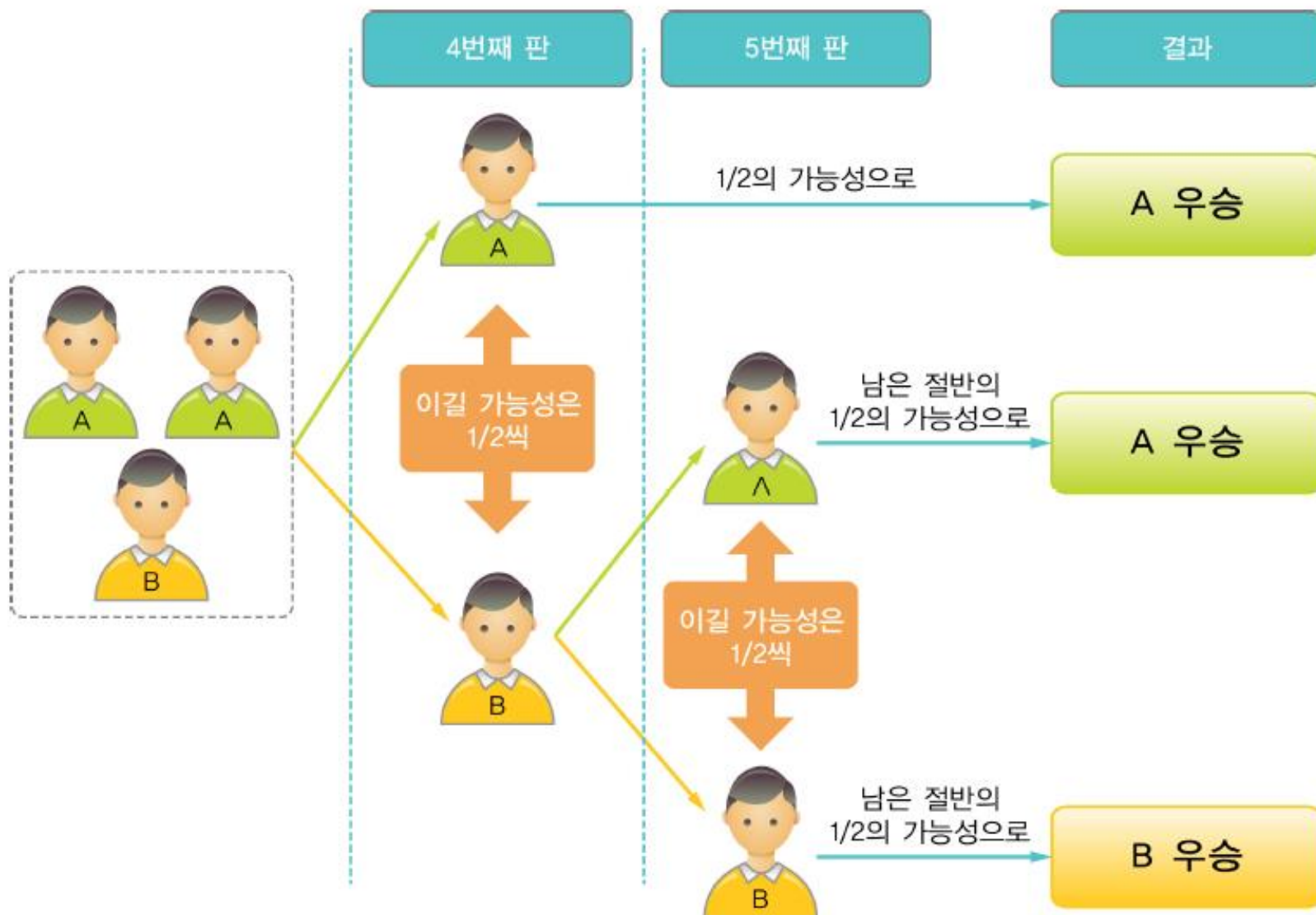


1047

- 어떻게 나눠야 불만이 없을까요?
  - 서로 능력이 동일한 두 명의 게이머 A와 B가 있습니다.
  - 5판 3선승제의 게임을 통해 우승자를 가리는 게임이 시작되었습니다.
    - 우승자에게는 64개의 게임 캐릭터 배지를 상품으로 주는 이벤트 대결입니다.
  - 3판이 지난 지금 A가 두번을 이겼고 B가 한번을 이겼습니다.
  - 4번째 판이 시작되려는 순간 게임 서버의 이상으로 더 이상 게임을 진행할 수 없게 되었습니다.
  - A가 우승했다고 하기도 곤란하고, B가 우승했다고 하기도 곤란한 상황입니다.
  - 상품을 어떻게 나누는 것이 좋을까요?



- 확률이라는 단어
  - 確率(확률) : 굳을 확, 비율 른
    - '(어떤 결정 등을) 굳힐 비율'
  - probability
    - probable : (명) '(어떤 일이) 있을 것 같은', '개연성 있는'
    - '개연성' 혹은 '개연성 있는 일'
      - 개연성의 사전적 의미 : 어떤 일이 일어날 수 있는 확실성의 정도

- 용어

- 확률실험 :  $E$

- 다음의 세가지를 만족할 때 확률실험 혹은 확률시행이라고 합니다.

- ❶ (결과를 구하기 위한) 어떤 실험을 통해 나타나는 결과를 알지 못한다.
- ❷ 결과는 알지 못하지만 결과로 나타날 수 있는 가능한 경우를 알고 있다.
- ❸ 동일한 실험을 몇 번이고 반복할 수 있다.

- ❶ 동전을 던지기 전에 '앞면'이 나올지 '뒷면'이 나올지 알 수 없습니다.
- ❷ 가능한 결과는 '앞면'과 '뒷면' 중에 하나임을 알고 있습니다.
- ❸ 동전을 던지는 실험은 몇 번이고 반복할 수 있습니다.

- 용어

- 표본공간 :  $\Omega$

- 확률실험으로부터 출현 가능한 모든 결과들의 모임을 표본공간이라 합니다.
    - 예제 : 동전던지기
      - $\Omega = \{\text{앞면, 뒷면}\} = \{H, T\}$

- 사건 : 기호 알파벳 대문자

- 표본공간의 각 원소(즉, 출현 가능한 개별 결과)들의 부분집합을 사건이라 합니다.
    - “(관심 있는)사건이 발생했다” : 시행 결과가 (관심있는)사건에 속하는 경우
    - 근원사건
      - 어떤 사건이 표본공간상의 하나의 원소로 구성된 사건

- 사건의 연산 : 임의의 두 사건  $A, B$ 에 대해
  - 합사건
    - 어떤 사건의 발생이 사건  $A$ 에서 일어나거나 혹은 사건  $B$ 에서 일어나는 사건
    - $A \cup B = \{\omega \mid \omega \in A \text{ or } \omega \in B\}$
  - 곱사건
    - 어떤 사건의 발생이 사건  $A$ 와 사건  $B$ 에서 동시에 일어나는 사건
    - $A \cap B = \{\omega \mid \omega \in A \text{ or } \omega \in B\}$
  - 여사건 :  $A^c$ 
    - 사건  $A$ 가 발생하지 않을 사건
    - $A^c = \{\omega \mid \omega \notin A\}$
  - 배반사건
    - 두 사건이 겹치는 부분이 없는 즉, 동시에 발생하지 않는 사건( $A \cap B = \phi$ )
- ❖ 독립사건
  - ❖ 두 사건이 서로의 발생에 영향을 끼치지 않는 사건

- 예 : 확률실험, 표본공간, 사건
  - $E_1$  : 동전을 2번 던져 나오는 면 관찰
    - 표본공간 :  $\Omega_1 = \{HH, HT, TH, TT\}$
    - 사건 : 첫 번째 동전이 앞면이 나오는 사건
      - $A_1 = \{HH, HT\}$
  - $E_2$  : 하루 중 인터넷 사용시간 관찰
    - 표본공간 :  $\Omega_2 = \{0 \leq t \leq 24\}$
    - 사건 : 사용시간이 1시간 이하인 사건
      - $A_2 = \{0 \leq t \leq 1\}$



- 확률의 정의

- 수학적 확률 (고전적 확률)

- ❶ 어떤 시행의 결과로 나타날 수 있는 가능한 결과의 수 :  $O$
- ❷ 각 결과들이 나타날 가능성은 동일하다는 가정
- ❸ 동일한 각 결과들의 확률 :  $\frac{1}{O}$

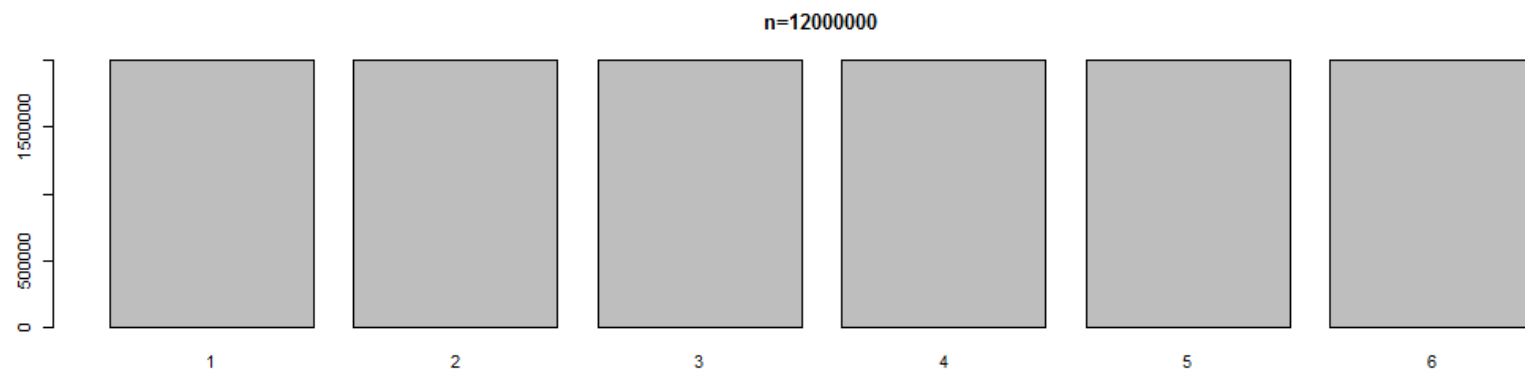
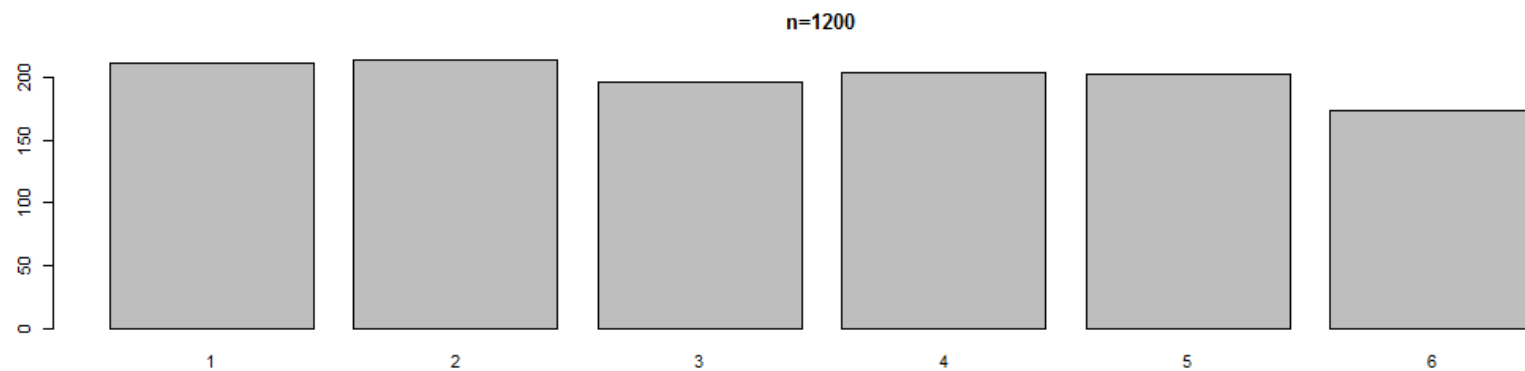
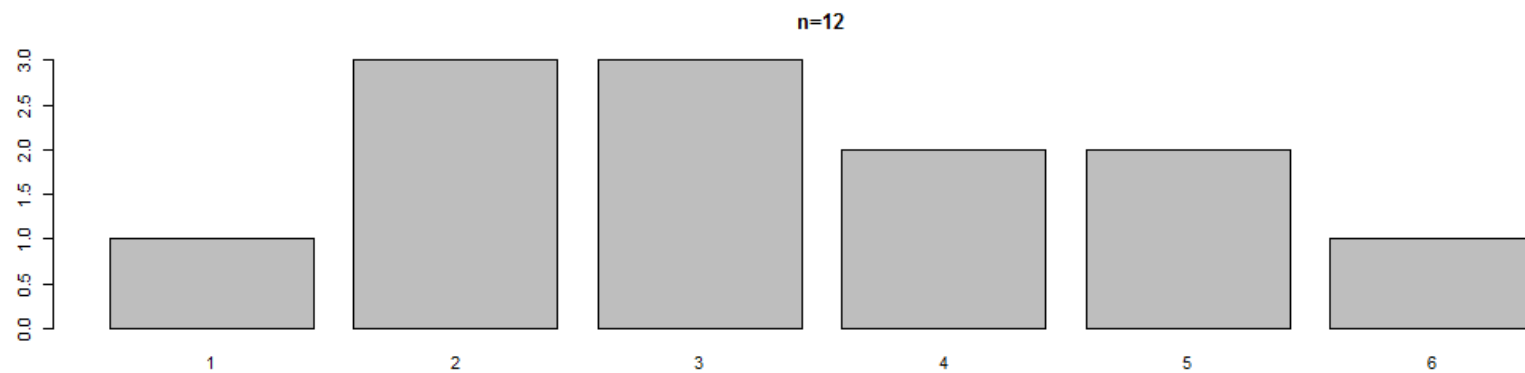
- 임의의 사건  $A$ 가 발생할 수학적 확률은 표본공간의 원소의 개수( $O$ ) 중 사건  $A$ 에 해당하는 근원사건의 개수( $n$ )입니다 ( $\frac{n}{O}$ ).
- 예) 주사위를 굴려 홀수가 나올 확률
  - 주사위의 각 눈이 나올 확률은 전체 6개의 눈으로 구성되어 있으며 각각이 나올 확률은 동일하다고 가정하면 확률은  $1/6$ 입니다.
  - 홀수인 사건을 구성하는 근원사건의 수는  $\{1, 3, 5\}$ 으로 세 개가 있습니다.
  - 전체 눈의 개수는 6이고 이로부터 홀수눈의 확률은  $3/6 = 1/2$  입니다.

- 통계적 확률 (고전적 확률)

- ❶ 동일한 조건에서 같은 실험을  $N$ 번 반복
- ❷ 사건  $A$ 가 모두 몇 번 발생했는지를 조사 :  $n$
- ❸ 사건  $A$ 가 발생할 확률 :  $P(A) = \frac{n}{N}$

- 실험의 반복횟수  $N$ 은 매우 커야 그 값을 받아들일 수 있으며, 반복횟수가 커짐에 따라 사건  $A$ 의 상대도수( $\frac{n}{N}$ )가 상수  $P(A)$ 로 접근해가는 경향을 보입니다.
- 예) 주사위를 여러 번 굴려 나온 눈을 관찰해 봅시다.

시행횟수	1의 눈	2의 눈	3의 눈	4의 눈	5의 눈	6의 눈
12	1	3	3	2	2	1
1200	211	214	196	204	202	173
12000000	2002632	1999749	2000328	1999958	1996037	2001296



- 확률 공리 (공리적 확률)

- 표본공간  $\Omega$ 상의 임의의 사건  $A$ 에 대한 실수치 함수에 대해

- ❶  $P(A)$ 는 0과 1사이의 값을 갖고( $0 \leq P(A) \leq 1$ ),
- ❷ 반드시 일어나는 사건(표본공간 전체)의 값은 1이며( $P(\Omega) = 1$ ),
- ❸ 서로 배반인 사건  $A_1, A_2, \dots, A_n, \dots$ 의 합집합에 대해 다음을 만족하면,

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

함숫값  $P(A)$ 를 사건  $A$ 의 확률이라 합니다.

- 확률 공리는 확률이 만족해야 하는 기본 성질이며 이를 통해 확률 계산을 합니다.

- 확률법칙

- 덧셈법칙

- 임의의 사건 A와 사건 B의 합사건에 대한 확률

- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

- 만일 두 사건 A와 B가 서로 배반이라면 ( $A \cap B = \phi$ )

- $P(A \cup B) = P(A) + P(B)$

- 곱셈법칙

- 조건부 확률

- 두 사건 A와 B에 대해

- $P(A | B)$ : 사건 B가 발생했을 때 사건 A가 발생할 확률

- $P(A|B) = \frac{P(A \cap B)}{P(B)}, \quad P(B) > 0$

- $P(B | A)$ : 사건 A가 발생했을 때 사건 B가 발생할 확률

- $P(B|A) = \frac{P(A \cap B)}{P(A)}, \quad P(A) > 0$

- 조건부 확률 예)

- 주사위를 던지는 실험에서 주사위의 눈이 짝수인 사건을 **A**, 주사위의 눈이 4이상인 사건을 **B**라 할 때  $P(A|B)$ 를 구해봅시다.

- $P(A|B) = \frac{P(A \cap B)}{P(B)}$  이므로  $P(B)$ 와  $P(A \cap B)$ 를 구합니다.

- $P(B)$  : 주사위의 눈이 4 이상이 나올 확률 =  $1/2$

- $P(A \cap B)$  : 주사위의 눈이 짝수이고 4 이상인 경우는 {4, 6} 이므로 확률은  $1/3$  입니다.

- $P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{1/3}{1/2} = \frac{2}{3}$

- 표본공간의 변화 : 표본공간으로 조건으로 주어진 사건으로 변화

- $\Omega = \{1, 2, 3, 4, 5, 6\} \rightarrow \Omega_B = \{4, 5, 6\}$

- 변화된 표본공간  $\Omega_B$ 에서 사건 **A**가 발생할 확률

- $\Omega_B$  상에서 짝수의 눈은 {4, 6} 이므로  $2/3$

- 곱셈법칙

- 두 사건 A와 B에 대해 조건부 확률을 이용하여

- $$P(A \cap B) = \begin{cases} P(A)P(A|B), & P(A) > 0 \\ P(B)P(B|A), & P(B) > 0 \end{cases}$$

- 만일 두 사건 A와 B가 독립이라면,

- $$P(A \cap B) = P(A)P(B)$$

- 두 사건이 서로의 발생에 영향을 끼치지 않는다면 곱사건의 확률은 두 사건의 곱이 됩니다. (예, 동전을 두 번 던져 처음 앞면이 나온 것이 두번째 던졌을 때 영향을 끼치지 않습니다.)

- 독립사건과 조건부 확률

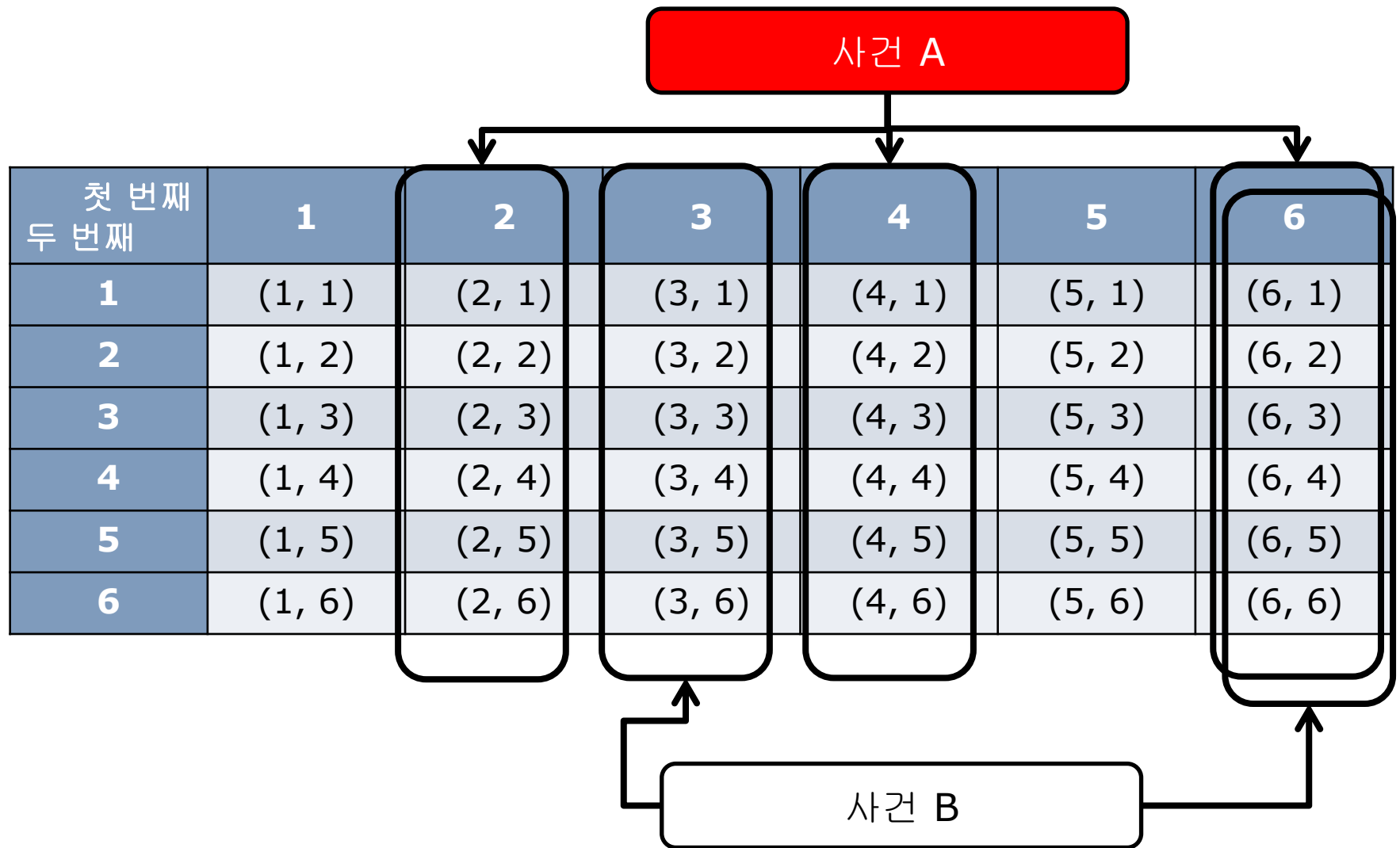
- 위에서 두 사건 A와 B가 독립이면,  $P(A \cap B) = P(A)P(B)$  이므로

- $$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)P(B)}{P(B)} = P(A)$$

- 즉, 사건 B의 발생여부가 사건 A의 발생에 영향을 끼치지 않습니다.

- 여사건의 확률
  - 사건  $A$ 의 여사건  $A^c$ 의 사건  $P(A^c)$ 는,
    - $P(A^c) = 1 - P(A)$
    - $P(A) + P(A^c) = 1$
- 예제) 주사위를 두 번 던지는 실험에서 다음의 두 사건에 대해 합사건, 곱사건의 확률을 구해 봅시다.
  - 사건  $A$ : 첫 번째 던진 주사위의 눈이 짝수인 사건,  $P(A) = \frac{18}{36} = \frac{1}{2}$
  - 사건  $B$ : 두 번째 던진 주사위의 눈이 3의 배수인 사건,  $P(B) = \frac{12}{36} = \frac{1}{3}$





- 합사건
  - $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
  - $A \cap B = \{ (6, 1), (6, 2), (6, 3), (6, 4), (6, 5), (6, 6) \}$  이므로  $P(A \cap B) = 1/6$
  - $P(A \cup B) = P(A) + P(B) - P(A \cap B) = 1/2 + 1/3 - 1/6 = 2/3$
- 곱사건 : 사건 B를 조건으로
  - $P(A \cap B) = P(B)P(A|B)$
  - $P(A|B) = P(A \cap B) / P(B)$ 로 사건 B를 새로운 표본공간으로 변화시켜 봅시다.

첫 번째 두 번째	1	2	3	4	5	6
3	(1, 3)	(2, 3)	(3, 3)	(4, 3)	(5, 3)	(6, 3)
6	(1, 6)	(2, 6)	(3, 6)	(4, 6)	(5, 6)	(6, 6)

- `prob` 패키지를 이용해 봅시다.
  - 동전던지기, 주사위 굴리기, 주머니 속에서 공 뽑기 등의 확률실험
  - 표본공간을 만들고 확률을 구해 봅시다.

```
① library(prob)
② tosscoin(1)
③ rolldie(1)
④ urnsamples(1:3, size=2)
⑤ urnsamples(1:3, size=2, replace=T)
⑥ urnsamples(c( rep("R", 3), rep("B", 2)), size=2)
⑦ tosscoin(2, makespace=T)
```

# 확률변수

- 동전을 두 번 던지는 실험을 생각해봅시다.
  - 이 실험에서 나올 수 있는 결과는  $\{H, H\}, \{H, T\}, \{T, H\}, \{T, T\}$ 입니다.
  - '(동전을 두 번 던져 앞면이 나오는 횟수)  $\times$  1,000원'의 상금이 주어지는 게임을 한다고 하면,
    - 우리의 관심사는 동전을 두 번 던져 앞면이 나오는 횟수가 될 것입니다.
    - 동전을 두 번 던져 앞면이 나오는 횟수는 다음의 표와 같습니다.

# 확률 변수

첫 번째 던진 동전	두 번째 던진 동전	표본공간	앞면이 나오는 횟수
		$\{H, H\}$	2
		$\{H, T\}$	1
		$\{T, H\}$	1
		$\{T, T\}$	0

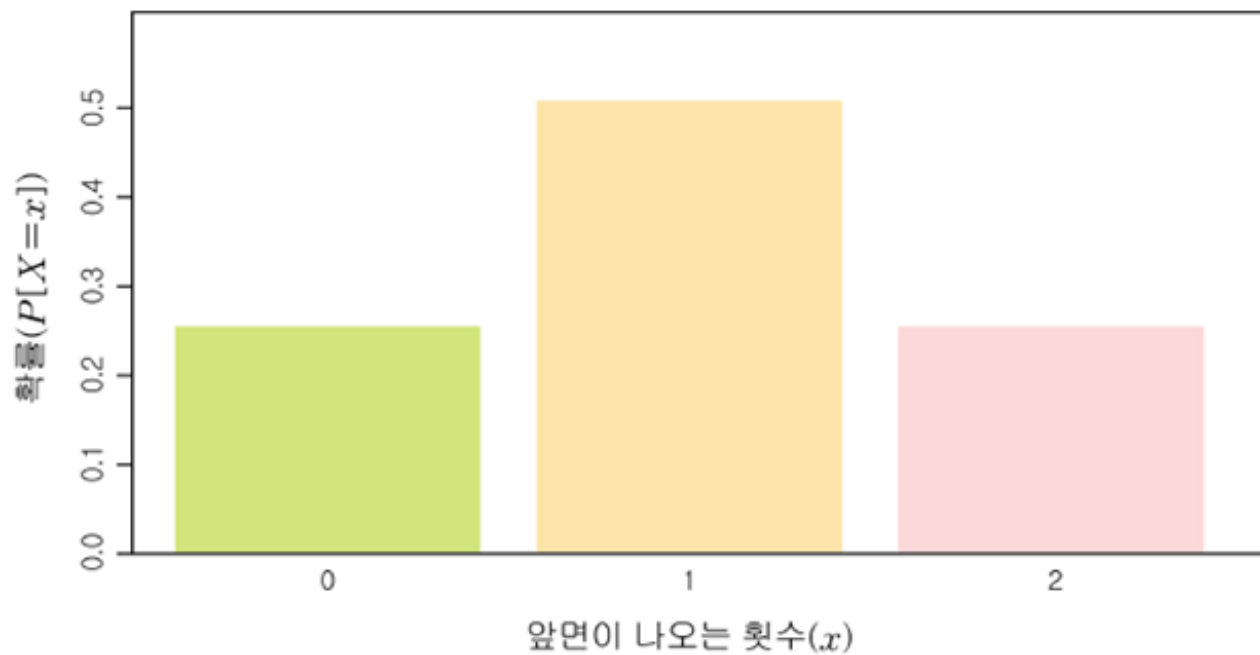
# 확률변수

- 확률변수

- '동전을 두 번 던져 앞면이 나오는 횟수'처럼 표본공간의 각 원소를 실숫값에 대응시키는 함수를 확률변수라 합니다.
  - 확률변수는 알파벳 대문자  $X, Y, Z, \dots$
  - 확률변수가 취하는 실숫값은 알파벳 소문자  $x, y, z \dots$
  - '확률변수  $X$ 가 값 실숫값  $x$ 를 가질 때'는  $X=x$ 로 표기합니다.
- 확률변수  $X$ 가 가질 수 있는 모든  $x_i$ 들에 확률이 대응되고, 확률변수는 이 확률에 따라 실숫값을 갖습니다.
- 확률변수  $X$ 가 특정 값  $x$ 를 가지는 사상  $X=x$ 의 확률을  $P(X=x)$ 로 표기합니다.
- 확률분포
  - 확률변수가 취할 수 있는 값과 각 값이 나타날 확률을 대응시킨 관계(함수) (표 3-4)
- 확률변수가 가질 수 있는 값에 따라 이산형 확률변수와 연속형 확률변수가 있습니다.

# 확률변수

표본공간	표본공간에서의 확률	$X=x$	$P(X=x)$
{H, H}	1/4	0	1/4
{H, T}	1/4	1	1/2
{T, H}	1/4		
{T, T}	1/4	2	1/4



# 확률변수

- 확률변수의 평균과 분산
  - 확률변수 X가 '동전을 두번 던져 앞면이 나오는 횟수'일 때의 평균과 분산을 구해 봅시다.
  - 확률변수의 평균
    - 확률변수 X가 가질 수 있는 값은 {0, 1, 2}
    - 상수 0, 1, 2에 대한 평균
      - $\bar{X} = \sum_{i=1}^n \frac{1}{n} \cdot x_i = \frac{1}{3}(0 + 1 + 2) = 1$
      - 상수의 평균에서  $\frac{1}{n}$ 은 각 자료들이 모두 동일하게  $\frac{1}{n}$ 의 비중을 갖고 있음을 나타냅니다.
      - 확률변수에서는  $\frac{1}{n}$ 에 해당하는 비중이 각 값이 나타날 확률로 바뀝니다.
    - 확률변수 X의 평균 :  $E(X) = \sum_{\text{모든 } x} x \cdot P(X = x)$



# 확률변수

- $E(X)$  는 확률변수  $X$ 의 평균을 나타내는 기호로, 확률변수의 평균을 기댓값이라 합니다.
- 확률변수  $X$ 의 평균 즉, 기댓값을 구해봅시다.

- $E(X) = \sum_{i=1}^x x_i \cdot P(X = x_i) = 0 \cdot \frac{1}{4} + 1 \cdot \frac{2}{4} + 2 \cdot \frac{1}{4} = \frac{(0 \cdot 1 + 1 \cdot 2 + 2 \cdot 1)}{4} = \frac{4}{4} = 1$

- 확률변수의 분산

- 분산은 편차 제곱의 평균입니다.
  - 확률변수의 분산에 그대로 적용해 봅시다.

- $Var(X) = E[(X - E(X))^2]$

- 확률변수의 분산은  $Var(X)$  혹은  $V(X)$ 로 나타냅니다.

- 기댓값을 구할 때와 마찬가지로 상수 자료들의 분산을 구할때 사용한  $1/n$ 이 확률로 바뀐다는 점 외에 나머지는 동일합니다.

- $E[(X - E(X))^2] = \sum_{\text{모든 } x} (x - E(X))^2 P(X = x)$

# 확률변수

- 분산의 간편식

- $Var(X) = E(X^2) - [E(X)]^2 = \sum_{\text{모든 } x} x^2 P(X = x) - [E(X)]^2$

- '확률변수의 제곱의 기댓값( $E(X^2)$ )'을 구한 후 '기댓값의 제곱( $[E(X)]^2$ )'을 뺍니다.

- 확률변수 X의 분산을 구해봅시다.

- 분산의 간편식을 이용합니다.

- 앞서 구한 기댓값( $E(X)$ )은 1입니다.

- 확률변수의 제곱의 기댓값을 구해봅시다.

- $E(X^2) = \sum_{i=1}^3 x_i^2 \cdot P(X = x_i) = 0^2 \frac{1}{4} + 1^2 \frac{2}{4} + 2^2 \frac{1}{4} = \frac{(0^2 \cdot 1 + 1^2 \cdot 2 + 2^2 \cdot 1)}{4} = \frac{6}{4}$

- $Var(X) = E(X^2) - [E(X)]^2 = \frac{6}{4} - 1^2 = \frac{2}{4} = \frac{1}{2}$

- 기댓값과 분산은 확률변수의 특성을 파악하는 좋은 모수입니다.

- R을 이용해 확률변수의 기댓값과 분산을 구해 봅시다.

# 확률변수

- 내용

- R로 확률변수의 평균과 분산을 구해 봅시다.
- 이를 통해 R에서 벡터 계산에 대해 알아보시다.
- 확률변수  $X$ 는 앞서 사용한 동전을 '두 번 던져 앞면이 나오는 횟수'로  $X$ 가 가지는 값은  $\{0, 1, 2\}$ 이고 각각 등장할 확률은  $\{1/4, 1/2, 1/4\}$  입니다.

① `x <- c(0, 1, 2)`

② `px <- c(1/4, 2/4, 1/4)`

③ `EX <- sum( x * px )`

④ `EX`

⑤ `x * 2`

⑥ `x * (1:6)`

⑦ `x * (1:4)`

⑧ `VX <- sum(x^2 * px) - EX^2`

⑨ `VX`