

Quiz Reports: 2025 Spring

이기원 (Kee-Won Lee)

2025-06-06

Contents

소개	13
1 1주차 데이터 실험 집계	15
1.1 실험의 목적	15
1.2 Q1. Dewey as good as elected, statistics convince Roper	17
1.3 Q2. Statistics is the science of learning from data,	18
1.4 Q3. How to lie with statistics	19
1.5 Q4. 종부세	20
1.6 마감 시간으로부터 제출 시간의 분포	24
2 2주차 데이터 실험 집계	27
2.1 실험의 목적	27
2.2 Q1. 춘추전국시대에 국가통계관리의 중요성 강조	30
2.3 Q2. 국가정책을 수립하는 데 통계의 역할	31
2.4 Q3. 우리나라 생산가능인구 감소 시기	32
2.5 Q4. 우리나라 총인구 최대 시기	33
2.6 Q5. 소멸위험 단계 개선 지역	34
2.7 Q6. 조출생률과 합계출산율	35
2.8 Q7. 눈속임 그래프(Cheating Charts)	36
2.9 마감 시간으로부터 제출 시간의 분포	40

3 3주차 데이터 실험 집계	45
3.1 실험의 목적	45
3.2 Q1. 국세와 지방세 비중	47
3.3 Q2. 조세부담률	48
3.4 Q3. OECD 국민부담률	49
3.5 Q4. 과세대상 균로소득 1,200만 원	50
3.6 Q5. 소득세 실효세율	51
3.7 Q6. 기업규모별 과세 현황	52
3.8 Q7. 국민부담률 적정 수준 : 아일랜드와 OECD 평균	54
3.9 마감 시간으로부터 제출 시간의 분포	57
4 4주차 데이터 실험 집계	61
4.1 실험의 목적	61
4.2 Q1. 세종대왕 시대 조세제도	62
4.3 Q2. 공법도입에 대한 대신들의 찬성율	64
4.4 Q3. 공법도입과 품관총민들의 찬반	65
4.5 Q4. 공법	66
4.6 Q5. 1423년 조선시대 호구와 인구	67
4.7 Q6. 지방관료와 품관총민	68
4.8 Q7. 부연설명의 효과 : 주당 균로 69시간제 도입 찬반	69
4.9 마감 시간으로부터 제출 시간의 분포	72
5 5주차 데이터 실험 집계	77
5.1 실험의 목적	77
5.2 Q1. 한글의 문자 유형	78
5.3 Q2. 정보혁명과 문자 체계	79
5.4 Q3. 알기 힘든 전문 용어	81
5.5 Q4. 해방직후 비문해율	82
5.6 Q5. 세대간 문해력 격차	83
5.7 Q6. 문해력 격차의 파급효과	84
5.8 Q7. 프레임을 설정하는 단어의 힘	85
5.9 마감 시간으로부터 제출 시간의 분포	88

CONTENTS	5
6 국민문해력조사 집계 결과	93
6.1 응답 집계	93
6.2 막대그래프	93
6.3 문해력 점수 계산	94
6.4 Red and Black 비교	95
6.5 문해력 등급 판정	100
6.6 유형별 정답률	100
6.7 어려운 문제?	101
6.8 정답률이 낮은 문제들	102
7 6주차 데이터 실험 집계	107
7.1 실험의 목적	107
7.2 Q1. 월간 독서율	109
7.3 Q2. 지역 및 지역크기별 가구수 비례 무작위추출법	110
7.4 Q3. 한달 독서량의 분포	111
7.5 Q4. 최근 1개월간 독서량	112
7.6 Q5. 20대의 연간독서율	113
7.7 Q6. 50대의 연간독서율	114
7.8 Q7. The more, the better? : 내가 남보다, 혹은 남이 나보다	115
7.9 마감 시간으로부터 제출 시간의 분포	118
8 7주차 데이터 실험 집계	123
8.1 실험의 목적	123
8.2 Q1. 통계학의 기본원리	125
8.3 Q2. 리터러리 디제스트의 실패	126
8.4 Q3. 1948년, 여론조사가 듀이를 당선시킨 해	127
8.5 Q4. 1948 미 대선 이후	128
8.6 Q5. 표본오차를 반으로 줄이려면?	129
8.7 Q6. 대선 여론조사의 목표모집단?	130
8.8 Wason Selection Task	131
8.9 Q7. Red에 추상적 질문, Black에 구체적 질문	132
8.10 Q8. Red에 구체적 질문, Black에 추상적 질문	134

8.11 Q9. 인지적 편향과 오류	136
8.12 학습 순서의 영향	139
8.13 마감 시간으로부터 제출 시간의 분포	142
9 8주차 데이터 실험 집계	147
9.1 실험의 목적	147
9.2 Q1. 랜덤화 플라시보 대조군 설계의 특징	148
9.3 Q2. 백신의 효과 확인	150
9.4 Q3. 3학년의 발병율	151
9.5 Cognitive Response Test	152
9.6 Q4. 야구 방망이와 공	154
9.7 Q5. 가발 만드는 기계	155
9.8 Q6. 호수에 수련	157
9.9 마감 시간으로부터 제출 시간의 분포	159
10 8주차 데이터 실험 집계	163
10.1 실험의 목적	163
10.2 Q1. 랜덤화 플라시보 대조군 설계의 특징	164
10.3 Q2. 백신의 효과 확인	166
10.4 Q3. 3학년의 발병율	167
10.5 Cognitive Response Test	168
10.6 Q4. 야구 방망이와 공	170
10.7 Q5. 가발 만드는 기계	172
10.8 Q6. 호수에 수련	173
10.9 마감 시간으로부터 제출 시간의 분포	175
11 옥스포드 행복도 조사 집계	179
11.1 실험의 목적	179
11.2 행복 점수 비교	180
11.3 소요시간 비교	183
11.4 마감 시간으로부터 제출 시간의 분포	185

CONTENTS	7
12 옥스포드 행복도 조사 집계	189
12.1 실험의 목적	189
12.2 행복 점수 비교	190
12.3 소요시간 비교	193
12.4 마감 시간으로부터 제출 시간의 분포	195
13 9주차 데이터 실험 집계	199
13.1 실험의 목적	199
13.2 Q1. OECD 2022년 9월 실업률 비교	200
13.3 Q2. 고용률의 산식	201
13.4 Q3. 취업자의 조건	202
13.5 Q4. 평균과 중위값	204
13.6 Q5. 평균과 중위값	205
13.7 Q6. 연 500만원 이하 근로소득자의 비율	206
13.8 Q7. 디폴트 효과 : “… 하면 표시하세요”의 효과	207
13.9 Thinking, Fast and Slow	210
13.10 마감 시간으로부터 제출 시간의 분포	211
14 9주차 데이터 실험 집계	215
14.1 실험의 목적	215
14.2 Q1. OECD 2022년 9월 실업률 비교	216
14.3 Q2. 고용률의 산식	217
14.4 Q3. 취업자의 조건	218
14.5 Q4. 평균과 중위값	220
14.6 Q5. 평균과 중위값	221
14.7 Q6. 연 500만원 이하 근로소득자의 비율	222
14.8 Thinking, Fast and Slow	226
14.9 마감 시간으로부터 제출 시간의 분포	227

15 10주차 데이터 실험 집계	231
15.1 실험의 목적	231
15.2 Q1. Wilkinson	232
15.3 Q2. 소득 상위1%의 둑과 시대적 사건들	233
15.4 Q3. Piketty의 21세기 자본	234
15.5 Q4. World Top Income Database	235
15.6 Q5. 황금기와 신자유주의 시기	237
15.7 Q6. 주역	239
15.8 Q7. Prospect Theory : 이득과 손실	240
15.9 마감 시간으로부터 제출 시간의 분포	243
16 10주차 데이터 실험 집계	247
16.1 실험의 목적	247
16.2 Q1. Wilkinson	248
16.3 Q2. 소득 상위1%의 둑과 시대적 사건들	249
16.4 Q3. Piketty의 21세기 자본	250
16.5 Q4. World Top Income Database	251
16.6 Q5. 황금기와 신자유주의 시기	253
16.7 Q6. 주역	255
16.8 Q7. Prospect Theory : 이득과 손실	256
16.9 마감 시간으로부터 제출 시간의 분포	259
17 식생활습관 교육동영상 시청 효과	263
17.1 실험의 목적	263
17.2 동영상 시청 전후 비교	263
17.3 Red and Black (랜덤화 효과)	267
18 식생활습관 교육동영상 시청 효과	273
18.1 실험의 목적	273
18.2 동영상 시청 전후 비교	273
18.3 Red and Black (랜덤화 효과)	277

CONTENTS	9
19 11주차 데이터실험 집계	285
19.1 실험의 목적	285
19.2 Q1. 1분위와 5분위의 평균소비성향 비교	286
19.3 Q2. 낙수효과	288
19.4 Q3. 최고한계세율과 상위1% 소득점유율	289
19.5 Q4. 최고한계세율과 상위1%, 하위99%의 소득증가	291
19.6 Q5. 생산성과 급료의 격차	292
19.7 Q6. LM3	294
19.8 Q7. 국공립대 등록금 국제 비교	295
19.9 마감 시간으로부터 제출 시간의 분포	298
20 11주차 데이터실험 집계	303
20.1 실험의 목적	303
20.2 Q1. 1분위와 5분위의 평균소비성향 비교	304
20.3 Q2. 낙수효과	306
20.4 Q3. 최고한계세율과 상위1% 소득점유율	307
20.5 Q4. 최고한계세율과 상위1%, 하위99%의 소득증가	309
20.6 Q5. 생산성과 급료의 격차	310
20.7 Q6. LM3	312
20.8 Q7. 국공립대 등록금 국제 비교	313
20.9 마감 시간으로부터 제출 시간의 분포	317
21 Big Five Personality and Multiple Intelligence	321
21.1 성격특성별 분포	321
21.2 3대 강점지능 분포	327
22 Big Five Personality and Multiple Intelligence	331
22.1 성격특성별 분포	331
22.2 3대 강점지능 분포	337

23 12주차 데이터 실험 집계	341
23.1 실험의 목적	341
23.2 Q1. Latane and Darley	342
23.3 Q2. Stanley Milgram	343
23.4 Q3. Solomon Asch	344
23.5 Q4. BFI 줄 굿기	345
23.6 Q5. 다중지능 짹 짓기	346
23.7 Q6. 성공한 사람들의 공통점	347
23.8 Q7. Framing Effect : 삶과 죽음	348
23.9 마감 시간으로부터 제출 시간의 분포	350
24 12주차 데이터 실험 집계	355
24.1 실험의 목적	355
24.2 Q1. Latane and Darley	356
24.3 Q2. Stanley Milgram	357
24.4 Q3. Solomon Asch	358
24.5 Q4. BFI 줄 굿기	359
24.6 Q5. 다중지능 짹 짓기	360
24.7 Q6. 성공한 사람들의 공통점	361
24.8 Q7. Framing Effect : 삶과 죽음	362
24.9 마감 시간으로부터 제출 시간의 분포	364
25 13주차 데이터 실험 집계	369
25.1 실험의 목적	369
25.2 Q1. Birthday Problem	371
25.3 Q2. Matching Problem	379
25.4 Q3. 직관과 어긋나는 용어	385
25.5 Q4. Monty Hall 문제	387
25.6 마감 시간으로부터 제출 시간의 분포	392

CONTENTS	11
26 직관의 한계 데이터 실험 집계	395
26.1 실험의 목적	395
26.2 Q1. Birthday Problem	397
26.3 Q2. Matching Problem	404
26.4 Q3. 직관과 어긋나는 용어	410
26.5 Q4. Monty Hall 문제	412
26.6 마감 시간으로부터 제출 시간의 분포	417
참고문헌	421

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.

제 1 장 1주차 데이터 실험 집계

1.1 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.

1.1.1 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 $1/20$ 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.

1.1.2 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 $1/2$ 을 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.

1.2. Q1. DEWEY AS GOOD AS ELECTED, STATISTICS CONVINCE ROPER17

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

이 헤드라인의 출처는 Freedman 등이 저술한 Statistics에서 1948년 대선 여론조사에 대해서 설명한 내용입니다 (Freedman et al. 1998).

-
1. 다음 문장에서 statistics의 의미는? *

Dewey as good as elected, statistics convince Roper.

- 통계학
- 통계
- 자료
- 통계청
- 정보

1.2.1 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

표 1.3: 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 0.09060이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

1.2.2 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.

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

이 문장의 출처는 미국통계협회 (ASA, American Statistical Association) 의 홈페이지 등입니다 (American Statistical Association n.d.) .

2. 다음 문장에서 statistics 의 의미는? *

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

- 통계학
 - 통계
 - 자료
 - 통계청
 - 정보
-

1.3.1 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

표 1.6: 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 0.24840이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

1.3.2 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.

1.4 Q3. How to lie with statistics

이 문장은 Darrel Huff가 저술한 책 제목입니다 (Huff 1954) .

3. 다음 문장에서 statistics 의 의미는? *

How to lie with statistics

- 통계학
- 통계
- 자료
- 통계청
- 정보

1.4.1 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

표 1.9: 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 0.20160이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

1.4.2 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.

1.5 Q4. 종부세

질문지의 선택지에 부연 설명을 추가하여 응답을 왜곡할 수 있다는 우려는 설문조사 설계에서 잘 알려진 문제입니다. 이러한 현상은 ‘응답 편향(response bias)’ 또는 ‘질문지 편향(questionnaire bias)’의 한 형태로 간주되며, 다양한 연구에서 이에 대한 사례와 분석이 이루어졌습니다.

“A Catalog of Biases in Questionnaires” 이 논문은 질문지에서 발생할 수 있는 48가지 편향 유형을 식별하고 분류하며, 각 유형에 대한 예시를 제공합니다. 특히, 질문의 형식과 응답 옵션의 설계가 응답자의 선택에 어떻게 영향을 미칠 수 있는지를 다루고 있습니다 (Boynton and Greenhalgh 2002).

“Response Biases in Standardised Surveys” by Kathrin Bogner & Uta Landrock 이 문서는 표준화된 설문조사에서 발생하는 응답 편향에 대해 설명하며, 질문의 형식과 응답 옵션이 응답자의 반응에 미치는 영향을 분석합니다 (Bogner and Landrock 2016).

“Biased Questions: How to Identify & Fix Them in Surveys” 이 자료는 편향된 질문을 식별하고 수정하는 방법에 대해 설명하며, 부연 설명이 포함된 선택지가 응답자의 선택에 어떻게 영향을 줄 수 있는지를 다룹니다 (Genroe 2020).

NBS(전국지표조사)가 2024년 7월25일 발표한 종합부동산세 관련 여론조사 문항을 편집하여 선택지에 부연설명을 붙였을 때 응답에 어느 정도의 영향을 미치는지 알아보고자 하였습니다 (중앙선거여론조사심의위원회 2024). 지난 학기에는 응답에 영향을 미친다는 것이 통계적으로 매우 유의한 수준으로 관찰되었습니다. 그러나, 이번 학기는 그렇지 않네요.

“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 차이를 관찰하지 못하였습니다.

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

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

1.5.1 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?

- 남세자들의 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

1.5.2 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

표 1.12: 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, Black에 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?

1.5.3 % 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

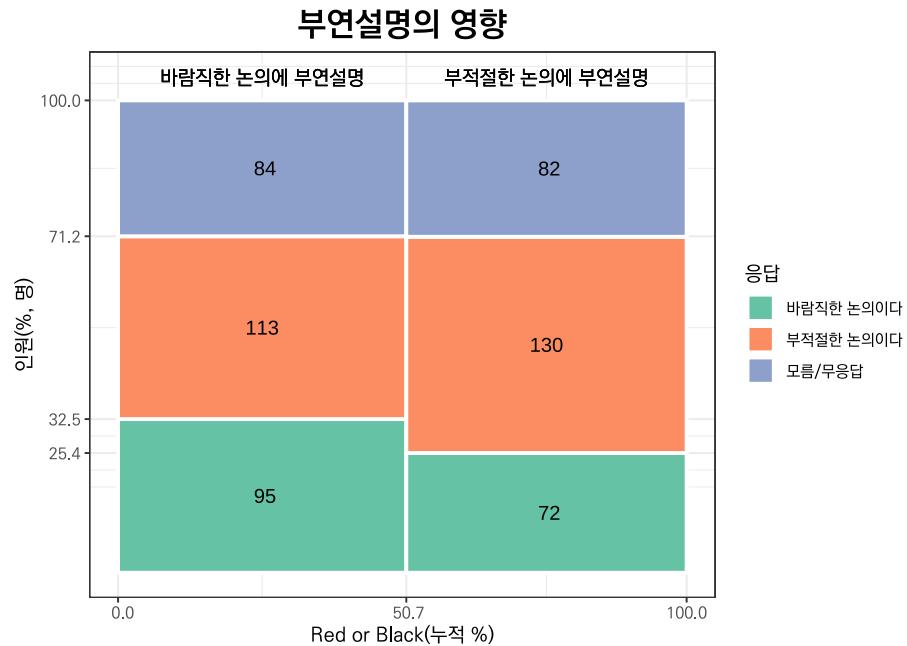
“바람직한 논의이다”에 부연설명을 붙인 Red에서 “바람직한 논의이다”라고 응답하는 사람들의 백분율, 32.5%은 “부적절한 논의이다”에 부연설명을 붙인 Black에서 “바람직한 논의이다”라고 응답하는 사람들의 백분율, 25.4% 보다 높습니다.

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.

1.5.4 Mosaic Plot

```
## Keep up to date with changes at https://tidyverse.org/blog/
```



Mosaic Plot 은 이 집계결과를 시각적으로 잘 보여줍니다.

“바람직한 논의이다”에 부연설명을 붙인 Red 에서 “바람직한 논의이다”라고 응답한 백분율이 “부적절한 논의이다”에 부연설명을 붙인 Black 에서 “바람직한 논의이다”라고 응답한 백분율보다 높고, Black 에서 “부적절한 논의이다”라고 응답한 백분율은 Red 에서 “부적절한 논의이다”라고 응답한 백분율보다 월등히 높습니다.

1.6 마감 시간으로부터 제출 시간의 분포

1.6.1 분포표

표 1.14: 일 단위

	14일	13일	12일	11일	10일	9일	8일	7일	6일	5일	4일	3일	2일	1일	계
Red	46	11	6	11	6	14	10	42	21	19	15	30	35	26	292
Black	35	10	6	7	5	12	12	34	22	17	26	37	29	32	284
계	81	21	12	18	11	26	22	76	43	36	41	67	64	58	576

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.

1.6.2 날마다 고르게 제출하는가?

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

표 1.16: 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.

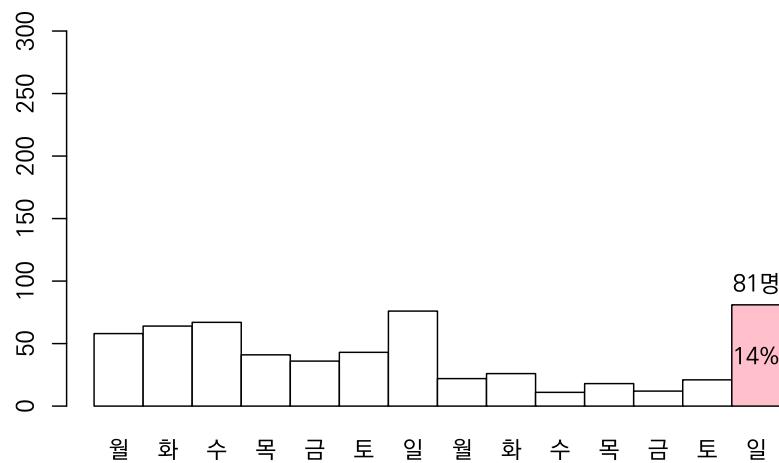
분포표 만으로도 쉽게 파악할 수 있지만 카이제곱테스트가 명확히 해 줍니다.

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

막대그래프로 살펴 보겠습니다.

1.6.3 막대그래프

Quiz250303 (576명 제출)



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

1.6.4 Red, Black 간에 닮았는가?

	14일	13일	12일	11일	10일	9일	8일	7일	6일	5일	4일	3일	2일	1일
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

표 1.18: Pearson's Chi-squared test: .

Test statistic	df	P value
8.59	13	0.8032

제출시간의 분포가 Red, Black 간에 닮았는지 알아 보았습니다.

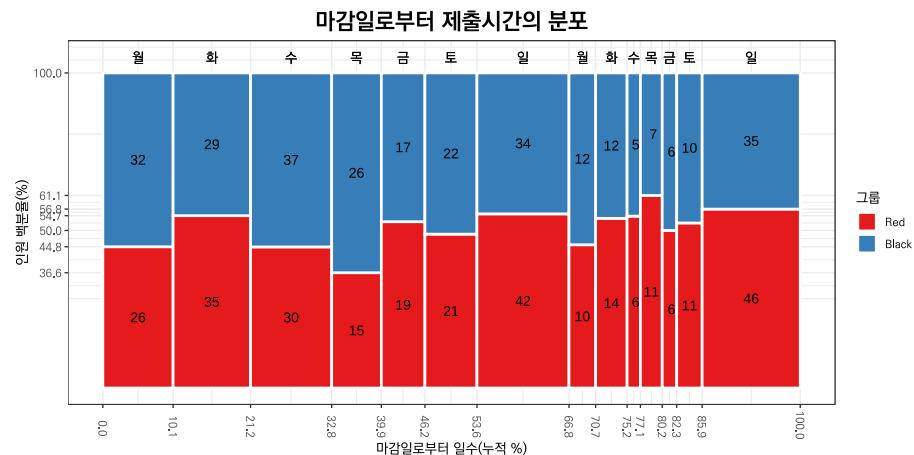
이번에는 분포표의 첫번째와 두번째 행, '계'열을 제외한 나머지 열에 대해서 카이제곱테스트를 수행합니다.

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.

닮았다고 느껴지나요?

1.6.5 Mosaic Plot



제 2 장 2주차 데이터 실험 집계

2.1 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 지각하는 데 차이가 있는지 알아봅니다.

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

2.1.1 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
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

제출시간	학번	랜덤화출석부	구글예습퀴즈
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.

2.1.2 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.

2.2 Q1. 춘추전국시대에 국가통계관리의 중요성 강조

춘추전국시대에 국가통계관리의 중요성을 역설한 사람은 누구인가? *

- 공자
- 맹자
- 관자
- 노자
- 장자

2.2.1 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

표 2.4: Pearson's Chi-squared test with simulated p-value (based on 2000 replicates): .

Test statistic	df	P value
9.064	NA	0.06247

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 0.06250이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

2.2.2 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.

2.3 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

2.3.1 통계의 중요성(집계표)

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

표 2.7: Pearson's Chi-squared test with simulated p-value (based on 2000 replicates): .

Test statistic	df	P value
2.056	NA	0.6617

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 0.6617이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

2.3.2 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.

2.4 Q3. 우리나라 생산가능인구 감소 시기

우리나라의 생산가능인구는 다음 중 어느 연도부터 감소하기 시작하였는가? *

- 2012
- 2017
- 2022
- 2027

2.4.1 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

표 2.10: 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?

2.4.2 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.

2.5 Q4. 우리나라 총인구 최대 시기

우리나라의 총인구는 어느 연도에 가장 많았는가? *

- 2018
- 2019
- 2020
- 2021

2.5.1 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

표 2.13: Pearson's Chi-squared test with simulated p-value
(based on 2000 replicates): .

Test statistic	df	P value
8.269	NA	0.03648 *

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

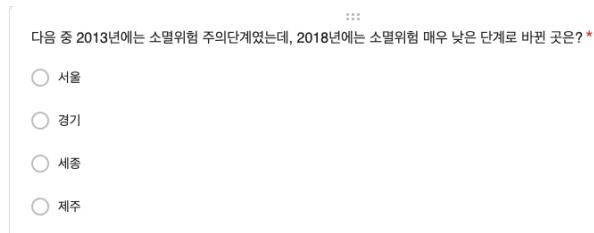
그 결과 카이제곱 통계량은 8.269, 자유도는 NA, p-value 는 0.03650이므로 Red, Black 간에 통계적으로 유의한 차이를 보이고 있습니다.

2.5.2 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.

2.6 Q5. 소멸위험 단계 개선 지역



2.6.1 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

표 2.16: Pearson's Chi-squared test with simulated p-value
(based on 2000 replicates): .

Test statistic	df	P value
0.7955	NA	0.8506

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 0.8506이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

2.6.2 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.

2.7 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

2.7.1 조출생률과 합계출산율(집계표)

	합계출산율 A			합계출산율 A :	합계출산율 A :	합계출산율 A :	계
	합계출산율 A :	B :	C :				
Red	0.25	0.125	0.5	0.125	0.125	0.5	282
Black	0.125	0.25	0.5	0.25	0.25	0.5	281
계	0.25	0.25	0.5	0.25	0.25	0.5	563

표 2.19: 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 0.03350이므로 Red, Black 간에 통계적으로 유의한 차이를 보이고 있습니다.

2.7.2 조출생률과 합계출산율(%)

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

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

2.8 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 인식하는 지 언론기관에서 발표한 눈속임 그래프를 보여줍니다.

Red에는 원형그래프의 각도를 속이고, Black에는 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 (모름)	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 요약하였는데, 막대그래프가 여론조사 결과를 제대로 보여주고 있나요?

후보자	지지율 (%)
문재인	33
안희정	22
황교안	9
안철수	9
이재명	5
유승민	2
기타	1

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

2.8.1 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

표 2.22: 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.

2.8.2 % 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

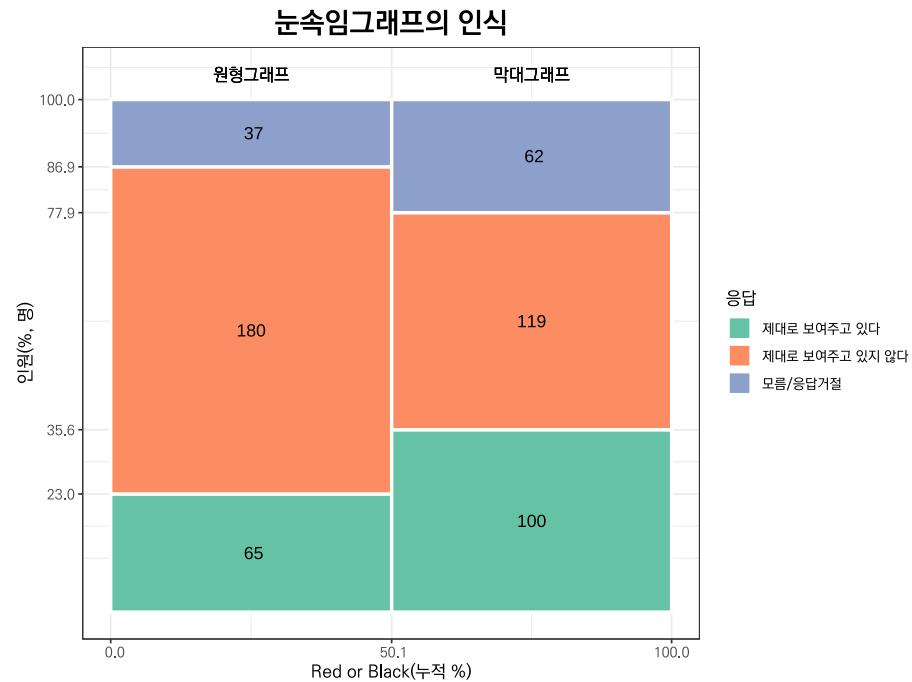
원형그래프의 각도를 왜곡한 Red에서 여론조사 결과를 “제대로 보여주고 있다”고 응답하는 사람들의 백분율, 23.0(%)은 “제대로 보여주고 있지 않다”고 응답하는 사람들의 백분율, 63.8(%) 보다 매우 낮습니다.

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.

2.8.3 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.

2.9 마감 시간으로부터 제출 시간의 분포

2.9.1 분포표

표 2.24: 일 단위

	14일	13일	12일	11일	10일	9일	8일	7일	6일	5일	4일	3일	2일	1일	계
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.

2.9.2 날마다 고르게 제출하는가?

14일	13일	12일	11일	10일	9일	8일	7일	6일	5일	4일	3일	2일	1일
111	28	20	9	18	15	21	94	39	31	41	34	47	55

표 2.26: Chi-squared test for given probabilities: .

Test statistic	df	P value
281.6	13	1.685e-52 ***

날마다 고르게 제출하는지 알아 보았습니다.

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

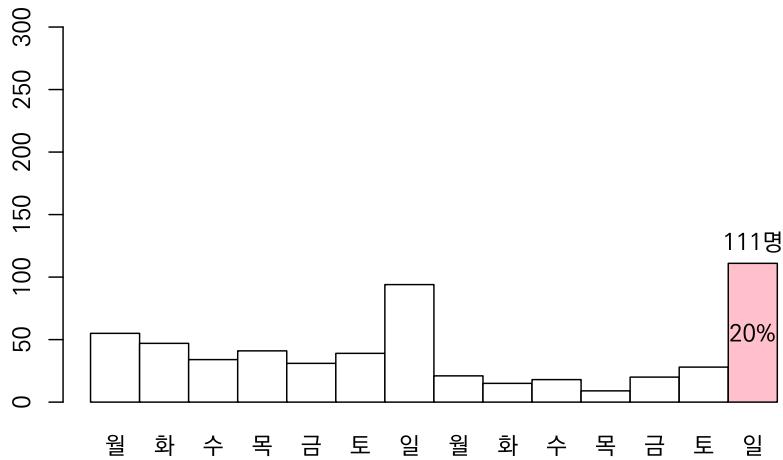
분포표 만으로도 쉽게 파악할 수 있지만 카이제곱테스트가 명확히 해 줍니다.

카이제곱 통계량은 281.60, 자유도는 13.00, p-value 는 1.7e-52 이므로 결코 고르게 제출한다고 말할 수 없겠습니다.

막대그래프로 살펴 보겠습니다.

2.9.3 막대그래프

Quiz250310 (563명 제출)



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

2.9.4 Red, Black 간에 닮았는가?

	14일	13일	12일	11일	10일	9일	8일	7일	6일	5일	4일	3일	2일	1일
Red	52	17	9	6	12	8	12	51	20	13	21	13	21	27
Black	59	11	11	3	6	7	9	43	19	18	20	21	26	28

표 2.28: Pearson's Chi-squared test: .

Test statistic	df	P value
9.39	13	0.7429

제출시간의 분포가 Red, Black 간에 닮았는지 알아 보았습니다.

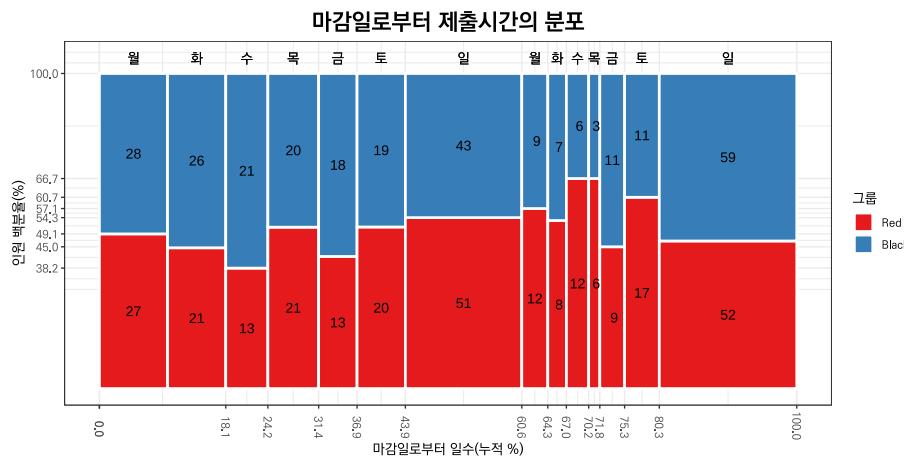
이번에는 분포표의 첫번째와 두번째 행, '계'열을 제외한 나머지 열에 대해서 카이제곱테스트를 수행합니다.

카이제곱 통계량은 9.390, 자유도는 13, p-value 는 0.7429 이므로 제출 시간의 분포는 Red, Black 간에 통계적으로 유의한 차이가 관찰되지 않습니다.

이 사실을 Mosaic Plot 을 이용하여 시각적으로 살펴보겠습니다.

답았다고 느껴지나요?

2.9.5 Mosaic Plot



제 3 장 3주차 데이터 실험 집계

3.1 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.

3.1.1 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 40하이거나 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.

3.1.2 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.

3.2 Q1. 국세와 지방세 비중

1. 2020년 우리나라 국세와 지방세의 세수 비중은?*

- 78:22
- 77:23
- 76:24
- 75:25
- 74:26

3.2.1 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

표 3.3: 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?

3.2.2 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.

3.3 Q2. 조세부담률

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

3.3.1 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

표 3.6: 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?

3.3.2 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.

3.4 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.
-

3.4.1 OECD 국민부담률(집계표)

조세 뿐 아니라 우리나라의 사회보장기여금을 국민부담률은 미국의 국민부담률은						
포함하는 모든 강제적 납부액을 대비 비율로 표시한 것이다	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.		계
Red	11	40	16	12	198	277
Black	8	32	18	19	199	276
계	19	72	34	31	397	553

표 3.9: 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 0.54760이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

3.4.2 OECD 국민부담률(%)

조세 뿐 아니라 사회보장기여금을 포함하는 모든 강제적 납부액을 명목 GDP 대비 비율로 표시한 것이다	우리나라의 국민부담률은 2010년 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.	계
3.4	13.0	6.1	5.6	71.8	100.0

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

3.5 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

3.5.1 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

표 3.12: 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?

3.5.2 과세대상 균로소득 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.

3.6 Q5. 소득세 실효세율

5. 2020년 우리나라 소득세의 실효세율은 얼마인가? *

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

3.6.1 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

표 3.15: 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?

3.6.2 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.

3.7 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.

3.7.1 기업규모별 과세 현황(집계표)

				그 외	법인당 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	상호출자제한기업은 신고법인수의 8.3%를 차지하는 데 부담하는 세액은 41.2%이다	상호출자제한기업 〉 그외 대기업 〉 중견기업 〉 중소기업의 순서로 많이 부담하였다.	법인당 평균세액은
Red	14	37	38	50	138
Black	18	33	21	61	143
계	32	70	59	111	281
					553

표 3.18: 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 0.14660이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

3.7.2 기업규모별 과세 현황(%)

			그 외	법인당 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	상호출자제한기업은 신고법인수의 8.3%를 차지하는 데 부담하는 세액은 41.2%이다	상호출자제한기업 〉 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.

3.8 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 Red 와 Black에 차이가 없어야 할텐데 집계결과는 어떻게 나오고 있나요?

분명히 영향을 미치고 있는 것으로 보입니다.

통계적으로 매우 유의한 차이가 관찰되고 있습니다.

감세정책의 효과가 긍정적이라고 설명한 Red 에서는 낮춰야 한다는 응답이, 감세정책의 효과가 부정적이라고 설명한 Black 에서는 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 1인 당 명목상 GDP는 10만 1,509달러로 세계 2위 수준입니다. 우리나라의 2020년 국민부담률 이 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 일부 격차를 확대하고 불평등을 줄이는 데 도움이 되는 사회 복지 프로그램 및 기타 정책에 자금을 지원하 는 정부의 능력을 제한하였습니다. 그 결과 상위 10%의 가구가 전체 부의 54%를 점유할 정 도로 불평등이 심화되었습니다. 우리나라의 2020년 국민부담률이 27.9%인 점에 비추어 어 떤 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

3.8.1 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

표 3.21: 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 둘다 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.

3.8.2 % 비교

	낮춰야 한다	높여야 한다	모름/무응답	계
Red(감세의 긍정적효과 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.

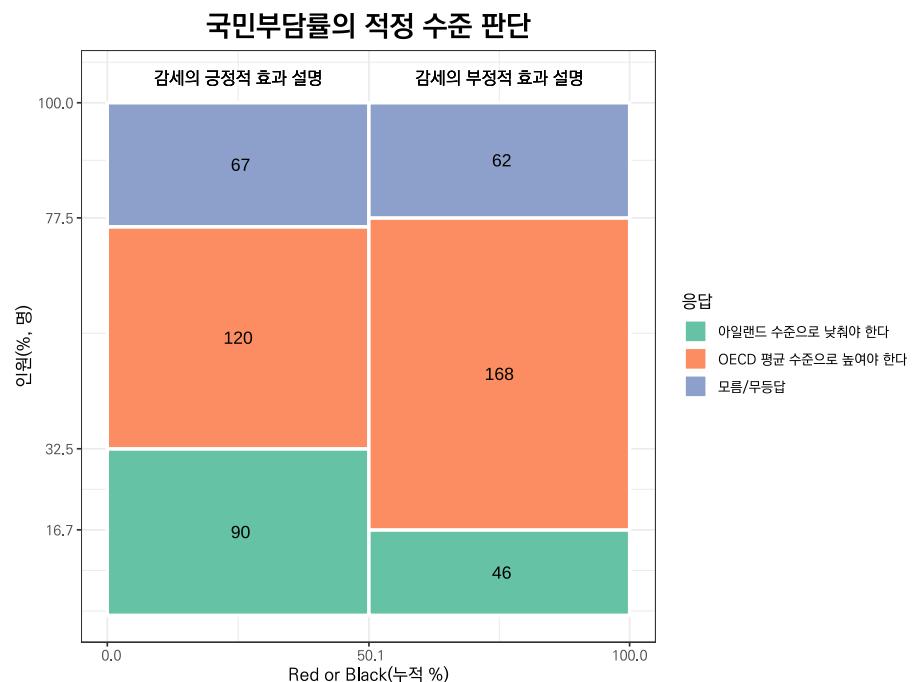
3.8.3 % 합계

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.

3.8.4 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.

3.9 마감 시간으로부터 제출 시간의 분포

3.9.1 분포표

표 3.24: 일 단위

	14일	13일	12일	11일	10일	9일	8일	7일	6일	5일	4일	3일	2일	1일	계
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.

3.9.2 날마다 고르게 제출하는가?

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

표 3.26: Chi-squared test for given probabilities: .

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

날마다 고르게 제출하는지 알아 보았습니다.

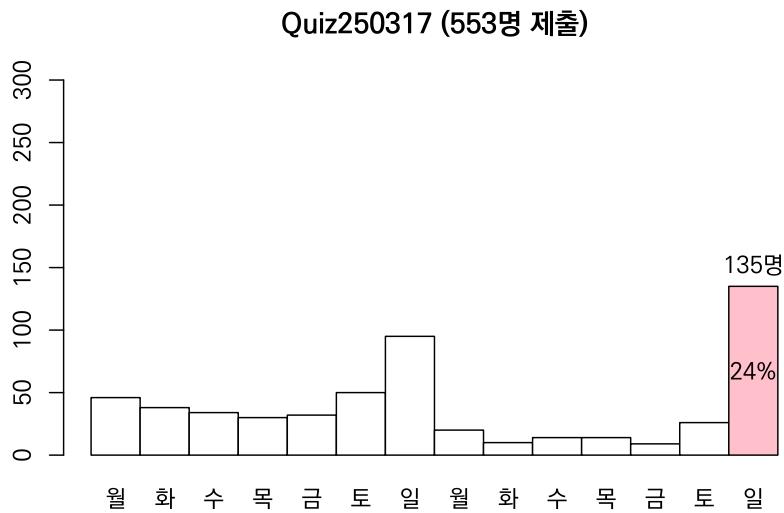
분포표의 “계”행에서 ’계’열을 제외하고 카이제곱테스트를 수행합니다.

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.

3.9.3 막대그래프



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

3.9.4 Red, Black 간에 닮았는가?

	14일	13일	12일	11일	10일	9일	8일	7일	6일	5일	4일	3일	2일	1일
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

표 3.28: Pearson's Chi-squared test: .

Test statistic	df	P value
16.98	13	0.2003

제출시간의 분포가 Red, Black 간에 닮았는지 알아 보았습니다.

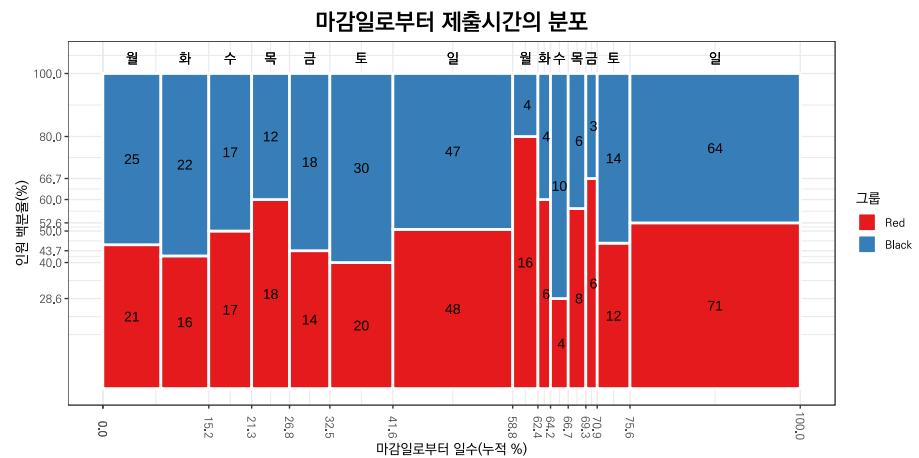
이번에는 분포표의 첫번째와 두번째 행, '계'열을 제외한 나머지 열에 대해서 카이제곱테스트를 수행합니다. 카이제곱 통계량은 16.978, 자유도는 13, p-value

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

이 사실을 Mosaic Plot을 이용하여 시각적으로 살펴보겠습니다.

답았다고 느껴지나요?

3.9.5 Mosaic Plot



제 4 장 4주차 데이터 실험 집계

4.1 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.

4.1.1 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 $1/2$ 을 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 40이고 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.

4.1.2 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.

4.2 Q1. 세종대왕 시대 조세제도

1. 세종대왕 즉위 당시의 조세제도에 대한 설명으로 적절치 않은 것은? *
- 과천법 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.

4.2.1 조선초기 조세제도

'조(租)'는 공전(公田)의 경작자가 국고에 과전법 체제에서 전국 토지를 세 등급으로 나누고 실제 수확량을 확인하여 징수하였다.		'세(稅)'는 사전(私田)의 경작자가 전주에게 바치는 지대를 지대를 뜻한다.		실제로 현장에 나가서 수확량을 파악하고 국가에 상납하는 지세를 뜻한다.		등급에 따라 일정한 비율로 세금을 감면해 주는 대동(大同)법을 실시하였다. 계
Red	22	28	26	22	180	278
Black	17	30	17	19	203	286
계	39	58	43	41	383	564

표 4.3: 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-value 는 0.3951이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

4.2.2 조선초기 조세제도(%)

과전법 체제에서 전국 토지를 세 등급으로 나누고 실제 수확량을 확인하여 징수하였다.	'조(租)'는 공전(公田)의 경작자가 국고에 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.

4.3 Q2. 공법도입에 대한 대신들의 찬성율

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

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

4.3.1 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

표 4.6: 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?

4.3.2 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.

4.4 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

4.4.1 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

표 4.9: 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 0.61300이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

4.4.2 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.

4.5 Q4. 공법

4. 공법의 기본세율은? *

- 10분의 1
- 15분의 1
- 20분의 1
- 30분의 1

4.5.1 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

표 4.12: 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 0.11480이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

4.5.2 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.

4.6 Q5. 1423년 조선시대 호구와 인구

5. 1422년 세종실록지리지에 실린 조선시대의 호수와 인구는? *

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

4.6.1 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

표 4.15: 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.

4.6.2 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.

4.7 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

4.7.1 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

표 4.18: 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 0.96700이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

4.7.2 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.

4.8 Q7. 부연설명의 효과 : 주당 근로 69시간제 도입 찬반

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

찬반여부에 대한 카이제곱테스트의 p-value를 높고 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> <ul style="list-style-type: none"> <input type="radio"/> 찬성한다. <input type="radio"/> 반대한다. <input type="radio"/> 모름/무응답
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> <ul style="list-style-type: none"> <input type="radio"/> 찬성한다. <input type="radio"/> 노동자가 과도한 연장근로를 받을 수 있어 반대한다. <input type="radio"/> 모름/무응답

4.8.1 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

표 4.21: 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.

4.8.2 % 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.

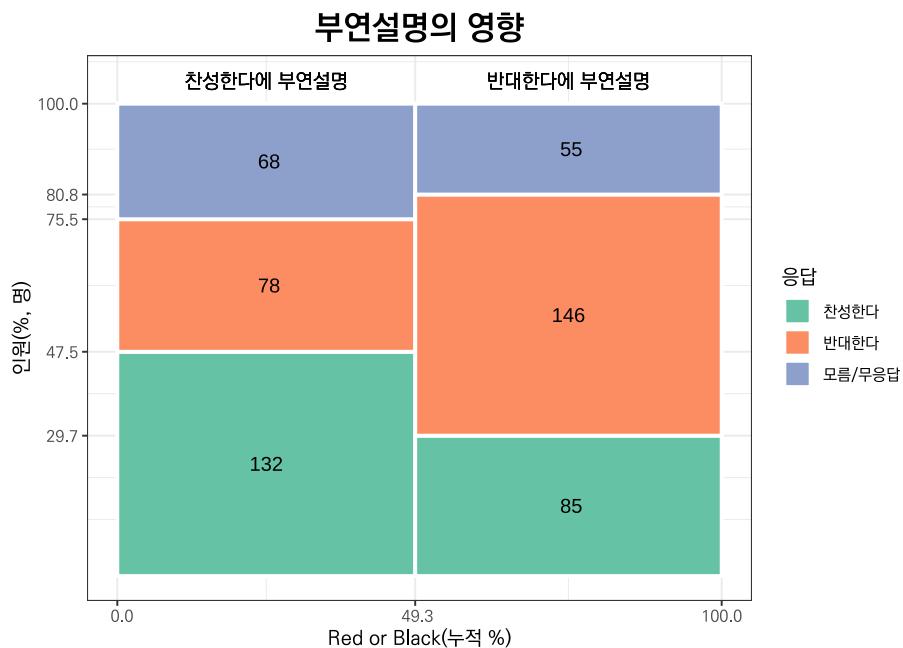
4.8.3 % 합계

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.

4.8.4 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.

4.9 마감 시간으로부터 제출 시간의 분포

4.9.1 분포표

표 4.24: 일 단위

	14일	13일	12일	11일	10일	9일	8일	7일	6일	5일	4일	3일	2일	1일	계
Red	71	15	13	4	10	3	6	46	20	22	14	20	16	18	278
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.

4.9.2 날마다 고르게 제출하는가?

14일	13일	12일	11일	10일	9일	8일	7일	6일	5일	4일	3일	2일	1일
142	29	20	11	21	7	12	96	40	36	26	44	33	47

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

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

날마다 고르게 제출하는지 알아 보았습니다.

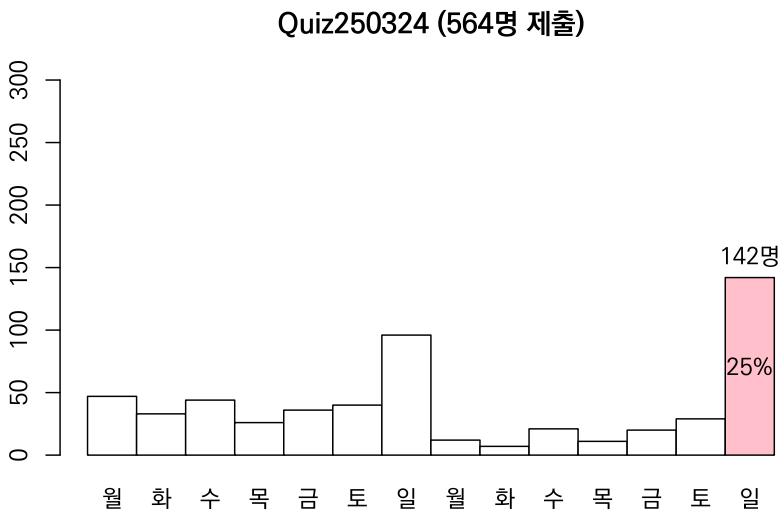
분포표의 “계”행에서 ‘계’열을 제외하고 카이제곱테스트를 수행합니다.

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.

4.9.3 막대그래프



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

4.9.4 Red, Black 간에 닮았는가?

	14일	13일	12일	11일	10일	9일	8일	7일	6일	5일	4일	3일	2일	1일
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

표 4.28: Pearson's Chi-squared test: .

Test statistic	df	P value
7.798	13	0.8565

제출시간의 분포가 Red, Black 간에 닮았는지 알아 보았습니다.

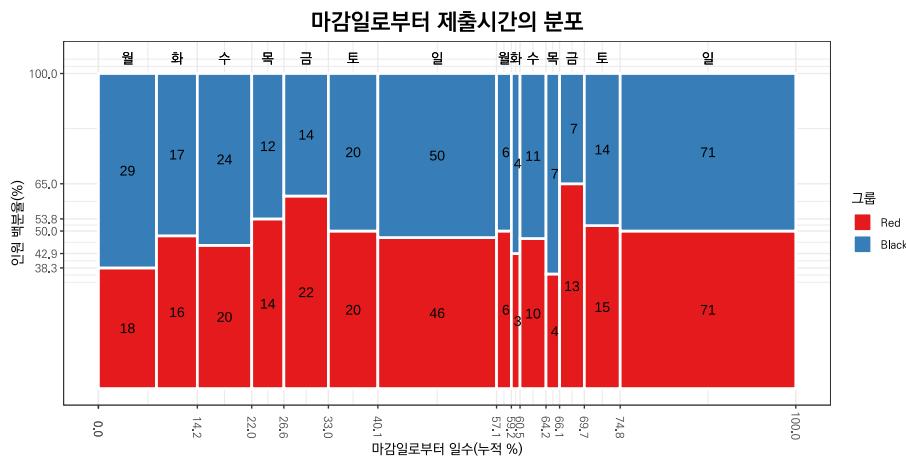
이번에는 분포표의 첫번째와 두번째 행, '계'열을 제외한 나머지 열에 대해서 카이제곱테스트를 수행합니다.

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?

4.9.5 Mosaic Plot



제 5 장 5주차 데이터 실험 집계

5.1 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.

5.1.1 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 $1/2$ 을 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.

5.1.2 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.

5.2 Q1. 한글의 문자 유형

1. 한글은 다음 중 어느 유형의 문자에 속하는가? *

- 민족 문자
- 엘리트 문자

5.2.1 한글은 민주 문자

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

표 5.3: 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 0.89350이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

5.2.2 한글은 민주 문자(%)

민주 문자	엘리트 문자	계
93.9	6.1	100.0

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

5.3 Q2. 정보혁명과 문자 체계

2. 다음 중 정보혁명을 이끄는데 유리한 문자체계는 어느 것인가? *

- 한자
- 한글

5.3.1 정보혁명을 이끄는 문자는 한글(집계표)

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

표 5.6: 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 0.300이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

5.3.2 정보혁명을 이끄는 문자는 한글(%)

한자	한글	계
7.0	93.0	100.0

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

5.4 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

5.4.1 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

표 5.9: 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?

5.4.2 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.

5.5 Q4. 해방직후 비문해율

4. 해방직후 우리나라의 비문해율은 대략 어느 수준이었는가? *

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

5.5.1 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

표 5.12: 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?

5.5.2 %

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.

5.6 Q5. 세대간 문해력 격차

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

- 대한민국
- 영국
- 이탈리아
- 미국

5.6.1 집계

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

표 5.15: 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 0.77320이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

5.6.2 %

대한민국	영국	이탈리아	미국	계
77.2	5.7	5.9	11.1	100.0

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

5.7 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

5.7.1 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

표 5.18: 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 달은 게 느껴집니까?

5.7.2 %

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.

5.8 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 보이기 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

5.8.1 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

표 5.21: 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 “반밖에” 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 “반밖에” 안 남았다는 반응을 보입니다.

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

랜덤화 효과일까요?

“모름/기타”의 응답이 유난히 깊은 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.

5.8.2 % 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.

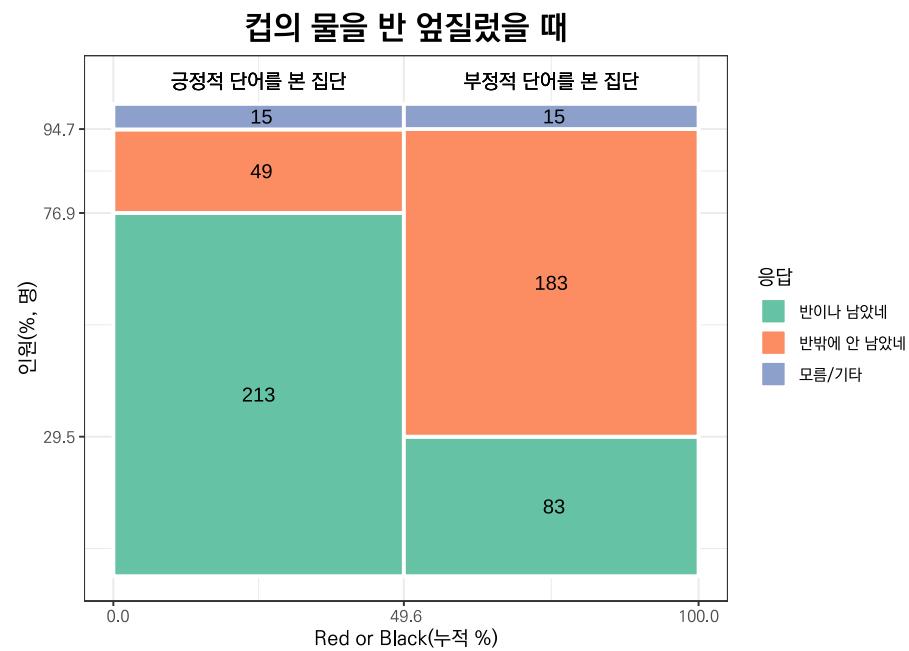
Red 와 Black 이 워낙 차이가 나지만 전체적으로 어느 정도가 “반이나” 남았다, 혹은 “반 밖에” 안 남았다고 응답하였는지 합쳐 보겠습니다.

5.8.3 % 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.

5.8.4 Mosaic Plot



Mosaic Plot 은 이 집계결과를 시각적으로 잘 보여줍니다. 긍정적 단어를 본 Red 에서 “반이나” 남아 있다고 응답한 백분율이 월등히 높고, 부정적 단어를 본 Black 에서 “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.

5.9 마감 시간으로부터 제출 시간의 분포

5.9.1 분포표

표 5.24: 일 단위

	14일	13일	12일	11일	10일	9일	8일	7일	6일	5일	4일	3일	2일	1일	계
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.

5.9.2 날마다 고르게 제출하는가?

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

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

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

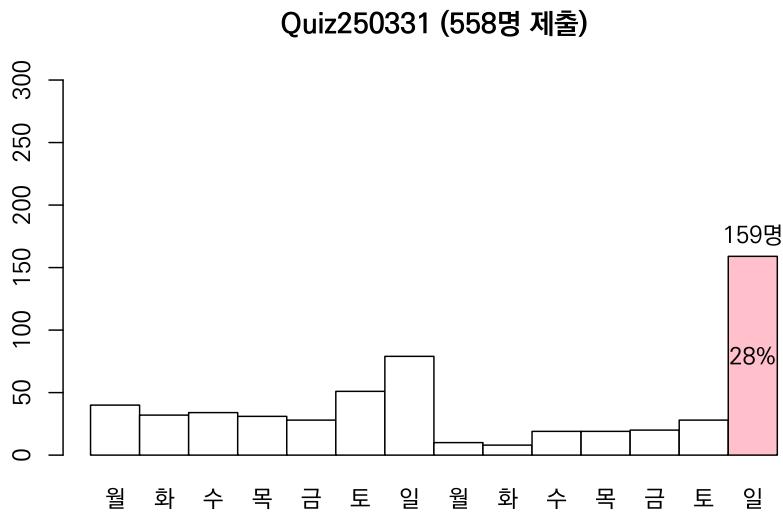
날마다 고르게 제출하는지 알아 보았습니다.

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.

5.9.3 막대그래프



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

5.9.4 Red, Black 간에 닮았는가?

	14일	13일	12일	11일	10일	9일	8일	7일	6일	5일	4일	3일	2일	1일
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

표 5.28: Pearson's Chi-squared test: .

Test statistic	df	P value
9.66	13	0.7215

제출시간의 분포가 Red, Black 간에 닮았는지 알아 보았습니다.

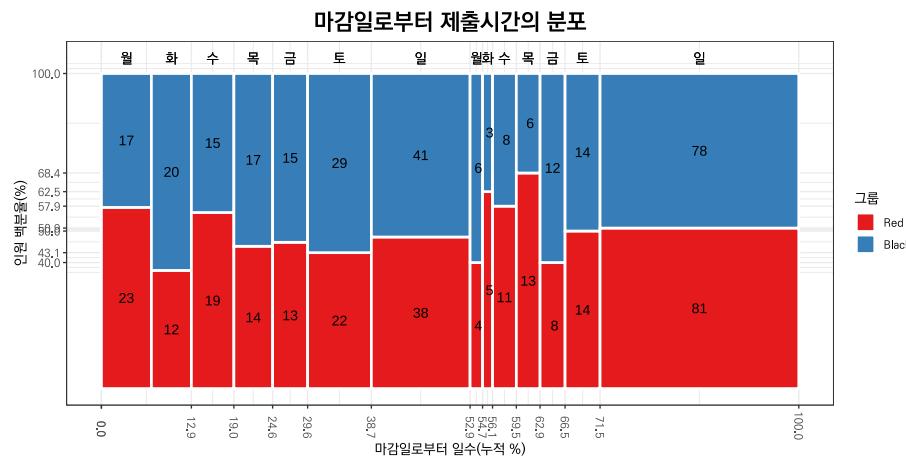
이번에는 분포표의 첫번째와 두번째 행, '계'열을 제외한 나머지 열에 대해서 카이제곱테스트를 수행합니다.

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?

5.9.5 Mosaic Plot



제 6 장 국민문해력조사 집계 결과

6.1 응답 집계

표 6.1: Counts

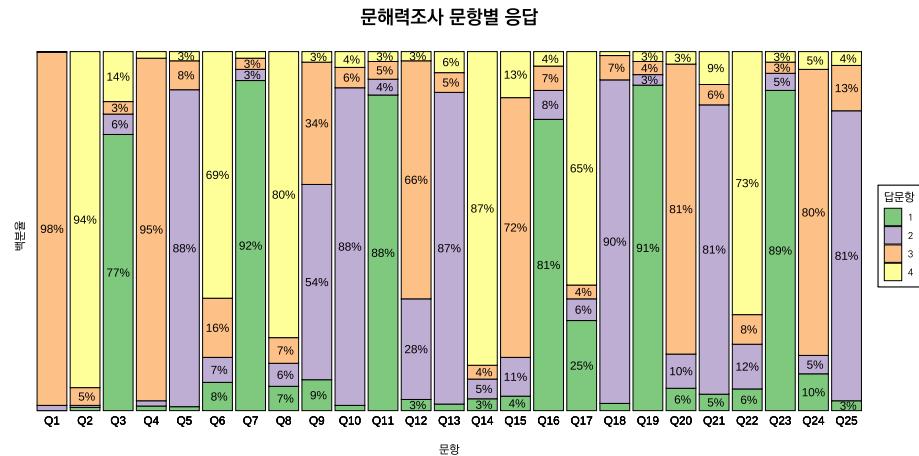
Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14	Q15	Q16	Q17	Q18	Q19	Q20	Q21	Q22	Q23	Q24	Q25		
0	5	42	0	7	6	43	50	27	47	8	48	0	7	10	18	22	44	31	37	1	49	54	25	33	48	
8	3	31	8	48	28	17	35	29	74	82	4	15	34	74	30	59	44	33	49	21	6	52	44	68	26	
53	27	19	52	44	90	17	39	18	63	1	27	36	20	21	39	57	21	37	20	44	31	45	17	43	59	
1	51	76	10	14	37	50	43	51	6	24	15	14	32	47	70	22	35	56	15	19	50	40	01	6	27	21

표 6.2: %

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
0	1	77	1	1	8	92	7	9	1	88	3	2	3	4	81	25	2	91	6	5	6	89	10	3
1	1	6	1	88	7	3	6	54	88	4	28	87	5	11	8	6	90	3	10	81	12	5	5	81
98	5	3	95	8	16	3	7	34	6	5	66	5	4	72	7	4	7	4	81	6	8	3	80	13
0	94	14	2	3	69	2	80	3	4	3	3	6	87	13	4	65	1	3	3	9	73	3	5	4

6.2 막대그래프

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



6.3 문해력 점수 계산

6.3.1 정답과 대조하여 R(Right)/W(Wrong) 표시

6.3.2 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, 88, 96, 96, 48, 84, 72, 100, 88, 88, 88, 92, 76, 96, 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, 84, 88, 88, 92, 88, 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, 36, 80, 92, 84, 68, 80, 76, 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, 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, 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, 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

6.4 Red and Black 비교

6.4.1 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

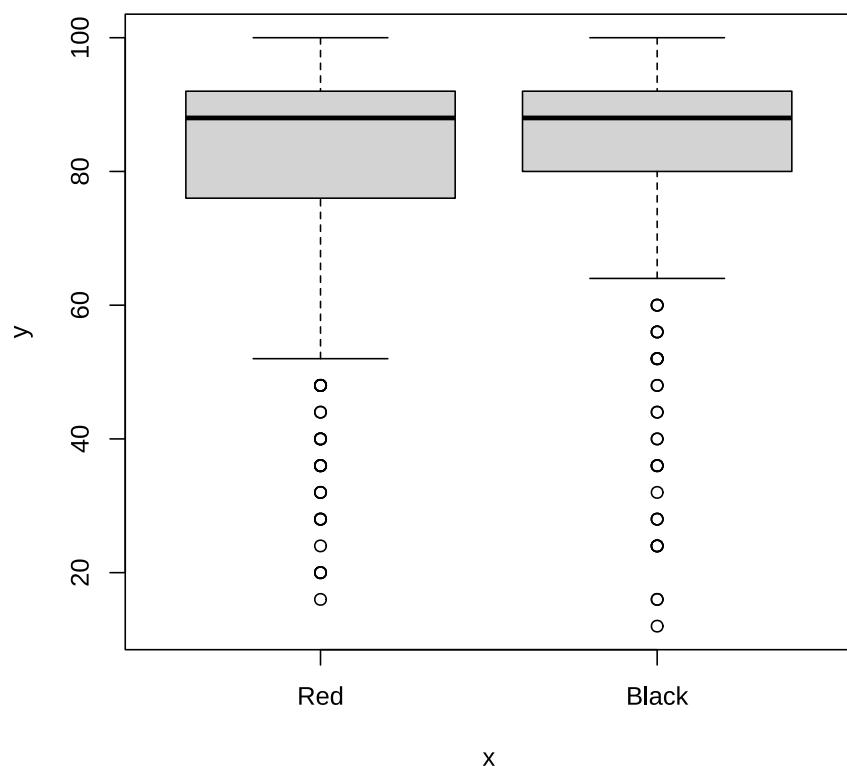
6.4.2 줄기-잎 그림

- Red

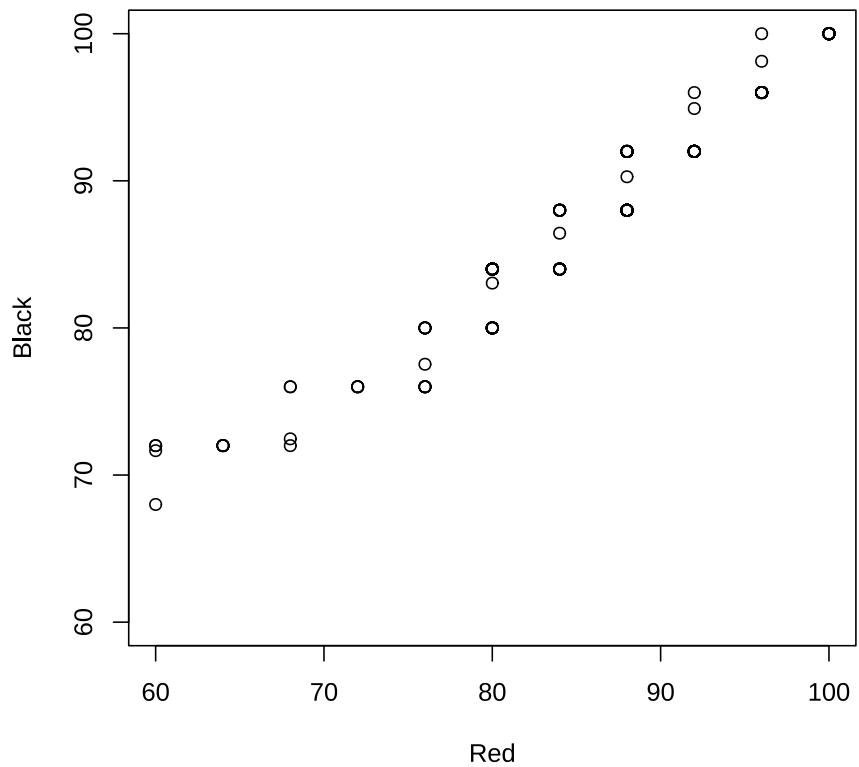
```
##  
##  The decimal point is 1 digit(s) to the right of the |  
##  
##      1 | 6
```

- Black

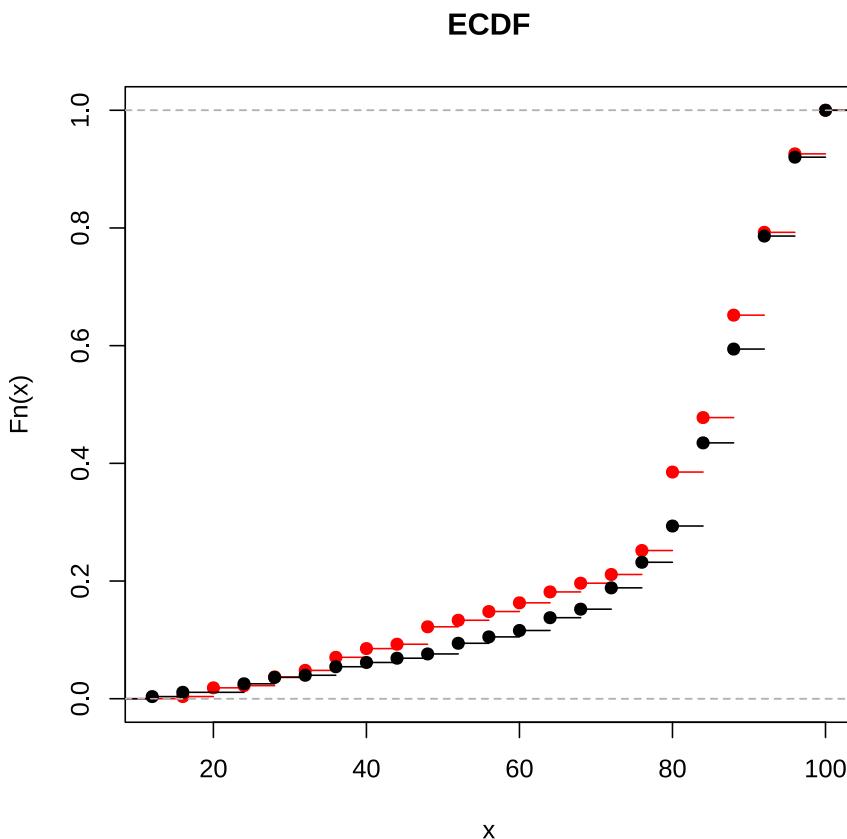
6.4.3 Box Plots



6.4.4 QQ plot



6.4.5 ECDF plot



6.4.6 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.

표 6.6: 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

6.5 문해력 등급 판정

6.5.1 분포표

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

표 6.7: 문해력 등급 분포

I	II	III	IV	계
13	41	55	437	546

표 6.8: 문해력 등급 분포(%)

I	II	III	IV	계
2.38	7.51	10.07	80.04	100

6.5.2 Red and Black

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

표 6.9: 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

표 6.10: Pearson's Chi-squared test: .

Test statistic	df	P value
5.301	3	0.151

6.6 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

6.7 어려운 문제?

6.7.1 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

6.7.2 정답률 70% 이하

	문6	문9	문12	문17
정답률	68.7	34.1	66.3	65

6.7.3 정답률 60% 이하

	문9
정답률	34.1

6.7.4 정답률 50% 이하

	문9
정답률	34.1

6.8 정답률이 낮은 문제들

6.8.1 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.

6.8.2 문9.

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

<우체국 택배 요금표>

무게 (kg)	2kg까지 (60cm까지)	5kg까지 (80cm까지)	8kg까지 (100cm까지)	10kg까지 (120cm까지)
동일자체	3,000원	4,000원	5,000원	6,000원
단지역	4,000원	5,000원	6,000원	7,000원
도착지역	5,000원	6,000원	7,000원	8,000원

☞ 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

6.8.3 문12.

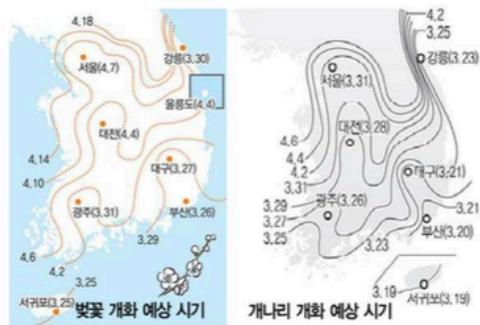
문 12) 다음은 2008년 4월 9일의 선거 때 배포된 '투표 확인증'에 대한 안내문입니다. 바르게 이해한 사람은 누구입니까?

투표 확인증 이용 안내	
사용 가능 시설	• 국립중앙박물관 및 소속 지방박물관(선거 당일만 이용 가능), 국립민속박물관 국립현대미술관, 국가지정문화재
유통 기간	2008년 4월 9일 ~ 4월 30일 (사설별 사용 기한이 다를 수 있으니 꼭 확인하시기 바랍니다.)
이용 방법	• '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.

6.8.4 문15.

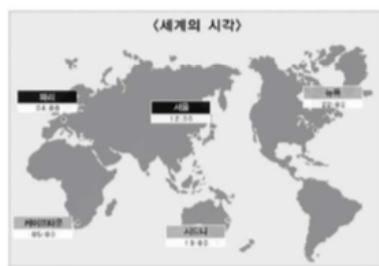
문 15) 다음 그림에 대한 설명으로 잘못된 것은 무엇입니까?



- 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.

6.8.5 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 |

6.8.6 문22.

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

실버 콘텐츠 활성화를 위한 교육 프로그램

⑥ 교육 개요

- 교육 대상: 노인일자리사업 담당자, 노인복지기관 관계자, 실버산업에 관심 있는 개인 및 단체
- 교육 기간: 2008년 7월 10일부터
- 교육 장소: 경기도청 3층 대회의실
- 교육 비용: 전과목 5만 원, 개별 과정 각 1만 원, 특강은 무료!!
- 교육 문의: 222-6206

⑦ 교육 내용

개최일자	강의주제	교사	내용
1 7월 10일(화) 15:00 ~ 17:00	문화 디자인 전략	김수희(문화부)	특강
2 7월 10일(화) 14:00 ~ 16:00	설로드(학) 시네 분석	서정수(영화언론)	
3 7월 10일(화) 16:00 ~ 18:00	문화콘텐츠의 대상·제작	윤정원(책제작)	
4 7월 10일(화) 16:00 ~ 18:00	노인의 눈으로 보는 영화감상	김경호(영화감독)	특강
5 7월 10일(화) 14:00 ~ 16:00	시민 미디어 활용	임경	
6 7월 10일(화) 14:00 ~ 16:00	제작 시제 및 방법		
7 7월 10일(화) 14:00 ~ 16:00	제작 일정 및 고려사항		

※ 개별 강연 선택 가능

① 이 프로그램은 노인들만을 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.

제 7 장 6주차 데이터 실험 집계

7.1 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.

7.1.1 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
2025-04-20 15:18:14	20246211	Red	Black
2025-04-20 20:50:23	20246250	Red	Black

	Red(구글예습퀴즈)	Black(구글예습퀴즈)
Red(랜덤화출석부)	273	4
Black(랜덤화출석부)	0	281
계	273	285

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

Red를 Black 이라고 한 사람이 4명, Black 을 Red 라고 한 사람이 0명입니다.

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, 4명이나 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 4명을 랜덤하게 둘로 나눌 때, 실제로 관찰된 4명 이상이나 0명이하로 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 0이하이거나 4이상을 관찰할 확률은 0.125입니다.

공평한 동전 던지기에서 앞면이 0개 이하 나오는 확률은 4개 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 273명, Black 에 응답한 인원은 285명입니다.

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

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

7.2 Q1. 월간 독서율

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

- 월간독서율
- 월간독서량
- 월간도서구입율
- 월간 도서구입량

7.2.1 집계

	월간독서율	월간독서량	월간도서구입율	월간 도서구입량	계
Red	238	25	9	1	273
Black	242	38	6	0	286
계	480	63	15	1	559

표 7.4: Pearson's Chi-squared test: .

Test statistic	df	P value
4.016	3	0.2598

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

그 결과 카이제곱 통계량은 4.02, 자유도는 3, p-value 는 0.2598이므로 Red, Black 간에 통계적으로 유의한 차이를 보이고 있습니다.

어디가 문제인가요?

7.2.2 %

월간독서율	월간독서량	월간도서구입율	월간 도서구입량	계
85.87	11.27	2.68	0.18	100.00

정답률은 Red, Black 을 합하여 계산하는데, 85.9(%) 입니다.

7.3 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

7.3.1 집계

	공평하게	소득 순으로	학력 순으로	연령 순으로	계
Red	237	19	10	7	273
Black	259	12	6	9	286
계	496	31	16	16	559

표 7.7: Pearson’s Chi-squared test: .

Test statistic	df	P value
3.506	3	0.32

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

그 결과 카이제곱 통계량은 3.51, 자유도는 3, p-value 는 0.32000이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

7.3.2 %

공평하게	소득 순으로	학력 순으로	연령 순으로	계
88.7	5.5	2.9	2.9	100.0

정답률은 Red, Black 을 합하여 계산하는데, 88.7(%) 입니다.

7.4 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 * 보았다.”에서 삭제하거나 교체해야 할 단어는?

- 책을
- 한달
- 평균
- 독서량
- 분포를

7.4.1 집계

	책을	한달	평균	독서량	분포를	계
Red	19	11	183	28	32	273
Black	13	13	171	37	52	286
계	32	24	354	65	84	559

표 7.10: Pearson's Chi-squared test: .

Test statistic	df	P value
7.408	4	0.1158

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

그 결과 카이제곱 통계량은 7.41, 자유도는 4, p-value 는 0.11580이므로 Red, Black 간에 통계적으로 유의한 차이를 보이고(지) 있(않)습니다.

실제로 닮은 게 느껴집니까?

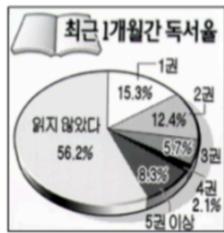
7.4.2 %

책을	한달	평균	독서량	분포를	계
5.7	4.3	63.3	11.6	15.0	100.0

정답률은 Red, Black 을 합하여 계산하는데, 63.3(%) 입니다.

7.5 Q4. 최근 1개월간 독서량

3. 다음 도표에서 잘못된 표현은? *



- 최근
- 1개월간
- 독서율
- 읽지 않았다
- 56.2%

7.5.1 집계

	최근	1개월간	독서율	읽지 않았다	56.2%	계
Red	24	15	199	20	15	273
Black	39	7	203	16	21	286
계	63	22	402	36	36	559

표 7.13: Pearson's Chi-squared test: .

Test statistic	df	P value
7.667	4	0.1046

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

그 결과 카이제곱 통계량은 7.67, 자유도는 4, p-value 는 0.1046이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

7.5.2 %

최근	1개월간	독서율	읽지 않았다	56.2%	계
11.3	3.9	71.9	6.4	6.4	100.0

정답률은 Red, Black 을 합하여 계산하는데, 71.9(%) 입니다.

7.6 Q5. 20대의 연간독서율

5. 2021년 우리나라 일반도서의 연간독서율은 어느 수준입니까?*

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

7.6.1 집계

	72.0%	76.7%	65.4%	52.1%	40.7%	계
Red	5	11	27	39	191	273
Black	7	17	24	45	193	286
계	12	28	51	84	384	559

표 7.16: Pearson's Chi-squared test: .

Test statistic	df	P value
1.933	4	0.748

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

그 결과 카이제곱 통계량은 1.93, 자유도는 4, p-value 는 0.74800이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

7.6.2 %

	72.0%	76.7%	65.4%	52.1%	40.7%	계
	2.1	5.0	9.1	15.0	68.7	100.0

정답률은 Red, Black 을 합하여 계산하는데, 68.7(%) 입니다.

7.7 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%

7.7.1 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	28	33	44	21	147	273
Black	31	36	54	16	149	286
계	59	69	98	37	296	559

표 7.19: Pearson's Chi-squared test: .

Test statistic	df	P value
1.691	4	0.7923

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

그 결과 카이제곱 통계량은 1.69, 자유도는 4, p-value 는 0.79230으로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

7.7.2 %

20~29세	20~29세	20~29세	20~29세	20~29세	계
31.3%, 60세	27.8%, 60세	34.7%, 60세	39%, 60세	50.5%, 60세	
이상 2.5%	이상 1%	이상 1.3%	이상 2%	이상 2.3%	
10.6	12.3	17.5	6.6	53.0	100.0

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

7.8 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? *

가. 내 연봉은 5천만원, 남들은 2천5백만원
 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

7.8.1 집계

	내가 남보다	남이 나보다	계
Red('내가 남보다' 먼저)	208	65	273
Black('남이 나보다' 먼저)	209	77	286
계	417	142	559

표 7.22: Pearson's Chi-squared test with Yates' continuity correction: .

Test statistic	df	P value
0.5597	1	0.4544

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

Black은 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 0.56, 자유도는 1, p-value 는 0.4544로 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, Black의 응답은 Q1~Q4에서와 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.

7.8.2 % 비교.

	내가 남보다	남이 나보다	계
Red(‘내가 남보다’ 먼저)	76.2	23.8	100.0
Black(‘남이 나보다’ 먼저)	73.1	26.9	100.0

이를 백분율로 살펴보면 “내가 남보다”를 앞 선택지로 한 Red에서 “내가 남보다”를 선택한 백분율, 76.2%(0)나 “남이 나보다”를 앞에 앞 선택지로 한 Black에서 “내가 남보다”를 선택한 백분율, 73.1%(0)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.

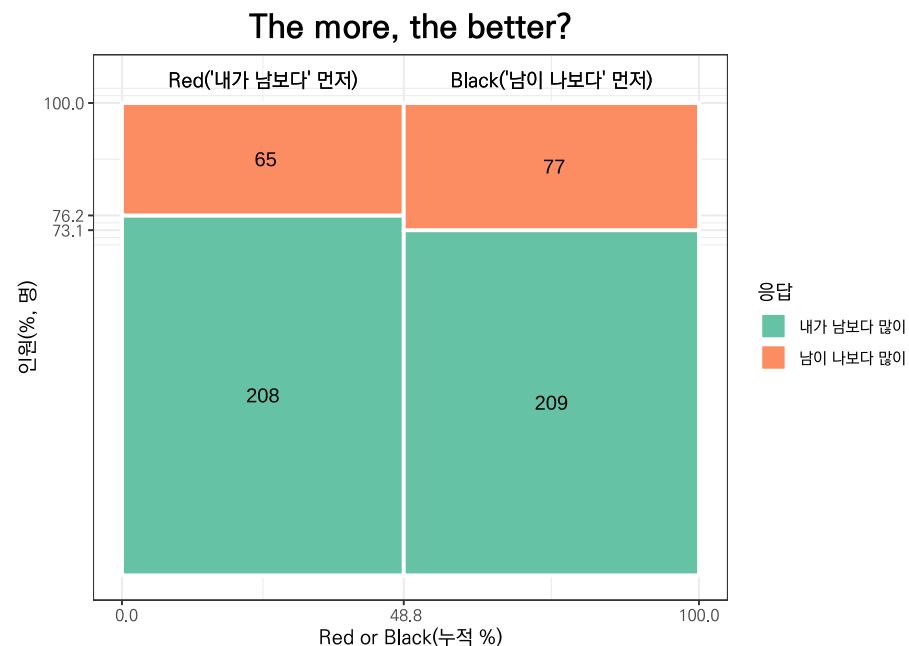
그 결과는 다음과 같습니다.

7.8.3 합산(%)

내가 남보다	남이 나보다	계
74.6	25.4	100.0

“내가 남보다” 많이 받는 세상을 원한다고 응답한 백분율은 선택지의 순서와는 무관하게 74.6%(으)로 매우 높습니다.

7.8.4 Mosaic Plot



Mosaic Plot 은 이 집계결과를 시각적으로 잘 보여줍니다.

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

7.9 마감 시간으로부터 제출 시간의 분포

7.9.1 분포표

표 7.25: 일 단위

	14일	13일	12일	11일	10일	9일	8일	7일	6일	5일	4일	3일	2일	1일	계
Red	77	11	10	6	4	6	9	34	20	16	10	26	18	26	273
Black	88	14	8	3	3	7	11	47	16	14	15	14	21	25	286
계	165	25	18	9	7	13	20	81	36	30	25	40	39	51	559

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?

14일	13일	12일	11일	10일	9일	8일	7일	6일	5일	4일	3일	2일	1일
165	25	18	9	7	13	20	81	36	30	25	40	39	51

표 7.27: Chi-squared test for given probabilities: .

Test statistic	df	P value
542.4	13	1.425e-107 * * *

날마다 고르게 제출하는지 알아 보았습니다.

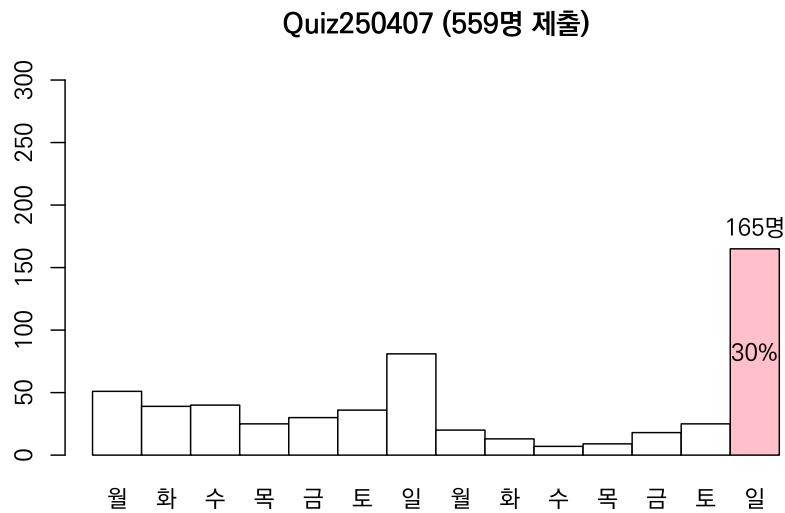
분포표의 “계”행에서 ‘계’열을 제외하고 카이제곱테스트를 수행합니다.

분포표 만으로도 쉽게 파악할 수 있지만 카이제곱테스트가 명확히 해 줍니다.

카이제곱 통계량은 542.39, 자유도는 13.00, p-value 는 1.4e-107 이므로 날짜별로 고르게 제출하고 있는지 살펴보기에는 천문학적으로 낮은 P-value 가 관찰되었습니다.

막대그래프로 살펴 보겠습니다.

7.9.3 막대그래프



7.9.4 Red, Black 간에 닮았는가?

	14일	13일	12일	11일	10일	9일	8일	7일	6일	5일	4일	3일	2일	1일
Red	77	11	10	6	4	6	9	34	20	16	10	26	18	26
Black	88	14	8	3	3	7	11	47	16	14	15	14	21	25

표 7.29: Pearson's Chi-squared test: .

Test statistic	df	P value
9.953	13	0.6978

제출시간의 분포가 Red, Black 간에 닮았는지 알아 보았습니다.

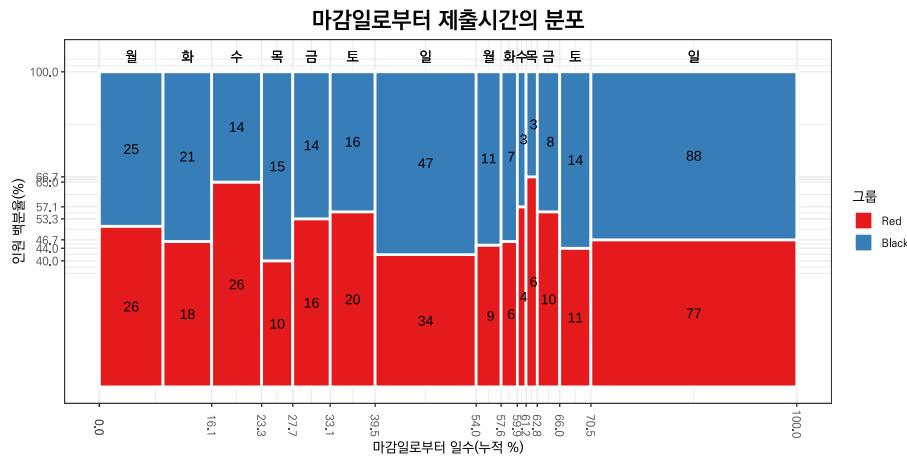
이번에는 분포표의 첫번째와 두번째 행, '계'열을 제외한 나머지 열에 대해서 카이제곱테스트를 수행합니다.

카이제곱 통계량은 9.95, 자유도는 13, p-value 는 0.70 이므로 제출 시간의 분포는 Red, Black 간에 통계적으로 유의한 차이가 관찰되고 있습니다.

이 사실을 Mosaic Plot 을 이용하여 시각적으로 살펴보겠습니다.

어디서 차이가 나고 있나요?

7.9.5 Mosaic Plot



제 8 장 7주차 데이터 실험 집계

8.1 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 쓸리는 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.

8.1.1 Red, Black을 잘못 표시한 사람들

	Red(구글예습퀴즈)	Black(구글예습퀴즈)
Red(랜덤화출석부)	265	2
Black(랜덤화출석부)	0	264
계	265	266

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 되는가입니다.

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

시행횟수가 20이고 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 265명, Black 에 응답한 인원은 266명입니다.

전체 응답인원 531 명을 랜덤하게 둘로 나눌 때 어느 한 쪽의 기대인원은 전체 응답인원의 절반인 265.5명이고, 표준오차는 전체 응답인원의 제곱근에 $1/2$ 을 곱해 준 11.5 명입니다.

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

8.2 Q1. 통계학의 기본원리

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

- 공평
- 무난
- 철저
- 균일

8.2.1 공평하게 추출하면 …

	공평	무난	철저	균일	계
Red	203	8	5	50	266
Black	211	9	6	40	266
계	414	17	11	90	532

표 8.3: Pearson's Chi-squared test: .

Test statistic	df	P value
1.415	3	0.7019

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

그 결과 카이제곱 통계량은 1.42, 자유도는 3, p-value 는 0.700이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

8.2.2 공평하게 추출하면 … (%)

공평	무난	철저	균일	계
77.8	3.2	2.1	16.9	100.0

정답률은 Red, Black 을 합하여 계산하는데, 77.8(%) 입니다.

8.3 Q2. 리터러리 디제스트의 실패

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

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

8.3.1 Selection Bias

	Selection Bias	Response Bias	Non-response Bias	Panel Bias	계
Red	191	33	38	4	266
Black	189	23	40	14	266
계	380	56	78	18	532

표 8.6: Pearson's Chi-squared test: .

Test statistic	df	P value
7.403	3	0.0601

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

그 결과 카이제곱 통계량은 7.40, 자유도는 3, p-value 는 0.0601이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

8.3.2 Selection Bias (%)

Selection Bias	Response Bias	Non-response Bias	Panel Bias	계
71.4	10.5	14.7	3.4	100.0

정답률은 Red, Black 을 합하여 계산하는데, 71.4(%) 입니다.

8.4 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

8.4.1 할당법의 문제점

	다단계 집락 추출방법	할당법	단순 랜덤추출방법	RDD	계
Red	20	197	39	10	266
Black	29	184	45	8	266
계	49	381	84	18	532

표 8.9: Pearson's Chi-squared test: .

Test statistic	df	P value
2.747	3	0.4322

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

그 결과 카이제곱 통계량은 2.75, 자유도는 3, p-value 는 0.430이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

8.4.2 할당법의 문제점(%)

다단계 집락 추출방법	할당법	단순 랜덤추출방법	RDD	계
9.2	71.6	15.8	3.4	100.0

정답률은 Red, Black 을 합하여 계산하는데, 71.6(%) 입니다.

8.5 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.
-

8.5.1 확률적 표본추출방법 도입

	확률적 표본추출	할당법	유선전화번호부	RDD도입	계
Red	196	30	22	18	266
Black	206	20	16	24	266
계	402	50	38	42	532

표 8.12: Pearson's Chi-squared test: .

Test statistic	df	P value
4.053	3	0.2558

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

그 결과 카이제곱 통계량은 4.05, 자유도는 3, p-value 는 0.260이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

8.5.2 확률적 표본추출방법 도입 … (%)

확률적 표본추출	할당법	유선전화번호부	RDD도입	계
75.6	9.4	7.1	7.9	100.0

정답률은 Red, Black 을 합하여 계산하는데, 75.6(%) 입니다.

8.6 Q5. 표본오차를 반으로 줄이려면?

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

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

8.6.1 4배로 늘려야

	2배로	4배로	1/2로	1/4로	계
Red	46	188	20	12	266
Black	57	182	21	6	266
계	103	370	41	18	532

표 8.15: Pearson's Chi-squared test: .

Test statistic	df	P value
3.296	3	0.3481

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

그 결과 카이제곱 통계량은 3.30, 자유도는 3, p-value 는 0.350이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

8.6.2 4배로 늘려야 (%)

2배로	4배로	1/2로	1/4로	계
19.4	69.5	7.7	3.4	100.0

정답률은 Red, Black 을 합하여 계산하는데, 69.5(%) 입니다.

8.7 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

8.7.1 선거당일 투표하는 유권자 전체

	국민 전체	18세 이상 국민 전체	등록된 유권자 전체	선거 당일 투표하는 유권자 전체	계
Red	14	59	35	158	266
Black	15	62	37	152	266
계	29	121	72	310	532

표 8.18: Pearson's Chi-squared test: .

Test statistic	df	P value
0.2805	3	0.9636

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

그 결과 카이제곱 통계량은 0.28, 자유도는 3, p-value 는 0.960이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

8.7.2 선거당일 투표하는 유권자 전체(%)

	국민 전체	18세 이상 국민 전체	등록된 유권자 전체	선거 당일 투표하는 유권자 전체	계
	5.5	22.7	13.5	58.3	100.0

정답률은 Red, Black 을 합하여 계산하는데, 58.3(%) 입니다.

8.8 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 정상입니다.

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.

8.8.1 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 카드를 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 _____

8.8.2 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, 2는 집계를 힘들게 합니다. 대소문자, 띄어쓰기 꼭 좀 지켜주세요.)



내 답변 _____

8.9 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 훌수이면 다른 쪽 면은 자음이다”가 됩니다.

쫙수가 아니면 훌수이고, 모음이 아니면 자음이니까요.

따라서 훌수 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 묻는 데 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.

8.9.1 집계

표 8.20: Red에 추상적 질문, Black에 구체적 질문

	정답	오답	계
Red(추상적 질문)	100	166	266
Black(구체적 질문)	168	98	266
계	268	264	532

{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 266(명)이 응답하였고 100(명)이 정답인 {A, 3}를 올렸습니다.

구체적인 상황에 놓인 Black의 경우 총 266(명)이 응답하였고 168(명)이 정답인 {Beer, 17}을 올려서 구체적인 질문에 압도적으로 많은 정답이 나온 것을 알 수 있습니다.

이를 백분율로 비교해 보면

8.9.2 % 비교

	정답	오답	계
Red(추상적 질문)	37.6	62.4	100.0
Black(구체적 질문)	63.2	36.8	100.0

추상적인 질문으로 이루어진 Red에서는 37.6(%)가 정답을 올렸고, 구체적인 질문으로 이루어진 Black에서는 63.2(%)가 정답을 올려서 구체적인 질문에 압도적으로 많은 정답이 올라왔다는 것을 알 수 있습니다.

이 상황을 Mosaic Plot으로 살펴보겠습니다.

8.9.3 Mosaic Plot



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

8.10 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 정답률과 어떤 관계가 있는지 파악할 수 있지 않을까 합니다.

8.10.1 집계

표 8.22: Red에 구체적 질문, Black에 추상적 질문

	정답	오답	계
Red(구체적 질문)	151	115	266
Black(추상적 질문)	72	194	266
계	223	309	532

구체적인 질문을 배치한 Red의 경우 총 266(명)이 응답하였고 151(명)이 정답인 {Beer, 17}(을)를 올렸습니다.

추상적인 질문을 배치한 Black의 경우 총 266(명)이 응답하였고 72(명)이 정답인 {A,3}(을)를 올려서 구체적인 질문에 압도적으로 많은 정답이 나온 것을 알 수 있습니다.

이를 백분율로 비교해 보면

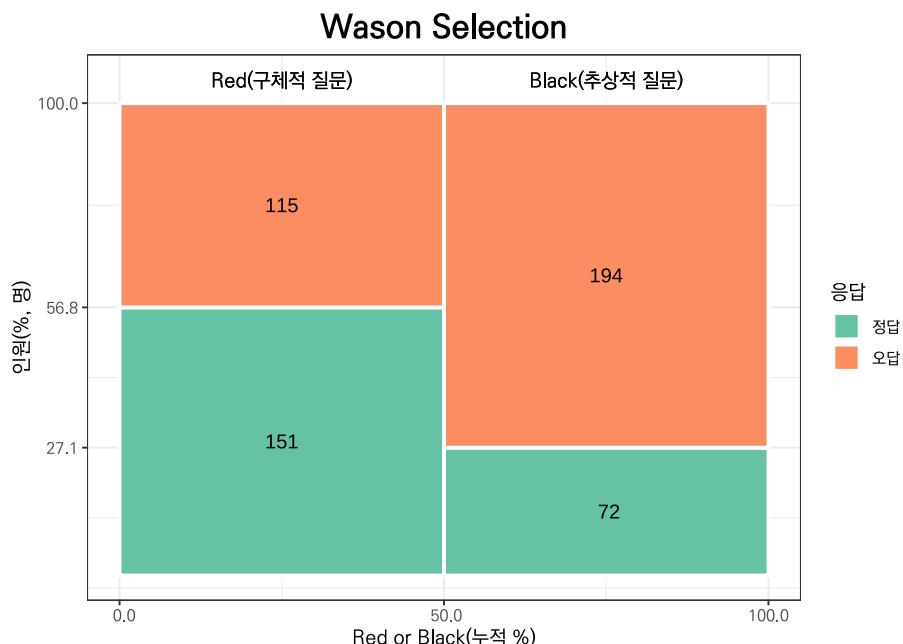
8.10.2 % 비교.

	정답	오답	계
Red(구체적 질문)	56.8	43.2	100.0
Black(추상적 질문)	27.1	72.9	100.0

구체적인 질문을 배치한 Red에서는 56.8(%)가 정답을 올렸고, 추상적인 질문을 배치한 Black에서는 27.1(%)가 정답을 올려서 구체적인 질문에 압도적으로 많은 정답이 올라왔다는 것을 알 수 있습니다.

이 상황을 Mosaic Plot으로 살펴보겠습니다.

8.10.3 Mosaic Plot



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

8.11 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.

8.11.1 집계

표 8.24: Wason Selection Task 인지편향 분석

	A,2	A,3	Other	계
Red(추상적 질문 먼저)	96	95	75	266
Black(구체적 질문 먼저)	119	61	86	266
계	215	156	161	532

{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 훌수 카드를 뒤집어야 합니다.

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

짜수가 아니면 훌수이니까 3을 뒤집어야 하는 것이죠.

추상적 질문이 먼저 Q7에 나온 Red의 경우 총 266(명)이 응답하였고 96(명)이 확증편향에서 비롯된 {A,2}를 올렸습니다.

정답인 {A,3}를 올린 95(명)보다 훨씬 많습니다.

추상적 질문이 Q8에 나온 Black의 경우 총 266(명)이 응답하였고 119(명)이 확증편향에서 비롯된 {A,2}를 올렸습니다.

정답인 {A,3}를 올린 61(명)보다 훨씬 많습니다.

합해서 215(명)이 확증편향에서 비롯된 {A,2}를 올렸고, 156(명)이 정답인 {A,3}를 올렸습니다.

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

백분율로 살펴 보겠습니다.

8.11.2 % 비교.

	A,2	A,3	Other	계
Red(추상적 질문 먼저)	36.1	35.7	28.2	100.0
Black(구체적 질문 먼저)	44.7	22.9	32.3	100.0
계	40.4	29.3	30.3	100.0

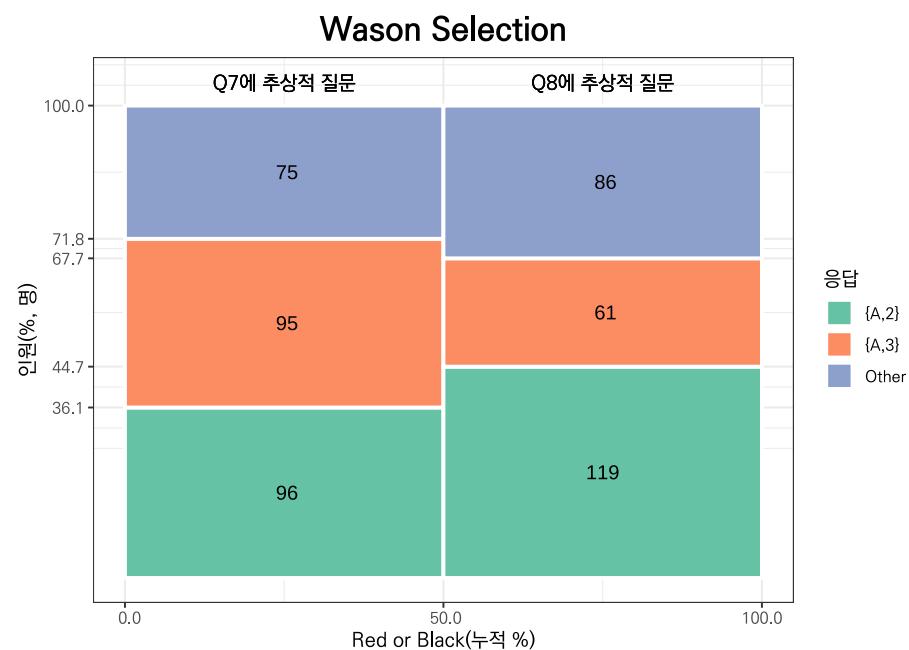
추상적인 질문이 먼저 Q7에 나온 Red에서는 36.1(%)가 확증편향에서 비롯된 응답 {A,2}를 올렸고, 35.7(%)가 정답인 {A,3}을 올렸는데, 추상적인 질문이 나중에 Q8에 나온 Black에서는 44.7(%)가 확증편향에서 비롯된 응답 {A,2}를 올렸고, 22.9(%)가 정답인 {A,3}을 올렸습니다.

합해서 보면 40.4(%)가 확증편향에서 비롯된 응답 {A,2}를 올렸고, 29.3(%)가 정답인 {A,3}을 올렸습니다.

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

이 상황을 Mosaic Plot으로 살펴보겠습니다.

8.11.3 Mosaic Plot



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

8.12 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?

8.12.1 집계표

표 8.26: Wason Selection

	추상적 질문 정답	구체적 질문 정답	계
Red(추상적 질문 먼저)	100	151	251
Black(구체적 질문 먼저)	72	168	240

표 8.27: Pearson's Chi-squared test with Yates' continuity correction: .

Test statistic	df	P value
4.797	1	0.02851 *

추상적 질문을 Q7에 배치하고 구체적 질문을 Q8에 배치한 Red의 경우 추상적 질문과 구체적 질문에 정답을 올린 사람은 총 251(명)이고 구체적 질문을 Q7에 배치하고 추상적 질문을 Q8에 배치한 Black의 경우 추상적 질문과 구체적 질문에 정답을 올린 사람은 총 240(명)으로 별로 차이가 나지 않습니다.

추상적 질문을 Q8에 배치한 Black의 경우 72(명)이 정답을 올려서 추상적 질문을 먼저 학습한 경우 정답을 더 많이 내었지만 통계적으로 유의한 차이는 아닌 것으로 나타나고 있습니다.

카이제곱 통계량은 4.80, p-value 는 0.029으로 통계적으로 유의한 차이를 관찰하지 못하였습니다.

따라서 학습 순서는 추상적 질문의 정답율에 영향을 미치지 못하고 있습니다.

백분율을 살펴 보겠습니다.

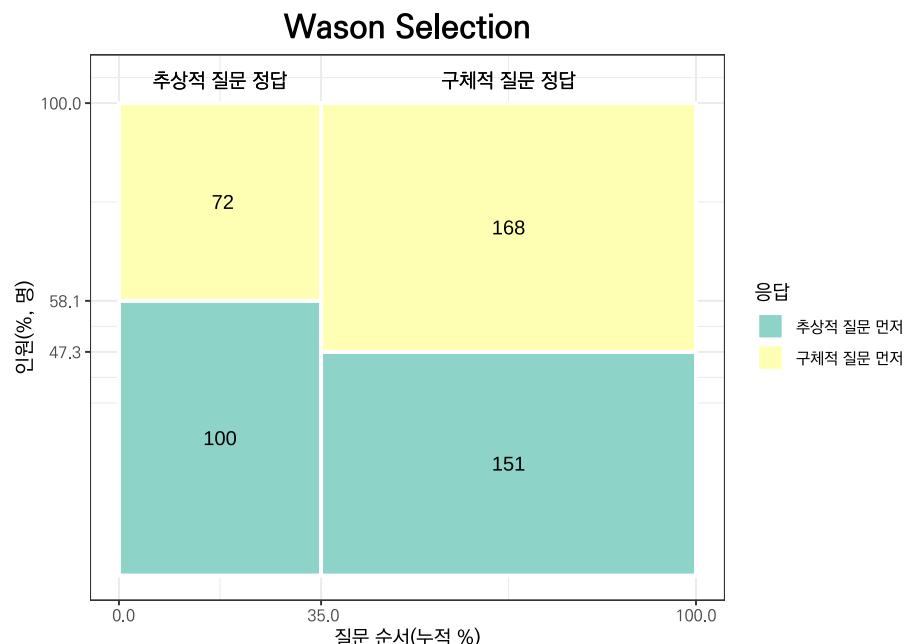
8.12.2 % 비교

표 8.28: Wason Selection

	추상적 질문 정답	구체적 질문 정답
Red(추상적 질문 먼저)	58.1	47.3
Black(구체적 질문 먼저)	41.9	52.7
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 58.1(%)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 정답률을 올린 인원이나 백분율이 비슷한다는 것을 시각적으로 파악할 수 있습니다.

8.12.3 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.

8.12.4 집계표

	정답	오답	계
추상적 문제	172	360	532
구체적 문제	319	213	532

추상적 질문에 답한 사람 총 532(명) 중에 정답을 올린 사람은 모두 172(명)이고 구체적 질문에 답한 사람 총 532(명) 중에 정답을 올린 사람은 모두 319(명)입니다. 백분율로 비교해 보면

8.12.5 % 비교

	정답	오답	계
추상적 문제	32.3	67.7	100.0
구체적 문제	60.0	40.0	100.0

추상적 질문에 정답을 올린 사람의 백분율은 32.3(%)이고 구체적 질문에 정답을 올린 사람의 백분율은 32.3(%)입니다.

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

8.12.6 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.

8.13 마감 시간으로부터 제출 시간의 분포

8.13.1 분포표

표 8.31: 일 단위

	14일	13일	12일	11일	10일	9일	8일	7일	6일	5일	4일	3일	2일	1일	계
Red	104	21	5	9	1	0	1	25	14	16	12	16	18	24	266
Black	98	12	7	1	3	3	7	35	13	14	11	14	16	32	266
계	202	33	12	10	4	3	8	60	27	30	23	30	34	56	532

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.

8.13.2 날마다 고르게 제출하는가?

14일	13일	12일	11일	10일	9일	8일	7일	6일	5일	4일	3일	2일	1일
202	33	12	10	4	3	8	60	27	30	23	30	34	56

표 8.33: Chi-squared test for given probabilities: .

Test statistic	df	P value
867.4	13	5.059e-177 * * *

날마다 고르게 제출하는지 알아 보았습니다.

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

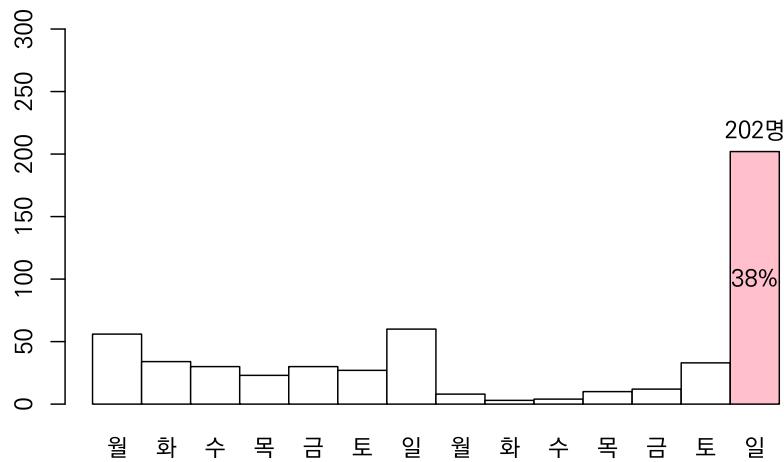
분포표 만으로도 쉽게 파악할 수 있지만 카이제곱테스트가 명확히 해 줍니다.

카이제곱 통계량은 867.37, 자유도는 13.00, p-value 는 5.1e-177 이므로 아직까지는 날짜별로 고르게 제출하고 있습니다.

막대그래프로 살펴 보겠습니다.

8.13.3 막대그래프

Quiz250414 (532명 제출)



8.13.4 Red, Black 간에 닮았는가?

	14일	13일	12일	11일	10일	9일	8일	7일	6일	5일	4일	3일	2일	1일
Red	104	21	5	9	1	0	1	25	14	16	12	16	18	24
Black	98	12	7	1	3	3	7	35	13	14	11	14	16	32

표 8.35: Pearson's Chi-squared test: .

Test statistic	df	P value
21.14	13	0.0702

제출시간의 분포가 Red, Black 간에 닮았는지 알아 보았습니다.

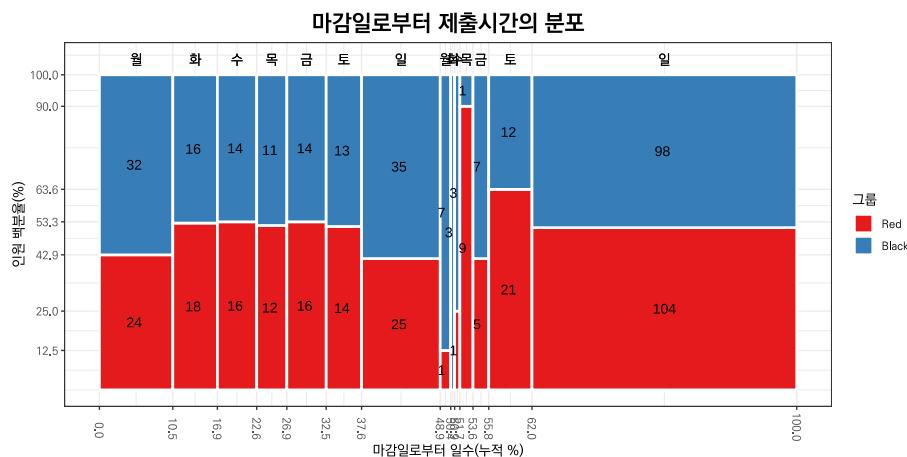
이번에는 분포표의 첫번째와 두번째 행, '계'열을 제외한 나머지 열에 대해서 카이제곱테스트를 수행합니다.

카이제곱 통계량은 21.14, 자유도는 13, p-value 는 0.0702 이므로 제출 시간의 분포는 Red, Black 간에 통계적으로 유의한 차이가 관찰되지 않습니다.

이 사실을 Mosaic Plot 을 이용하여 시각적으로 살펴보겠습니다.

답았다고 느껴지나요?

8.13.5 Mosaic Plot



제 9 장 8주차 데이터 실험 집계

9.1 실험의 목적

8주차 구글 예습 설문지 집계결과를 분석합니다.

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

Q4~Q6에서는 인지반응테스트를 수행하면서 1번효과에 대해서 알아봅니다.

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

9.1.1 Red, Black을 잘못 표시한 사람들

	Red(구글예습퀴즈)	Black(구글예습퀴즈)
Red(랜덤화출석부)	363	3
Black(랜덤화출석부)	3	370
계	366	373

랜덤화출석부에 있는 Red, Black 과 실제 구글설문에 올린 Red, Black 이 다른 사람들의 수효는 6명입니다.

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

두 가지 방법으로 분석합니다.

우선 Red, Black 을 잘못 선택한 6명을 랜덤하게 둘로 나누면 어느 한 쪽 집단에 들어갈 기대인원은 6명을 둘로 나눈 3(명)이고, 표준오차는 6의 제곱근에 $1/2$ 을 곱해 준 1.2명이 됩니다.

실제로 Red를 Black 이라고 한 사람수, 3명이나 Black 을 Red 라고 한 사람수, 3명은 기대인원으로부터 표준오차 범위에 아주 잘 들어갑니다.

두 번째 분석 방법은 확률을 계산해 보는 것입니다.

Red, Black 을 잘못 선택한 6명을 랜덤하게 둘로 나눌 때, 실제로 관찰된 3명 이상이나 3명이하로 잘못 선택한 사람수가 나올 가능성은 얼마나 되는가 입니다.

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

시행횟수가 6이고 한 번 시행에서 성공확률이 $1/2$ 인 이항분포에서 성공횟수가 3이하이거나 3이상을 관찰할 확률은 1입니다.

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

다만, 이번 실험과 같이 3명씩 동일한 결과가 나온 경우에는 중복되는 확률을 빼 주어야 합니다.

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 366명, Black 에 응답한 인원은 373명입니다.

전체 응답인원 739 명을 랜덤하게 둘로 나눌 때 어느 한 쪽의 기대인원은 전체 응답인원의 절반인 369.5명이고, 표준오차는 전체 응답인원의 제곱근에 $1/2$ 을 곱해 준 13.6 명입니다.

따라서 Red, Black 각 그룹에 관찰된 인원은 기대인원으로부터 표준오차 범위 안에 들어갑니다. 랜덤화출석체부에 있는 Red, Black 과 실제 구글설문에 올린 Red, Black 이 다른 사람들의 수효는 6명입니다. Red를 Black 이라고 한 사람이 3명, Black 을 Red 라고 한 사람이 3명입니다.

9.2 Q1. 랜덤화 플라시보 대조군 설계의 특징

1. Salk Vaccine Field Trial에서 Randomized Placebo Control Approach 의 특징 중 부모 동의 받은 어린이들을 치료군과 대조군으로 공평하게 나누기 위하여 도입된 방법은 무엇인가?

- 가) 랜덤화 나) 가짜약 대조군 다) 이중눈가림 라) 층화

9.2.1 집계

	랜덤화	가짜약 대조군	이중눈가림	총화	계
Red	227	90	42	7	366
Black	224	98	46	5	373
계	451	188	88	12	739

표 9.3: Pearson's Chi-squared test: .

Test statistic	df	P value
0.8093	3	0.8472

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

그 결과 카이제곱 통계량은 0.81, 자유도는 3, p-value 는 0.85이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

9.2.2 집계 (%)

랜덤화	가짜약 대조군	이중눈가림	총화	계
61.0	25.4	11.9	1.6	100.0

정답률은 Red, Black 을 합하여 계산하는데, 61.0(%) 입니다.

9.3 Q2. 백신의 효과 확인

2. Salk Vaccine Field Trail의 결과를 요약한 표로부터 어느 숫자들이 소아마비 예방주사의 효과를 명확히 보여주고 있는가?

임상시험결과 요약				
시행 대상	수행	케이스	스	반재료
전체	1,829,916	863	47	
플라시보 커트를 설정:	749,236	358	48	
백신 전종	200,745	57	28	
생리식염수 접종	201,229	142	71	
부모동의 안함	338,778	157	46	
불완전 접종	8,484	2	24	
NFIP설계:	1,080,680	505	47	
백신 전종	221,998	56	25	
대조군	725,173	391	54	
부모동의 안함	123,605	54	44	
불완전 접종	9,904	4	40	

9.3.1 집계

	28 vs 46	28 vs 71	28 vs 25	25 vs 54	계
Red	41	240	62	23	366
Black	26	266	64	17	373
계	67	506	126	40	739

표 9.6: Pearson's Chi-squared test: .

Test statistic	df	P value
5.56	3	0.1351

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

그 결과 카이제곱 통계량은 5.56, 자유도는 3, p-value 는 0.140이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닭은 게 느껴집니까?

9.3.2 집계 (%)

28 vs 46	28 vs 71	28 vs 25	25 vs 54	계
9.1	68.5	17.1	5.4	100.0

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

9.4 Q3. 3학년의 발병율

3. 3학년의 전반적인 소아마비 발병율을 잘 나타내고 있는 집단은 다음 중 어느 것인가? *

- NFIP 설계의 백신 접종 집단
- 플라시보 컨트롤 설계의 백신 접종 집단
- 플라시보 컨트롤 설계의 생리식염수 접종 집단
- NFIP 설계의 대조군 집단

9.4.1 집계

	NFIP 설계의 백신 접종 집단	플라시보 컨트롤 설계의 백신 접종 집단	플라시보 컨트롤 설계의 생리식염수 접종 집단	NFIP 설계의 대조군 집단	계
		Red	Black	Red	Black
Red	43	55	63	205	366
Black	59	60	62	192	373
계	102	115	125	397	739

표 9.9: Pearson's Chi-squared test: .

Test statistic	df	P value
3.095	3	0.3772

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

그 결과 카이제곱 통계량은 3.09, 자유도는 3, p-value 는 0.38이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

9.4.2 집계 (%)

NFIP 설계의 백신 접종 집단	플라시보 컨트롤 설계의 백신 접종 집단	플라시보 컨트롤 설계의 생리식염수 접종 집단	NFIP 설계의 대조군 집단	계
13.8	15.6	16.9	53.7	100.0

정답률은 Red, Black 을 합하여 계산하는데, 53.7(%) 입니다.

9.5 Cognitive Response Test

3초 안에 답하려면 틀리기 쉬운 문제들로 구성한 인지반응 테스트인데 정답률이 상당히 높습니다.

Red 는 가)에 끌리는 답, 나)에 생각이 필요한 답을 배치하고 Black 은 가)에 생각이 필요한 답, 나)에 끌리는 답을 배치하였습니다만 이로 인한 차이는 통계적으로 유의하지 않습니다.

Kahneman 의 Thinking, Fast and Slow 에 의하면 하버드, MIT, 프린스턴 대학 학생의 50%가 직관적인 오답을 말했고, 그 외 대학생들의 오답률은 80퍼센트가 넘었답니다.

스스로 자랑스러워 해도 됩니다.

3초 제한을 둔 이유는 깊이 생각하지 말고 직관적으로 풀어보라는 것이었는데 앞에 나온 댓글들을 보니까 시간을 들여 가면서 정답을 구하려 한 것 같습니다.

해외 명문대 학생들은 가벼운 마음으로 생각나는 대로 답하는 데 우리 수강생들은 3초 제한을 넘겨가면서 정답을 구하려 하는 걸까요?

Red

4. 야구 방망이와 공 세트가 1달러10센트다. 방망이는 공보다 1달러가 비싸다. 공은 얼마겠는 * 가? (3초 이내 응답)

가. 10센트
 나. 5센트

5. 기계 다섯 대가 5분 동안 가발 다섯 개를 만든다면, 기계 100대로 가발 100개를 만들려면 * 몇 분이 걸리겠는가? (3초 이내 응답)

가. 100분
 나. 5분

6. 호수에 수련 잎이 한 무더기 떠 있다. 수련 잎이 차지하는 면적은 날마다 두 배로 늘어난다. * 수련 잎이 호수 전체를 뒤덮는 데 48일이 걸린다면, 호수 절반을 뒤덮는 데는 며칠이 걸리겠는 가? (3초 이내 응답)

가. 24일
 나. 47일

Black

4. 야구 방망이와 공 세트가 1달러10센트다. 방망이는 공보다 1달러가 비싸다. 공은 얼마겠는 * 가? (3초 이내 응답)

가. 5센트
 나. 10센트

5. 기계 다섯 대가 5분 동안 가발 다섯 개를 만든다면, 기계 100대로 가발 100개를 만들려면 * 몇 분이 걸리겠는가? (3초 이내 응답)

가. 5분
 나. 100분

6. 호수에 수련 잎이 한 무더기 떠 있다. 수련 잎이 차지하는 면적은 날마다 두 배로 늘어난다. * 수련 잎이 호수 전체를 뒤덮는 데 48일이 걸린다면, 호수 절반을 뒤덮는 데는 며칠이 걸리겠는 가? (3초 이내 응답)

가. 47일
 나. 24일

9.6 Q4. 야구 방망이와 공

9.6.1 집계

	10센트	5센트	계
Red(10센트 먼저)	182	184	366
Black(5센트 먼저)	196	177	373
계	378	361	739

표 9.12: Pearson's Chi-squared test with Yates' continuity correction: .

Test statistic	df	P value
0.4806	1	0.4882

Q4는 야구 방망이와 공 세트가 1달러10센트이고 방망이는 공보다 1달러가 비싸다는 설명으로부터 얼핏 그 차이인 10센트가 바로 공의 가격이라고 착각하기 쉽습니다.

3초 안에 답을 해야 하기 때문에 방망이 + 공 = 1달러10센트, 방망이 - 공 = 1달러 를 계산할 시간이 없습니다. Red 와 Black 은 선택지의 순서만 바꾼 것인데 아무런 영향을 주지 않고 있습니다.

여기서 수행하고 있는 카이제곱 테스트는 이 순서를 바꾼 것의 효과가 통계적으로 유의한지 알아보기 위한 것입니다.

그 결과 카이제곱 통계량은 0.48, 자유도는 1, p-value 는 0.490이므로 선택지의 순서를 바꾼 것은 통계적으로 유의한 차이를 보이지 않습니다.

9.6.2 % 비교.

	10센트	5센트	계
Red(10센트 먼저)	49.7	50.3	100.0
Black(5센트 먼저)	52.5	47.5	100.0

이를 백분율로 살펴보면 10센트를 먼저 선택지에 올린 Red 에서 정답 5센트를 고른 백분율, 50.3%(0)이나 5센트를 먼저 선택지에 올린 Black 에서 정답 5센트를 고른 백분율, 47.5%(0)이나 큰 차이가 없어서 선택지를 바꾸어도 응답에는 아무런 영향을 주지 못한다는 것을 명확히 알 수 있습니다.

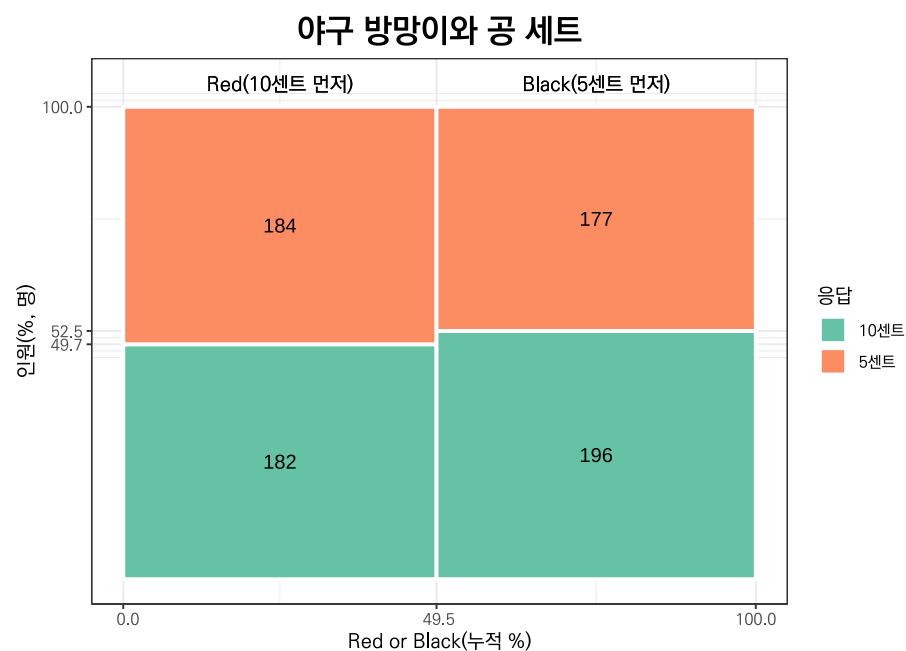
따라서 Red, Black 은 합쳐서 분석하는 것이 타당합니다. 그 결과는 다음과 같습니다.

9.6.3 합산(%)

10센트	5센트	계
51.2	48.8	100.0

정답 5센트를 고른 백분율은 선택지의 순서 바꿈과는 무관하게 48.8%(으)로 상당히 높은 편이라고 할 수 있습니다.

9.6.4 Mosaic Plot



Mosaic Plot 은 이 집계결과를 시각적으로 잘 보여줍니다.

선택지의 순서 바꿈에 거의 무관할 정도로 응답이 같았기 때문에 Red, Black 0이 닮았고, 정답인 5센트를 고른 비율도 닮은 것을 잘 알 수 있습니다.

9.7 Q5. 가발 만드는 기계

9.7.1 집계

	100분	5분	계
Red(100분 먼저)	117	249	366
Black(5분 먼저)	121	252	373
계	238	501	739

표 9.16: Pearson's Chi-squared test with Yates' continuity correction: .

Test statistic	df	P value
0.003446	1	0.9532

기계 다섯 대가 5분 동안 가발 다섯 개를 만들면 기계 한 대가 5분 동안 가발 한 개를 만드는 셈이므로 기계 100대를 동원하더라도 5분이면 가발 100 개를 만들 수 있습니다.

다만, 다섯 대, 5분, 다섯 개가 함께 나오니까 저도 모르게 즉각적으로 100대, 100분, 100개로 반응하게 됩니다.

Red 와 Black 은 선택지의 순서만 바꾼 것인데 아무런 영향을 주지 않고 있습니다.

여기서 수행하고 있는 카이제곱 테스트는 이 순서를 바꾼 것의 효과가 통계적으로 유의한지 알아보기 위한 것입니다.

그 결과 카이제곱 통계량은 0.0034, 자유도는 1, p-value 는 0.950이므로 선택지의 순서를 바꾼 것은 통계적으로 유의한 차이를 보이지 않습니다.

적어도 지금 자료로는 닮아도 이렇게 닮기가 힘들 정도입니다.

9.7.2 % 비교.

	100분	5분	계
Red(100분 먼저)	32.0	68.0	100.0
Black(5분 먼저)	32.4	67.6	100.0

이를 백분율로 살펴보면 100분을 먼저 선택지에 올린 Red에서 정답 5분을 고른 백분율, 68.0%(0)이나 5분을 먼저 선택지에 올린 Black 에서 정답 5분을 고른 백분율, 67.6%(0)이나 큰 차이가 없어서 선택지를 바꾸어도 응답에는 아무런 영향을 주지 못한다는 것을 명확히 알 수 있습니다.

따라서 Red, Black 은 합쳐서 분석하는 것이 타당합니다.

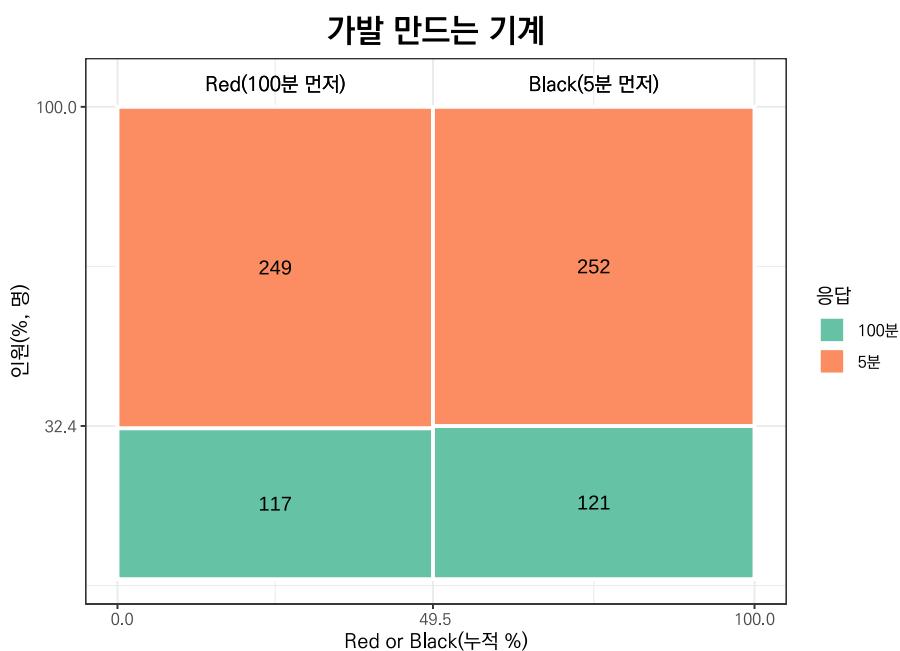
그 결과는 다음과 같습니다.

9.7.3 합산(%)

100분	5분	계
32.2	67.8	100.0

정답 5분을 고른 백분율은 선택지의 순서 바꿈과는 무관하게 67.8%(으)로 상당히 높은 편이라고 할 수 있습니다.

9.7.4 Mosaic Plot



Mosaic Plot 은 이 집계결과를 시각적으로 잘 보여줍니다.

선택지의 순서 바꿈에 거의 무관할 정도로 응답이 같았기 때문에 Red, Black 0이 닮았고, 정답인 5분을 고른 비율도 닮은 것을 잘 알 수 있습니다.

9.8 Q6. 호수에 수련

9.8.1 집계

	24일	47일	계
Red(24일 먼저)	183	183	366
Black(47일 먼저)	168	205	373
계	351	388	739

표 9.20: Pearson's Chi-squared test with Yates' continuity correction: .

Test statistic	df	P value
1.629	1	0.2019

수련 잎이 차지하는 면적이 날마다 두 배로 늘어나기 때문에 호수 전체를 뒤덮는데 48일이 걸린다면, 호수 절반을 뒤덮는 데는 그 절반인 24일로 답하기 쉽습니다.

그러나 절반이 뒤덮인 상태에서 하루면 전체를 뒤덮게 되므로 정답은 47일입니다.

Red 와 Black 은 선택지의 순서만 바꾼 것인데 아무런 영향을 주지 않고 있습니다.

여기서 수행하고 있는 카이제곱 테스트는 이 순서를 바꾼 것의 효과가 통계적으로 유의한지 알아보기 위한 것입니다.

그 결과 카이제곱 통계량은 1.63, 자유도는 1, p-value 는 0.200이므로 선택지의 순서를 바꾼 것은 통계적으로 유의한 차이를 보이지 않습니다.

9.8.2 % 비교.

	24일	47일	계
Red(24일 먼저)	50.0	50.0	100.0
Black(47일 먼저)	45.0	55.0	100.0

이를 백분율로 살펴보면 24일을 먼저 선택지에 올린 Red에서 정답 47일을 고른 백분율, 50.0(%)(이)나 47일을 먼저 선택지에 올린 Black 에서 정답 47일을 고른 백분율, 55.0(%)(이)나 큰 차이가 없어서 선택지를 바꾸어도 응답에는 아무런 영향을 주지 못한다는 것을 명확히 알 수 있습니다.

따라서 Red, Black 은 합쳐서 분석하는 것이 타당합니다.

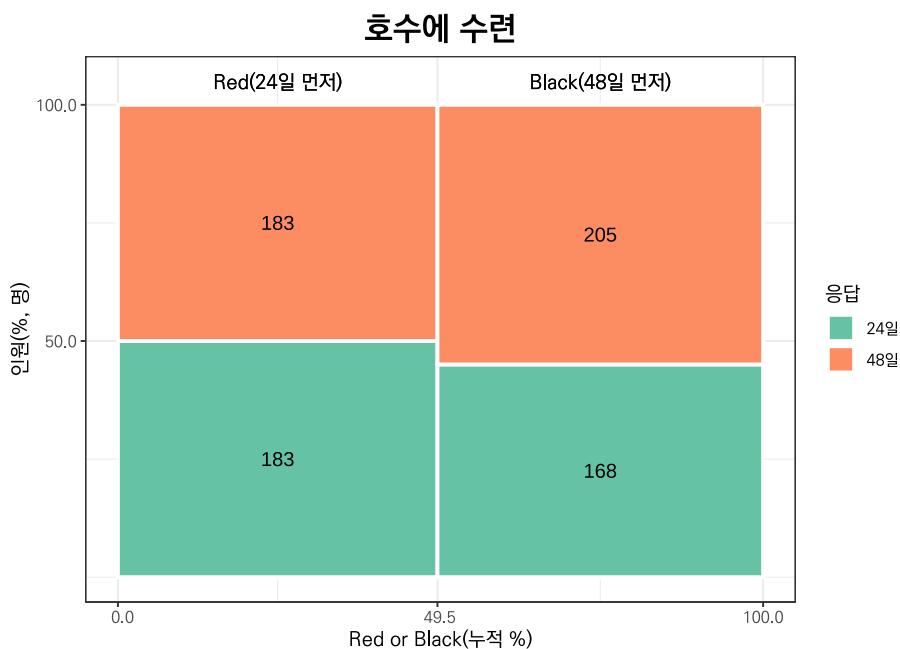
그 결과는 다음과 같습니다.

9.8.3 합산(%)

24일	47일	계
47.5	52.5	100.0

정답 47일을 고른 백분율은 선택지의 순서 바꿈과는 무관하게 52.5%(으)로 상당히 높은 편이라고 할 수 있습니다.

9.8.4 Mosaic Plot



Mosaic Plot 은 이 집계결과를 시각적으로 잘 보여줍니다.

선택지의 순서 바꿈에 거의 무관할 정도로 응답이 같았기 때문에 Red, Black 0이 닮았고, 정답인 47일을 고른 비율도 닮은 것을 잘 알 수 있습니다.

9.9 마감 시간으로부터 제출 시간의 분포

9.9.1 분포표

표 9.23: 일 단위

	14일	13일	12일	11일	10일	9일	8일	7일	6일	5일	4일	3일	2일	1일	계
Red	119	23	8	13	12	15	18	44	28	17	25	11	14	19	366
Black	121	23	18	8	8	6	13	52	37	18	18	14	14	22	372
계	240	46	26	21	20	21	31	96	65	35	43	25	28	41	738

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.

9.9.2 날마다 고르게 제출하는가?

14일	13일	12일	11일	10일	9일	8일	7일	6일	5일	4일	3일	2일	1일
240	46	26	21	20	21	31	96	65	35	43	25	28	41

표 9.25: Chi-squared test for given probabilities: .

Test statistic	df	P value
822.1	13	2.538e-167 ***

날마다 고르게 제출하는지 알아 보았습니다.

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

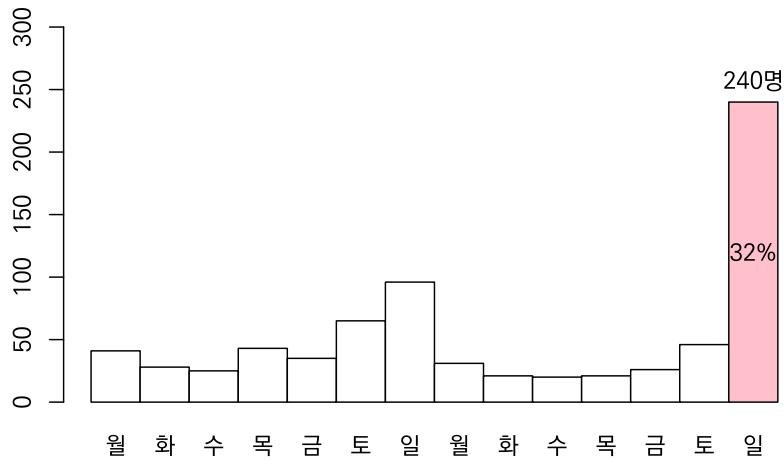
분포표 만으로도 쉽게 파악할 수 있지만 카이제곱테스트가 명확히 해 줍니다.

카이제곱 통계량은 822.11, 자유도는 13.00, p-value 는 2.5e-167 이므로 제출은 고르지 않고 특정 날짜에 치우쳐 있습니다.

막대그래프로 살펴 보겠습니다.

9.9.3 막대그래프

Quiz241021 (739명 제출)



9.9.4 Red, Black 간에 닮았는가?

	14일	13일	12일	11일	10일	9일	8일	7일	6일	5일	4일	3일	2일	1일
Red	119	23	8	13	12	15	18	44	28	17	25	11	14	19
Black	121	23	18	8	8	6	13	52	37	18	18	14	14	22

표 9.27: Pearson's Chi-squared test: .

Test statistic	df	P value
14.13	13	0.3648

제출시간의 분포가 Red, Black 간에 닮았는지 알아 보았습니다.

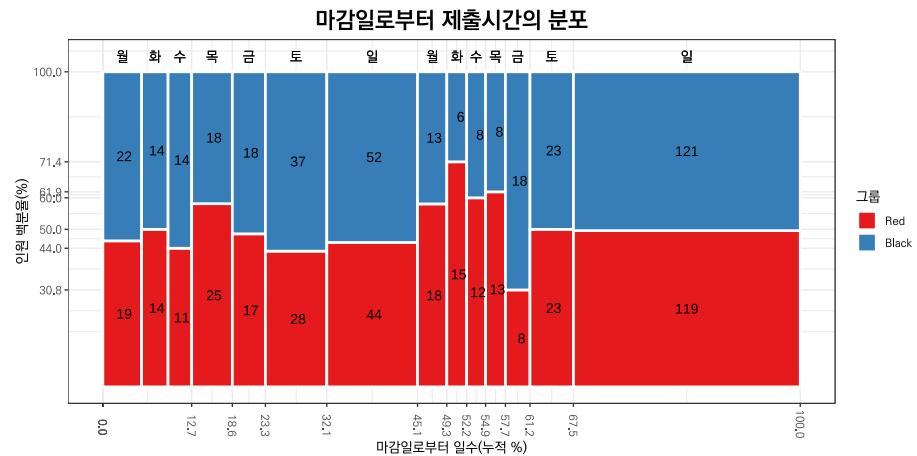
이번에는 분포표의 첫번째와 두번째 행, '계'열을 제외한 나머지 열에 대해서 카이제곱테스트를 수행합니다.

카이제곱 통계량은 14.13, 자유도는 13, p-value 는 0.36 이므로 제출 시간의 분포는 Red, Black 간에 통계적으로 유의한 차이가 관찰되지 않습니다.

이 사실을 Mosaic Plot 을 이용하여 시각적으로 살펴보겠습니다.

닮았다고 느껴지나요?

9.9.5 Mosaic Plot



제 10 장 8주차 데이터 실험 집계

10.1 실험의 목적

8주차 구글 예습 설문지 집계결과를 분석합니다.

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

Q4 ~ Q6에서는 인지반응테스트를 수행하면서 1번효과에 대해서 알아봅니다.

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

10.1.1 Red, Black을 잘못 표시한 사람들

	Red(구글예습퀴즈)	Black(구글예습퀴즈)
Red(랜덤화출석부)	288	1
Black(랜덤화출석부)	2	287
계	290	288

랜덤화출석부에 있는 Red, Black 과 실제 구글설문에 올린 Red, Black 이 다른 사람들의 수효는 3명입니다.

Red를 Black 이라고 한 사람이 1명, Black 을 Red 라고 한 사람이 2명입니다.

두 가지 방법으로 분석합니다.

우선 Red, Black 을 잘못 선택한 3명을 랜덤하게 둘로 나누면 어느 한 쪽 집단에 들어갈 기대인원은 3명을 둘로 나눈 1.5(명)이고, 표준오차는 3의 제곱근에 1/2을 곱해 준 0.9명이 됩니다.

실제로 Red를 Black 이라고 한 사람수, 1명이나 Black 을 Red 라고 한 사람수, 2명은 기대인원으로부터 표준오차 범위에 아주 잘 들어갑니다.

두 번째 분석 방법은 확률을 계산해 보는 것입니다.

Red, Black 을 잘못 선택한 3명을 랜덤하게 둘로 나눌 때, 실제로 관찰된 2명 이상이나 1명이하로 잘못 선택한 사람수가 나올 가능성은 얼마나 되는가 입니다.

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

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

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

다만, 이번 실험과 같이 3명씩 동일한 결과가 나온 경우에는 중복되는 확률을 빼 주어야 합니다.

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.

10.1.2 응답인원의 Red, Black

Red 로 응답한 인원은 290명, Black 에 응답한 인원은 288명입니다.

전체 응답인원 578 명을 랜덤하게 둘로 나눌 때 어느 한 쪽의 기대인원은 전체 응답인원의 절반인 289명이고, 표준오차는 전체 응답인원의 제곱근에 $1/2$ 을 곱해 준 12 명입니다.

따라서 Red, Black 각 그룹에 관찰된 인원은 기대인원으로부터 표준오차 범위 안에 들어갑니다. 랜덤화출석부에 있는 Red, Black 과 실제 구글설문에 올린 Red, Black 이 다른 사람들의 수효는 3명입니다. Red를 Black 이라고 한 사람이 1명, Black 을 Red 라고 한 사람이 2명입니다.

10.2 Q1. 랜덤화 플라시보 대조군 설계의 특징

1. Salk Vaccine Field Trial에서 Randomized Placebo Control Approach 의 특징 중 부모 동의 받은 어린이들을 치료군과 대조군으로 공평하게 나누기 위하여 도입된 방법은 무엇인가?

- 가) 랜덤화 나) 가짜약 대조군 다) 이중눈가림 라) 층화

10.2.1 집계

	랜덤화	가짜약 대조군	이중눈가림	총화	계
Red	191	66	23	10	290
Black	191	69	26	2	288
계	382	135	49	12	578

표 10.3: Pearson's Chi-squared test: .

Test statistic	df	P value
5.577	3	0.1341

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

그 결과 카이제곱 통계량은 5.58, 자유도는 3, p-value 는 0.130이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

10.2.2 집계 (%)

랜덤화	가짜약 대조군	이중눈가림	총화	계
66.1	23.4	8.5	2.1	100.0

정답률은 Red, Black 을 합하여 계산하는데, 66.1(%) 입니다.

10.3 Q2. 백신의 효과 확인

2. Salk Vaccine Field Trail의 결과를 요약한 표로부터 어느 숫자들이 소아
마비 예방주사의 효과를 명확히 보여주고 있는가?

임상시험결과 요약			
시행 대상	스증	케이스	스 반색률
74,411	1,870,016	863	47
플라시보 컨트롤 설계:	749,236	358	48
백신 전 총	200,745	57	28
생리식염수 접종	201,229	142	71
부모동의 안함	338,778	157	46
불완전 접종	8,484	2	24
NFIP설계:	1,080,680	505	47
백신 전 총	221,998	56	25
대조군	725,173	391	54
부모동의 안함	123,605	54	44
불완전 접종	9,904	4	40

- 가) 28 vs 46 나) 28 vs 71
 다) 28 vs 25 라) 25 vs 54

10.3.1 집계

	28 vs 46	28 vs 71	28 vs 25	25 vs 54	계
Red	19	204	49	18	290
Black	19	214	35	20	288
계	38	418	84	38	578

표 10.6: Pearson's Chi-squared test: .

Test statistic	df	P value
2.671	3	0.4452

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

그 결과 카이제곱 통계량은 2.67, 자유도는 3, p-value 는 0.450이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

10.3.2 집계 (%)

28 vs 46	28 vs 71	28 vs 25	25 vs 54	계
6.6	72.3	14.5	6.6	100.0

정답률은 Red, Black 을 합하여 계산하는데, 72.3(%) 입니다.

10.4 Q3. 3학년의 발병율

3. 3학년의 전반적인 소아마비 발병율을 잘 나타내고 있는 집단은 다음 중 어느 것인가? *

- NFIP 설계의 백신 접종 집단
- 플라시보 컨트롤 설계의 백신 접종 집단
- 플라시보 컨트롤 설계의 생리식염수 접종 집단
- NFIP 설계의 대조군 집단

10.4.1 집계

	NFIP 설계의 백신 접종 집단	플라시보 컨트롤 설계의 백신 접종 집단		플라시보 컨트롤 설계의 생리식염수 접종 집단		NFIP 설계의 대조군 집단	계
		접종 집단	생리식염수 접종 집단	접종 집단	생리식염수 접종 집단		
Red	25	43	47	175	175	290	
Black	29	47	43	169	169	288	
계	54	90	90	344	344	578	

표 10.9: Pearson's Chi-squared test: .

Test statistic	df	P value
0.7496	3	0.8615

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

그 결과 카이제곱 통계량은 0.75, 자유도는 3, p-value 는 0.86이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

10.4.2 집계 (%)

NFIP 설계의 백신 접종 집단	플라시보 컨트롤 설계의 백신 접종 집단	플라시보 컨트롤 설계의 생리식염수 접종 집단	NFIP 설계의 대조군 집단	계
9.3	15.6	15.6	59.5	100.0

정답률은 Red, Black 을 합하여 계산하는데, 59.5(%) 입니다.

10.5 Cognitive Response Test

3초 안에 답하려면 틀리기 쉬운 문제들로 구성한 인지반응 테스트인데 정답률이 상당히 높습니다.

Red 는 가)에 끌리는 답, 나)에 생각이 필요한 답을 배치하고 Black 은 가)에 생각이 필요한 답, 나)에 끌리는 답을 배치하였습니다만 이로 인한 차이는 통계적으로 유의하지 않습니다.

Kahneman 의 Thinking, Fast and Slow 에 의하면 하버드, MIT, 프린스턴 대학 학생의 50%가 직관적인 오답을 말했고, 그 외 대학생들의 오답률은 80퍼센트가 넘었답니다.

스스로 자랑스러워 해도 됩니다.

3초 제한을 둔 이유는 깊이 생각하지 말고 직관적으로 풀어보라는 것이었는데 앞에 나온 댓글들을 보니까 시간을 들여 가면서 정답을 구하려 한 것 같습니다.

해외 명문대 학생들은 가벼운 마음으로 생각나는 대로 답하는 데 우리 수강생들은 3초 제한을 넘겨가면서 정답을 구하려 하는 걸까요?

Red

4. 야구 방망이와 공 세트가 1달러10센트다. 방망이는 공보다 1달러가 비싸다. 공은 얼마겠는 * 가? (3초 이내 응답)

가. 10센트
 나. 5센트

5. 기계 다섯 대가 5분 동안 가발 다섯 개를 만든다면, 기계 100대로 가발 100개를 만들려면 * 몇 분이 걸리겠는가? (3초 이내 응답)

가. 100분
 나. 5분

6. 호수에 수련 잎이 한 무더기 떠 있다. 수련 잎이 차지하는 면적은 날마다 두 배로 늘어난다. * 수련 잎이 호수 전체를 뒤덮는 데 48일이 걸린다면, 호수 절반을 뒤덮는 데는 며칠이 걸리겠는 가? (3초 이내 응답)

가. 24일
 나. 47일

Black

4. 야구 방망이와 공 세트가 1달러10센트다. 방망이는 공보다 1달러가 비싸다. 공은 얼마겠는 * 가? (3초 이내 응답)

가. 5센트
 나. 10센트

5. 기계 다섯 대가 5분 동안 가발 다섯 개를 만든다면, 기계 100대로 가발 100개를 만들려면 * 몇 분이 걸리겠는가? (3초 이내 응답)

가. 5분
 나. 100분

6. 호수에 수련 잎이 한 무더기 떠 있다. 수련 잎이 차지하는 면적은 날마다 두 배로 늘어난다. * 수련 잎이 호수 전체를 뒤덮는 데 48일이 걸린다면, 호수 절반을 뒤덮는 데는 며칠이 걸리겠는 가? (3초 이내 응답)

가. 47일
 나. 24일

10.6 Q4. 야구 방망이와 공

10.6.1 집계

	10센트	5센트	계
Red(10센트 먼저)	147	143	290
Black(5센트 먼저)	124	164	288
계	271	307	578

표 10.12: Pearson's Chi-squared test with Yates' continuity correction: .

Test statistic	df	P value
3.082	1	0.07916

Q4는 야구 방망이와 공 세트가 1달러10센트이고 방망이는 공보다 1달러가 비싸다는 설명으로부터 얼핏 그 차이인 10센트가 바로 공의 가격이라고 착각하기 쉽습니다.

3초 안에 답을 해야 하기 때문에 방망이 + 공 = 1달러10센트, 방망이 - 공 = 1달러 를 계산할 시간이 없습니다.

Red 와 Black 은 선택지의 순서만 바꾸었는데 앞에 놓은 선택지를 고르는 1번효과가 나타나(지 않고) 통계적으로 유의한 차이를 보이지(고) 않(있)습니다.

여기서 수행하고 있는 카이제곱 테스트는 이 순서를 바꾼 것의 효과가 통계적으로 유의한지 알아보기 위한 것입니다.

그 결과 카이제곱 통계량은 3.08, 자유도는 1, p-value 는 0.07920이므로 선택지의 순서를 바꾼 것이 앞에 배치한 선택지를 고르는 데 있어서 통계적으로 유의한 차이를 보이지 않습니다.

10.6.2 % 비교.

	10센트	5센트	계
Red(10센트 먼저)	50.7	49.3	100.0
Black(5센트 먼저)	43.1	56.9	100.0

이를 백분율로 살펴보면 10센트를 먼저 선택지에 올린 Red 에서 정답 5센트를 고른 백분율, 49.3%(0)나 5센트를 먼저 선택지에 올린 Black 에서 정답 5센트를 고른

백분율, 56.9(%) 간에 차이가 관찰되었지만 선택지를 바꾸는 것이 응답에 큰 영향을 주지 않고 있다는 것을 알 수 있습니다.

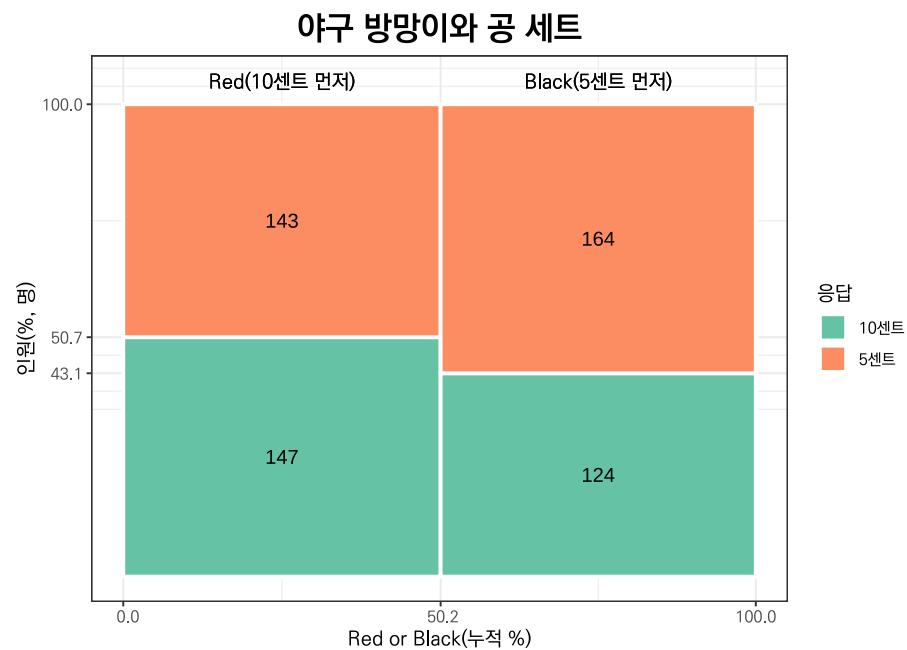
Red, Black 간에 통계적으로 유의한 차이가 관찰되지 않았기에 합쳐서 분석한 결과는 다음과 같습니다.

10.6.3 합산(%)

10센트	5센트	계
46.9	53.1	100.0

정답 5센트를 고른 백분율은 Red, Black 을 합쳤을 때 53.1%(으)로 제법 높은 편이라고 할 수 있습니다.

10.6.4 Mosaic Plot



Mosaic Plot 은 이 집계결과를 시각적으로 잘 보여줍니다.

선택지의 순서 바꿈에 거의 무관할 정도로 응답이 같았기 때문에 Red, Black 이 닮았고, 정답인 5센트를 고른 비율도 닮은 것을 잘 알 수 있습니다.

10.7 Q5. 가발 만드는 기계

10.7.1 집계

	100분	5분	계
Red(100분 먼저)	89	201	290
Black(5분 먼저)	72	216	288
계	161	417	578

표 10.16: Pearson's Chi-squared test with Yates' continuity correction: .

Test statistic	df	P value
2.053	1	0.1519

기계 다섯 대가 5분 동안 가발 다섯 개를 만들면 기계 한 대가 5분 동안 가발 한 개를 만드는 셈이므로 기계 100대를 동원하더라도 5분이면 가발 100 개를 만들 수 있습니다.

다만, 다섯 대, 5분, 다섯 개가 함께 나오니까 저도 모르게 즉각적으로 100대, 100분, 100개로 반응하게 됩니다.

Red 와 Black 은 선택지의 순서만 바꾼 것인데 응답에 영향을 미치는지 알아 봅니다.

여기서 수행하고 있는 카이제곱 테스트는 이 순서를 바꾼 것의 효과가 통계적으로 유의한지 알아보기 위한 것입니다.

그 결과 카이제곱 통계량은 2.05, 자유도는 1, p-value 는 0.150이므로 선택지의 순서를 바꾼 것은 통계적으로 유의한 차이를 보이지 않습니다.

10.7.2 % 비교.

	100분	5분	계
Red(100분 먼저)	30.7	69.3	100.0
Black(5분 먼저)	25.0	75.0	100.0

이를 백분율로 살펴보면 100분을 먼저 선택지에 올린 Red에서 정답 5분을 고른 백분율, 69.3%(이)나 5분을 먼저 선택지에 올린 Black 에서 정답 5분을 고른 백분율, 75.0%(이)나 큰 차이가 없어서 선택지를 바꾸는 것이 응답에 영향을 준다는 것을 명확히 알 수 있습니다.

Red, Black 을 합쳐서 분석합니다.

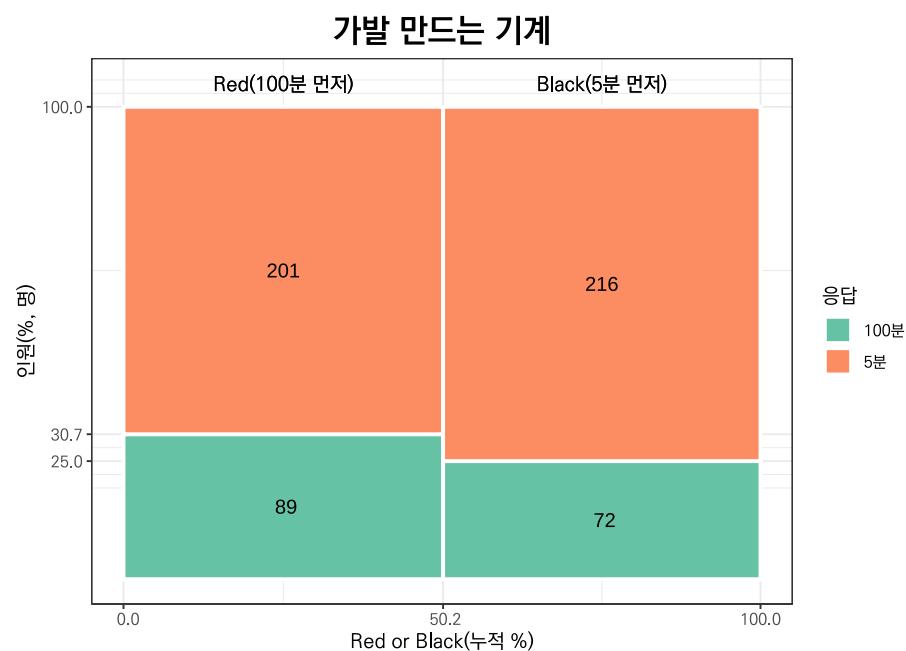
그 결과는 다음과 같습니다.

10.7.3 합산(%)

100분	5분	계
27.9	72.1	100.0

정답 5분을 고른 백분율은 선택지의 순서 바꿈과는 무관하게 72.1%(으)로 상당히 높은 편이라고 할 수 있습니다.

10.7.4 Mosaic Plot



Mosaic Plot 은 이 집계결과를 시각적으로 잘 보여줍니다.

선택지의 순서 바꿈에 거의 무관할 정도로 응답이 같았기 때문에 Red, Black 0이 닮았고, 정답인 5분을 고른 비율도 닮은 것을 잘 알 수 있습니다.

10.8 Q6. 호수에 수련

10.8.1 집계

	24일	47일	계
Red(24일 먼저)	133	157	290
Black(47일 먼저)	108	180	288
계	241	337	578

표 10.20: Pearson's Chi-squared test with Yates' continuity correction: .

Test statistic	df	P value
3.819	1	0.05066

수련 잎이 차지하는 면적이 날마다 두 배로 늘어나기 때문에 호수 전체를 뒤덮는데 48일이 걸린다면, 호수 절반을 뒤덮는 데는 그 절반인 24일로 답하기 쉽습니다.

그러나 절반이 뒤덮인 상태에서 하루면 전체를 뒤덮게 되므로 정답은 47일입니다.

Red 와 Black 은 선택지의 순서만 바꾼 것인데 아무런 영향을 주지 않고 있습니다.

여기서 수행하고 있는 카이제곱 테스트는 이 순서를 바꾼 것의 효과가 통계적으로 유의한지 알아보기 위한 것입니다.

그 결과 카이제곱 통계량은 3.82, 자유도는 1, p-value 는 0.0510으로 선택지의 순서를 바꾼 것은 통계적으로 유의한 차이를 보이지 않습니다.

10.8.2 % 비교.

	24일	47일	계
Red(24일 먼저)	45.9	54.1	100.0
Black(47일 먼저)	37.5	62.5	100.0

이를 백분율로 살펴보면 24일을 먼저 선택지에 올린 Red에서 정답 47일을 고른 백분율, 54.1(%)(이)나 47일을 먼저 선택지에 올린 Black 에서 정답 47일을 고른 백분율, 62.5(%)(이)나 큰 차이가 없어서 선택지를 바꾸어도 응답에는 아무런 영향을 주지 못한다는 것을 명확히 알 수 있습니다.

따라서 Red, Black 은 합쳐서 분석하는 것이 타당합니다.

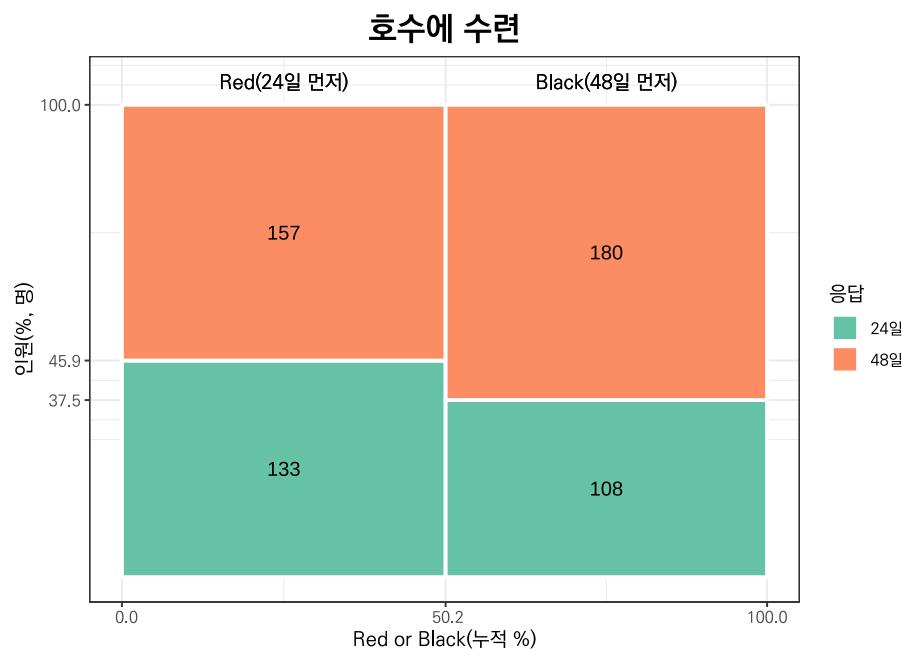
그 결과는 다음과 같습니다.

10.8.3 합산(%)

24일	47일	계
41.7	58.3	100.0

정답 47일을 고른 백분율은 선택지의 순서 바꿈과는 무관하게 58.3%(으)로 상당히 높은 편이라고 할 수 있습니다.

10.8.4 Mosaic Plot



Mosaic Plot 은 이 집계결과를 시각적으로 잘 보여줍니다.

선택지의 순서 바꿈에 거의 무관할 정도로 응답이 같았기 때문에 Red, Black 0이 닮았고, 정답인 47일을 고른 비율도 닮은 것을 잘 알 수 있습니다.

10.9 마감 시간으로부터 제출 시간의 분포

10.9.1 분포표

표 10.23: 일 단위

	14일	13일	12일	11일	10일	9일	8일	7일	6일	5일	4일	3일	2일	1일	계
Red	123	16	14	8	4	5	8	37	18	12	12	14	4	15	290
Black	127	18	11	5	3	5	7	40	18	7	7	12	8	20	288
계	250	34	25	13	7	10	15	77	36	19	19	26	12	35	578

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.

10.9.2 날마다 고르게 제출하는가?

14일	13일	12일	11일	10일	9일	8일	7일	6일	5일	4일	3일	2일	1일
250	34	25	13	7	10	15	77	36	19	19	26	12	35

표 10.25: Chi-squared test for given probabilities: .

Test statistic	df	P value
1234	13	7.921e-256 ***

날마다 고르게 제출하는지 알아 보았습니다.

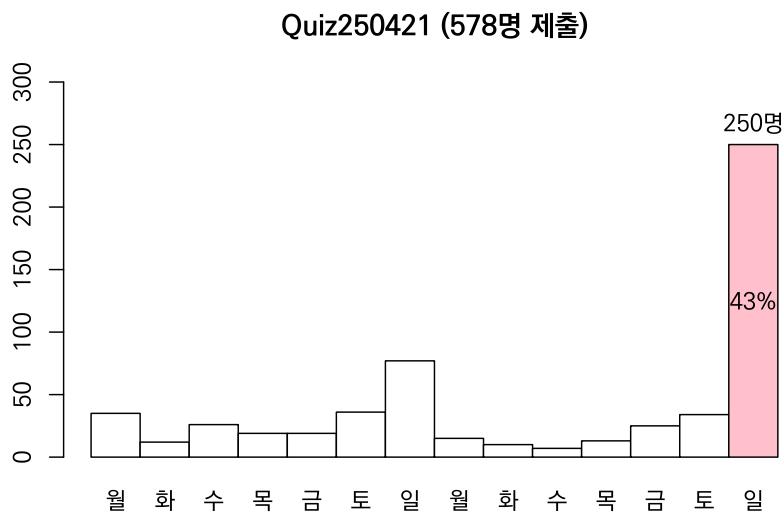
분포표의 “계”행에서 ’계’열을 제외하고 카이제곱테스트를 수행합니다.

분포표 만으로도 쉽게 파악할 수 있지만 카이제곱테스트가 명확히 해 줍니다.

카이제곱 통계량은 1234.15, 자유도는 13.00, p-value 는 7.9e-256 이므로 제출은 고르지 않고 특정 날짜에 치우쳐 있습니다.

막대그래프로 살펴 보겠습니다.

10.9.3 막대그래프



10.9.4 Red, Black 간에 닮았는가?

	14일	13일	12일	11일	10일	9일	8일	7일	6일	5일	4일	3일	2일	1일
Red	123	16	14	8	4	5	8	37	18	12	12	14	4	15
Black	127	18	11	5	3	5	7	40	18	7	7	12	8	20

표 10.27: Pearson's Chi-squared test: .

Test statistic	df	P value
6.387	13	0.9309

제출시간의 분포가 Red, Black 간에 닮았는지 알아 보았습니다.

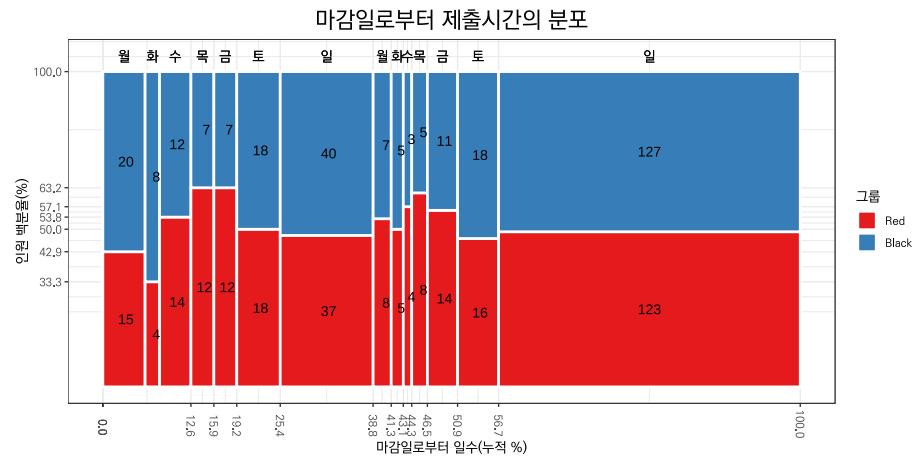
이번에는 분포표의 첫번째와 두번째 행, '계'열을 제외한 나머지 열에 대해서 카이제곱테스트를 수행합니다.

카이제곱 통계량은 6.39, 자유도는 13, p-value 는 0.93 이므로 제출 시간의 분포는 Red, Black 간에 통계적으로 유의한 차이가 관찰되지 않습니다.

이 사실을 Mosaic Plot 을 이용하여 시각적으로 살펴보겠습니다.

닮았다고 느껴지나요?

10.9.5 Mosaic Plot



제 11 장 옥스포드 행복도 조사 집계

11.1 실험의 목적

옥스포드 행복 설문지 실험결과를 분석합니다.

옥스포드 행복 설문지(Oxford Happiness Questionnaire)는 긍정심리학(Psychology of Happiness) 분야에서 널리쓰이는 자기보고식 설문도구로 “행복감(subjective well-being) 을 정량적으로 측정하기 위한 목적으로 1989년 Michael Argyle 과 Peter Hill 이 개발하였습니다 (Hills and Argyle 2002) .

총 29개 문항으로 구성되어 있으며 각 문항은 Likert 척도 (1~6점)로 응답하게 되어 있고, 일부는 역문항이 포함되어 있어서 점수를 환산할 때 주의가 필요합니다.

채점 방식은 문항당 1~6점으로 주어지고 전체 평균 점수를 계산하여 점수가 높을 수록 행복감이 높은 것으로 해석합니다.

평균 점수 4.00이하이면 상대적으로 낮은 행복감, 4.0~4.90이면 보통 수준, 5.0 이상이면 높은 행복감이라고 평가합니다.

이 실험에는 두 가지 목적이 있습니다.

첫째는 있는 그대로 옥스포드 행복 설문지로 측정하는 여러분의 행복 수준은 어느 정도인지,

둘째는 영문판으로 응답할 때와 국문판으로 응답할 때 어떤 차이가 있는지, 특히 응답에 소요되는 시간을 측정해서 비교해 보았습니다.

11.1.1 Red, Black을 잘못 표시한 사람들

	Red(옥스포드행복설문지)	Black(옥스포드행복설문지)
Red(랜덤화출석체부)	361	2
Black(랜덤화출석체부)	4	370
계	365	372

랜덤화출석부에 있는 Red, Black 과 실제 구글설문에 올린 Red, Black 이 다른 사람들의 수효는 6명입니다.

Red를 Black 이라고 한 사람이 2명, Black 을 Red 라고 한 사람이 4명입니다.

11.1.2 응답인원의 Red, Black

Red(국문) 로 응답한 인원은 365명, Black(영문) 에 응답한 인원은 372명입니다.

전체 응답인원 737 명을 랜덤하게 둘로 나눌 때 어느 한 쪽의 기대인원은 이를 반으로 나눠 준 368.5명이고, 표준오차는 이의 제곱근에 $1/2$ 을 곱해 준 13.6 명입니다.

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

11.2 행복 점수 비교

국문 설문지와 영문 설문지의 행복점수를 비교합니다.

먼저 기초통계입니다.

행복점수 단계별로 비교했을 때나 기초통계로 비교했을 때나 국문 설문지로 작성한 그룹 (Red)의 행복점수가 높게 나옵니다.

그 차이는 통계적으로 유의한 수준입니다.

11.2.1 `summary()`

- 국문:

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
1.345	3.483	4	4.035	4.552	5.655

- 영문:

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
1.793	3.448	3.707	3.855	4.207	5.655

Red(국문) 집단의 평균 행복점수는 4.04, Black(영문) 집단의 평균 행복점수는 3.86 으로 국문 설문지에 응답한 행복점수의 평균이 0.18점 더 높게 나옵니다.

중위수를 비교하더라도 Red(국문) 집단의 행복점수 중위수는 4.00, Black(영문) 집단의 행복점수 중위수는 3.71 으로 국문 설문지에 응답한 행복점수의 중위수가 0.29 점 더 높게 나옵니다.

11.2.2 t.test() on means

표 11.4: Welch Two Sample t-test: Scores by group

Test statistic	df	P value	Alternative hypothesis	mean in group 국문	mean in group 영문
3.695	693.6	0.000237 * * *	two.sided	4.035	3.855

국문 설문지에 대한 행복점수의 평균과 영문 설문지에 대한 행복점수의 평균에서 관찰된 차이가 통계적으로 유의한지 t-test로 알아 봅니다.

양 집단의 표준편차를 알 수 없기 때문에 Welch의 근사식을 활용합니다.

이 때 t-통계량은 3.70, 자유도는 693.62, p-value는 0.00024라고 나옵니다.

즉, 통계적으로 유의한 차이를 관찰하였습니다.

국문 행복도 조사지와 영문 행복도 조사지의 행복점수에는 차이가 있고, 국문 행복도의 평균점수가 더 높은 것으로 나왔습니다.

결과 요약표에는 각 집단의 평균 행복점수도 나와 있습니다.

국문으로 물어 봤을 때하고 영문으로 물어 봤을 때는 명확히 차이가 있다는 것을 알 수 있습니다.

왜 그런 차이가 관찰되는 걸까요?

11.2.3 table()

	행복하지 않다	다소 행복하지 않다	그저 그렇다	꽤 행복하다	매우 행복하다	계
국문	2	15	165	142	41	365
영문	1	12	228	113	18	372
계	3	27	393	255	59	737

이번에는 국문 행복도 조사 결과와 영문 행복도 조사 결과를 등급별로 나누어 교차표를 계산해 보았습니다.

총 737명 응답한 가운데 796명이 그저 그렇다 이상으로 행복하다고 응답하였고 그저 그렇다고 응답 한 사람들이 393명이었습니다.

국문, 영문의 차이는 어떤가요?

11.2.4 chisq.test()

표 11.6: Pearson's Chi-squared test with simulated p-value
(based on 2000 replicates): .

Test statistic	df	P value
22.61	NA	0.0009995 * * *

국문, 영문 행복도 조사 결과의 등급별 차이를 파악하기 위하여 카이제곱 테스트를 수행하였습니다.

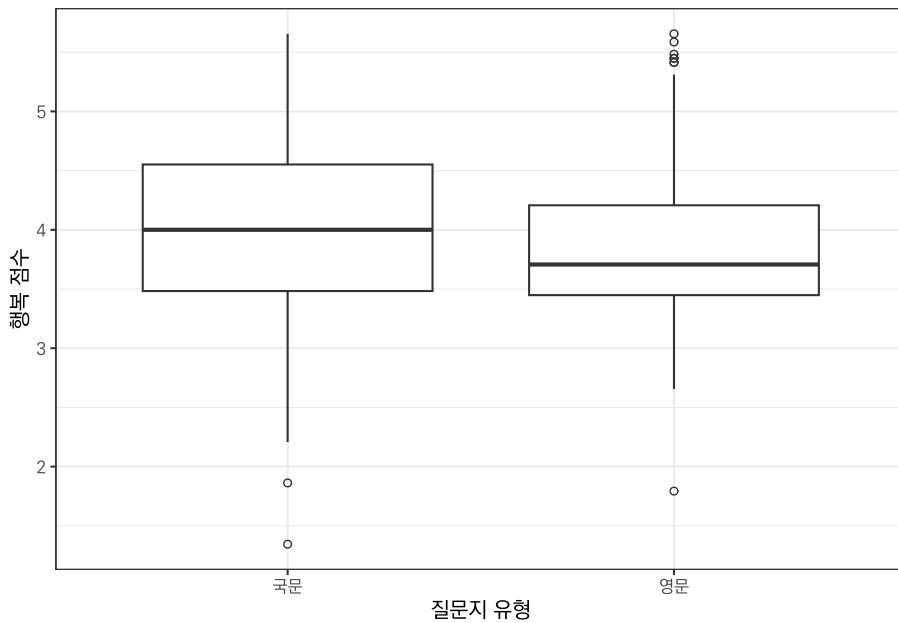
그 결과 카이제곱 통계량은 22.61, 반복수 2,0000의 븋스트랩 근사를 사용하였기 때문에 자유도는 알 수 없고, p-value 는 0.001이므로 통계적으로 유의한 차이를 보입니다.

11.2.5 백분율(%)

행복하지 않다	다소 행 복하지 않다	그저 그렇다	꽤 행복하다	매우 행복하다	계
0.4	3.7	53.3	34.6	8.0	100.0

행복점수의 등급별 백분율을 파악해 보면 꽤 행복하다 이상이 42.61%이고, 그저 그렇다가 53.32%, 다소 행복하지 않다 이하가 4.07 입니다.

다행히도 대부분 행복합니다.

11.2.6 `geom_boxplot()`

국문, 영문 행복도 조사 결과를 시각적으로 비교하기 위하여 Boxplot 을 그려 보았습니다.

국문 행복도 조사의 중위수가 영문 행복도 조사의 중위수보다 높다는 게 눈에 띕니다.

그 밖에 기초 통계에 나오는 통계값들을 시각적으로 비교할 수 있습니다.

11.3 소요시간 비교

초 단위로 측정한 소요시간에서는 오전, 오후를 잘못 쓴다거나 시작 시간을 제출 시간보다 늦게 입력한 사람들도 있고, 과도하게 (한 시간 이상) 소요된 것으로 나오는 사람들이 있어서 일일이 찾아서 수정하였고, 한 시간 이상은 사실상 설문지를 작성하다가 다른 일을 하고 다시 돌아와 마치는 것으로 판단하여 NA 로 처리하였습니다.

기초 통계부터 살펴보겠습니다.

11.3.1 `summary()`

- 국문:

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	NA's
	1	141	184	286.7	280	2907	32

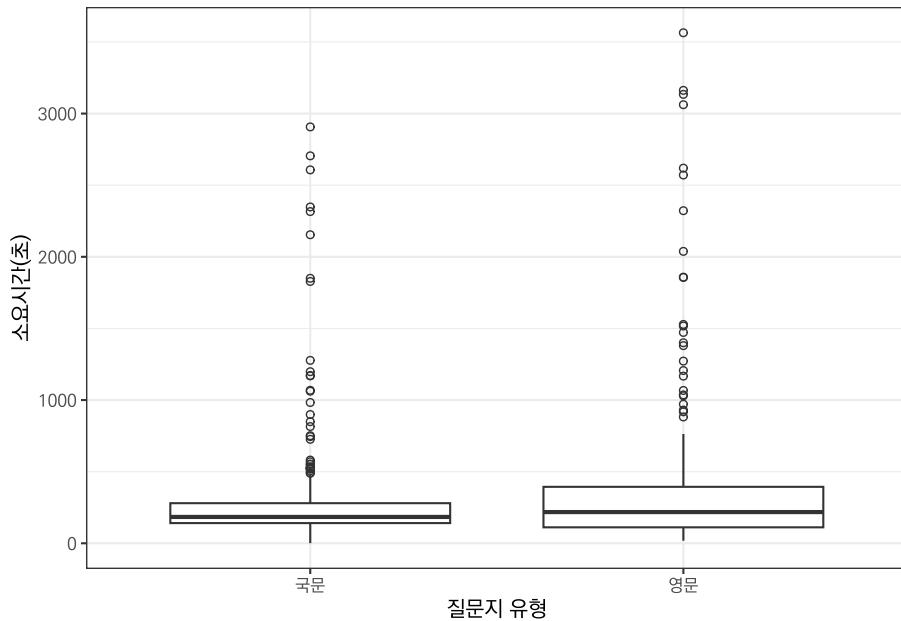
- 영문:

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	NA's
	18	111.5	218	353.1	394.2	3564	32

국문 행복도 조사의 평균 소요시간은 286.71초, 영문 행복도 조사의 평균 소요시간은 353.08초이므로 평균적으로 66.4초, 즉 1분 6초 정도 국문 행복도 조사 시간이 덜 걸립니다.

소요시간의 중위수를 비교해 보아도 국문 행복도 조사 소요시간의 중위수는 184.00초, 영문 행복도 조사 소요시간의 중위수는 218.00초이므로 34 초, 약 0분 34초 만큼 국문 행복도 조사의 소요시간이 덜 걸립니다.

11.3.2 geom_boxplot()



국문, 영문 행복도 조사 소요시간의 Boxplot 을 그려 보았습니다.

워낙 큰 값들이 많이 나오기 때문에 중위수 비교 등이 좀 어려운 측면이 있습니다.

11.3.3 t.test() on means

표 11.10: Welch Two Sample t-test: as.numeric(Duration) by group

Test statistic	df	P value	Alternative hypothesis	mean in group 국문	mean in group 영문
-1.996	641	0.04632*	two.sided	286.7	353.1

국문, 영문 행복도 조사의 소요 시간에 대해서도 t-test 를 수행하였는데 t-통계량은 -2.00, 자유도는 640.95, p-value 는 0.046이어서 통계적으로 유의한 차이가 관찰되고 있습니다.

11.4 마감 시간으로부터 제출 시간의 분포

11.4.1 분포표

표 11.11: 일 단위

	14일	13일	12일	11일	10일	9일	8일	7일	6일	5일	4일	3일	2일	1일	계
국문	116	27	8	13	12	13	15	46	29	16	21	13	15	21	365
영문	116	21	18	10	7	7	13	56	37	19	19	13	14	22	372
계	232	48	26	23	19	20	28	102	66	35	40	26	29	43	737

분포표로부터 두 가지 문제를 살펴보겠습니다.

첫째, 날마다 고르게 제출하는가?

둘째, Red(국문), Black(영문) 간에 통계적으로 유의한 차이가 있는가?

각 문제를 살펴보기 위해서는 분포표의 일부분을 대상으로 카이제곱 테스트를 수행합니다.

11.4.2 날마다 고르게 제출하는가?

	14일	13일	12일	11일	10일	9일	8일	7일	6일	5일	4일	3일	2일	1일
	232	48	26	23	19	20	28	102	66	35	40	26	29	43

표 11.13: Chi-squared test for given probabilities: .

Test statistic	df	P value
779.4	13	3.514e-158 ***

날마다 고르게 제출하는지 알아 보았습니다.

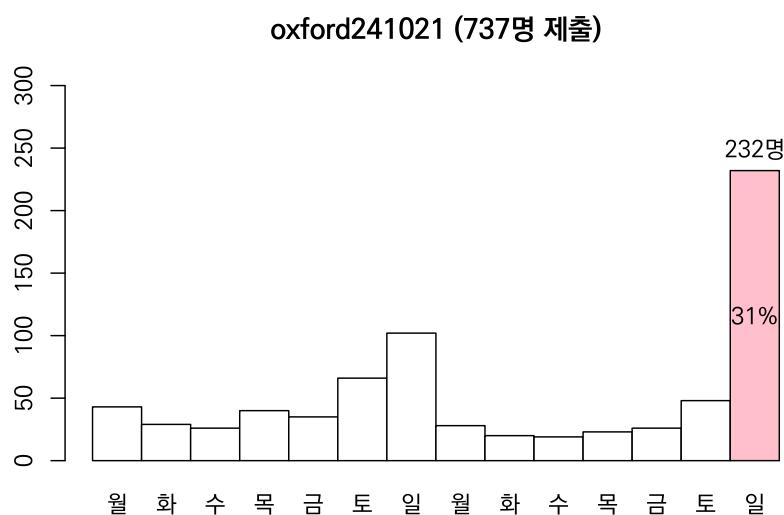
분포표의 “계”행에서 ’계’열을 제외하고 카이제곱테스트를 수행합니다.

분포표 만으로도 쉽게 파악할 수 있지만 카이제곱테스트가 명확히 해 줍니다.

카이제곱 통계량은 779.43, 자유도는 13.00, p-value 는 3.5e-158 이므로 제출은 고르지 않고 특정 날짜에 치우쳐 있습니다.

막대그래프로 살펴 보겠습니다.

11.4.3 막대그래프



11.4.4 Red(국문), Black(영문) 간에 닮았는가?

	14일	13일	12일	11일	10일	9일	8일	7일	6일	5일	4일	3일	2일	1일
국문	116	27	8	13	12	13	15	46	29	16	21	13	15	21
영문	116	21	18	10	7	7	13	56	37	19	19	13	14	22

표 11.15: Pearson's Chi-squared test: .

Test statistic	df	P value
10.55	13	0.6488

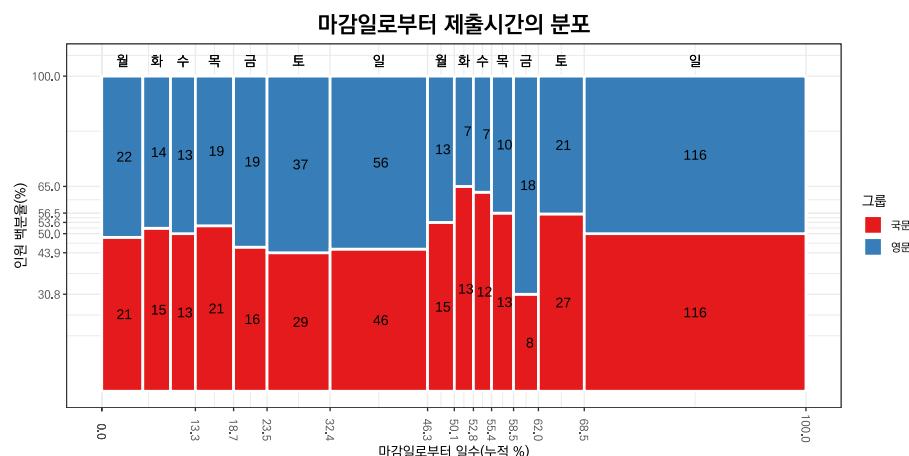
제출시간의 분포가 Red(국문), Black(영문) 간에 닮았는지 알아 보았습니다.

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

카이제곱 통계량은 10.55, 자유도는 13, p-value 는 0.65 이므로 제출 시간의 분포는 Red(국문), Black(영문) 간에 통계적으로 유의한 차이가 관찰되지 않습니다.

이 사실을 Mosaic Plot 을 이용하여 시각적으로 살펴보겠습니다. 닮았다고 느껴지나요?

11.4.5 Mosaic Plot



제 12 장 옥스포드 행복도 조사 집계

12.1 실험의 목적

옥스포드 행복 설문지 실험결과를 분석합니다.

옥스포드 행복 설문지(Oxford Happiness Questionnaire)는 긍정심리학(Psychology of Happiness) 분야에서 널리쓰이는 자기보고식 설문도구로 “행복감(subjective well-being)”을 정량적으로 측정하기 위한 목적으로 1989년 Michael Argyle 과 Peter Hill이 개발하였습니다.

총 29개 문항으로 구성되어 있으며 각 문항은 Likert 척도(1~6점)로 응답하게 되어 있고, 일부는 역문항이 포함되어 있어서 점수를 환산할 때 주의가 필요합니다.

채점 방식은 문항당 1 ~ 6점으로 주어지고 전체 평균 점수를 계산하여 점수가 높을 수록 행복감이 높은 것으로 해석합니다.

평균 점수 1점 이상 2점 미만이면 “행복하지 않다”, 2점 이상 3점 미만이면 “다소 행복하지 않다”, 3점 이상 4점 미만이면 “그저 그렇다”, 4점 이상 5점 미만이면 “꽤 행복하다”, 5점 이상이면 “매우 행복하다”로 분류합니다.

이 실험에는 두 가지 목적이 있습니다.

첫째는 있는 그대로 옥스포드 행복 설문지로 측정하는 여러분의 행복 수준은 어느 정도인지,

둘째는 영문판으로 응답할 때와 국문판으로 응답할 때 어떤 차이가 있는지, 특히 응답에 소요되는 시간을 측정해서 비교해 보았습니다.

12.1.1 Red, Black을 잘못 표시한 사람들

	Red(옥스포드행복설문지)	Black(옥스포드행복설문지)
Red(랜덤화출석체부)	287	0
Black(랜덤화출석체부)	1	284
계	288	284

랜덤화출석부에 있는 Red, Black 과 실제 구글설문에 올린 Red, Black 이 다른 사람들의 수효는 1명입니다.

Red를 Black 이라고 한 사람이 0명, Black 을 Red 라고 한 사람이 1명입니다.

12.1.2 응답인원의 Red, Black

Red(국문)로 응답한 인원은 288명, Black(영문)에 응답한 인원은 284명입니다.

전체 응답인원 572 명을 랜덤하게 둘로 나눌 때 어느 한 쪽의 기대인원은 이를 반으로 나눠 준 286명이고, 표준오차는 이의 제곱근에 $1/2$ 을 곱해 준 12 명입니다.

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

12.2 행복 점수 비교

국문 설문지와 영문 설문지의 행복점수를 비교합니다.

먼저 기초통계입니다.

행복점수 단계별로 비교했을 때나 기초통계로 비교했을 때 국문 설문지로 작성한 그룹 (Red)의 행복점수가 다소 높게 나오지만, 그 차이가 통계적으로 유의한 수준은 아닙니다.

12.2.1 `summary()`

- 국문:

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
2	3.474	3.897	3.987	4.457	5.655

- 영문:

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
1.483	3.483	3.862	3.915	4.319	5.724

Red(국문) 집단의 평균 행복점수는 3.99, Black(영문) 집단의 평균 행복점수는 3.92으로 국문 설문지에 응답한 행복점수의 평균이 0.072점 더 높게 나옵니다.

중위수를 비교하더라도 Red(국문) 집단의 행복점수 중위수는 3.90, Black(영문) 집단의 행복점수 중위수는 3.86으로 국문 설문지에 응답한 행복점수의 중위수가 0.034 점 더 높게 나옵니다.

12.2.2 t.test() on means

표 12.4: Welch Two Sample t-test: Scores by group

Test statistic	df	P value	Alternative hypothesis	mean in group 국문	mean in group 영문
1.343	563.6	0.1797	two.sided	3.987	3.915

국문 설문지에 대한 행복점수의 평균과 영문 설문지에 대한 행복점수의 평균에서 관찰된 차이가 통계적으로 유의한지 t-test로 알아 봅니다.

양 집단의 표준편차를 알 수 없기 때문에 Welch의 근사식을 활용합니다.

이 때 t-통계량은 1.34, 자유도는 563.63, p-value는 0.1797라고 나옵니다.

즉, 통계적으로 유의한 차이를 관찰하지 못하였습니다.

국문 행복도 조사와 영문 행복도 조사의 평균 행복점수 차이는 랜덤화의 효과로 충분히 나올 수 있는 수준이라는 것입니다.

결과 요약표에는 각 집단의 평균 행복점수도 나와 있습니다.

지난 학기의 자료들을 살펴보면 국문으로 물어 봤을 때하고 영문으로 물어 봤을 때의 행복 점수에서 통계적으로 유의한 차이가 있었다는 것을 알 수 있습니다.

지난 자료와 금번 자료에서는 왜 이러한 차이가 관찰되는 걸까요?

12.2.3 table()

	행복하지 않다	다소 행복하지 않다	그저 그렇다	꽤 행복하다	매우 행복하다	계
국문	0	11	148	100	29	288
영문	2	6	155	108	13	284
계	2	17	303	208	42	572

이번에는 국문 행복도 조사 결과와 영문 행복도 조사 결과를 등급별로 나누어 교차표를 계산해 보았습니다.

총 572 명 응답한 가운데 250명이 꽤 행복하다 이상으로 행복하다고 응답하였고 그저 그렇다고 응답 한 사람들이 303 명이었습니다.

국문, 영문의 차이는 어떤가요?

12.2.4 Cochran-ArmitageTest()

표 12.6: Cochran-Armitage test for trend: .

Test statistic	dim	P value	Alternative hypothesis
1.278	5	0.2012	two.sided

국문, 영문 행복도 조사 결과의 등급별 차이가 통계적으로 유의한지 파악하기 위하여 순서형 변수 분석에 적합한 코크란-아미티지 테스트를 수행하였습니다.

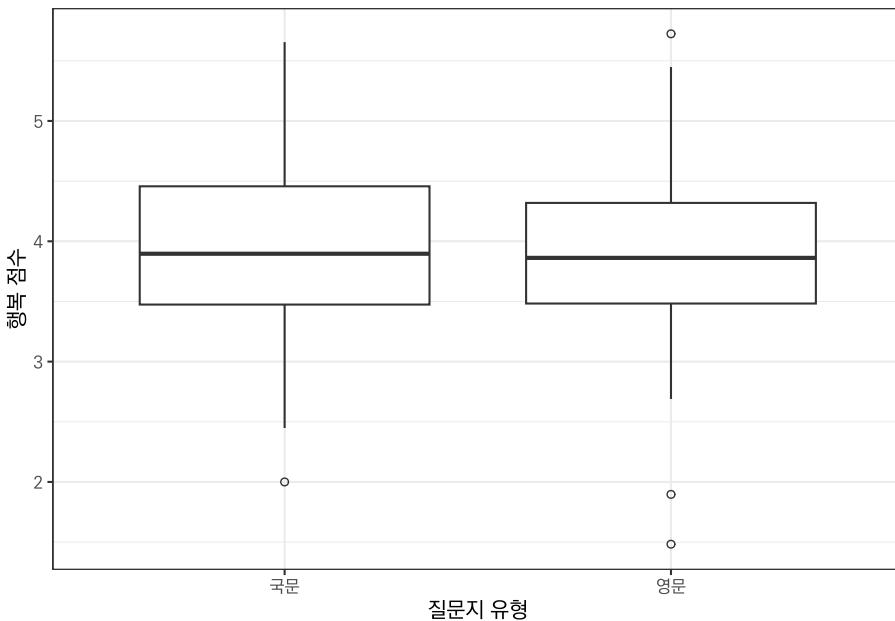
그 결과 코크란-아미키지 통계량은 1.28, dim은 5, p-value 는 0.2012이므로 통계적으로 유의한 차이를 보이고 있지 않습니다.

12.2.5 백분율(%)

행복하지 않다	다소 행복하지 않다	그저 그렇다	꽤 행복하다	매우 행복하다	계
0.3	3.0	53.0	36.4	7.3	100.0

행복점수의 등급별 백분율을 파악해 보면 꽤 행복하다 이상이 43.71%이고, 그저 그렇다가 52.97%, 다소 행복하지 않다 이하가 3.32%입니다.

다행히도 대부분 행복합니다.

12.2.6 `geom_boxplot()`

국문, 영문 행복도 조사 결과를 시각적으로 비교하기 위하여 Boxplot 을 그려 보았습니다.

국문 행복도 조사의 중위수가 영문 행복도 조사의 중위수보다 높다는 게 눈에 들어옵니다.

그 밖에 기초 통계에 나오는 통계값들을 시각적으로 비교할 수 있습니다.

12.3 소요시간 비교

초 단위로 측정한 소요시간에서는 오전, 오후를 잘못 쓴다거나 시작 시간을 제출 시간보다 늦게 입력한 사람들도 있고, 과도하게 (한 시간 이상) 소요된 것으로 나오는 사람들이 있어서 일일이 찾아서 수정하였고, 한 시간 이상은 사실상 설문지를 작성하다가 다른 일을 하고 다시 돌아와 마치는 것으로 판단하여 NA 로 처리하였습니다.

기초 통계부터 살펴보겠습니다.

12.3.1 `summary()`

- 국문:

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
17	130	190.5	292.6	296	3268

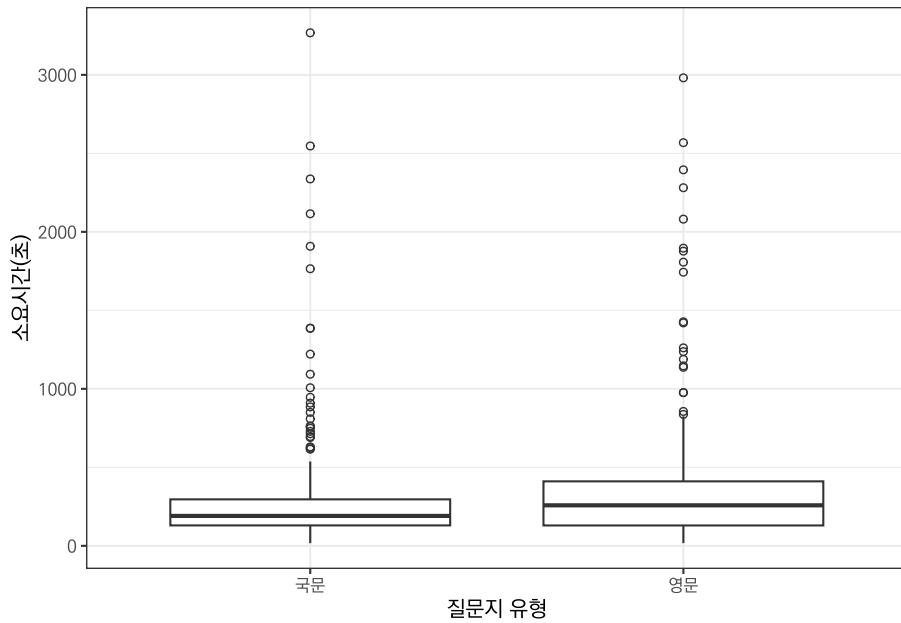
- 영문:

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
17	129.8	258	362.5	410.8	2981

국문 행복도 조사의 평균 소요시간은 292.58초, 영문 행복도 조사의 평균 소요시간은 362.47초이므로 평균적으로 69.9초, 즉 1분 10초 정도 국문 행복도 조사 시간이 덜 걸립니다.

소요시간의 중위수를 비교해 보아도 국문 행복도 조사 소요시간의 중위수는 190.50초, 영문 행복도 조사 소요시간의 중위수는 258.00초이므로 67.5 초, 약 1분 8초 만큼 국문 행복도 조사의 소요시간이 덜 걸립니다.

12.3.2 geom_boxplot()



국문, 영문 행복도 조사 소요시간의 Boxplot 을 그려 보았습니다.

워낙 큰 값들이 많이 나오기 때문에 중위수 비교 등이 좀 어려운 측면이 있습니다.

12.3.3 t.test() on means

표 12.10: Welch Two Sample t-test: as.numeric(Duration) by group

Test statistic	df	P value	Alternative hypothesis	mean in group 국문	mean in group 영문
-2.12	560.4	0.03447*	two.sided	292.6	362.5

국문, 영문 행복도 조사의 소요 시간에 대해서도 t-test 를 수행하였는데 t-통계량은 -2.12, 자유도는 560.42, p-value 는 0.0340이어서 통계적으로 유의한 차이가 관찰되고 있습니다.

12.4 마감 시간으로부터 제출 시간의 분포

12.4.1 분포표

표 12.11: 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]														
국문	120	16	12	8	4	5	6	39	19	9	14	15	7	14	288
영문	117	20	15	6	2	6	7	36	20	8	7	12	8	20	284
계	237	36	27	14	6	11	13	75	39	17	21	27	15	34	572

분포표로부터 두 가지 문제를 살펴보겠습니다.

첫째, 날마다 고르게 제출하는가?

둘째, Red(국문), Black(영문) 간에 통계적으로 유의한 차이가 있는가?

각 문제를 살펴보기 위해서는 분포표의 일부분을 대상으로 카이제곱 테스트를 수행합니다.

12.4.2 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]														
	237	36	27	14	6	11	13	75	39	17	21	27	15	34	

표 12.13: Chi-squared test for given probabilities: .

Test statistic	df	P value
1110	13	5.117e-229 ***

날마다 고르게 제출하는지 알아 보았습니다.

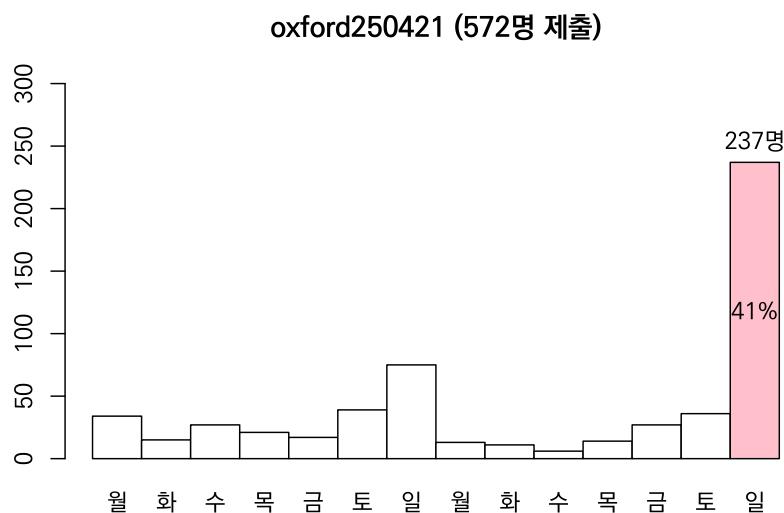
분포표의 “계”행에서 ‘계’열을 제외하고 카이제곱테스트를 수행합니다.

분포표 만으로도 쉽게 파악할 수 있지만 카이제곱테스트가 명확히 해 줍니다.

카이제곱 통계량은 1109.52, 자유도는 13.00, p-value 는 5.1e-229 이므로 제출은 고르지 않고 특정 날짜에 치우쳐 있습니다.

막대그래프로 살펴 보겠습니다.

12.4.3 막대그래프



12.4.4 Red(국문), Black(영문) 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]													
국문	120	16	12	8	4	5	6	39	19	9	14	15	7	14
영문	117	20	15	6	2	6	7	36	20	8	7	12	8	20

표 12.15: Pearson's Chi-squared test: .

Test statistic	df	P value
5.905	13	0.9495

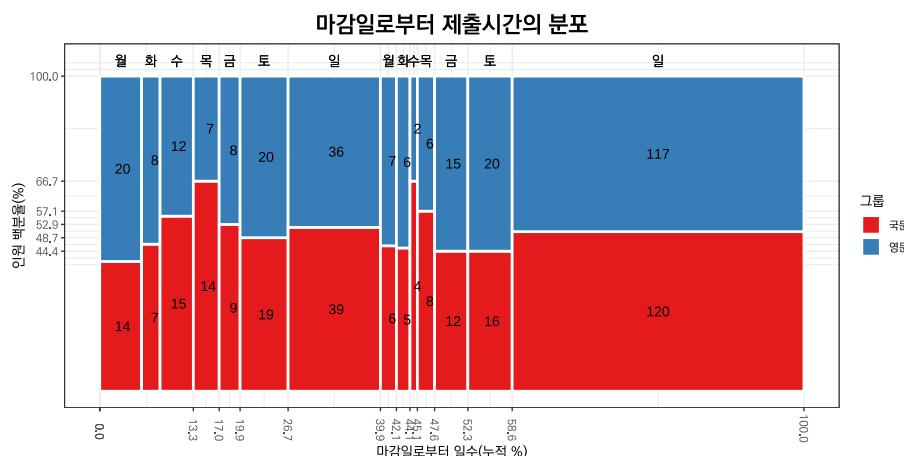
제출시간의 분포가 Red(국문), Black(영문) 간에 닮았는지 알아 보았습니다.

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

카이제곱 통계량은 5.90, 자유도는 13, p-value 는 0.95 이므로 제출 시간의 분포는 Red(국문), Black(영문) 간에 통계적으로 유의한 차이가 관찰되지 않습니다.

이 사실을 Mosaic Plot 을 이용하여 시각적으로 살펴보겠습니다. 닮았다고 느껴지나요?

12.4.5 Mosaic Plot



제 13 장 9주차 데이터 실험 집계

13.1 실험의 목적

9주차 구글 예습 설문지 집계결과를 분석합니다.

Q1~Q6에서는 랜덤화의 효과로 Red, Black이 얼마나 닮았는지 알아봅니다.

Q7에서는 Default Effects에 대한 데이터실험 결과에 대하여 분석합니다.

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

13.1.1 Red, Black을 잘못 표시한 사람들

	Red(구글예습퀴즈)	Black(구글예습퀴즈)
Red(랜덤화출석부)	355	0
Black(랜덤화출석부)	1	363
계	356	363

13.1.2 응답인원의 Red, Black

Red로 응답한 인원은 356명, Black에 응답한 인원은 363명입니다.

전체 응답인원 719명을 랜덤하게 둘로 나눌 때 어느 한 쪽의 기대인원은 전체 응답인원의 절반인 359.5명이고, 표준오차는 전체 응답인원의 제곱근에 $1/2$ 을 곱해 준 13.4명입니다.

따라서 Red, Black 각 그룹에 관찰된 인원은 기대인원으로부터 표준오차 범위 안에 들어갑니다. 랜덤화출석부에 있는 Red, Black과 실제 구글설문에 올린 Red, Black이 다른 사람들의 수효는 1명입니다. Red를 Black이라고 한 사람이 0명, Black을 Red라고 한 사람이 1명입니다.

13.2 Q1. OECD 2022년 9월 실업률 비교

1. 다음 중에서 2022년 9월 기준으로 실업률이 가장 낮은 곳은 어디인가? *



- 스페인
- 미국
- OECD 전체
- 대한민국

13.2.1 실업률 가장 낮은 나라 (집계))

	스페인	미국	OECD 전체	대한민국	계
Red	17	26	10	303	356
Black	23	29	14	297	363
계	40	55	24	600	719

표 13.3: Pearson's Chi-squared test: .

Test statistic	df	P value
1.722	3	0.632

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

그 결과 카이제곱 통계량은 1.722, 자유도는 3, p-value 는 0.63200으로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

13.2.2 실업률 가장 낮은 나라 (%)

스페인	미국	OECD 전체	대한민국	계
5.6	7.6	3.3	83.4	100.0

정답률은 Red, Black 을 합하여 계산하는데, 83.4(%) 입니다.

13.3 Q2. 고용률의 산식



13.3.1 고용률?

	경제활동인구 / 15세이상 인구	취업자 / 경제활동인구	취업자 / 15세 이상 인구	실업자 / 경제활동인구	계
Red	48	62	231	15	356
Black	57	70	218	18	363
계	105	132	449	33	719

표 13.6: Pearson's Chi-squared test: .

Test statistic	df	P value
1.837	3	0.6068

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

그 결과 카이제곱 통계량은 1.837, 자유도는 3, p-value 는 0.60680이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

13.3.2 고용률? (%)

경제활동인구 / 15세이상 인구	취업자 / 경제활동인구	취업자 / 15세 이상 인구	실업자 / 경제활동인구	계
14.6	18.4	62.4	4.6	100.0

정답률은 Red, Black 을 합하여 계산하는데, 62.4(%) 입니다.

13.4 Q3. 취업자의 조건

3. 취업자의 조건에 해당하지 않는 것은 다음 중 어느 것인가? *

- 조사 대상 주간 중 수입을 목적으로 1시간 이상 일한 자
- 자기 가구에서 경영하는 농장이나 사업체의 수입을 높이는 데 도운 가족종사자로서 주당 18시간 이상 일한 자
- 조사대상 주간에 일시적으로 병이 나거나, 날씨가 안 좋거나, 휴가 또는 연기, 노동쟁의 등의 이유로 일을 못한 일시 휴직자
- 조사대상 주간에 수입 있는 일을 하지 않았으나, 지난 4주간 구직활동하였으며 즉시 취업 가능한자

13.4.1 취업자란?

	조사 대상 주간 중 수입을 목적으로 1시간 이상 일한 자	자기 가구에서 경영하는 농장이나 사업체의 수입을 높이는 데 도운 가족종사자로서 주당 18시간 이상 일한 자	조사대상 주간에 일시적으로 병이 나거나, 날씨가 안 좋거나, 휴가 또는 연가, 노동쟁의 등의 이유로 일을 못한 일시 휴직자	조사대상 주간에 수입 있는 일을 하지 않았으나, 지난 4주간 구직활동하였으며 즉시 취업 가능한자	계
Red	31	28	52	245	356
Black	29	41	51	242	363
계	60	69	103	487	719

표 13.9: Pearson's Chi-squared test: .

Test statistic	df	P value
2.476	3	0.4796

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

그 결과 카이제곱 통계량은 2.476, 자유도는 3, p-value 는 0.47960이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

13.4.2 취업자란? (%)

	조사 대상 주간 중 수입을 목적으로 1시간 이상 일한 자	자기 가구에서 경영하는 농장이나 사업체의 수입을 높이는 데 도운 가족종사자로서 주당 18시간 이상 일한 자	조사대상 주간에 일시적으로 병이 나거나, 날씨가 안 좋거나, 휴가 또는 연가, 노동쟁의 등의 이유로 일을 못한 일시 휴직자	조사대상 주간에 수입 있는 일을 하지 않았으나, 지난 4주간 구직활동하였으며 즉시 취업 가능한자	계
	8.3	9.6	14.3	67.7	100.0

정답률은 Red, Black 을 합하여 계산하는데, 67.7(%) 입니다.

13.5 Q4. 평균과 중위값

4. 5명이 다니는 회사에서 4명은 월 100만원을 받고, 한 명은 월 600만원을 받을 때, 월 급여의 평균과 중위값을 * 제대로 짹 지은 것은?

- 평균 : 100만원, 중위값 : 200만원
- 평균 : 200만원, 중위값 : 100만원
- 평균 : 1,000만원, 중위값 : 100만원
- 평균 : 100만원, 중위값 : 1,000만원

13.5.1 집계

	평균 : 100만원, 중위값 : 200만원	평균 : 200만원, 중위값 : 100만원	평균 : 1,000만원, 중위값 : 100만원	평균 : 100만원, 중위값 : 1,000만원	계
Red	41	261	36	18	356
Black	25	279	44	15	363
계	66	540	80	33	719

표 13.12: Pearson's Chi-squared test: .

Test statistic	df	P value
5.484	3	0.1396

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

그 결과 카이제곱 통계량은 5.484, 자유도는 3, p-value 는 0.1396이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

13.5.2 %

평균 : 100만원, 중위값 : 200만원	평균 : 200만원, 중위값 : 100만원	평균 : 1,000만원, 중위값 : 100만원	평균 : 100만원, 중위값 : 1,000만원	계
9.2	75.1	11.1	4.6	100.0

정답률은 Red, Black 을 합하여 계산하는데, 75.1(%) 입니다.

13.6 Q5. 평균과 중위값

5. 4번 회사에서 4명의 월 급여는 그대로인데 다른 한 사람은 월 4,600만원을 받게 될 때, 월 급여의 평균과 중위값을 제대로 짹 지은 것은? *

- 평균 : 100만원, 중위값 : 200만원
- 평균 : 200만원, 중위값 : 100만원
- 평균 : 1,000만원, 중위값 : 100만원
- 평균 : 100만원, 중위값 : 1,000만원

13.6.1 집계

평균 : 100만원, 중위값 : 200만원	평균 : 200만원, 중위값 : 100만원	평균 : 1,000만원, 중위값 : 100만원	평균 : 100만원, 중위값 : 1,000만원	계
Red	26	52	252	26
Black	20	46	267	30
계	46	98	519	719

표 13.15: Pearson's Chi-squared test: .

Test statistic	df	P value
1.801	3	0.6147

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

그 결과 카이제곱 통계량은 1.801, 자유도는 3, p-value 는 0.6147이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

13.6.2 %

평균 : 100만원, 중위값 : 200만원	평균 : 200만원, 중위값 : 100만원	평균 : 1,000만원, 중위값 : 100만원	평균 : 100만원, 중위값 : 1,000만원	계
6.4	13.6	72.2	7.8	100.0

정답률은 Red, Black 을 합하여 계산하는데, 72.2(%) 입니다.

13.7 Q6. 연 500만원 이하 근로소득자의 비율

6. 2010년 기준 우리나라 근로소득자 중 연 500만 원 이하의 소득을 신고한 사람은 몇% 인가? *

<표2> 급여규모별 인원과 근로소득 총액 비중(연말정산 신고기준, 2009~10년, 단위:%)

급여규모별	2009년				2010년			
	인원		근로소득 총액		인원		근로소득 총액	
전체	비중	누적비	비중	누적비	비중	누적비	비중	누적비
5백만원 이하	18.6	18.6	1.8	1.8	19.1	19.1	1.7	1.7
5백-1천만원	13.0	31.6	3.9	5.7	12.3	31.4	3.6	5.3
1천-2천만원	23.1	54.6	13.3	19.0	22.8	54.3	12.8	18.1
2천-3천만원	14.7	69.3	14.3	33.3	14.4	68.7	13.6	31.7
3천-4천만원	10.1	79.4	13.9	47.1	9.8	78.5	13.0	44.7
4천-6천만원	12.0	91.4	23.3	70.4	12.0	90.5	22.5	67.2
6천-8천만원	5.3	96.7	14.3	84.7	5.4	95.9	14.1	81.3
8천만-1억원	1.9	98.6	6.6	91.3	2.3	98.2	7.8	89.0
1-2억원 이하	1.2	99.8	5.9	97.2	1.6	99.8	7.7	96.7
2-3억원 이하	0.1	99.9	0.9	98.1	0.1	99.9	1.2	97.9
3-5억원 이하	0.0	100.0	0.7	98.8	0.1	100.0	0.8	98.7
5-10억원 이하	0.0	100.0	0.6	99.4	0.0	100.0	0.6	99.4
10억원 초과	0.0	100.0	0.6	100.0	0.0	100.0	0.6	100.0
상위10%	10.0		32.9		10.0		33.9	

자료: 국세청, 국제통계연보 2010-11년

- "5%"
- "10%"
- "15%"
- "19.1%"
- "25%"

13.7.1 집계

	5%	10%	15%	19.1%	25%	계
Red	20	43	37	244	12	356
Black	32	42	44	233	12	363
계	52	85	81	477	24	719

표 13.18: Pearson's Chi-squared test: .

Test statistic	df	P value
3.569	4	0.4674

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

그 결과 카이제곱 통계량은 3.569, 자유도는 4, p-value 는 0.46740이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

13.7.2 %

5%	10%	15%	19.1%	25%	계
7.2	11.8	11.3	66.3	3.3	100.0

정답률은 Red, Black 을 합하여 계산하는데, 66.3(%) 입니다.

13.8 Q7. 디폴트 효과 : “… 하면 표시하세요”의 효과

표시하지 않으면 사후 장기기증 의사가 있는 것으로 간주하는 경우(Red)와 표시하면 사후 장기기증 의사가 있는 것으로 간주하는 경우(Black), 표시하지 않으면 사후 장기기증 의사가 있다고 보는 질문지(Red)에 사후 장기기증 의사 비율이 더 높게 나오는 것으로 알려져 있으나 관찰 결과는 반대로 나옵니다.

어디에 디폴트를 설정하느냐에 따라 조사결과가 달라지는 것으로 알려져 있으나 응답 결과는 랜덤화 효과에 가깝게 나와서 당황스럽습니다.

Red
<p>7. (Johnson & Goldstein 2003) 여러분이 새로 이사온 나라(덴마크, 네덜란드, 영국, 독일) *에서는 사후에 장기를 기증하는 것이 기본(디폴트, default)이고 원하지 않으면 직접 의사 표시를 하여야 합니다. 사후 장기 기증의사가 없으면 표시하세요. 사후 장기 기증의사가 있으면 빈 칸으로 놓아 두면 됩니다.</p> <p><input type="checkbox"/> 나는 사후 장기기증에 참여할 뜻이 없다</p>
Black
<p>7. (Johnson & Goldstein 2003) 여러분이 새로 이사온 나라(오스트리아, 벨기에, 프랑스, *헝가리, 폴란드, 포르투갈, 스웨덴)에서는 사후에 장기를 기증하지 않는 것이 기본(디폴트, default)이고 원하면 직접 의사 표시를 하여야 합니다. 사후 장기 기증의사가 있으면 표시하세요. 사후 장기 기증의사가 없으면 빈 칸으로 놓아 두면 됩니다.</p> <p><input type="checkbox"/> 나는 사후 장기기증에 참여할 뜻이 있다</p>

13.8.1 집계 (전체)

	사후 장기기증 의사 있음	사후 장기기증 의사 없음	계
Red(사후 장기기증 의사 없으면 표시)	170	186	356
Black(사후 장기기증 의사 있으면 표시)	180	183	363
계	350	369	719

표 13.21: Pearson's Chi-squared test with Yates' continuity correction: .

Test statistic	df	P value
0.1741	1	0.6765

Q7의 Red 는 default 가 사후 장기기증 의사가 있는 것으로서 사후 장기기증 의사가 없으면 표시도록 하고, Black 은 default 가 사후 장기기증 의사가 없는 것으로서 사후 장기기증 의사가 있으면 표시도록 합니다.

표시하는 것조차 귀찮은 사람들은 당연히 default 를 따를 것으로 기대되어 Red 에서는 장기기증 의사가 있다는 응답 비율이 더 높고, Black 에서는 장기기증 의사가 없다는 응답 비율이 더 높을 것으로 기대되었지만 지난 학기들의 기록이나 이번 학기의 집계결과를 보면 그렇지 않아서 당황스럽습니다.

여기서 수행하고 있는 카이제곱 테스트는 default 를 어떻게 설정하느냐의 효과가 통계적으로 유의한지 알아보기 위한 것입니다.

그 결과 카이제곱 통계량은 0.17, 자유도는 1, p-value 는 0.680이므로 default의 설정을 바꾼 것은 사후 장기기증 의사에 통계적으로 유의한 차이를 보이지 않습니다.

13.8.2 % 비교

	사후 장기기증 의사 있음	사후 장기기증 의사 없음	계
Red(사후 장기기증 의사 없으면 표시)	47.8	52.2	100.0
Black(사후 장기기증 의사 있으면 표시)	49.6	50.4	100.0

이를 백분율로 살펴보면 사후 장기기증 의사가 없으면 표시하라는 Red에서 사후 장기기증 의사가 있다고 응답한 백분율, 47.8%(0이)나 사후 장기기증 의사가 있으면 표시하라는 Black에서 사후 장기기증 의사가 있다고 응답한 백분율, 49.6%(0이)나 큰 차이가 없어서 default의 설정이 응답에는 통계적으로 유의한 영향을 주지 못한다는 것을 명확히 알 수 있습니다.

오히려 기대와는 반대로 장기기증 의사가 있으면 표시해야 하는 Black 의 사후 장기기증 의사가 있다는 응답 백분율이 더 높습니다.

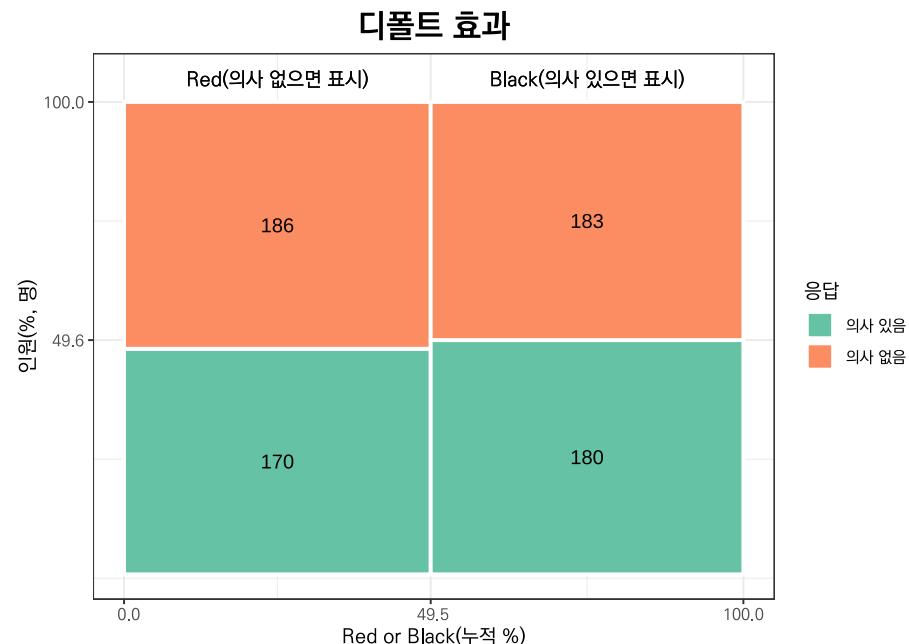
어쨌든 통계적으로 유의한 차이를 관찰하지 못했으므로 Red, Black 은 합쳐서 분석하는 것이 타당합니다. 그 결과는 다음과 같습니다.

13.8.3 합산(%)

사후 장기기증 의사 있음	사후 장기기증 의사 없음	계
48.7	51.3	100.0

사후 장기기증 의사가 있다고 응답한 백분률은 default의 설정과는 무관하게 48.7%(으)로 절반이 좀 안 되는군요.

13.8.4 Mosaic Plot



Mosaic Plot은 이 집계결과를 시각적으로 잘 보여줍니다.

default의 설정에 거의 무관할 정도로 응답이 비슷하기 때문에 Red, Black이 닮았고, 사후 장기기증 의사가 있다고 응답한 백분율도 닮은 것을 잘 알 수 있습니다.

13.9 Thinking, Fast and Slow

13.9.1 Article

The evidence comes from a comparison of the rate of organ donation in European countries, which reveals startling differences between neighboring and culturally similar countries. An article published in 2003 noted that the rate of organ donation was close to 100% in Austria but only 12% in Germany, 86% in Sweden but only 4% in Denmark.

우리의 관찰은 Thinking, Fast and Slow (Daniel Kahneman, 2011)의 default effects에서 소개한 유럽의 사례와는 아주 차이가 많이 납니다.

2003년 오스트리아는 100%, 스웨덴은 86%가 사후 장기기증 의사가 있다고 한 반면에 독일은 12%, 덴마크는 4% 밖에 사후 장기기증 의사가 있다는 조사 결과가 나왔습니다.

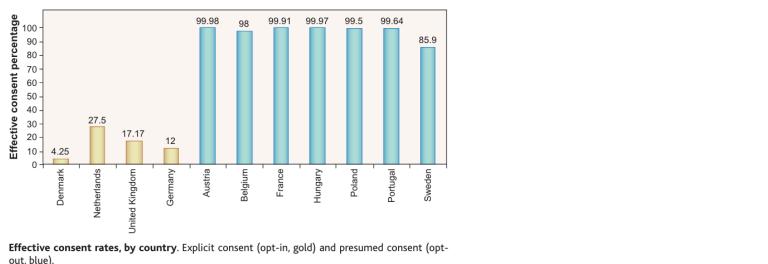
13.9.2 Facts Found

These enormous differences are a framing effect, which is caused by the format of the critical question. The high-donation countries have an opt-out form, where individuals who wish not to donate must check an appropriate box. Unless they take this simple action, they are considered willing donors. The low-contribution countries have an opt-in form: you must check a box to become a donor. That is all. The best single predictor of whether or not people will donate their organs is the designation of the default option that will be adopted without having to check a box.

여기서 결정적인 차이는 오스트리아와 스웨덴은 사후 장기기증 의사가 없으면 표시를 해야 하고, 독일과 덴마크는 사후 장기기증 의사가 있으면 표시를 해야 하는 시스템이었다는 점입니다.

여러분들의 응답과는 아주 다른 상황이죠?

13.9.3 Original Article (Johnson & Goldstein 2003)



Johnson and Goldstein의 원전에 나오는 그래프를 옮겨 왔습니다.

황금색으로 표시한 나라들은 사후 장기기증 의사를 직접 표시해야 하는 나라들이고 하늘색으로 표시한 나라들은 사후 장기기증 의사를 default로 하여 사후 장기기증 의사가 없으면 직접 표시해야 하는 나라들입니다.

여기서 오스트리아, 독일, 스웨덴, 덴마크의 수치를 확인해 보세요.

13.10 마감 시간으로부터 제출 시간의 분포

13.10.1 분포표

표 13.24: 일 단위

	14일	13일	12일	11일	10일	9일	8일	7일	6일	5일	4일	3일	2일	1일	계
Red	112	22	10	13	8	6	13	47	20	13	27	20	20	25	356
Black	106	24	15	8	8	11	10	46	16	25	18	25	25	26	363
계	218	46	25	21	16	17	23	93	36	38	45	45	45	51	719

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.

13.10.2 날마다 고르게 제출하는가?

14일	13일	12일	11일	10일	9일	8일	7일	6일	5일	4일	3일	2일	1일
218	46	25	21	16	17	23	93	36	38	45	45	45	51

표 13.26: Chi-squared test for given probabilities: .

Test statistic	df	P value
679.9	13	6.699e-137 ***

날마다 고르게 제출하는지 알아 보았습니다.

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

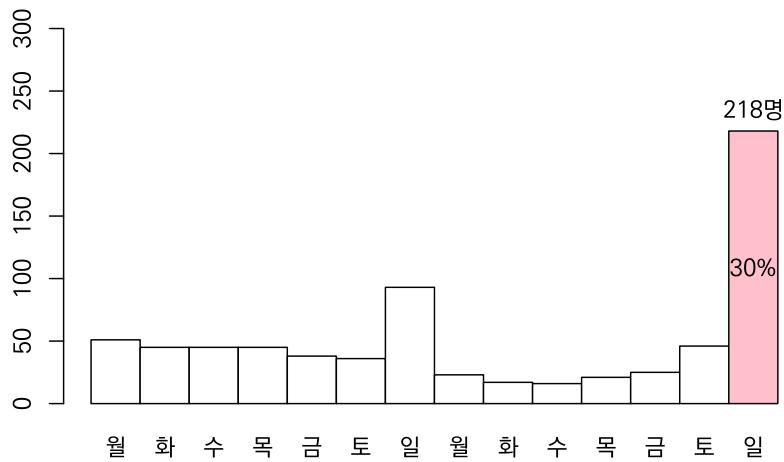
분포표 만으로도 쉽게 파악할 수 있지만 카이제곱테스트가 명확히 해 줍니다.

카이제곱 통계량은 679.93, 자유도는 13.00, p-value 는 6.7e-137 이므로 날짜별로 고르게 제출하고 있지 않음을 시사합니다.

막대그래프로 살펴 보겠습니다.

13.10.3 막대그래프

Quiz241028 (719명 제출)



13.10.4 Red, Black 간에 닮았는가?

	14일	13일	12일	11일	10일	9일	8일	7일	6일	5일	4일	3일	2일	1일
Red	112	22	10	13	8	6	13	47	20	13	27	20	20	25
Black	106	24	15	8	8	11	10	46	16	25	18	25	25	26

표 13.28: Pearson's Chi-squared test: .

Test statistic	df	P value
11.41	13	0.5763

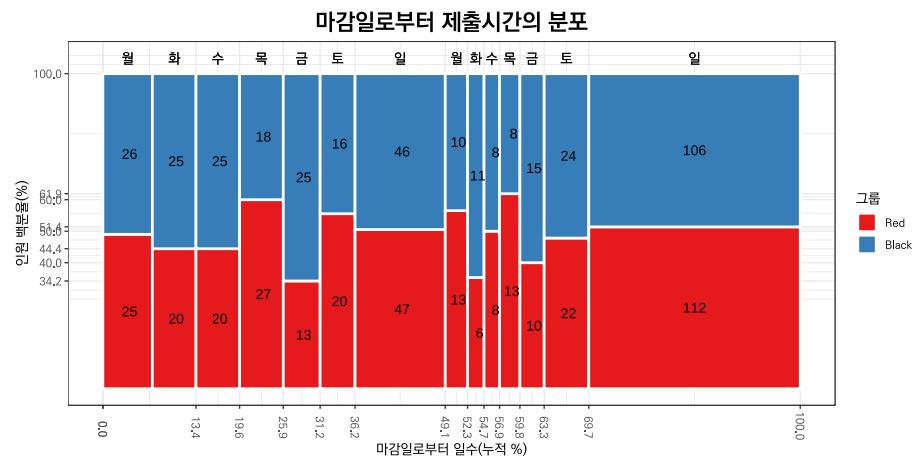
제출시간의 분포가 Red, Black 간에 닮았는지 알아 보았습니다.

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

카이제곱 통계량은 11.413, 자유도는 13, p-value 는 0.57630이므로 제출 시간의 분포는 Red, Black 간에 통계적으로 유의한 차이가 관찰되지 않습니다.

이 사실을 Mosaic Plot 을 이용하여 시각적으로 살펴보겠습니다. 닮았다고 느껴지나요?

13.10.5 Mosaic Plot



제 14 장 9주차 데이터 실험 집계

14.1 실험의 목적

9주차 구글 예습 설문지 집계결과를 분석합니다.

Q1 ~ Q6에서는 랜덤화의 효과로 Red, Black이 얼마나 닮았는지 알아봅니다.

Q7에서는 Default Effects에 대한 데이터실험 결과에 대하여 분석합니다.

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

14.1.1 Red, Black을 잘못 표시한 사람들

	Red(구글예습퀴즈)	Black(구글예습퀴즈)
Red(랜덤화출석부)	285	0
Black(랜덤화출석부)	2	285
계	287	285

14.1.2 응답인원의 Red, Black

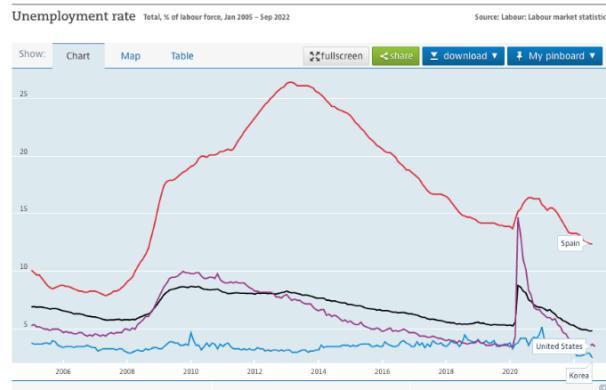
Red로 응답한 인원은 287명, Black에 응답한 인원은 285명입니다.

전체 응답인원 572명을 랜덤하게 둘로 나눌 때 어느 한 쪽의 기대인원은 전체 응답인원의 절반인 286명이고, 표준오차는 전체 응답인원의 제곱근에 $1/2$ 을 곱해 준 12명입니다.

따라서 Red, Black 각 그룹에 관찰된 인원은 기대인원으로부터 표준오차 범위 안에 들어갑니다. 랜덤화출석부에 있는 Red, Black과 실제 구글설문에 올린 Red, Black이 다른 사람들의 수효는 2명입니다. Red를 Black이라고 한 사람이 0명, Black을 Red라고 한 사람이 2명입니다.

14.2 Q1. OECD 2022년 9월 실업률 비교

1. 다음 중에서 2022년 9월 기준으로 실업률이 가장 낮은 곳은 어디인가? *



- 스페인
- 미국
- OECD 전체
- 대한민국

14.2.1 실업률 가장 낮은 나라 (집계))

	스페인	미국	OECD 전체	대한민국	계
Red	12	32	8	235	287
Black	19	20	7	239	285
계	31	52	15	474	572

표 14.3: Pearson's Chi-squared test: .

Test statistic	df	P value
4.443	3	0.2174

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

그 결과 카이제곱 통계량은 4.443, 자유도는 3, p-value 는 0.21740|으로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

14.2.2 실업률 가장 낮은 나라 (%)

스페인	미국	OECD 전체	대한민국	계
5.4	9.1	2.6	82.9	100.0

정답률은 Red, Black 을 합하여 계산하는데, 82.9(%) 입니다.

14.3 Q2. 고용률의 산식



14.3.1 고용률?

	경제활동인구 / 15세이상 인구	취업자 / 경제활동인구	취업자 / 15세 이상 인구	실업자 / 경제활동인구	계
Red	36	47	188	16	287
Black	48	49	177	11	285
계	84	96	365	27	572

표 14.6: Pearson's Chi-squared test: .

Test statistic	df	P value
3.006	3	0.3906

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

그 결과 카이제곱 통계량은 3.006, 자유도는 3, p-value 는 0.3906이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

14.3.2 고용률? (%)

경제활동인구 / 15세이상 인구	취업자 / 경제활동인구	취업자 / 15세 이상 인구	실업자 / 경제활동인구	계
14.7	16.8	63.8	4.7	100.0

정답률은 Red, Black 을 합하여 계산하는데, 63.8(%) 입니다.

14.4 Q3. 취업자의 조건

3. 취업자의 조건에 해당하지 않는 것은 다음 중 어느 것인가? *

- 조사 대상 주간 중 수입을 목적으로 1시간 이상 일한 자
- 자기 가구에서 경영하는 농장이나 사업체의 수입을 높이는 데 도운 가족종사자로서 주당 18시간 이상 일한 자
- 조사대상 주간에 일시적으로 병이 나거나, 날씨가 안 좋거나, 휴가 또는 연기, 노동쟁의 등의 이유로 일을 못한 일시 휴직자
- 조사대상 주간에 수입 있는 일을 하지 않았으나, 지난 4주간 구직활동하였으며 즉시 취업 가능한자

14.4.1 취업자란?

조사 대상 주간 중 수입을 목적으로 1시간 이상 일한 자	자기 가구에서 경영하는 농장이나 사업체의 수입을 높이는 데 도운 가족종사자로서 주당 18시간 이상 일한 자		조사대상 주간에 일시적으로 병이 나거나, 날씨가 안 좋거나, 휴가 또는 연가, 노동쟁의 등의 이유로 일을 못한 일시 휴직자		조사대상 주간에 수입 있는 일을 하지 않았으나, 지난 4주간 구직활동하였으며 즉시 취업 가능한자	계
Red	18	34	36	199	287	
Black	26	25	46	188	285	
계	44	59	82	387	572	

표 14.9: Pearson's Chi-squared test: .

Test statistic	df	P value
4.353	3	0.2258

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

그 결과 카이제곱 통계량은 4.353, 자유도는 3, p-value 는 0.22580이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

14.4.2 취업자란? (%)

조사 대상 주간 중 수입을 목적으로 1시간 이상 일한 자	자기 가구에서 경영하는 농장이나 사업체의 수입을 높이는 데 도운 가족종사자로서 주당 18시간 이상 일한 자		조사대상 주간에 일시적으로 병이 나거나, 날씨가 안 좋거나, 휴가 또는 연가, 노동쟁의 등의 이유로 일을 못한 일시 휴직자		조사대상 주간에 수입 있는 일을 하지 않았으나, 지난 4주간 구직활동하였으며 즉시 취업 가능한자	계
	7.7	10.3	14.3	67.7	100.0	

정답률은 Red, Black 을 합하여 계산하는데, 67.7(%) 입니다.

14.5 Q4. 평균과 중위값

4. 5명이 다니는 회사에서 4명은 월 100만원을 받고, 한 명은 월 600만원을 받을 때, 월 급여의 평균과 중위값을 * 제대로 짹 지은 것은?

- 평균 : 100만원, 중위값 : 200만원
 - 평균 : 200만원, 중위값 : 100만원
 - 평균 : 1,000만원, 중위값 : 100만원
 - 평균 : 100만원, 중위값 : 1,000만원
-

14.5.1 집계

	평균 : 100만원, 중위값 : 200만원	평균 : 200만원, 중위값 : 100만원	평균 : 1,000만원, 중위값 : 100만원	평균 : 100만원, 중위값 : 1,000만원	계
Red	23	218	29	17	287
Black	26	214	31	14	285
계	49	432	60	31	572

표 14.12: Pearson's Chi-squared test: .

Test statistic	df	P value
0.5707	3	0.9031

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

그 결과 카이제곱 통계량은 0.571, 자유도는 3, p-value 는 0.9031이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

14.5.2 %

평균 : 100만원,	평균 : 200만원,	평균 : 1,000만원,	평균 : 100만원,
중위값 :	중위값 :	중위값 :	중위값 :
200만원	100만원	100만원	1,000만원
8.6	75.5	10.5	5.4
			계 100.0

정답률은 Red, Black 을 합하여 계산하는데, 75.5(%) 입니다.

14.6 Q5. 평균과 중위값

5. 4번 회사에서 4명의 월 급여는 그대로인데 다른 한 사람은 월 4,600만원을 받게 될 때, 월 급여의 평균과 중위 * 값을 제대로 짹 지은 것은?

- 평균 : 100만원, 중위값 : 200만원
- 평균 : 200만원, 중위값 : 100만원
- 평균 : 1,000만원, 중위값 : 100만원
- 평균 : 100만원, 중위값 : 1,000만원

14.6.1 집계

	평균 : 100만원, 중위값 : 200만원	평균 : 200만원, 중위값 : 100만원	평균 : 1,000만원, 중위값 : 100만원	평균 : 100만원, 중위값 : 1,000만원	계
Red	21	33	214	19	287
Black	14	35	214	22	285
계	35	68	428	41	572

표 14.15: Pearson's Chi-squared test: .

Test statistic	df	P value
1.671	3	0.6433

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

그 결과 카이제곱 통계량은 1.671, 자유도는 3, p-value 는 0.6433이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

14.6.2 %

평균 : 100만원,	평균 : 200만원,	평균 : 1,000만원,	평균 : 100만원,
중위값 :	중위값 :	중위값 :	중위값 :
200만원	100만원	100만원	1,000만원
6.1	11.9	74.8	7.2
			100.0

정답률은 Red, Black 을 합하여 계산하는데, 74.8(%) 입니다.

14.7 Q6. 연 500만원 이하 근로소득자의 비율

6. 2010년 기준 우리나라 근로소득자 중 연 500만 원 이하의 소득을 신고한 사람은 몇% 인가? *

<표2> 급여규모별 인원과 근로소득 총액 비중(연말정산 신고기준, 2009-10년, 단위:%)

급여규모별	2009년			2010년		
	인원	비중	누적비	인원	비중	누적비
전체						
5백만원 이하	18.6	18.6	1.8	19.1	19.1	1.7
5백-1천만원	13.0	31.6	3.9	12.3	31.4	3.6
1천-2천만원	23.1	54.6	13.3	19.0	22.8	12.8
2천-3천만원	14.7	69.3	14.3	33.3	14.4	68.7
3천-4천만원	10.1	79.4	13.9	47.1	9.8	78.5
4천-6천만원	12.0	91.4	23.3	70.4	12.0	90.5
6천-8천만원	5.3	96.7	14.3	84.7	5.4	95.9
8천만-1억원	1.9	98.6	6.6	91.3	2.3	98.2
1-2억원 이하	1.2	99.8	5.9	97.2	1.6	99.8
2-3억원 이하	0.1	99.9	0.9	98.1	0.1	99.9
3-5억원 이하	0.0	100.0	0.7	98.8	0.1	100.0
5-10억원 이하	0.0	100.0	0.6	99.4	0.0	100.0
10억원 초과	0.0	100.0	0.6	100.0	0.0	100.0
상위10%	10.0		32.9		10.0	
						33.9

자료: 국세청, 국제통계연보 2010-11년

- "5%"
- "10%"
- "15%"
- "19.1%"
- "25%"

14.7.1 집계

	5%	10%	15%	19.1%	25%	계
Red	19	34	35	192	7	287
Black	13	32	36	199	5	285
계	32	66	71	391	12	572

표 14.18: Pearson's Chi-squared test: .

Test statistic	df	P value
1.328	4	0.8565

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

그 결과 카이제곱 통계량은 1.328, 자유도는 4, p-value 는 0.8565이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

14.7.2 %

5%	10%	15%	19.1%	25%	계
5.6	11.5	12.4	68.4	2.1	100.0

정답률은 Red, Black 을 합하여 계산하는데, 68.4(%) 입니다.

##Q7. 디폴트 효과 : “… 하면 표시하세요”의 효과

표시하지 않으면 사후 장기기증 의사가 있는 것으로 간주하는 경우(Red)와 표시하면 사후 장기기증 의사가 있는 것으로 간주하는 경우(Black), 표시하지 않으면 사후 장기기증 의사가 있다고 보는 질문지(Red)에 사후 장기기증 의사 비율이 더 높게 나오는 것으로 알려져 있으나 관찰 결과는 반대로 나옵니다.

어디에 디폴트를 설정하느냐에 따라 조사결과가 달라지는 것으로 알려져 있으나 응답 결과는 랜덤화 효과에 가깝게 나와서 당황스럽습니다.

Red
<p>7. (Johnson & Goldstein 2003) 여러분이 새로 이사온 나라(덴마크, 네덜란드, 영국, 독일) *에서는 사후에 장기를 기증하는 것이 기본(디폴트, default)이고 원하지 않으면 직접 의사 표시를 하여야 합니다. 사후 장기 기증의사가 없으면 표시하세요. 사후 장기 기증의사가 있으면 빈 칸으로 놓아 두면 됩니다.</p> <p><input type="checkbox"/> 나는 사후 장기기증에 참여할 뜻이 없다</p>
Black
<p>7. (Johnson & Goldstein 2003) 여러분이 새로 이사온 나라(오스트리아, 벨기에, 프랑스, *헝가리, 폴란드, 포르투갈, 스웨덴)에서는 사후에 장기를 기증하지 않는 것이 기본(디폴트, default)이고 원하면 직접 의사 표시를 하여야 합니다. 사후 장기 기증의사가 있으면 표시하세요. 사후 장기 기증의사가 없으면 빈 칸으로 놓아 두면 됩니다.</p> <p><input type="checkbox"/> 나는 사후 장기기증에 참여할 뜻이 있다</p>

14.7.3 집계 (전체)

	사후 장기기증 의사 있음	사후 장기기증 의사 없음	계
Red(사후 장기기증 의사 없으면 표시)	126	161	287
Black(사후 장기기증 의사 있으면 표시)	141	144	285
계	267	305	572

표 14.21: Pearson's Chi-squared test with Yates' continuity correction: .

Test statistic	df	P value
1.566	1	0.2107

Q7의 Red 는 default 가 사후 장기기증 의사가 있는 것으로서 사후 장기기증 의사가 없으면 표시도록 하고, Black 은 default 가 사후 장기기증 의사가 없는 것으로서 사후 장기기증 의사가 있으면 표시도록 합니다.

표시하는 것조차 귀찮은 사람들은 당연히 default 를 따를 것으로 기대되어 Red 에서는 장기기증 의사가 있다는 응답 비율이 더 높고, Black 에서는 장기기증 의사가 없다는 응답 비율이 더 높을 것으로 기대되었지만 지난 학기들의 기록이나 이번 학기의 집계결과를 보면 그렇지 않아서 당황스럽습니다.

여기서 수행하고 있는 카이제곱 테스트는 default 를 어떻게 설정하느냐의 효과가 통계적으로 유의한지 알아보기 위한 것입니다.

그 결과 카이제곱 통계량은 1.57, 자유도는 1, p-value 는 0.21이므로 default의 설정을 바꾼 것은 사후 장기기증 의사에 통계적으로 유의한 차이를 보이지 않습니다.

14.7.4 % 비교

	사후 장기기증 의사 있음	사후 장기기증 의사 없음	계
Red(사후 장기기증 의사 없으면 표시)	43.9	56.1	100.0
Black(사후 장기기증 의사 있으면 표시)	49.5	50.5	100.0

이를 백분율로 살펴보면 사후 장기기증 의사가 없으면 표시하라는 Red에서 사후 장기기증 의사가 있다고 응답한 백분율, 43.9%(이)나 사후 장기기증 의사가 있으면 표시하라는 Black에서 사후 장기기증 의사가 있다고 응답한 백분율, 49.5%(이)나 큰 차이가 없어서 default의 설정이 응답에는 통계적으로 유의한 영향을 주지 못한다는 것을 명확히 알 수 있습니다.

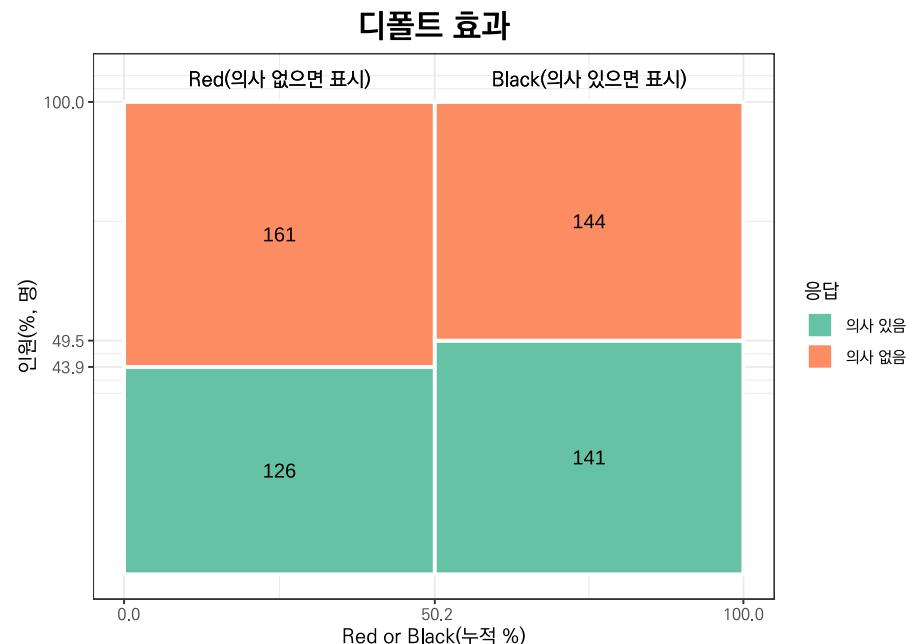
어쨌든 통계적으로 유의한 차이를 관찰하지 못했으므로 Red, Black 은 합쳐서 분석하는 것이 타당합니다. 그 결과는 다음과 같습니다.

14.7.5 합산(%)

사후 장기기증 의사 있음	사후 장기기증 의사 없음	계
46.7	53.3	100.0

사후 장기기증 의사가 있다고 응답한 백분율은 default의 설정과는 무관하게 46.7%(으)로 절반이 좀 안 되는군요.

14.7.6 Mosaic Plot



Mosaic Plot 은 이 집계결과를 시각적으로 잘 보여줍니다.

default 의 설정에 거의 무관할 정도로 응답이 비슷하기 때문에 Red, Black 이 닮았고, 사후 장기기증 의사가 있다고 응답한 백분율도 닮은 것을 잘 알 수 있습니다.

14.8 Thinking, Fast and Slow

14.8.1 Article

The evidence comes from a comparison of the rate of organ donation in European countries, which reveals startling differences between neighboring and culturally similar countries. An article published in 2003 noted that the rate of organ donation was close to 100% in Austria but only 12% in Germany, 86% in Sweden but only 4% in Denmark.

우리의 관찰은 Thinking, Fast and Slow (Daniel Kahneman, 2011) 의 default effects에서 소개한 유럽의 사례와는 아주 차이가 많이 납니다.

2003년 오스트리아는 100%, 스웨덴은 86%가 사후 장기기증 의사가 있다고 한 반면에 독일은 12%, 덴마크는 4% 밖에 사후 장기기증 의사가 있다는 조사 결과가 나왔습니다.

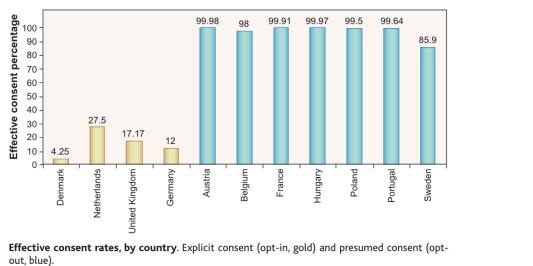
14.8.2 Facts Found

These enormous differences are a framing effect, which is caused by the format of the critical question. The high-donation countries have an opt-out form, where individuals who wish not to donate must check an appropriate box. Unless they take this simple action, they are considered willing donors. The low-contribution countries have an opt-in form: you must check a box to become a donor. That is all. The best single predictor of whether or not people will donate their organs is the designation of the default option that will be adopted without having to check a box.

여기서 결정적인 차이는 오스트리아와 스웨덴은 사후 장기기증 의사가 없으면 표시를 해야 하고, 독일과 덴마크는 사후 장기기증 의사가 있으면 표시를 해야 하는 시스템이었다는 점입니다.

여러분들의 응답과는 아주 다른 상황이죠?

14.8.3 Original Article (Johnson & Goldstein 2003)



Johnson and Goldstein의 원전에 나오는 그래프를 옮겨 왔습니다.

황금색으로 표시한 나라들은 사후 장기기증 의사를 직접 표시해야 하는 나라들이고 하늘색으로 표시한 나라들은 사후 장기기증 의사를 default로 하여 사후 장기기증 의사가 없으면 직접 표시해야 하는 나라들입니다.

여기서 오스트리아, 독일, 스웨덴, 덴마크의 수치를 확인해 보세요.

14.9 마감 시간으로부터 제출 시간의 분포

14.9.1 분포표

표 14.24: 일 단위

	14일	13일	12일	11일	10일	9일	8일	7일	6일	5일	4일	3일	2일	1일	계
Red	122	21	10	10	5	4	8	42	18	7	11	3	9	17	287
Black	128	20	11	3	5	6	8	33	14	8	7	12	15	15	285
계	250	41	21	13	10	10	16	75	32	15	18	15	24	32	572

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.

14.9.2 날마다 고르게 제출하는가?

14일	13일	12일	11일	10일	9일	8일	7일	6일	5일	4일	3일	2일	1일
250	41	21	13	10	10	16	75	32	15	18	15	24	32

표 14.26: Chi-squared test for given probabilities: .

Test statistic	df	P value
1246	13	2.469e-258 ***

날마다 고르게 제출하는지 알아 보았습니다.

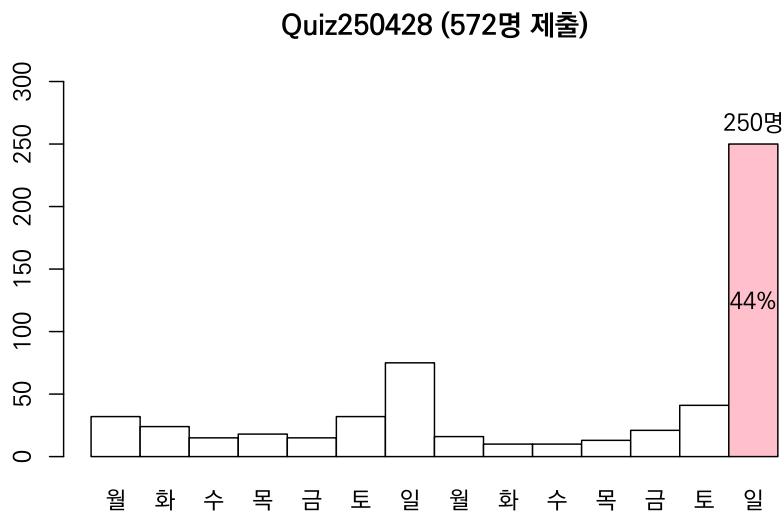
분포표의 “계”행에서 ‘계’열을 제외하고 카이제곱테스트를 수행합니다.

분포표 만으로도 쉽게 파악할 수 있지만 카이제곱테스트가 명확히 해 줍니다.

카이제곱 통계량은 1245.80, 자유도는 13.00, p-value 는 2.5e-258 이므로 날짜별로 고르게 제출하고 있지 않음을 시사합니다.

막대그래프로 살펴 보겠습니다.

14.9.3 막대그래프



14.9.4 Red, Black 간에 닮았는가?

	14일	13일	12일	11일	10일	9일	8일	7일	6일	5일	4일	3일	2일	1일
Red	122	21	10	10	5	4	8	42	18	7	11	3	9	17
Black	128	20	11	3	5	6	8	33	14	8	7	12	15	15

표 14.28: Pearson's Chi-squared test: .

Test statistic	df	P value
13.94	13	0.3782

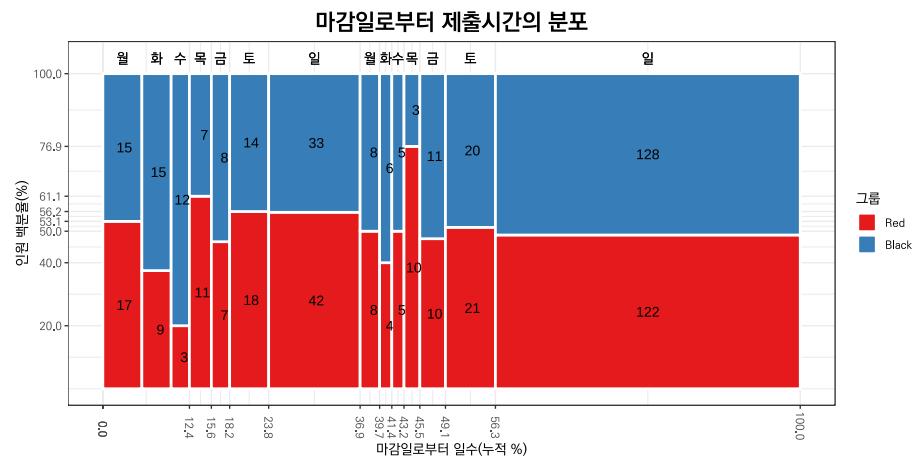
제출시간의 분포가 Red, Black 간에 닮았는지 알아 보았습니다.

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

카이제곱 통계량은 13.939, 자유도는 13, p-value 는 0.3782이므로 제출 시간의 분포는 Red, Black 간에 통계적으로 유의한 차이가 관찰되지 않습니다.

이 사실을 Mosaic Plot 을 이용하여 시각적으로 살펴보겠습니다. 닮았다고 느껴지나요?

14.9.5 Mosaic Plot



제 15 장 10주차 데이터 실험 집계

15.1 실험의 목적

10주차 구글 예습 설문지 집계결과를 분석합니다.

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

Q7에서는 Prospect Theory에서 이득은 확고히 하고 손실은 회피하려는 심리에 대해서 살펴봅니다.

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

15.1.1 Red, Black을 잘못 표시한 사람들

	Red(구글예습퀴즈)	Black(구글예습퀴즈)
Red(랜덤화출석부)	360	2
Black(랜덤화출석부)	0	358
계	360	360

15.1.2 응답인원의 Red, Black

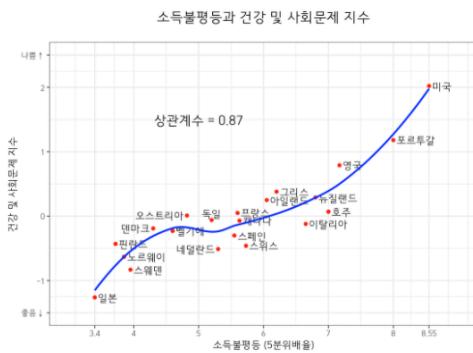
Red로 응답한 인원은 360명, Black에 응답한 인원은 360명입니다.

전체 응답인원 720명을 랜덤하게 둘로 나눌 때 어느 한 쪽의 기대인원은 전체 응답인원의 절반인 360명이고, 표준오차는 전체 응답인원의 제곱근에 $1/2$ 을 곱해 준 13.4명입니다.

따라서 Red, Black 각 그룹에 관찰된 인원은 기대인원으로부터 표준오차 범위 안에 들어갑니다. 랜덤화출석부에 있는 Red, Black과 실제 구글설문에 올린 Red, Black이 다른 사람들의 수효는 2명입니다. Red를 Black이라고 한 사람이 2명, Black을 Red라고 한 사람이 0명입니다.

15.2 Q1. Wilkinson

1. Richard Wilkinson 교수의 그래프는 무엇을 의미하는가? *



- 소득불평등이 심한 나라에서 건강 및 사회문제지수가 나쁘게 나온다.
- 소득이 많은 나라에서 건강 및 사회문제지수가 나쁘게 나온다.
- 소득불평등이 심하면 건강 및 사회문제지수가 나빠진다.
- 소득이 많아지면 건강 및 사회문제지수가 나빠진다.

15.2.1 소득불평등과 건강 및 사회문제 지표

	소득불평등이 심한 나라에서 건강 및 사회문제지수가 나쁘게 나온다.	소득이 많은 나라에서 건강 및 사회문제지수가 나쁘게 나온다.	소득불평등이 심하면 건강 및 사회문제지수가 나빠진다.	소득이 많아지면 건강 및 사회문제지수가 나빠진다.	계
Red	216	52	79	13	360
Black	209	48	87	16	360
계	425	100	166	29	720

표 15.3: Pearson's Chi-squared test: .

Test statistic	df	P value
0.9712	3	0.8082

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

그 결과 카이제곱 통계량은 0.971, 자유도는 3, p-value 는 0.80820이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

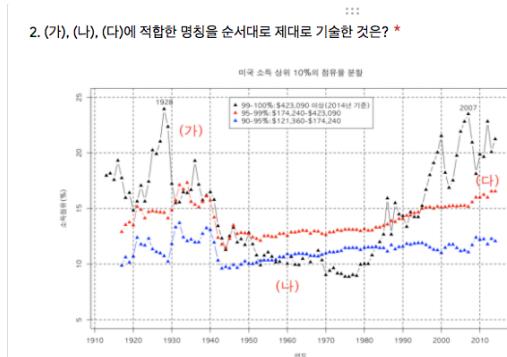
실제로 닮은 게 느껴집니까?

15.2.2 소득불평등과 건강 및 사회문제 지표 (%)

소득불평등이 심한 나라에서 건강 및 사회문제지수가 나쁘게 나온다.	소득이 많은 나라에서 건강 및 사회문제지수가 나쁘게 나온다.	소득불평등이 심하면 건강 및 사회문제지수가 나빠진다.	소득이 많아지면 건강 및 사회문제지수가 나빠진다.	계
59.0	13.9	23.1	4.0	100.0

정답률은 Red, Black 을 합하여 계산하는데, 59.0(%) 입니다.

15.3 Q2. 소득 상위1%의 몫과 시대적 사건들



- (가) 대공황, 대침체, 대번영
- (나) 대침체, 대공황, 대번영
- (다) 대번영, 대공황, 대침체
- (나) 대공황, 대번영, 대침체

15.3.1 대공황. 대번영, 대침체

	대공황, 대침체, 대번영	대침체, 대공황, 대번영	대번영, 대공황, 대침체	대공황, 대번영, 대침체	계
Red	51	43	64	202	360
Black	68	52	56	184	360
계	119	95	120	386	720

표 15.6: Pearson's Chi-squared test: .

Test statistic	df	P value
4.654	3	0.199

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

그 결과 카이제곱 통계량은 4.654, 자유도는 3, p-value 는 0.1990이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

15.3.2 대공황. 대번영, 대침체(%)

대공황, 대침체, 대번영	대침체, 대공황, 대번영	대번영, 대공황, 대침체	대공황, 대번영, 대침체	계
16.5	13.2	16.7	53.6	100.0

정답률은 Red, Black 을 합하여 계산하는데, 53.6(%) 입니다.

15.4 Q3. Piketty의 21세기 자본

3. Piketty 교수의 <21세기 자본>에 의하면 소득불평등이 2010년대 미국이나 1910년대 유럽 수준으로 심 *
할 경우 상위 10%의 총소득 점유율은 몇 %정도에 달하였는가?

- "25%"
- "35%"
- "50%"
- "60%"

15.4.1 상위10%의 총소득 점유율

	25%	35%	50%	60%	계
Red	47	76	163	74	360
Black	42	73	157	88	360
계	89	149	320	162	720

표 15.9: Pearson's Chi-squared test: .

Test statistic	df	P value
1.664	3	0.645

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

그 결과 카이제곱 통계량은 1.664, 자유도는 3, p-value 는 0.64500이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

15.4.2 상위10%의 총소득 점유율 (%)

25%	35%	50%	60%	계
12.4	20.7	44.4	22.5	100.0

정답률은 Red, Black 을 합하여 계산하는데, 44.4(%) 입니다.

15.5 Q4. World Top Income Database

4. World Top Income Database 에 의하면 2012년 우리나라 상위 10%의 소득점유율은 어느 정도에 이 * 르렀는가?
- "25%"
 - "35%"
 - "45%"
 - "60%"

15.5.1 우리나라 상위10%의 소득점유율

	25%	35%	45%	60%	계
Red	36	67	226	31	360
Black	24	78	226	32	360
계	60	145	452	63	720

표 15.12: Pearson's Chi-squared test: .

Test statistic	df	P value
3.25	3	0.3546

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

그 결과 카이제곱 통계량은 3.250, 자유도는 3, p-value 는 0.35460이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

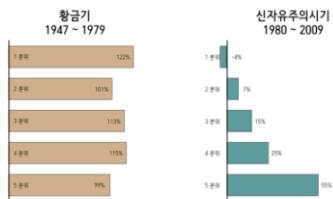
15.5.2 우리나라 상위10%의 소득점유율(%)

25%	35%	45%	60%	계
8.3	20.1	62.8	8.8	100.0

정답률은 Red, Black 을 합하여 계산하는데, 62.8(%) 입니다.

15.6 Q5. 황금기와 신자유주의 시기

5. 다음 도표로부터 알 수 있는 사실은? *



- 황금기에 소득이 가장 많이 늘어난 계층은 하위 20%(1분위)이다.
- 황금기에 소득증가율이 가장 높은 계층은 하위 20%(1분위)이다.
- 신자유주의시기에 소득이 가장 많이 늘어난 계층은 하위 20%(1분위)이다.
- 신자유주의시기에 소득증가율이 가장 높은 계층은 하위 20%(1분위)이다.

15.6.1 5분위별 소득증가율 비교

	황금기에 소득이 가장 많이 늘어난 계층은 하위 20%(1분위) 이다.	황금기에 소득증가율이 가장 높은 계층은 하위 20%(1분위) 이다.	신자유주의시기에 소득이 가장 많이 늘어난 계층은 하위 20%(1분위) 이다.	신자유주의시기에 소득증가율이 가장 높은 계층은 하위 20%(1분위) 이다.	계
Red	61	221	40	38	360
Black	62	217	52	29	360
계	123	438	92	67	720

표 15.15: Pearson's Chi-squared test: .

Test statistic	df	P value
2.819	3	0.4204

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

그 결과 카이제곱 통계량은 2.819, 자유도는 3, p-value 는 0.4204이므로 Red, Black 간에 통계적으로 유의한 차이를 보이고 있습니다.

실제로 닮은 게 느껴집니까?

15.6.2 5분위별 소득증가율 비교 (%)

황금기에 소득이 가장 많이 늘어난 계층은 하위 20%(1분위)이다.	황금기에 소득증가율이 가장 높은 계층은 하위 20%(1분위)이다.	신자유주의시기에 소득이 가장 많이 늘어난 계층은 하위 20%(1분위)이다.	신자유주의시기에 소득증가율이 가장 높은 계층은 하위 20%(1분위)이다.	계
17.1	60.8	12.8	9.3	100.0

정답률은 Red, Black 을 합하여 계산하는데, 60.8(%) 입니다.

15.7 Q6. 주역

6. 다음 표에 나오는 상황을 주역에 나오는 64괘 중에서 찾는다면 어떤 괘에 가장 가까운가?*

1993~2012 실질 소득 증가 분석(미국)

전체 기간 1993-2012	평균소득증가율 17.9%	상위1%의 소득증가율 86.1%	하위99%의 소득증가율 6.6%	증가(감소)에서 상위1%의비중 66%
확장기 1993-2000	31.5%	98.7%	20.3%	45%
2001경기침체 2000-2002	-11.7%	-30.8%	-6.5%	57%
부시膨창기 2002-2007	16.1%	61.8%	6.8%	65%
대침체기 2007-2009	-17.4%	-36.3%	-11.6%	49%
회복기 2009-2012	6.0%	31.4%	0.4%	95% 

지천태



천지비



풍뢰익



산택손



15.7.1 천지비

	지천태	천지비	풍뢰익	산택손	계
Red	56	210	51	43	360
Black	56	201	69	34	360
계	112	411	120	77	720

표 15.18: Pearson's Chi-squared test: .

Test statistic	df	P value
3.949	3	0.267

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

그 결과 카이제곱 통계량은 3.949, 자유도는 3, p-value 는 0.26700이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

15.7.2 천지비 (%)

지천태	천지비	풍뢰익	산택순	계
15.6	57.1	16.7	10.7	100.0

정답률은 Red, Black 을 합하여 계산하는데, 57.1(%) 입니다.

15.8 Q7. Prospect Theory : 이득과 손실

이득은 확고히 하고, 손실은 피하고자 하는 보통 사람들의 심리적 경향을 확인합니다.

기대값을 계산해 보면 어느 게임을 선택하더라도 상관이 없어야 하는 데 이득 게임인 Red에서는 조금 덜 받더라도 이득을 확고히 하려는 게임(A)에, 손실 게임인 Black에서는 손실을 확률적으로 줄이려는 게임(B)을 많이 선택하여 Kahneman과 Tversky의 Prospect Theory 에 부합하는 결과를 보여줍니다.

Red
7. (Thinking, Fast and Slow) 다음 중 어떤 게임을 선택하겠는가? *
<input type="radio"/> A : 100% 확률로 900불 따기
<input type="radio"/> B : 90% 확률로 1,000불 따기

Black

7. (Thinking, Fast and Slow) 다음 중 어떤 게임을 선택하겠는가? *

- A : 100% 확률로 900불 잊기
- B : 90% 확률로 1,000불 잊기

15.8.1 집계

	확률 100%	확률 90%	계
이득 게임	272	88	360
손실 게임	130	230	360
계	402	318	720

표 15.21: Pearson's Chi-squared test with Yates' continuity correction: .

Test statistic	df	P value
112	1	3.62e-26 ***

Q7의 Red 는 100%의 확률로 900달러를 따거나 90%의 확률로 1000달러를 따는 게임 중에서 많은 사람들이 900달러를 확고하게 갖는 것이 혹시라도 모를 10%의 확률로 아무 것도 안 생기는 것보다 선호한다는 것을 보여줍니다.

Black 은 100%의 확률로 900달러를 잊거나 90%의 확률로 1000달러를 잊는 게임 중에서는 손실을 확고히 하기 보다 10%의 확률이라도 1000달러의 손실을 회피하는 게임을 선호한다는 것을 보여줍니다.

이론적으로는 손실 게임이나 이득 게임이나 기대값이 같기 때문에 100% 게임이나 90% 게임을 비슷하게 선택할 것으로 생각되지만 우리의 실제 행동은 전혀 그렇지 않다는 것을 보여 줍니다.

여기서 수행하고 있는 카이제곱 테스트는 이득은 확고히 하고 손실은 운에 기대어 회피하려는 경향을 잘 보여 줍니다.

그 결과 카이제곱 통계량은 111.97, 자유도는 1, p-value 는 3.6e-26 합리적 행동을 한다는 가설로부터 어떤 선택도 비슷하게 나오리라는 기대와는 통계적으로 매우, 매우 유의하게 차이나는 결과를 보여줍니다.

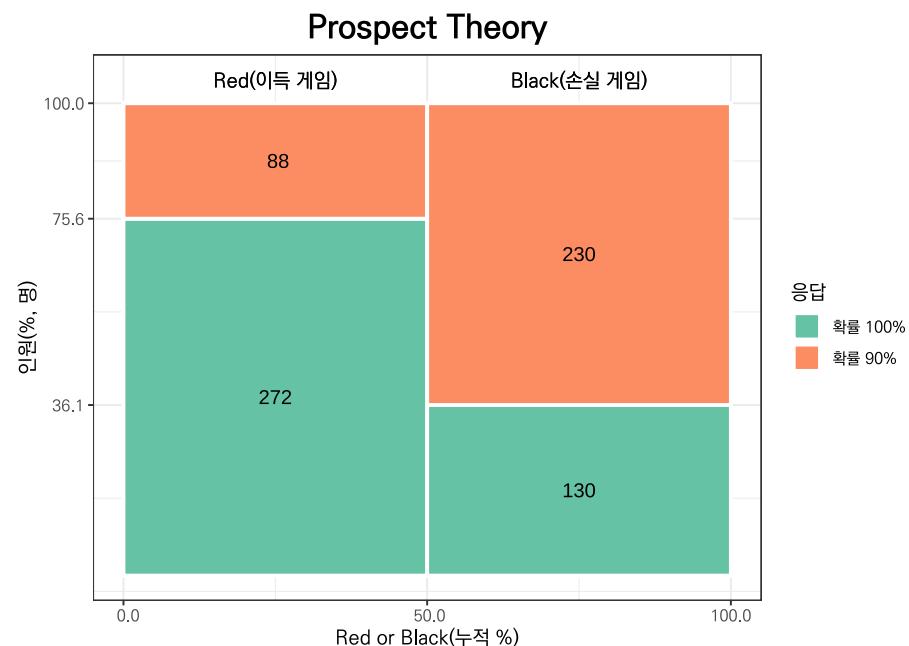
15.8.2 % 비교

	확률 100%	확률 90%	계
이득 게임	75.6	24.4	100.0
손실 게임	36.1	63.9	100.0

이를 백분율로 살펴보면 이득 게임인 Red에서 100% 확률로 900달러를 따려는 사람들의 백분율, 75.6(%)은 90% 확률로 1000달러를 따려는 사람들의 백분율, 24.4(%) 보다 월등히 높습니다. 반면 손실 게임인 Black에서 100% 확률로 900달러를 잃겠다는 사람들의 백분율, 36.1(%)은 90%의 확률로 1000달러를 잃겠다는 사람들의 백분율, 63.9(%) 보다 적습니다.

이득을 확고히 하려는 사람들의 백분율보다 손실을 운에 기대어 회피하려는 사람들의 백분율이 월등히 적은 것을 관찰할 수 있습니다.

15.8.3 Mosaic Plot



Mosaic Plot 은 이 집계결과를 시각적으로 잘 보여줍니다.

이득을 확고히 하려는 사람들의 백분율이 손실을 운에 기대어 회피하려는 사람들의 백분율보다 월등히 많은 것을 쉽게 파악할 수 있습니다.

15.9 마감 시간으로부터 제출 시간의 분포

15.9.1 분포표

표 15.23: 일 단위

	14일	13일	12일	11일	10일	9일	8일	7일	6일	5일	4일	3일	2일	1일	계
Red	122	29	6	5	13	9	9	36	22	20	17	13	25	34	360
Black	118	31	10	6	6	10	19	35	18	21	13	18	29	26	360
계	240	60	16	11	19	19	28	71	40	41	30	31	54	60	720

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.

15.9.2 날마다 고르게 제출하는가?

14일	13일	12일	11일	10일	9일	8일	7일	6일	5일	4일	3일	2일	1일
240	60	16	11	19	19	28	71	40	41	30	31	54	60

표 15.25: Chi-squared test for given probabilities: .

Test statistic	df	P value
831.3	13	2.701e-169 * * *

날마다 고르게 제출하는지 알아 보았습니다.

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

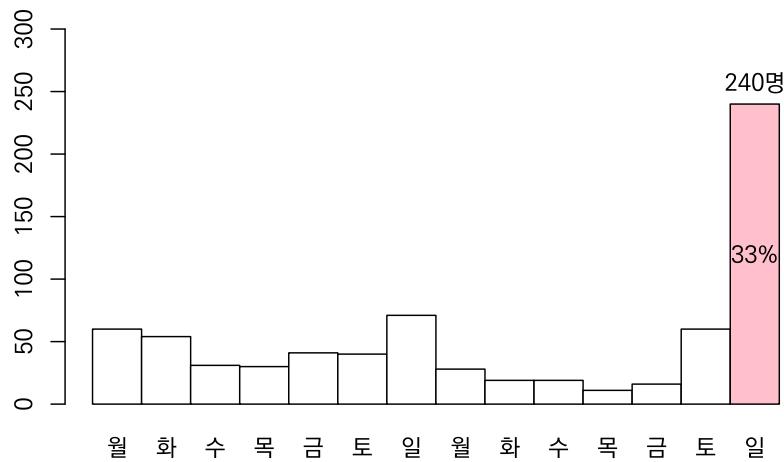
분포표 만으로도 쉽게 파악할 수 있지만 카이제곱테스트가 명확히 해 줍니다.

카이제곱 통계량은 831.317, 자유도는 13.00, p-value 는 2.7e-169 이므로 날짜별로 고르게 제출하지 않고 있음을 강력히 시사합니다.

막대그래프로 살펴 보겠습니다.

15.9.3 막대그래프

Quiz241104 (720명 제출)



15.9.4 Red, Black 간에 닮았는가?

	14일	13일	12일	11일	10일	9일	8일	7일	6일	5일	4일	3일	2일	1일
Red	122	29	6	5	13	9	9	36	22	20	17	13	25	34
Black	118	31	10	6	6	10	19	35	18	21	13	18	29	26

표 15.27: Pearson's Chi-squared test: .

Test statistic	df	P value
10.57	13	0.6469

제출시간의 분포가 Red, Black 간에 닮았는지 알아 보았습니다.

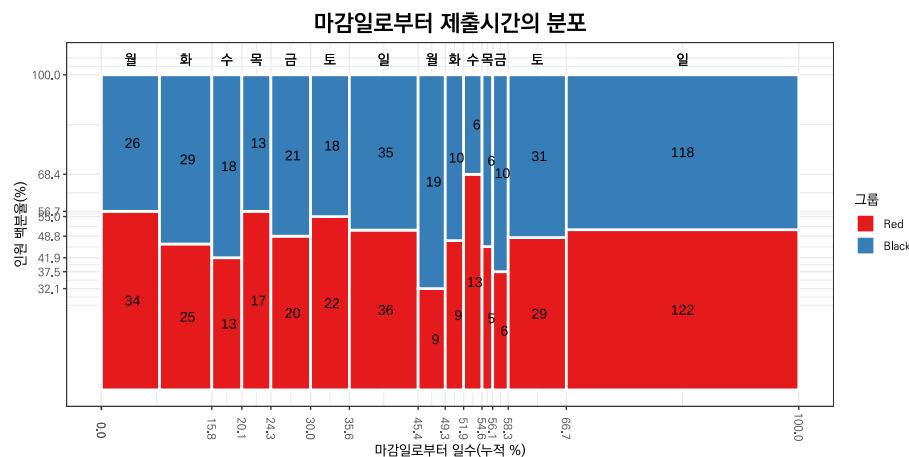
이번에는 분포표의 첫번째와 두번째 행, '계'열을 제외한 나머지 열에 대해서 카이제곱테스트를 수행합니다.

카이제곱 통계량은 10.57, 자유도는 13, p-value 는 0.6469 이므로 제출 시간의 분포는 Red, Black 간에 통계적으로 유의한 차이가 관찰되지 않습니다.

이 사실을 Mosaic Plot 을 이용하여 시각적으로 살펴보겠습니다.

답았다고 느껴지나요?

15.9.5 Mosaic Plot



제 16 장 10주차 데이터 실험 집계

16.1 실험의 목적

10주차 구글 예습 설문지 집계결과를 분석합니다.

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

Q7에서는 Prospect Theory에서 이득은 확고히 하고 손실은 회피하려는 심리에 대해서 살펴봅니다.

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

16.1.1 Red, Black을 잘못 표시한 사람들

	Red(구글예습퀴즈)	Black(구글예습퀴즈)
Red(랜덤화출석부)	280	2
Black(랜덤화출석부)	1	278
계	281	280

16.1.2 응답인원의 Red, Black

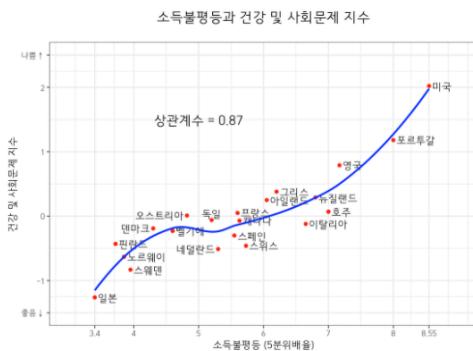
Red로 응답한 인원은 281명, Black에 응답한 인원은 280명입니다.

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

따라서 Red, Black 각 그룹에 관찰된 인원은 기대인원으로부터 표준오차 범위 안에 들어갑니다. 랜덤화출석부에 있는 Red, Black과 실제 구글설문에 올린 Red, Black이 다른 사람들의 수효는 3명입니다. Red를 Black이라고 한 사람이 2명, Black을 Red라고 한 사람이 1명입니다.

16.2 Q1. Wilkinson

1. Richard Wilkinson 교수의 그래프는 무엇을 의미하는가? *



- 소득불평등이 심한 나라에서 건강 및 사회문제지수가 나쁘게 나온다.
- 소득이 많은 나라에서 건강 및 사회문제지수가 나쁘게 나온다.
- 소득불평등이 심하면 건강 및 사회문제지수가 나빠진다.
- 소득이 많아지면 건강 및 사회문제지수가 나빠진다.

16.2.1 소득불평등과 건강 및 사회문제 지표

	소득불평등이 심한 나라에서 건강 및 사회문제지수가 나쁘게 나온다.	소득이 많은 나라에서 건강 및 사회문제지수가 나쁘게 나온다.	소득불평등이 심하면 건강 및 사회문제지수가 나빠진다.	소득이 많아지면 건강 및 사회문제지수가 나빠진다.	계
Red	162	37	75	7	281
Black	180	29	62	9	280
계	342	66	137	16	561

표 16.3: Pearson's Chi-squared test: .

Test statistic	df	P value
3.134	2	0.2086

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

그 결과 카이제곱 통계량은 3.134, 자유도는 2, p-value 는 0.20860이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

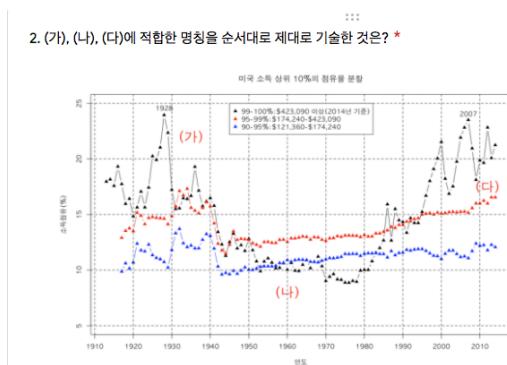
실제로 닮은 게 느껴집니까?

16.2.2 소득불평등과 건강 및 사회문제 지표 (%)

소득불평등이 심한 나라에서 건강 및 사회문제지수가 나쁘게 나온다.	소득이 많은 나라에서 건강 및 사회문제지수가 나쁘게 나온다.	소득불평등이 심하면 건강 및 사회문제지수가 나빠진다.	소득이 많아지면 건강 및 사회문제지수가 나빠진다.	계
61.0	11.8	24.4	2.9	100.0

정답률은 Red, Black 을 합하여 계산하는데, 61.0(%) 입니다.

16.3 Q2. 소득 상위1%의 몫과 시대적 사건들



- 대공황, 대침체, 대번영
- 대침체, 대공황, 대번영
- 대번영, 대공황, 대침체
- 대공황, 대번영, 대침체

16.3.1 대공황, 대번영, 대침체

	대공황, 대침체, 대번영	대침체, 대공황, 대번영	대번영, 대공황, 대침체	대공황, 대번영, 대침체	계
Red	44	40	34	163	281
Black	58	23	37	162	280
계	102	63	71	325	561

표 16.6: Pearson's Chi-squared test: .

Test statistic	df	P value
6.637	3	0.08442

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

그 결과 카이제곱 통계량은 6.637, 자유도는 3, p-value 는 0.08440이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

16.3.2 대공황, 대번영, 대침체(%)

대공황, 대침체, 대번영	대침체, 대공황, 대번영	대번영, 대공황, 대침체	대공황, 대번영, 대침체	계
18.2	11.2	12.7	57.9	100.0

정답률은 Red, Black 을 합하여 계산하는데, 57.9(%) 입니다.

16.4 Q3. Piketty의 21세기 자본

3. Piketty 교수의 <21세기 자본>에 의하면 소득불평등이 2010년대 미국이나 1910년대 유럽 수준으로 심 *
할 경우 상위 10%의 총소득 점유율은 몇 %정도에 달하였는가?

- "25%"
- "35%"
- "50%"
- "60%"

16.4.1 상위10%의 총소득 점유율

	25%	35%	50%	60%	계
Red	28	63	131	59	281
Black	35	66	127	52	280
계	63	129	258	111	561

표 16.9: Pearson's Chi-squared test: .

Test statistic	df	P value
1.349	3	0.7175

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

그 결과 카이제곱 통계량은 1.349, 자유도는 3, p-value 는 0.71750이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

16.4.2 상위10%의 총소득 점유율 (%)

25%	35%	50%	60%	계
11.2	23.0	46.0	19.8	100.0

정답률은 Red, Black 을 합하여 계산하는데, 46.0(%) 입니다.

16.5 Q4. World Top Income Database *

4. World Top Income Database 에 의하면 2012년 우리나라 상위 10%의 소득점유율은 어느 정도에 이 *
르렀는가?
- "25%"
 - "35%"
 - "45%"
 - "60%"

16.5.1 우리나라 상위10%의 소득점유율

	25%	35%	45%	60%	계
Red	26	46	181	28	281
Black	34	45	181	20	280
계	60	91	362	48	561

표 16.12: Pearson's Chi-squared test: .

Test statistic	df	P value
2.409	3	0.4919

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

그 결과 카이제곱 통계량은 2.409, 자유도는 3, p-value 는 0.49190이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

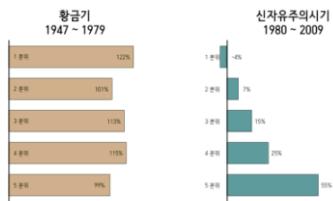
16.5.2 우리나라 상위10%의 소득점유율(%)

25%	35%	45%	60%	계
10.7	16.2	64.5	8.6	100.0

정답률은 Red, Black 을 합하여 계산하는데, 64.5(%) 입니다.

16.6 Q5. 황금기와 신자유주의 시기

5. 다음 도표로부터 알 수 있는 사실은? *



- 황금기에 소득이 가장 많이 늘어난 계층은 하위 20%(1분위)이다.
- 황금기에 소득증가율이 가장 높은 계층은 하위 20%(1분위)이다.
- 신자유주의시기에 소득이 가장 많이 늘어난 계층은 하위 20%(1분위)이다.
- 신자유주의시기에 소득증가율이 가장 높은 계층은 하위 20%(1분위)이다.

16.6.1 5분위별 소득증가율 비교

	황금기에 소득이 가장 많이 늘어난 계층은 하위 20%(1분위) 이다.	황금기에 소득증가율이 가장 높은 계층은 하위 20%(1분위) 이다.	신자유주의시기에 소득이 가장 많이 늘어난 계층은 하위 20%(1분위) 이다.	신자유주의시기에 소득증가율이 가장 높은 계층은 하위 20%(1분위) 이다.	계
Red	41	182	40	18	281
Black	48	170	31	31	280
계	89	352	71	49	561

표 16.15: Pearson's Chi-squared test: .

Test statistic	df	P value
5.548	3	0.1358

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

그 결과 카이제곱 통계량은 5.548, 자유도는 3, p-value 는 0.13580이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

16.6.2 5분위별 소득증가율 비교 (%)

황금기에 소득이 가장 많이 늘어난 계층은 하위 20%(1분위)이다.	황금기에 소득증가율이 가장 높은 계층은 하위 20%(1분위)이다.	신자유주의시기에 소득이 가장 많이 늘어난 계층은 하위 20%(1분위)이다.	신자유주의시기에 소득증가율이 가장 높은 계층은 하위 20%(1분위)이다.	계
15.9	62.7	12.7	8.7	100.0

정답률은 Red, Black 을 합하여 계산하는데, 62.7(%) 입니다.

16.7 Q6. 주역

6. 다음 표에 나오는 상황을 주역에 나오는 64괘 중에서 찾는다면 어떤 괘에 가장 가까운가?*

1993~2012 실질 소득 증가 분석(미국)

전체 기간 1993-2012	평균소득증가율 17.9%	상위1%의 소득증가율 86.1%	하위99%의 소득증가율 6.6%	증가(감소)에서 상위1%의비중 66%
확장기 1993-2000	31.5%	98.7%	20.3%	45%
2001경기 침체 2000-2002	-11.7%	-30.8%	-4.5%	57%
부시 팸胀기 2002-2007	16.1%	61.8%	6.8%	65%
대침체기 2007-2009	-17.4%	-36.3%	-11.6%	49%
회복기 2009-2012	6.0%	31.4%	0.4%	95% 

지천태



천지비



풍뢰익



산택손



16.7.1 천지비

	지천태	천지비	풍뢰익	산택손	계
Red	41	170	39	31	281
Black	38	163	46	33	280
계	79	333	85	64	561

표 16.18: Pearson's Chi-squared test: .

Test statistic	df	P value
0.8983	3	0.8258

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

그 결과 카이제곱 통계량은 0.898, 자유도는 3, p-value 는 0.82580이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

16.7.2 천지비 (%)

지천태	천지비	풍뢰익	산택순	계
14.1	59.4	15.2	11.4	100.0

정답률은 Red, Black 을 합하여 계산하는데, 59.4(%) 입니다.

16.8 Q7. Prospect Theory : 이득과 손실

이득은 확고히 하고, 손실은 피하고자 하는 보통 사람들의 심리적 경향을 확인합니다.

기대값을 계산해 보면 어느 게임을 선택하더라도 상관이 없어야 하는 데 이득 게임인 Red에서는 조금 덜 받더라도 이득을 확고히 하려는 게임(A)에, 손실 게임인 Black에서는 손실을 확률적으로 줄이려는 게임(B)을 많이 선택하여 Kahneman과 Tversky의 Prospect Theory 에 부합하는 결과를 보여줍니다.

Red
7. (Thinking, Fast and Slow) 다음 중 어떤 게임을 선택하겠는가? * <input type="radio"/> A : 100% 확률로 900불 따기 <input type="radio"/> B : 90% 확률로 1,000불 따기

Black

7. (Thinking, Fast and Slow) 다음 중 어떤 게임을 선택하겠는가? *

- A : 100% 확률로 900불 잊기
- B : 90% 확률로 1,000불 잊기

16.8.1 집계

	확률 100%	확률 90%	계
이득 게임	205	76	281
손실 게임	84	196	280
계	289	272	561

표 16.21: Pearson's Chi-squared test with Yates' continuity correction: .

Test statistic	df	P value
101.9	1	5.875e-24 ***

Q7의 Red 는 100%의 확률로 900달러를 따거나 90%의 확률로 1000달러를 따는 게임 중에서 많은 사람들이 900달러를 확고하게 갖는 것이 혹시라도 모를 10%의 확률로 아무 것도 안 생기는 것보다 선호한다는 것을 보여줍니다.

Black 은 100%의 확률로 900달러를 잊거나 90%의 확률로 1000달러를 잊는 게임 중에서는 손실을 확고히 하기 보다 10%의 확률이라도 1000달러의 손실을 회피하는 게임을 선호한다는 것을 보여줍니다.

이론적으로는 손실 게임이나 이득 게임이나 기대값이 같기 때문에 100% 게임이나 90% 게임을 비슷하게 선택할 것으로 생각되지만 우리의 실제 행동은 전혀 그렇지 않다는 것을 보여 줍니다.

여기서 수행하고 있는 카이제곱 테스트는 이득은 확고히 하고 손실은 운에 기대어 회피하려는 경향을 잘 보여 줍니다.

그 결과 카이제곱 통계량은 101.89, 자유도는 1, p-value 는 5.9e-24 합리적 행동을 한다는 가설로부터 어떤 선택도 비슷하게 나오리라는 기대와는 통계적으로 매우, 매우 유의하게 차이나는 결과를 보여줍니다.

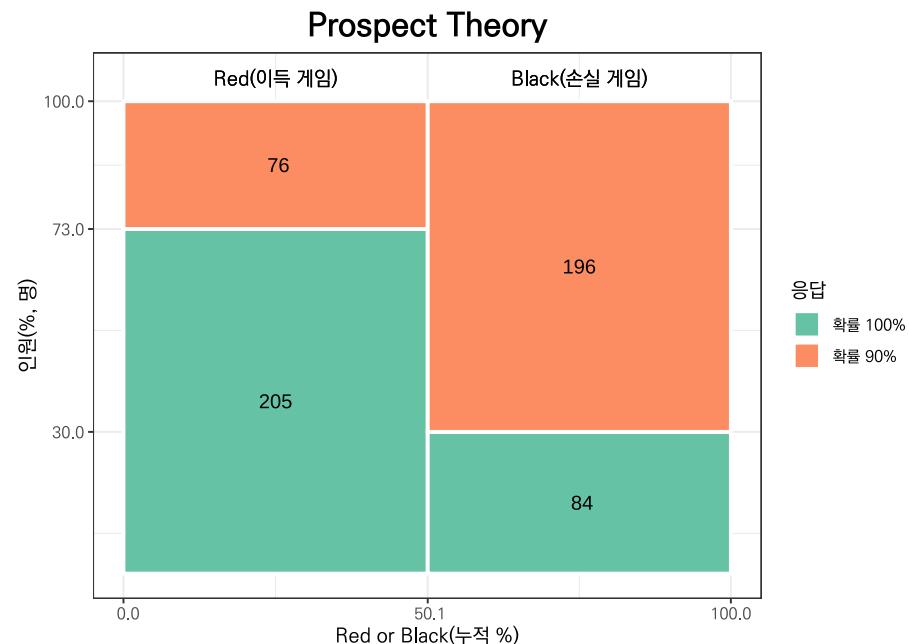
16.8.2 % 비교

	확률 100%	확률 90%	계
이득 게임	73.0	27.0	100.0
손실 게임	30.0	70.0	100.0

이를 백분율로 살펴보면 이득 게임인 Red에서 100% 확률로 900달러를 따려는 사람들의 백분율, 73.0(%)은 90% 확률로 1000달러를 따려는 사람들의 백분율, 27.0(%) 보다 월등히 높습니다. 반면 손실 게임인 Black에서 100% 확률로 900달러를 잃겠다는 사람들의 백분율, 30.0(%)은 90%의 확률로 1000달러를 잃겠다는 사람들의 백분율, 70.0(%) 보다 적습니다.

이득을 확고히 하려는 사람들의 백분율보다 손실을 운에 기대어 회피하려는 사람들의 백분율이 월등히 적은 것을 관찰할 수 있습니다.

16.8.3 Mosaic Plot



Mosaic Plot 은 이 집계결과를 시각적으로 잘 보여줍니다.

이득을 확고히 하려는 사람들의 백분율이 손실을 운에 기대어 회피하려는 사람들의 백분율보다 월등히 많은 것을 쉽게 파악할 수 있습니다.

16.9 마감 시간으로부터 제출 시간의 분포

16.9.1 분포표

표 16.23: 일 단위

	14일	13일	12일	11일	10일	9일	8일	7일	6일	5일	4일	3일	2일	1일	계
Red	118	23	11	6	1	7	8	23	18	14	16	14	8	14	281
Black	136	20	13	5	3	5	6	26	11	10	10	15	9	11	280
계	254	43	24	11	4	12	14	49	29	24	26	29	17	25	561

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.

16.9.2 날마다 고르게 제출하는가?

14일	13일	12일	11일	10일	9일	8일	7일	6일	5일	4일	3일	2일	1일
254	43	24	11	4	12	14	49	29	24	26	29	17	25

표 16.25: Chi-squared test for given probabilities: .

Test statistic	df	P value
1277	13	3.904e-265 * * *

날마다 고르게 제출하는지 알아 보았습니다.

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

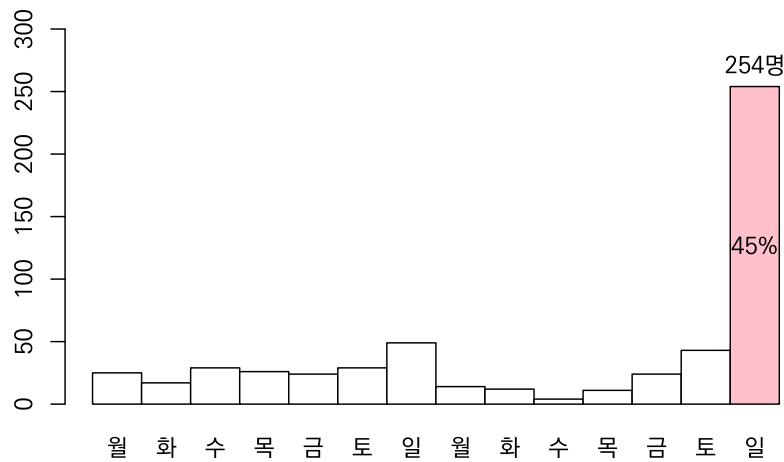
분포표 만으로도 쉽게 파악할 수 있지만 카이제곱테스트가 명확히 해 줍니다.

카이제곱 통계량은 1277.392, 자유도는 13.00, p-value 는 3.9e-265 이므로 날짜별로 고르게 제출하지 않고 있음을 강력히 시사합니다.

막대그래프로 살펴 보겠습니다.

16.9.3 막대그래프

Quiz250505 (561명 제출)



16.9.4 Red, Black 간에 닮았는가?

	14일	13일	12일	11일	10일	9일	8일	7일	6일	5일	4일	3일	2일	1일
Red	118	23	11	6	1	7	8	23	18	14	16	14	8	14
Black	136	20	13	5	3	5	6	26	11	10	10	15	9	11

표 16.27: Pearson's Chi-squared test: .

Test statistic	df	P value
7.738	13	0.8603

제출시간의 분포가 Red, Black 간에 닮았는지 알아 보았습니다.

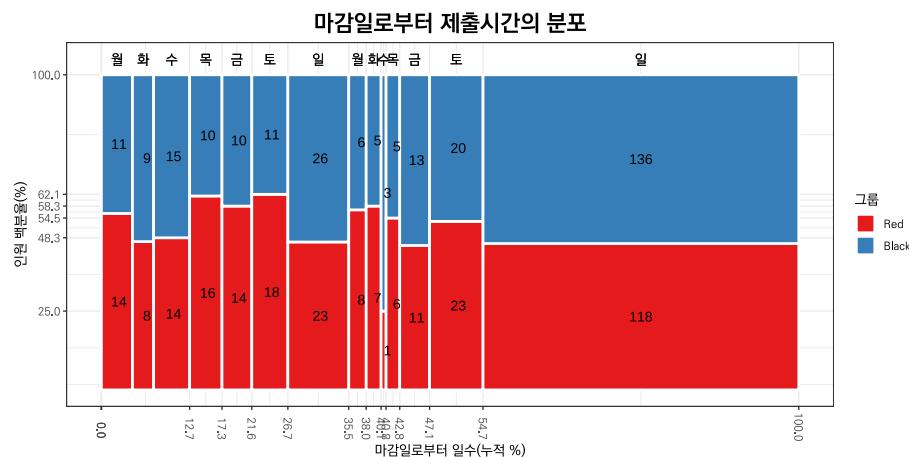
이번에는 분포표의 첫번째와 두번째 행, '계'열을 제외한 나머지 열에 대해서 카이제곱테스트를 수행합니다.

카이제곱 통계량은 7.74, 자유도는 13, p-value 는 0.8603 이므로 제출 시간의 분포는 Red, Black 간에 통계적으로 유의한 차이가 관찰되지 않습니다.

이 사실을 Mosaic Plot 을 이용하여 시각적으로 살펴보겠습니다.

답았다고 느껴지나요?

16.9.5 Mosaic Plot



제 17 장 식생활습관 교육동영상 시청 효과

17.1 실험의 목적

국민문해력조사, 옥스포드 행복도 조사에 이어 여러분의 식생활 습관에 대하여 알아보고 있습니다.

식생활습관 현황을 묻는 질문 후에 식생활 개선 관련 동영상을 시청하면 앞으로 식생활습관을 바꾸겠다고 생각할지 묻는 실험입니다.

즉, 동영상 시청의 효과를 파악하는 실험인데 동일한 대상에게 시청 전과 시청 후의 인식 차이를 물어 보는 것입니다.

집계 결과를 여러 가지 통계로 제시하였습니다.

17.2 동영상 시청 전후 비교

교차표를 빈도수와 백분율로 비교하고, 개별 수강생의 점수 차이에 대하여 기초 통계량과 t-test를 수행하였습니다.

17.2.1 Cross Tables

	식생활을 고치세요	식생활이 보통입니다	식생활이 대체로 좋아요	불참	계
시청 전	396	292	40	72	800
시청 후	82	380	262	76	800

전체 수강생 800(명) 중에 72(명)이 불참하여 총 728(명)이 식생활습관 현황 조사에 참여한 가운데 “식생활이 대체로 좋아요”인 수강생은 40(명)에 불과하였고, “식생활을

“고치세요”인 수강생은 무려 396(명)에 달했었는데, 영국의 식생활개선 활동가 제이미 올리버의 학교 급식 동영상 시청 후 “식생활이 대체로 좋아요”로 개선하겠다는 수강생은 262(명)으로 급격히 늘어났고, 여전히 “식생활을 고치세요” 수준에 머무르게 되는 수강생은 82(명)으로 대폭 줄어듭니다.

17.2.2 Percentages

	식생활을 고치세요	식생활이 보통입니다	식생활이 대체로 좋아요	계
시청 전	54.4	40.1	5.5	100
시청 후	11.3	52.5	36.2	100

이를 백분율로 살펴보면 동영상 시청 전에 “식생활이 대체로 좋아요”인 수강생은 5.5%(%)에 불과하였고, “식생활을 고치세요”인 수강생은 무려 54.4%(%)에 달했었는데, 동영상 시청 후에는 “식생활이 대체로 좋아요”로 개선하겠다는 수강생은 36.2%(%)으로 급격히 늘어났고, 여전히 “식생활을 고치세요” 수준에 머무르게 되는 수강생은 11.3%(%)으로 대폭 줄어듭니다.

17.2.3 Summary Statistics for diff

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	NA's
-22	8	16	16.67	26	74	76

수강생 각자 동영상 시청 전 식생활습관 현황조사로부터 점수를 계산하였고, 시청 후에 식생활습관 계획조사로부터 습관을 개선했을 때의 점수를 계산하였기 때문에 그 차이를 개선점수라고 할 수 있습니다.

만약 개선점수가マイナス라면 동영상 시청의 효과가 없거나 오히려 부정적이라는 뜻이죠.

여기 최소값으로 나와 있는 -22.00(은)는 그 경우를 의미합니다.

차이가 플러스일 수록 동영상 시청의 효과를 입증하는 것이죠.

그 차이가 어느 정도 인지를 살펴보기 위하여 먼저 기초통계를 제시하였습니다.

여기서 개별 개선점수의 산술평균은 16.67로 상당히 큰 값입니다.

개별 개선점수의 중위수는 16.00로 상당히 많이 개선된 것으로 볼 수 있습니다.

얼마나 개선되었는지를 파악하려면 표준편차를 알아야 하고 이를 자유도의 제곱근으로 나눠 준 표준오차로 나눠 준 t-통계량이 나오게 됩니다.

17.2.4 Standard Deviation of **diff**

$$\begin{array}{c} \hline \text{SD}_\text{diff} \\ \hline 13.49 \end{array}$$

개선점수의 표준편차는 13.49입니다.

따라서 개선점수의 산술평균은 차이가 없다고 가정할 때 보다 약 1.24 배수만큼 큰 개선효과를 의미합니다.

17.2.5 t-test for **diff** (One Sample)

표 17.5: One Sample t-test: **diff**

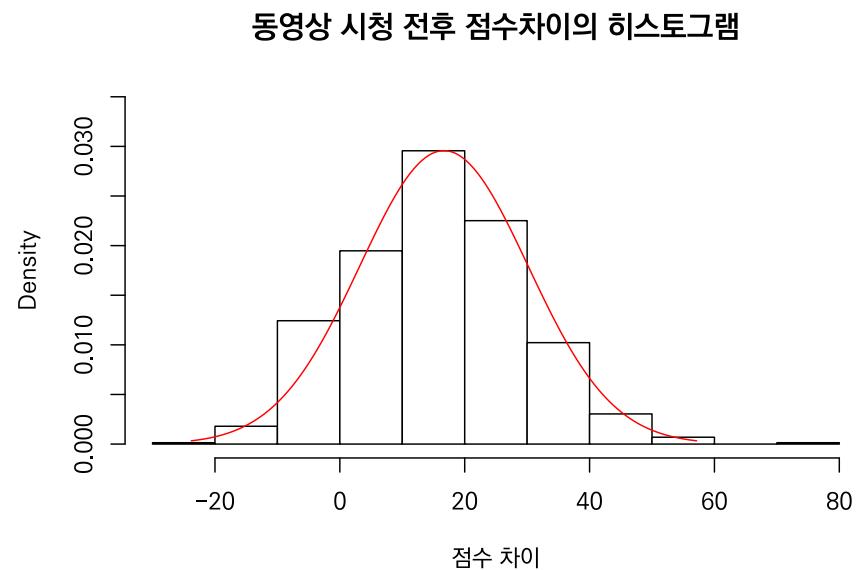
Test statistic	df	P value	Alternative hypothesis	mean of x
33.26	723	3.695e-148 * *	greater	16.67

동영상 시청의 효과가 없다면 개선점수가 0이거나 평균적으로マイナス가 나올 것입니다.

이를 통계적으로 검증하기 위한 t-통계값은 33.26이고 자유도는 시청 전후의 식생활습관 점수가 모두 나와 있는 724(명)에서 한 명을 뺀 723이어서 사실상 정규분포라고 보아도 무방합니다.

따라서 p-value 는 3.7e-148(0이)라는 엄청나게 작은 값이 나오고 따라서 동영상시청의 효과는 통계적으로 매우, 매우 유의한 것으로 결론내릴 수 있습니다.

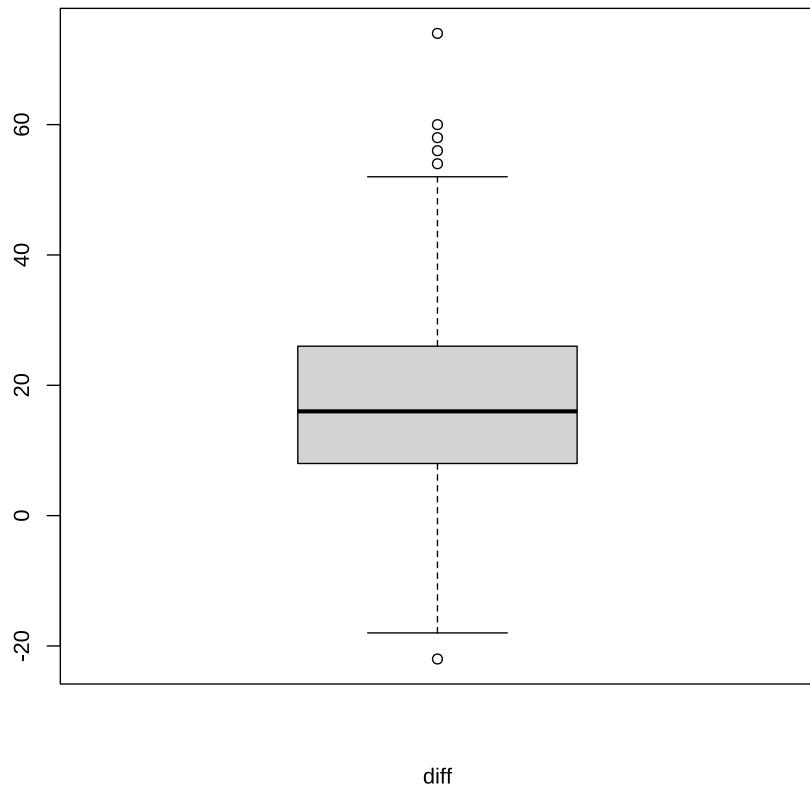
17.2.6 Histogram for `diff`



개선점수의 분포를 히스토그램으로 나타내었습니다.

그리고 붉은색 선은 개선점수의 산술평균을 평균, 표준편차를 개선점수의 표준편차로 갖는 정규분포의 확률밀도함수입니다.

자유도 723인 t-분포는 사실상 정규분포와 같기 때문입니다.

17.2.7 Box plot for **diff**

개선점수의 Box Plot입니다.

중위수가 16.00인데, 20보다 작다는 것과 제1사분위수가 0과 20의 중간쯤으로 실제는 8.00, 제3사분위수가 20을 좀 넘는 26.00임을 대략 알 수 있습니다.

물론, 최소값이 0보다 작고 최대값은 40을 넘는다는 것도 시각적으로 파악 가능합니다.

17.3 Red and Black (랜덤화 효과)

비록 Red, Black 을 직접 적어 넣으라고 하지는 않았어도 기존에 부여받은 그룹이 있기 때문에 랜덤화 효과를 파악할 수 있습니다.

Red, Black 별로 기초통계값을 비교하고 t-test, 그룹별 히스토그램과 Box Plot을
그려서 시각적으로 비교합니다.

17.3.1 Summary by Red and Black

- Red:

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	NA's
-22	8	16	16.86	26	60	38

- Black:

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	NA's
-18	6	16	16.49	24	74	38

Red 그룹의 개선점수 평균은 16.86, Black 그룹의 개선점수 평균은 16.49로 크게 차이가 나지 않습니다.

Red 그룹의 개선점수 중위수는 16.00, Black 그룹의 개선점수 중위수는 16.00로 크게 차이가 나지 않습니다.

다른 기초통계값들은 각자 비교해 보세요.

17.3.2 Standard Deviation by Red and Black

Red	Black
13.12	13.86

Red 그룹 개선점수의 표준편차는 13.12이고, Black 그룹 개선점수의 표준편차는 13.86이어서 상당히 달았습니다.

17.3.3 t-test (Welch's) Red vs Black (Two Sample)

표 17.9: Welch Two Sample t-test: diff by group

Test statistic	df	P value	Alternative hypothesis	mean in group Red	mean in group Black
0.3672	720.2	0.7136	two.sided	16.86	16.49

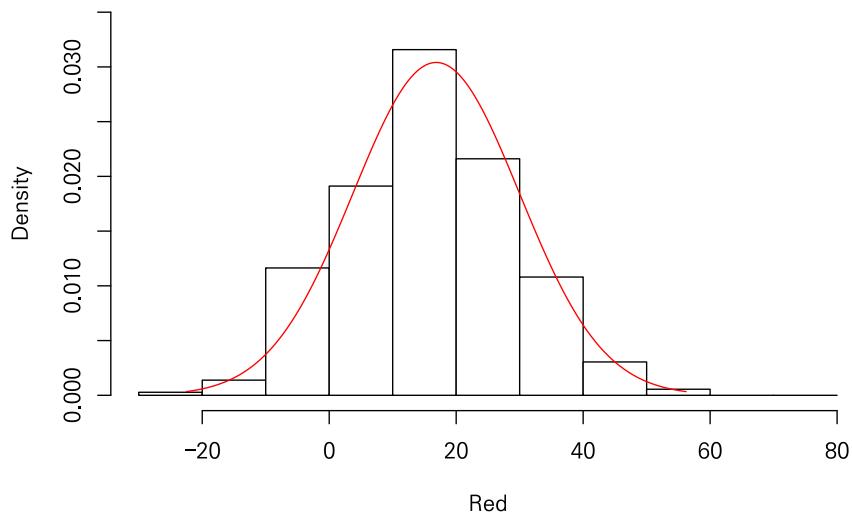
Red 그룹 개선점수 모집단의 표준편차와 Black 그룹 개선점수 모집단의 표준편차를 알 수 없기 때문에 Welch의 근사식을 적용한 t-test (Two Sample) (을)를 수행하였습니다.

근사식으로 계산한 자유도는 720.23, t-통계값은 0.36720이고 따라서 p-value 는 0.71360이 나오므로 Red, Black 개선점수 간의 차이는 통계적으로 유의하지 않고 닮았다고 결론내릴 수 있습니다.

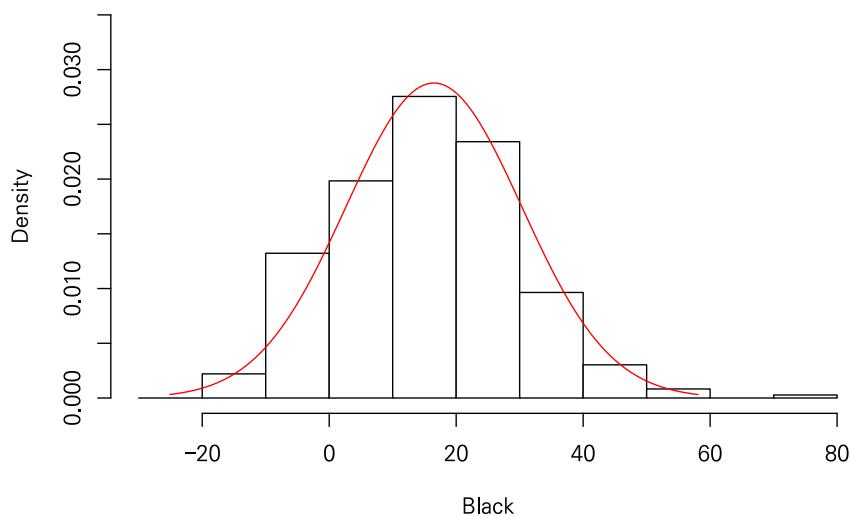
이 비교는 서로 다른지를 묻고 있기 때문에 two.sided, 즉 양측 검증이라는 것을 알 수 있고 Red 그룹 개선점수의 평균, 16.86(와)과 Black 그룹 개선점수의 평균, 16.49 도 요약표에서 확인할 수 있습니다.

17.3.4 Histogram by Red and Black

Red 그룹 개선점수의 히스토그램



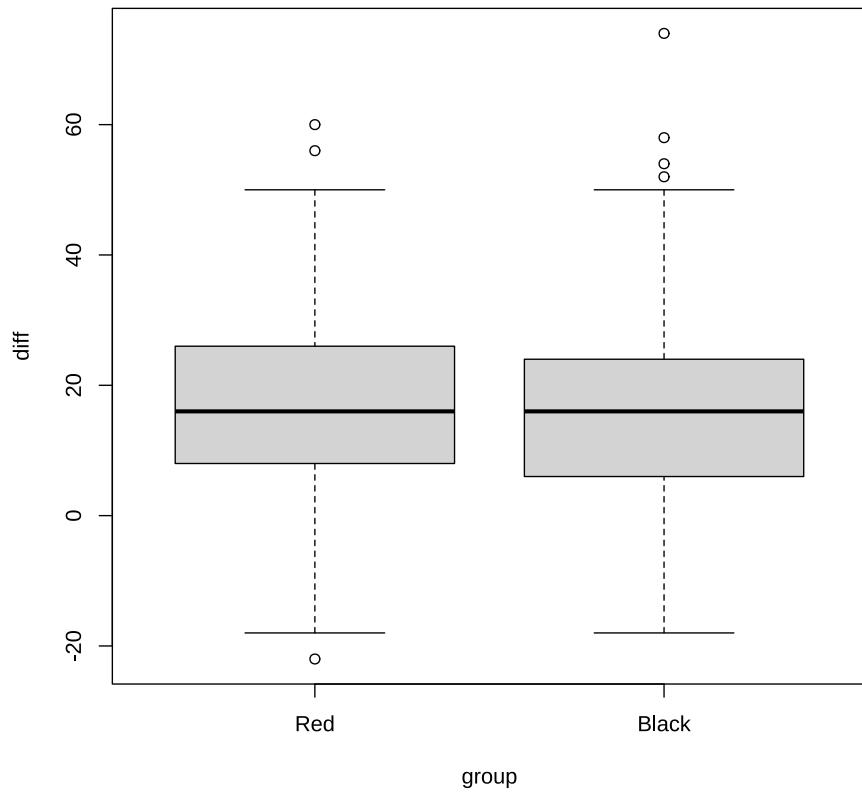
Black 그룹 개선점수의 히스토그램



Red 그룹 개선점수의 히스토그램과 Black 그룹 개선점수의 히스토그램을 아래 위로
그려서 비교하였습니다.

최소값, 제1사분위수, 중위수, 제3사분위수, 최대값이 상당히 닮은 것을 시각적으로 느낄 수 있습니다.

17.3.5 Box plot by Red and Black



Red 그룹 개선점수의 Boxplot과 Black 그룹 개선점수의 Boxplot을 화면 안에 나란히 배치하여 비교하였습니다.

이 또한 상당히 닮은 것을 시각적으로 느낄 수 있습니다.

17.3.6 개별 식생활습관 등급 변화

식생활습관 현황과 계획 조사에 참여하거나 불참한 모든 수강생의 점수입니다.

시청전 점수와 시청후 점수가 어떻게 바뀌고 시청전등급이 시청후등급으로 어떻게 바뀌는지 살펴보기 바랍니다.

동영상을 시청하고도 개선점수가マイ너스인 사람들은 전체 응답자 724(명) 중에서 총 49(명)입니다.

제 18 장 식생활습관 교육동영상 시청 효과

18.1 실험의 목적

국민문해력조사, 옥스포드 행복도 조사에 이어 여러분의 식생활 습관에 대하여 알아보고 있습니다.

식생활습관 현황을 묻는 질문 후에 식생활 개선 관련 동영상을 시청하면 앞으로 식생활습관을 바꾸겠다고 생각할지 묻는 실험입니다.

즉, 동영상 시청의 효과를 파악하는 실험인데 동일한 대상에게 시청 전과 시청 후의 인식 차이를 물어 보는 것입니다.

집계 결과를 여러 가지 통계로 제시하였습니다.

18.2 동영상 시청 전후 비교

교차표를 빈도수와 백분율로 비교하고, 개별 수강생의 점수 차이에 대하여 기초 통계량과 t-test를 수행하였습니다.

18.2.1 Cross Tables

	식생활을 고치세요	식생활이 보통입니다	식생활이 대체로 좋아요	불참	계
시청 전	292	235	44	27	598
시청 후	57	262	245	34	598

전체 수강생 598(명) 중에 27(명)이 불참하여 총 571(명)이 식생활습관 현황 조사에 참여한 가운데 “식생활이 대체로 좋아요”인 수강생은 44(명)에 불과하였고, “식생활을

고치세요”인 수강생은 무려 292(명)에 달했었는데, 영국의 식생활개선 활동가 제이미 올리버의 학교 급식 동영상 시청 후 “식생활이 대체로 좋아요”로 개선하겠다는 수강생은 245(명)으로 급격히 늘어났고, 여전히 “식생활을 고치세요” 수준에 머무르게 되는 수강생은 57(명)으로 대폭 줄어듭니다.

18.2.2 Percentages

	식생활을 고치세요	식생활이 보통입니다	식생활이 대체로 좋아요	계
시청 전	51.1	41.2	7.7	100
시청 후	10.1	46.5	43.4	100

이를 백분율로 살펴보면 동영상 시청 전에 “식생활이 대체로 좋아요”인 수강생은 7.7%(%)에 불과하였고, “식생활을 고치세요”인 수강생은 무려 51.1%(%)에 달했었는데, 동영상 시청 후에는 “식생활이 대체로 좋아요”로 개선하겠다는 수강생은 43.4%(%)으로 급격히 늘어났고, 여전히 “식생활을 고치세요” 수준에 머무르게 되는 수강생은 10.1%(%)으로 대폭 줄어듭니다.

18.2.3 Summary Statistics for diff

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	NA's
-16	5	16	16.36	26	62	35

수강생 각자 동영상 시청 전 식생활습관 현황조사로부터 점수를 계산하였고, 시청 후에 식생활습관 계획조사로부터 습관을 개선했을 때의 점수를 계산하였기 때문에 그 차이를 개선점수라고 할 수 있습니다.

만약 개선점수가マイナス라면 동영상 시청의 효과가 없거나 오히려 부정적이라는 뜻이죠.

여기 최소값으로 나와 있는 -16.00(은)는 그 경우를 의미합니다.

차이가 플러스일 수록 동영상 시청의 효과를 입증하는 것이죠.

그 차이가 어느 정도 인지를 살펴보기 위하여 먼저 기초통계를 제시하였습니다.

여기서 개별 개선점수의 산술평균은 16.36로 상당히 큰 값입니다.

개별 개선점수의 중위수는 16.00로 상당히 많이 개선된 것으로 볼 수 있습니다.

얼마나 개선되었는지를 파악하려면 표준편차를 알아야 하고 이를 자유도의 제곱근으로 나눠 준 표준오차로 나눠 준 t-통계량이 나오게 됩니다.

18.2.4 Standard Deviation of **diff**

$$\begin{array}{c} \hline \text{SD}_\text{diff} \\ \hline \hline 13.48 \end{array}$$

개선점수의 표준편차는 13.48입니다.

따라서 개선점수의 산술평균은 차이가 없다고 가정할 때 보다 약 1.21 배수만큼 큰 개선효과를 의미합니다.

18.2.5 t-test for **diff** (One Sample)

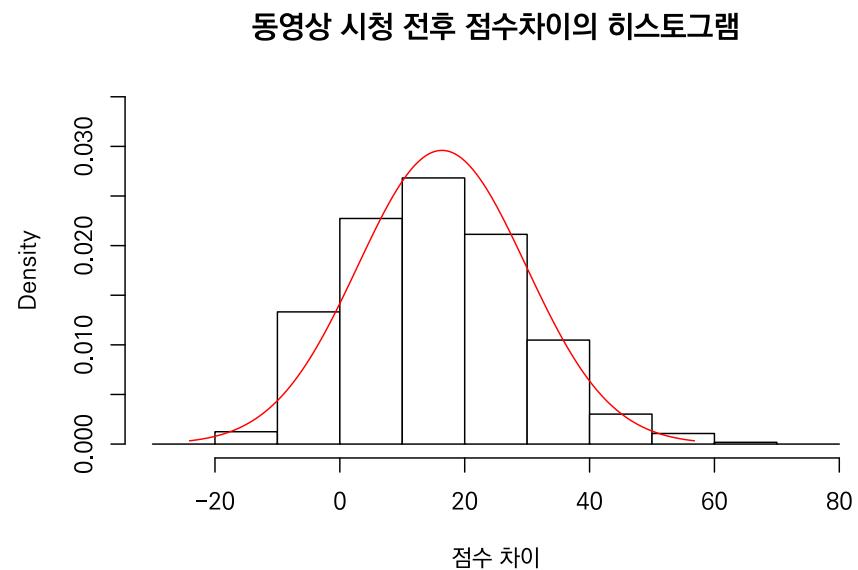
표 18.5: One Sample t-test: **diff**

Test statistic	df	P value	Alternative hypothesis	mean of x
28.79	562	5.644e-113 * *	greater	16.36

동영상 시청의 효과가 없다면 개선점수가 0이거나 평균적으로マイナス가 나올 것입니다.

이를 통계적으로 검증하기 위한 t-통계값은 28.79이고 자유도는 시청 전후의 식생활습관 점수가 모두 나와 있는 563(명)에서 한 명을 뺀 562이어서 사실상 정규분포라고 보아도 무방합니다.

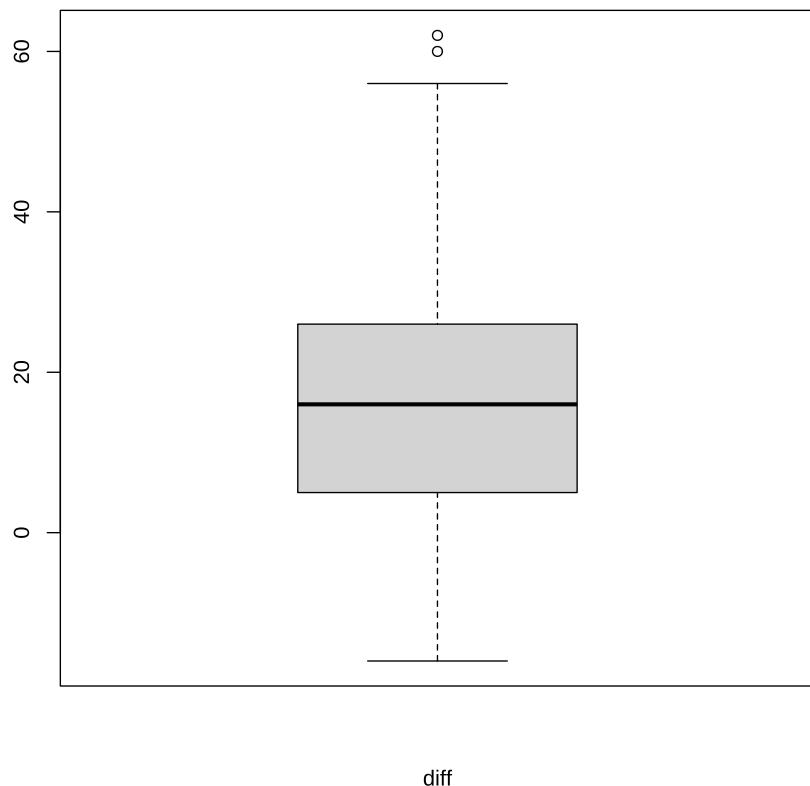
따라서 p-value 는 5.6e-113(0이)라는 엄청나게 작은 값이 나오고 따라서 동영상시청의 효과는 통계적으로 매우, 매우 유의한 것으로 결론내릴 수 있습니다.

18.2.6 Histogram for `diff`

개선점수의 분포를 히스토그램으로 나타내었습니다.

그리고 붉은색 선은 개선점수의 산술평균을 평균, 표준편차를 개선점수의 표준편차로 갖는 정규분포의 확률밀도함수입니다.

자유도 562인 t-분포는 사실상 정규분포와 같기 때문입니다.

18.2.7 Box plot for **diff**

개선점수의 Box Plot입니다.

중위수가 16.00인데, 20보다 작다는 것과 제1사분위수가 0과 20의 중간쯤으로 실제는 5.00, 제3사분위수가 20을 좀 넘는 26.00임을 대략 알 수 있습니다.

물론, 최소값이 0보다 작고 최대값은 40을 넘는다는 것도 시각적으로 파악 가능합니다.

18.3 Red and Black (랜덤화 효과)

비록 Red, Black 을 직접 적어 넣으라고 하지는 않았어도 기존에 부여받은 그룹이 있기 때문에 랜덤화 효과를 파악할 수 있습니다.

Red, Black 별로 기초통계값을 비교하고 t-test, 그룹별 히스토그램과 Box Plot을 통해서 시각적으로 비교합니다.

18.3.1 Summary by Red and Black

- Red:

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	NA's
-16	6	16	16.52	26	60	20

- Black:

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	NA's
-14	4	16	16.19	24	62	15

Red 그룹의 개선점수 평균은 16.52, Black 그룹의 개선점수 평균은 16.19로 크게 차이가 나지 않습니다.

Red 그룹의 개선점수 중위수는 16.00, Black 그룹의 개선점수 중위수는 16.00로 크게 차이가 나지 않습니다.

다른 기초통계값들은 각자 비교해 보세요.

18.3.2 Standard Deviation by Red and Black

Red	Black
13.81	13.17

Red 그룹 개선점수의 표준편차는 13.81이고, Black 그룹 개선점수의 표준편차는 13.17이어서 상당히 닮았습니다.

18.3.3 t-test (Welch's) Red vs Black (Two Sample)

표 18.9: Welch Two Sample t-test: diff by group

Test statistic	df	P value	Alternative hypothesis	mean in group Red	mean in group Black
0.2928	558.6	0.7697	two.sided	16.52	16.19

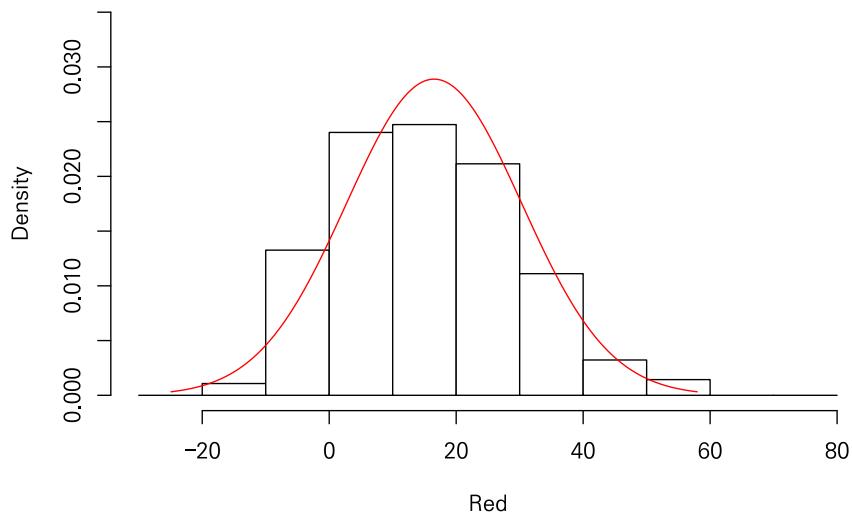
Red 그룹 개선점수 모집단의 표준편차와 Black 그룹 개선점수 모집단의 표준편차를 알 수 없기 때문에 Welch 의 근사식을 적용한 t-test (Two Sample) (을)를 수행하였습니다.

근사식으로 계산한 자유도는 558.59, t-통계값은 0.29280이고 따라서 p-value 는 0.76970이 나오므로 Red, Black 개선점수 간의 차이는 통계적으로 유의하지 않고 닮았다고 결론내릴 수 있습니다.

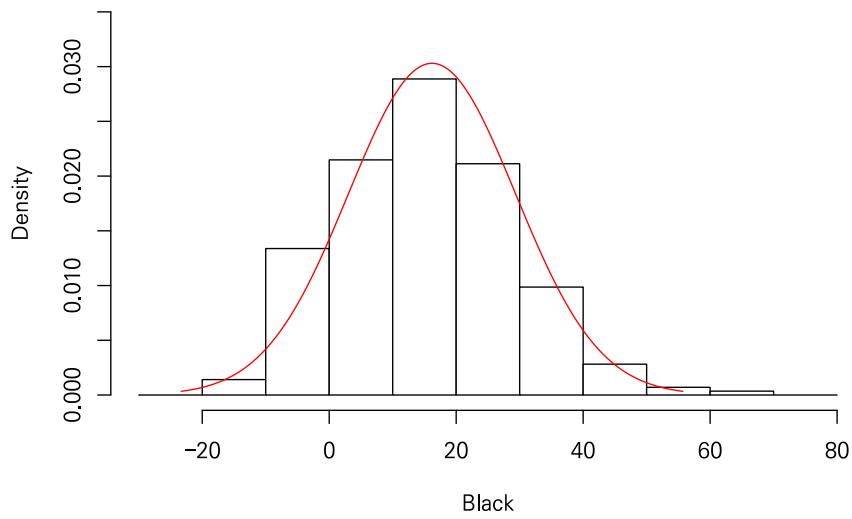
이 비교는 서로 다른지를 묻고 있기 때문에 two.sided, 즉 양측 검증이라는 것을 알 수 있고 Red 그룹 개선점수의 평균, 16.52(와)과 Black 그룹 개선점수의 평균, 16.19 도 요약표에서 확인할 수 있습니다.

18.3.4 Histogram by Red and Black

Red 그룹 개선점수의 히스토그램



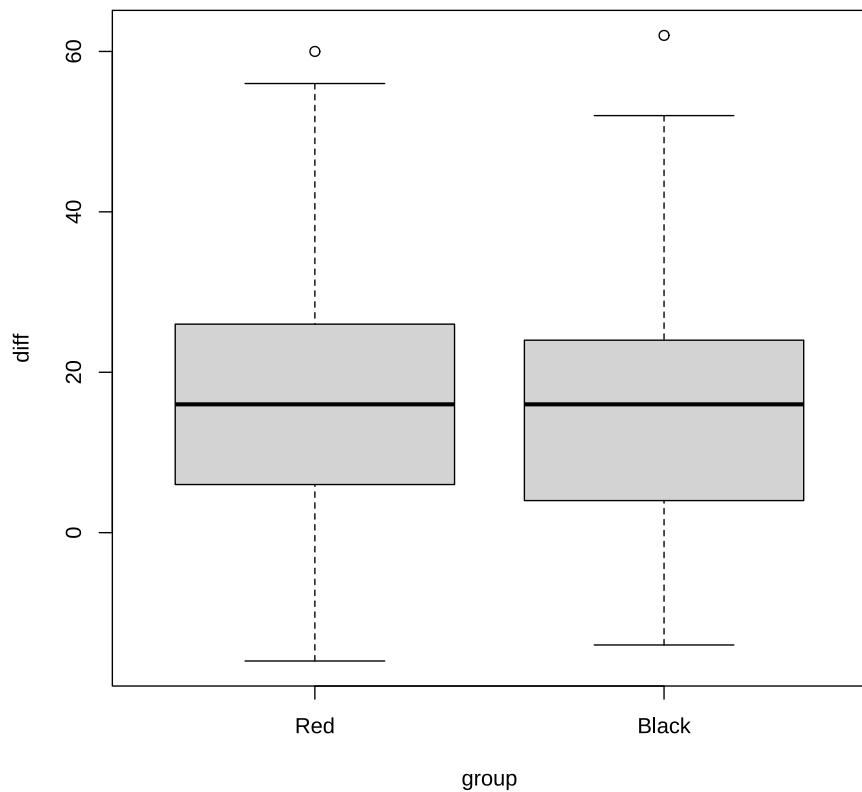
Black 그룹 개선점수의 히스토그램



Red 그룹 개선점수의 히스토그램과 Black 그룹 개선점수의 히스토그램을 아래 위로
그려서 비교하였습니다.

최소값, 제1사분위수, 중위수, 제3사분위수, 최대값이 상당히 닮은 것을 시각적으로 느낄 수 있습니다.

18.3.5 Box plot by Red and Black



Red 그룹 개선점수의 Boxplot과 Black 그룹 개선점수의 Boxplot을 화면 안에 나란히 배치하여 비교하였습니다.

이 또한 상당히 닮은 것을 시각적으로 느낄 수 있습니다.

18.3.6 개별 식생활습관 등급 변화

식생활습관 현황과 계획 조사에 참여하거나 불참한 모든 수강생의 점수입니다.

시청전 점수와 시청후 점수가 어떻게 바뀌고 시청전등급이 시청후등급으로 어떻게 바뀌는지 살펴보기 바랍니다.

동영상을 시청하고도 개선점수가マイ너스인 사람들은 전체 응답자 563(명) 중에서 총 33(명)입니다.

학번	그룹	시청전 점수	시청후 점수	점수차이	시청전등급	시청후등급
20171054	Red	72	72	0	식생활이 보통입니다	식생활이 보통입니다
20172727	Black	62	74	12	식생활이 보통입니다	식생활이 보통입니다
20181117	Red	40	100	60	식생활을 고치세요	식생활이 대체로 좋아요
20182101	Black	NA	NA	NA	NA	NA
20182936	Black	42	68	26	식생활을 고치세요	식생활이 보통입니다
20182982	Black	82	82	0	식생활이 대체로 좋아요	식생활이 대체로 좋아요
20183016	Black	82	88	6	식생활이 대체로 좋아요	식생활이 대체로 좋아요
20183242	Black	70	94	24	식생활이 보통입니다	식생활이 대체로 좋아요
20191035	Black	62	76	14	식생활이 보통입니다	식생활이 보통입니다
20191615	Black	52	80	28	식생활을 고치세요	식생활이 대체로 좋아요
20191702	Red	34	62	28	식생활을 고치세요	식생활이 보통입니다
20192201	Black	74	90	16	식생활이 보통입니다	식생활이 대체로 좋아요
20192204	Red	60	60	0	식생활이 보통입니다	식생활이 보통입니다
20192216	Red	42	72	30	식생활을 고치세요	식생활이 보통입니다
20192231	Red	48	58	10	식생활을 고치세요	식생활을 고치세요
20192742	Red	56	66	10	식생활을 고치세요	식생활이 보통입니다
20192844	Black	64	84	20	식생활이 보통입니다	식생활이 대체로 좋아요
20192854	Black	76	100	24	식생활이 보통입니다	식생활이 대체로 좋아요
20192957	Black	74	NA	NA	식생활이 보통입니다	NA
20193224	Black	52	64	12	식생활을 고치세요	식생활이 보통입니다
20193262	Red	52	84	32	식생활을 고치세요	식생활이 대체로 좋아요
20193716	Black	50	76	26	식생활을 고치세요	식생활이 보통입니다
20194143	Red	44	82	38	식생활을 고치세요	식생활이 대체로 좋아요
20195109	Red	56	74	18	식생활을 고치세요	식생활이 보통입니다
20195122	Red	68	88	20	식생활이 보통입니다	식생활이 대체로 좋아요
20195148	Black	76	86	10	식생활이 보통입니다	식생활이 대체로 좋아요
20195174	Red	66	82	16	식생활이 보통입니다	식생활이 대체로 좋아요
20196285	Red	74	90	16	식생활이 보통입니다	식생활이 대체로 좋아요
20196405	Black	62	88	26	식생활이 보통입니다	식생활이 대체로 좋아요
20201044	Red	60	92	32	식생활이 보통입니다	식생활이 대체로 좋아요
20201207	Black	62	76	14	식생활이 보통입니다	식생활이 보통입니다
20201603	Red	56	72	16	식생활을 고치세요	식생활이 보통입니다
20201610	Red	38	84	46	식생활을 고치세요	식생활이 대체로 좋아요
20201620	Red	38	60	22	식생활을 고치세요	식생활이 보통입니다
20201731	Black	74	60	-14	식생활이 보통입니다	식생활이 보통입니다
20202133	Red	50	70	20	식생활을 고치세요	식생활이 보통입니다
20202215	Red	36	60	24	식생활을 고치세요	식생활이 보통입니다
20202327	Red	NA	NA	NA	NA	NA
20202348	Black	46	98	52	식생활을 고치세요	식생활이 대체로 좋아요
20202362	Red	44	60	16	식생활을 고치세요	식생활이 보통입니다
20202405	Black	56	54	-2	식생활을 고치세요	식생활을 고치세요
20202514	Red	56	60	4	식생활을 고치세요	식생활이 보통입니다
20202524	Red	76	92	16	식생활이 보통입니다	식생활이 대체로 좋아요
20202531	Red	34	60	26	식생활을 고치세요	식생활이 보통입니다
20202537	Black	70	84	14	식생활이 보통입니다	식생활이 대체로 좋아요
20202581	Red	60	70	10	식생활이 보통입니다	식생활이 보통입니다
20202584	Black	64	74	10	식생활이 보통입니다	식생활이 보통입니다
20202591	Black	46	76	30	식생활을 고치세요	식생활이 보통입니다
20202616	Black	56	62	6	식생활을 고치세요	식생활이 보통입니다
20202905	Red	66	80	14	식생활이 보통입니다	식생활이 대체로 좋아요
20202913	Black	64	96	32	식생활이 보통입니다	식생활이 대체로 좋아요
20202917	Black	68	68	0	식생활이 보통입니다	식생활이 보통입니다
20202927	Black	60	60	0	식생활이 보통입니다	식생활이 보통입니다

제 19 장 11주차 데이터실험 집계

19.1 실험의 목적

11주차 구글 예습 설문지 집계결과를 분석합니다.

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

Q7에서는 Anchor Effects 를 알아 보기 위하여 Red 에서는 대학 등록금이 무상인 나라들의 예를 들고 나서 우리나라 국공립 대학 등록금의 적정 수준에 대해서 묻고, Black에서는 대학 등록금이 고액인 나라들의 예를 들고 나서 우리나라 국공립 대학 등록금의 적정 수준에 대해서 물어 보아 차이가 통계적으로 유의한지 살핍니다.

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

19.1.1 Red, Black을 잘못 표시한 사람들

	Red(구글예습퀴즈)	Black(구글예습퀴즈)
Red(랜덤화출석부)	345	1
Black(랜덤화출석부)	1	355
계	346	356

19.1.2 응답인원의 Red, Black

Red로 응답한 인원은 346명, Black에 응답한 인원은 356명입니다.

전체 응답인원 702 명을 랜덤하게 둘로 나눌 때 어느 한 쪽의 기대인원은 전체 응답인원의 절반인 351명이고, 표준오차는 전체 응답인원의 제곱근에 $1/2$ 을 곱해 준 13.2 명입니다.

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

19.2 Q1. 1분위와 5분위의 평균소비성향 비교

1. 다음 2021년 4/4분기 가계동향조사에서 1분위와 5분위의 평균소비성향을 제대로 비교한 것은? *

<표 5> 소득 5분위별 가계수지(2021. 4/4)

	(단위: 천원, %, %p, 전년 동분기대비)									
	1분위		2분위		3분위					
	증감률	증감률	증감률	증감률	증감률					
가구원수(명)	1.49	1.81	2.40	2.90	3.22					
가구주연령(세)	61.5	51.9	49.9	49.7	50.4					
소득	1,058	8.3	2,531	6.0	3,879	6.9	5,612	5.3	10,130	6.9
경상소득	1,048	7.9	2,512	5.8	3,847	6.7	5,562	5.3	9,726	6.1
근로소득	262	17.1	1,452	4.7	2,394	3.4	3,676	4.0	6,680	7.3
사업소득	142	7.0	485	17.3	834	14.9	1,229	14.4	2,402	2.7
재산소득	10	-16.8	15	-24.7	23	14.4	28	1.1	53	19.7
이전소득	634	5.2	560	1.1	596	9.5	628	-2.3	592	6.8
공적이전	462	9.0	404	5.2	438	21.6	407	-2.7	362	26.5
사적이전	172	-3.9	156	-8.1	157	-14.2	221	-1.5	230	-14.3
비경상소득	10	81.2	19	36.4	32	33.7	51	0.8	404	28.9
가계지출	1,363	5.5	2,175	6.9	3,063	10.1	4,140	3.3	6,288	4.5
소비지출	1,195	5.5	1,759	6.5	2,383	10.0	3,131	5.4	4,266	3.7
비소비지출	168	5.5	417	8.9	679	10.8	1,008	-2.7	2,022	6.1
처분가능소득 ¹⁾	890	8.8	2,114	5.4	3,199	6.1	4,604	7.2	8,108	7.1
흑자액 ²⁾	-305	3.2	355	0.4	816	-3.8	1,473	11.4	3,843	11.1
흑자율(%) ³⁾	-34.3	4.3p	16.8	-0.8p	25.5	-2.6p	32.0	1.2p	47.4	1.7p
평균소비성향(%) ⁴⁾	134.3	-4.3p	83.2	0.8p	74.5	2.6p	68.0	-1.2p	52.6	-1.7p

- 1분위의 평균소비성향이 5분위의 평균소비성향보다 낮다
- 1분위의 평균소비성향이 5분위의 평균소비성향보다 높다
- 1분위의 평균소비성향이 5분위의 평균소비성향과 같다
- 1분위의 평균소비성향과 5분위의 평균소비성향은 비교할 수 없다

19.2.1 집계

	1분위의 평균소비성향이 5분위의 평균소비성향보다 낮다	1분위의 평균소비성향이 5분위의 평균소비성향보다 높다	1분위의 평균소비성향과 5분위의 평균소비성향과 같다	1분위의 평균소비성향과 5분위의 평균소비성향은 비교할 수 없다	계
Red	23	257	52	14	346
Black	24	243	67	22	356
계	47	500	119	36	702

표 19.3: Pearson's Chi-squared test: .

Test statistic	df	P value
3.94	3	0.268

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

그 결과 카이제곱 통계량은 3.940, 자유도는 3, p-value 는 0.26800이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

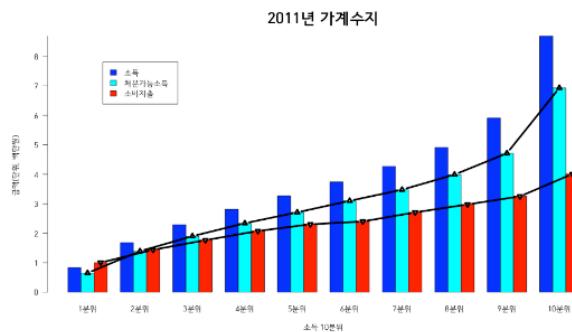
19.2.2 %

1분위의 평균소비성향이 5분위의 평균소비성향보다 낮다	1분위의 평균소비성향이 5분위의 평균소비성향보다 높다	1분위의 평균소비성향과 5분위의 평균소비성향과 같다	1분위의 평균소비성향과 5분위의 평균소비성향은 비교할 수 없다	계
6.7	71.2	17.0	5.1	100.0

정답률은 Red, Black 을 합하여 계산하는데, 71.2(%) 입니다.

19.3 Q2. 낙수효과

2. 다음 도표를 보고 낙수효과에 대하여 옳게 설명한 것은? *



- 낙수효과를 기대하기 어렵다
- 낙수효과를 잘 보여주고 있다
- 소비지출이 급격히 늘어나고 있다
- 고소득층의 평균소비성향이 급격히 늘어나고 있다

19.3.1 집계

	낙수효과를 기대하기 어렵다	낙수효과를 잘 보여주고 있다	소비지출이 급격히 늘어나고 있다	고소득층의 평균소비성향이 급격히 늘어나고 있다	계
Red	219	75	31	21	346
Black	210	81	42	23	356
계	429	156	73	44	702

표 19.6: Pearson's Chi-squared test: .

Test statistic	df	P value
2.026	3	0.567

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

그 결과 카이제곱 통계량은 2.026, 자유도는 3, p-value 는 0.56700이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

19.3.2 %

낙수효과를 기대하기 어렵다	낙수효과를 잘 보여주고 있다	소비지출이 급격히 늘어나고 있다	고소득층의 평균소비성향이 급격히 늘어나고 있다	계
61.1	22.2	10.4	6.3	100.0

정답률은 Red, Black 을 합하여 계산하는데, 61.1(%) 입니다.

19.4 Q3. 최고한계세율과 상위1% 소득점유율



19.4.1 집계

	최고한계세율이 높을 때 상위1%의 소득점유율도 높다.	최고한계세율이 낮을 때 상위1%의 소득점유율도 낮다.	최고한계세율이 높을 때 상위1%의 소득점유율은 낮다.	최고한계세율이 높아지면 상위1%의 소득점유율은 낮아진다.	계
Red	26	58	230	32	346
Black	28	62	213	53	356
계	54	120	443	85	702

표 19.9: Pearson's Chi-squared test: .

Test statistic	df	P value
5.907	3	0.1162

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

그 결과 카이제곱 통계량은 5.907, 자유도는 3, p-value 는 0.1162이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

19.4.2 %

	최고한계세율이 높을 때 상위1%의 소득점유율도 높다.	최고한계세율이 낮을 때 상위1%의 소득점유율도 낮다.	최고한계세율이 높을 때 상위1%의 소득점유율은 낮다.	최고한계세율이 높아지면 상위1%의 소득점유율은 낮아진다.	계
	7.7	17.1	63.1	12.1	100.0

정답률은 Red, Black 을 합하여 계산하는데, 63.1(%) 입니다.

19.5 Q4. 최고한계세율과 상위1%, 하위99%의 소득증가



19.5.1 집계

최고한계세율이 높았던 시기에 하위 99%의 소득증가율이 상위1%의 소득증가율 보다 높았다.	최고한계세율을 높이면 하위 99%의 소득증가율이 상위1%의 소득증가율보다 높아진다.	최고한계세율이 높았던 시기에 하위 99%의 소득이 상위1%의 소득보다 많았다.	최고한계세율을 높이면 하위 99%의 소득이 상위1%의 소득보다 많아진다.	계
Red 172	79	70	25	346
Black 175	76	71	34	356
계 347	155	141	59	702

표 19.12: Pearson's Chi-squared test: .

Test statistic	df	P value
1.322	3	0.724

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

그 결과 카이제곱 통계량은 1.322, 자유도는 3, p-value 는 0.72400이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

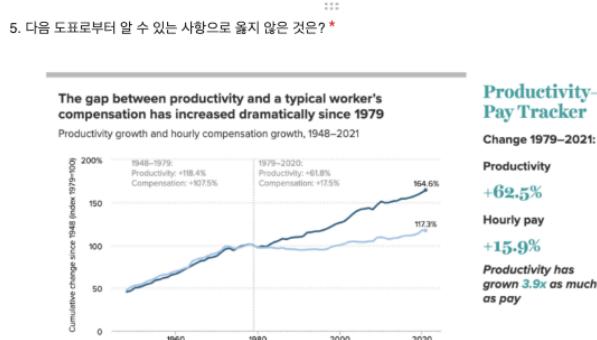
실제로 닮은 게 느껴집니까?

19.5.2 %

최고한계세율이 높았던 시기에 하위 99%의 소득증가율이 상위1%의 소득증가율 보다 높았다.	최고한계세율을 높이면 하위 99%의 소득증가율이 상위1%의 소득증가율보다 높아진다.	최고한계세율이 높았던 시기에 하위 99%의 소득이 상위1%의 소득보다 많았다.	최고한계세율을 높이면 하위 99%의 소득이 상위1%의 소득보다 많아진다.	계
49.4	22.1	20.1	8.4	100.0

정답률은 Red, Black 을 합하여 계산하는데, 20.1(%) 입니다.

19.6 Q5. 생산성과 급료의 격차



- 1948년부터 1979년까지는 생산성과 급료가 함께 상승하였다.
- 1979년 이후 생산성과 급료가 함께 상승하였다.
- 1979년 이후 생산성은 62.5%의 상승을 보였다.
- 1979년 이후 급료는 15.9% 의 상승을 보였다.

19.6.1 집계

1948년부터 1979년까지는 생산성과 급료가 함께 상승하였다.					1979년 이후 생산성과 급료가 함께 상승하였다.	1979년 이후 생산성은 62.5%의 상승을 보였다.	1979년 이후 급료는 15.9% 의 상승을 보였다.	계
Red	36	238	46	26	346			
Black	40	224	70	22	356			
계	76	462	116	48	702			

표 19.15: Pearson's Chi-squared test: .

Test statistic	df	P value
5.792	3	0.1222

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

그 결과 카이제곱 통계량은 5.792, 자유도는 3, p-value 는 0.1222이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

19.6.2 %

1948년부터 1979년까지는 생산성과 급료가 함께 상승하였다.					1979년 이후 생산성과 급료가 함께 상승하였다.	1979년 이후 생산성은 62.5%의 상승을 보였다.	1979년 이후 급료는 15.9% 의 상승을 보였다.	계
10.8	65.8	16.5	6.8	100.0				

정답률은 Red, Black 을 합하여 계산하는데, 65.8(%) 입니다.

19.7 Q6. LM3

6. 어느 기관의 LM3가 2.0이라고 알려져 있다. 이 기관이 10억의 사업비를 받는다면 지역경제에 파급효과 * 는 얼마라고 할 수 있는가?

- 10억원
- 20억원
- 25억원
- 30억원

19.7.1 집계

	10억원	20억원	25억원	30억원	계
Red	13	249	56	28	346
Black	22	234	74	26	356
계	35	483	130	54	702

표 19.18: Pearson's Chi-squared test: .

Test statistic	df	P value
5.205	3	0.1574

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

그 결과 카이제곱 통계량은 5.205, 자유도는 3, p-value 는 0.1574이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

19.7.2 %

10억원	20억원	25억원	30억원	계
5.0	68.8	18.5	7.7	100.0

정답률은 Red, Black 을 합하여 계산하는데, 68.8(%) 입니다.

19.8 Q7. 국공립대 등록금 국제 비교

이 질문은 기준점이 어디에 있느냐에 따라서 응답이 달라진다는 Anchor Effects 를 보여줍니다.

국공립대학의 등록금이 무상인 나라들을 소개한 Red 에서는 적정 수준이 “무상” 또는 100만원이라고 응답한 비율이 국공립대학의 등록금이 우리나라보다 높거나 대등한 나라들의 예를 먼저 들고 적정 수준을 물어본 경우(Black)에 “무상” 또는 100만원이라고 응답한 비율보다 월등히 높은 것을 알 수 있습니다.

500만원 이상을 적정하다고 응답한 비율의 비교도 흥미롭습니다.

Red

7. (Thinking, Fast and Slow) 스웨덴, 노르웨이, 아일랜드, 아이슬란드, 핀란드, 체코 등에 *
서는 국공립대의 등록금이 전혀 없습니다(2006/2007). 우리나라 국공립대학 등록금의 수준
은 어느 정도가 적절하다고 생각하십니까?

- 무상
- 100만원
- 200만원
- 500만원
- 1,000만원

Black

7. (Thinking, Fast and Slow) 미국 국공립대학의 등록금은 5,700달러, 일본은 4,300달러, *
오스트레일리아는 4,000달러, 캐나다는 3,700달러입니다(2006/2007). 우리나라 국공립대
학 등록금의 수준은 어느 정도가 적절하다고 생각하십니까?

- 무상
- 100만원
- 200만원
- 500만원
- 1,000만원

19.8.1 국공립대 등록금의 적정 수준 (집계)

	무상	100만원	200만원	500만원	1,000만원	계
Red(등록금무상국가 소개)	66	125	120	29	6	346
Black(등록금고액국가 소개)	25	64	197	57	13	356
계	91	189	317	86	19	702

표 19.21: Pearson's Chi-squared test: .

Test statistic	df	P value
68.43	4	4.867e-14 ***

Q7의 Red에는 스웨덴, 노르웨이, 아일랜드, 아이슬란드, 핀란드, 체코 등 국공립대의 등록금이 전혀 없는 나라들의 예를 들고 나서 우리나라 국공립대 등록금의 적정 수준을 묻고, Black에는 미국, 일본, 오스트레일리아, 캐나다 등 국공립대학의 등록금이 높은 수준인 나라들의 예를 들고 나서 우리나라 국공립대 등록금의 적정 수준을 물었습니다.

Daniel Kahneman 의 Think, Fast and Slow 에서 소개한 바와 같이 이와 설정하였을 때 소개한 나라들의 등록금 수준이 마치 낮을 내린 것과 같은 역할을 하여 등록금이 무상인 나라들을 소개한 Red 에서는 우리나라 국공립대 등록금의 적정 수준을 낮게 잡고, 대학 등록금 수준이 높은 나라들을 소개한 Black 에서는 우리나라 국공립대 등록금의 적정 수준을 높게 잡는 것을 관찰할 수 있습니다.

그 결과 적정 등록금 수준을 다섯 단계로 나누어 차이를 분석하기 위한 카이제곱 통계량은 68.430, 자유도는 4, p-value 는 4.9e-14 으로 등록금의 적정 수준의 차이가 통계적으로 유의함을 보여 줍니다.

즉, Anchor 효과가 있는 것으로 파악됩니다.

19.8.2 국공립대 등록금의 적정 수준 (%)

	무상	100만원	200만원	500만원	1,000만원	계
Red(등록금무상국가 소개)	19.1	36.1	34.7	8.4	1.7	100.0
Black(등록금고액국가 소개)	7.0	18.0	55.3	16.0	3.7	100.0

이를 백분율로 살펴보면 대학 등록금 무상 국가들을 소개한 Red에서 국공립대 적정 등록금 수준을 무상이라고 답한 백분율은 19.1(%), 100만원이라고 답한 백분율은 36.1(%)입니다.

반면 대학 등록금이 고액인 국가들을 소개한 Black에서 국공립대 적정 등록금 수준을 무상이라고 답한 백분율은 7.0(%), 100만원이라고 답한 백분율은 18.0(%)입니다.

Red에서 국공립대 적정 등록금 수준을 500만원이라고 답한 백분율은 8.4(%), 1,000만원이라고 답한 백분율은 1.7(%)입니다.

반면 Black에서 국공립대 적정 등록금 수준을 500만원이라고 답한 백분율은 16.0(%), 1,000만원이라고 답한 백분율은 3.7(%)입니다.

19.8.3 국공립대 등록금의 적정 수준 (2단계) (%)

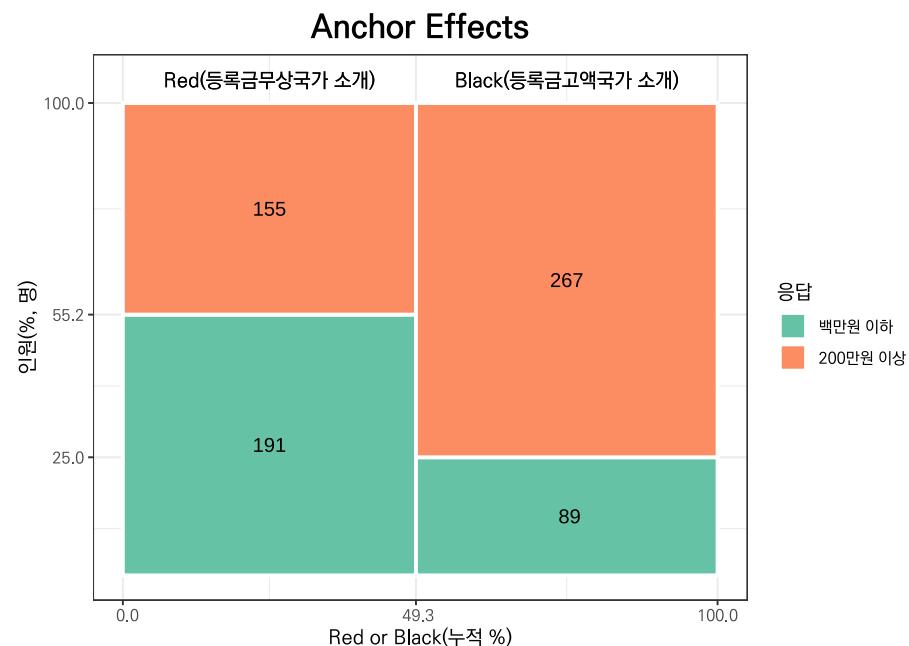
무상, 100만원을 100만원 이하, 200만원, 500만원, 1,000만원을 200만원 이상으로 나눠 보겠습니다.

	100만원 이하	200만원 이상	계
Red(등록금무상국가 소개)	55.2	44.8	100.0
Black(등록금고액국가 소개)	25.0	75.0	100.0

이를 백분율로 살펴보면 Red에서 100만원 이하를 적정 수준이라고 답한 백분율, 55.2(%)은 200만원 이상을 적정 수준이라고 답한 백분율, 36.1(%) 보다 월등히 높습니다. 반면 Black에서 100만원 이하를 적정 수준이라고 답한 백분율, 7.0(%)은 200만원 이상을 적정 수준이라고 답한 백분율, 18.0(%) 보다 낮습니다.

Anchor Effects 를 확인할 수 있는 것입니다.

19.8.4 Mosaic Plot



Mosaic Plot 은 이 집계결과를 시각적으로 잘 보여줍니다.

대학 등록금의 적정 수준을 예로 들은 나라들의 수준에 맞춰 설정하는 것을 시각적으로 파악할 수 있습니다.

19.9 마감 시간으로부터 제출 시간의 분포

19.9.1 분포표

표 19.24: 일 단위

	14일	13일	12일	11일	10일	9일	8일	7일	6일	5일	4일	3일	2일	1일	계
Red	117	20	14	8	9	10	10	35	12	19	24	17	19	32	346
Black	126	25	9	6	4	11	6	38	24	13	25	18	24	27	356
계	243	45	23	14	13	21	16	73	36	32	49	35	43	59	702

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.

19.9.2 날마다 고르게 제출하는가?

14일	13일	12일	11일	10일	9일	8일	7일	6일	5일	4일	3일	2일	1일
243	45	23	14	13	21	16	73	36	32	49	35	43	59

표 19.26: Chi-squared test for given probabilities: .

Test statistic	df	P value
878.9	13	1.718e-179 * * *

날마다 고르게 제출하는지 알아 보았습니다.

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

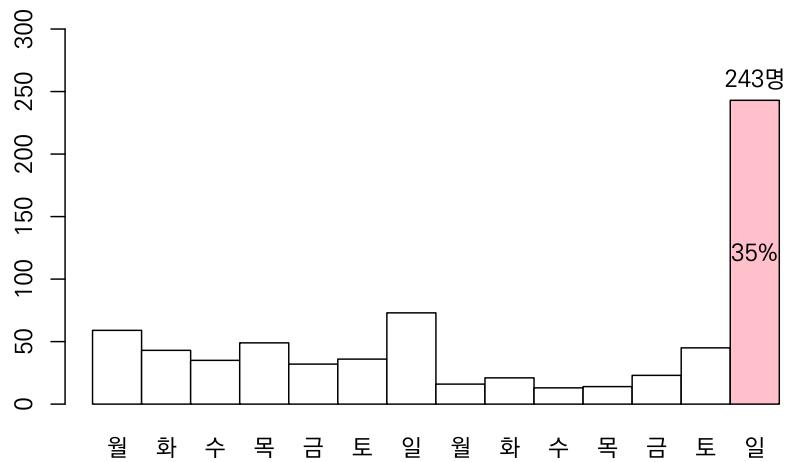
분포표 만으로도 쉽게 파악할 수 있지만 카이제곱테스트가 명확히 해 줍니다.

카이제곱 통계량은 878.883, 자유도는 13.00, p-value 는 1.7e-179 이므로 날짜별로 제출이 고르지 않다는 점을 강력히 시사합니다.

막대그래프로 살펴 보겠습니다.

19.9.3 막대그래프

Quiz241111 (702명 제출)



19.9.4 Red, Black 간에 닮았는가?

	14일	13일	12일	11일	10일	9일	8일	7일	6일	5일	4일	3일	2일	1일
Red	117	20	14	8	9	10	10	35	12	19	24	17	19	32
Black	126	25	9	6	4	11	6	38	24	13	25	18	24	27

표 19.28: Pearson's Chi-squared test: .

Test statistic	df	P value
11.39	13	0.5778

제출시간의 분포가 Red, Black 간에 닮았는지 알아 보았습니다.

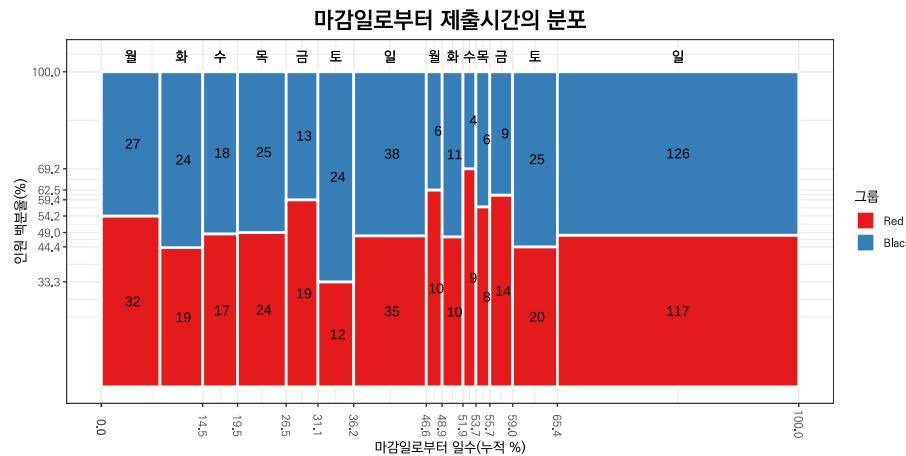
이번에는 분포표의 첫번째와 두번째 행, '계'열을 제외한 나머지 열에 대해서 카이제곱테스트를 수행합니다.

카이제곱 통계량은 11.39, 자유도는 13, p-value 는 0.5778 이므로 제출 시간의 분포는 Red, Black 간에 통계적으로 유의한 차이가 관찰되지 않습니다.

이 사실을 Mosaic Plot을 이용하여 시각적으로 살펴보겠습니다.

답았다고 느껴지나요?

19.9.5 Mosaic Plot



제 20 장 11주차 데이터실험 집계

20.1 실험의 목적

11주차 구글 예습 설문지 집계결과를 분석합니다.

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

Q7에서는 Anchor Effects 를 알아 보기 위하여 Red 에서는 대학 등록금이 무상인 나라들의 예를 들고 나서 우리나라 국공립 대학 등록금의 적정 수준에 대해서 묻고, Black에서는 대학 등록금이 고액인 나라들의 예를 들고 나서 우리나라 국공립 대학 등록금의 적정 수준에 대해서 물어 보아 차이가 통계적으로 유의한지 살핍니다.

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

20.1.1 Red, Black을 잘못 표시한 사람들

	Red(구글예습퀴즈)	Black(구글예습퀴즈)
Red(랜덤화출석부)	279	0
Black(랜덤화출석부)	0	277
계	279	277

20.1.2 응답인원의 Red, Black

Red로 응답한 인원은 279명, Black에 응답한 인원은 277명입니다.

전체 응답인원 556 명을 랜덤하게 둘로 나눌 때 어느 한 쪽의 기대인원은 전체 응답인원의 절반인 278명이고, 표준오차는 전체 응답인원의 제곱근에 $1/2$ 을 곱해 준 11.8 명입니다.

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

20.2 Q1. 1분위와 5분위의 평균소비성향 비교

1. 다음 2021년 4/4분기 가계동향조사에서 1분위와 5분위의 평균소비성향을 제대로 비교한 것은? *

<표 5> 소득 5분위별 가계수지(2021. 4/4)

	(단위: 천원, %, %p, 전년 동분기대비)									
	1분위		2분위		3분위					
	증감률	증감률	증감률	증감률	증감률					
가구원수(명)	1.49	1.81	2.40	2.90	3.22					
가구주연령(세)	61.5	51.9	49.9	49.7	50.4					
소득	1,058	8.3	2,531	6.0	3,879	6.9	5,612	5.3	10,130	6.9
경상소득	1,048	7.9	2,512	5.8	3,847	6.7	5,562	5.3	9,726	6.1
근로소득	262	17.1	1,452	4.7	2,394	3.4	3,676	4.0	6,680	7.3
사업소득	142	7.0	485	17.3	834	14.9	1,229	14.4	2,402	2.7
재산소득	10	-16.8	15	-24.7	23	14.4	28	1.1	53	19.7
이전소득	634	5.2	560	1.1	596	9.5	628	-2.3	592	6.8
공적이전	462	9.0	404	5.2	438	21.6	407	-2.7	362	26.5
사적이전	172	-3.9	156	-8.1	157	-14.2	221	-1.5	230	-14.3
비경상소득	10	81.2	19	36.4	32	33.7	51	0.8	404	28.9
가계지출	1,363	5.5	2,175	6.9	3,063	10.1	4,140	3.3	6,288	4.5
소비지출	1,195	5.5	1,759	6.5	2,383	10.0	3,131	5.4	4,266	3.7
비소비지출	168	5.5	417	8.9	679	10.8	1,008	-2.7	2,022	6.1
처분가능소득 ¹⁾	890	8.8	2,114	5.4	3,199	6.1	4,604	7.2	8,108	7.1
흑자액 ²⁾	-305	3.2	355	0.4	816	-3.8	1,473	11.4	3,843	11.1
흑자율(%) ³⁾	-34.3	4.3p	16.8	-0.8p	25.5	-2.6p	32.0	1.2p	47.4	1.7p
평균소비성향(%) ⁴⁾	134.3	-4.3p	83.2	0.8p	74.5	2.6p	68.0	-1.2p	52.6	-1.7p

- 1분위의 평균소비성향이 5분위의 평균소비성향보다 낮다
- 1분위의 평균소비성향이 5분위의 평균소비성향보다 높다
- 1분위의 평균소비성향이 5분위의 평균소비성향과 같다
- 1분위의 평균소비성향과 5분위의 평균소비성향은 비교할 수 없다

20.2.1 집계

	1분위의 평균소비성향이 5분위의 평균소비성향보다 낮다	1분위의 평균소비성향이 5분위의 평균소비성향보다 높다	1분위의 평균소비성향이 5분위의 평균소비성향과 같다	1분위의 평균소비성향과 5분위의 평균소비성향은 비교할 수 없다	계
Red	20	206	40	13	279
Black	19	199	41	18	277
계	39	405	81	31	556

표 20.3: Pearson's Chi-squared test: .

Test statistic	df	P value
0.9582	3	0.8114

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

그 결과 카이제곱 통계량은 0.958, 자유도는 3, p-value 는 0.81140이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

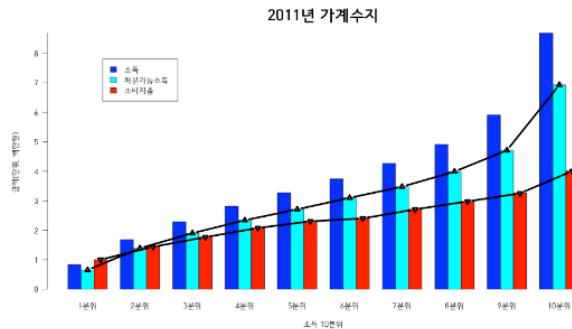
20.2.2 %

1분위의 평균소비성향이 5분위의 평균소비성향보다 낮다	1분위의 평균소비성향이 5분위의 평균소비성향보다 높다	1분위의 평균소비성향이 5분위의 평균소비성향과 같다	1분위의 평균소비성향과 5분위의 평균소비성향은 비교할 수 없다	계
7.0	72.8	14.6	5.6	100.0

정답률은 Red, Black 을 합하여 계산하는데, 72.8(%) 입니다.

20.3 Q2. 낙수효과

2. 다음 도표를 보고 낙수효과에 대하여 옳게 설명한 것은? *



- 낙수효과를 기대하기 어렵다
- 낙수효과를 잘 보여주고 있다
- 소비지출이 급격히 늘어나고 있다
- 고소득층의 평균소비성향이 급격히 늘어나고 있다

20.3.1 집계

	낙수효과를 기대하기 어렵다	낙수효과를 잘 보여주고 있다	소비지출이 급격히 늘어나고 있다	고소득층의 평균소비성향이 급격히 늘어나고 있다	
				계	고소득층의 평균소비성향이 급격히 늘어나고 있다
Red	174	52	33	20	279
Black	180	63	21	13	277
계	354	115	54	33	556

표 20.6: Pearson's Chi-squared test: .

Test statistic	df	P value
5.298	3	0.1512

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

그 결과 카이제곱 통계량은 5.298, 자유도는 3, p-value 는 0.15120이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

20.3.2 %

낙수효과를 기대하기 어렵다	낙수효과를 잘 보여주고 있다	소비지출이 급격히 늘어나고 있다	평균소비성향이 급격히 늘어나고 있다	고소득층의 계
63.7	20.7	9.7	5.9	100.0

정답률은 Red, Black 을 합하여 계산하는데, 63.7(%) 입니다.

20.4 Q3. 최고한계세율과 상위1% 소득점유율

3. 다음 도표로부터 알 수 있는 사항을 가장 잘 기술한 것은? *



- 최고한계세율이 높을 때 상위1%의 소득점유율도 높다.
- 최고한계세율이 낮을 때 상위1%의 소득점유율도 낮다.
- 최고한계세율이 높을 때 상위1%의 소득점유율은 낮다.
- 최고한계세율이 높아지면 상위1%의 소득점유율은 낮아진다.

20.4.1 집계

	최고한계세율이 높을 때 상위1%의 소득점유율도 높다.	최고한계세율이 낮을 때 상위1%의 소득점유율도 낮다.	최고한계세율이 높을 때 상위1%의 소득점유율은 낮다.	최고한계세율이 높아지면 상위1%의 소득점유율은 낮아진다.	계
Red	27	37	174	41	279
Black	27	31	183	36	277
계	54	68	357	77	556

표 20.9: Pearson's Chi-squared test: .

Test statistic	df	P value
1.074	3	0.7834

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

그 결과 카이제곱 통계량은 1.074, 자유도는 3, p-value 는 0.78340이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

20.4.2 %

	최고한계세율이 높을 때 상위1%의 소득점유율도 높다.	최고한계세율이 낮을 때 상위1%의 소득점유율도 낮다.	최고한계세율이 높을 때 상위1%의 소득점유율은 낮다.	최고한계세율이 높아지면 상위1%의 소득점유율은 낮아진다.	계
	9.7	12.2	64.2	13.8	100.0

정답률은 Red, Black 을 합하여 계산하는데, 64.2(%) 입니다.

20.5 Q4. 최고한계세율과 상위1%, 하위99%의 소득증가



20.5.1 집계

	최고한계세율이 높았던 시기에 하위 99%의 소득증가율이 상위1%의 소득증가율 보다 높았다.	최고한계세율을 높이면 하위 99%의 소득증가율이 상위1%의 소득증가율보다 높아진다.	최고한계세율이 높았던 시기에 하위 99%의 소득이 상위1%의 소득보다 많았다.	최고한계세율을 높이면 하위 99%의 소득이 상위1%의 소득보다 많아진다.	계
Red	159	44	56	20	279
Black	154	57	39	27	277
계	313	101	95	47	556

표 20.12: Pearson's Chi-squared test: .

Test statistic	df	P value
5.831	3	0.1201

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

그 결과 카이제곱 통계량은 5.831, 자유도는 3, p-value 는 0.12010이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

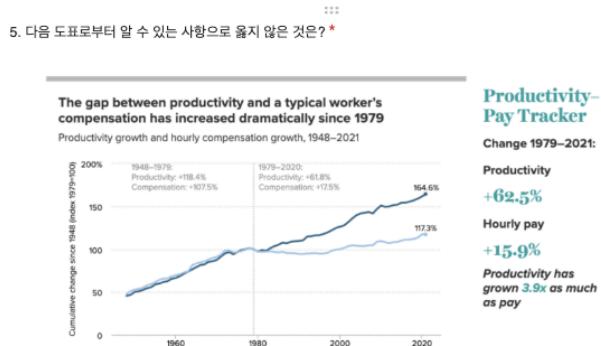
실제로 닮은 게 느껴집니까?

20.5.2 %

최고한계세율이 높았던 시기에 하위 99%의 소득증가율이 상위1%의 소득증가율 보다 높았다.	최고한계세율을 높이면 하위 99%의 소득증가율이 상위1%의 소득증가율보다 높아진다.	최고한계세율이 높았던 시기에 하위 99%의 소득이 상위1%의 소득보다 많았다.	최고한계세율을 높이면 하위 99%의 소득보다 많아진다.	계
56.3	18.2	17.1	8.5	100.0

정답률은 Red, Black 을 합하여 계산하는데, 17.1(%) 입니다.

20.6 Q5. 생산성과 급료의 격차



- 1948년부터 1979년까지는 생산성과 급료가 함께 상승하였다.
- 1979년 이후 생산성과 급료가 함께 상승하였다.
- 1979년 이후 생산성은 62.5%의 상승을 보였다.
- 1979년 이후 급료는 15.9% 의 상승을 보였다.

20.6.1 집계

1948년부터 1979년까지는 생산성과 급료가 함께 상승하였다.					1979년 이후 생산성과 급료가 함께 상승하였다.	1979년 이후 생산성은 62.5%의 상승을 보였다.	1979년 이후 급료는 15.9% 의 상승을 보였다.	계
Red	32	187	39	21	279			
Black	31	194	35	17	277			
계	63	381	74	38	556			

표 20.15: Pearson's Chi-squared test: .

Test statistic	df	P value
0.7746	3	0.8555

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

그 결과 카이제곱 통계량은 0.775, 자유도는 3, p-value 는 0.8555이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

20.6.2 %

1948년부터 1979년까지는 생산성과 급료가 함께 상승하였다.					1979년 이후 생산성과 급료가 함께 상승하였다.	1979년 이후 생산성은 62.5%의 상승을 보였다.	1979년 이후 급료는 15.9% 의 상승을 보였다.	계
11.3	68.5	13.3	6.8	100.0				

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

20.7 Q6. LM3

6. 어느 기관의 LM3가 2.0이라고 알려져 있다. 이 기관이 10억의 사업비를 받는다면 지역경제에 파급효과 * 는 얼마라고 할 수 있는가?

- 10억원
- 20억원
- 25억원
- 30억원

20.7.1 집계

	10억원	20억원	25억원	30억원	계
Red	12	200	52	15	279
Black	14	215	38	10	277
계	26	415	90	25	556

표 20.18: Pearson's Chi-squared test: .

Test statistic	df	P value
3.867	3	0.2762

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

그 결과 카이제곱 통계량은 3.867, 자유도는 3, p-value 는 0.2762이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

20.7.2 %

	10억원	20억원	25억원	30억원	계
	4.7	74.6	16.2	4.5	100.0

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

20.8 Q7. 국공립대 등록금 국제 비교

이 질문은 기준점이 어디에 있느냐에 따라서 응답이 달라진다는 Anchor Effects 를 보여줍니다.

국공립대학의 등록금이 무상인 나라들을 소개한 Red 에서는 적정 수준이 “무상” 또는 100만원이라고 응답한 비율이 국공립대학의 등록금이 우리나라보다 높거나 대등한 나라들의 예를 먼저 들고 적정 수준을 물어본 경우(Black)에 “무상” 또는 100만원이라고 응답한 비율보다 월등히 높은 것을 알 수 있습니다.

500만원 이상을 적정하다고 응답한 비율의 비교도 흥미롭습니다.

Red

7. (Thinking, Fast and Slow) 스웨덴, 노르웨이, 아일랜드, 아이슬란드, 핀란드, 체코 등에 *
서는 국공립대의 등록금이 전혀 없습니다(2006/2007). 우리나라 국공립대학 등록금의 수준
은 어느 정도가 적절하다고 생각하십니까?

- 무상
- 100만원
- 200만원
- 500만원
- 1,000만원

Black

7. (Thinking, Fast and Slow) 미국 국공립대학의 등록금은 5,700달러, 일본은 4,300달러, *
오스트레일리아는 4,000달러, 캐나다는 3,700달러입니다(2006/2007). 우리나라 국공립대
학 등록금의 수준은 어느 정도가 적절하다고 생각하십니까?

- 무상
- 100만원
- 200만원
- 500만원
- 1,000만원

20.8.1 국공립대 등록금의 적정 수준 (집계)

	무상	100만원	200만원	500만원	1,000만원	계
Red(등록금무상국가 소개)	66	99	90	19	5	279
Black(등록금고액국가 소개)	11	47	163	53	3	277
계	77	146	253	72	8	556

표 20.21: Cochran-Armitage test for trend: .

Test statistic	dim	P value	Alternative hypothesis
-8.755	5	0 * * *	one.sided

Q7의 Red에는 스웨덴, 노르웨이, 아일랜드, 아이슬란드, 핀란드, 체코 등 국공립대학의 등록금이 전혀 없는 나라들의 예를 든 후에 우리나라 국공립대학 등록금의 적정 수준을 묻고, Black에는 미국, 일본, 오스트레일리아, 캐나다 등 국공립대학 등록금이 높은 수준인 나라들의 예를 든 후에 우리나라 국공립대학 등록금의 적정 수준을 물었습니다.

Daniel Kahneman 의 Thinking, Fast and Slow 에서 설명한 것처럼 이와 같이 설정하였을 때 소개한 나라들의 등록금 수준이 마치 닻을 내린 것과 같은 역할을 하여 등록금이 무상인 나라들을 소개한 Red 에서는 우리나라 국공립대학 등록금의 적정 수준을 낮게 잡고, 대학 등록금 수준이 높은 나라들을 소개한 Black 에서는 우리나라 국공립대학 등록금의 적정 수준을 높게 잡는 것을 관찰할 수 있습니다.

응답결과를 분석하는 데 있어서 유의해야 할 사항은 등록금 수준이 무상에서 1000만원까지 순서가 있다는 점입니다. 따라서 명목형 변수에 대하여 사용한 카이제곱 동일성 테스트를 적용하는 것은 적절치 않고 코크란-아미티지 테스트가 이 상황에 적합한 분석 도구입니다.

적정 등록금 수준을 다섯 단계로 나누어 차이를 분석하기 위한 코크란-아미티지 통계량은 -8.755, dim은 5, p-value 는 0.0000 으로 등록금의 적정 수준의 차이가 통계적으로 유의함을 보여 줍니다.

즉, Anchor 효과가 있는 것으로 파악됩니다.

20.8.2 국공립대 등록금의 적정 수준 (%)

	무상	100만원	200만원	500만원	1,000만원	계
Red(등록금무상국가 소개)	23.7	35.5	32.3	6.8	1.8	100.0
Black(등록금고액국가 소개)	4.0	17.0	58.8	19.1	1.1	100.0

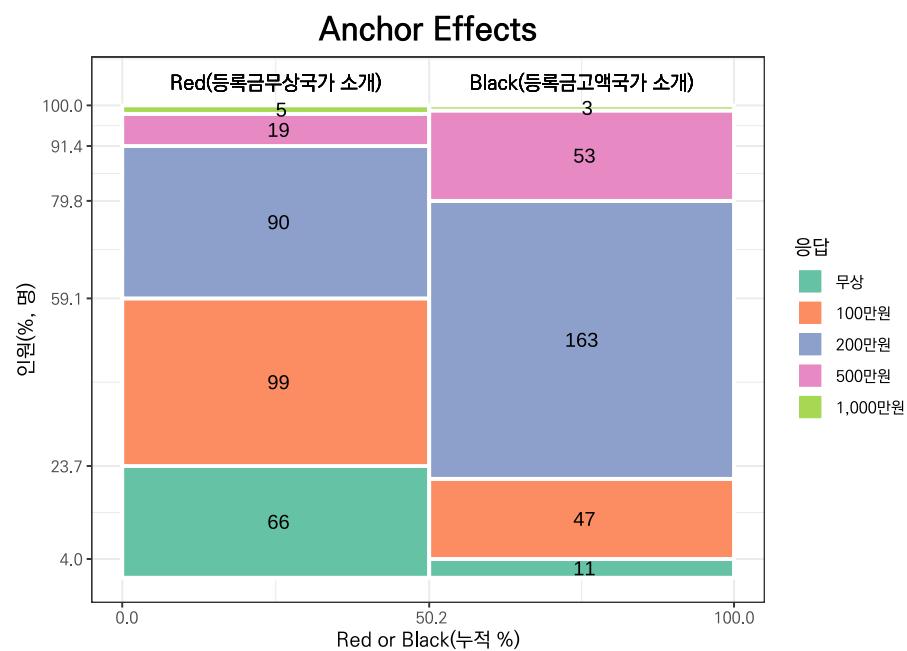
이를 백분율로 살펴보면 대학 등록금 무상 국가들을 소개한 Red에서 국공립대 적정 등록금 수준을 무상이라고 답한 백분율은 23.7(%), 100만원이라고 답한 백분율은 35.5(%)입니다.

반면 대학 등록금이 고액인 국가들을 소개한 Black에서 국공립대 적정 등록금 수준을 무상이라고 답한 백분율은 4.0(%), 100만원이라고 답한 백분율은 17.0(%)입니다.

Red에서 국공립대 적정 등록금 수준을 500만원이라고 답한 백분율은 6.8(%), 1,000만원이라고 답한 백분율은 1.8(%)입니다.

반면 Black에서 국공립대 적정 등록금 수준을 500만원이라고 답한 백분율은 19.1(%), 1,000만원이라고 답한 백분율은 1.1(%)입니다.

20.8.3 Mosaic Plot



Mosaic Plot 은 이 집계결과를 시각적으로 잘 보여줍니다.

대학 등록금의 적정 수준을 예로 들은 나라들의 수준에 맞춰 설정하는 것을 시각적으로 파악할 수 있습니다.

20.8.4 국공립대 등록금의 적정 수준 (2단계) (%)

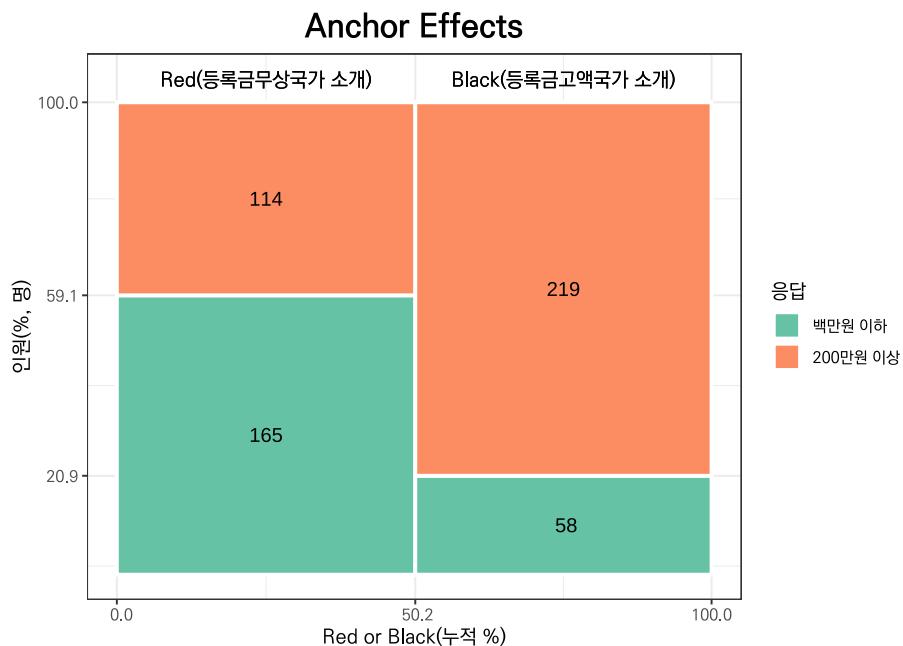
무상, 100만원을 100만원 이하, 200만원, 500만원, 1,000만원을 200만원 이상으로 나눠 보겠습니다.

	100만원 이하	200만원 이상	계
Red(등록금무상국가 소개)	59.1	40.9	100.0
Black(등록금고액국가 소개)	20.9	79.1	100.0

이를 백분율로 살펴보면 Red 에서 100만원 이하를 적정 수준이라고 답한 백분율, 59.1(%)은 200만원 이상을 적정 수준이라고 답한 백분율, 35.5(%) 보다 월등히 높습니다. 반면 Black에서 100만원 이하를 적정 수준이라고 답한 백분율, 4.0(%)은 200만원 이상을 적정 수준이라고 답한 백분율, 17.0(%) 보다 낮습니다.

Anchor Effects 를 확인할 수 있는 것입니다.

20.8.5 Mosaic Plot



Mosaic Plot 은 이 집계결과를 시각적으로 잘 보여줍니다.

대학 등록금의 적정 수준을 예로 들은 나라들의 수준에 맞춰 설정하는 것을 시각적으로 파악할 수 있습니다.

20.9 마감 시간으로부터 제출 시간의 분포

20.9.1 분포표

표 20.24: 일 단위

	14일	13일	12일	11일	10일	9일	8일	7일	6일	5일	4일	3일	2일	1일	계
Red	127	19	7	8	2	12	6	31	16	10	13	7	6	15	279
Black	146	15	8	11	4	3	8	21	12	8	7	11	7	16	277
계	273	34	15	19	6	15	14	52	28	18	20	18	13	31	556

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.

20.9.2 날마다 고르게 제출하는가?

14일	13일	12일	11일	10일	9일	8일	7일	6일	5일	4일	3일	2일	1일
273	34	15	19	6	15	14	52	28	18	20	18	13	31

표 20.26: Chi-squared test for given probabilities: .

Test statistic	df	P value
1519	13	4.088e-317 * * *

날마다 고르게 제출하는지 알아 보았습니다.

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

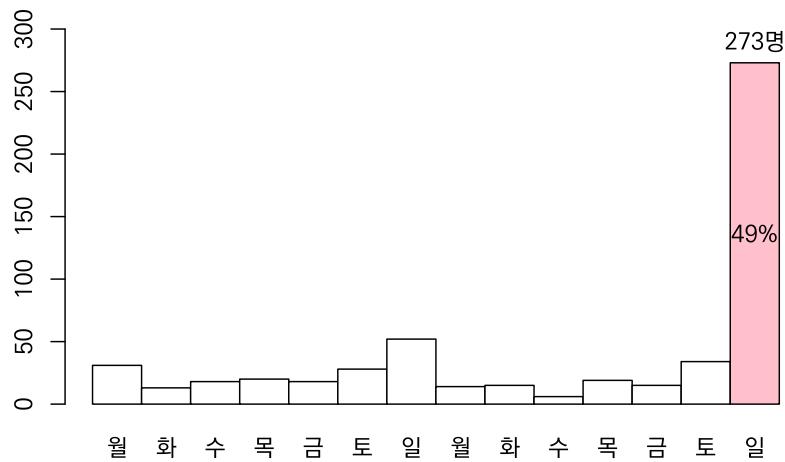
분포표 만으로도 쉽게 파악할 수 있지만 카이제곱테스트가 명확히 해 줍니다.

카이제곱 통계량은 1518.669, 자유도는 13.00, p-value 는 4.1e-317 이므로 날짜별로 제출이 고르지 않다는 점을 강력히 시사합니다.

막대그래프로 살펴 보겠습니다.

20.9.3 막대그래프

Quiz250512 (556명 제출)



20.9.4 Red, Black 간에 닮았는가?

	14일	13일	12일	11일	10일	9일	8일	7일	6일	5일	4일	3일	2일	1일
Red	127	19	7	8	2	12	6	31	16	10	13	7	6	15
Black	146	15	8	11	4	3	8	21	12	8	7	11	7	16

표 20.28: Pearson's Chi-squared test: .

Test statistic	df	P value
14.19	13	0.3604

제출시간의 분포가 Red, Black 간에 닮았는지 알아 보았습니다.

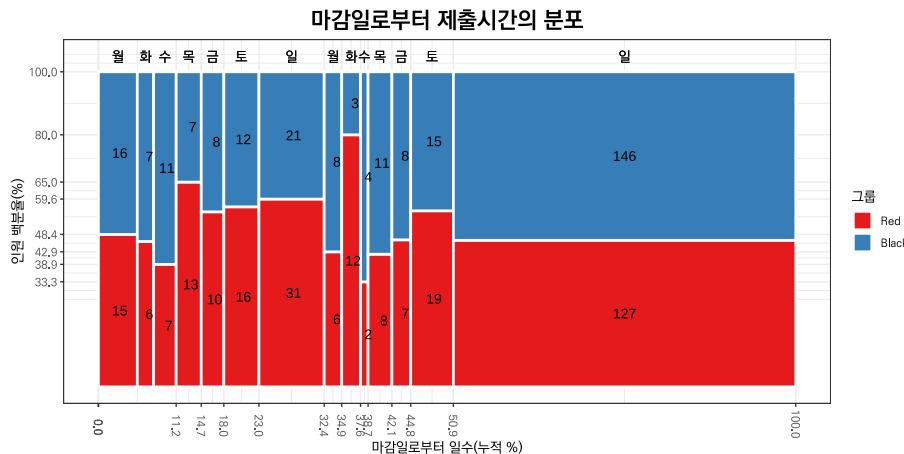
이번에는 분포표의 첫번째와 두번째 행, '계'열을 제외한 나머지 열에 대해서 카이제곱테스트를 수행합니다.

카이제곱 통계량은 14.19, 자유도는 13, p-value 는 0.3604 이므로 제출 시간의 분포는 Red, Black 간에 통계적으로 유의한 차이가 관찰되지 않습니다.

이 사실을 Mosaic Plot을 이용하여 시각적으로 살펴보겠습니다.

답았다고 느껴지나요?

20.9.5 Mosaic Plot



제 21 장 Big Five Personality and Multiple Intelligence

21.1 성격특성별 분포

이 수업에서 성격 5요인(Big Five Personality)을 파악하기 위하여 사용하고 있는 뉴캐슬 성격 평가(NPA, Newcastle Personality Assessor 12 Questions)는 영국 뉴캐슬 대학교의 대니얼 네틀(Daniel Nettle) 박사가 개발한 간단한 성격 평가 도구로, 12개의 문항을 통해 개인의 성격을 평가합니다. 대니얼 네틀의 저서 'Personality'를 2009년에 이라는 제목으로 번역판이 나왔는데 번역도 매끄럽고 읽기 좋은 편이었지만 책 31쪽의 성격진단표에서 “7~9번에 대한 응답점수”는 “7, 9번에 대한 응답점수”의 잘못된 번역이라는 것을 알게 되었습니다. 꽤 오랜 기간 올려 놓았던 자료들을 모두 수정해야 했습니다. 그 과정에 도움을 준 책은 2018년에 초판이 나온 최현석 프라임요양병원 원장의 저서 이었습니다.

이 평가 도구는 성격의 다섯 가지 주요 요소인 개방성(Openness), 성실성(Conscientiousness), 외향성(Extraversion), 친화성(Agreeableness), 신경성(Neuroticism)을 측정합니다. 각 요소는 2~3개의 문항으로 구성되어 있으며, 응답자는 각 문항에 대해 5점 척도로 자신에게 얼마나 해당하는지를 평가합니다. NPA는 짧은 시간 내에 성격 특성을 평가할 수 있어 연구나 개인적인 자기 이해를 위한 도구로 활용됩니다. 다만, 문항 수가 적어 심층적인 분석보다는 전반적인 성격 경향을 파악하는 데 유용합니다.

1. 외향성(Extraversion)은 사회적 상황에서의 활력과 적극성을 나타내며, 타인과의 상호작용을 즐기는 성향입니다. 외향성 점수가 높은 사람은 사교적이고 활동적이며, 긍정적 감정을 잘 표현합니다. 외향성 점수가 낮은 사람은 내성적이고, 혼자 있는 것을 선호하며, 에너지를 외부보다는 내부에서 얻는 경향이 있습니다.
2. 신경성(Neuroticism)은 감정적 안정성에 대한 성향으로, 불안정성과 부정적 감정에 쉽게 영향을 받는 정도를 나타냅니다. 신경성 점수가 높은 사람은 스트레스와 불안에 민감하고, 감정의 기복이 심할 수 있습니다. 신경성 점수가 낮은 사람은 감정적으로 안정적이고, 스트레스에 강하며, 감정을 잘 다스리는 성향을 보입니다.
3. 성실성(Conscientiousness)은 계획적이고 신중하며, 목표를 향해 꾸준히 노력하는 성향을 나타냅니다. 성실성 점수가 높은 사람은 자기 통제가 강하고,

322 CHAPTER 21. BIG FIVE PERSONALITY AND MULTIPLE INTELLIGENCE

책임감이 있으며, 일관성 있게 목표를 추구합니다. 성실성 점수가 낮은 사람은 충동적이고 계획적이지 않은 성향이 있으며, 조직적이지 않다고 여겨질 수 있습니다.

4. 친화성(Agreeableness)은 타인에 대한 공감과 협력적 성향, 신뢰도를 의미합니다. 친화성 점수가 높은 사람은 이타적이고, 타인과 협력하는 데 능하며, 신뢰감이 높습니다. 친화성 점수가 낮은 사람은 자기주장이 강하고, 경쟁적이며, 때로는 타인과 대립적인 성향을 보일 수 있습니다.
5. 개방성(Openness)은 새로운 경험, 아이디어, 예술적 감각에 대한 개방성과 수용성을 의미합니다. 개방성 점수가 높은 사람은 상상력이 풍부하고 창의적이며, 새로운 것을 탐구하려는 경향이 강합니다. 개방성 점수가 낮은 사람은 전통을 중시하고, 보수적이며, 변화를 덜 선호하는 경향이 있습니다.

이제, 여러분들의 응답결과를 집계해 보겠습니다.

21.1.1 집계

	외향성	신경성	성실성	친화성	개방성
높음	51	136	136	104	30
중상	174	233	242	111	91
중하	200	222	234	286	157
낮음	294	128	107	218	441
계	719	719	719	719	719

집계결과를 성격요인별 등급별로 나누어 교차표를 만들었습니다. 전체 719명이 응답한 가운데 외향성의 경우 “높음”은 51명에 불과하고 “낮음”은 294명이나 됩니다. 외향성이 “중하”이하인 사람은 494명으로 외향성이 “중상” 이상인 사람, 225명 보다 월등히 많습니다. 즉, 내향적인 사람들이 훨씬 많습니다.

신경성의 경우 “높음”은 136명, “낮음”은 128명으로 비슷합니다. 신경성이 “중하”이하인 사람은 350명으로 신경성이 “중상” 이상인 사람, 369명과 비슷합니다. 신경성의 분포는 상당히 대칭이라고 할 수 있겠습니다.

성실성의 경우 “높음”은 136명, “낮음”은 107명으로 “높음”이 월등히 많습니다. 성실성이 “중하”이하인 사람은 341명으로 성실성이 “중상” 이상인 사람, 378명 보다 적습니다. 성실성의 분포는 마감일이 다가옴에 따라 큰 변화를 보입니다. 특히 “낮음”이 얼마나 늘어나는지 관찰해 보세요.

친화성의 경우 “높음”은 104명으로 친화성이 “낮음”인 218명 보다 적습니다. 친화성이 “중하”이하인 사람은 504명으로 친화성이 “중상” 이상인 사람, 215명 보다 월등히 많습니다.

개방성의 경우 “높음”은 30명으로 절대적으로 해당 인원이 적고, “낮음”인 441명에 많이 몰려 있습니다. 개방성이 “중하”이하인 사람은 598명으로 친화성이 “중상” 이상인 사람, 121명 보다 월등히 많습니다.

백분율로 비교해 보겠습니다.

21.1.2 %

	외향성	신경성	성실성	친화성	개방성
높음	7.1	18.9	18.9	14.5	4.2
중상	24.2	32.4	33.7	15.4	12.7
중하	27.8	30.9	32.5	39.8	21.8
낮음	40.9	17.8	14.9	30.3	61.3
계	100.0	100.0	100.0	100.0	100.0

외향성의 경우 “높음”的 백분율은 7.1(%)에 불과하고 “낮음”的 백분율은 40.9(%)나 됩니다. 외향성이 “중하”이하인 사람의 백분율은 68.7(%)로 외향성이 “중상”이상인 사람의 백분율 31.3(%) 보다 월등히 많습니다. 즉, 내향적인 사람들이 훨씬 많습니다.

신경성의 경우 “높음”的 백분율은 18.9(%), “낮음”的 백분율은 17.8(%)로 비슷합니다. 신경성이 “중하”이하인 사람의 백분율은 48.7(%)로 신경성이 “중상”이상인 사람의 백분율 51.3(%)와 비슷합니다. 신경성의 분포는 상당히 대칭이라고 할 수 있겠습니다.

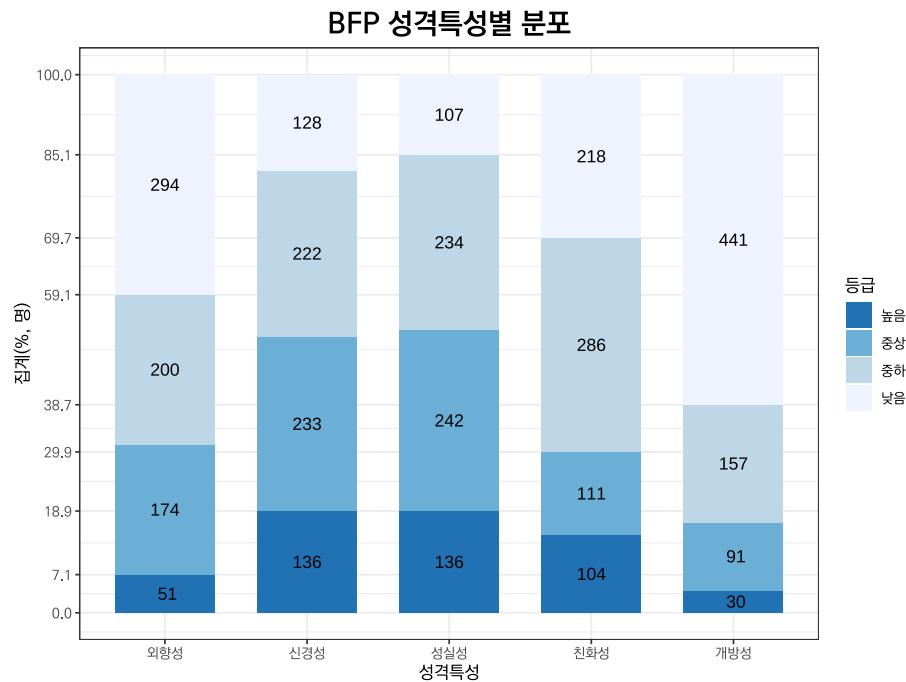
성실성의 경우 “높음”的 백분율은 18.9(%), “낮음”的 백분율은 14.9(%)로 “높음”이 다소 많습니다. 성실성이 “중하”이하인 사람의 백분율은 47.4(%)로 성실성이 “중상”이상인 사람의 백분율 52.6(%) 보다 다소 적습니다. 신경성과 마찬가지로 분포가 대칭에 가깝습니다. 성실성의 분포는 마감일이 다가옴에 따라 큰 변화를 보입니다. 특히 “낮음”이 얼마나 늘어나는지 관찰해 보세요.

친화성의 경우 “높음”的 백분율은 14.5(%)로 친화성이 “낮음”的 백분율 30.3(%) 보다 낮습니다. 친화성이 “중하”이하인 사람의 백분율은 70.1(%)로 친화성이 “중상”이상인 사람의 백분율 29.9(%)보다 월등히 많습니다.

개방성의 경우 “높음”的 백분율은 4.2(%)로 매우 낮은 수준이고, “낮음”的 백분율인 61.3(%)는 압도적으로 높습니다. 개방성이 “중하”이하인 사람의 백분율은 83.2(%)로 개방성이 “중상”이상인 사람의 백분율 16.8(%) 보다 압도적으로 높습니다. 왜 그런 걸까요?

이 분포표를 막대그래프로 시각화하였습니다. 외향성 낮은 사람들, 즉 내향적인 사람들이 압도적으로 많고, 신경성의 분포는 대칭이며, 성실성의 분포는 제출시간이 마감일에 가까워질수록 “낮음”的 비중이 늘어나는 것을 알 수 있습니다. 친화성이 ‘중하’ 이하인 사람들도 압도적으로 많지만 개방성이 “중하”이하인 사람들의 비중과는 비교도 되지 않습니다.

21.1.3 Barplot



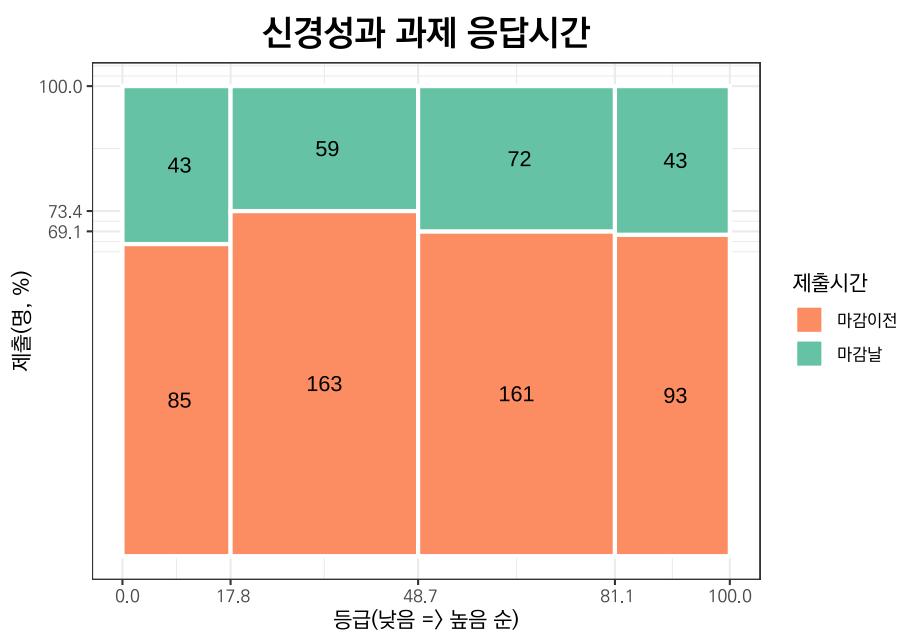
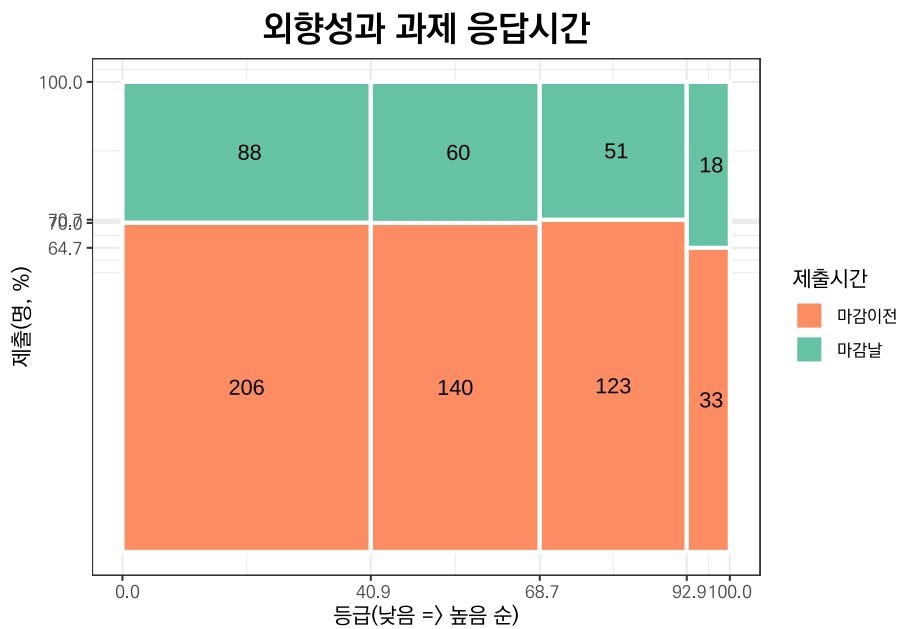
21.1.4 BFP와 과제제출 시간

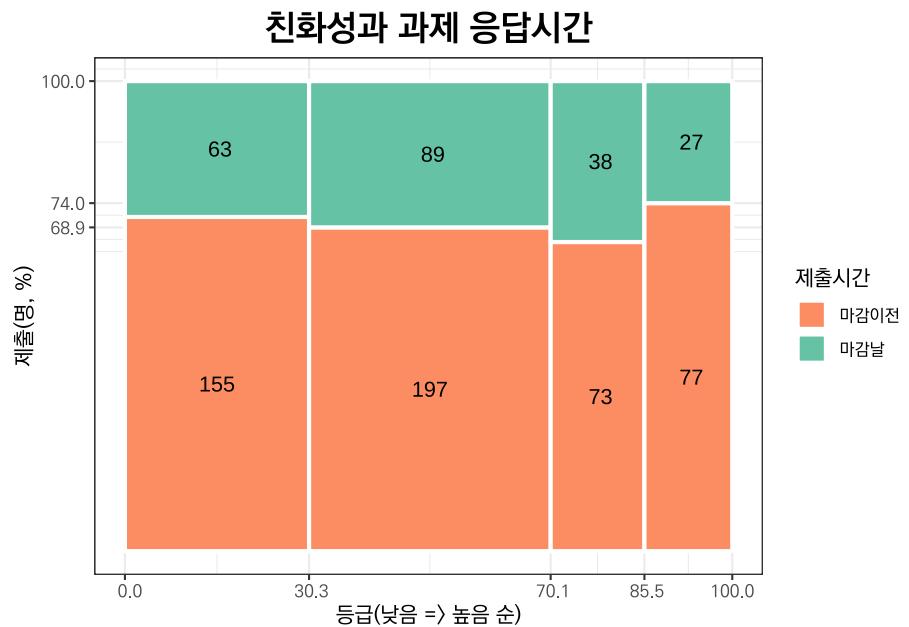
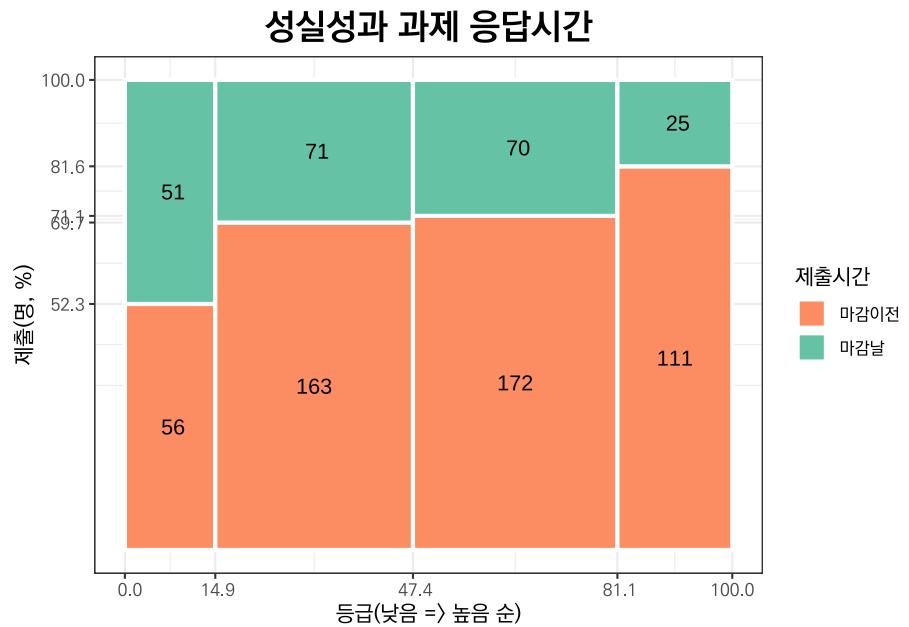
마감일이 되어서야 질문지에 응답하는 사람들과 그 이전에 응답하는 사람들이 BFP 와 어떤 관련이 있는지 알아 보겠습니다.

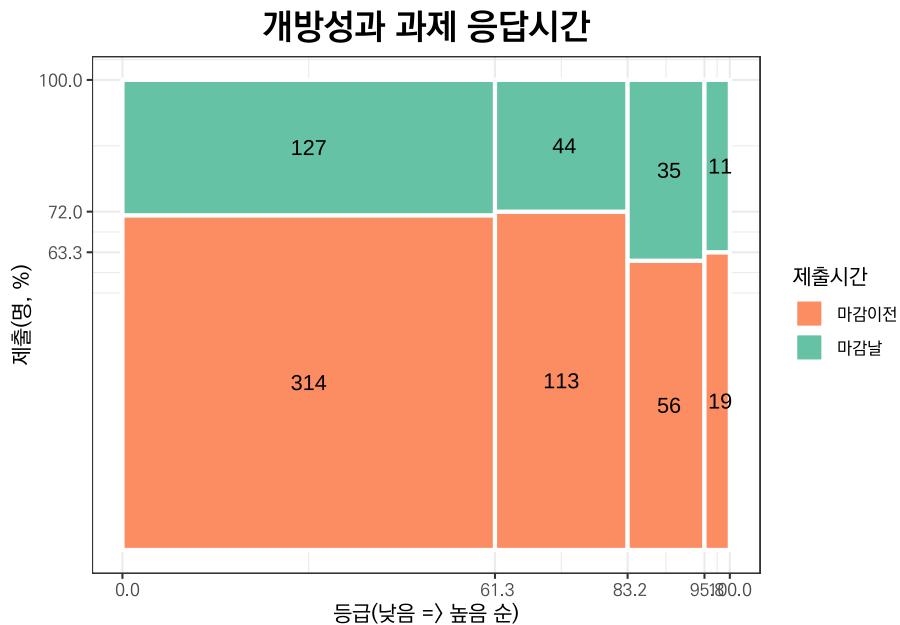
다른 성격 특성과 제출 시간과는 뚜렷한 관계가 나타나지 않지만 성실성은 “낮음”의 경우 마감날에 제출하는 비중이 40%에 가까울 정도로 높고, “중하”에서 “중상”의 경우는 약 30% 내외로 “낮음”에서 “높음” 순으로 갈수록 비중이 줄어드는 것을 알 수 있습니다.

성실성 “높음”의 경우는 마감날에 제출하는 사람들의 비중이 20%가 채 안됩니다.

성실성 “중상” 이상인 사람들이 마감일 이전에 제출하는 비중이 높다는 것을 알 수 있습니다.







21.2 3대 강점지능 분포

다중지능 이론은 하버드 대학교의 심리학자 하워드 가드너(Howard Gardner)가 1983년에 처음 제안한 개념으로, 지능이 단일한 능력이 아닌 다양한 형태로 존재한다는 이론입니다. 가드너는 사람들이 각기 다른 유형의 지능을 가지고 있으며, 교육에서도 이 다중지능을 고려해야 한다고 주장했습니다. 그는 원래 7가지 지능을 제안했으나 이후 8번째, 9번째 지능이 추가되었습니다. 우리 수업에서는 8개의 지능을 다룹니다. 다중지능의 8가지 유형입니다.

1. 음악 지능 (Musical Intelligence) 은 리듬, 멜로디, 음색 등을 인지하고 음악적으로 표현할 수 있는 능력입니다. 예시 직업으로는 음악가, 작곡가, 지휘자 등이 있습니다.
2. 신체운동 지능 (Bodily-Kinesthetic Intelligence) 은 신체를 자유롭게 조절하고 동작을 통해 자신의 생각을 표현할 수 있는 능력입니다. 예시 직업으로는 운동선수, 무용가, 배우, 외과의사 등이 있습니다.
3. 논리수학 지능 (Logical-Mathematical Intelligence) 은 수학적 연산과 논리적 사고에 뛰어난 능력으로 문제를 체계적으로 해결하는 데 강점이 있습니다. 예시 직업으로는 수학자, 과학자, 프로그래머, 공자 등이 있습니다.
4. 공간 지능 (Spatial Intelligence) 은 공간에서 물체를 인식하고 시각화하는 능력으로 도형이나 물체의 관계를 파악하고 상상할 수 있는 능력입니다. 예시 직업으로는 예술가, 건축가, 조각가, 디자이너 등이 있습니다.

328 CHAPTER 21. BIG FIVE PERSONALITY AND MULTIPLE INTELLIGENCE

5. 언어 지능 (Linguistic Intelligence)은 언어를 이해하고 표현하는 능력으로 읽기, 쓰기, 말하기, 단어의 의미 이해 능력이 뛰어납니다. 예시 직업으로는 작가, 언어학자, 변호사, 기자 등이 있습니다.
6. 인간친화 지능 (Interpersonal Intelligence)은 타인의 감정, 의도, 동기를 잘 파악하고 효과적으로 소통할 수 있는 능력입니다. 예시 직업으로는 교사, 상담사, 정치인, 마케팅 전문가 등이 있습니다.
7. 자기성찰 지능 (Intrapersonal Intelligence)은 자신의 감정과 욕구를 잘 인식하고 통제할 수 있는 능력으로 자기 이해와 성찰을 통해 목표를 성취하는 데 도움을 줍니다. 예시 직업으로는 철학자, 심리학자, 작가 등이 있습니다.
8. 자연친화 지능 (Naturalistic Intelligence)은 자연 환경, 생물학적 다양성, 동식물에 대한 민감성을 나타내며 생태계와의 상호작용에 관심을 갖습니다. 예시 직업으로는 생물학자, 환경운동가, 농업전문가 등이 있습니다.

21.2.1 집계

	음악	신체운동	논리수학	공간	언어	인간친화	자기성찰	자연친화	계
제1순위	81	48	91	71	65	183	154	19	712
제2순위	49	73	120	66	69	152	159	24	712
제3순위	59	87	121	76	68	143	117	41	712

3대 강점지능의 분포를 살펴 보면 인간친화 지능과 자기성찰 지능이 압도적으로 많이 포함되어 있음을 알 수 있습니다.

총 712명이 응답한 가운데 인간친화 지능이 제1순위에 포함된 사람은 183명, 제2순위에 포함된 사람은 152명, 제3순위에 포함된 사람은 143명으로 모두 478(명)이나 됩니다.

8가지 지능이 3대 강점지능에 골고루 포함된다면 어느 특정 지능이 3대 강점지능에 들어갈 기대인원은 전체인원을 8로 나눈 후 3을 곱해 준 267.0명 정도 기대되는 데 이 보다 훨씬 많은 숫자임을 알 수 있습니다.

그 다음으로 자기성찰 지능이 제1순위에 포함된 사람은 154명, 제2순위에 포함된 사람은 159명, 제3순위에 포함된 사람은 117명으로 모두 430(명)이나 됩니다.

그 다음으로는 논리수학 지능이 눈에 띄는 데 논리수학 지능이 제1순위에 포함된 사람은 91명, 제2순위에 포함된 사람은 120명, 제3순위에 포함된 사람은 121명으로 모두 332(명)입니다.

이제 백분율을 살펴보겠습니다.

21.2.2 %

	음악	신체운동	논리수학	공간	언어	인간친화	자기성찰	자연친화	계
제1순위	11.4	6.7	12.8	10.0	9.1	25.7	21.6	2.7	100.0
제2순위	6.9	10.3	16.9	9.3	9.7	21.3	22.3	3.4	100.0
제3순위	8.3	12.2	17.0	10.7	9.6	20.1	16.4	5.8	100.0

3대 강점지능의 분포를 백분율로 살펴 보아도 인간친화 지능과 자기성찰 지능이 압도적으로 많이 포함되어 있음을 알 수 있습니다.

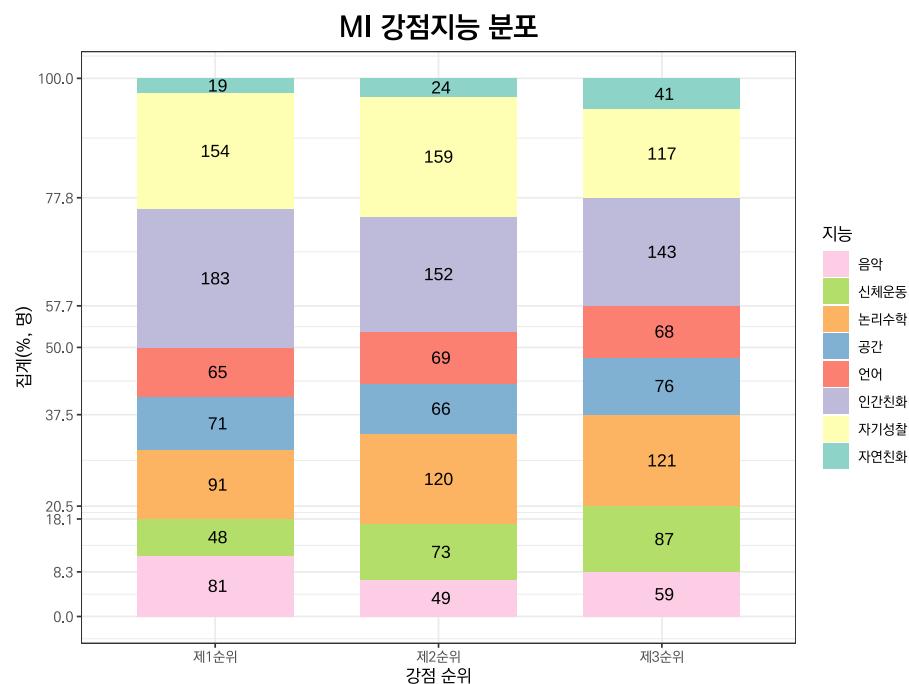
인간친화 지능이 제1순위에 포함된 사람의 백분율은 25.7(%), 제2순위에 포함된 사람은 21.3(%), 제3순위에 포함된 사람은 20.1(%)입니다.

그 다음으로 자기성찰 지능이 제1순위에 포함된 사람은 21.6(%), 제2순위에 포함된 사람은 22.3(%), 제3순위에 포함된 사람은 16.4(%입니다.

그 다음으로 논리수학 지능이 제1순위에 포함된 사람은 12.8(%), 제2순위에 포함된 사람은 16.9(%), 제3순위에 포함된 사람은 17.0(%입니다.

막대그래프로 이 상황을 시각적으로 요약해 보겠습니다.

21.2.3 Bar Plot



330 CHAPTER 21. BIG FIVE PERSONALITY AND MULTIPLE INTELLIGENCE

제 22 장 Big Five Personality and Multiple Intelligence

22.1 성격특성별 분포

이 수업에서 성격 5요인(Big Five Personality)을 파악하기 위하여 사용하고 있는 뉴캐슬 성격 평가(NPA, Newcastle Personality Assessor 12 Questions)는 영국 뉴캐슬 대학교의 대니얼 네틀(Daniel Nettle) 박사가 개발한 간단한 성격 평가 도구로, 12개의 문항을 통해 개인의 성격을 평가합니다. 대니얼 네틀의 저서 'Personality'를 2009년에 이라는 제목으로 번역판이 나왔는데 번역도 매끄럽고 읽기 좋은 편이었지만 책 31쪽의 성격진단표에서 “7~9번에 대한 응답점수”는 “7, 9번에 대한 응답점수”의 잘못된 번역이라는 것을 알게 되었습니다. 꽤 오랜 기간 올려 놓았던 자료들을 모두 수정해야 했습니다. 그 과정에 도움을 준 책은 2018년에 초판이 나온 최현석 프라임요양병원 원장의 저서 이었습니다.

이 평가 도구는 성격의 다섯 가지 주요 요소인 개방성(Openness), 성실성(Conscientiousness), 외향성(Extraversion), 친화성(Agreeableness), 신경성(Neuroticism)을 측정합니다. 각 요소는 2~3개의 문항으로 구성되어 있으며, 응답자는 각 문항에 대해 5점 척도로 자신에게 얼마나 해당하는지를 평가합니다. NPA는 짧은 시간 내에 성격 특성을 평가할 수 있어 연구나 개인적인 자기 이해를 위한 도구로 활용됩니다. 다만, 문항 수가 적어 심층적인 분석보다는 전반적인 성격 경향을 파악하는 데 유용합니다.

1. 외향성(Extraversion)은 사회적 상황에서의 활력과 적극성을 나타내며, 타인과의 상호작용을 즐기는 성향입니다. 외향성 점수가 높은 사람은 사교적이고 활동적이며, 긍정적 감정을 잘 표현합니다. 외향성 점수가 낮은 사람은 내성적이고, 혼자 있는 것을 선호하며, 에너지를 외부보다는 내부에서 얻는 경향이 있습니다.
2. 신경성(Neuroticism)은 감정적 안정성에 대한 성향으로, 불안정성과 부정적 감정에 쉽게 영향을 받는 정도를 나타냅니다. 신경성 점수가 높은 사람은 스트레스와 불안에 민감하고, 감정의 기복이 심할 수 있습니다. 신경성 점수가 낮은 사람은 감정적으로 안정적이고, 스트레스에 강하며, 감정을 잘 다스리는 성향을 보입니다.
3. 성실성(Conscientiousness)은 계획적이고 신중하며, 목표를 향해 꾸준히 노력하는 성향을 나타냅니다. 성실성 점수가 높은 사람은 자기 통제가 강하고,

책임감이 있으며, 일관성 있게 목표를 추구합니다. 성실성 점수가 낮은 사람은 충동적이고 계획적이지 않은 성향이 있으며, 조직적이지 않다고 여겨질 수 있습니다.

4. 친화성(Agreeableness)은 타인에 대한 공감과 협력적 성향, 신뢰도를 의미합니다. 친화성 점수가 높은 사람은 이타적이고, 타인과 협력하는 데 능하며, 신뢰감이 높습니다. 친화성 점수가 낮은 사람은 자기주장이 강하고, 경쟁적이며, 때로는 타인과 대립적인 성향을 보일 수 있습니다.
5. 개방성(Openness)은 새로운 경험, 아이디어, 예술적 감각에 대한 개방성과 수용성을 의미합니다. 개방성 점수가 높은 사람은 상상력이 풍부하고 창의적이며, 새로운 것을 탐구하려는 경향이 강합니다. 개방성 점수가 낮은 사람은 전통을 중시하고, 보수적이며, 변화를 덜 선호하는 경향이 있습니다.

이제, 여러분들의 응답결과를 집계해 보겠습니다.

22.1.1 집계

	외향성	신경성	성실성	친화성	개방성
높음	34	88	83	91	23
중상	139	202	200	115	79
중하	170	170	183	177	119
낮음	219	102	96	179	341
계	562	562	562	562	562

집계결과를 성격요인별 등급별로 나누어 교차표를 만들었습니다. 전체 562명이 응답한 가운데 외향성의 경우 “높음”은 34명에 불과하고 “낮음”은 219명이나 됩니다. 외향성이 “중하”이하인 사람은 389명으로 외향성이 “중상” 이상인 사람, 173명 보다 월등히 많습니다. 즉, 내향적인 사람들이 훨씬 많습니다.

신경성의 경우 “높음”은 88명, “낮음”은 102명으로 비슷합니다. 신경성이 “중하”이하인 사람은 272명으로 신경성이 “중상” 이상인 사람, 290명과 비슷합니다. 신경성의 분포는 상당히 대칭이라고 할 수 있겠습니다.

성실성의 경우 “높음”은 83명, “낮음”은 96명으로 “높음”이 월등히 많습니다. 성실성이 “중하”이하인 사람은 279명으로 성실성이 “중상” 이상인 사람, 283명 보다 적습니다. 성실성의 분포는 마감일이 다가옴에 따라 큰 변화를 보입니다. 특히 “낮음”이 얼마나 늘어나는지 관찰해 보세요.

친화성의 경우 “높음”은 91명으로 친화성이 “낮음”인 179명 보다 적습니다. 친화성이 “중하”이하인 사람은 356명으로 친화성이 “중상” 이상인 사람, 206명 보다 월등히 많습니다.

개방성의 경우 “높음”은 23명으로 절대적으로 해당 인원이 적고, “낮음”인 341명에 많이 몰려 있습니다. 개방성이 “중하”이하인 사람은 460명으로 친화성이 “중상” 이상인 사람, 102명 보다 월등히 많습니다.

백분율로 비교해 보겠습니다.

22.1.2 %

	외향성	신경성	성실성	친화성	개방성
높음	6.0	15.7	14.8	16.2	4.1
중상	24.7	35.9	35.6	20.5	14.1
중하	30.2	30.2	32.6	31.5	21.2
낮음	39.0	18.1	17.1	31.9	60.7
계	100.0	100.0	100.0	100.0	100.0

외향성의 경우 “높음”的 백분율은 6.0(%)에 불과하고 “낮음”的 백분율은 39.0(%)나 됩니다. 외향성이 “중하”이하인 사람의 백분율은 69.2(%)로 외향성이 “중상”이상인 사람의 백분율 30.8(%) 보다 월등히 많습니다. 즉, 내향적인 사람들이 훨씬 많습니다.

신경성의 경우 “높음”的 백분율은 15.7(%), “낮음”的 백분율은 18.1(%)로 비슷합니다. 신경성이 “중하”이하인 사람의 백분율은 48.4(%)로 신경성이 “중상”이상인 사람의 백분율 51.6(%)와 비슷합니다. 신경성의 분포는 상당히 대칭이라고 할 수 있겠습니다.

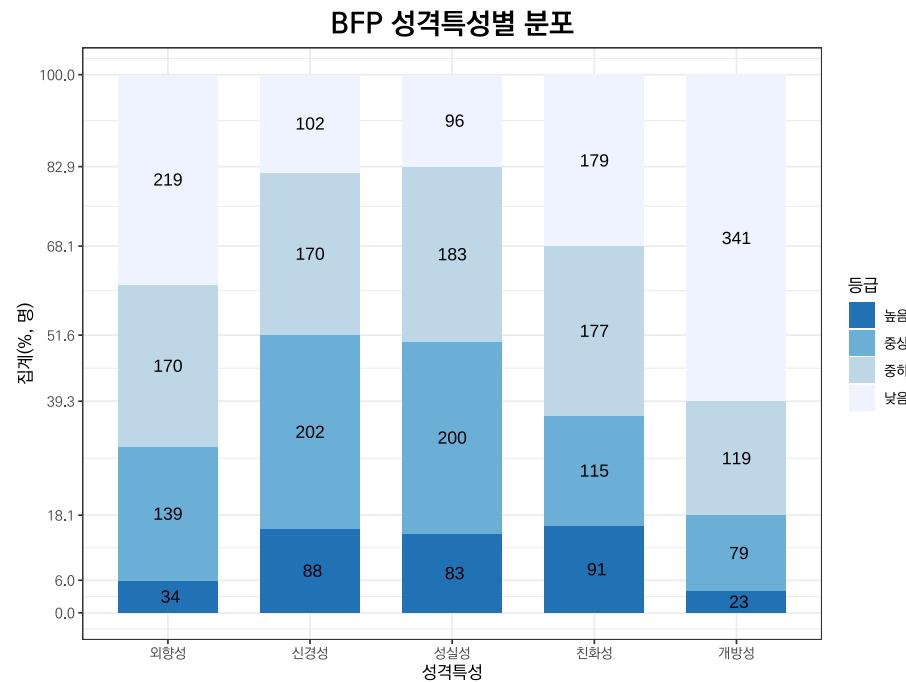
성실성의 경우 “높음”的 백분율은 14.8(%), “낮음”的 백분율은 17.1(%)로 “높음”이 다소 많습니다. 성실성이 “중하”이하인 사람의 백분율은 49.6(%)로 성실성이 “중상”이상인 사람의 백분율 50.4(%) 보다 다소 적습니다. 신경성과 마찬가지로 분포가 대칭에 가깝습니다. 성실성의 분포는 마감일이 다가옴에 따라 큰 변화를 보입니다. 특히 “낮음”이 얼마나 늘어나는지 관찰해 보세요.

친화성의 경우 “높음”的 백분율은 16.2(%)로 친화성이 “낮음”的 백분율 31.9(%) 보다 낮습니다. 친화성이 “중하”이하인 사람의 백분율은 63.3(%)로 친화성이 “중상”이상인 사람의 백분율 36.7(%)보다 월등히 많습니다.

개방성의 경우 “높음”的 백분율은 4.1(%)로 매우 낮은 수준이고, “낮음”的 백분율인 60.7(%)는 압도적으로 높습니다. 개방성이 “중하”이하인 사람의 백분율은 81.9(%)로 개방성이 “중상”이상인 사람의 백분율 18.1(%) 보다 압도적으로 높습니다. 왜 그런 걸까요?

이 분포표를 막대그래프로 시각화하였습니다. 외향성 낮은 사람들, 즉 내향적인 사람들이 압도적으로 많고, 신경성의 분포는 대칭이며, 성실성의 분포는 제출시간이 마감일에 가까워질수록 “낮음”的 비중이 늘어나는 것을 알 수 있습니다. 친화성이 ‘중하’ 이하인 사람들도 압도적으로 많지만 개방성이 “중하”이하인 사람들의 비중과는 비교도 되지 않습니다.

22.1.3 Barplot



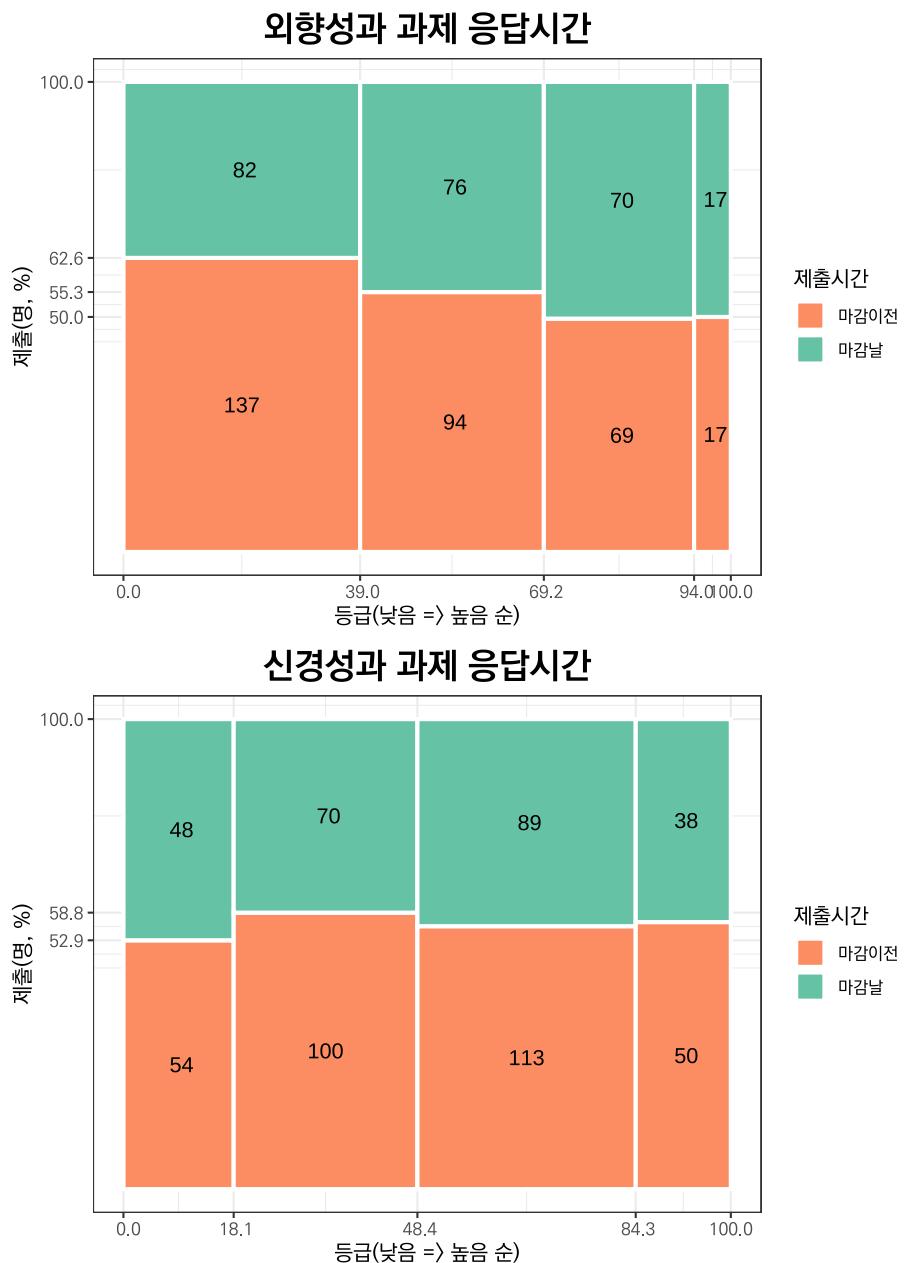
22.1.4 BFP와 과제제출 시간

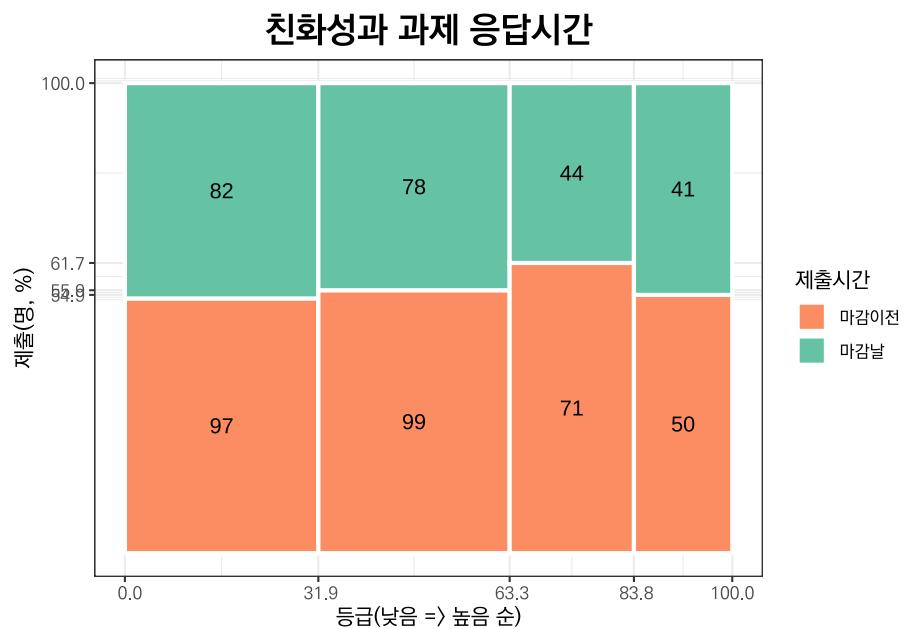
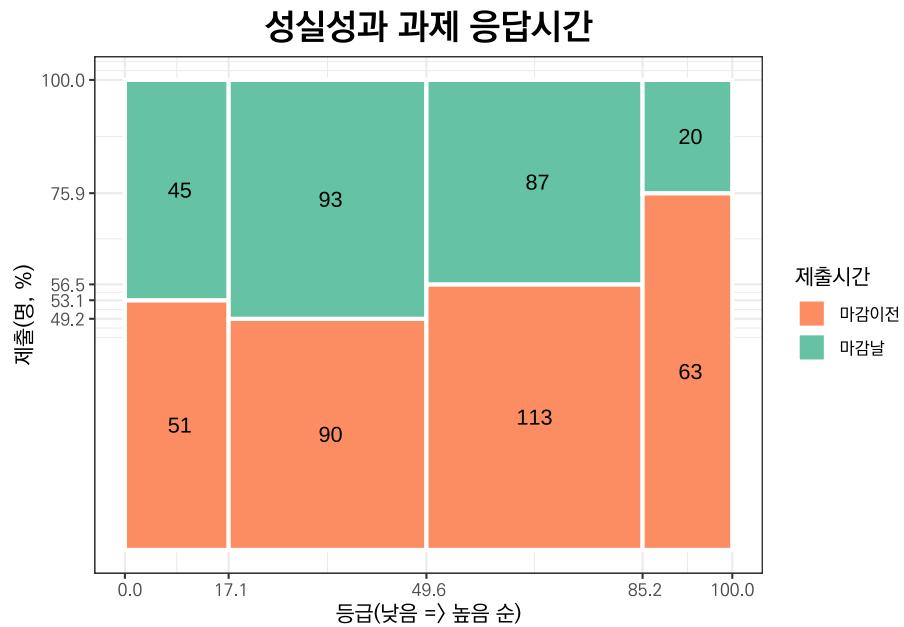
마감일이 되어서야 질문지에 응답하는 사람들과 그 이전에 응답하는 사람들이 BFP 와 어떤 관련이 있는지 알아 보겠습니다.

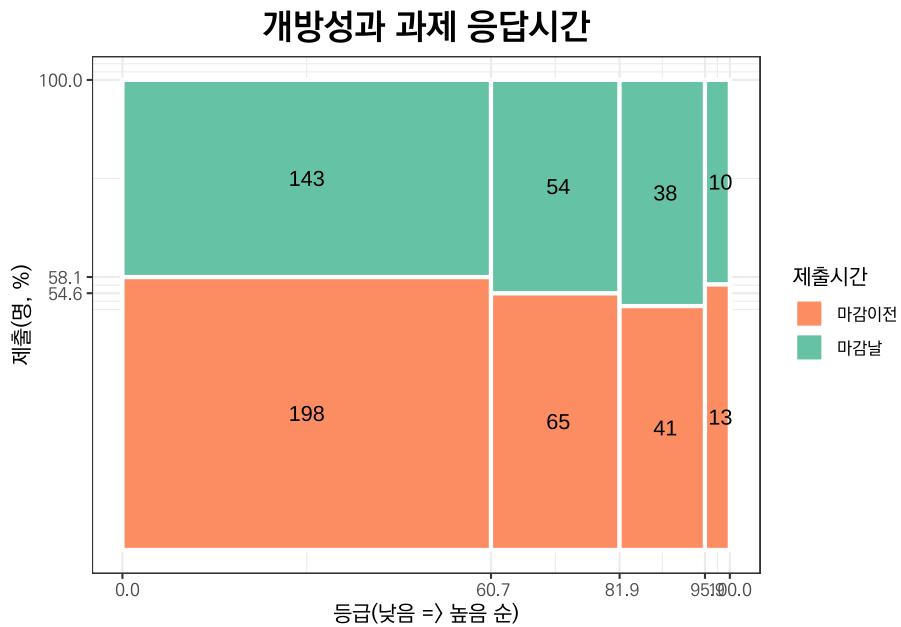
다른 성격 특성과 제출 시간과는 뚜렷한 관계가 나타나지 않지만 성실성은 “낮음”의 경우 마감날에 제출하는 비중이 40%에 가까울 정도로 높고, “중하”에서 “중상”의 경우는 약 30% 내외로 “낮음”에서 “높음” 순으로 갈수록 비중이 줄어드는 것을 알 수 있습니다.

성실성 “높음”의 경우는 마감날에 제출하는 사람들의 비중이 20%가 채 안됩니다.

성실성 “중상” 이상인 사람들이 마감일 이전에 제출하는 비중이 높다는 것을 알 수 있습니다.







22.2 3대 강점지능 분포

다중지능 이론은 하버드 대학교의 심리학자 하워드 가드너(Howard Gardner)가 1983년에 처음 제안한 개념으로, 지능이 단일한 능력이 아닌 다양한 형태로 존재한다는 이론입니다. 가드너는 사람들이 각기 다른 유형의 지능을 가지고 있으며, 교육에서도 이 다중지능을 고려해야 한다고 주장했습니다. 그는 원래 7가지 지능을 제안했으나 이후 8번째, 9번째 지능이 추가되었습니다. 우리 수업에서는 8개의 지능을 다룹니다. 다중지능의 8가지 유형입니다.

1. 음악 지능 (Musical Intelligence) 은 리듬, 멜로디, 음색 등을 인지하고 음악적으로 표현할 수 있는 능력입니다. 예시 직업으로는 음악가, 작곡가, 지휘자 등이 있습니다.
2. 신체운동 지능 (Bodily-Kinesthetic Intelligence) 은 신체를 자유롭게 조절하고 동작을 통해 자신의 생각을 표현할 수 있는 능력입니다. 예시 직업으로는 운동선수, 무용가, 배우, 외과의사 등이 있습니다.
3. 논리수학 지능 (Logical-Mathematical Intelligence) 은 수학적 연산과 논리적 사고에 뛰어난 능력으로 문제를 체계적으로 해결하는 데 강점이 있습니다. 예시 직업으로는 수학자, 과학자, 프로그래머, 공자 등이 있습니다.
4. 공간 지능 (Spatial Intelligence) 은 공간에서 물체를 인식하고 시각화하는 능력으로 도형이나 물체의 관계를 파악하고 상상할 수 있는 능력입니다. 예시 직업으로는 예술가, 건축가, 조각가, 디자이너 등이 있습니다.

338 CHAPTER 22. BIG FIVE PERSONALITY AND MULTIPLE INTELLIGENCE

5. 언어 지능 (Linguistic Intelligence)은 언어를 이해하고 표현하는 능력으로 읽기, 쓰기, 말하기, 단어의 의미 이해 능력이 뛰어납니다. 예시 직업으로는 작가, 언어학자, 변호사, 기자 등이 있습니다.
6. 인간친화 지능 (Interpersonal Intelligence)은 타인의 감정, 의도, 동기를 잘 파악하고 효과적으로 소통할 수 있는 능력입니다. 예시 직업으로는 교사, 상담사, 정치인, 마케팅 전문가 등이 있습니다.
7. 자기성찰 지능 (Intrapersonal Intelligence)은 자신의 감정과 욕구를 잘 인식하고 통제할 수 있는 능력으로 자기 이해와 성찰을 통해 목표를 성취하는 데 도움을 줍니다. 예시 직업으로는 철학자, 심리학자, 작가 등이 있습니다.
8. 자연친화 지능 (Naturalistic Intelligence)은 자연 환경, 생물학적 다양성, 동식물에 대한 민감성을 나타내며 생태계와의 상호작용에 관심을 갖습니다. 예시 직업으로는 생물학자, 환경운동가, 농업전문가 등이 있습니다.

22.2.1 집계

	음악	신체운동	논리수학	공간	언어	인간친화	자기성찰	자연친화	계
제1순위	65	50	68	41	56	147	123	6	556
제2순위	29	50	97	51	51	133	123	22	556
제3순위	41	54	100	65	64	117	94	21	556

3대 강점지능의 분포를 살펴 보면 인간친화 지능과 자기성찰 지능이 압도적으로 많이 포함되어 있음을 알 수 있습니다.

총 556명이 응답한 가운데 인간친화 지능이 제1순위에 포함된 사람은 147명, 제2순위에 포함된 사람은 133명, 제3순위에 포함된 사람은 117명으로 모두 397(명)이나 됩니다.

8가지 지능이 3대 강점지능에 골고루 포함된다면 어느 특정 지능이 3대 강점지능에 들어갈 기대인원은 전체인원을 8로 나눈 후 3을 곱해 준 208.5명 정도 기대되는 데 이 보다 훨씬 많은 숫자임을 알 수 있습니다.

그 다음으로 자기성찰 지능이 제1순위에 포함된 사람은 123명, 제2순위에 포함된 사람은 123명, 제3순위에 포함된 사람은 94명으로 모두 340(명)이나 됩니다.

그 다음으로는 논리수학 지능이 눈에 띄는 데 논리수학 지능이 제1순위에 포함된 사람은 68명, 제2순위에 포함된 사람은 97명, 제3순위에 포함된 사람은 100명으로 모두 265(명)입니다.

이제 백분율을 살펴보겠습니다.

22.2.2 %

	음악	신체운동	논리수학	공간	언어	인간친화	자기성찰	자연친화	계
제1순위	11.7	9.0	12.2	7.4	10.1	26.4	22.1	1.1	100.0
제2순위	5.2	9.0	17.4	9.2	9.2	23.9	22.1	4.0	100.0
제3순위	7.4	9.7	18.0	11.7	11.5	21.0	16.9	3.8	100.0

3대 강점지능의 분포를 백분율로 살펴 보아도 인간친화 지능과 자기성찰 지능이 압도적으로 많이 포함되어 있음을 알 수 있습니다.

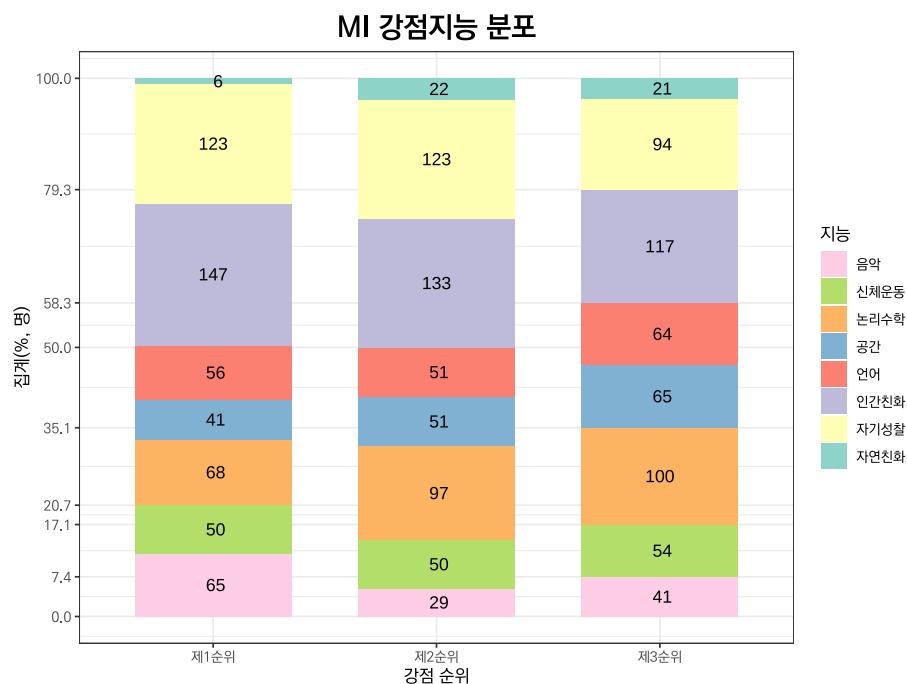
인간친화 지능이 제1순위에 포함된 사람의 백분율은 26.4(%), 제2순위에 포함된 사람은 23.9(%), 제3순위에 포함된 사람은 21.0(%)입니다.

그 다음으로 자기성찰 지능이 제1순위에 포함된 사람은 22.1(%), 제2순위에 포함된 사람은 22.1(%), 제3순위에 포함된 사람은 16.9(%입니다.

그 다음으로 논리수학 지능이 제1순위에 포함된 사람은 12.2(%), 제2순위에 포함된 사람은 17.4(%), 제3순위에 포함된 사람은 18.0(%입니다.

막대그래프로 이 상황을 시각적으로 요약해 보겠습니다.

22.2.3 Bar Plot



제 23 장 12주차 데이터 실험 집계

23.1 실험의 목적

12주차 구글 예습 설문지 집계결과를 분석합니다.

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

Q7에서는 Framing Effects 를 알아 보기 위하여 동일한 내용을 Red에서는 생존을 프레임으로 프로그램을 비교하고, Black에서는 사망을 프레임으로 프로그램을 비교할 때 생존 프레임에서는 확실한 생존 인원을 선호하고, 사망 프레임에서는 운에 기대어 모두 사망하지 않는 프로그램을 선호한다는 것을 보여 줍니다.

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

23.1.1 Red, Black을 잘못 표시한 사람들

	Red(구글예습퀴즈)	Black(구글예습퀴즈)
Red(랜덤화출석부)	345	3
Black(랜덤화출석부)	1	350
계	346	353

23.1.2 응답인원의 Red, Black

Red로 응답한 인원은 346명, Black에 응답한 인원은 353명입니다.

전체 응답인원 699명을 랜덤하게 둘로 나눌 때 어느 한 쪽의 기대인원은 전체 응답인원의 절반인 349.5명이고, 표준오차는 전체 응답인원의 제곱근에 $1/2$ 을 곱해 준 13.2명입니다.

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

23.2 Q1. Latane and Darley

1. Latane 와 Darley 의 연기실험에서 여럿이 있을 때 연기가 난다고 신고한 경우는 얼마나 되었는가? *

- 가. 열에 한명
- 나. 열에 세명
- 다. 열에 다섯명
- 라. 열에 여덟명

23.2.1 방관자 효과

	열에 한명	열에 세명	열에 다섯명	열에 여덟명	계
Red	218	86	33	9	346
Black	207	83	52	11	353
계	425	169	85	20	699

표 23.3: Pearson's Chi-squared test: .

Test statistic	df	P value
4.715	3	0.1939

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

그 결과 카이제곱 통계량은 4.715, 자유도는 3 , p-value 는 0.1939이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

23.2.2 방관자 효과(%)

열에 한명	열에 세명	열에 다섯명	열에 여덟명	계
60.8	24.2	12.2	2.9	100.0

정답률은 Red, Black 을 합하여 계산하는데, 60.8(%) 입니다.

23.3 Q2. Stanley Milgram

2. 스탠리 밀그램의 복종 실험에서 마지막 450V까지 누른 사람은 어느 정도였는가? *

- 가. "15%"
- 나. "25%"
- 다. "45%"
- 라. "65%"

23.3.1 450V까지 누르는 사람

	15%	25%	45%	65%	계
Red	17	48	262	19	346
Black	19	56	259	19	353
계	36	104	521	38	699

표 23.6: Pearson's Chi-squared test: .

Test statistic	df	P value
0.6737	3	0.8794

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

그 결과 카이제곱 통계량은 0.674, 자유도는 3, p-value 는 0.8794이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

23.3.2 450V까지 누르는 사람 (%)

	15%	25%	45%	65%	계
	5.2	14.9	74.5	5.4	100.0

정답률은 Red, Black 을 합하여 계산하는데, 74.5(%) 입니다.

23.4 Q3. Solomon Asch

3. 솔로몬 애쉬의 가짜 시지각 실험에서 공모자가 몇 명일 때 동조율이 가장 급격히 올라가는가? *

- 가. 1명
- 나. 2명
- 다. 3명
- 라. 4명

23.4.1 Power of 3

	1명	2명	3명	4명	계
Red	18	64	47	217	346
Black	23	47	69	214	353
계	41	111	116	431	699

표 23.9: Pearson's Chi-squared test: .

Test statistic	df	P value
7.337	3	0.06189

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

그 결과 카이제곱 통계량은 7.337, 자유도는 3, p-value 는 0.06190이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

23.4.2 Power of 3 (%)

1명	2명	3명	4명	계
5.9	15.9	16.6	61.7	100.0

정답률은 Red, Black 을 합하여 계산하는데, 61.7(%) 입니다.

23.5 Q4. BFI 줄 긋기

1. 성격유형의 영문명칭과 한글명칭을 맞추시오 *					
	외향성	성실성	친화성	신경성	개방성
Openness	<input type="radio"/>				
Conscientious...	<input type="radio"/>				
Extraversion	<input type="radio"/>				
Agreeableness	<input type="radio"/>				
Neuroticism	<input type="radio"/>				

23.5.1 몇 개나 맞추나?

	0	1	2	3	5	계
Red	16	34	16	29	251	346
Black	21	34	27	36	235	353
계	37	68	43	65	486	699

표 23.12: Pearson's Chi-squared test: .

Test statistic	df	P value
4.701	4	0.3194

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

그 결과 카이제곱 통계량은 4.701, 자유도는 4, p-value 는 0.3194이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

23.5.2 몇 개나 맞추나? (%)

	0	1	2	3	5	계
	5.3	9.7	6.2	9.3	69.5	100.0

모두 맞춘 사람들의 백분율은 69.5(%) 입니다.

23.6 Q5. 다중지능 짹 짓기

2. 다중지능의 영문명칭과 한글명칭을 연결하시오. *								
	음악	신체운동	논리수학	공간	언어	인간친화	자기성찰	자연친화
Bodily-kin...	○	○	○	○	○	○	○	○
Musical	○	○	○	○	○	○	○	○
Visual-sp...	○	○	○	○	○	○	○	○
Intrapers...	○	○	○	○	○	○	○	○
Naturalist...	○	○	○	○	○	○	○	○
Logical...	○	○	○	○	○	○	○	○
Verbal-lin...	○	○	○	○	○	○	○	○
Interpers...	○	○	○	○	○	○	○	○

23.6.1 몇 개나 맞추나?

	0	1	2	3	4	5	6	8	계
Red	18	12	10	4	10	15	52	225	346
Black	26	15	8	6	7	12	67	212	353
계	44	27	18	10	17	27	119	437	699

표 23.15: Pearson's Chi-squared test: .

Test statistic	df	P value
5.481	7	0.6015

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

그 결과 카이제곱 통계량은 5.481, 자유도는 7, p-value 는 0.6015이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

23.6.2 몇 개나 맞추나? (%)

0	1	2	3	4	5	6	8	계
6.3	3.9	2.6	1.4	2.4	3.9	17.0	62.5	100.0

모두 맞춘 사람들의 백분율은 62.5(%) 입니다.

23.7 Q6. 성공한 사람들의 공통점

3. 성공한 사람들이 갖고 있는 3대 강점지능 중 거의 공통적으로 발견되는 지능은? *
- 언어
 - 논리수학
 - 인간친화
 - 자기성찰
 - 자연친화

23.7.1 자기성찰지능의 중요성

	언어	논리수학	인간친화	자기성찰	자연친화	계
Red	24	16	49	257	0	346
Black	35	15	68	231	4	353
계	59	31	117	488	4	699

표 23.18: Pearson's Chi-squared test: .

Test statistic	df	P value
6.541	3	0.08806

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

그 결과 카이제곱 통계량은 6.541, 자유도는 3, p-value 는 0.08810이므로 Red, Black 간에 통계적으로 유의한 차이를 (보이고 있습니다.) 보이지 않습니다.

실제로 닮은 게 느껴집니까?

23.7.2 자기성찰지능 (%)

언어	논리수학	인간친화	자기성찰	자연친화	계
8.44	4.43	16.74	69.81	0.57	100.00

정답률은 Red, Black 을 합하여 계산하는데, 69.8(%) 입니다.

23.8 Q7. Framing Effect : 삶과 죽음

이 질문은 프레임을 어떻게 설정하느냐에 따라 반응이 확연하게 달라진다는 Framing effects를 보여 줍니다.

600명 중 200명이 사는 거나 400명이 죽는 것, 모두 살리는 거나 모두 죽지 않는 것이나 같은 내용임에도 반응이 다르게 나타납니다.

생존 프레임에서는 확실한 생존자 수효율 보장을 프로그램을 선호하고, 사망 프레임에서는 모두 사망하지 않기를 운에 기대는 프로그램을 선호하는 반응이 많다는 것을 확인합니다.

Red

7. (Thinking, Fast and Slow) 미국 정부는 아시아에서 발생한 희귀병으로 600명이 사망할 *
것이라고 예상하고, 이 질병을 박멸하려 한다. 그러기 위해 2개의 프로그램이 물량에 올랐다.
어느 쪽이 더 희망적인가. 이 병의 생사에 대한 확률은 과학적으로 정확하다. 다음의 선택 대안
에서 당신은 어느 것을 선택할 것인가?

가. 200명은 살린다.
 나. 600명 모두 살릴 확률 1/3, 한 명도 못 살릴 확률 2/3

Black

7. (Thinking, Fast and Slow) 미국 정부는 아시아에서 발생한 희귀병으로 600명이 사망할 *
것이라고 예상하고, 이 질병을 박멸하려 한다. 그러기 위해 2개의 프로그램이 물량에 올랐다.
어느 쪽이 더 희망적인가. 이 병의 생사에 대한 확률은 과학적으로 정확하다. 다음의 선택 대안
에서 당신은 어느 것을 선택할 것인가?

가. 400명은 죽는다.
 나. 한 명도 사망하지 않을 확률 1/3, 600명이 사망할 확률 2/3

23.8.1 집계

	확실히	확률적으로	계
생존 프레임	208	138	346
사망 프레임	119	234	353
계	327	372	699

표 23.21: Pearson's Chi-squared test with Yates' continuity correction: .

Test statistic	df	P value
47.88	1	4.537e-12 ***

Q7의 Red는 생존 프레임으로 600명 중 200명을 확실히 살리는 프로그램하고 600명 모두 살릴 확률은 1/3이고 모두 사망 확률은 2/3로 사실상 살릴 수 있는 기대인원은 200명으로 같지만 확실히 살린다는 프로그램을 더 선호하고, Black은 사망 프레임으로 Red 와 동일한 프로그램임에도 설명하기를 600명 중 400명이 확실히 죽는 프로그램하고 모두 사망하지 않을 확률은 1/3이고, 모두 사망할 확률은 2/3인 프로그램 중에서 선택하라고 했을 때 사망 프레임이 씌워져 있는 프로그램 중에서는 그나마 운에 기댈 수 있는 프로그램을 더 선호한다는 것을 명확히 보여 줍니다.

모두 살린다는 것이나 모두 사망하지 않는다는 것은 같은 말입니다.

그 결과 Red, Black 의 차이를 분석하기 위한 카이제곱 통계량은 47.877, 자유도는 1, p-value 는 4.5e-12 으로 프레임의 차이가 통계적으로 유의함을 보여 줍니다.

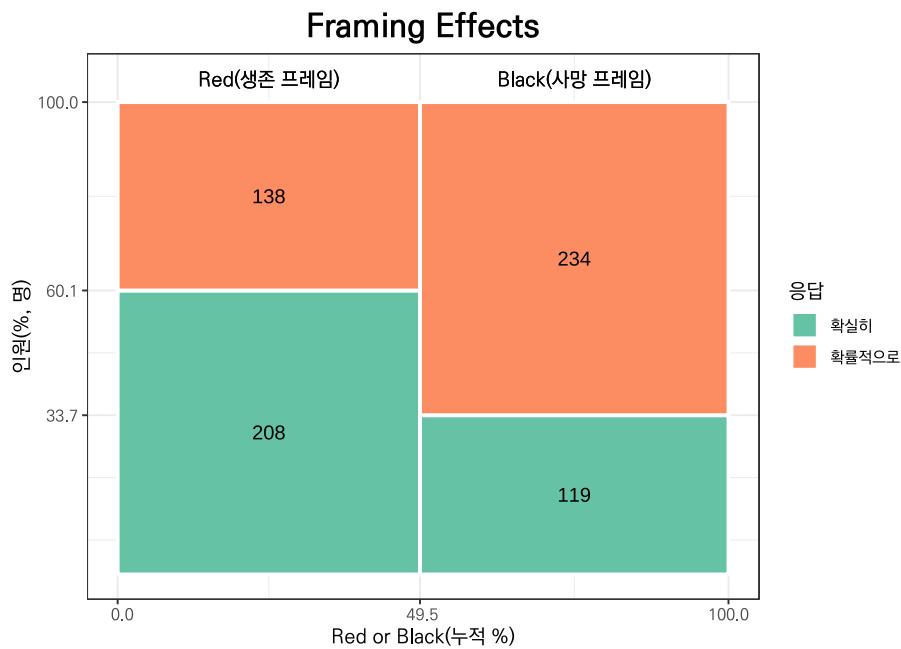
즉, Framing Effects가 확실히 영향을 주고 있는 것으로 파악됩니다.

23.8.2 % 비교.

	확실히	확률적으로	계
생존 프레임	60.1	39.9	100.0
사망 프레임	33.7	66.3	100.0

이를 백분율로 살펴보면 생존 프레임인 Red에서 확실히 200명을 살리는 프로그램을 선호하는 백분율, 60.1%(은)는 운에 기대어 모두 살리는 프로그램을 선호하는 백분율, 39.9% 보다 높고, 사망 프레임인 Black에서 확실히 400명이 사망한다는 프로그램을 선호하는 백분율, 33.7%(은)는 운에 기대어 모두 사망하지 않을 프로그램을 선호하는 백분율, 66.3%보다 낮다는 것을 알 수 있습니다.

23.8.3 Mosaic Plot



Mosaic Plot 은 이 집계결과를 시각적으로 잘 보여줍니다.

생존 프레임과 사망 프레임에 따라 프로그램의 선호가 달라짐을 명확히 보여줍니다.

23.9 마감 시간으로부터 제출 시간의 분포

23.9.1 분포표

표 23.23: 일 단위

	14일	13일	12일	11일	10일	9일	8일	7일	6일	5일	4일	3일	2일	1일	계
Red	106	30	11	10	15	6	4	41	24	15	17	22	13	32	346
Black	134	27	7	13	7	5	11	37	20	17	15	14	19	27	353
계	240	57	18	23	22	11	15	78	44	32	32	36	32	59	699

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.

23.9.2 날마다 고르게 제출하는가?

14일	13일	12일	11일	10일	9일	8일	7일	6일	5일	4일	3일	2일	1일
240	57	18	23	22	11	15	78	44	32	32	36	32	59

표 23.25: Chi-squared test for given probabilities: .

Test statistic	df	P value
871.3	13	7.397e-178 * * *

날마다 고르게 제출하는지 알아 보았습니다.

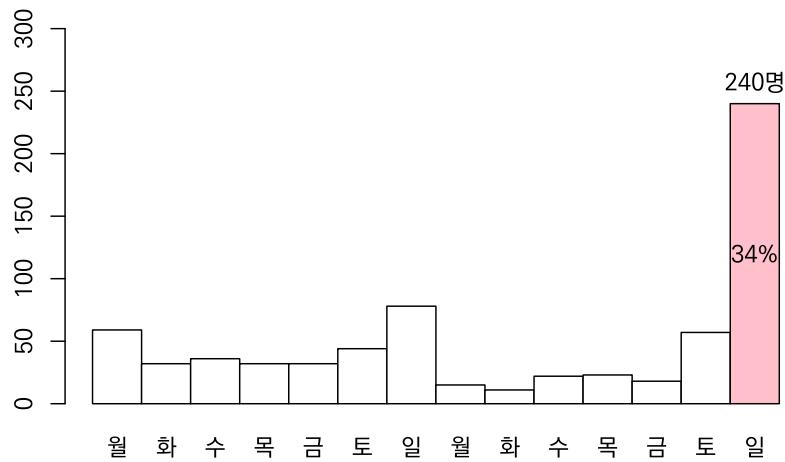
분포표의 “계”행에서 ‘계’열을 제외하고 카이제곱테스트를 수행합니다.

분포표 만으로도 쉽게 파악할 수 있지만 카이제곱테스트가 명확히 해 줍니다. 카이제곱 통계량은 871.263, 자유도는 13.00, p-value 는 7.4e-178 이므로 날짜 별로 고르게 제출하지 않았다는 점을 강력히 시사합니다.

막대그래프로 살펴 보겠습니다.

23.9.3 막대그래프

Quiz241118 (699명 제출)



23.9.4 Red, Black 간에 닮았는가?

	14일	13일	12일	11일	10일	9일	8일	7일	6일	5일	4일	3일	2일	1일
Red	106	30	11	10	15	6	4	41	24	15	17	22	13	32
Black	134	27	7	13	7	5	11	37	20	17	15	14	19	27

표 23.27: Pearson's Chi-squared test: .

Test statistic	df	P value
15.05	13	0.3044

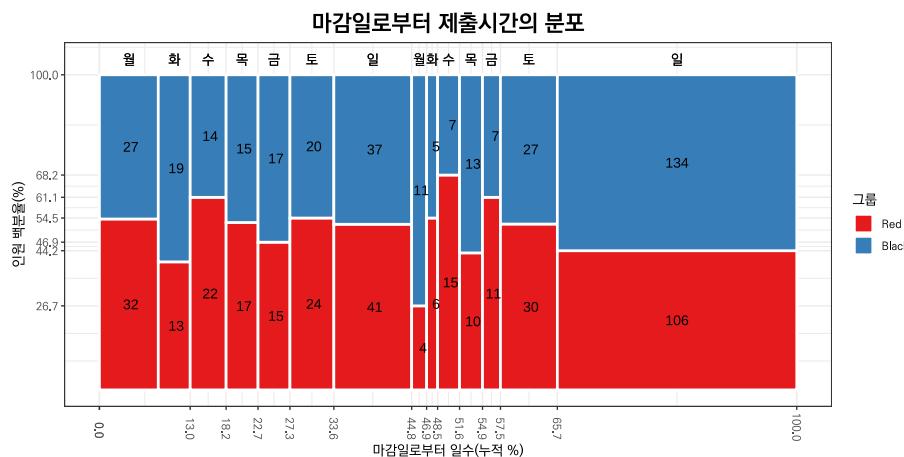
제출시간의 분포가 Red, Black 간에 닮았는지 알아 보았습니다.

이번에는 분포표의 첫번째와 두번째 행, '계'열을 제외한 나머지 열에 대해서 카이제곱테스트를 수행합니다. 카이제곱 통계량은 15.05, 자유도는 13, p-value는 0.3044 이므로 제출 시간의 분포는 Red, Black 간에 통계적으로 유의한 차이가 관찰되지 않습니다.

이 사실을 Mosaic Plot 을 이용하여 시각적으로 살펴보겠습니다.

답았다고 느껴지나요?

23.9.5 Mosaic Plot



제 24 장 12주차 데이터 실험 집계

24.1 실험의 목적

12주차 구글 예습 설문지 집계결과를 분석합니다.

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

Q7에서는 Framing Effects 를 알아 보기 위하여 동일한 내용을 Red에서는 생존을 프레임으로 프로그램을 비교하고, Black에서는 사망을 프레임으로 프로그램을 비교할 때 생존 프레임에서는 확실한 생존 인원을 선호하고, 사망 프레임에서는 운에 기대어 모두 사망하지 않는 프로그램을 선호한다는 것을 보여 줍니다.

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

24.1.1 Red, Black을 잘못 표시한 사람들

	Red(구글예습퀴즈)	Black(구글예습퀴즈)
Red(랜덤화출석부)	280	0
Black(랜덤화출석부)	0	284
계	280	284

24.1.2 응답인원의 Red, Black

Red로 응답한 인원은 280명, Black에 응답한 인원은 284명입니다.

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 안에 들어갑니다.

24.2 Q1. Latane and Darley

1. Latane 와 Darley 의 연기실험에서 여럿이 있을 때 연기가 난다고 신고한 경우는 얼마나 되었는가? *

- 가. 열에 한명
- 나. 열에 세명
- 다. 열에 다섯명
- 라. 열에 여덟명

24.2.1 방관자 효과

	열에 한명	열에 세명	열에 다섯명	열에 여덟명	계
Red	173	71	29	9	282
Black	165	72	37	10	284
계	338	143	66	19	566

표 24.3: Pearson's Chi-squared test: .

Test statistic	df	P value
1.212	3	0.7502

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

그 결과 카이제곱 통계량은 1.212, 자유도는 3 , p-value 는 0.7502이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

24.2.2 방관자 효과(%)

열에 한명	열에 세명	열에 다섯명	열에 여덟명	계
59.7	25.3	11.7	3.4	100.0

정답률은 Red, Black 을 합하여 계산하는데, 59.7(%) 입니다.

24.3 Q2. Stanley Milgram

2. 스탠리 밀그램의 복종 실험에서 마지막 450V까지 누른 사람은 어느 정도였는가? *

- 가. "15%"
- 나. "25%"
- 다. "45%"
- 라. "65%"

24.3.1 450V까지 누르는 사람

	15%	25%	45%	65%	계
Red	12	43	212	15	282
Black	13	43	216	12	284
계	25	86	428	27	566

표 24.6: Pearson's Chi-squared test: .

Test statistic	df	P value
0.4037	3	0.9395

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

그 결과 카이제곱 통계량은 0.404, 자유도는 3, p-value 는 0.9395이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

24.3.2 450V까지 누르는 사람 (%)

	15%	25%	45%	65%	계
	4.4	15.2	75.6	4.8	100.0

정답률은 Red, Black 을 합하여 계산하는데, 75.6(%) 입니다.

24.4 Q3. Solomon Asch

3. 솔로몬 애쉬의 가짜 시지각 실험에서 공모자가 몇 명일 때 동조율이 가장 급격히 올라가는가? *

- 가. 1명
- 나. 2명
- 다. 3명
- 라. 4명

24.4.1 Power of 3

	1명	2명	3명	4명	계
Red	32	40	42	168	282
Black	20	39	45	180	284
계	52	79	87	348	566

표 24.9: Pearson's Chi-squared test: .

Test statistic	df	P value
3.292	3	0.3487

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

그 결과 카이제곱 통계량은 3.292, 자유도는 3, p-value 는 0.3487이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

24.4.2 Power of 3 (%)

1명	2명	3명	4명	계
9.2	14.0	15.4	61.5	100.0

정답률은 Red, Black 을 합하여 계산하는데, 61.5(%) 입니다.

24.5 Q4. BFI 줄 긋기

1. 성격유형의 영문명칭과 한글명칭을 맞추시오 *					
	외향성	성실성	친화성	신경성	개방성
Openness	<input type="radio"/>				
Conscientious...	<input type="radio"/>				
Extraversion	<input type="radio"/>				
Agreeableness	<input type="radio"/>				
Neuroticism	<input type="radio"/>				

24.5.1 몇 개나 맞추나?

	0	1	2	3	5	계
Red	17	38	19	26	182	282
Black	18	31	20	36	179	284
계	35	69	39	62	361	566

표 24.12: Pearson's Chi-squared test: .

Test statistic	df	P value
2.395	4	0.6635

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

그 결과 카이제곱 통계량은 2.395, 자유도는 4, p-value 는 0.6635이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

24.5.2 몇 개나 맞추나? (%)

	0	1	2	3	5	계
	6.2	12.2	6.9	11.0	63.8	100.0

모두 맞춘 사람들의 백분율은 63.8(%) 입니다.

24.6 Q5. 다중지능 짹 짓기

2. 다중지능의 영문명칭과 한글명칭을 연결하시오. *								
	음악	신체운동	논리수학	공간	언어	인간친화	자기성찰	자연친화
Bodily-kin...	○	○	○	○	○	○	○	○
Musical	○	○	○	○	○	○	○	○
Visual-sp...	○	○	○	○	○	○	○	○
Intrapers...	○	○	○	○	○	○	○	○
Naturalist...	○	○	○	○	○	○	○	○
Logical...	○	○	○	○	○	○	○	○
Verbal-lin...	○	○	○	○	○	○	○	○
Interpers...	○	○	○	○	○	○	○	○

24.6.1 몇 개나 맞추나?

	0	1	2	3	4	5	6	8	계
Red	29	11	6	4	10	9	56	157	282
Black	15	13	11	6	11	9	53	166	284
계	44	24	17	10	21	18	109	323	566

표 24.15: Pearson's Chi-squared test: .

Test statistic	df	P value
6.866	7	0.443

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

그 결과 카이제곱 통계량은 6.866, 자유도는 7, p-value 는 0.44300이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

실제로 닮은 게 느껴집니까?

24.6.2 몇 개나 맞추나? (%)

	0	1	2	3	4	5	6	8	계
	7.8	4.2	3.0	1.8	3.7	3.2	19.3	57.1	100.0

모두 맞춘 사람들의 백분율은 57.1(%) 입니다.

24.7 Q6. 성공한 사람들의 공통점

3. 성공한 사람들이 갖고 있는 3대 강점지능 중 거의 공통적으로 발견되는 지능은? *
- 언어
 - 논리수학
 - 인간친화
 - 자기성찰
 - 자연친화

24.7.1 자기성찰지능의 중요성

	언어	논리수학	인간친화	자기성찰	자연친화	계
Red	20	18	39	198	7	282
Black	23	15	46	196	4	284
계	43	33	85	394	11	566

표 24.18: Pearson's Chi-squared test: .

Test statistic	df	P value
1.88	4	0.7579

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

그 결과 카이제곱 통계량은 1.880, 자유도는 4, p-value 는 0.75790이므로 Red, Black 간에 통계적으로 유의한 차이를 (보이고 있습니다.) 보이지 않습니다.

실제로 닮은 게 느껴집니까?

24.7.2 자기성찰지능 (%)

언어	논리수학	인간친화	자기성찰	자연친화	계
7.6	5.8	15.0	69.6	1.9	100.0

정답률은 Red, Black 을 합하여 계산하는데, 69.6(%) 입니다.

24.8 Q7. Framing Effect : 삶과 죽음

이 질문은 프레임을 어떻게 설정하느냐에 따라 반응이 확연하게 달라진다는 Framing effects를 보여 줍니다.

600명 중 200명이 사는 거나 400명이 죽는 것, 모두 살리는 거나 모두 죽지 않는 것이나 같은 내용임에도 반응이 다르게 나타납니다.

생존 프레임에서는 확실한 생존자 수효율 보장을 프로그램을 선호하고, 사망 프레임에서는 모두 사망하지 않기를 운에 기대는 프로그램을 선호하는 반응이 많다는 것을 확인합니다.

Red

7. (Thinking, Fast and Slow) 미국 정부는 아시아에서 발생한 희귀병으로 600명이 사망할 *
것이라고 예상하고, 이 질병을 박멸하려 한다. 그러기 위해 2개의 프로그램이 물량에 올랐다.
어느 쪽이 더 희망적인가. 이 병의 생사에 대한 확률은 과학적으로 정확하다. 다음의 선택 대안
에서 당신은 어느 것을 선택할 것인가?

가. 200명은 살린다.
 나. 600명 모두 살릴 확률 1/3, 한 명도 못 살릴 확률 2/3

Black

7. (Thinking, Fast and Slow) 미국 정부는 아시아에서 발생한 희귀병으로 600명이 사망할 *
것이라고 예상하고, 이 질병을 박멸하려 한다. 그러기 위해 2개의 프로그램이 물량에 올랐다.
어느 쪽이 더 희망적인가. 이 병의 생사에 대한 확률은 과학적으로 정확하다. 다음의 선택 대안
에서 당신은 어느 것을 선택할 것인가?

가. 400명은 죽는다.
 나. 한 명도 사망하지 않을 확률 1/3, 600명이 사망할 확률 2/3

24.8.1 집계

	확실히	확률적으로	계
생존 프레임	179	103	282
사망 프레임	100	184	284
계	279	287	566

표 24.21: Pearson's Chi-squared test with Yates' continuity correction: .

Test statistic	df	P value
44.1	1	3.121e-11 ***

Q7의 Red는 생존 프레임으로 600명 중 200명을 확실히 살리는 프로그램하고 600명 모두 살릴 확률은 1/3이고 모두 사망 확률은 2/3로 사실상 살릴 수 있는 기대인원은 200명으로 같지만 확실히 살린다는 프로그램을 더 선호하고, Black은 사망 프레임으로 Red 와 동일한 프로그램임에도 설명하기를 600명 중 400명이 확실히 죽는 프로그램하고 모두 사망하지 않을 확률은 1/3이고, 모두 사망할 확률은 2/3인 프로그램 중에서 선택하라고 했을 때 사망 프레임이 씌워져 있는 프로그램 중에서는 그나마 운에 기댈 수 있는 프로그램을 더 선호한다는 것을 명확히 보여 줍니다.

모두 살린다는 것이나 모두 사망하지 않는다는 것은 같은 말입니다.

그 결과 Red, Black 의 차이를 분석하기 위한 카이제곱 통계량은 44.100, 자유도는 1, p-value 는 3.1e-11 으로 프레임의 차이가 통계적으로 유의함을 보여 줍니다.

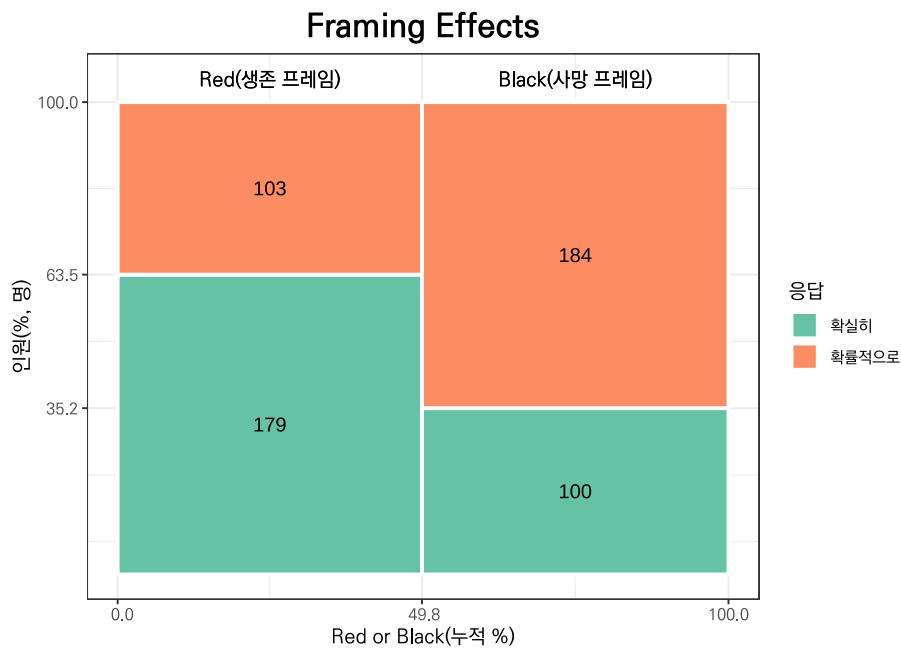
즉, Framing Effects가 확실히 영향을 주고 있는 것으로 파악됩니다.

24.8.2 % 비교.

	확실히	확률적으로	계
생존 프레임	63.5	36.5	100.0
사망 프레임	35.2	64.8	100.0

이를 백분율로 살펴보면 생존 프레임인 Red에서 확실히 200명을 살리는 프로그램을 선호하는 백분율, 63.5%(은)는 운에 기대어 모두 살리는 프로그램을 선호하는 백분율, 36.5% 보다 높고, 사망 프레임인 Black에서 확실히 400명이 사망한다는 프로그램을 선호하는 백분율, 35.2%(은)는 운에 기대어 모두 사망하지 않을 프로그램을 선호하는 백분율, 64.8%보다 낮다는 것을 알 수 있습니다.

24.8.3 Mosaic Plot



Mosaic Plot 은 이 집계결과를 시각적으로 잘 보여줍니다.

생존 프레임과 사망 프레임에 따라 프로그램의 선호가 달라짐을 명확히 보여줍니다.

24.9 마감 시간으로부터 제출 시간의 분포

24.9.1 분포표

표 24.23: 일 단위

	14일	13일	12일	11일	10일	9일	8일	7일	6일	5일	4일	3일	2일	1일	계
Red	128	19	9	6	4	3	5	32	10	9	14	8	15	20	282
Black	133	21	14	4	1	3	6	30	7	13	8	10	11	23	284
계	261	40	23	10	5	6	11	62	17	22	22	18	26	43	566

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.

24.9.2 날마다 고르게 제출하는가?

14일	13일	12일	11일	10일	9일	8일	7일	6일	5일	4일	3일	2일	1일
261	40	23	10	5	6	11	62	17	22	22	18	26	43

표 24.25: Chi-squared test for given probabilities: .

Test statistic	df	P value
1375	13	3.294e-286 * * *

날마다 고르게 제출하는지 알아 보았습니다.

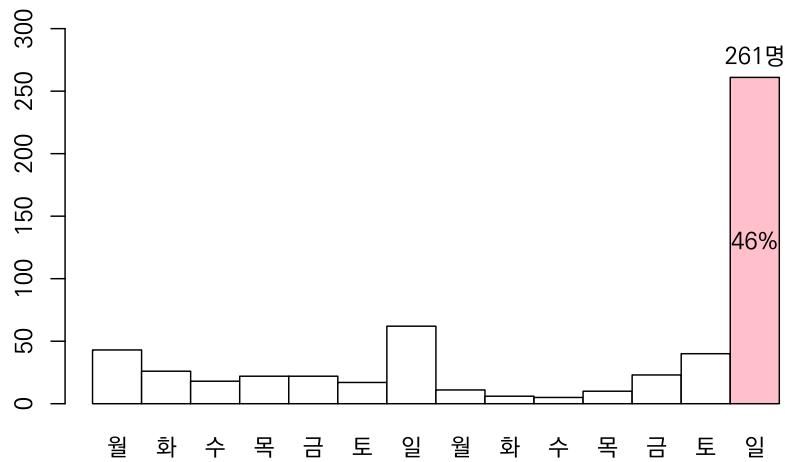
분포표의 “계”행에서 ‘계’열을 제외하고 카이제곱테스트를 수행합니다.

분포표 만으로도 쉽게 파악할 수 있지만 카이제곱테스트가 명확히 해 줍니다. 카이제곱 통계량은 1375.251, 자유도는 13.00, p-value 는 3.3e-286 이므로 날짜 별로 고르게 제출하지 않았다는 점을 강력히 시사합니다.

막대그래프로 살펴 보겠습니다.

24.9.3 막대그래프

Quiz250519 (566명 제출)



24.9.4 Red, Black 간에 닮았는가?

	14일	13일	12일	11일	10일	9일	8일	7일	6일	5일	4일	3일	2일	1일
Red	128	19	9	6	4	3	5	32	10	9	14	8	15	20
Black	133	21	14	4	1	3	6	30	7	13	8	10	11	23

표 24.27: Pearson's Chi-squared test: .

Test statistic	df	P value
7.571	13	0.8704

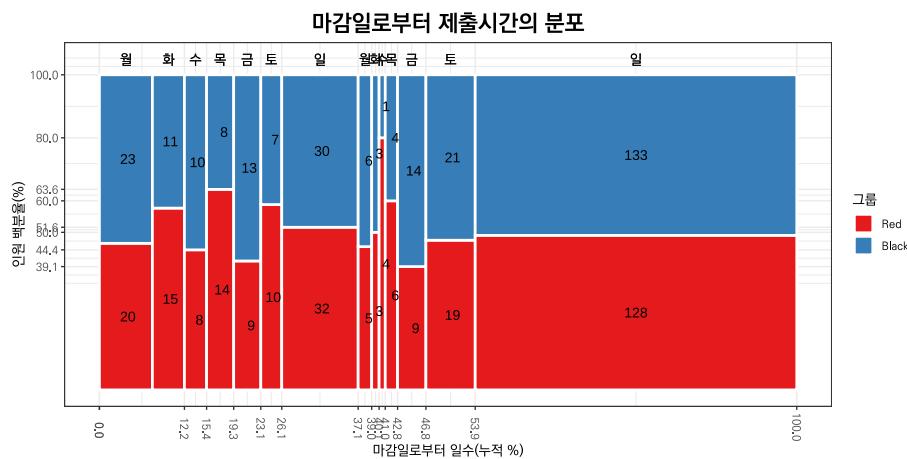
제출시간의 분포가 Red, Black 간에 닮았는지 알아 보았습니다.

이번에는 분포표의 첫번째와 두번째 행, '계'열을 제외한 나머지 열에 대해서 카이제곱테스트를 수행합니다. 카이제곱 통계량은 7.57, 자유도는 13, p-value는 0.8704 이므로 제출 시간의 분포는 Red, Black 간에 통계적으로 유의한 차이가 관찰되지 않습니다.

이 사실을 Mosaic Plot 을 이용하여 시각적으로 살펴보겠습니다.

답았다고 느껴지나요?

24.9.5 Mosaic Plot



제 25 장 13주차 데이터 실험 집계

25.1 실험의 목적

13주차 구글 예습 설문지 집계결과를 분석합니다.

Q1에서는 잘 알려진 생일 문제를 수강생들에게 적용하여 보았습니다.

상당히 많은 인원이기 때문에 한쌍 이상의 생일이 같은 확률은 당연히 1이고 생일이 같은 사람이 얼마나 되는지 알아보겠습니다.

이론적으로 기대되는 인원과 표준오차를 계산하여 실제 관찰된 인원과 비교합니다.

생일의 월별분포에 대해서도 분석합니다.

Q2에서는 맷칭문제의 사례로 가수와 노래를 짹짓는 문제를 살펴봅니다.

랜덤하게 고르도록 하였는데 왜 이론적으로 기대하는 분포와 다른 결과가 나오는 것인지 생각해 봅니다.

Q3에서는 연비라는 용어가 직관적으로 주는 오류에 대해서 알아 봅니다.

연비가 높은 차를 연비가 좀 더 높은 차로 바꾸는 것과 연비가 낮은 차를 연비가 좀 더 높은 차로 바꾸는 것 중에서 어떤 선택이 좀더 연료비를 절감할 수 있는지 알아 봅니다.

Q4에서는 잘 알려진 Monty Hall 문제를 수강생들에게 물어 본 결과를 분석합니다.

Red 와 Black 은 “고수한다”와 “바꾼다”的 순서를 바꿔 보았는데 그 효과는 통계적으로 유의한 차이를 보이지 않습니다.

그리고 여러분들이 직관적으로 고른 답은 대부분 정답이 아닌 것으로 드러납니다.

25.1.1 Red, Black을 잘못 표시한 사람들

	Red(구글예습퀴즈)	Black(구글예습퀴즈)
Red(랜덤화출석체부)	354	3
Black(랜덤화출석체부)	1	360
계	355	363

Red(구글예습퀴즈)	Black(구글예습퀴즈)
-------------	---------------

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

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

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, 3명이나 Black 을 Red 라고 한 사람수, 1명은 기대인원으로부터 표준오차 범위에 아주 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.

다만, 이번 실험과 같이 3명씩 동일한 결과가 나온 경우에는 중복되는 확률을 빼 주어야 합니다.

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.

25.1.2 응답인원의 Red, Black

Red 로 응답한 인원은 355명, Black 에 응답한 인원은 363명입니다.

전체 응답인원 718 명을 랜덤하게 둘로 나눌 때 어느 한 쪽의 기대인원은 전체 응답인원의 절반인 359명이고, 표준오차는 전체 응답인원의 제곱근에 $1/2$ 을 곱해 준 13.4 명입니다.

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

25.2 Q1. Birthday Problem

1. 나의 생일은 _____ 월 _____ 일이다 *
예시 : 7월7일(X), 7/7(X), 0707(X), 07/07(0)

내 답변

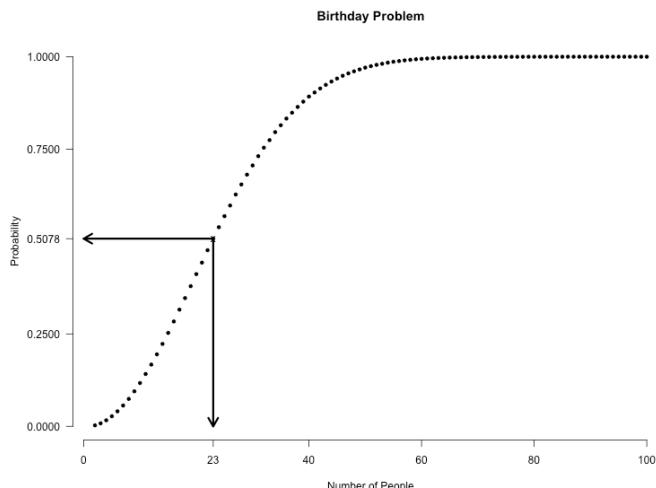
Q1은 생일 문제라고 잘 알려져 있습니다.

23명만 모여 있어도 생일이 같은 사람이 한쌍 이상 있을 확률이 $1/2$ 을 넘어갑니다.

다음 그림은 1에서 365의 숫자 중 x-축에 나온 숫자 만큼 랜덤하게 뽑아서 (복원 추출) 같은 숫자가 있는지 여부를 백만번씩 돌려서 나온 결과들입니다.

23명을 뽑았을 때 백만번 중에 50만7천8백번 같은 숫자가 나왔다는 뜻입니다.

큰수의 법칙에 따라서 그 비율은 23명 중에 생일이 같은 사람이 최소한 한 쌍 있을 확률로 수렴해 갑니다.



다음은 집계결과로부터 어느 날짜에 몇 명의 생일이 같은지 순서대로 정리한 것입니다.

가끔 이 결과를 보고 “나도 제출했는데 내 생일이 빠졌다”고 의문을 제기하는 사람들이 있습니다만 그 사람과 생일 같은 사람이 없는 것 뿐입니다.

집계가 진행되면서 당연히도 생일 같은 사람들의 수효는 늘어나게 마련입니다.

그럼에도 그 인원은 이론적으로 계산한 기대인원과 표준오차 범위에 거의 들어가는 것을 관찰할 수 있습니다.

보고서의 맨 위에 집계 시점이 나옵니다.

시간 경과에 따라 생일이 같은 사람의 수효가 늘어나는 것과 그 수효가 이론적으로 예측한 범위에 들어가는지 살펴보기 바랍니다.

25.2.1 어느 날에 몇 명씩 생일이 같은가?

01월01일	2
01월03일	4
01월04일	3
01월05일	5
01월06일	5
01월08일	4
01월09일	3
01월10일	4
01월16일	3
01월19일	2
01월20일	2
01월21일	2
01월22일	4
01월23일	4
01월24일	3
01월30일	2
01월31일	2
02월01일	4
02월03일	3
02월04일	3
02월05일	2
02월06일	2
02월07일	5
02월08일	3
02월09일	3
02월10일	2
02월11일	2
02월12일	4
02월16일	3
02월18일	2
02월19일	2
02월20일	5
02월21일	3

02월22일	2
02월23일	3
03월01일	2
03월03일	2
03월04일	4
03월05일	2
03월06일	4
03월08일	2
03월10일	2
03월12일	3
03월15일	4
03월16일	2
03월17일	2
03월18일	3
03월19일	7
03월20일	2
03월21일	2
03월22일	5
03월23일	3
03월24일	3
03월25일	2
03월27일	2
03월30일	3
04월02일	2
04월04일	4
04월05일	2
04월06일	4
04월07일	3
04월08일	2
04월10일	3
04월13일	2
04월14일	2
04월17일	5
04월19일	4
04월20일	2
04월21일	6
04월22일	2
04월23일	3
04월24일	2
04월27일	2
04월28일	5
04월29일	4
04월30일	2
05월03일	2
05월06일	3

05월07일	3
05월08일	2
05월09일	2
05월10일	2
05월15일	2
05월16일	3
05월17일	2
05월18일	3
05월21일	2
05월23일	3
05월24일	3
05월26일	2
05월28일	2
05월30일	5
06월03일	3
06월04일	3
06월06일	2
06월07일	4
06월10일	2
06월11일	4
06월13일	2
06월14일	3
06월15일	4
06월16일	3
06월17일	2
06월18일	2
06월19일	3
06월21일	2
06월22일	2
06월25일	2
06월26일	2
06월27일	4
06월28일	2
06월29일	2
06월30일	2
07월04일	2
07월05일	4
07월06일	2
07월07일	4
07월09일	3
07월10일	3
07월11일	3
07월13일	3
07월16일	2
07월17일	4

07월18일	2
07월23일	2
07월25일	2
07월26일	5
07월27일	2
07월30일	2
08월02일	3
08월04일	2
08월05일	2
08월07일	2
08월08일	2
08월09일	2
08월10일	2
08월11일	2
08월12일	3
08월13일	3
08월15일	3
08월16일	2
08월18일	3
08월19일	3
08월23일	3
08월24일	2
08월26일	3
08월27일	4
08월30일	4
08월31일	4
09월01일	3
09월02일	3
09월06일	2
09월09일	3
09월10일	3
09월13일	3
09월15일	3
09월16일	3
09월17일	2
09월20일	5
09월23일	7
09월26일	2
10월03일	2
10월05일	3
10월06일	4
10월07일	2
10월09일	2
10월10일	4
10월11일	2

10월12일	5
10월13일	3
10월14일	2
10월16일	2
10월19일	3
10월20일	2
10월21일	2
10월26일	4
10월27일	2
10월28일	2
10월29일	2
10월30일	3
11월01일	2
11월02일	4
11월04일	4
11월06일	2
11월07일	4
11월08일	2
11월11일	3
11월12일	4
11월13일	2
11월14일	4
11월15일	2
11월17일	3
11월18일	3
11월20일	2
11월21일	3
11월24일	3
11월25일	4
11월27일	3
11월28일	3
11월30일	4
12월02일	3
12월03일	2
12월06일	3
12월07일	4
12월10일	2
12월11일	3
12월12일	3
12월14일	3
12월15일	4
12월18일	3
12월20일	3
12월21일	2
12월22일	3

12월26일	3
12월27일	2
12월28일	2
12월30일	2
계	622

25.2.2 생일이 같은 사람은 몇 명 정도 기대되는가?

전체 응답인원 719(명) 중에 생일이 같은 사람은 622(명)이고 생일이 같은 날은 217(일)입니다.

N 을 전체 인원이라 할 때, 기대 인원은 $N \times \{1 - (\frac{364}{365})^{N-1}\}$, 분산은 $N \times \{1 - (\frac{364}{365})^{N-1}\} + N \times (N-1) \times \{1 - (\frac{363}{365})^{N-2}\}$ 로 계산됩니다.

무응답이거나 결석한 학생을 제외한 응답 인원 719명에 대하여 공식에 따라 기대인원을 계산하면 618.7명, 표준오차는 24.9명으로 계산되어 생일이 같은 사람들의 수효 622(명)은 기대인원으로부터 표준오차, 혹은 두 배의 표준오차 범위 안에 잘 들어감을 알 수 있습니다.

25.2.2.1 기대되는 인원

기대인원
618.7

25.2.2.2 표준오차

표준오차
24.9

이전 학기 자료들에서는 10명이 생일이 같은 경우도 있었고, 8명이 생일이 같은 경우는 여럿 나오기도 했는데 이번 학기에는 03월19일, 09월23일에 7(명)의 생일이 같습니다.

2024년 1학기, 2023년 1, 2학기, 2022년 1, 2학기, 2021년 1, 2학기 모두 이론적으로 기대하는 값과 관찰값이 잘 들어 맞았습니다.

2020년 1~2학기에는 기대에 약간 못 미치는 인원을 관찰하였지만 통상적으로 얘기하는 표준오차의 두배 이내에는 잘 들어맞는 인원입니다.

여러분의 생일은 몇 명이나 같은 사람이 있나요?

이론적으로 기대하는 인원과 실제 관찰된 인원이 잘 부합한다는 점에 대해서 어떤 생각이 듭니까?

25.2.3 태어난 달의 분포는?

제출한 생일 날짜들을 월별로 정리하였습니다.

아래 교차표로부터 두 가지 질문을 던져볼 수 있겠습니다.

응답자들은 월별로 고르게 출생하였을까?

출생한 달의 분포는 Red, Black 간에 닮았는가?

이 질문에 답하기 위해서 카이제곱 테스트를 수행합니다.

	1월	2월	3월	4월	5월	6월	7월	8월	9월	10월	11월	12월	계
Red	32	34	39	39	23	20	25	25	20	30	42	27	356
Black	31	27	28	26	26	41	29	39	29	33	26	28	363
계	63	61	67	65	49	61	54	64	49	63	68	55	719

월별로 고르게 출생하였는지 알아보려면 위의 교차표에서 “계” 행의 1월부터 12월까지를
잘라 내어 카이제곱 균일성 테스트를 수행해야 합니다.

이 때 1월부터 12월까지 12개의 범주가 있으니까 자유도는 하나를 뺀 11이 됩니다.

25.2.4 월별로 고르게 출생하였는가?

1월	2월	3월	4월	5월	6월	7월	8월	9월	10월	11월	12월
63	61	67	65	49	61	54	64	49	63	68	55

표 25.6: Chi-squared test for given probabilities: .

Test statistic	df	P value
7.96	11	0.7169

월별로 고르게 출생하였는지 알아보기 위하여 수행한 카이제곱 테스트에서 카이제곱 통계량은 7.960, 자유도는 11.00, p-value 는 0.7169 이므로 월별로 고르게 출생하고 있음을 시사합니다.

가장 많이 태어난 달은 11월에 68(명)이고 가장 적게 태어난 달은 5월, 9월에 49, 49(명)입니다.

태어난 달의 분포가 Red, Black 간에 닮았는지 살펴보기 위하여 “계”행과 “계”열을 삭제한 다음 교차표에 카이제곱 테스트를 수행합니다.

25.2.5 Red and Black

	1월	2월	3월	4월	5월	6월	7월	8월	9월	10월	11월	12월
Red	32	34	39	39	23	20	25	25	20	30	42	27
Black	31	27	28	26	26	41	29	39	29	33	26	28

표 25.8: Pearson's Chi-squared test: .

Test statistic	df	P value
21.51	11	0.02846 *

Red, Black 간에 월별 출생의 분포가 닮았는지 알아보기 위하여 수행한 카이제곱 테스트에서 카이제곱 통계량은 21.510, 자유도는 11, p-value 는 0.0285 이므로 Red, Black 간에 출생의 분포는 통계적으로 유의한 차이를 보이고 있습니다.

여기서 자유도는 $(2 - 1) \times (12 - 1) = 11$ 로 계산한 것입니다.

앞의 출생이 월별로 고른가에 대한 테스트와 결과적으로 자유도가 같지만 과정은 다릅니다.

매 학기 Red 와 Black 의 차이를 살펴보는 랜덤화효과는 거의 예외없이 잘 나타납니다.

p-value 가 0.05보다 대부분 훨씬 큰 값으로 나오는 것이죠.

그런데 월별 출생인원의 분포는 고르지 않은 경우가 제법 있습니다.

바로 2023년 1, 2학기와 2021년 1학기가 그런 경우이죠.

확인해 보기 바랍니다.

25.3 Q2. Matching Problem

다음은 어느 가수가 어느 노래를 불렀는지 짹 짓는 (matching) 문제입니다.

수강생들이 태어나기 훨씬 전에 활동하던 옛날 가수와 노래들이기 때문에 누가 무엇을 불렀는지 알 길이 없고 운에 기대어 랜덤하게 골라야 합니다.

이 때 몇 개나 맞출 수 있을까요?

맞춘 갯수가 이론적으로 기대하는 갯수와 잘 들어맞는지 알아보기 위하여 카이제곱 테스트를 수행합니다.

하나도 못 맞추는 경우를 derangement 라고 합니다.

완전히 엉클어진 경우이죠.

ABCD 를 완전히 엉클어 놓기 위해서 A에 주목합시다.

A를 갖다 놓을 수 있는 곳은 A위치를 제외한 나머지 3군데 중 하나입니다.

그 위치를 B라고 했을 때 두 가지 경우가 생깁니다.

원래의 B를 어디에 놓느냐 하는 것이죠.

B를 A의 위치에 놓는 경우와 A아닌 다른 위치에 놓는 방법이 있습니다.

B를 A의 위치에 놓으면 C와 D를 엉클어 놓으면 됩니다.

방법은 하나밖에 없습니다.

그런데 B를 A가 아닌 다른 위치에 놓겠다고 하면 선택은 ACD 를 엉클어 놓는 방법의 수만큼 있습니다.

두 개입니다.

CDA와 DAC이죠.

따라서 ABCD를 엉클어 놓는 방법의 수는 $(4 - 1) \times (1 + 2) = 9(\text{개})$ 입니다.

맺칭이 두개 일어나는 경우는 서로 맞는 2개, 예를 들어서 AB를 고르고 나머지 두 개, CD는 자동적으로 서로 맞지 않게 DC로 배치하면 되니까 서로 맞는 2개를 고르는 방법의 수, 즉 4개에서 2개를 고르는 방법의 수 $\binom{4}{2} = 6(\text{개})$ 가 나옵니다.

1개를 맞추려면 나머지 3개를 서로 엇갈리게 배치하는 방법이 2개 밖에 없습니다.

예를 들어서 ABCD가 바른 순서일 때 A를 고정시키면 BCD 를 엇갈리게 배치하는 방법은 CDB와 DBC 밖에 없습니다.

따라서 $4 \times 2 = 8(\text{개})$ 의 배치 방법이 있습니다.

여기까지 잘 따라왔으면 1개도 맞추지 못하는 경우의 수는 $24 - (8 + 6 + 1) = 9(\text{개})$ 가 되기 때문에 또 다른 방법으로 derangement 의 갯수를 확인하게 됩니다.

이 9개가 나오는 과정을 잘 알려진 공식으로 표현하면 $4! \left(1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!}\right) 0$ 됩니다.

2. 가수와 곡을 맞추시오 (생각 금지! 인터넷 검색 절대 금지!!, 그냥 랜덤하게 골라주세요). *

	Cruel War	Famous Blue Raincoat	And Yours is Piece of Mine	Rain
Uriah Heep	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Leonard Cohen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Peter, Paul and Mary	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Marmalade	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Q2는 맺칭 (matching)문제라고 잘 알려져 있는데 스마트폰이 대중화된 이후로 이상하게도 가수와 노래를 짹짓는 문제에서 수강생들의 정답 비율이, 특히 네개를 다 맞히는, 매우 높았었습니다.

예전 학기들 자료를 보세요.

가수와 노래가 60년대에 속하기 때문에 여러분이 전혀 알 수 없는 것들인데 이 높은 정답률이 인터넷 검색의 위력이라는 것을 2018년 2학기에 알게 되었습니다.

여러분들이 이 노래를 전혀 모르는 상태에서 랜덤하게 고른다면 하나도 못 맞출 확률이 $9/24$ 로 가장 높습니다. 1개 맞출 확률은 $8/24$, 2개는 $6/24$, 다 맞출 확률은 $1/24$ 에 불과합니다.

따라서 맞추는 갯수의 기대값과 표준편차는 모두 한 개입니다.

재수 없으면 하나도 못 맞히고, 재수 좋으면 두 개 정도 맞힌다는 것입니다.

“A이면 B이다”와 논리적으로 동등한 것은 “B가 아니면 A가 아니다”라는 것을 기억해 보면, 집계결과가 이론적으로 예측한 바와 잘 맞지 않으므로 랜덤하게 고르지 않았다는 얘기가 됩니다.

그래서 2018년 2학기와 2019년 2학기에 Matching이라는 퀴즈를 준비했는데 이마저도 기대한 대로 결과가 나오지 않았습니다.

주역의 꽤를 어느 정도 알고 있거나 이 정도는 인터넷 검색이 가능한 것 같습니다.

주역 꽤 구분하기와 비교해 보십시오.

그러던 중에 2020년 1학기 온라인 수업으로 소통이 원활하지 않다고 생각하던 중에 구글 설문지에다 여러분의 인터넷 검색을 금지하고, 랜덤하게 골라달라는 부탁을 하였더니 이론과 너무나도 잘 들어맞는 결과가 나왔습니다.

그런데, 지난 학기에도 예전 학기들 만큼이나 1개를 맞추는 사람들의 수효 352명이 맷칭 모델로부터 기대되는 인원 234명보다 압도적으로 많이 관찰되면서 4곡 다는 아니더라도 한 곡 정도는 인터넷 검색이 많이 있었던 게 아닌가 싶습니다.

인터넷 검색을 금하지 않았던 시기에는 4개 다 맞히는 인원이 가장 많았던 점과 비교해 보면 덜하긴 한 것이죠.

각자 자신의 경험을 댓글로 올려 보세요.

랜덤하게 골랐나요?

몇 개나 맞췄나요?

70년대초 즐겨 듣던 노래들인 데 … 한 학기 댓글 올리느라고 힘든 몸과 마음을 음악으로 차분하게 다스려 보세요.

음악에 대한 의견은 어떤가요?

음악 감상이 끝나면 분석으로 들어갑니다.

25.3.1 응답 분포

	Cruel War	Famous Blue Raincoat	And Yours is Piece of Mine	Rain	계
Uriah Heep	116	256	216	131	719
Leonard Cohen	201	134	322	62	719
Peter, Paul and Mary	302	161	122	134	719
Marmalade	100	168	59	392	719
계	719	719	719	719	2876

가수별, 노래별로 응답 빈도의 교차표를 만들었습니다.

행의 소계나 열의 소계가 모두 719(명)으로 같습니다.

Uriah Heep 의 노래가 Rain이니까 정답을 맞춘 사람은 131(명)입니다.

랜덤하게 골랐다면 719(명)을 4로 나눠 준 179.75(명) 정도 나와야 합니다.

어떨습니까?

Leonard Cohen의 노래는 Famous Blue Raincoat 입니다.

정답을 맞춘 사람은 134(명)입니다.

Peter, Paul and Mary 의 노래는 Cruel War 이니까 정답을 맞춘 사람은 302(명)입니다.

Marmalade 의 노래는 And Yours is Piece of Mine 이니까 정답을 맞춘 사람은 59(명)입니다.

유난히 많이 맞춘 가수의 노래가 있군요.

검색을 한 걸까요?

가수별로 응답 분포의 백분율을 구해 보겠습니다.

25.3.2 응답분포(가수별 %)

	Cruel War	Famous Blue Raincoat	And Yours is Piece of Mine	Rain	계
Uriah Heep (Rain)	16.1	35.6	30	18.2	100
Leonard Cohen (Famous Blue Raincoat)	28	18.6	44.8	8.6	100
Peter, Paul and Mary (Cruel War)	42	22.4	17	18.6	100
Marmalade (And Yours is Piece of Mine)	13.9	23.4	8.2	54.5	100

가수별 백분율을 살펴보면 어느 가수의 노래를 유난히 잘 맞추고 있는지 파악할 수 있습니다.

우선 Uriah Heep 의 노래는 정답을 맞춘 백분율이 18.2(%) 입니다.

간단히 25%가 평균이라고 할 수 있는데 어떻습니까?

Leonard Cohen 의 노래는 정답을 맞춘 백분율이 18.6(%) 입니다. Peter, Paul and Mary 의 노래는 정답을 맞춘 백분율이 42(%) 입니다.

Marmalade 의 노래는 정답을 맞춘 백분율이 8.2(%) 입니다.

누구의 노래를 유난히 잘 맞히고 있습니까?

랜덤하게 고르라고 한 지시를 잘 지키고 있는지 알아 봅시다.

25.3.3 정답갯수의 분포

0개	1개	2개	4개	계
209	356	129	25	719

응답자별로 정답 맞춘 갯수를 세어서 테이블로 정리한 것입니다.

전체 719(명) 중에서 하나도 못 맞춘 사람은 209(명), 1개를 맞춘 사람은 356(명), 2개를 맞춘 사람은 129(명), 4개 모두 맞춘 사람은 25(명)입니다.

랜덤하게 고르라고 한 지시를 잘 지켰는지를 파악하기 위하여 카이제곱 테스트를 수행합니다.

맷칭 문제의 확률분포로부터 각 맞춘 개수의 기대인원을 계산해 보겠습니다.

25.3.4 Observed vs Expected

표 25.12: Observed vs Expected

	0개	1개	2개	4개	계
Observed	209.0	356.0	129.0	25.0	719.0
Expected	269.6	239.7	179.8	30.0	719.0
Difference	-60.6	116.3	-50.8	-5.0	0.0

네명의 가수와 노래를 짹짓는 맷칭 문제에서 맷칭 갯수 {0, 1, 2, 4} 각각의 확률은 {9/24, 8/24, 6/24, 1/24}입니다.

응답인원 719명을 각 확률에 곱해보면 이론적으로 기대되는 인원이 계산됩니다.

매칭 갯수 별로 관찰된 인원을 Observed 행에 올리고, 매칭 문제의 확률분포로부터 계산한 기대 인원을 Expected 행에 올렸습니다.

하나도 못 맞춘 인원은 209(명)인데 확률분포로부터 기대되는 인원은 269.6(명)이어서 그 차이가 -60.6(명)입니다. 1개를 맞춘 인원은 356(명)인데 확률분포로부터 기대되는 인원은 239.7(명)이어서 그 차이가 116.3(명)입니다.

2개를 맞춘 인원은 129(명)인데 확률분포로부터 기대되는 인원은 179.8(명)이어서 그 차이가 -50.8(명)입니다.

4개를 다 맞춘 인원은 25(명)인데 확률분포로부터 기대되는 인원은 30.0(명)이어서 그 차이가 -5.0(명)입니다.

4개를 맞춘 인원은 거의 기대인원과 같은데, 1개를 맞춘 인원은 기대인원보다 유난히 많습니다.

그래서 하나도 못 맞추거나 2개를 맞춘 인원이 기대인원보다 적은 결과를 빚습니다.

인터넷 검색하지 말라고 해서 네 개 다 검색하지는 못하고 한 개만 검색한 걸까요? 카이제곱 테스트를 수행합니다.

25.3.5 맷칭 모델 카이제곱 적합도 테스트

표 25.13: Chi-squared test for given probabilities: .

Test statistic	df	P value
85.25	3	2.295e-18 ***

맷칭 문제의 집계 결과가 맷칭 문제의 확률분포로부터 기대되는 인원과 통계적으로 유의한 차이가 있는지 알아보기 위하여 카이제곱 테스트를 수행하였습니다.

그 결과 카이제곱 통계량은 85.248, 자유도는 3.00, p-value 는 2.3e-180이므로 매우 통계적으로 유의한 차이를 보입니다.

앞에서 관찰한 것처럼 1개 맞춘 사람들이 유난히 많은 게 카이제곱 통계량이 커지는 데 크게 기여하였습니다.

그 의미를 각자 새겨보기 바랍니다.

응답결과를 Red 와 Black으로 나눠 보았을 때 통계적으로 유의한 차이가 있을까요?

25.3.6 Red and Black

	0개	1개	2개	4개	계
Red	103	176	62	15	356
Black	106	180	67	10	363
계	209	356	129	25	719

표 25.15: Pearson's Chi-squared test: .

Test statistic	df	P value
1.214	3	0.7497

맷칭 문제의 집계 결과를 Red, Black으로 나누어 보았습니다.

Red, Black 간에 통계적으로 유의한 차이가 있는지 알아보기 위하여 카이제곱 테스트를 수행하였더니 카이제곱 통계량은 1.214, 자유도는 3, p-value 는 0.74970이므로 Red, Black 간에는 통계적으로 유의한 차이를 보이지 않습니다.

25.4 Q3. 직관과 어긋나는 용어

25.4.1 연비

Q3는 Thinking Fast and Slow 에 나오는 예로 연비의 정의를 단위 연료당 갈 수 있는 거리로 정의하다보니 생길 수 있는 오류를 지적하고 있습니다.

연비라는 용어가 주는 직관과는 잘 맞지 않다는 것을 여러분의 응답에서 잘 알 수 있습니다.

연비는 1리터의 연료로 갈 수 있는 거리입니다.

연비가 높으면 높을수록 적은 연료를 넣고도 먼 거리를 갈 수 있습니다.

그런데 원래 연비가 10인 차량을 연비 12인 차량으로 업그레이드하는 것과 연비가 30인 차량을 연비가 40인 차량으로 업그레이드 하는 것 중에서 어떤 선택이 더 연료를 절감할 수 있는지 계산해 보면 우리가 직관적으로 생각하는 것과는 다른 정답이 기다리고 있음을 알게 됩니다.

이 문제에 대하여 우리 수강생들의 응답을 분석해 봅니다.

3. 동일한 거리를 갈 때 다음 두 사람 중 누가 주유비를 이전보다 절감할 수 있는가? *

- 가. 연비가 10인 차량을 연비가 12인 차량으로 바꾼다
- 나. 연비가 30인 차량을 연비가 40인 차량으로 바꾼다

1,200 킬로미터를 주행한다고 해 봅시다.

'가'는 120리터에서 100리터로 20리터를 절감하고, '나'는 40리터에서 30리터로 10리터를 절감하게 됩니다.

따라서 '가'운전자가 이전보다 더 절감합니다.

연비라는 용어가 주는 직관과는 잘 맞지 않다는 것을 여러분의 응답에서 잘 알 수 있습니다.

연비 높은 차량으로 바꾸는 것이 더 절감할 것이라는 응답이 압도적입니다.

악마는 디테일에 있습니다.

Red, Black 부터 살펴봅니다.

25.4.2 집계

	연비 10 => 12	연비 30 => 40	계
Red	155	201	356
Black	160	203	363
계	315	404	719

표 25.16: Pearson's Chi-squared test with Yates' continuity correction: .

Test statistic	df	P value
0.004921	1	0.9441

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

그 결과 카이제곱 통계량은 0.0049, 자유도는 1, p-value 는 0.9441이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

따라서 합쳐서 백분율을 비교해 보겠습니다.

25.4.3 % 비교.

	연비 10 => 12	연비 30 => 40	계
Red	43.5	56.5	100.0
Black	44.1	55.9	100.0

합쳐서 보겠습니다.

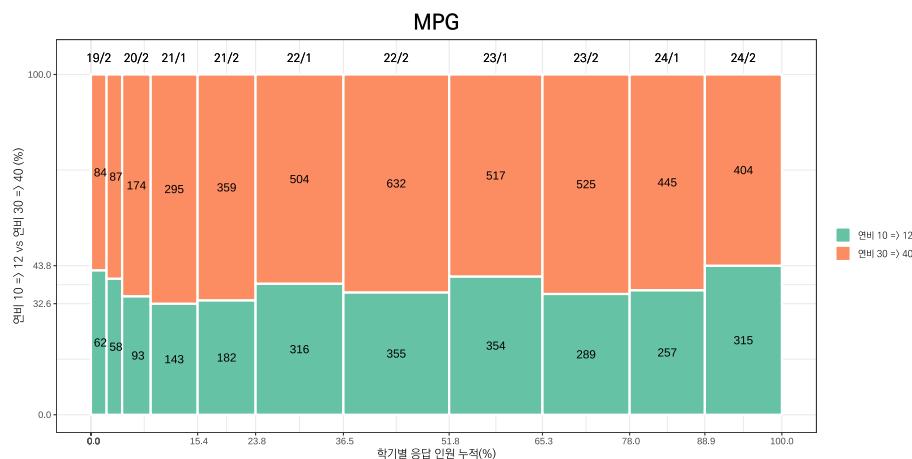
25.4.4 % (통합)

연비 10 => 12	연비 30 => 40	계
		100.0
43.8	56.2	

Red, Black을 합쳐서 계산할 때 56.2(%)가 정답을 고르지 못하였습니다.

이러한 경향은 학기마다 거의 비슷합니다.

25.4.5 학기별 응답 요약



데이터로 확인해 보세요.

정답을 맞춘 백분율은 32.6%에서 42.5% 사이에 있습니다. 절반을 넘어가지 못하는 것을 Mosaic Plot이 잘 보여주고 있습니다.

용어에서 직관적으로 받는 인상과 실제는 다른 것이죠.

25.5 Q4. Monty Hall 문제

퀴즈 쇼 진행자인 Monty Hall 이 세 개의 문을 보여줍니다.

세 문 중 하나는 자동차를 숨기고 있고, 나머지 두 문 뒤에는 염소가 있습니다.

참가자는 세 문 중 하나를 선택합니다.

참가자가 선택한 문을 열기 전에, Monty Hall 은 항상 염소가 있는 다른 문 하나를 열어 보여줍니다.

이제 Monty Hall 은 참가자에게 선택을 바꿀 기회를 줍니다.

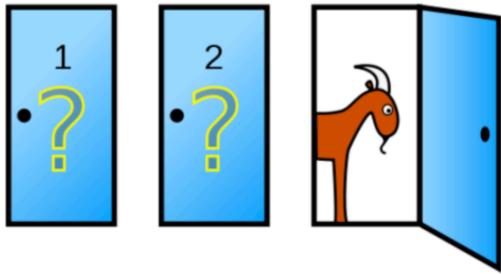
1. 참가자는 선택한 문을 고수하거나,
2. 남아 있는 다른 문으로 바꿀 수 있습니다.

질문 : 선택을 바꾸는 것이 더 나은 전략일까요?

25.5.1 문항 배열 효과?

Red

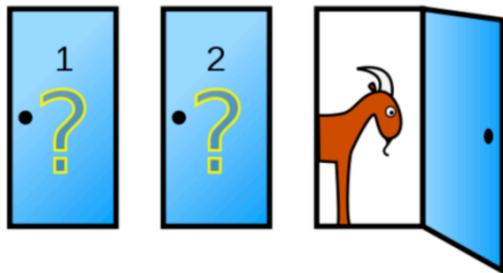
4. 세 개의 문 중에 하나를 선택하여 문 뒤에 있는 선물을 가질 수 있는 게임쇼에 참가했다. 한 문 뒤에는 자동차가 있고, 나머지 두 문 뒤에는 염소가 있다. 이때 어떤 참가자가 1번 문을 선택 했을 때, 게임쇼 진행자는 3번 문을 열어 문 뒤에 염소가 있음을 보여주면서 1번 대신 2번을 선택하겠냐고 물었다. 이 때 어떤 선택을 해야 하는가?



- 가. 선택을 고수한다.
- 나. 선택을 바꾼다.

Black

4. 세 개의 문 중에 하나를 선택하여 문 뒤에 있는 선물을 가질 수 있는 게임쇼에 참가했다. 한 문 ***** 뒤에는 자동차가 있고, 나머지 두 문 뒤에는 염소가 있다. 이때 어떤 참가자가 1번 문을 선택 했을 때, 게임쇼 진행자는 3번 문을 열어 문 뒤에 염소가 있음을 보여주면서 1번 대신 2번을 선택하겠냐고 물었다. 이 때 어떤 선택을 해야 하는가?



- 가. 선택을 바꾼다.
- 나. 선택을 고수한다.

많은 사람들은 선택을 바꾸나 유지하거나 확률이 동일하다고 생각하지만, 선택을 바꾸는 것이 항상 유리합니다.

처음 선택한 문이 자동차일 확률은 $1/3$ 이고 처음 선택한 문이 염소일 확률은 $2/3$ 입니다.

Monty는 항상 염소가 있는 문을 엽니다.

즉, 선택을 바꾸는 경우, 처음에 염소를 선택했을 확률 $2/3$ 가 남은 문에 자동차가 있을 확률로 전환됩니다.

따라서 처음 선택을 유지할 경우 자동차일 확률은 $1/3$ 그대로이지만 선택을 바꿀 경우 자동차일 확률은 $2/3$ 로 늘어나게 됩니다.

자동차는 세 문 중 하나에 무작위로 배치되어 있습니다.

참가자가 문1을 선택했을 때, 자동차가 문1에 있을 확률은 $1/3$ 이고 Monty는 문2나 문3 중에 하나를 고를 수 밖에 업습니다.

이 때 참가자가 선택을 바꾸면 염소를 고르게 됩니다.

자동차가 문2에 있다면, Monty는 반드시 문3을 열 수 밖에 없고 이때 참가자가 선택을 바꾸면 자동차가 있는 문2를 선택하여 승리하게 됩니다.

자동차가 문3에 있더라도 같은 논리로 참가자가 선택을 바꾸면 자동차가 있는 문3을 선택하여 승리하게 됩니다.

즉, 선택을 바꾸면 승리할 확률이 $2/3$ 가 되는 것입니다.

사람들은 “문이 두 개 남았으니, 확률은 $1/2$ 로 동일할 것이다”라고 생각하기 쉽습니다.

그러나 Monty가 문을 열 때 이미 정보를 반영한다는 점에서 조건부 확률이 적용됩니다.

따라서 이 문제는 우리의 직관이 얼마나 확률을 잘못 이해할 수 있는지를 보여줍니다.

Red 와 Black 의 차이는 “바꾼다”와 “고수한다”의 순서를 바꾼 것으로 “바꾼다”를 앞에 놓은 Black 집단에서 바꾼다는 응답이 다소 높게 나왔으나 통계적으로 유의한 수준은 아닙니다.

25.5.2 집계

	고수한다	바꾼다	계
Red	251	105	356
Black	231	132	363
계	482	237	719

표 25.19: Pearson's Chi-squared test with Yates' continuity correction: .

Test statistic	df	P value
3.533	1	0.06014

Q4의 Red는 “선택을 고수한다”를 앞에, “선택을 바꾼다”를 뒤에 나오도록 하였고, Black은 “선택을 바꾼다”를 앞에, “선택을 고수한다”를 뒤에 나오도록 하여 소위 1번효과가 작동하는지를 살펴 보았습니다.

그 결과 Red, Black 의 차이를 분석하기 위한 카이제곱 통계량은 3.533, 자유도는 1, p-value 는 0.0601 으로 1번효과는 통계적으로 유의하지 않음을 보여 줍니다.

그런데 2023년 1학기에는 특이하게도 “바꾼다”를 앞에 놓은 Black 에서 “바꾼다”를 선택한 인원이 Red 보다 통계적으로 유의한 수준으로 많이 나와서 소위 1번효과를 관찰할 수 있습니다.

2021년 2학기에도 통계적으로 유의한 차이를 관찰하였고, 2017년 2학기에도 통계적으로 유의한 차이가 나온 적이 있지만 대체로 1번효과는 잘 나타나지 않습니다.

백분율로 비교합니다.

25.5.3 % 비교.

	고수한다	바꾼다	계
Red	70.5	29.5	100.0
Black	63.6	36.4	100.0

이를 백분율로 살펴보면 Red에서 “선택을 고수한다”는 백분율, 70.5%(은)는 “선택을 바꾼다”는 백분율, 29.5%보다 월등히 높고, Black에서 “선택을 고수한다”는 백분율, 63.6%(은)는 “선택을 바꾼다”는 백분율, 36.4%보다 역시 월등히 높다는 것을 알 수 있습니다.

통계적으로 유의한 차이를 관찰하지 않았기 때문에 합쳐서 보는 게 타당합니다.

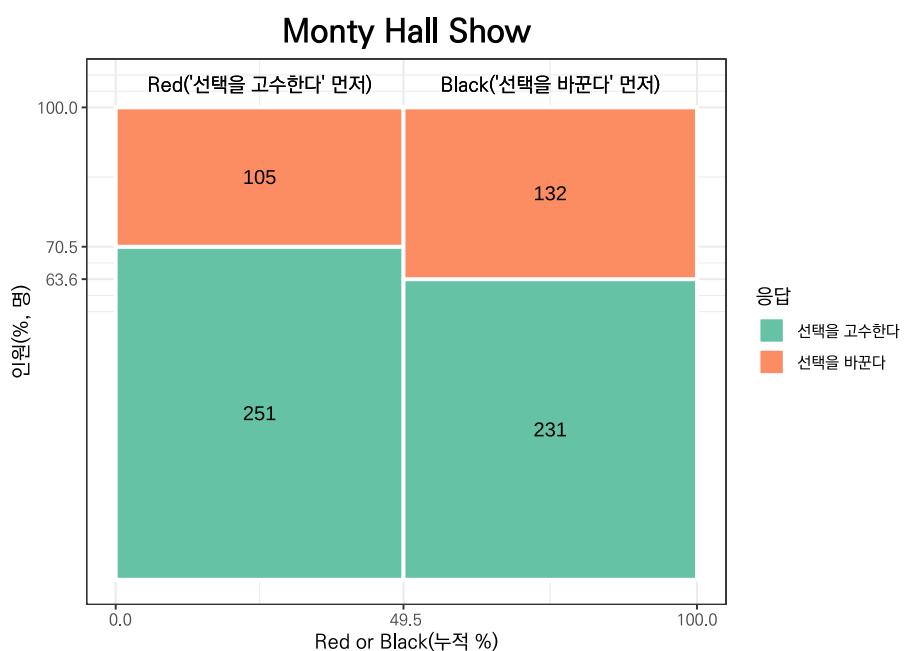
25.5.4 합산(%)

	고수한다	바꾼다	계
계	67.0	33.0	100.0

Red, Black을 통합하여 보면 “선택을 고수한다”는 백분율 67.0%(은)는 ‘선택을 바꾼다’는 백분율 33.0%보다 월등히 높습니다.

그만큼 우리의 직관이 취약하다는 것을 알 수 있습니다.

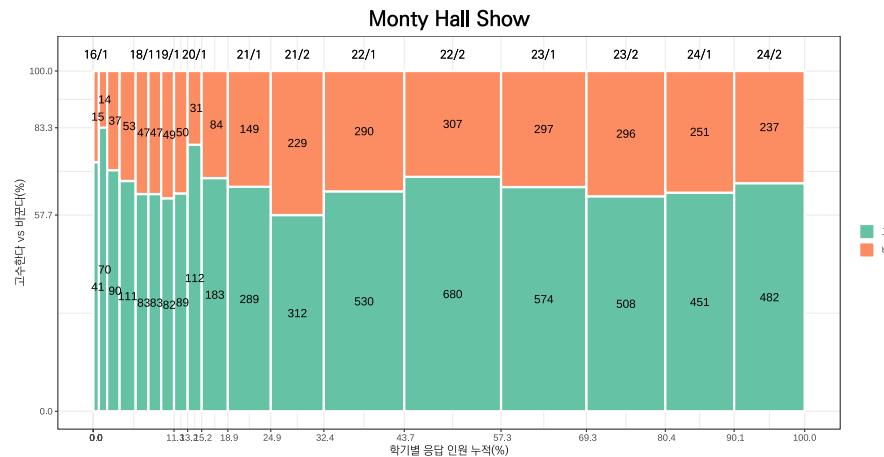
25.5.5 Mosaic Plot



25.5.6 학기별 응답 요약

학기별로 '고수한다'와 '바꾼다'의 비율을 Mosaic Plot으로 요약해 보았습니다.

여러분의 직관에 대해서 그리고 학기별로 꾸준히 관찰되는 거의 고정된 비율에 대해서 댓글을 올려 주세요.



25.6 마감 시간으로부터 제출 시간의 분포

25.6.1 분포표

표 25.22: 일 단위

	14일	13일	12일	11일	10일	9일	8일	7일	6일	5일	4일	3일	2일	1일	계
Red	112	19	14	9	9	10	8	48	24	16	25	9	14	39	356
Black	119	21	14	6	6	3	11	44	22	17	18	17	29	36	363
계	231	40	28	15	15	13	19	92	46	33	43	26	43	75	719

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.

25.6.2 날마다 고르게 제출하는가?

14일	13일	12일	11일	10일	9일	8일	7일	6일	5일	4일	3일	2일	1일
231	40	28	15	15	13	19	92	46	33	43	26	43	75

표 25.24: Chi-squared test for given probabilities: .

Test statistic	df	P value
807.4	13	3.542e-164 * * *

날마다 고르게 제출하는지 알아 보았습니다.

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

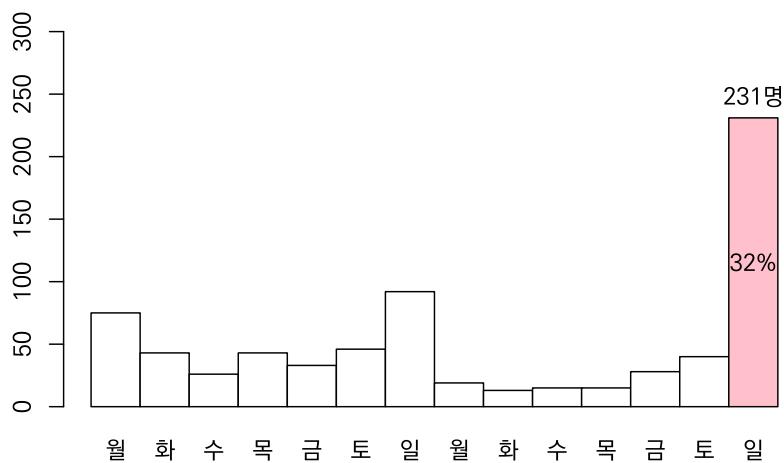
분포표 만으로도 쉽게 파악할 수 있지만 카이제곱테스트가 명확히 해 줍니다.

카이제곱 통계량은 807.428, 자유도는 13.00, p-value 는 3.5e-164 이므로 날짜별로 고르게 제출하고 있지 않다는 것을 강력히 시사합니다.

막대그래프로 살펴 보겠습니다.

25.6.3 막대그래프

Quiz241125 (719명 제출)



25.6.4 Red, Black 간에 닮았는가?

	14일	13일	12일	11일	10일	9일	8일	7일	6일	5일	4일	3일	2일	1일
Red	112	19	14	9	9	10	8	48	24	16	25	9	14	39
Black	119	21	14	6	6	3	11	44	22	17	18	17	29	36

표 25.26: Pearson's Chi-squared test: .

Test statistic	df	P value
14.93	13	0.3115

제출시간의 분포가 Red, Black 간에 닮았는지 알아 보았습니다.

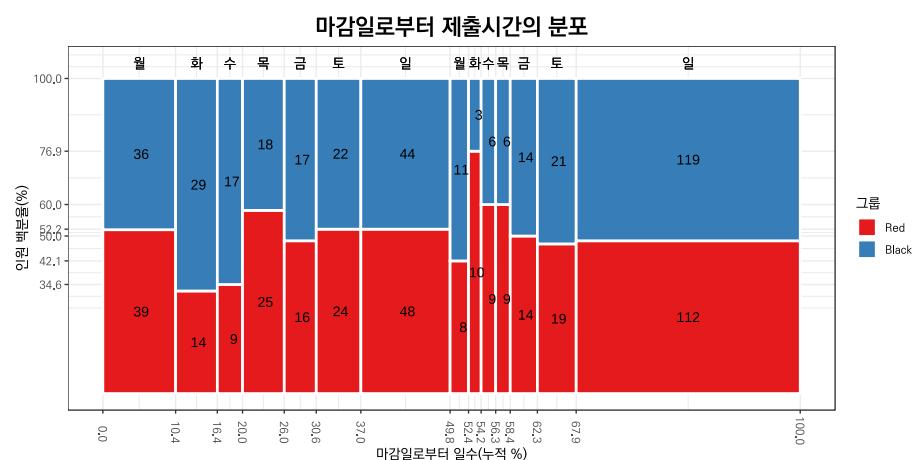
이번에는 분포표의 첫번째와 두번째 행, '계'열을 제외한 나머지 열에 대해서 카이제곱테스트를 수행합니다.

카이제곱 통계량은 14.93, 자유도는 13, p-value 는 0.3115 이므로 제출 시간의 분포는 Red, Black 간에 통계적으로 유의한 차이가 관찰되지 않습니다.

이 사실을 Mosaic Plot 을 이용하여 시각적으로 살펴보겠습니다.

닮았다고 느껴지나요?

25.6.5 Mosaic Plot



제 26 장 직관의 한계 데이터 실험 집계

26.1 실험의 목적

13주차 구글 예습 설문지 집계결과를 분석합니다.

Q1에서는 잘 알려진 생일 문제를 수강생들에게 적용하여 보았습니다.

상당히 많은 인원이기 때문에 한쌍 이상의 생일이 같은 확률은 당연히 1이고 생일이 같은 사람이 얼마나 되는지 알아보겠습니다.

이론적으로 기대되는 인원과 표준오차를 계산하여 실제 관찰된 인원과 비교 합니다.

생일의 월별분포에 대해서도 분석합니다.

Q2에서는 맷칭문제의 사례로 가수와 노래를 짹짓는 문제를 살펴봅니다.

랜덤하게 고르도록 하였는데 왜 이론적으로 기대하는 분포와 다른 결과가 나오는 것인지 생각해 봅니다.

Q3에서는 연비라는 용어가 직관적으로 주는 오류에 대해서 알아 봅니다.

연비가 높은 차를 연비가 좀 더 높은 차로 바꾸는 것과 연비가 낮은 차를 연비가 좀 더 높은 차로 바꾸는 것 중에서 어떤 선택이 좀더 연료비를 절감할 수 있는지 알아 봅니다.

Q4에서는 잘 알려진 Monty Hall 문제를 수강생들에게 물어 본 결과를 분석합니다.

Red 와 Black 은 “고수한다”와 “바꾼다”的 순서를 바꿔 보았는데 그 효과는 통계적으로 유의한 차이를 보이지 않습니다.

그리고 여러분들이 직관적으로 고른 답은 대부분 정답이 아닌 것으로 드러납니다.

26.1.1 Red, Black을 잘못 표시한 사람들

	Red(구글예습퀴즈)	Black(구글예습퀴즈)
Red(랜덤화출석체부)	282	0
Black(랜덤화출석체부)	1	283
계	283	283

Red(구글예습퀴즈)	Black(구글예습퀴즈)
-------------	---------------

랜덤화출석부에 있는 Red, Black 과 실제 구글설문에 올린 Red, Black 이 다른 사람들의 수효는 1명입니다.

Red를 Black 이라고 한 사람이 0명, Black 을 Red 라고 한 사람이 1명입니다.

두 가지 방법으로 분석합니다.

우선 Red, Black 을 잘못 선택한 1명을 랜덤하게 둘로 나누면 어느 한 쪽 집단에 들어갈 기대인원은 1명을 둘로 나눈 0.5(명)이고, 표준오차는 1의 제곱근에 $1/2$ 을 곱해 준 0.5명이 됩니다.

실제로 Red를 Black 이라고 한 사람수, 0명이나 Black 을 Red 라고 한 사람수, 1명은 기대인원으로부터 표준오차 범위에 아주 잘 들어갑니다.

두 번째 분석 방법은 확률을 계산해 보는 것입니다.

Red, Black 을 잘못 선택한 1명을 랜덤하게 둘로 나눌 때, 실제로 관찰된 1명 이상이나 0명이하로 잘못 선택한 사람수가 나올 가능성은 얼마나 되는가 입니다.

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

시행횟수가 10이고 한 번 시행에서 성공확률이 $1/2$ 인 이항분포에서 성공횟수가 0이하이거나 1이상을 관찰할 확률은 1입니다.

공평한 동전 던지기에서 앞면이 0개 이하 나오는 확률은 1개 이상 나오는 확률과 같기 때문에 사실상 한쪽만 계산해서 2배 해 주면 됩니다.

다만, 이번 실험과 같이 3명씩 동일한 결과가 나온 경우에는 중복되는 확률을 빼 주어야 합니다.

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.

26.1.2 응답인원의 Red, Black

Red 로 응답한 인원은 283명, Black 에 응답한 인원은 283명입니다.

전체 응답인원 566 명을 랜덤하게 둘로 나눌 때 어느 한 쪽의 기대인원은 전체 응답인원의 절반인 283명이고, 표준오차는 전체 응답인원의 제곱근에 $1/2$ 을 곱해 준 11.9 명입니다.

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

26.2 Q1. Birthday Problem

1. 나의 생일은 _____ 월 _____ 일이다 *
예시 : 7월7일(X), 7/7(X), 0707(X), 07/07(0)

내 답변

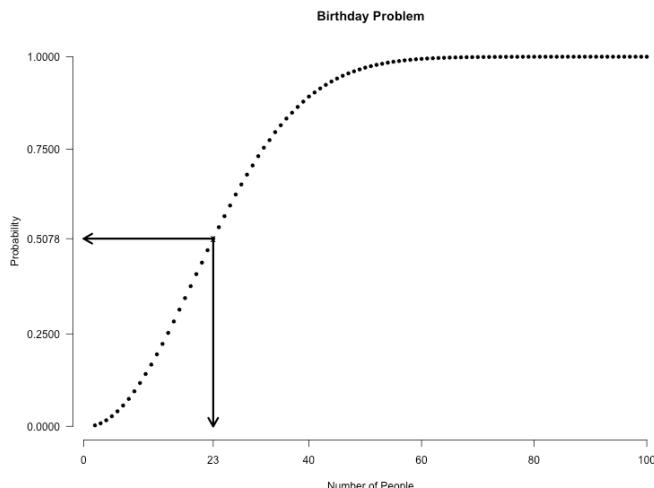
Q1은 생일 문제라고 잘 알려져 있습니다.

23명만 모여 있어도 생일이 같은 사람이 한 쌍 이상 있을 확률이 $1/2$ 을 넘어갑니다.

다음 그림은 1에서 365의 숫자 중 x-축에 나온 숫자 만큼 랜덤하게 뽑아서 (복원 추출) 같은 숫자가 있는지 여부를 백만번씩 돌려서 나온 결과들입니다.

23명을 뽑았을 때 백만번 중에 50만7천8백번 같은 숫자가 나왔다는 뜻입니다.

큰수의 법칙에 따라서 그 비율은 23명 중에 생일이 같은 사람이 최소한 한 쌍 있을 확률로 수렴해 갑니다.



다음은 집계결과로부터 어느 날짜에 몇 명의 생일이 같은지 순서대로 정리한 것입니다.

가끔 이 결과를 보고 “나도 제출했는데 내 생일이 빠졌다”고 의문을 제기하는 사람들이 있습니다만 그 사람과 생일 같은 사람이 없는 것 뿐입니다.

집계가 진행되면서 당연히도 생일 같은 사람들의 수효는 늘어나게 마련입니다.

그럼에도 그 인원은 이론적으로 계산한 기대인원과 표준오차 범위에 거의 들어가는 것을 관찰할 수 있습니다.

보고서의 맨 위에 집계 시점이 나옵니다.

시간 경과에 따라 생일이 같은 사람의 수효가 늘어나는 것과 그 수효가 이론적으로 예측한 범위에 들어가는지 살펴보기 바랍니다.

26.2.1 어느 날에 몇 명씩 생일이 같은가?

01월02일	3
01월04일	2
01월07일	4
01월10일	3
01월11일	2
01월12일	2
01월14일	3
01월17일	2
01월21일	3
01월23일	2
01월24일	2
02월01일	2
02월06일	3
02월09일	3
02월10일	2
02월17일	3
02월18일	2
02월19일	2
02월20일	2
02월21일	3
02월23일	2
02월24일	3
02월25일	3
02월26일	2
02월28일	2
03월01일	2
03월02일	3
03월04일	2
03월05일	4
03월08일	2
03월10일	2
03월11일	5
03월14일	2

03월18일	2
03월20일	2
03월21일	3
03월22일	4
03월24일	5
03월25일	5
03월26일	2
03월27일	2
03월30일	4
04월05일	4
04월09일	2
04월10일	4
04월12일	4
04월13일	2
04월15일	2
04월16일	3
04월18일	2
04월19일	3
04월25일	3
04월26일	3
04월27일	2
04월28일	2
04월29일	3
05월03일	4
05월06일	2
05월07일	4
05월09일	2
05월10일	2
05월11일	5
05월12일	2
05월14일	2
05월15일	2
05월18일	2
05월19일	3
05월20일	4
05월21일	2
05월22일	3
05월24일	3
05월27일	2
05월28일	4
06월09일	2
06월12일	4
06월14일	2
06월15일	3
06월17일	3

06월20일	2
06월22일	2
06월25일	2
06월26일	5
06월29일	3
07월01일	3
07월05일	4
07월07일	4
07월08일	2
07월09일	3
07월10일	2
07월13일	3
07월15일	3
07월16일	2
07월19일	3
07월21일	2
07월25일	2
07월26일	3
07월27일	2
07월30일	2
08월01일	2
08월02일	2
08월03일	3
08월06일	3
08월08일	2
08월10일	4
08월12일	2
08월14일	3
08월17일	3
08월21일	2
08월22일	2
08월23일	2
08월24일	2
08월26일	3
08월30일	3
08월31일	3
09월01일	3
09월03일	2
09월04일	3
09월06일	4
09월07일	3
09월08일	3
09월15일	2
09월16일	2
09월17일	2

09월19일	2
09월21일	4
09월24일	2
09월25일	2
09월29일	2
09월30일	3
10월01일	2
10월06일	2
10월07일	5
10월08일	3
10월09일	2
10월10일	2
10월11일	3
10월20일	2
10월21일	4
10월24일	2
10월25일	3
10월26일	2
10월27일	2
11월01일	2
11월02일	3
11월03일	3
11월04일	2
11월07일	2
11월08일	3
11월09일	2
11월11일	2
11월14일	3
11월15일	2
11월17일	2
11월18일	3
11월20일	2
11월29일	3
11월30일	2
12월02일	2
12월03일	2
12월05일	2
12월06일	3
12월07일	2
12월11일	4
12월12일	3
12월13일	5
12월20일	3
12월23일	2
12월26일	3

12월30일	3
12월31일	2
계	454

26.2.2 생일이 같은 사람은 몇 명 정도 기대되는가?

전체 응답인원 568(명) 중에 생일이 같은 사람은 454(명)이고 생일이 같은 날은 170(일)입니다.

N 을 전체 인원이라 할 때, 기대 인원은 $N \times \{1 - (\frac{364}{365})^{N-1}\}$, 분산은 $N \times \{1 - (\frac{364}{365})^{N-1}\} + N \times (N-1) \times \{1 - (\frac{363}{365})^{N-2}\}$ 로 계산됩니다.

무응답이거나 결석한 학생을 제외한 응답 인원 568명에 대하여 공식에 따라 기대인원을 계산하면 448.1명, 표준오차는 21.2명으로 계산되어 생일이 같은 사람들의 수효 454(명)은 기대인원으로부터 표준오차, 혹은 두 배의 표준오차 범위 안에 잘 들어감을 알 수 있습니다.

26.2.2.1 기대되는 인원

기대인원
448.1

26.2.2.2 표준오차

표준오차
21.2

이전 학기 자료들에서는 10명이 생일이 같은 경우도 있었고, 8명이 생일이 같은 경우는 여럿 나오기도 했는데 이번 학기에는 03월11일, 03월24일, 03월25일, 05월11일, 06월26일, 10월07일, 12월13일에 5(명)의 생일이 같습니다.

2024년 1학기, 2023년 1, 2학기, 2022년 1, 2학기, 2021년 1, 2학기 모두 이론적으로 기대하는 값과 관찰값이 잘 들어 맞았습니다.

2020년 1-2학기에는 기대에 약간 못 미치는 인원을 관찰하였지만 통상적으로 얘기하는 표준오차의 두배 이내에는 잘 들어맞는 인원입니다.

여러분의 생일은 몇 명이나 같은 사람이 있나요?

이론적으로 기대하는 인원과 실제 관찰된 인원이 잘 부합한다는 점에 대해서 어떤 생각이 듭니까?

26.2.3 태어난 달의 분포는?

제출한 생일 날짜들을 월별로 정리하였습니다.

아래 교차표로부터 두 가지 질문을 던져볼 수 있겠습니다.

응답자들은 월별로 고르게 출생하였을까?

출생한 달의 분포는 Red, Black 간에 닮았는가?

이 질문에 답하기 위해서 카이제곱 테스트를 수행합니다.

	1월	2월	3월	4월	5월	6월	7월	8월	9월	10월	11월	12월	계
Red	27	19	29	21	26	20	26	22	26	18	30	20	284
Black	14	26	33	29	30	18	19	29	21	27	14	24	284
계	41	45	62	50	56	38	45	51	47	45	44	44	568

월별로 고르게 출생하였는지 알아보려면 위의 교차표에서 “계” 행의 1월부터 12월까지를
잘라 내어 카이제곱 균일성 테스트를 수행해야 합니다.

이 때 1월부터 12월까지 12개의 범주가 있으니까 자유도는 하나를 뺀 11이 됩니다.

26.2.4 월별로 고르게 출생하였는가?

1월	2월	3월	4월	5월	6월	7월	8월	9월	10월	11월	12월
41	45	62	50	56	38	45	51	47	45	44	44

표 26.6: Chi-squared test for given probabilities: .

Test statistic	df	P value
10.07	11	0.5241

월별로 고르게 출생하였는지 알아보기 위하여 수행한 카이제곱 테스트에서 카이제곱 통계량은 10.070, 자유도는 11.00, p-value 는 0.5241 이므로 월별로 고르게 출생하고 있음을 시사합니다.

가장 많이 태어난 달은 3월에 62(명)이고 가장 적게 태어난 달은 6월에 38(명)입니다.

태어난 달의 분포가 Red, Black 간에 닮았는지 살펴보기 위하여 “계” 행과 “계” 열을 삭제한 다음 교차표에 카이제곱 테스트를 수행합니다.

26.2.5 Red and Black

	1월	2월	3월	4월	5월	6월	7월	8월	9월	10월	11월	12월
Red	27	19	29	21	26	20	26	22	26	18	30	20
Black	14	26	33	29	30	18	19	29	21	27	14	24

표 26.8: Pearson's Chi-squared test: .

Test statistic	df	P value
17.7	11	0.08873

Red, Black 간에 월별 출생의 분포가 닮았는지 알아보기 위하여 수행한 카이제곱 테스트에서 카이제곱 통계량은 17.703, 자유도는 11, p-value 는 0.0887 이므로 Red, Black 간에 출생의 분포는 통계적으로 유의한 차이를 보이고 있습니다.

여기서 자유도는 $(2 - 1) \times (12 - 1) = 11$ 로 계산한 것입니다.

앞의 출생이 월별로 고른가에 대한 테스트와 결과적으로 자유도가 같지만 과정은 다릅니다.

매 학기 Red 와 Black 의 차이를 살펴보는 랜덤화효과는 거의 예외없이 잘 나타납니다.

p-value 가 0.05보다 대부분 훨씬 큰 값으로 나오는 것이죠.

그런데 월별 출생인원의 분포는 고르지 않은 경우가 제법 있습니다.

바로 2023년 1, 2학기와 2021년 1학기가 그런 경우이죠.

확인해 보기 바랍니다.

26.3 Q2. Matching Problem

다음은 어느 가수가 어느 노래를 불렀는지 짹 짓는 (matching) 문제입니다.

수강생들이 태어나기 훨씬 전에 활동하던 옛날 가수와 노래들이기 때문에 누가 무엇을 불렀는지 알 길이 없고 운에 기대어 랜덤하게 골라야 합니다.

이 때 몇 개나 맞출 수 있을까요?

맞춘 갯수가 이론적으로 기대하는 갯수와 잘 들어맞는지 알아보기 위하여 카이제곱 테스트를 수행합니다.

하나도 못 맞추는 경우를 derangement 라고 합니다.

완전히 엉클어진 경우이죠.

ABCD 를 완전히 엉클어 놓기 위해서 A에 주목합시다.

A를 갖다 놓을 수 있는 곳은 A위치를 제외한 나머지 3군데 중 하나입니다.

그 위치를 B라고 했을 때 두 가지 경우가 생깁니다.

원래의 B를 어디에 놓느냐 하는 것이죠.

B를 A의 위치에 놓는 경우와 A아닌 다른 위치에 놓는 방법이 있습니다.

B를 A의 위치에 놓으면 C와 D를 엉클어 놓으면 됩니다.

방법은 하나밖에 없습니다.

그런데 B를 A가 아닌 다른 위치에 놓겠다고 하면 선택은 ACD를 엉클어 놓는 방법의 수만큼 있습니다.

두 개입니다.

CDA와 DAC이죠.

따라서 ABCD를 엉클어 놓는 방법의 수는 $(4 - 1) \times (1 + 2) = 9$ (개)입니다.

맷칭이 두개 일어나는 경우는 서로 맞는 2개, 예를 들어서 AB를 고르고 나머지 두 개, CD는 자동적으로 서로 맞지 않게 DC로 배치하면 되니까 서로 맞는 2개를 고르는 방법의 수, 즉 4개에서 2개를 고르는 방법의 수 $\binom{4}{2} = 6$ (개)가 나옵니다.

1개를 맞추려면 나머지 3개를 서로 엇갈리게 배치하는 방법이 2개 밖에 없습니다.

예를 들어서 ABCD가 바른 순서일 때 A를 고정시키면 BCD를 엇갈리게 배치하는 방법은 CDB와 DBC 밖에 없습니다.

따라서 $4 \times 2 = 8$ (개)의 배치 방법이 있습니다.

여기까지 잘 따라왔으면 1개도 맞추지 못하는 경우의 수는 $24 - (8 + 6 + 1) = 9$ (개)가 되기 때문에 또 다른 방법으로 derangement의 갯수를 확인하게 됩니다.

이 9개가 나오는 과정을 잘 알려진 공식으로 표현하면 $4! \left(1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!}\right)$ 이 됩니다.

	Cruel War	Famous Blue Raincoat	And Yours is Piece of Mine	Rain
Uriah Heep	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Leonard Cohen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Peter, Paul and Mary	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Marmalade	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Q2는 맷칭 (matching)문제라고 잘 알려져 있는데 스마트폰이 대중화된 이후로 이상하게도 가수와 노래를 짹짓는 문제에서 수강생들의 정답 비율이, 특히 네개를 다 맞히는, 매우 높았었습니다.

예전 학기들 자료를 보세요.

가수와 노래가 60년대에 속하기 때문에 여러분이 전혀 알 수 없는 것들인데 이 높은 정답률이 인터넷 검색의 위력이라는 것을 2018년 2학기에 알게 되었습니다.

여러분들이 이 노래를 전혀 모르는 상태에서 랜덤하게 고른다면 하나도 못 맞출 확률이 9/24로 가장 높습니다. 1개 맞출 확률은 8/24, 2개는 6/24, 다 맞출 확률은 1/24에 불과합니다.

따라서 맞추는 갯수의 기대값과 표준편차는 모두 한 개입니다.

재수 없으면 하나도 못 맞히고, 재수 좋으면 두 개 정도 맞힌다는 것입니다.

“A이면 B이다”와 논리적으로 동등한 것은 “B가 아니면 A가 아니다”라는 것을 기억해 보면, 집계결과가 이론적으로 예측한 바와 잘 맞지 않으므로 랜덤하게 고르지 않았다는 얘기가 됩니다.

그래서 2018년 2학기와 2019년 2학기에 Matching이라는 퀴즈를 준비했는데 이마저도 기대한 대로 결과가 나오지 않았습니다.

주역의 꽈를 어느 정도 알고 있거나 이 정도는 인터넷 검색이 가능한 것 같습니다.

주역 꽈 구분하기와 비교해 보십시오.

그러던 중에 2020년 1학기 온라인 수업으로 소통이 원활하지 않다고 생각하던 중에 구글 설문지에다 여러분의 인터넷 검색을 금지하고, 랜덤하게 골라달라는 부탁을 하였더니 이론과 너무나도 잘 들어맞는 결과가 나왔습니다.

그런데, 지난 학기에도 예전 학기들 만큼이나 1개를 맞추는 사람들의 수효 352명이 맷칭 모델로부터 기대되는 인원 234명보다 압도적으로 많이 관찰되면서 4곡 다는 아니더라도 한 곡 정도는 인터넷 검색이 많이 있었던 게 아닌가 싶습니다.

인터넷 검색을 금하지 않았던 시기에는 4개 다 맞히는 인원이 가장 많았던 점과 비교해 보면 덜하긴 한 것이죠.

각자 자신의 경험을 댓글로 올려 보세요.

랜덤하게 골랐나요?

몇 개나 맞췄나요?

70년대초 즐겨 듣던 노래들인 데 … 한 학기 댓글 올리느라고 힘든 몸과 마음을 음악으로 차분하게 다스려 보세요.

음악에 대한 의견은 어떤가요?

음악 감상이 끝나면 분석으로 들어갑니다.

26.3.1 응답 분포

	Cruel War	Famous Blue Raincoat	And Yours is Piece of Mine	Rain	계
Uriah Heep	119	173	170	106	568

	Cruel War	Famous Blue Raincoat	And Yours is Piece of Mine	Rain	계
Leonard Cohen	157	116	227	68	568
Peter, Paul and Mary	215	128	97	128	568
Marmalade	77	151	74	266	568
계	568	568	568	568	2272

가수별, 노래별로 응답 빈도의 교차표를 만들었습니다.

행의 소계나 열의 소계가 모두 568(명)으로 같습니다.

Uriah Heep 의 노래가 Rain이니까 정답을 맞춘 사람은 106(명)입니다.

랜덤하게 골랐다면 568(명)을 4로 나눠 준 142(명) 정도 나와야 합니다.

어떻습니까?

Leonard Cohen의 노래는 Famous Blue Raincoat 입니다.

정답을 맞춘 사람은 116(명)입니다.

Peter, Paul and Mary 의 노래는 Cruel War 이니까 정답을 맞춘 사람은 215(명)입니다.

Marmalade 의 노래는 And Yours is Piece of Mine 이니까 정답을 맞춘 사람은 74(명)입니다.

유난히 많이 맞춘 가수의 노래가 있군요.

검색을 한 걸까요?

가수별로 응답 분포의 백분율을 구해 보겠습니다.

26.3.2 응답분포(가수별 %)

	Cruel War	Famous Blue Raincoat	And Yours is Piece of Mine	Rain	계
Uriah Heep (Rain)	21	30.5	29.9	18.7	100
Leonard Cohen (Famous Blue Raincoat)	27.6	20.4	40	12	100
Peter, Paul and Mary (Cruel War)	37.9	22.5	17.1	22.5	100
Marmalade (And Yours is Piece of Mine)	13.6	26.6	13	46.8	100

가수별 백분율을 살펴보면 어느 가수의 노래를 유난히 잘 맞추고 있는지 파악할 수 있습니다.

우선 Uriah Heep 의 노래는 정답을 맞춘 백분율이 18.7(%) 입니다.

간단히 25%가 평균이라고 할 수 있는데 어떻습니까?

Leonard Cohen 의 노래는 정답을 맞춘 백분율이 20.4(%) 입니다. Peter, Paul and Mary 의 노래는 정답을 맞춘 백분율이 37.9(%) 입니다.

Marmalade 의 노래는 정답을 맞춘 백분율이 13(%) 입니다.

누구의 노래를 유난히 잘 맞히고 있습니까?

랜덤하게 고르라고 한 지시를 잘 지키고 있는지 알아 봅시다.

26.3.3 정답갯수의 분포

0개	1개	2개	4개	계
189	254	108	17	568

응답자별로 정답 맞춘 갯수를 세어서 테이블로 정리한 것입니다.

전체 568(명) 중에서 하나도 못 맞춘 사람은 189(명), 1개를 맞춘 사람은 254(명), 2개를 맞춘 사람은 108(명), 4개 모두 맞춘 사람은 17(명)입니다.

랜덤하게 고르라고 한 지시를 잘 지켰는지를 파악하기 위하여 카이제곱 테스트를 수행합니다.

맷칭 문제의 확률분포로부터 각 맞춘 개수의 기대인원을 계산해 보겠습니다.

26.3.4 Observed vs Expected

표 26.12: Observed vs Expected

	0개	1개	2개	4개	계
Observed	189.0	254.0	108.0	17.0	568.0
Expected	213.0	189.3	142.0	23.7	568.0
Difference	-24.0	64.7	-34.0	-6.7	0.0

네명의 가수와 노래를 짹짓는 맷칭 문제에서 맷칭 갯수 {0, 1, 2, 4} 각각의 확률은 {9/24, 8/24, 6/24, 1/24}입니다.

응답인원 568명을 각 확률에 곱해보면 이론적으로 기대되는 인원이 계산됩니다.

매칭 갯수 별로 관찰된 인원을 Observed 행에 올리고, 매칭 문제의 확률분포로부터 계산한 기대 인원을 Expected 행에 올렸습니다.

하나도 못 맞춘 인원은 189(명)인데 확률분포로부터 기대되는 인원은 213.0(명)이어서 그 차이가 -24.0(명)입니다. 1개를 맞춘 인원은 254(명)인데 확률분포로부터 기대되는 인원은 189.3(명)이어서 그 차이가 64.7(명)입니다.

2개를 맞춘 인원은 108(명)인데 확률분포로부터 기대되는 인원은 142.0(명)이어서 그 차이가 -34.0(명)입니다.

4개를 다 맞춘 인원은 17(명)인데 확률분포로부터 기대되는 인원은 23.7(명)이어서 그 차이가 -6.7(명)입니다.

4개를 맞춘 인원은 거의 기대인원과 같은데, 1개를 맞춘 인원은 기대인원보다 유난히 많습니다.

그래서 하나도 못 맞추거나 2개를 맞춘 인원이 기대인원보다 적은 결과를 빚습니다.

인터넷 검색하지 말라고 해서 네 개 다 검색하지는 못하고 한 개만 검색한 걸까요? 카이제곱 테스트를 수행합니다.

26.3.5 매칭 모델 카이제곱 적합도 테스트

표 26.13: Chi-squared test for given probabilities: .

Test statistic	df	P value
34.81	3	1.336e-07 ***

매칭 문제의 집계 결과가 매칭 문제의 확률분포로부터 기대되는 인원과 통계적으로 유의한 차이가 있는지 알아보기 위하여 카이제곱 테스트를 수행하였습니다.

그 결과 카이제곱 통계량은 34.810, 자유도는 3.00, p-value 는 1.3e-07이므로 매우 통계적으로 유의한 차이를 보입니다.

앞에서 관찰한 것처럼 1개 맞춘 사람들이 유난히 많은 게 카이제곱 통계량이 커지는 데 크게 기여하였습니다.

그 의미를 각자 새겨보기 바랍니다.

응답결과를 Red 와 Black으로 나눠 보았을 때 통계적으로 유의한 차이가 있을까요?

26.3.6 Red and Black

	0개	1개	2개	4개	계
Red	96	130	51	7	284
Black	93	124	57	10	284
계	189	254	108	17	568

표 26.15: Pearson's Chi-squared test: .

Test statistic	df	P value
1.052	3	0.7886

맷칭 문제의 집계 결과를 Red, Black으로 나누어 보았습니다.

Red, Black 간에 통계적으로 유의한 차이가 있는지 알아보기 위하여 카이제곱 테스트를 수행하였더니 카이제곱 통계량은 1.052, 자유도는 3, p-value 는 0.78860이므로 Red, Black 간에는 통계적으로 유의한 차이를 보이지 않습니다.

26.4 Q3. 직관과 어긋나는 용어

26.4.1 연비

Q3는 Thinking Fast and Slow 에 나오는 예로 연비의 정의를 단위 연료당 갈 수 있는 거리로 정의하다보니 생길 수 있는 오류를 지적하고 있습니다.

연비라는 용어가 주는 직관과는 잘 맞지 않다는 것을 여러분의 응답에서 잘 알 수 있습니다.

연비는 1리터의 연료로 갈 수 있는 거리입니다.

연비가 높으면 높을수록 적은 연료를 넣고도 먼 거리를 갈 수 있습니다.

그런데 원래 연비가 10인 차량을 연비 12인 차량으로 업그레이드하는 것과 연비가 30인 차량을 연비가 40인 차량으로 업그레이드 하는 것 중에서 어떤 선택이 더 연료를 절감할 수 있는지 계산해 보면 우리가 직관적으로 생각하는 것과는 다른 정답이 기다리고 있음을 알게 됩니다.

이 문제에 대하여 우리 수강생들의 응답을 분석해 봅니다.

3. 동일한 거리를 갈 때 다음 두 사람 중 누가 주유비를 이전보다 절감할 수 있는가? *

- 가. 연비가 10인 차량을 연비가 12인 차량으로 바꾼다
- 나. 연비가 30인 차량을 연비가 40인 차량으로 바꾼다

1,200 킬로미터를 주행한다고 해 봅시다.

'가'는 120리터에서 100리터로 20리터를 절감하고, '나'는 40리터에서 30리터로 10리터를 절감하게 됩니다.

따라서 '가'운전자가 이전보다 더 절감합니다.

연비라는 용어가 주는 직관과는 잘 맞지 않다는 것을 여러분의 응답에서 잘 알 수 있습니다.

연비 높은 차량으로 바꾸는 것이 더 절감할 것이라는 응답이 압도적입니다.

악마는 디테일에 있습니다.

Red, Black 부터 살펴봅니다.

26.4.2 집계

	연비 10 => 12	연비 30 => 40	계
Red	119	165	284
Black	129	155	284
계	248	320	568

표 26.16: Pearson's Chi-squared test with Yates' continuity correction: .

Test statistic	df	P value
0.5797	1	0.4464

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

그 결과 카이제곱 통계량은 0.580, 자유도는 1, p-value 는 0.44640이므로 Red, Black 간에 통계적으로 유의한 차이를 보이지 않습니다.

따라서 합쳐서 백분율을 비교해 보겠습니다.

26.4.3 % 비교.

	연비 10 => 12	연비 30 => 40	계
Red	41.9	58.1	100.0
Black	45.4	54.6	100.0

합쳐서 보겠습니다.

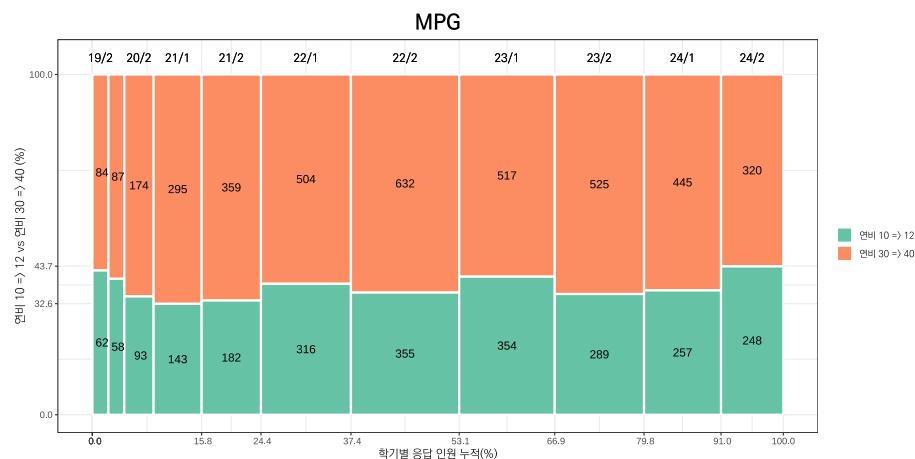
26.4.4 % (통합)

연비 10 => 12	연비 30 => 40	계
43.7	56.3	100.0

Red, Black을 합쳐서 계산할 때 56.3(%)가 정답을 고르지 못하였습니다.

이러한 경향은 학기마다 거의 비슷합니다.

26.4.5 학기별 응답 요약



데이터로 확인해 보세요.

정답을 맞춘 백분율은 32.6%에서 42.5% 사이에 있습니다. 절반을 넘어가지 못하는 것을 Mosaic Plot이 잘 보여주고 있습니다.

용어에서 직관적으로 받는 인상과 실제는 다른 것이죠.

26.5 Q4. Monty Hall 문제

퀴즈 쇼 진행자인 Monty Hall 이 세 개의 문을 보여줍니다.

세 문 중 하나는 자동차를 숨기고 있고, 나머지 두 문 뒤에는 염소가 있습니다.

참가자는 세 문 중 하나를 선택합니다.

참가자가 선택한 문을 열기 전에, Monty Hall 은 항상 염소가 있는 다른 문 하나를 열어 보여줍니다.

이제 Monty Hall 은 참가자에게 선택을 바꿀 기회를 줍니다.

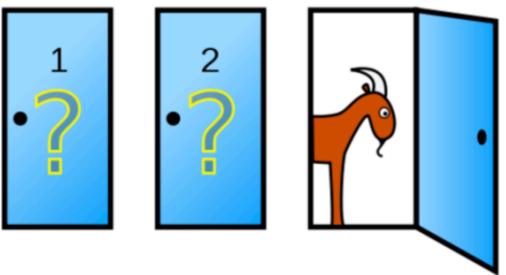
1. 참가자는 선택한 문을 고수하거나,
2. 남아 있는 다른 문으로 바꿀 수 있습니다.

질문 : 선택을 바꾸는 것이 더 나은 전략일까요?

26.5.1 문항 배열 효과?

Red

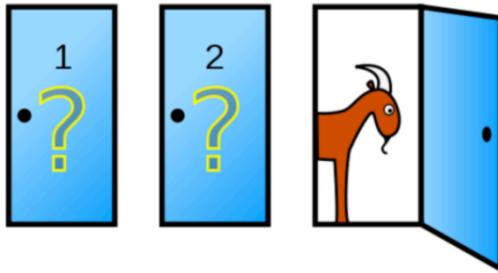
4. 세 개의 문 중에 하나를 선택하여 문 뒤에 있는 선물을 가질 수 있는 게임쇼에 참가했다. 한 문 뒤에는 자동차가 있고, 나머지 두 문 뒤에는 염소가 있다. 이때 어떤 참가자가 1번 문을 선택 했을 때, 게임쇼 진행자는 3번 문을 열어 문 뒤에 염소가 있음을 보여주면서 1번 대신 2번을 선택하겠냐고 물었다. 이 때 어떤 선택을 해야 하는가?



- 가. 선택을 고수한다.
- 나. 선택을 바꾼다.

Black

4. 세 개의 문 중에 하나를 선택하여 문 뒤에 있는 선물을 가질 수 있는 게임쇼에 참가했다. 한 문 뒤에는 자동차가 있고, 나머지 두 문 뒤에는 염소가 있다. 이때 어떤 참가자가 1번 문을 선택 했을 때, 게임쇼 진행자는 3번 문을 열어 문 뒤에 염소가 있음을 보여주면서 1번 대신 2번을 선택하겠냐고 물었다. 이 때 어떤 선택을 해야 하는가? *



- 가. 선택을 바꾼다.
- 나. 선택을 고수한다.

많은 사람들은 선택을 바꾸나 유지하거나 확률이 동일하다고 생각하지만, 선택을 바꾸는 것이 항상 유리합니다.

처음 선택한 문이 자동차일 확률은 1/3이고 처음 선택한 문이 염소일 확률은 2/3입니다.

Monty는 항상 염소가 있는 문을 엽니다.

즉, 선택을 바꾸는 경우, 처음에 염소를 선택했을 확률 2/3가 남은 문에 자동차가 있을 확률로 전환됩니다.

따라서 처음 선택을 유지할 경우 자동차일 확률은 1/3 그대로이지만 선택을 바꿀 경우 자동차일 확률은 2/3로 늘어나게 됩니다.

자동차는 세 문 중 하나에 무작위로 배치되어 있습니다.

참가자가 문1을 선택했을 때, 자동차가 문1에 있을 확률은 1/3이고 Monty는 문2나 문3 중에 하나를 고를 수 밖에 업습니다.

이 때 참가자가 선택을 바꾸면 염소를 고르게 됩니다.

자동차가 문2에 있다면, Monty는 반드시 문3을 열 수 밖에 없고 이때 참가자가 선택을 바꾸면 자동차가 있는 문2를 선택하여 승리하게 됩니다.

자동차가 문3에 있더라도 같은 논리로 참가자가 선택을 바꾸면 자동차가 있는 문3을 선택하여 승리하게 됩니다.

즉, 선택을 바꾸면 승리할 확률이 2/3가 되는 것입니다.

사람들은 “문이 두 개 남았으니, 확률은 1/2로 동일할 것이다”라고 생각하기 쉽습니다.

그러나 Monty가 문을 열 때 이미 정보를 반영한다는 점에서 조건부 확률이 적용됩니다.

따라서 이 문제는 우리의 직관이 얼마나 확률을 잘못 이해할 수 있는지를 보여줍니다.

Red 와 Black 의 차이는 “바꾼다”와 “고수한다”의 순서를 바꾼 것으로 “바꾼다”를 앞에 놓은 Black 집단에서 바꾼다는 응답이 다소 높게 나왔으나 통계적으로 유의한 수준은 아닙니다.

26.5.2 집계

	고수한다	바꾼다	계
Red	199	85	284
Black	184	100	284
계	383	185	568

표 26.19: Pearson's Chi-squared test with Yates' continuity correction: .

Test statistic	df	P value
1.571	1	0.21

Q4의 Red는 “선택을 고수한다”를 앞에, “선택을 바꾼다”를 뒤에 나오도록 하였고, Black은 “선택을 바꾼다”를 앞에, “선택을 고수한다”를 뒤에 나오도록 하여 소위 1번효과가 작동하는지를 살펴 보았습니다.

그 결과 Red, Black 의 차이를 분석하기 위한 카이제곱 통계량은 1.571, 자유도는 1, p-value 는 0.2100 으로 1번효과는 통계적으로 유의하지 않음을 보여 줍니다.

그런데 2023년 1학기에는 특이하게도 “바꾼다”를 앞에 놓은 Black 에서 “바꾼다”를 선택한 인원이 Red 보다 통계적으로 유의한 수준으로 많이 나와서 소위 1번효과를 관찰할 수 있습니다.

2021년 2학기에도 통계적으로 유의한 차이를 관찰하였고, 2017년 2학기에도 통계적으로 유의한 차이가 나온 적이 있지만 대체로 1번효과는 잘 나타나지 않습니다.

백분율로 비교합니다.

26.5.3 % 비교.

	고수한다	바꾼다	계
Red	70.1	29.9	100.0
Black	64.8	35.2	100.0

이를 백분율로 살펴보면 Red에서 “선택을 고수한다”는 백분율, 70.1%(은)는 “선택을 바꾼다”는 백분율, 29.9% 보다 월등히 높고, Black에서 “선택을 고수한다”는 백분율, 64.8%(은)는 “선택을 바꾼다”는 백분율, 35.2%보다 역시 월등히 높다는 것을 알 수 있습니다.

통계적으로 유의한 차이를 관찰하지 않았기 때문에 합쳐서 보는 게 타당합니다.

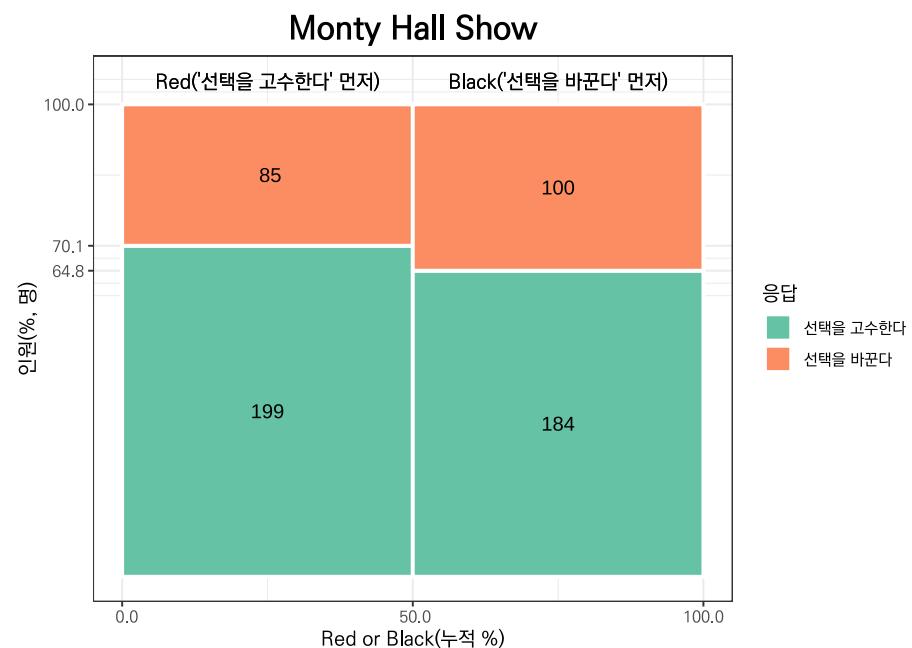
26.5.4 합산(%)

	고수한다	바꾼다	계
계	67.4	32.6	100.0

Red, Black 을 통합하여 보면 “선택을 고수한다”는 백분율 67.4%(은)는 ’선택을 바꾼다’는 백분율 32.6%보다 월등히 높습니다.

그만큼 우리의 직관이 취약하다는 것을 알 수 있습니다.

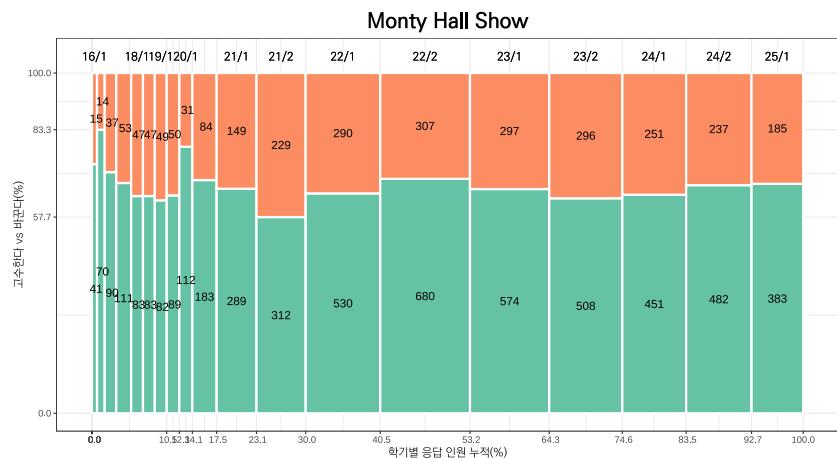
26.5.5 Mosaic Plot



26.5.6 학기별 응답 요약

학기별로 '고수한다'와 '바꾼다'의 비율을 Mosaic Plot으로 요약해 보았습니다.

여러분의 직관에 대해서 그리고 학기별로 꾸준히 관찰되는 거의 고정된 비율에 대해서 댓글을 올려 주세요.



26.6 마감 시간으로부터 제출 시간의 분포

26.6.1 분포표

표 26.22: 일 단위

	14일	13일	12일	11일	10일	9일	8일	7일	6일	5일	4일	3일	2일	1일	계
Red	119	21	9	4	4	3	8	39	9	8	17	7	16	20	284
Black	126	21	15	6	3	2	6	28	7	14	8	12	11	25	284
계	245	42	24	10	7	5	14	67	16	22	25	19	27	45	568

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.

26.6.2 날마다 고르게 제출하는가?

14일	13일	12일	11일	10일	9일	8일	7일	6일	5일	4일	3일	2일	1일
245	42	24	10	7	5	14	67	16	22	25	19	27	45

표 26.24: Chi-squared test for given probabilities: .

Test statistic	df	P value
1199	13	2.44e-248 * * *

날마다 고르게 제출하는지 알아 보았습니다.

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

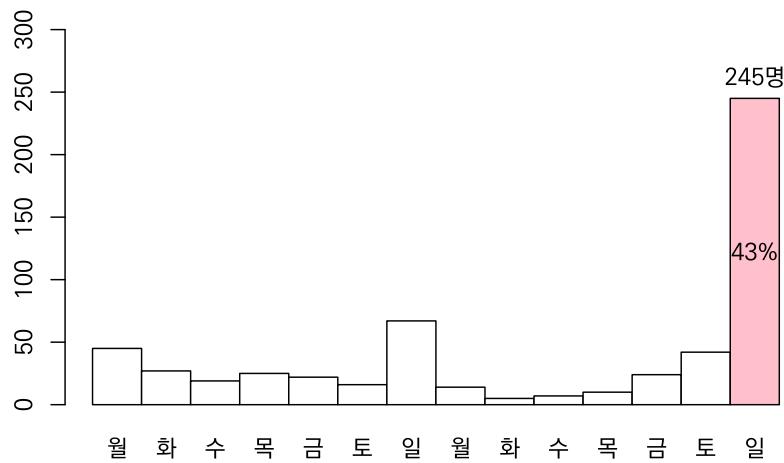
분포표 만으로도 쉽게 파악할 수 있지만 카이제곱테스트가 명확히 해 줍니다.

카이제곱 통계량은 1199.352, 자유도는 13.00, p-value 는 2.4e-248 이므로 날짜별로 고르게 제출하고 있지 않다는 것을 강력히 시사합니다.

막대그래프로 살펴 보겠습니다.

26.6.3 막대그래프

MH250519 (568명 제출)



26.6.4 Red, Black 간에 닮았는가?

	14일	13일	12일	11일	10일	9일	8일	7일	6일	5일	4일	3일	2일	1일
Red	119	21	9	4	4	3	8	39	9	8	17	7	16	20
Black	126	21	15	6	3	2	6	28	7	14	8	12	11	25

표 26.26: Pearson's Chi-squared test: .

Test statistic	df	P value
12.46	13	0.4905

제출시간의 분포가 Red, Black 간에 닮았는지 알아 보았습니다.

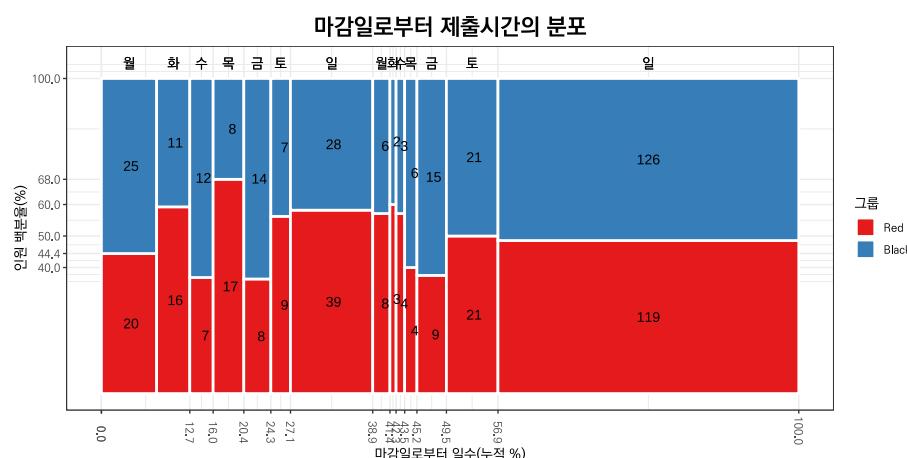
이번에는 분포표의 첫번째와 두번째 행, '계'열을 제외한 나머지 열에 대해서 카이제곱테스트를 수행합니다.

카이제곱 통계량은 12.46, 자유도는 13, p-value 는 0.4905 이므로 제출 시간의 분포는 Red, Black 간에 통계적으로 유의한 차이가 관찰되지 않습니다.

이 사실을 Mosaic Plot 을 이용하여 시각적으로 살펴보겠습니다.

닮았다고 느껴지나요?

26.6.5 Mosaic Plot



참고문헌

- American Statistical Association. n.d. "Statistics Is the Science of Learning from Data." <https://www.amstat.org/asa-newsroom>.
- Bogner, Kathrin, and Uta Landrock. 2016. "Response Biases in Standardised Surveys." GESIS – Leibniz Institute for the Social Sciences. https://www.gesis.org/fileadmin/admin/Dateikatalog/pdf/guidelines/response_biases_standardized_surveys_bogner_landrock_2016.pdf.
- Boynton, Peter M., and Trisha Greenhalgh. 2002. "A Catalogue of Bias in Questionnaires." BMJ: British Medical Journal 324 (7339): 752–55. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1122072/>.
- Freedman, David, Robert Pisani, Roger Purves, and Ani Adhikari. 1998. "Statistics." W. W. Norton & Company.
- Genroe. 2020. "Biased Questions: How to Identify & Fix Them in Surveys." <https://www.genroe.com/blog/biased-questions/15422>.
- Hills, Peter, and Michael Argyle. 2002. "The Oxford Happiness Questionnaire: A Compact Scale for the Measurement of Psychological Well-Being." Personality and Individual Differences 33 (7): 1073–82.
- Huff, Darrell. 1954. How to Lie with Statistics. W. W. Norton & Company.
- 중앙선거여론조사심의위원회. 2024. "종합부동산세 관련 여론조사 결과 (NBS, 2024년 7월 조사)." <https://www.nesdc.go.kr/>.