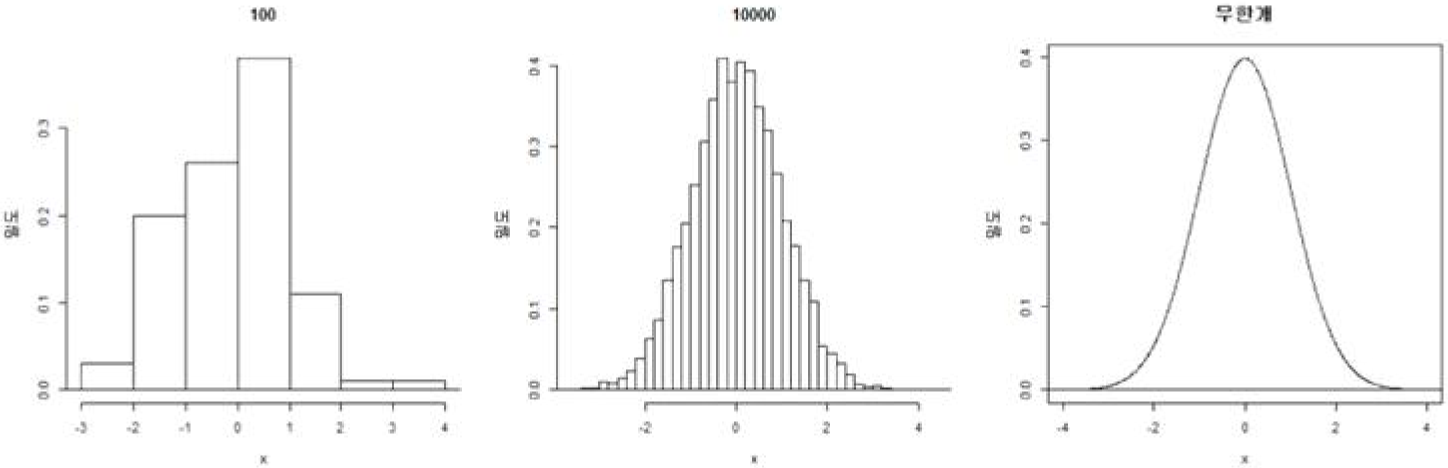
■ 확률밀도함수(probability density function)

- 연속확률변수: 확률변수의 치역이 실수
- 히스토그램
  - 밀도(density): 히스토그램의 높이
  - 전체 면적 = 1

SOOKMYUNG WOMEN'S UNIVERSITY K-MOOC

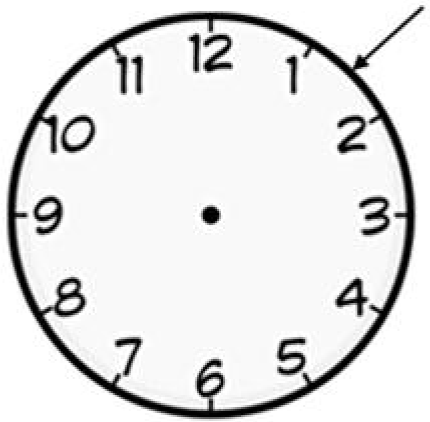
● 연속자료로 이루어진 모집단에서 표본추출



- ○ n= 100, 10000 ⇨ 표본
- ○ $n=\infty$ 모집단: x에서의 높이(밀도) = $f(x)$

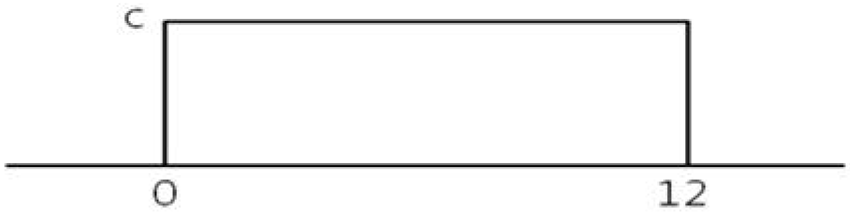
확률밀도함수 ↵

◉ 0~12까지의 숫자가 표시된 돌림판



- 표본공간: $\Omega = \{x : 0 < x \le 12\}$
- X: 바늘이 지적하는 위치
- 0에서 12사이에서 발생가능성이 동일

⇨ 밀도는 이 구간에서 동일: $f(x) = c$



- 전체 면적 = 1 ⇒ $f(x) = \frac{1}{12}, \quad 0 < x \le 12$

- 확률밀도함수에서의 확률
  - 히스토그램의 면적 = 해당 구간에서의 비율(상대도수)
  - 확률밀도함수의 면적 = 해당 구간에서의 확률
  - X가 구간 $[a, b]$에 속할 확률:

$$P(a \le X \le b) = \int_a^b f(x)\,dx$$

◉ 0~12까지의 숫자가 표시된 돌림판

○ X가 3에서 6사이에 있을 확률

$$P(3 \leq X \leq 6) = \frac{3}{12} = \frac{1}{4}$$

Q. X=3 일 확률은? P(X=3)=0

- X가 연속확률변수일 때

  - 모든 x에 대해 $P(X=x)=0$
  - $$P(a<X<b)=P(a<X\leq b)$$
    $$=P(a\leq X<b)=P(a\leq X\leq b)$$

  - 확률밀도함수 $f(x)$는 x에서의 확률이 아니라 그 위치에서 상대적으로 얼마나 밀집되어 있는지를 나타낸 것

● 확률밀도함수의 성질

① 모든 x에 대해, $f(x) \geq 0$

② $\int_{-\infty}^{\infty} f(x)\,dx = 1$

③ $P(a < X \leq b) = \int_{a}^{b} f(x)\,dx$

SOOKMYUNG WOMEN'S UNIVERSITY K-MOOC

- 누적분포함수(cumulative distribution function)

  - 성질 ③의 특수한 형태

$$P(X \le x) = \int_{-\infty}^{x} f(u)\,du = F(x)$$

  - 예】 0~12까지의 돌림판

    - $F(x) = \begin{cases} 0, & x \le 0 \\ \dfrac{x}{12}, & 0 < x < 12 \\ 1, & x > 12 \end{cases}$

    - 점프가 없음 ⇒ 임의의 점에서의 확률은 0

- 요약
  - 확률밀도함수: 해당지점에서의 상대적 밀도
    - X가 구간 $[a, b]$에 속할 확률:

$$P(a \le X \le b) = \int_a^b f(x)\,dx$$

    - 모든 x에 대해 $P(X = x) = 0$

  - 누적분포함수

$$P(X \le x) = \int_{-\infty}^{x} f(u)\,du = F(x)$$