청	계
8.9	13.5	16.8	14.7	46.1	100.0

정답률은 Red, Black 을 합하여 계산하는데, 46.1(%) 입니다.

6.8 Q7. 부연설명의 효과 : 주당 근로 69시간제 도입 찬반

부연설명을 찬성 쪽에 불이는가(Red), 또는 반대 쪽에 불이는가(Black)에 따라 응답이 영향을 받는 것으로 관찰됩니다.

찬반여부에 대한 카이제곱테스트의 p-value를 놓고 볼 때 그 차이가 통계적으로 매우 유의합니다.

Red
<p>7. (23/03/16 NBS 선택지 수정) 정부가 현행 주 52시간으로 제한한 근로시간을 최대 69시간까지 허용하되, 그만큼 다른 주의 연장 근로시간을 줄이는 방안을 추진 중인데요. 선생님께서는 이에 대해 어떻게 생각하십니까?</p> <p><input type="radio"/> 찬성한다. <input type="radio"/> 반대한다. <input type="radio"/> 모름/무응답</p>
Black
<p>7. (23/03/16 NBS 선택지 수정) 정부가 현행 주 52시간으로 제한한 근로시간을 최대 69시간까지 허용하되, 그만큼 다른 주의 연장 근로시간을 줄이는 방안을 추진 중인데요. 선생님께서는 이에 대해 어떻게 생각하십니까?</p> <p><input type="radio"/> 찬성한다. <input type="radio"/> 노동자가 과도한 연장근로를 받을 수 있어 반대한다. <input type="radio"/> 모름/무응답</p>

6.8.1 집계

	찬성한다.	반대한다.	모름/무응답	계
Red(찬성한다에 부연설명)	132	78	68	278
Black(반대한다에 부연설명)	85	146	55	286
계	217	224	123	564

Table 6.21: Pearson's Chi-squared test: .

Test statistic	df	P value
32.09	2	1.076e-07 * * *

Q7의 Red는 주당 근로 69시간제의 도입에 찬반을 묻는 질문 중 찬성을 유도하는 부연설명을 붙였을 때 278명이 응답한 가운데 132명이 “찬성한다”는 반응을 보이고, 78명이 “반대한다”는 반응을 보입니다.

Black은 같은 상황에서 반대를 유도하는 부연설명을 붙였을 때 286명이 응답한 가운데 85명이 “찬성한다”는 반응을 보이고, 146명이 “반대한다”는 반응을 보입니다.

그리고 “모름/무응답”에 답한 인원은 Red에 68명, Black에 55명이 응답하였습니다.

카이제곱 테스트는 이와 같은 상황에서 찬성을 유도하는 부연설명을 붙인 경우와 반대를 유도하는 부연설명을 붙인 경우에 그 응답의 차이가 통계적으로 유의하다는 것을 보여 줍니다.

카이제곱 통계량은 32.090, 자유도는 2, p-value 는 1.1e-07으로 부연설명을 어디에 붙이느냐에 따라 그 차이가 통계적으로 유의한 것으로 나왔습니다.

여기서 부연설명이 응답에 영향을 끼치지 않는다고 가정해 봅시다.

그렇다면 Red, Black 의 응답은 대부분의 Q1 ~ Q6 에서와 같이 랜덤화 효과에 의하여 통계적으로 유의한 차이를 보이지 않을 것입니다.

그런데 실제로 관찰된 카이제곱 통계값은 통계적으로 유의한 차이를 보여 주고 있습니다.

따라서 부연설명이 영향을 끼치지 않는다는 가정은 티당치 않은 것으로 볼 수밖에 없습니다.

이러한 논증 방식을 귀류법이라 합니다.

6.8.2 % 비교

	찬성한다.	반대한다.	모름/무응답	계
Red(찬성한다에 부연설명)	47.5	28.1	24.5	100.0
Black(반대한다에 부연설명)	29.7	51.0	19.2	100.0

찬성을 유도하는 부연설명을 붙인 Red에서 “찬성한다”고 응답하는 사람들의 백분율, 47.5(%)은 “반대한다”고 응답하는 사람들의 백분율, 28.1(%) 보다 높습니다.

반면 반대를 유도하는 부연설명을 붙인 Black에서 “찬성한다”고 응답하는 사람들의 백분율, 29.7(%)은 “반대한다”고 응답하는 사람들의 백분율, 51.0(%) 보다 훨씬 적습니다.

찬성을 유도하는 부연설명을 붙이느냐, 반대를 유도하는 부연설명을 붙이느냐에 따라 반응이 달라진다는 것을 잘 알 수 있습니다.

Red 와 Black 이 워낙 차이가 나지만 전체적으로 어느 정도가 “찬성한다”하고 어느 정도가 “반대한다”고 응답하였는지 합쳐 보겠습니다.

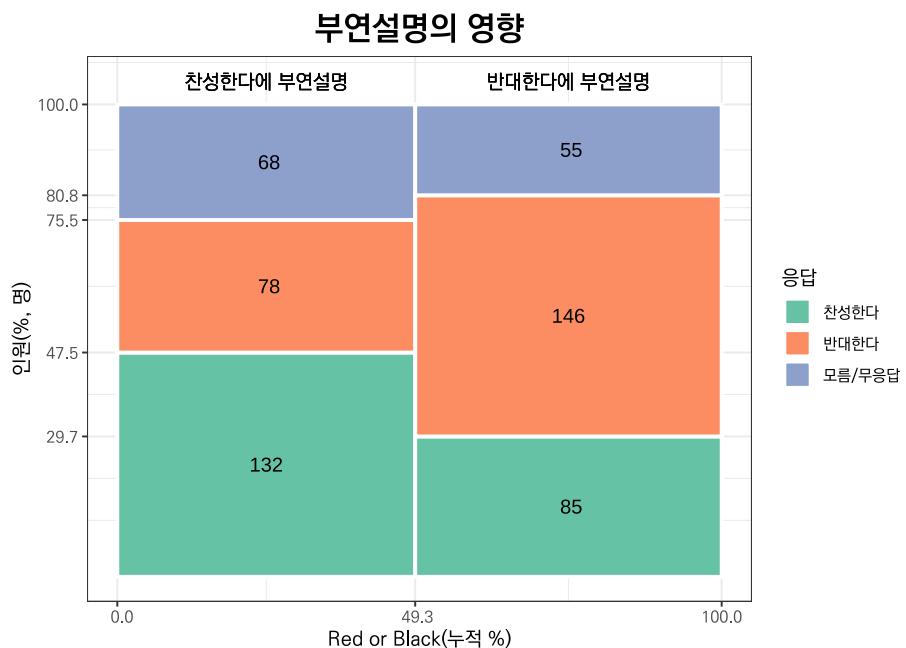
6.8.3 % 합계

찬성한다.	반대한다.	모름/무응답	계
38.5	39.7	21.8	100.0

“찬성한다”고 응답한 백분율은 Red, Black 합쳐서 38.5%(으)로 ‘반대한다’고 응답한 백분율, 39.7% 과 약간 적습니다.

그리고, 모름/무응답이 21.8%로 적지 않습니다.

6.8.4 Mosaic Plot



Mosaic Plot 은 이 집계결과를 시각적으로 잘 보여줍니다.

찬성을 유도하는 부연설명을 붙인 Red 에서 “찬성한다”고 응답한 백분율이 높고, 반대를 유도하는 부연설명을 붙인 Black 에서 “반대한다”고 응답한 백분율이 월등히 높은 것을 시각적으로 알 수 있습니다.

6.9 마감 시간으로부터 제출 시간의 분포

6.9.1 분포표

Table 6.24: 일 단위

	[0,1]	(1,2]	(2,3]	(3,4]	(4,5]	(5,6]	(6,7]	(7,8]	(8,9]	(9,10]	[10,11]	[11,12]	[12,13]	[13,14]	
Red	71	15	13	4	10	3	6	46	20	22	14	20	16	18	278

	[0,1]	(1,2]	(2,3]	(3,4]	(4,5]	(5,6]	(6,7]	(7,8]	(8,9]	(9,10]	(10,11]	(11,12]	(12,13]	(13,14]	(14,15]
Black	71	14	7	7	11	4	6	50	20	14	12	24	17	29	286
계	142	29	20	11	21	7	12	96	40	36	26	44	33	47	564

분포표로부터 두 가지 문제를 살펴보겠습니다.

첫째, 날마다 고르게 제출하는가?

둘째, Red, Black 간에 통계적으로 유의한 차이가 있는가?

각 문제를 살펴보기 위해서는 분포표의 일부분을 대상으로 카이제곱 테스트를 수행합니다.

6.9.2 날마다 고르게 제출하는가?

[0,1]	(1,2]	(2,3]	(3,4]	(4,5]	(5,6]	(6,7]	(7,8]	(8,9]	(9,10]	(10,11]	(11,12]	(12,13]	(13,14]	(14,15]
142	29	20	11	21	7	12	96	40	36	26	44	33	47	

Table 6.26: Chi-squared test for given probabilities: .

Test statistic	df	P value
433.4	13	1.915e-84 * * *

날마다 고르게 제출하는지 알아 보았습니다.

분포표의 “계”행에서 ’계’열을 제외하고 카이제곱테스트를 수행합니다.

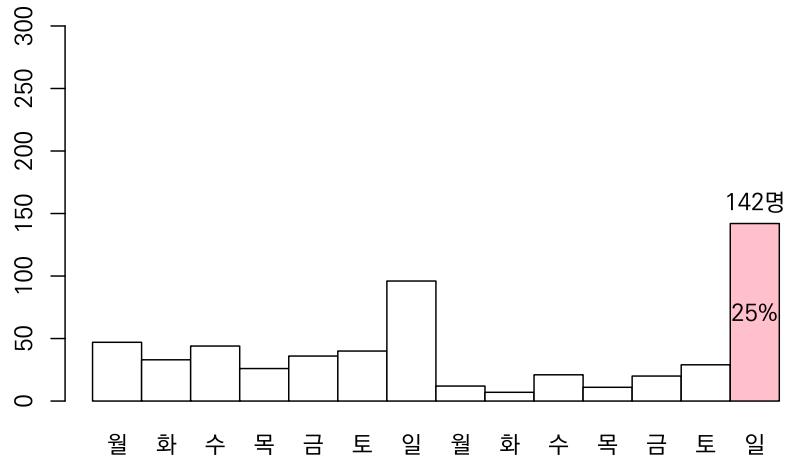
분포표 만으로도 쉽게 파악할 수 있지만 카이제곱테스트가 명확히 해 줍니다.

카이제곱 통계량은 433.43, 자유도는 13.00, p-value 는 1.9e-84 이므로 결코 고르게 제출한다고 말할 수 없습니다.

막대그래프로 살펴 보겠습니다.

6.9.3 막대그래프

Quiz250324 (564명 제출)



막대그래프는 총 제출인원 564(명) 중에 142(명), 25(%)가 마감일에 몰리는 것을 명확히 보여주고 있습니다.

6.9.4 Red, Black 간에 닮았는가?

	[0,1](1,2)(2,3)(3,4)(4,5)(5,6)(6,7](7,8](8,9](9,10]10,11]11,12]12,13]13,14]													
Red	71	15	13	4	10	3	6	46	20	22	14	20	16	18
Black	71	14	7	7	11	4	6	50	20	14	12	24	17	29

Table 6.28: Pearson's Chi-squared test: .

Test statistic	df	P value
7.798	13	0.8565

제출시간의 분포가 Red, Black 간에 닮았는지 알아 보았습니다.

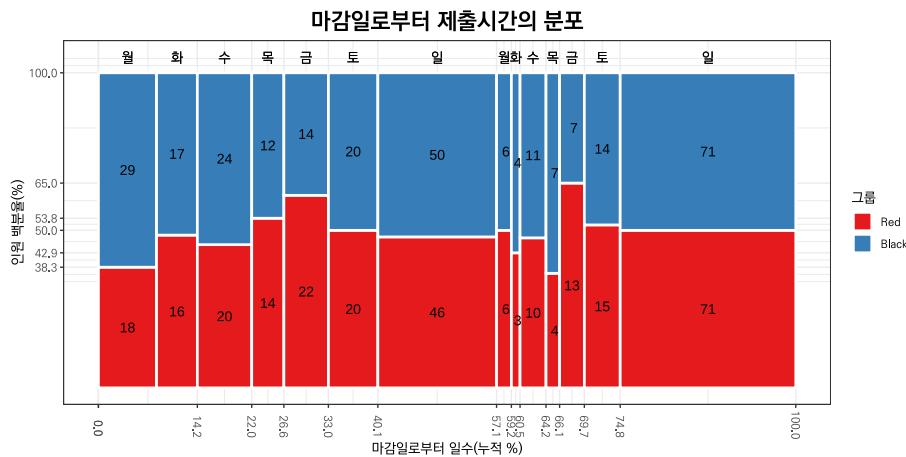
이번에는 분포표의 첫번째와 두번째 행, '계'열을 제외한 나머지 열에 대해서 카이제곱테스트를 수행합니다.

카이제곱 통계량은 7.798, 자유도는 13, p-value 는 0.8565 이므로 제출 시간의 분포는 Red, Black 간에 통계적으로 유의한 차이가 관찰되지 않습니다.

이 사실을 Mosaic Plot 을 이용하여 시각적으로 살펴보겠습니다.

닮았다고 느껴지나요?

6.9.5 Mosaic Plot



Chapter 7

5주차 데이터 실험 집계

7.1 실험의 목적

5주차 구글 예습 설문지 집계결과를 분석합니다.

Q1~Q6에서는 랜덤화의 효과로 Red, Black 이 얼마나 닮았는지 알아봅니다.

Q7에서는 컵에 우유를 반을 쏟은 상황에서 긍정적인 단어를 읽은 그룹하고 부정적인 단어를 읽은 그룹 사이에 어떠한 인식의 차이가 발생하는지 살펴봅니다.

끝으로 제출시간의 분포가 날마다 고른지, Red, Black 간에는 닮았는지 알아봅니다.

7.1.1 Red, Black을 잘못 표시한 사람들

	Red(구글예습퀴즈)	Black(구글예습퀴즈)
Red(랜덤화출석부)	274	1
Black(랜덤화출석부)	3	280
계	277	281

랜덤화출석부에 있는 Red, Black 과 실제 구글설문에 올린 Red, Black 이 다른 사람들의 수효는 4명입니다.

Red를 Black 이라고 한 사람이 1명, Black 을 Red 라고 한 사람이 3명입니다.

두 가지 방법으로 분석합니다.

우선 Red, Black 을 잘못 선택한 4명을 랜덤하게 둘로 나누면 어느 한 쪽 집단에 들어갈 기대인원은 4명을 둘로 나눈 2(명)이고, 표준오차는 4의 제곱근에 1/2을 곱해 준 1명이 됩니다.

실제로 Red를 Black 이라고 한 사람수, 1명이나 Black 을 Red 라고 한 사람수, 3명은 기대인원으로부터 표준오차 범위는 벗어 나지만 표준오차 두 배 범위에는 잘 들어갑니다.

두 번째 분석 방법은 확률을 계산해 보는 것입니다.

Red, Black 을 잘못 선택한 4명을 랜덤하게 둘로 나눌 때, 실제로 관찰된 3명 이상이나 1명이하로 잘못 선택한 사람수가 나올 가능성은 얼마나 되는가 입니다.

이 경우 공평한 동전던지기를 확률 법칙으로 표현한 이항분포로부터 계산할 수 있습니다.

시행횟수가 4이고 한 번 시행에서 성공확률이 $1/2$ 인 이항분포에서 성공횟수가 1이하이거나 3이상을 관찰할 확률은 0.625입니다.

공평한 동전 던지기에서 앞면이 1개 이하 나오는 확률은 3개 이상 나오는 확률과 같기 때문에 사실상 한쪽만 계산해서 2배 해 주면 됩니다.

이 값을 p-value 라고 하는데, p-value가 0.05보다 작을 때 통계적으로 유의한 차이를 관찰하였다고 말합니다.

즉, 공평한 동전을 던지는 것과 같은 과정이라고 가정하였을 때 실제로 관찰된 값들이 가정으로부터 얼마나 떨어져 있는지를 표현한 것입니다.

0.05는 이런 실험을 스무 번 정도 반복하면 1번 나올 정도로 드문 사건을 의미합니다.

즉 가정이 잘못되었다는 것입니다.

그런데 Red, Black 을 잘못 표시한 사람들의 분포에서 관찰된 p-value 는 0.05와는 비교도 안될 정도로 큰 값입니다.

따라서 두 집단이 랜덤화 효과가 작동하여 통계적으로 유의한 차이를 보이지 않는다고 할 수 있습니다.

7.1.2 응답인원의 Red, Black

Red 로 응답한 인원은 277명, Black 에 응답한 인원은 281명입니다.

전체 응답인원 558 명을 랜덤하게 둘로 나눌 때 어느 한 쪽의 기대인원은 전체 응답인원의 절반인 279명이고, 표준오차는 전체 응답인원의 제곱근에 $1/2$ 을 곱해 준 11.8 명입니다.

따라서 Red, Black 각 그룹에 관찰된 인원은 기대인원으로부터 표준오차 범위 안에 들어갑니다.

7.2 Q1. 한글의 문자 유형

1. 한글은 다음 중 어느 유형의 문자에 속하는가? *

- 민주 문자
- 엘리트 문자

7.2.1 한글은 민주 문자

	민주 문자	엘리트 문자	계
Red	261	16	277
Black	263	18	281
계	524	34	558

Table 7.3: Pearson's Chi-squared test with Yates' continuity correction: .

Test statistic	df	P value
0.01791	1	0.8935

Q1의 집계 결과가 Red, Black 간에 통계적으로 유의한 차이가 있는지 알아보기 위하여 카이제곱 테스트를 수행하였습니다.

그 결과 카이제곱 통계량은 0.018, 자유도는 1, p-value 는 0.8935이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

7.2.2 한글은 민주 문자(%)

민주 문자	엘리트 문자	계
93.9	6.1	100.0

정답률은 Red, Black 을 합하여 계산하는데, 93.9(%) 입니다.

7.3 Q2. 정보혁명과 문자 체계

2. 다음 중 정보혁명을 이끄는 데 유리한 문자체계는 어느 것인가? *

- 한자
- 한글

7.3.1 정보혁명을 이끄는 문자는 한글(집계표)

	한자	한글	계
Red	23	254	277
Black	16	265	281
계	39	519	558

Table 7.6: Pearson's Chi-squared test with Yates' continuity correction: .

Test statistic	df	P value
1.087	1	0.2971

Q2의 집계 결과가 Red, Black 간에 통계적으로 유의한 차이가 있는지 알아보기 위하여 카이제곱 테스트를 수행하였습니다.

그 결과 카이제곱 통계량은 1.09, 자유도는 1, p-value 는 0.30이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

7.3.2 정보혁명을 이끄는 문자는 한글(%)

한자	한글	계
7.0	93.0	100.0

정답률은 Red, Black 을 합하여 계산하는데, 93.0(%) 입니다.

7.4 Q3. 알기 힘든 전문 용어

3. 다음 중 그 뜻을 정확히 아는 단어는 몇 개나 되는가? *

몽리, 결궤, 저치, 장리

하나도 없다
 1개
 2개
 3개
 4개

7.4.1 몇 개나 아나요?(집계표)

	하나도 없다	1개	2개	3개	4개	계
Red	138	72	41	15	11	277
Black	150	70	40	11	10	281
계	288	142	81	26	21	558

Table 7.9: Pearson's Chi-squared test: .

Test statistic	df	P value
1.175	4	0.8822

Q3의 집계 결과가 Red, Black 간에 통계적으로 유의한 차이가 있는지 알아보기 위하여 카이제곱 테스트를 수행하였습니다.

그 결과 카이제곱 통계량은 1.175, 자유도는 4, p-value 는 0.8822이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

7.4.2 몇 개나 아나요?(%)

하나도 없다	1개	2개	3개	4개	계
51.6	25.4	14.5	4.7	3.8	100.0

물론, 이 문제에는 정답이 없으므로 정답률은 의미가 없습니다.

가장 많은 비율로 응답한 것이 “하나도 없다”이고 51.6(%) 입니다.

7.5 Q4. 해방직후 비문해율

4. 해방직후 우리나라의 비문해율은 대략 어느 수준이었는가? *

- 90%
- 80%
- 50%
- 20%
- 10%

7.5.1 집계

	90%	80%	50%	20%	10%	계
Red	13	198	27	28	11	277
Black	15	193	35	23	15	281
계	28	391	62	51	26	558

Table 7.12: Pearson's Chi-squared test: .

Test statistic	df	P value
2.316	4	0.6778

Q4의 집계 결과가 Red, Black 간에 통계적으로 유의한 차이가 있는지 알아보기 위하여 카이제곱 테스트를 수행하였습니다.

그 결과 카이제곱 통계량은 2.316, 자유도는 4, p-value 는 0.6778이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

7.5.2 %

90%	80%	50%	20%	10%	계
5.0	70.1	11.1	9.1	4.7	100.0

정답률은 Red, Black 을 합하여 계산하는데, 70.1(%) 입니다.

7.6 Q5. 세대간 문해력 격차

5. 2013년 OECD 보고서에 따르면 OECD 국가 중 20대와 50대의 문해력 격차가 가장 큰 나라는 어디인가? *

- 대한민국
- 영국
- 이탈리아
- 미국

7.6.1 집계

	대한민국	영국	이탈리아	미국	계
Red	212	18	18	29	277
Black	219	14	15	33	281
계	431	32	33	62	558

Table 7.15: Pearson's Chi-squared test: .

Test statistic	df	P value
1.116	3	0.7732

Q5의 집계 결과가 Red, Black 간에 통계적으로 유의한 차이가 있는지 알아보기 위하여 카이제곱 테스트를 수행하였습니다.

그 결과 카이제곱 통계량은 1.116, 자유도는 3, p-value 는 0.7732이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

7.6.2 %

대한민국	영국	이탈리아	미국	계
77.2	5.7	5.9	11.1	100.0

정답률은 Red, Black 을 합하여 계산하는데, 77.2(%) 입니다.

7.7 Q6. 문해력 격차의 파급효과

6. 2013년 OECD 보고서에 따르면 문해력이 낮은 집단은 문해력이 높은 집단에 비하여 어떤 차이를 보이고 있는 * 적절치 않은 것을 하나 고르시오?

- 시간당 중위 임금이 60% 낮다
- 2배 정도 실직할 가능성이 높다
- 건강이 좋지 않다
- 사회활동이나 자원활동에 덜 참여한다
- 남들을 덜 신뢰한다

7.7.1 집계

	60% 낮은 임금	실직 가능성	나쁜 건강	활동 불참	덜 신뢰	계
Red	153	24	40	23	37	277
Black	152	20	39	20	50	281
계	305	44	79	43	87	558

Table 7.18: Pearson's Chi-squared test: .

Test statistic	df	P value
2.503	4	0.6441

Q6의 집계 결과가 Red, Black 간에 통계적으로 유의한 차이가 있는지 알아보기 위하여 카이제곱 테스트를 수행하였습니다.

그 결과 카이제곱 통계량은 2.503, 자유도는 4, p-value 는 0.6441이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

7.7.2 %

60% 낮은 임금	실직 가능성	나쁜 건강	활동 불참	덜 신뢰	계
54.7	7.9	14.2	7.7	15.6	100.0

정답률은 Red, Black 을 합하여 계산하는데, 54.7(%) 입니다.

7.8 Q7. 프레임을 설정하는 단어의 힘

컵 가득 음료수를 채웠는데 미끄러지면서 반을 쏟았습니다.

이 상황에서 어떤 사람은 “그래도 반이 남았네”라고 긍정적 반응을 하고 또 다른 사람은 “어찌나 반 밖에 안 남았네”라고 부정적 반응을 보일 수 있습니다.

만약, 반응을 보기 전에 Red 에는 긍정적 단어들을 읽게 하고, Black 에는 부정적 단어들을 읽게 한 후 반응을 물어보면 어떻게 될까요?

Red, Black 의 성격상 단어를 보기 전에는 긍정적 반응의 비율과 부정적 반응의 비율이 닮을 것으로 기대됩니다.

그러나 단어를 보고 나서는 반응이 극명하게 나뉘는 것을 관찰하게 됩니다.

통계적으로 매우, 매우, … 유의한 차이를 보여 줍니다.

이것도 일종의 프레이밍이 인식에 미치는 영향이라고 할 수 있겠습니다.

힘든 상황이라도 긍정적인 생각을 가져야겠죠?

Red

7. 컵 가득 음료수를 채웠는데 미끄러지면서 반을 쏟았다. 다음에 주어진 단어들을 읽어보고 *
나라면 어떻게 반응할지 답하시오.

견딜만하다.

기대된다.

자극이 된다.

도약을 꿈꾼다.

더 나은 선택이 기다리고 있다.

휴우, 그래도 반이나 남았네.

웰, 반 밖에 안 남았네.

모름/기타

Black

7. 컵 가득 음료수를 채웠는데 미끄러지면서 반을 쏟았다. 다음에 주어진 단어들을 읽어보고 *
나라면 어떻게 반응할지 답하시오.

매우 화가 난다.

죽도록 힘들다.

불안하다.

우울하다.

외롭다.

월, 반 밖에 안 남았네.

휴우, 그래도 반이나 남았네.

모름/기타

7.8.1 집계

	반이나	반밖에	모름/기타	계
Red(긍정적 단어)	213	49	15	277
Black(부정적 단어)	83	183	15	281
계	296	232	30	558

Table 7.21: Pearson's Chi-squared test: .

Test statistic	df	P value
134.5	2	6.315e-30 ***

Q7의 Red는 음료수를 쏟은 상황에서 “견딜만 하다” 등의 긍정적인 단어들을 본 사람들인 데 277명이 응답한 가운데 213명이 “반이나” 남아 있다는 반응을 보이고, 49명이 “반 밖에” 안 남았다는 반응을 보입니다.

Black은 같은 상황에서 “매우 화가 난다” 등의 부정적인 단어들을 본 사람들인 데 281명이 응답한 가운데 83명이 “반이나” 남아 있다는 반응을 보이고, 183명이 “반 밖에” 안 남았다는 반응을 보입니다.

그리고 “모름/기타”에 답한 인원은 Red에 15명, Black 에 15명이었습니다.

랜덤화 효과일까요?

“모름/기타”의 응답이 유난히 많은 게 이런 해석을 불러 일으킵니다.

카이제곱 테스트는 이와 같은 상황에서 단어들을 이용하여 긍정적인 프레임을 구성한 경우와 부정적인 프레임을 구성한 경우에 그 차이가 통계적으로 매우, 매우, … 유의하다는 것을 보여 줍니다.

카이제곱 통계량은 134.469, 자유도는 2, p-value 는 6.3e-30 긍정적인 단어나 부정적인 단어를 단지 읽는 것만으로도 프레임이 설정되어 반응이 다르게 나온다는 것을 보여줍니다.

여기서 긍정이나 부정 프레이밍이 응답에 영향을 끼치지 않는다고 가정해 봅시다.

그렇다면 Red, Black 의 응답은 Q1 ~ Q6에서와 같이 랜덤화 효과에 의하여 통계적으로 유의한 차이를 보이지 않을 것입니다.

그런데 실제로 관찰된 카이제곱 통계값은 통계적으로 매우 유의한 차이를 보여 줍니다.

따라서 긍정이나 부정 프레이밍이 영향을 끼치지 않는다는 가정이 잘못되었다는 것을 논리적으로 입증할 수 있습니다.

이러한 논증 방식을 귀류법이라 합니다.

7.8.2 % 비교.

	반이나	반밖에	모름/기타	계
Red(긍정적 단어)	76.9	17.7	5.4	100.0
Black(부정적 단어)	29.5	65.1	5.3	100.0

긍정적인 단어들을 읽어 본 Red에서 “반이나” 남았다고 응답하는 사람들의 백분율, 76.9(%)은 “반 밖에” 안 남았다고 응답하는 사람들의 백분율, 17.7(%) 보다 월등히 높습니다.

반면 부정적인 단어들을 읽어 본 Black에서 “반이나” 남았다고 응답하는 사람들의 백분율, 29.5(%)은 “반 밖에” 안 남았다고 응답하는 사람들의 백분율, 65.1(%) 보다 훨씬 적습니다.

긍정적인 단어를 읽느냐, 부정적인 단어를 읽느냐에 따라 형성된 프레임의 영향이 막대하다는 것을 잘 알 수 있습니다.

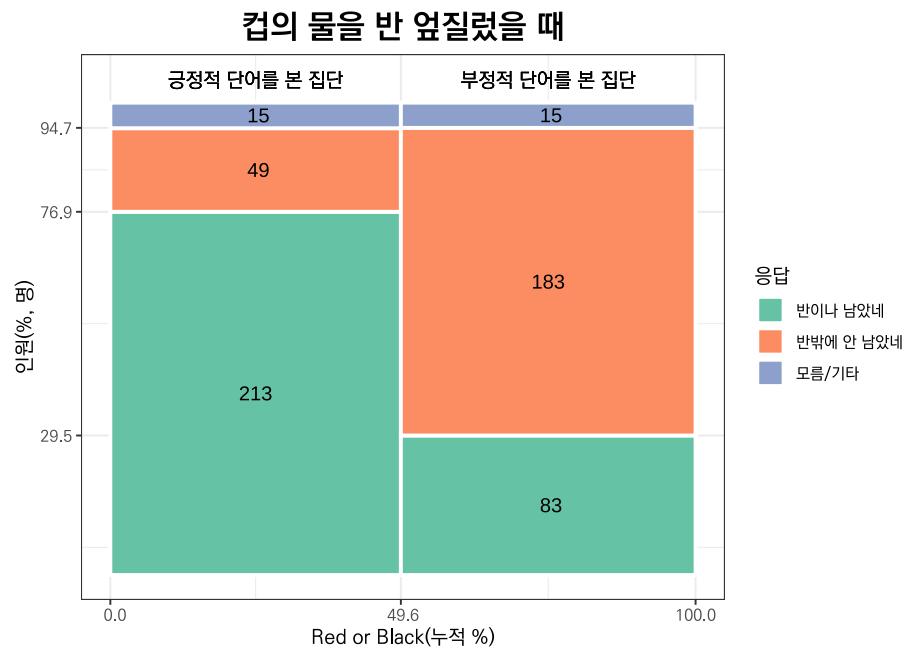
Red 와 Black 이 워낙 차이가 나지만 전체적으로 어느 정도가 “반이나” 남았다, 혹은 “반 밖에” 안 남았다고 응답하였는지 합쳐 보겠습니다.

7.8.3 % 합산

반이나	반밖에	모름/기타	계
53.0	41.6	5.4	100.0

“반이나” 남아 있다고 응답한 백분율은 Red, Black 합쳐서 53.0%(으)로 ‘반 밖에’ 안 남아 있다고 응답한 백분율, 41.6%() 보다 상당히 많습니다.

7.8.4 Mosaic Plot



Mosaic Plot은 이 집계결과를 시각적으로 잘 보여줍니다. 긍정적 단어를 본 Red에서 “반이나” 남아 있다고 응답한 백분율이 월등히 높고, 부정적 단어를 본 Black에서 “반 밖에” 안 남았다고 응답한 백분율이 월등히 높은 것을 시각적으로 알 수 있습니다.

덤으로 “모름/기타”的 백분율은 Red, Black 간에 매우 닮았다는 점도 시각적으로 쉽게 파악할 수 있습니다.

7.9 마감 시간으로부터 제출 시간의 분포

7.9.1 분포표

Table 7.24: 일 단위

	[0,1](1,2](2,3](3,4](4,5](5,6](6,7](7,8](8,9](9,10](10,11](11,12](12,13](13,14]														
Red	81	14	8	13	11	5	4	38	22	13	14	19	12	23	277
Black	78	14	12	6	8	3	6	41	29	15	17	15	20	17	281
계	159	28	20	19	19	8	10	79	51	28	31	34	32	40	558

분포표로부터 두 가지 문제를 살펴보겠습니다.

첫째, 날마다 고르게 제출하는가?

둘째, Red, Black 간에 통계적으로 유의한 차이가 있는가?

각 문제를 살펴보기 위해서는 분포표의 일부분을 대상으로 카이제곱 테스트를 수행합니다.

7.9.2 날마다 고르게 제출하는가?

[0,1] (1,2] (2,3] (3,4] (4,5] (5,6] (6,7] (7,8] (8,9] (9,10] (10,11] (11,12] (12,13] (13,14]															
159	28	20	19	19	8	10	79	51	28	31	34	32	40		

Table 7.26: Chi-squared test for given probabilities: .

Test statistic	df	P value
488.7	13	3.693e-96 * * *

날마다 고르게 제출하는지 알아 보았습니다.

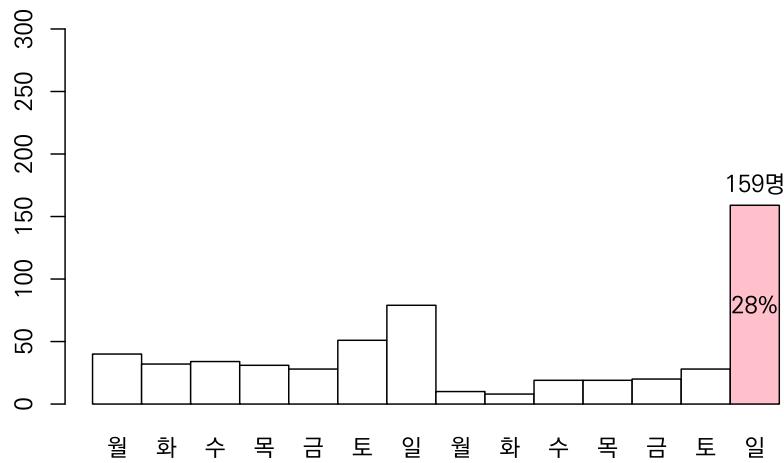
분포표의 “계”행에서 ‘계’열을 제외하고 카이제곱테스트를 수행합니다. 분포표 만으로도 쉽게 파악할 수 있지만 카이제곱테스트가 명확히 해 줍니다.

카이제곱 통계량은 488.69, 자유도는 13.00, p-value 는 3.7e-96 이므로 결코 고르게 제출한다고 말할 수 없겠습니다.

막대그래프로 살펴 보겠습니다.

7.9.3 막대그래프

Quiz250331 (558명 제출)



막대그래프는 총 제출인원 558(명) 중에 159(명), 28(%)가 마감일에 몰리는 것을 명확히 보여주고 있습니다.

7.9.4 Red, Black 간에 닮았는가?

	[0,1](1,2](2,3](3,4](4,5](5,6](6,7](7,8](8,9](9,10](10,11](11,12](12,13](13,14]													
Red	81	14	8	13	11	5	4	38	22	13	14	19	12	23
Black	78	14	12	6	8	3	6	41	29	15	17	15	20	17

Table 7.28: Pearson's Chi-squared test: .

Test statistic	df	P value
9.66	13	0.7215

제출시간의 분포가 Red, Black 간에 닮았는지 알아 보았습니다.

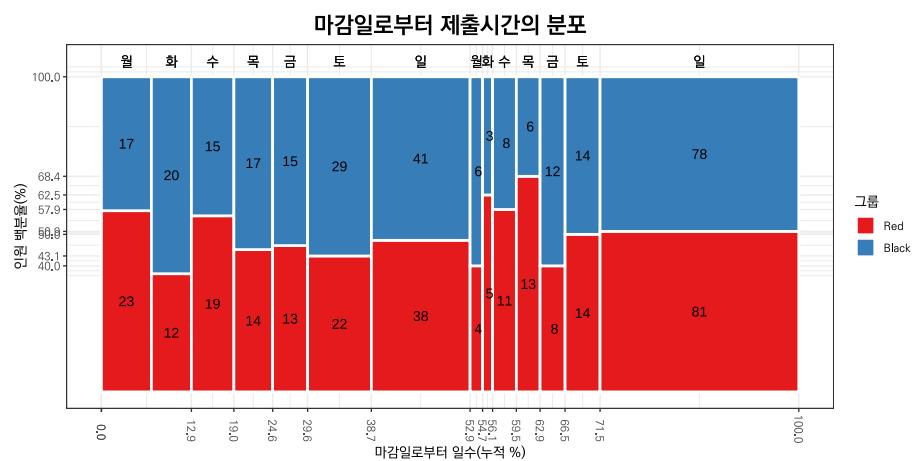
이번에는 분포표의 첫번째와 두번째 행, '계'열을 제외한 나머지 열에 대해서 카이제곱테스트를 수행합니다.

카이제곱 통계량은 9.660, 자유도는 13, p-value 는 0.7215 이므로 제출 시간의 분포는 Red, Black 간에 통계적으로 유의한 차이가 관찰되지 않습니다.

이 사실을 Mosaic Plot을 이용하여 시각적으로 살펴보겠습니다.

답았다고 느껴지나요?

7.9.5 Mosaic Plot



Chapter 8

6주차 데이터 실험 집계

8.1 실험의 목적

6주차 구글 예습 설문지 집계결과를 분석합니다.

Q1~Q6에서는 랜덤화의 효과로 Red, Black 이 얼마나 닮았는지 알아봅니다.

Q7에서는 소득의 절대값보다 상대적 비교가 중시된다는 실험 결과를 분석합니다.

제출시간의 분포가 날마다 고른지, Red, Black 간에는 닮았는지 알아봅니다.

8.1.1 Red, Black을 잘못 표시한 사람들

제출시간	학번	랜덤화출석부	구글예습퀴즈
2025-04-13 21:41:05	20202133	Red	Black
2025-04-13 21:50:35	20241011	Red	Black

	Red(구글예습퀴즈)	Black(구글예습퀴즈)
Red(랜덤화출석부)	189	2
Black(랜덤화출석부)	0	188
계	189	190

랜덤화출석부에 있는 Red, Black 과 실제 구글설문에 올린 Red, Black 이 다른 사람들의 수효는 2명입니다.

Red를 Black 이라고 한 사람이 2명, Black 을 Red 라고 한 사람이 0명입니다.

두 가지 방법으로 분석합니다.

우선 Red, Black 을 잘못 선택한 2명을 랜덤하게 둘로 나누면 어느 한 쪽 집단에 들어갈 기대인원은 2명을 둘로 나눈 1(명)이고, 표준오차는 2의 제곱근에 $1/2$ 을 곱해 준 0.7명이 됩니다.

실제로 Red를 Black 이라고 한 사람수, 2명이나 Black 을 Red 라고 한 사람수, 0명은 기대인원으로부터 표준오차 범위는 벗어 나지만 표준오차 두 배 범위에는 잘 들어갑니다.

두 번째 분석 방법은 확률을 계산해 보는 것입니다.

Red, Black 을 잘못 선택한 2명을 랜덤하게 둘로 나눌 때, 실제로 관찰된 2명 이상이나 0명이하로 잘못 선택한 사람수가 나올 가능성은 얼마나 되는가 입니다.

이 경우 공평한 동전던지기를 확률 법칙으로 표현한 이항분포로부터 계산할 수 있습니다.

시행횟수가 2이고 한 번 시행에서 성공확률이 $1/2$ 인 이항분포에서 성공횟수가 0이하이거나 2이상을 관찰할 확률은 0.5입니다.

공평한 동전 던지기에서 앞면이 0개 이하 나오는 확률은 2개 이상 나오는 확률과 같기 때문에 사실상 한쪽만 계산해서 2배 해 주면 됩니다.

이 값을 p-value 라고 하는데, p-value가 0.05보다 작을 때 통계적으로 유의한 차이를 관찰하였다고 말합니다.

즉, 공평한 동전을 던지는 것과 같은 과정이라고 가정하였을 때 실제로 관찰된 값들이 가정으로부터 얼마나 떨어져 있는지를 표현한 것입니다.

0.05는 이런 실험을 스무 번 정도 반복하면 1번 나올 정도로 드문 사건을 의미합니다.

즉 가정이 잘못되었다는 것입니다.

그런데 Red, Black 을 잘못 표시한 사람들의 분포에서 관찰된 p-value 는 0.05와는 비교도 안될 정도로 큰 값입니다.

따라서 두 집단이 랜덤화 효과가 작동하여 통계적으로 유의한 차이를 보이지 않는다고 할 수 있습니다.

8.1.2 응답인원의 Red, Black

Red 로 응답한 인원은 189명, Black 에 응답한 인원은 190명입니다.

전체 응답인원 379 명을 랜덤하게 둘로 나눌 때 어느 한 쪽의 기대인원은 전체 응답인원의 절반인 189.5명이고, 표준오차는 전체 응답인원의 제곱근에 $1/2$ 을 곱해 준 9.7 명입니다.

따라서 Red, Black 각 그룹에 관찰된 인원은 기대인원으로부터 표준오차 범위 안에 들어갑니다.

8.2 Q1. 월간 독서율

4. “조사 대상 가운데 한 달에 한 권 이상 책을 읽는 사람은 32%이며 ...” 라는 기사로부터 알 수 있는 사실은 무엇 * 입니까?

- 월간독서율
- 월간독서량
- 월간도서구입율
- 월간 도서구입량

8.2.1 집계

	월간독서율	월간독서량	월간도서구입율	도서구입량	월간 계
Red	164	18	6	1	189
Black	164	23	3	0	190
계	328	41	9	1	379

Table 8.4: Pearson's Chi-squared test: .

Test statistic	df	P value
2.607	3	0.4562

Q1의 집계 결과가 Red, Black 간에 통계적으로 유의한 차이가 있는지 알아보기 위하여 카이제곱 테스트를 수행하였습니다.

그 결과 카이제곱 통계량은 2.61, 자유도는 3, p-value 는 0.4562이므로 Red, Black 간에 통계적으로 유의한 차이를 보이고 있습니다.

어디가 문제인가요?

8.2.2 %

월간독서율	월간독서량	월간도서구입율	월간 도서구입량	계
86.54	10.82	2.37	0.26	100.00

정답률은 Red, Black 을 합하여 계산하는데, 86.5(%) 입니다.

8.3 Q2. 지역 및 지역크기별 가구수 비례 무작위추출법

1. “표본은 ‘지역 및 지역크기별 가구수 비례 무작위추출법으로’ 남자 7백42명, 여자 7백57명을 뽑았다.”에서 홀따 *
- 음표 친 부분을 쉽게 바꾸면 다음 중 어느 것이 적합한가?

- 공평하게
- 소득 순으로
- 학력 순으로
- 연령 순으로

8.3.1 집계

	공평하게	소득 순으로	학력 순으로	연령 순으로	계
Red	164	10	9	6	189
Black	175	7	3	5	190
계	339	17	12	11	379

Table 8.7: Pearson’s Chi-squared test: .

Test statistic	df	P value
3.975	3	0.2642

Q2의 집계 결과가 Red, Black 간에 통계적으로 유의한 차이가 있는지 알아보기 위하여 카이제곱 테스트를 수행하였습니다.

그 결과 카이제곱 통계량은 3.97, 자유도는 3, p-value 는 0.2642이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

8.3.2 %

공평하게	소득 순으로	학력 순으로	연령 순으로	계
89.4	4.5	3.2	2.9	100.0

정답률은 Red, Black 을 합하여 계산하는데, 89.4(%) 입니다.

8.4 Q3. 한달 독서량의 분포

2. “책을 읽은 사람들의 한달 평균 독서량은 1권 15.3%, 2권 12.4%, 3권 5.7%, 4권 2.1%, 5권 이상 8.3%의 분포를 보였다.”에서 삭제하거나 교체하여 여야 할 단어는?

- 책을
- 한달
- 평균
- 독서량
- 분포를

8.4.1 집계

	책을	한달	평균	독서량	분포를	계
Red	15	9	131	15	19	189
Black	8	4	117	24	37	190
계	23	13	248	39	56	379

Table 8.10: Pearson's Chi-squared test: .

Test statistic	df	P value
12.7	4	0.01282 *

Q3의 집계 결과가 Red, Black 간에 통계적으로 유의한 차이가 있는지 알아보기 위하여 카이제곱 테스트를 수행하였습니다.

그 결과 카이제곱 통계량은 12.70, 자유도는 4, p-value 는 0.01280이므로 Red, Black 간에 통계적으로 유의한 차이를 보이고(지) 있(않)습니다.

실제로 닮은 게 느껴집니까?

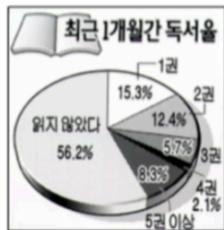
8.4.2 %

	책을	한달	평균	독서량	분포를	계
	6.1	3.4	65.4	10.3	14.8	100.0

정답률은 Red, Black 을 합하여 계산하는데, 65.4(%) 입니다.

8.5 Q4. 최근 1개월간 독서량

3. 다음 도표에서 잘못된 표현은? *



- 최근
- 1개월간
- 독서율
- 읽지 않았다
- 56.2%

8.5.1 집계

	최근	1개월간	독서율	읽지 않았다	56.2%	계
Red	15	13	139	11	11	189
Black	30	5	134	9	12	190
계	45	18	273	20	23	379

Table 8.13: Pearson's Chi-squared test: .

Test statistic	df	P value
8.888	4	0.06396

Q4의 집계 결과가 Red, Black 간에 통계적으로 유의한 차이가 있는지 알아보기 위하여 카이제곱 테스트를 수행하였습니다.

그 결과 카이제곱 통계량은 8.89, 자유도는 4, p-value 는 0.06400이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

8.5.2 %

최근	1개월간	독서율	읽지 않았다	56.2%	계
11.9	4.7	72.0	5.3	6.1	100.0

정답률은 Red, Black 을 합하여 계산하는데, 72.0(%) 입니다.

8.6 Q5. 20대의 연간독서율

5. 2021년 우리나라 일반도서의 연간독서율은 어느 수준입니까?*

- 72%
- 76.7%
- 65.4%
- 52.1%
- 40.7%

8.6.1 집계

	72.0%	76.7%	65.4%	52.1%	40.7%	계
Red	4	8	17	26	134	189
Black	3	10	17	31	129	190
계	7	18	34	57	263	379

Table 8.16: Pearson's Chi-squared test: .

Test statistic	df	P value
0.8961	4	0.9251

Q5의 집계 결과가 Red, Black 간에 통계적으로 유의한 차이가 있는지 알아보기 위하여 카이제곱 테스트를 수행하였습니다.

그 결과 카이제곱 통계량은 0.90, 자유도는 4, p-value 는 0.9251이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

8.6.2 %

72.0%	76.7%	65.4%	52.1%	40.7%	계
1.8	4.7	9.0	15.0	69.4	100.0

정답률은 Red, Black 을 합하여 계산하는데, 69.4(%) 입니다.

8.7 Q6. 50대의 연간독서율

6. 2021년 우리나라 20-29세의 전자책 독서율과 60세 이상의 전자책 독서율을 정확하게 비교한 것은? *

- 20-29세 31.3%, 60세 이상 2.5%
- 20-29세 27.8%, 60세 이상 1%
- 20-29세 34.7%, 60세 이상 1.3%
- 20-29세 39%, 60세 이상 2%
- 20-29세 50.5%, 60세 이상 2.3%

8.7.1 집계

	20-29세 31.3%, 60세 이상 2.5%	20-29세 27.8%, 60세 이상 1%	20-29세 34.7%, 60세 이상 1.3%	20-29세 39%, 60세 이상 2%	20-29세 50.5%, 60세 이상 2.3%	계
Red	19	24	28	15	103	189
Black	21	26	33	12	98	190
계	40	50	61	27	201	379

Table 8.19: Pearson's Chi-squared test: .

Test statistic	df	P value
1.045	4	0.9029

Q6의 집계 결과가 Red, Black 간에 통계적으로 유의한 차이가 있는지 알아보기 위하여 카이제곱 테스트를 수행하였습니다.

그 결과 카이제곱 통계량은 1.04, 자유도는 4, p-value 는 0.9029이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 답은 게 느껴집니까?

8.7.2 %

20-29세 31.3%, 60세 이상 2.5%	20-29세 27.8%, 60세 이상 1%	20-29세 34.7%, 60세 이상 1.3%	20-29세 39%, 60세 이상 2%	20-29세 50.5%, 60세 이상 2.3%	계
10.6	13.2	16.1	7.1	53.0	100.0

정답률은 Red, Black 을 합하여 계산하는데, 53.0(%) 입니다.

8.8 Q7. The more, the better? : 내가 남보다, 혹은 남이 나보다

Q7의 응답결과는 매우 중요한 시사점을 보여줍니다.

다다익선, 많으면 많을수록 좋다, 영어로는 The more, the better 가 아니라 절대 금액은 적더라도 상대적 비교에서 많은 것을 더 선호한다는 것입니다.

Red 와 Black 의 차이는 순서를 바꿔서 소위 1번효과가 나타나는지를 살펴보고자 하였으나 P-value 가 시사하듯이 그러한 차이는 관찰되지 않습니다.

소득의 절대값이 아니라 상대 비교가 중요하다는 Solnick and Hemenway(1998)의 연구결과와 일치합니다.

랜덤화하였지만 응답에는 차이가 없음을 보여주고 있습니다.

Red
7. (Solnick & Hemenway 1998) 내가 살고 싶은 세상은 다음 중 어디입니까? *
<input type="radio"/> 가. 내 연봉은 5천만원 남들은 2천5백만원
<input type="radio"/> 나. 내 연봉은 1억원, 남들은 2억원

Black

7. (Solnick & Hemenway 1998) 내가 살고 싶은 세상은 다음 중 어디입니까? *

가. 내 연봉은 1억원, 남들은 2억원
 나. 내 연봉은 5천만원, 남들은 2천5백만원

8.8.1 집계

	내가 남보다	남이 나보다	계
Red('내가 남보다' 먼저)	140	49	189
Black('남이 나보다' 먼저)	142	48	190
계	282	97	379

Table 8.22: Pearson's Chi-squared test with Yates' continuity correction: .

Test statistic	df	P value
0.0009076	1	0.976

Q7의 Red 는 앞 선택지에 내 연봉은 5천만원, 남들은 2천5백만원을 받는 세상을 제시하고 뒷 선택지에 내 연봉은 1억원, 남들은 2억원을 받는 세상을 제시하였습니다.

Black 은 각 선택지의 순서만 바꿔서 소위 1번효과를 측정하려 하였습니다.
 집계결과는 절대금액에서 많이 받는 세상보다 상대적으로 더 많이 받는 세상을 원한다는 점을 명확히 보여 줍니다.

여기서 수행하고 있는 카이제곱 테스트는 선택지의 순서만 바꿨을 때 통계적으로 유의한 차이를 보여주는지 파악하는 것으로서 카이제곱 통계량은 0.00091, 자유도는 1, p-value 는 0.9760로 깜짝 놀랄만큼 선택지의 순서는 관계 없다는 것을 보여 줍니다.

여러분이 보기에도 그렇게 닮았다고 느껴집니까?

이번에도 귀류법의 논리를 적용해 봅니다.

여기서 선택지의 순서를 바꾸는 것이 응답에 영향을 끼치지 않는다고 가정해 봅시다.

그렇다면 Red, Black 의 응답은 Q1~Q4 에서와 같이 랜덤화 효과에 의하여 통계적으로 유의한 차이를 보이지 않을 것입니다.

그리고 실제로 관찰된 카이제곱 통계값의 p-value 도 0.05와는 비교도 안 되는 수준으로 크기 때문에 통계적으로 유의한 차이를 보여주지 못합니다.

따라서 선지의 순서를 바꾸는 것은 응답에 영향을 끼치지 못한다는 가설을 받아들이게 됩니다.

8.8.2 % 비교.

	내가 남보다	남이 나보다	계
Red('내가 남보다' 먼저)	74.1	25.9	100.0
Black('남이 나보다' 먼저)	74.7	25.3	100.0

이를 백분율로 살펴보면 “내가 남보다”를 앞 선택지로 한 Red에서 “내가 남보다”를 선택한 백분율, 74.1%(이)나 “남이 나보다”를 앞에 앞 선택지로 한 Black에서 “내가 남보다”를 선택한 백분율, 74.7%(이)나 거의 차이가 없어서 선택지의 순서를 바꿔 봐야 응답에 통계적으로 유의한 영향을 주지 못한다는 것을 명확히 알 수 있습니다.

따라서 통계적으로 유의한 차이를 관찰하지 못했으므로 Red, Black은 합쳐서 분석하는 것이 타당합니다.

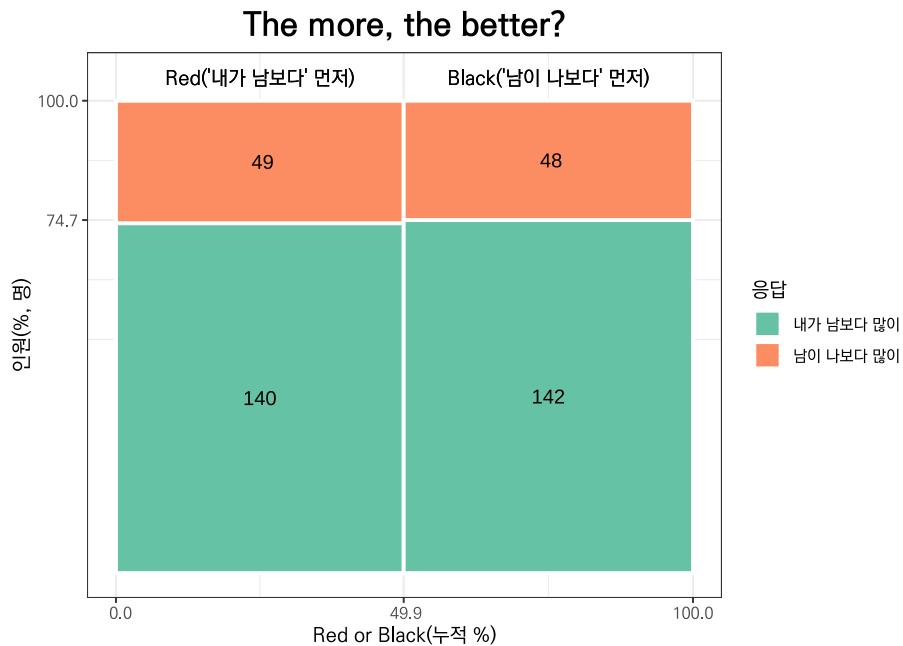
그 결과는 다음과 같습니다.

8.8.3 합산(%)

내가 남보다	남이 나보다	계
74.4	25.6	100.0

“내가 남보다” 많이 받는 세상을 원한다고 응답한 백분률은 선택지의 순서와는 무관하게 74.4%(으)로 매우 높습니다.

8.8.4 Mosaic Plot



Mosaic Plot 은 이 집계결과를 시각적으로 잘 보여줍니다.

선택지의 순서에 거의 무관할 정도로 응답이 비슷하기 때문에 Red, Black 이 많았고, “내가 남보다” 많이 받아야 한다고 응답한 백분율도 높은 것을 잘 알 수 있습니다.

8.9 마감 시간으로부터 제출 시간의 분포

8.9.1 분포표

Table 8.25: 일 단위

	[0,1]	(1,2]	(2,3]	(3,4]	(4,5]	(5,6]	(6,7]	(7,8]	(8,9]	(9,10]	[10,11]	[11,12]	[12,13]	[13,14]
Red	0	3	10	6	4	6	9	35	20	16	10	26	18	26
Black	0	6	8	3	3	7	11	47	16	14	15	14	21	25
계	0	9	18	9	7	13	20	82	36	30	25	40	39	51
														379

분포표로부터 두 가지 문제를 살펴보겠습니다.

첫째, 날마다 고르게 제출하는가?

둘째, Red, Black 간에 통계적으로 유의한 차이가 있는가?

각 문제를 살펴보기 위해서는 분포표의 일부분을 대상으로 카이제곱 테스트를 수행합니다.

8.9.2 날마다 고르게 제출하는가?

(1,2]	(2,3]	(3,4]	(4,5]	(5,6]	(6,7]	(7,8]	(8,9]	(9,10]	(10,11]	11,12]	12,13]	13,14]
9	18	9	7	13	20	82	36	30	25	40	39	51

Table 8.27: Chi-squared test for given probabilities: .

Test statistic	df	P value
182.5	12	1.285e-32 ***

날마다 고르게 제출하는지 알아 보았습니다.

분포표의 “계”행에서 ’계’열을 제외하고 카이제곱테스트를 수행합니다.

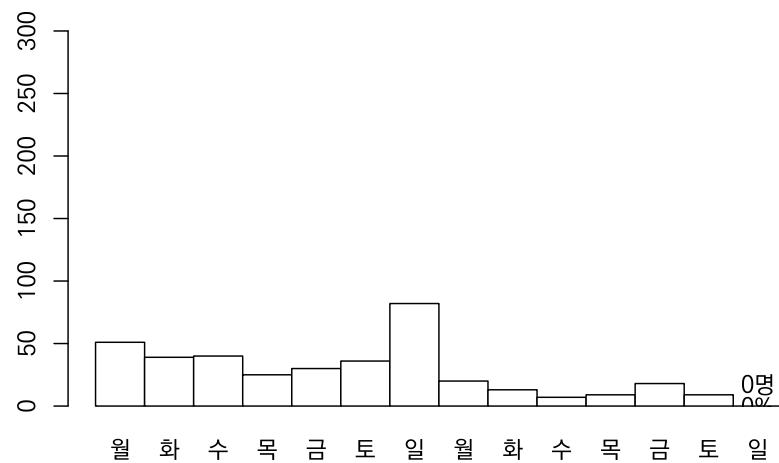
분포표 만으로도 쉽게 파악할 수 있지만 카이제곱테스트가 명확히 해 줍니다.

카이제곱 통계량은 182.54, 자유도는 12.00, p-value 는 1.3e-32 이므로 날짜별로 고르게 제출하고 있는지 살펴보기에는 천문학적으로 낮은 P-value 가 관찰되었습니다.

막대그래프로 살펴 보겠습니다.

8.9.3 막대그래프

Quiz250407 (379명 제출)



8.9.4 Red, Black 간에 닮았는가?

(1,2]	(2,3]	(3,4]	(4,5]	(5,6]	(6,7]	(7,8]	(8,9]	(9,10]	(10,11]	(11,12]	(12,13]	(13,14]
Red	3	10	6	4	6	9	35	20	16	10	26	18
Black	6	8	3	3	7	11	47	16	14	15	14	21

Table 8.29: Pearson's Chi-squared test: .

Test statistic	df	P value
9.824	12	0.6314

제출시간의 분포가 Red, Black 간에 닮았는지 알아 보았습니다.

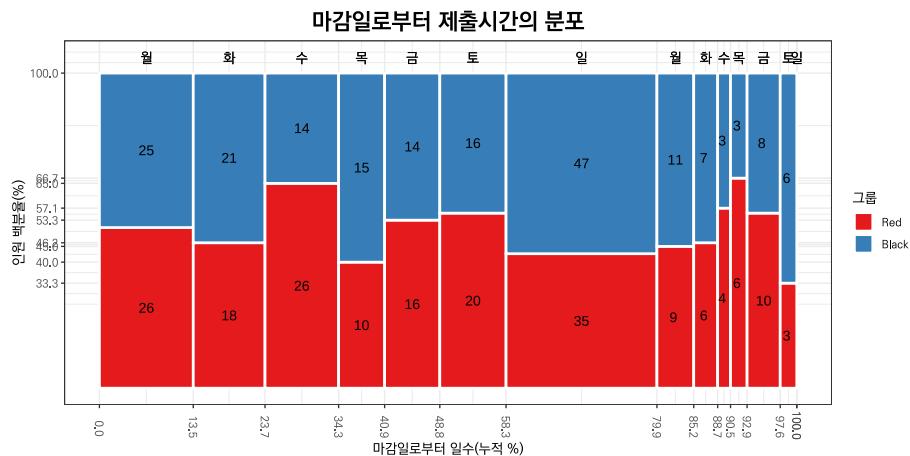
이번에는 분포표의 첫번째와 두번째 행, '계'열을 제외한 나머지 열에 대해서 카이제곱테스트를 수행합니다.

카이제곱 통계량은 9.82, 자유도는 12, p-value 는 0.63 이므로 제출 시간의 분포는 Red, Black 간에 통계적으로 유의한 차이가 관찰되고 있습니다.

이 사실을 Mosaic Plot 을 이용하여 시각적으로 살펴보겠습니다.

어디서 차이가 나고 있나요?

8.9.5 Mosaic Plot



Chapter 9

7주차 데이터 실험 집계

9.1 실험의 목적

7주차 구글 예습 설문지 집계결과를 분석합니다.

Q1 ~ Q6에서는 랜덤화의 효과로 Red, Black 이 얼마나 닮았는지 알아봅니다.

Q7에서는 Wason Selection Task에서 추상적 문제에 취약하고 인지적 편향에 쓰러는 우리의 모습을 파악합니다. 같은 구조의 문제를 추상적으로 표현할 때와 구체적인 사례를 들어 표현할 때 정답률이 매우 차이나는 것을 살펴보고 인지적 편향을 어떻게 확인하는지 그리고 학습 방법에 대한 추론까지 진행해 봅니다.

제출시간의 분포가 날마다 고른지, Red, Black 간에는 닮았는지 알아봅니다.

9.1.1 Red, Black을 잘못 표시한 사람들

	Red(구글예습퀴즈)	Black(구글예습퀴즈)
Red(랜덤화출석부)	91	0
Black(랜덤화출석부)	0	93
계	91	93

랜덤화출석부에 있는 Red, Black 과 실제 구글설문에 올린 Red, Black 이 다른 사람들의 수호는 0명입니다.

Red를 Black 이라고 한 사람이 0명, Black 을 Red 라고 한 사람이 0명입니다.

두 가지 방법으로 분석합니다.

우선 Red, Black 을 잘못 선택한 0명을 랜덤하게 둘로 나누면 어느 한 쪽 집단에 들어갈 기대인원은 0명을 둘로 나눈 0(명)이고, 표준오차는 0의 제곱근에 1/2을 곱해 준 0명이 됩니다.

실제로 Red를 Black 이라고 한 사람수, 0명이나 Black 을 Red 라고 한 사람수, 0명은 기대인원으로부터 표준오차 범위는 벗어 나지만 표준오차 두 배 범위에는 잘 들어갑니다.

두 번째 분석 방법은 확률을 계산해 보는 것입니다.

Red, Black 을 잘못 선택한 0명을 랜덤하게 둘로 나눌 때, 실제로 관찰된 0명 이상이나 0명이하로 잘못 선택한 사람수가 나올 가능성은 얼마나 되는가 입니다.

이 경우 공평한 동전던지기를 확률 법칙으로 표현한 이항분포로부터 계산할 수 있습니다.

시행횟수가 0이고 한 번 시행에서 성공확률이 1/2 인 이항분포에서 성공횟수가 0이하이거나 0이상을 관찰할 확률은 2입니다.

공평한 동전 던지기에서 앞면이 0개 이하 나오는 확률은 0개 이상 나오는 확률과 같기 때문에 사실상 한쪽만 계산해서 2배 해 주면 됩니다.

이 값을 p-value 라고 하는데, p-value가 0.05보다 작을 때 통계적으로 유의한 차이를 관찰하였다고 말합니다.

즉, 공평한 동전을 던지는 것과 같은 과정이라고 가정하였을 때 실제로 관찰된 값들이 가정으로부터 얼마나 떨어져 있는지를 표현한 것입니다.

0.05는 이런 실험을 스무 번 정도 반복하면 1번 나올 정도로 드문 사건을 의미합니다.

즉 가정이 잘못되었다는 것입니다.

그런데 Red, Black 을 잘못 표시한 사람들의 분포에서 관찰된 p-value 는 0.05와는 비교도 안될 정도로 큰 값입니다.

따라서 두 집단이 랜덤화 효과가 작동하여 통계적으로 유의한 차이를 보이지 않는다고 할 수 있습니다.

9.1.2 응답인원의 Red, Black

Red 로 응답한 인원은 91명, Black 에 응답한 인원은 93명입니다.

전체 응답인원 184 명을 랜덤하게 둘로 나눌 때 어느 한 쪽의 기대인원은 전체 응답인원의 절반인 92명이고, 표준오차는 전체 응답인원의 제곱근에 1/2을 곱해 준 6.8 명입니다.

따라서 Red, Black 각 그룹에 관찰된 인원은 기대인원으로부터 표준오차 범위 안에 들어갑니다.

9.2 Q1. 통계학의 기본원리

1. 통계학의 기본원리입니다. 괄호 안에 적당한 단어를 넣는다면? 표본을 ()하게 추출하면 모집단의 특성을 잘 * 닮는다.

- 공평
- 무난
- 철저
- 균일

9.2.1 공평하게 추출하면 ...

	공평	무난	철저	균일	계
Red	67	2	1	21	91
Black	73	3	0	17	93
계	140	5	1	38	184

Table 9.3: Pearson's Chi-squared test: .

Test statistic	df	P value
1.857	3	0.6027

Q1의 집계 결과가 Red, Black 간에 통계적으로 유의한 차이가 있는지 알아보기 위하여 카이제곱 테스트를 수행하였습니다.

그 결과 카이제곱 통계량은 1.86, 자유도는 3, p-value 는 0.60이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

9.2.2 공평하게 추출하면 ... (%)

공평	무난	철저	균일	계
76.09	2.72	0.54	20.65	100.00

정답률은 Red, Black 을 합하여 계산하는데, 76.1(%) 입니다.

9.3 Q2. 리터러리 디이제스트의 실패

2. 1936년 리터러리 디이제스트의 여론조사가 실패로 돌아간 가장 큰 원인은 무엇입니까? *

- Selection Bias
- Response Bias
- Non-response Bias
- Panel Bias

9.3.1 Selection Bias

	Selection Bias	Response Bias	Non-response Bias	Panel Bias	계
Red	63	10	18	0	91
Black	63	8	17	5	93
계	126	18	35	5	184

Table 9.6: Pearson's Chi-squared test: .

Test statistic	df	P value
5.23	3	0.1557

Q2의 집계 결과가 Red, Black 간에 통계적으로 유의한 차이가 있는지 알아보기 위하여 카이제곱 테스트를 수행하였습니다.

그 결과 카이제곱 통계량은 5.23, 자유도는 3, p-value 는 0.16이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

9.3.2 Selection Bias (%)

Selection Bias	Response Bias	Non-response Bias	Panel Bias	계
68.5	9.8	19.0	2.7	100.0

정답률은 Red, Black 을 합하여 계산하는데, 68.5(%) 입니다.

9.4 Q3. 1948년, 여론조사가 듀이를 당선시킨 해

3. 1948년 미국 대선 당시 활약하던 3대 여론조사기관이 사용하던 표본 추출 방법은 무엇입니까? *

- 다단계 집락 추출방법
- 할당법
- 단순 랜덤추출방법
- RDD

9.4.1 할당법의 문제점

	다단계 집락 추출방법	할당법	단순 랜덤추출방법	RDD	계
Red	7	66	13	5	91
Black	10	58	22	3	93
계	17	124	35	8	184

Table 9.9: Pearson's Chi-squared test: .

Test statistic	df	P value
3.839	3	0.2794

Q3의 집계 결과가 Red, Black 간에 통계적으로 유의한 차이가 있는지 알아보기 위하여 카이제곱 테스트를 수행하였습니다.

그 결과 카이제곱 통계량은 3.84, 자유도는 3, p-value 는 0.28이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

9.4.2 할당법의 문제점(%)

다단계 집락 추출방법	할당법	단순 랜덤추출방법	RDD	계
9.2	67.4	19.0	4.3	100.0

정답률은 Red, Black 을 합하여 계산하는데, 67.4(%) 입니다.

9.5 Q4. 1948 미 대선 이후

4. 1948년도 미국 대선의 선거여론조사와 그 이후 선거여론조사의 결정적 차이는 무엇입니까? *

- 확률적 표본추출방법이 도입되었다.
- 할당법이 도입되었다.
- 유선전화번호부를 활용하게 되었다.
- Random Digit Dialling이 전면적으로 도입되었다.

9.5.1 확률적 표본추출방법 도입

	확률적 표본추출	할당법	유선전화번호부	RDD도입	계
Red	66	10	8	7	91
Black	72	7	4	10	93
계	138	17	12	17	184

Table 9.12: Pearson's Chi-squared test: .

Test statistic	df	P value
2.632	3	0.452

Q4의 집계 결과가 Red, Black 간에 통계적으로 유의한 차이가 있는지 알아보기 위하여 카이제곱 테스트를 수행하였습니다.

그 결과 카이제곱 통계량은 2.63, 자유도는 3, p-value 는 0.45이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

9.5.2 확률적 표본추출방법 도입 … (%)

확률적 표본추출	할당법	유선전화번호부	RDD도입	계
75.0	9.2	6.5	9.2	100.0

정답률은 Red, Black 을 합하여 계산하는데, 75.0(%) 입니다.

9.6 Q5. 표본오차를 반으로 줄이려면?

5. 표본오차를 $1/2$ 로 줄이려면 표본의 크기를 어떻게 해야 합니까? *

- 2배로 늘린다.
- 4배로 늘린다.
- $1/2$ 로 줄인다.
- $1/4$ 로 줄인다.

9.6.1 4배로 늘려야

	2배로	4배로	1/2로	1/4로	계
Red	18	58	11	4	91
Black	26	55	11	1	93
계	44	113	22	5	184

Table 9.15: Pearson's Chi-squared test: .

Test statistic	df	P value
3.313	3	0.3459

Q5의 집계 결과가 Red, Black 간에 통계적으로 유의한 차이가 있는지 알아보기 위하여 카이제곱 테스트를 수행하였습니다.

그 결과 카이제곱 통계량은 3.31, 자유도는 3, p-value 는 0.350이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

9.6.2 4배로 늘려야 (%)

2배로	4배로	1/2로	1/4로	계
23.9	61.4	12.0	2.7	100.0

정답률은 Red, Black 을 합하여 계산하는데, 61.4(%) 입니다.

9.7 Q6. 대선 여론조사의 목표모집단?

6. 다음 중 대선 여론조사에서 목표모집단으로 적절한 것은? *

- 국민 전체
- 18세 이상 국민 전체
- 등록된 유권자 전체
- 선거 당일 투표하는 유권자 전체

9.7.1 선거당일 투표하는 유권자 전체

	국민 전체	18세 이상 국민 전체	등록된 유권자 전체	선거 당일 투표하는 유권자 전체	계
Red	6	19	14	52	91
Black	3	26	16	48	93
계	9	45	30	100	184

Table 9.18: Pearson's Chi-squared test: .

Test statistic	df	P value
2.361	3	0.501

Q6의 집계 결과가 Red, Black 간에 통계적으로 유의한 차이가 있는지 알아보기 위하여 카이제곱 테스트를 수행하였습니다.

그 결과 카이제곱 통계량은 2.36, 자유도는 3, p-value 는 0.50이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

9.7.2 선거당일 투표하는 유권자 전체(%)

국민 전체	18세 이상 국민 전체	등록된 유권자 전체	선거 당일 투표하는 유권자 전체	계
4.9	24.5	16.3	54.3	100.0

정답률은 Red, Black 을 합하여 계산하는데, 54.3(%) 입니다.

9.8 Wason Selection Task

같은 구조의 문제를 추상적으로 물어볼 때와 구체적으로 사례를 들어서 물어볼 때의 정답률에 큰 차이가 있음에 유의하세요.

Red 집단에게는 추상적 질문을 먼저 던지고, 구체적 사례를 듣 질문을 나중에 던졌으며 Black 집단에게는 구체적 사례를 듣 질문을 먼저 던지고, 추상적 질문을 나중에 던졌습니다.

추상적인 질문에 대해서는 매우 낮은 정답률을 보이지만 구체적인 질문에 대해서는 정답률이 훨씬 올라가는 것을 관찰할 수 있습니다.

추상적인 질문에 찔찔매는 것이 경상입니다.

Wason Selection Task 는 인지 편향, 그 중에서도 확증 편향이 많은 사람들에게 공통적으로 나타난다는 것을 보여줍니다.

반증의 근거가 되는 자료는 잘 들여다 보려 하지 않습니다.

이 실험 결과의 어느 부분이 이를 입증하는 지 살펴 봅니다.

9.8.1 Red. Q7에 추상적 문제, Q8에 구체적 문제

Red

7. 다음 각 카드의 한쪽 면에는 숫자, 다른 쪽 면에는 알파벳이 있다. “한쪽 면이 모음이면 다른 * 쪽은 짹수가 있다”라는 규칙을 아래 네 장의 카드가 지키고 있는지 알아보려면 최소한 어느(어느) 카드를 들쳐봐야 하는가? (A,2와 같은 방식으로 답하세요. a,2 나 A, 2는 집계를 힘들게 합니다. 대소문자, 띄어쓰기 꼭 좀 지켜주세요.)

내 답변 _____

8. 술집에서 미성년자의 음주를 단속하고 있다. 아래 카드는 한쪽 면에는 고객의 나이, 다른 쪽 * 면에는 고객이 마시고 있는 음료의 종류를 기입한 것이다. “맥주를 마시려면, 21세 이상이어야 한다”라는 규칙을 지키고 있는지 알아보려면 아래 네 명 중 최소한 누구(누구)를 검문하여야 하는가?(Beer,31 와 같은 방식으로 답하세요. BEER,31이나 beer,31 또는 beer, 31 등은 집계를 힘들게 합니다. 대소문자, 띄어쓰기 꼭 좀 지켜주세요.)

내 답변 _____

9.8.2 Black. Q7에 구체적 문제, Q8에 추상적 문제

Black

7. 술집에서 미성년자의 음주를 단속하고 있다. 아래 카드는 한쪽 면에는 고객의 나이, 다른 쪽 면에는 고객이 마시고 있는 음료의 종류를 기입한 것이다. “맥주를 마시려면, 21세 이상이어야 한다”라는 규칙을 지키고 있는지 알아보려면 아래 네 명 중 최소한 누구(누구)를 검문하여야 하는가?(Beer,31 와 같은 방식으로 답하세요. BEER,31 이나 beer,31 또는 beer, 31 등은 집계를 힘들게 합니다. 대소문자, 띄어쓰기 꼭 좀 지켜주세요.)



내 답변 _____

8. 다음 각 카드의 한쪽 면에는 숫자, 다른 쪽 면에는 알파벳이 있다. “한쪽 면이 모음이면 다른 쪽은 짹수가 있다”라는 규칙을 아래 네 장의 카드가 지키고 있는지 알아보려면 최소한 어느(어느) 카드를 들쳐봐야 하는가?(A,2와 같은 방식으로 답하세요. a,2 나 A, 2 는 집계를 힘들게 합니다. 대소문자, 띄어쓰기 꼭 좀 지켜주세요.)



내 답변 _____

9.9 Q7. Red에 추상적 질문, Black에 구체적 질문

“한쪽 면이 모음이면 다른 쪽은 짹수가 있다.”

이 규칙은 “X이면 Y이다”의 형식으로 되어 있습니다.

이 논리식과 동등한 것은 대우인 “Y가 아니면 X가 아니다”입니다.

매우 불편한 구조이죠.

그렇다 보니까 이게 잘 떠오를 리가 없습니다.

‘선거여론조사의 발달’에서 학습한 바 있는 “표본을 공평하게 뽑으면 모집단의 특성을 잘 닮는다”의 대우가 바로 “모집단을 닮지 않으면 표본을 공평하게 뽑지 않은 것이다”입니다.

즉, 표본을 공평하게 뽑지 않아서 모집단을 제대로 닮지 않은 표본을 뽑았다는 것이죠.

주어진 네 장의 카드 중에서 한쪽 면이 모음인 것은 A입니다.

따라서 A는 우선 들쳐봐야 하는 카드이고, “한쪽 면이 모음이면 다른 쪽은 짹수가 있다”의 대우는 “한쪽 면이 짹수가 아니면 다른 쪽 면이 모음이 아니다”, 즉 “한쪽 면이 홀수이면 다른 쪽 면은 자음이다”가 됩니다.

짜수가 아니면 홀수이고, 모음이 아니면 자음이니까요.

따라서 홀수 카드를 들쳐봐야 합니다.

그래서 A, 3 두장을 들쳐보면 됩니다.

맥주와 연령 문제는 실생활과 밀접한 구체적인 사안이어서 “어, 맥주 마시는 사람 신분증 좀 보여주세요, 17살 미성년자는 지금 마시는 것이 맥주인가요?”하고 묻는데 익숙하지만 직관적으로 Beer와 17을 검문해야 한다고 추론하였는지 논증하는 연습이 필요합니다.

“맥주를 마시려면, 21세 이상이어야 한다”라는 규칙으로부터 “맥주”를 검문해야 하고, 검문으로부터 나이를 확인합니다.

그리고 이 규칙과 동등한 대우인 “21세 이상이 아니면, 맥주를 마실 수 없다”, 즉, “21세 미만이면 맥주를 마실 수 없다”로부터 “21세 미만”인 “17세”를 검문해야 하는 것입니다.

물론 실생활에서 접할 수 있는 문제이기 때문에 미성년자가 맥주를 마시고 있는 것은 아닌지 Beer와 17을 골라야 한다고 쉽게 답할 수 있지만 그 배경에는 이러한 논리가 숨어 있습니다.

9.9.1 집계

Table 9.20: Red에 추상적 질문, Black에 구체적 질문

	정답	오답	계
Red(추상적 질문)	32	59	91
Black(구체적 질문)	56	37	93
계	88	96	184

{A, 2, B, 3}에서 어느 카드를(들을) 골라야 “한쪽 면이 모음이면, 다른 쪽 면은 짹수이다”라는 규칙을 지키고 있는지 확인할 수 있는가? 라는 질문을 Red에 배치하고, {Beer, 31, Coke, 17}에서 누구를(들을) 검문해야 하는가라는 질문을 Black에 배치했습니다.

Red의 경우 총 91(명)이 응답하였고 32(명)이 정답인 {A, 3}를 올렸습니다.

구체적인 상황에 놓인 Black의 경우 총 93(명)이 응답하였고 56(명)이 정답인 {Beer, 17}을 올려서 구체적인 질문에 압도적으로 많은 정답이 나온 것을 알 수 있습니다.

이를 백분율로 비교해 보면

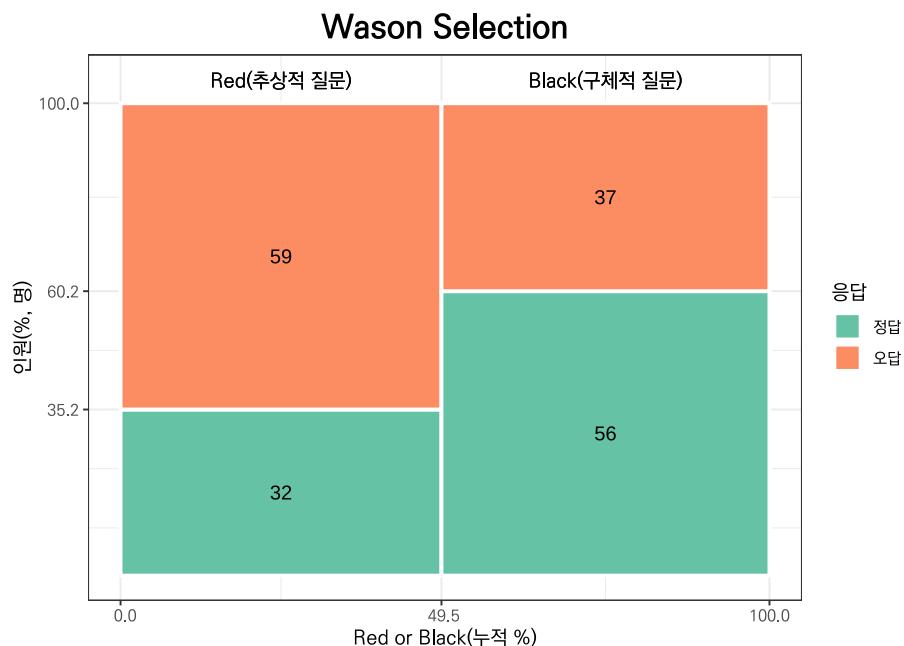
9.9.2 % 비교

	정답	오답	계
Red(추상적 질문)	35.2	64.8	100.0
Black(구체적 질문)	60.2	39.8	100.0

추상적인 질문으로 이루어진 Red에서는 35.2(%)가 정답을 올렸고, 구체적인 질문으로 이루어진 Black에서는 60.2(%)가 정답을 올려서 구체적인 질문에 압도적으로 많은 정답이 올라왔다는 것을 알 수 있습니다.

이 상황을 Mosaic Plot으로 살펴보겠습니다.

9.9.3 Mosaic Plot



Mosaic Plot으로부터 추상적 질문이 주어진 Red에서 정답 비율이 구체적 질문이 주어진 Black에서 정답 비율에 비해서 매우 적다는 것을 시각적으로 파악할 수 있습니다.

9.10 Q8. Red에 구체적 질문, Black에 추상적 질문

Q8에서는 Q7과 반대로 Red에 구체적 질문, Black에 추상적 질문을 배치하였습니다.

이렇게 하므로써 질문지에 응답한 모든 사람은 한 번씩 구체적 질문과 추상적 질문에 답할 수 있게 되었습니다.

집계 결과는 비슷합니다.

다만, 이렇게 추상적 질문을 먼저 배치하고 구체적 질문을 나중에 배치하느냐, 혹은 그 반대로 구체적 질문을 먼저 배치하고 추상적 질문을 나중에 배치한 것의 영향이 있는지를 파악한다면 학습 순서가 경답률과 어떤 관계가 있는지 파악할 수 있지 않을까 합니다.

9.10.1 집계

Table 9.22: Red에 구체적 질문, Black에 추상적 질문

	정답	오답	계
Red(구체적 질문)	58	33	91
Black(추상적 질문)	27	66	93
계	85	99	184

구체적인 질문을 배치한 Red의 경우 총 91(명)이 응답하였고 58(명)이 정답인 {Beer, 17}(을)를 올렸습니다.

추상적인 질문을 배치한 Black의 경우 총 93(명)이 응답하였고 27(명)이 정답인 {A,3}(을)를 올려서 구체적인 질문에 압도적으로 많은 정답이 나온 것을 알 수 있습니다.

이를 백분율로 비교해 보면

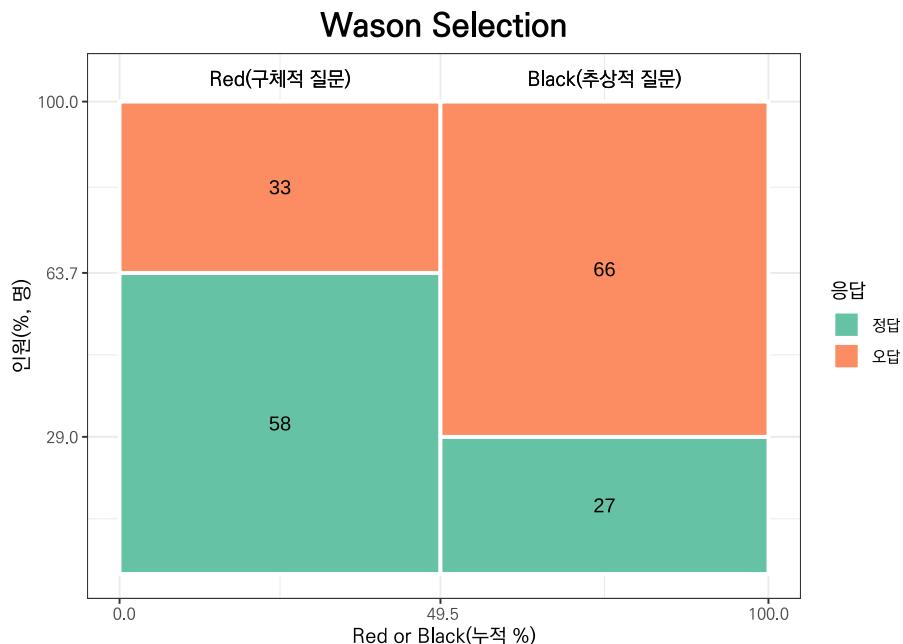
9.10.2 % 비교.

	정답	오답	계
Red(구체적 질문)	63.7	36.3	100.0
Black(추상적 질문)	29.0	71.0	100.0

구체적인 질문을 배치한 Red에서는 63.7(%)가 정답을 올렸고, 추상적인 질문을 배치한 Black에서는 29.0(%)가 정답을 올려서 구체적인 질문에 압도적으로 많은 정답이 올라왔다는 것을 알 수 있습니다.

이 상황을 Mosaic Plot으로 살펴보겠습니다.

9.10.3 Mosaic Plot



Mosaic Plot으로부터 구체적 질문이 주어진 Red의 정답 비율이 추상적 질문이 주어진 Black의 정답 비율에 비해서 매우 높다는 것을 시각적으로 파악할 수 있습니다.

9.11 Q9. 인지적 편향과 오류

Wason Selection Task에서 많은 사람들이 겪는 흔한 오류(예: 확증편향)을 설명합니다.

사람들은 보통 자신의 가설을 확인하기 위한 정보만 찾고, 반례가 될 수 있는 카드는 무시하려는 경향이 있습니다.

Peter C. Wason (1924-2003)의 연구에 의하면 정답을 찾아내는 백분율은 10%에 불과합니다.

여러분의 응답과 비교해 보세요.

9.11.1 집계

Table 9.24: Wason Selection Task 인지편향 분석

	A,2	A,3	Other	계
Red(추상적 질문 먼저)	36	30	25	91
Black(구체적 질문 먼저)	47	23	23	93
계	83	53	48	184

{A, 2, B, 3}에서 어느 카드를(들을) 골라야 “한쪽 면이 모음이면, 다른 쪽 면은 짹수이다”라는 규칙을 지키고 있는지 확인할 수 있는가?라는 질문이 Q7에 먼저 나오는 것을 Red에 배치하고, Black에서는 {A, 2, B, 3}에 대한 질문이 Q8에 나오도록 배치했습니다.

많은 사람들은 이 질문에 대해서 A와 2를 뒤집으려 합니다.

A는 모음이니까 확인해야 할 것 같고, 2는 짹수이니까 확인하려고 듭니다.

여기서 확증 편향이 나타납니다.

사람들은 주어진 규칙을 확인하기 위해 당장 눈에 들어오는 모음과 짹수, 즉 A와 2를 확인하려는 경향이 강합니다.

그러나 논리적으로 규칙을 검증하려면 짹수가 아닌 홀수 카드를 뒤집어야 합니다.

“한쪽 면이 모음이면, 다른 쪽 면은 짹수이다”와 동등한 규칙은 “한쪽 면이 짹수가 아니면, 다른 쪽 면은 모음이 아니다”이기 때문입니다.

쫙수가 아니면 홀수이니까 3을 뒤집어야 하는 것이죠.

추상적 질문이 먼저 Q7에 나온 Red의 경우 총 91(명)이 응답하였고 36(명)이 확증편향에서 비롯된 {A,2}를 올렸습니다.

정답인 {A,3}를 올린 30(명)보다 훨씬 많습니다.

추상적 질문이 Q8에 나온 Black의 경우 총 93(명)이 응답하였고 47(명)이 확증편향에서 비롯된 {A,2}를 올렸습니다.

정답인 {A,3}를 올린 23(명)보다 훨씬 많습니다.

합해서 83(명)이 확증편향에서 비롯된 {A,2}를 올렸고, 53(명)이 정답인 {A,3}를 올렸습니다.

이는 확증편향에서 비롯된 응답이 정답의 2배를 넘을 정도로 많다는 것을 보여줍니다.

백분율로 살펴 보겠습니다.

9.11.2 % 비교.

	A,2	A,3	Other	계
Red(추상적 질문 먼저)	39.6	33.0	27.5	100.0
Black(구체적 질문 먼저)	50.5	24.7	24.7	100.0
계	45.1	28.8	26.1	100.0

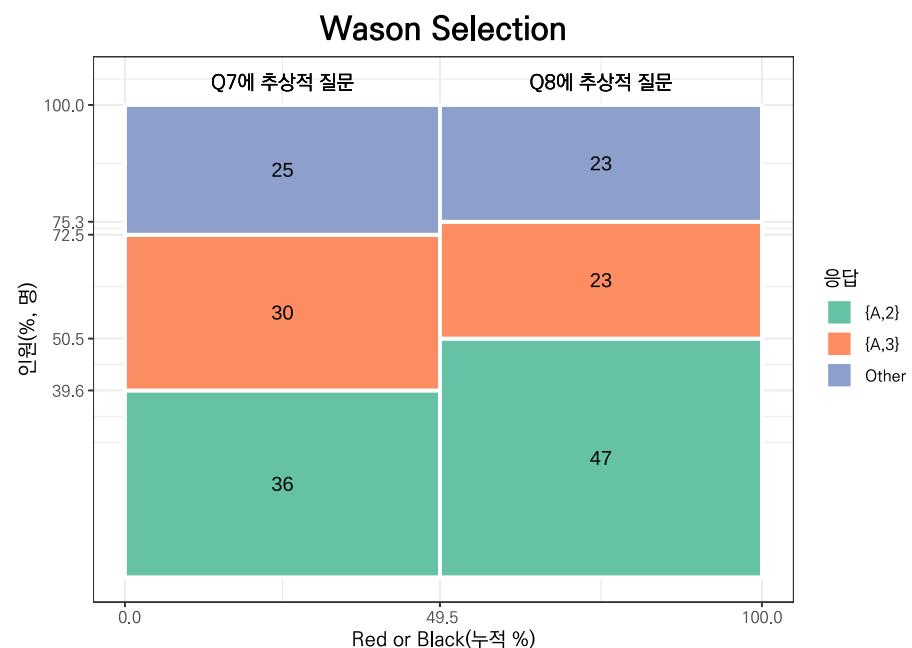
추상적인 질문이 먼저 Q7에 나온 Red에서는 39.6(%)가 확증편향에서 비롯된 응답 {A,2}를 올렸고, 33.0(%)가 정답인 {A,3}을 올렸는데, 추상적인 질문이 나중에 Q8에 나온 Black에서는 50.5(%)가 확증편향에서 비롯된 응답 {A,2}를 올렸고, 24.7(%)가 정답인 {A,3}을 올렸습니다.

합해서 보면 45.1(%)가 확증편향에서 비롯된 응답 {A,2}를 올렸고, 28.8(%)가 정답인 {A,3}을 올렸습니다.

확증편향으로 인한 응답이 정답보다 2배를 넘어가는 것을 다시 확인할 수 있습니다.

이 상황을 Mosaic Plot으로 살펴보겠습니다.

9.11.3 Mosaic Plot



Mosaic Plot으로부터 확증편향에서 비롯된 응답의 비율이 정답의 비율이나 기타 응답의 비율보다 월등히 높다는 것을 시각적으로 파악할 수 있습니다.

9.12 학습 순서의 영향

구체적 질문을 먼저 학습하고 추상적 질문을 학습하는 것과 추상적 질문을 먼저 학습하고 구체적 질문을 학습하는 방식 중에 어느 것이 더 나은지 비교한 결과 정답 인원은 매우 닮았는데, 순서에 따라 정답인원의 차이에는 통계적으로 유의한 차이가 관찰되지 않았습니다.

어떻게 해석할 수 있을까요?

9.12.1 집계표

Table 9.26: Wason Selection

	추상적 질문 정답	구체적 질문 정답	계
Red(추상적 질문 먼저)	32	58	90
Black(구체적 질문 먼저)	27	56	83

Table 9.27: Pearson's Chi-squared test with Yates' continuity correction: .

Test statistic	df	P value
0.06701	1	0.7957

추상적 질문을 Q7에 배치하고 구체적 질문을 Q8에 배치한 Red의 경우 추상적 질문과 구체적 질문에 정답을 올린 사람은 총 90(명)이고 구체적 질문을 Q7에 배치하고 추상적 질문을 Q8에 배치한 Black의 경우 추상적 질문과 구체적 질문에 정답을 올린 사람은 총 83(명)으로 별로 차이가 나지 않습니다.

추상적 질문을 Q8에 배치한 Black의 경우 27(명)이 정답을 올려서 추상적 질문을 먼저 학습한 경우 정답을 더 많이 내었지만 통계적으로 유의한 차이는 아닌 것으로 나타나고 있습니다.

카이제곱 통계량은 0.067, p-value는 0.80으로 통계적으로 유의한 차이를 관찰하지 못하였습니다.

따라서 학습 순서는 추상적 질문의 정답률에 영향을 미치지 못하고 있습니다.

백분율을 살펴 보겠습니다.

9.12.2 % 비교

Table 9.28: Wason Selection

	추상적 질문 정답	구체적 질문 정답
Red(추상적 질문 먼저)	54.2	50.9
Black(구체적 질문 먼저)	45.8	49.1
계	100.0	100.0

추상적 질문에 대한 Red, Black 간 정답률 차이와 구체적 질문에 대한 Red, Black 간 정답률 차이를 비교하였습니다.

추상적 질문에 대한 전체 정답 중에서 추상적 질문을 먼저 제시한 Red 가 54.2(%)를 차지하여 추상적 질문을 나중에 제시한 Black 보다 높습니다만 그 차이는 앞에서 살펴 본 것처럼 통계적으로 유의하지는 않습니다.



Mosaic Plot으로부터 구체적 질문이 먼저 주어진 Red나 구체적 질문이 나중에 주어진 Black이나 정답을 올린 인원이나 백분율이 비슷한다는 것을 시각적으로 파악할 수 있습니다.

9.12.3 합산

실험에 참여한 어느 누구나 추상적 문제와 구체적 문제를 한 번씩 풀게 됩니다.

학습 순서의 영향은 없는 것으로 파악되었으니까 추상적 문제의 정답률과 구체적 문제의 정답률을 합쳐서 비교하는 것이 합리적입니다.

9.12.4 집계표

	정답	오답	계
추상적 문제	59	125	184
구체적 문제	114	70	184

추상적 질문에 답한 사람 총 184(명) 중에 정답을 올린 사람은 모두 59(명)이고
 구체적 질문에 답한 사람 총 184(명) 중에 정답을 올린 사람은 모두
 114(명)입니다. 백분율로 비교해 보면

9.12.5 % 비교

	정답	오답	계
추상적 문제	32.1	67.9	100.0
구체적 문제	62.0	38.0	100.0

추상적 질문에 정답을 올린 사람의 백분율은 32.1(%)이고 구체적 질문에 정답을
 올린 사람의 백분율은 32.1(%)입니다.

추상적 질문의 정답율이 구체적 질문의 정답율에 비하여 월등히 낮다는 것을 알 수
 있습니다. 이를 시각적으로 비교해 보겠습니다.

9.12.6 Barplot



이 경우에는 막대그래프로 표현하는 것이 보다 시각적으로 두 상황을 비교하기에 더 효과적입니다.

추상적 질문의 응답 중에서 정답의 비율이 구체적 질문의 응답 중 정답의 비율보다 월등히 적다는 것이 시각적으로 잘 드러나고 있습니다.

9.13 마감 시간으로부터 제출 시간의 분포

9.13.1 분포표

Table 9.31: 일 단위

	[0,1]	(1,2]	(2,3]	(3,4]	(4,5]	(5,6]	(6,7]	(7,8]	(8,9]	(9,10]	10,11]	11,12]	12,13]	13,14]
Red	0	0	0	0	0	0	0	5	16	12	16	18	24	91
Black	0	0	0	0	0	0	0	6	14	11	14	16	32	93
계	0	0	0	0	0	0	0	11	30	23	30	34	56	184

분포표로부터 두 가지 문제를 살펴보겠습니다.

첫째, 날마다 고르게 제출하는가?

둘째, Red, Black 간에 통계적으로 유의한 차이가 있는가?

각 문제를 살펴보기 위해서는 분포표의 일부분을 대상으로 카이제곱 테스트를 수행합니다.

9.13.2 날마다 고르게 제출하는가?

(8,9]	(9,10]	(10,11]	(11,12]	(12,13]	(13,14]
11	30	23	30	34	56

Table 9.33: Chi-squared test for given probabilities: .

Test statistic	df	P value
35.85	5	1.019e-06 * * *

날마다 고르게 제출하는지 알아 보았습니다.

분포표의 “계”행에서 ’계’열을 제외하고 카이제곱테스트를 수행합니다.

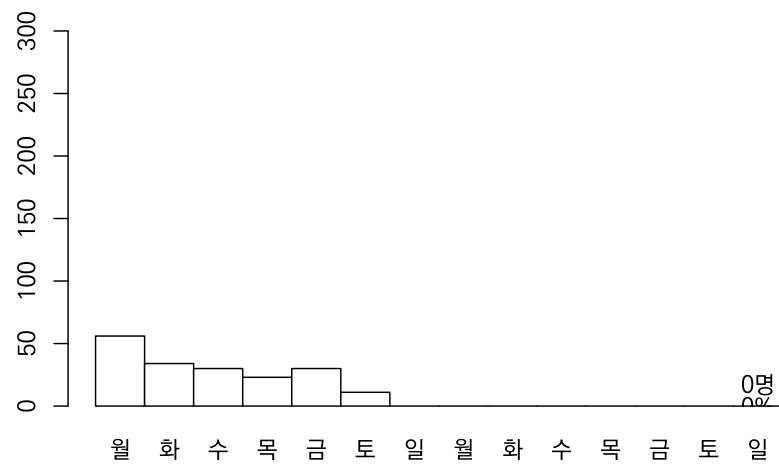
분포표 만으로도 쉽게 파악할 수 있지만 카이제곱테스트가 명확히 해 줍니다.

카이제곱 통계량은 35.85, 자유도는 5.00, p-value 는 1e-06 이므로 아직까지는 날짜별로 고르게 제출하고 있습니다.

막대그래프로 살펴 보겠습니다.

9.13.3 막대그래프

Quiz250414 (184명 제출)



9.13.4 Red, Black 간에 닮았는가?

	(8,9]	(9,10]	(10,11]	(11,12]	(12,13]	(13,14]
Red	5	16	12	16	18	24
Black	6	14	11	14	16	32

Table 9.35: Pearson's Chi-squared test: .

Test statistic	df	P value
1.64	5	0.8964

제출시간의 분포가 Red, Black 간에 닮았는지 알아 보았습니다.

이번에는 분포표의 첫번째와 두번째 행, '계'열을 제외한 나머지 열에 대해서 카이제곱테스트를 수행합니다.

카이제곱 통계량은 1.64, 자유도는 5, p-value 는 0.8964 이므로 제출 시간의 분포는 Red, Black 간에 통계적으로 유의한 차이가 관찰되지 않습니다.

이 사실을 Mosaic Plot 을 이용하여 시각적으로 살펴보겠습니다.

답았다고 느껴지나요?

9.13.5 Mosaic Plot

