제 5장

확률변수와 확률함수

고려대학교 경영대학 박 광태

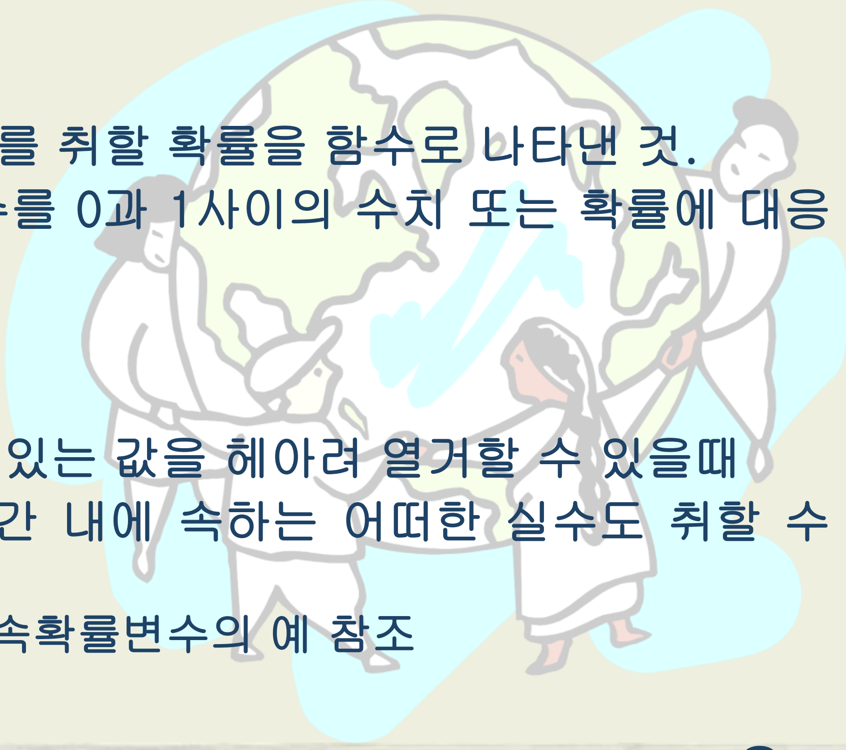
확률변수와 확률함수

- 확률변수
  - 변수 X가 갖는 값을 확실히 예측할 수 없을 때 그 변수를 확률변수라 함.
  - 실험의 결과치를 실수에 대응시키는 함수

- 확률함수
  - 확률변수 X가 어떤 특정 실수 x를 취할 확률을 함수로 나타낸 것.
  - 확률변수에 대하여 정의된 실수를 0과 1사이의 수치 또는 확률에 대응시키는 함수

- 이산확률변수와 연속확률변수
  - 이산확률변수 : 변수가 취할 수 있는 값을 헤아려 열거할 수 있을때
  - 연속확률변수 : 주어진 실수구간 내에 속하는 어떠한 실수도 취할 수 있을 때

    ※ p.106의 이산확률변수와 연속확률변수의 예 참조



확률변수와 확률함수

- P.106의 이산확률변수와 연속확률변수의 예

이산확률변수의 예:

① 예제 1의 변수 X; $X=0, 1, 2, 3$

② 한 주일 동안에 발생한 특정제품의 판매량; $X=0, 1, 2, 3, \cdots$

③ 회계장부의 각 쪽당 오류횟수; $X=0, 1, 2, 3, \cdots$

연속확률변수의 예:

① 예제 2의 변수 Z; $0 \le Z < \infty$

② 집에서 회사까지의 운전시간; 만일 집에서 회사까지 운전하는데 걸리는 시간이 최소 20분에서 최대 1시간 30분 사이에서 발생한다면, $20 \le X \le 90$

③ 생산라인에서 출하되는 완제품의 불량률; $0 \le X \le 1$

확률변수와 확률함수

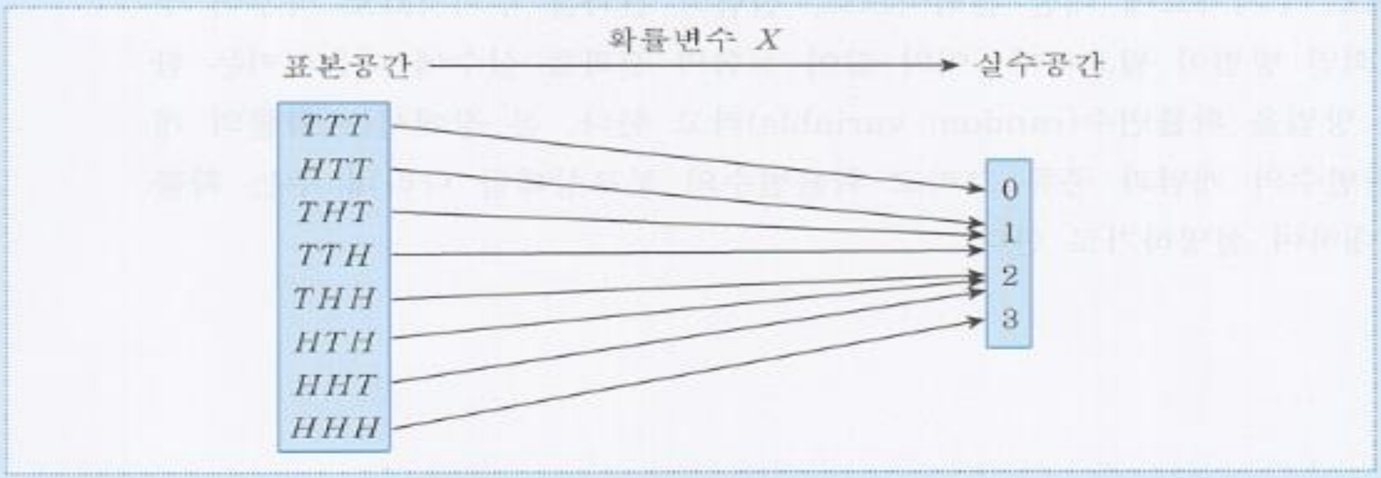
예제 5-1

동전을 세 번 던지고 그 결과를 기록하는 실험의 표본공간 S는 다음과 같다.

$$S=\{HHH,\ HHT,\ HTH,\ THH,\ TTH,\ THT,\ HTT,\ TTT\}$$

위의 표본공간에 대하여 확률변수 X를 '동전 앞면의 수'라고 정의하자. 그러면 표본공간 S는 X라고 하는 함수 또는 방법에 의하여 그림 5-1과 같이 여러 개의 실수값에 대응된다.

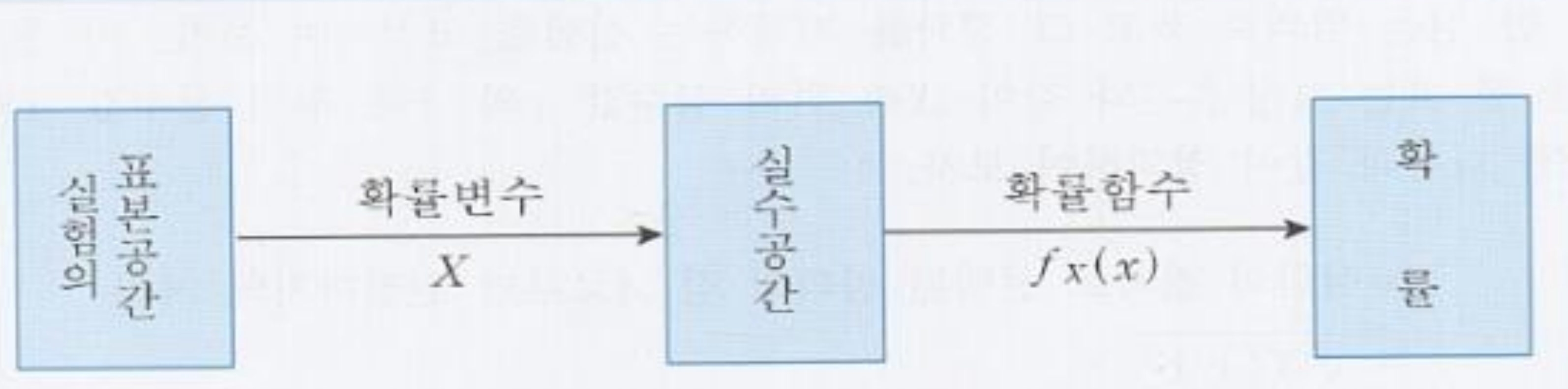
그림 5-1 확률변수



즉, 확률변수 X가 취할 수 있는 실수, 또는 X의 표본공간 S_X는 $\{0,\ 1,\ 2,\ 3\}$이다.

확률변수와 확률함수

그림 5-3 확률함수



확률밀도함수

- 확률밀도함수
  - 확률밀도함수는 연속확률변수 X가 취할 수 있는 실수구간에 대하여 확률을 대응시키는 방법 또는 규칙을 말한다.

- 확률밀도함수의 조건
  - 모든 X에 대하여, $f(x) \geq 0$
  - 모든 X에 대한 확률의 합은 1이다.
    즉,

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

확률밀도함수

예제 5-5

한국과 중국과의 관계가 개선되어 양국 간의 경제교류가 활발해짐에 따라 제일컴퓨터 주식회사는 중국에 퍼스널 컴퓨터를 수출하기로 결정하였다. 제일기업은 연간 판매량 X(단위: 1,000)가 다음의 확률함수에 따라 발생할 것으로 추정하고 있다.

$$f_X(x) = \begin{cases} \dfrac{(x-30)}{450}, & 30 \le x \le 60 \\ 0, & \text{다른 구간에서} \end{cases}$$

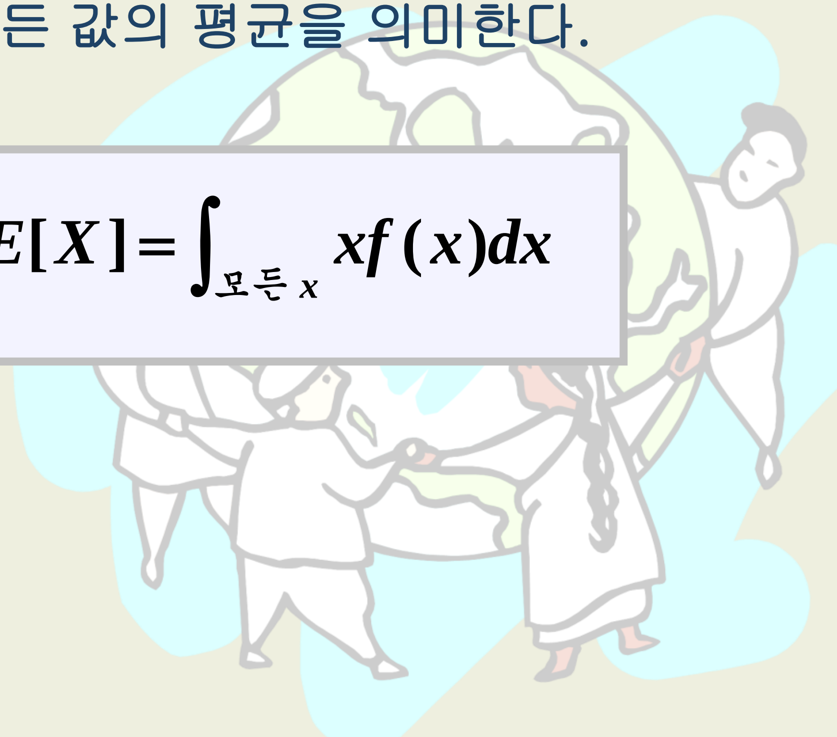
위의 확률함수를 그래프로 나타내고, 연간 판매량이 50,000대 이상일 확률을 구하시오.

기대값 (1)

- 기대값의 정의

  - 기대값은 자료의 중심적 경향을 나타내 주는 수치적 척도로써, 확률변수가 취할 수 있는 모든 값의 평균을 의미한다.

  - X가 연속변수인 경우: $$E[X] = \int_{\text{모든 } x} x f(x)\,dx$$



기대값 (2)

- 기대값의 특성 (a와 b는 상수, X와 Y는 확률변수)

  1. $E(a)=a$
  2. $E(bX)=bE(X)$
  3. $E(a+bX)=a+bE(X)$
  4. $E(X+Y)=E(X)+E(Y)$
  5. $E(aX+bY)=aE(X)+bE(Y)$

기대값 (2)

예제 5-15

변수 X는 표본공간이 $S_X = \{-2, 1, 2\}$인 이산변수이며, 각 값에 대한 확률은 $p_X(-2) = \frac{5}{8}$, $p_X(1) = \frac{1}{8}$, $p_X(2) = \frac{2}{8}$이다. 이를 이용하여 함수 $g(X) = X^2$의 기대값을 구하시오.

기대값 (2)

예제 5-16

다음과 같은 밀도함수를 가지는 연속함수가 있다고 하자.

$$f_X(x) = \begin{cases} 2(1-x), & 0 \le x \le 1 \\ 0, & \text{다른 구간에서} \end{cases}$$

$g(X) = X^2$의 기대값을 구하시오.

기대값 (2)

예제 5-17

동화기업은 MDF를 제조하여 가구업체에 판매하는 가공목재 제조업체이다. 동화에서 제조되는 MDF는 두께가 모두 5cm이며, 가로와 세로의 길이는 제품에 따라 상이하다. 가로(Y)는 1m, 2m, 3m, 5m의 네 가지가 있으며, 세로(X)는 1m, 2m, 3m의 세 가지 길이로 판매한다. 지난달에 판매된 MDF 제품은 10,000개이며, 제품별 분포는 다음의 표와 같다고 하자.

세로(X) \ 가로(Y)	1	2	3	5	$P(X)$
1	0.10	0.15	0.12	0.08	0.45
2	0.05	0.10	0.08	0.02	0.25
3	0.06	0.14	0.06	0.04	0.30
$P(Y)$	0.21	0.39	0.26	0.14	1.00

기대값 (2)

- 예제 5-17(계속)

$g(X, Y) = xy$일 때, $g(X, Y)$의 기대값을 구하라. 그리고 $g(X, Y)$의 기대값은 무엇을 의미하는가?

분산

분산

$$Var(X) = \int [x - E(X)]^2 f(x)dx$$
$$= \int x^2 f(x)dx - 2E(X)\int xf(x)dx + [E(X)]^2 \int f(x)dx$$
$$= E(X^2) - E^2(X)$$

분산의 성질

- a와 b가 상수일 때,

$$Var(aX+b) = a^2\, Var(X)$$

- 변수 X와 Y가 독립일 때,

$$Var(X+Y) = Var(X) + Var(Y)$$

분산

예제 5-20

변수 X는 기대값이 μ_X이고 분산이 σ_X^2인 연속확률변수이다. 그리고 변수 Z는 X의 함수로써 다음과 같이 정의된다.

$$Z = \frac{X - \mu_X}{\sigma_X}$$

변수 Z의 기대값과 분산은 무엇인가?

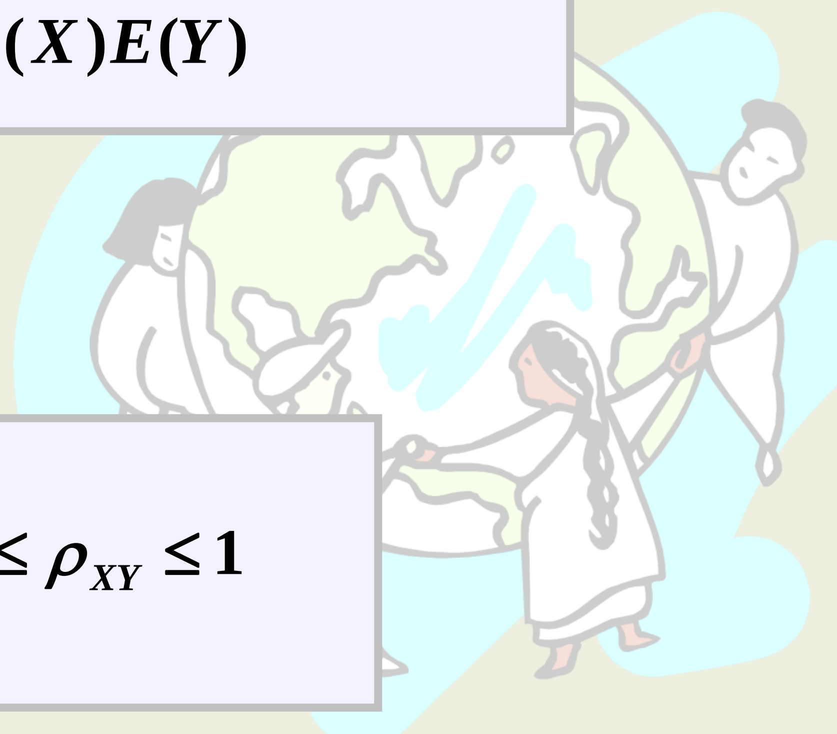
공분산

- 공분산

$$Cov(X,Y) = E[(X - E(X))(Y - E(Y))]$$
$$= E(XY) - E(X)E(Y)$$

- 상관계수

$$\rho_{XY} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}; \quad -1 \le \rho_{XY} \le 1$$



공분산

그림 5-6 변수 X와 Y의 관계