가설검정의 원리

사례: 심판의 동전던지기 조작

A매치 축구경기에서 동전던지기에 흔히 “피파 동전”이라는 것을 사용한다. 이는 국제축구연맹인 FIFA가 만들어서 국제심판들에게 배포한 것이다. 동전 앞면에는 FIFA 로고가 새겨져 있으며, 뒷면에는 페어플레이라는 영문이 새겨져 있다.

국가		승	무	패	승점	득점	실점
1	코스타리카	0	2	0	2	4	4
2	캐나다	0	2	0	2	2	2
3	대한민국	0	2	0	2	2	2

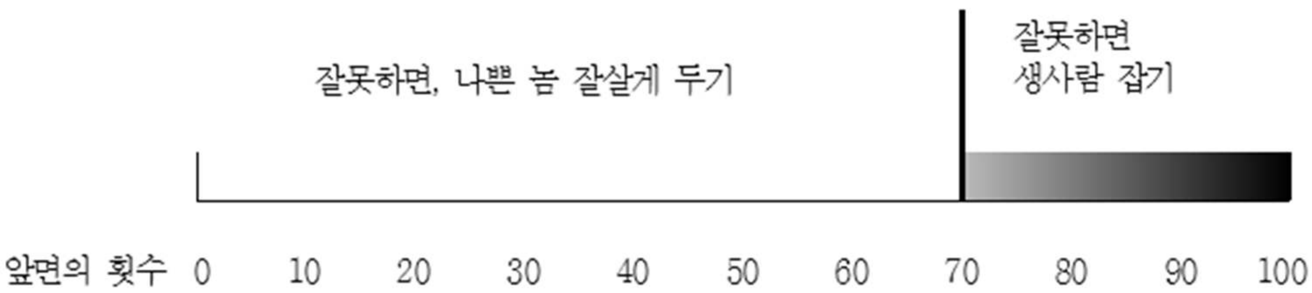
사례: 심판의 동전던지기 조작

- ◆ 앞면이 더 많이 나오는 동전으로 사용한다고 의심

연구문제	동전 조작이 있었는가? 그렇다면 어느 것이 조작된 동전인가?
연구가설	조작된 동전이다. 즉 앞면이 뒷면보다 더 많이 나오는 동전이다.

- ◆ 국제심판들이 사용하고 있는 동전을 100번씩 던져서, 앞면이 나오는 횟수를 관찰하기로 함

의사결정	조작된 동전이라고 판단하는 기준을 얼마로 해야 하는가?



가설의 설정: 귀무가설과 대립가설

- ◆ 통계적 검정을 수행하는 목적은 자신의 '새로운 주장'을 입증하려는 것
- ◆ '대립(對立)'이라는 말은 기존에 당연하다고 여겨지던 사실과 상반된다는 의미
- ◆ 당연하다고 여겨지던 사실은 귀무가설(영가설)
- ◆ 새로운 주장은 대립가설(연구가설)
- ◆ 가설검정은 대립가설을 받아들일지 여부, 또는 귀무가설을 충분히 기각할 사유가 있는지 확인하는 과정
- ◆ 질문
  - ●연구자가 관심있는 가설은 무엇인가?
  - ●연구자가 기존에 사실로 받아들여지던 귀무가설을 믿는다면 가설검정이 필요한가?
  - ●연구자의 머릿속에 진실로 생각하고 있는 것은?
  - ●그럼 새로운 사실을 쉽게 기각하고 연구가설을 택하면 되나?

가설의 설정: 귀무가설과 대립가설

◆ 가설검정시 주의 사항

- ●가설의 검정은 모집단에 대한 검정이다.
- ●다만 모집단에 대해 검정을 할 수 없으므로 표본을 이용한다.
- ●논문이나 보고서의 관심사항은 연구가설이다. 따라서 연구가설이 무엇인지 명확하게 기술해야 한다.

◆ 귀무가설의 설정

- ●연구자가 연구를 수행하기 이전까지 당연하게 여겨지던 상식
- ●일반적이고 평범하며 정상적인 상태
- ●변화 이전에 계속 유지되어온 상태

- ●참고: 귀무가설은 가설검정이 가능한 기준이 있어야 한다.
- ●예: ‘두 사람의 키는 다르다’ 라는 귀무가설?

P-값의 이해

◆ **법정의 가설**

- 귀무가설, H_0 : 피고는 무죄이다.
- 대립가설, H_1 : 피고는 유죄이다.

◆ **검사의 관심사**

- 유죄입증
- 유죄입증은 무죄가 아니라는 증거를 제시하여 귀무가설을 기각하게 만드는 것

◆ **변호사의 관심사**

- 무죄입증
- 무죄 입증은 유죄의 증거가 충분하지 않아 귀무가설을 기각하지 못하게 만드는 것

◆ **귀무가설은 채택의 대상이 아님, 기각이 가능한지 불가능하지의 대상**

◆ **대립가설은 귀무가설의 기각으로 인해 채택하게 되는 가설**

P-값의 이해

◆ P 값 (p-value)

● 귀무가설을 기각하고 대립가설을 채택한 결정이 틀릴 확률

p-값 (p-value)
❑ 주어진 관찰값에서의 p-값은 귀무가설이 맞는다는 가정 하에서 관찰값 이상의 극단적인 값이 얻어질 확률이다. ❑ 따라서 특정 관찰값을 귀무가설을 기각하는 기준으로 사용할 경우, 귀무가설이 맞는데 귀무가설을 기각하는 오류를 범할 확률을 의미한다. ❑ p-값이 작으면 작을수록, 관찰값은 귀무가설을 기각할 수 있는 강력한 증거가 된다.

가설검정의 두 가지 오류

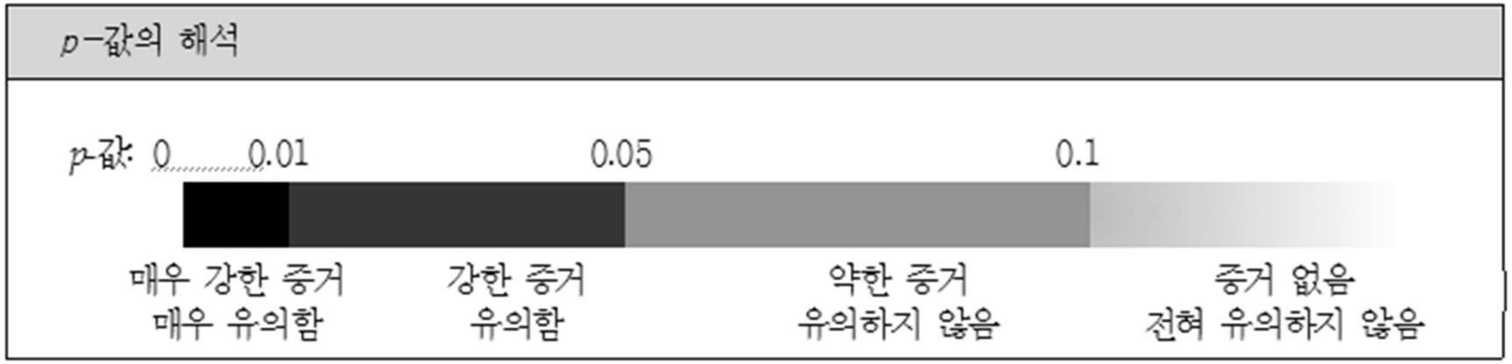
◆ 제1종 오류와 제2종 오류

실제로는 검정결과	귀무가설이 사실일 때 "피고가 무죄일 때"	대립가설이 사실일 때 "피고가 유죄일 때"
귀무가설을 기각한다 "피고에게 유죄를 판결한다"	제1종 오류 (α)	올바른 결정, 검정력 ($1-\beta$)
귀무가설을 기각하지 못한다 "피고에게 무죄를 판결한다"	올바른 결정 ($1-\alpha$)	제2종 오류(β)

유의수준의 설정과 *P*-값의 해석

◆ **유의수준의 설정과 *p*-값의 해석**

p-값과 귀무가설의 기각
❑ p-값 ≤ α이면, 유의수준 α에서 귀무가설을 기각한다. 이때 관찰값 혹은 분석결과는 통계적으로 유의하다. ❑ p-값 > α이면, 유의수준 α에서 귀무가설을 기각하지 못한다. 이때 관찰값 혹은 분석결과는 통계적으로 유의하지 않다.



◆ **보고서작성 Tip**

● 검정통계량이라는 값 위에 별을 붙이는데 *p*-값이 0.1과 0.05 사이에 있으면 별 하나★를 달고, 0.01보다도 작으면 별 두 개★★를 단다. 별 세 개★★★ 는 *p*-값이 0.001보다도 작다는 것을 의미

1종 오류와 2종 오류에 대한 대책

- 1종 오류와 2종 오류는 하나가 줄어들면 다른 오류가 줄어드는 trade-off가 있음
- 따라서 기각역의 설정만으로는 효과가 없음
- 1종 오류와 2종 오류에 가장 좋은 대책은 표본의 수를 늘이는 것임

가설검정의 절차

가설검정의 절차

1. 당신이 주장하는 대립가설과 함께, 이와 상반되는 귀무가설을 설정한다.

2. 동전던지기에서 앞면의 실제 관찰값과 같은 실증자료를 획득한다.

3. 귀무가설이 옳은데 이를 기각하는 오류의 허용범위인 유의수준을 설정한다.

4. 귀무가설이 옳다면 나왔어야 할 이론적 기대값을 계산한다.

5. 관찰값과 기대값의 차이를 확률적으로 설명할 수 있는 검정통계량의 값을 구한다.

6. 이론적 확률분포에 근거하여, 관찰된 유의수준인 **p-**값을 계산한다.

7. *p*-값과 유의수준을 비교하여, 귀무가설의 기각 여부를 결정한다.

주사위 던지기와 다항분포

- **베르누이 실험: 성공과 실패라는 두 가지 결과만 가능한 실험, 각 실험은 독립**
- **예) 동전던지기, 성별, 시험 합격여부**
- **베르누이 실험을 여러 번 반복하는 경우, 성공횟수는 이항분포를 따름**
  - X~Binomial(n,p)
- **다항실험: 2개 이상의 가능한 실험결과가 존재하며 가 각 실험은 독립**
  - 예) 주사위 던지기, 성적(A+, …, F)

카이제곱 독립성 검정을 통한 가설검정의 이해

◆ **연구문제1**

어느 날 한 친구가 도박장에서 주사위 놀이를 하였다. 게임은 그 친구 스스로 내기 돈의 액수와 함께, 자신이 승리하는 주사위 눈을 택하는 것으로 시작된다. 주사위 눈의 선택은 홀수 혹은 짝수로 결정한다. 그리고 주사위를 던져 자신이 선택한 눈이 나오면 베팅한 만큼 돈을 따고, 반대의 눈이 나오면 그 만큼 돈을 잃는 간단한 방식이었다. 게임은 밤새 계속되었고 처음에는 돈을 따고 잃기를 계속 반복했다. 하지만 어느 순간부터 대부분의 게임에서 잃기 시작하더니 결국 빈털터리가 되었다. 집까지 날렸다.

X_j	1	2	3	4	5	6	sum
O_j	11	4	15	9	14	7	60
%	18.3	6.7	25.0	15.0	23.3	11.7	100

◆ **Step 1. 가설의 설정**

●H_1 : 주사위에 적어도 하나의 눈의 발생비율은 1/6이 아니다.

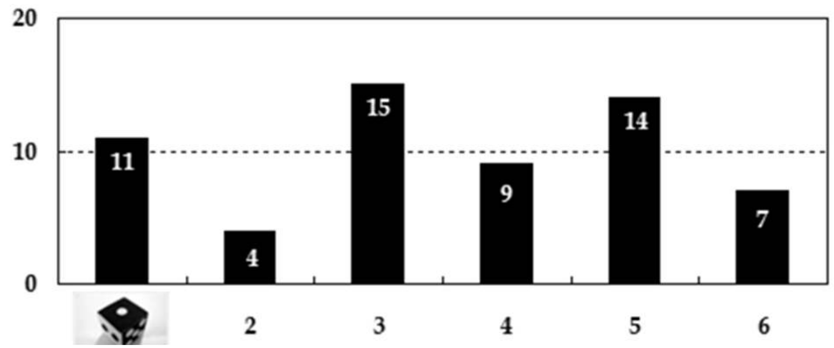
●H_0 : 주사위에 각 눈의 발생비율은 1/6 이다.

카이제곱 독립성 검정을 통한 가설검정의 이해

◆ **Step 2. 실증자료의 획득**

자신의 주장을 입증하기 위해, 도박에서 돈을 몽땅 날린 친구는 실증자료를 어떻게 획득할 수 있을까? 그 친구는 약간의 돈을 마련하여, 도박장을 한 번 더 방문했다. 목적은 주사위를 빼내오기 위한 것이었다. 하지만 주사위 관리요원이 따로 있을 정도로 보완은 삼엄했다. 마지막 게임에서 간신히 주사위 하나를 가지고 나올 수 있었다. 집에 돌아온 그 친구는 주사위를 던지는 실험을 시작했다. 주사위 각 눈에 대한 관찰값을 얻기 위한 것이었다. 시간이 흘러가면서 시행횟수는 증가했다. 확률에 의존하여 각 실험의 결과가 발생하는 확률과정(stochastic process)이 거듭되었다. 주사위를 60회 던진 결과, 주사위 각 눈 별로 아래와 같은 관찰빈도를 얻을 수 있었다.

X_j	O_j
1	11
2	4
3	15
4	9
5	14
6	7
sum	60



카이제곱 독립성 검정을 통한 가설검정의 이해

◆ **Step 3. 유의수준의 설정**

관례적으로 흔히 사용되는 유의수준 α = 0.05이다. 그러면 도박으로 알거지가 된 친구도 유의수준으로 0.05를 사용해야 하는가? 유의수준과 관련된 제1종 오류는 정상적인 주사위를 조작된 주사위라고 말하는 오류이다. 주사위에 조작이 있었던 것으로 판단되면, 도박장에 가서 항의하려는 친구가 이러한 오류를 범한다면 어떤 일이 벌어지겠는가? 조폭에게 피 터지게 얻어맞고, 문 밖으로 내던져지는 영화 속의 한 장면이 떠오르지 않는가. 이 친구는 유의수준을 0.01로 낮추기로 했다..

카이제곱 독립성 검정을 통한 가설검정의 이해

◆ **Step 4. 기대값의 계산**

X_j	1	2	3	4	5	6	sum
O_j	11	4	15	9	14	7	60
E_j	10	10	10	10	10	10	60
$O_j - E_j$	1	-6	5	-1	4	-3	0

카이제곱 독립성 검정을 통한 가설검정의 이해

◆ **Step 5. 검정통계량의 계산**

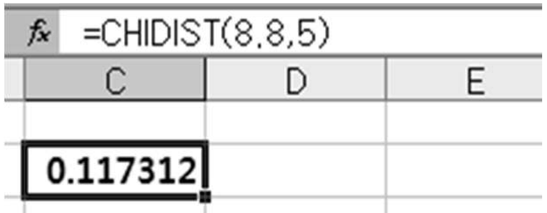
Xj	1	2	3	4	5	6	sum
O_j	11	4	15	9	14	7	60
E_j	10	10	10	10	10	10	60
$(Oj-Ej)2$	1	36	25	1	16	9	88

Xj	1	2	3	4	5	6	sum
O_j	11	4	15	9	14	7	60
E_j	10	10	10	10	10	10	60
$(O_j - E_j)^2 / E_j$	0.1	3.6	2.5	0.1	1.6	0.9	8.8

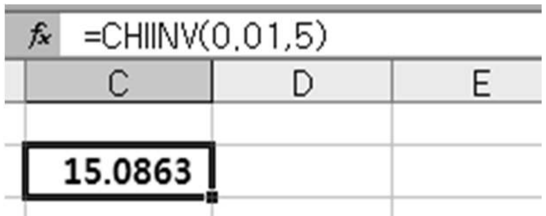
카이제곱 독립성 검정을 통한 가설검정의 이해

◆ **Step 5. 검정통계량의 계산**

엑셀함수를 이용하여, 주사위던지기와 관련된 검정통계량 = $\sum(O_j - E_j)^2 / E_j$ 의 값이 8.8보다 크게 나올 확률을 구해보자. 주사위 눈의 범주가 6가지이므로 자유도는 5이다.

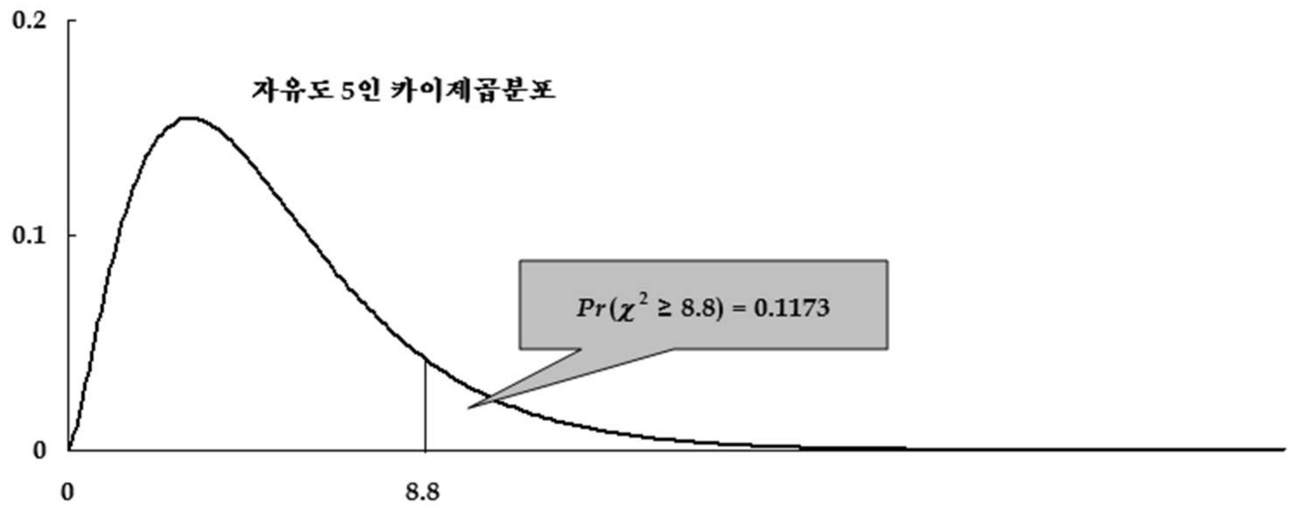


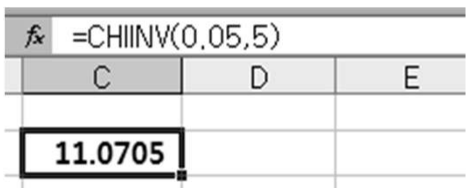
$Pr(\chi^2 \geq x) = 0.01$ 인 x 의 값을 구해보자. 즉 검정통계량 의 값이 얼마일 때, 그 값보다 클 확률이 0.01이 되겠는가? 이를 $\chi^2_{0.01}$ 로 표현한다.



카이제곱 독립성 검정을 통한 가설검정의 이해

◆ **Step 6. p-값의 계산**





카이제곱 독립성 검정을 통한 가설검정의 이해

◆ **Step 7. 귀무가설의 기각여부 결정**

?

카이제곱 독립성 검정을 통한 가설검정의 이해

◆ **Excel에서 쉽게 p-value 구하기**

관찰빈도와 기대빈도에 관한 정보를 가지고, 엑셀에서 바로 카이제곱 검정을 수행할 수 있다. 이때 사용되는 통계함수는 다음과 같으며, 결과로 *p*-값이 계산된다.

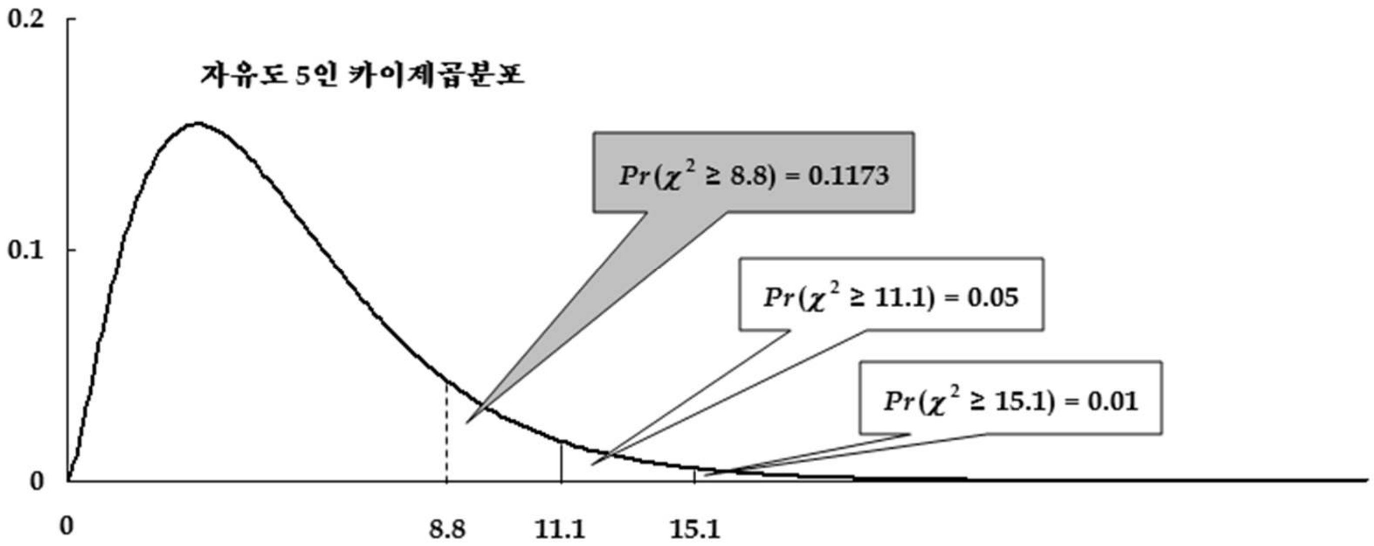
=CHITEST (actual_range, expected_range)

여기서 'actual_range'에는 관찰빈도가 있는 셀들을 지정하고, 'expected_range'에는 기대빈도가 있는 셀들을 지정하면 된다.

C2 =CHITEST(D5:I5,D6:I6)

	A	B	C	D	E	F	G	H	I	J
1										
2			0.117312							
3										
4			X_j	1	2	3	4	5	6	sum
5			O_j	11	4	15	9	14	7	60
6			E_j	10	10	10	10	10	10	60

기각역과 P-VALUE



0.2
0.1
0
자유도 5인 카이제곱분포
$Pr(\chi^2 \geq 8.8) = 0.1173$
$Pr(\chi^2 \geq 11.1) = 0.05$
$Pr(\chi^2 \geq 15.1) = 0.01$
0
8.8
11.1
15.1

연구문제 6-2

주사위를 60번 던진 결과, 각 눈이 나온 빈도가 다음과 같았다. 1의 눈 11회, 2의 눈 4회, 3의 눈 15회, 4의 눈 9회, 5의 눈 14회, 그리고 6의 눈 7회였다. 이러한 결과를 근거로, 홀수의 눈이 짝수의 눈보다 더 많이 나오는 주사위라고 주장할 수 있는가?

연구문제 6-2 풀이

C2 f_x =CHITEST(D5:E5,D6:E6)

	A	B	C	D	E	F
1						
2			0.0098			
3						
4			X_j	0	1	sum
5			O_j	20	40	60
6			E_j	30	30	60

연구문제 6-3

두 가지 형태의 완두콩 교배에서 나타나는 유전법칙에 대해 생각해보자. 멘델의 법칙(Mendelian theory)에 따르면, 다음 4가지 형태의 완두가 '9:3:3:1'이라고 한다. (1) 둥글고 황색(RY), (2) 주름지고 황색(WY), (3) 둥글고 녹색(RG), 그리고 (4) 주름지고 녹색(WG)이 그 분류이다. 만일 160개의 관찰도수가 각각 83, 36, 25, 16이었다면, 이 결과는 멘델의 이론과 일치한다고 말할 수 있는가? 유의수준 0.01에서 검정하라.

연구문제 6-3 풀이

C2 f_x =CHITEST(D5:G5,D6:G6)

	A	B	C	D	E	F	G	H
1								
2			0.1033					
3								
4			X_j	RY	WY	RG	WG	sum
5			O_j	83	36	25	16	160
6			E_j	90	30	30	10	160