2장 확률변수

§2.1. 이산 확률 변수

> ***Definition***
>
> 확률실험의 표본공간 Ω 위에서 정의된 함수
>
> $$X : \Omega \to \mathbb{R}$$
> $$\omega \mapsto X(\omega) = x_\omega$$
>
> 를 Ω 위에서 정의되는 **확률변수**(random variable)라 하고, 확률변수 X가 취하는 모든 실수 값들의 집합
>
> $$R_X = \{ X(\omega) = x_\omega \mid \omega \in \Omega \}$$
>
> 을 X의 **치역공간**(range space)이라 한다.

참고 (1) 치역공간 R_X가 유한집합이거나 가산무한집합일 때, X를 **이산 확률 변수**(discrete random variable)이라 한다.

$$R_X = \{ x_{\omega_1}, x_{\omega_2}, \cdots, x_{\omega_n} \} \text{ 또는 } \{ x_{\omega_i} \mid 0 \le i < \infty \}$$

(2) 치역공간 R_X가 하나의 구간 또는 유한개 구간들의 집합족으로 주어지면, X를 **연속 확률 변수**(continuous random variable)이라 한다.

$$R_X = (-\infty, \infty),\ [0, \infty) \text{ 또는 } (-1, 0) \cup (0, 1)$$

Example

(1) 동전을 세 번 던지는 확률실험에서의 표본공간은

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

이다. 이 실험에서 표본점에 나타나는 '앞면(H)의 개수'로 확률변수 X를 정의할 때, X의 치역공간을 구하여라.

(2) 한 쌍의 주사위를 던지는 확률실험에서의 표본공간은

$$\Omega = \{ (i, j) \mid i, j = 1, 2, 3, 4, 5, 6 \}$$

이다. 이 실험에서 두 주사위의 눈의 차를 변수 X라 할 때, X의 치역공간을 구하여라.

(3) 주사위 눈이 1이 나올 때까지 반복하여 주사위를 던진 횟수로 확률변수 X를 정의 할 때, X의 치역공간을 구하여라.

(4) 어떤 공장에서 생산되는 상품의 수명을 관찰하고, 그 값을 확률변수 X라 할 때, X의 치역공간을 구하여라.

▣ 확률함수와 확률분포

> ***Definition***
>
> X를 확률공간 $(\Omega, \mathfrak{F}, P)$ 위에서 정의된 이산 확률 변수이라 하자.
>
> 확률변수 X의 치역공간 R_X 안의 모든 값 x에 대하여
>
> $$A_x = \{ \omega \in \Omega \mid X(\omega) = x \}$$
>
> 일 때, 다음과 같이 정의된 함수
>
> $$p_X : R_X \to \mathbb{R}$$
> $$x \mapsto p_X(x) = P(X = x) = P(A_x)$$
>
> 를 X의 **확률 함수**(probability function)라 하고, 집합
>
> $$\{ (x, p_X(x)) \mid x \in R_X \}$$
>
> 를 X의 **확률 분포**(probability distribution)라 한다.

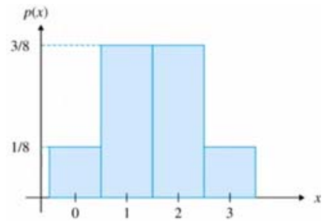
참고 (1) 모든 $x \in R_X$에 대하여, $p_X(x) \ge 0$이다.

(2) $\displaystyle\sum_{x \in R_X} p_X(x) = 1$

(3) 이산확률변수 X의 치역공간 R_X안의 각 점 x에 대한 다음 표를 X의 **확률분포표**라 한다.

x	x_1	x_2	$\cdots$	x_n	$\cdots$	합
$P(X=x)$	$p_X(x_1)$	$p_X(x_2)$	$\cdots$	$p_X(x_n)$	$\cdots$	1

(4) 이산확률변수 X의 치역공간 R_X안의 각 점 x를 중심으로 밑면의 길이가 1이고 높이가 $p_X(x)$인 사각형들의 그림을 X의 **확률히스토그램**이라 한다.



Example

동전을 세 번 던지는 확률실험에서의 '앞면(H)의 개수'로 확률변수 X를 정의할 때, X의 확률분포표를 완성하여라.

x					
$P(X=x)$					

Example

한 쌍의 주사위를 던지는 확률실험에서 두 주사위의 눈의 차를 확률변수 X로 정의할 때, X의 확률분포표를 완성하여라.

x							
$P(X=x)$							

Example

확률변수 X가 취할 수 있는 값 $-2, -1, 0, 1, 2$에 대하여,

$$p_X(-2)=p_X(2)=0.1,$$

$$p_X(-1)=p_X(1)=0.25$$

일 때, $P(X=0)$을 구하여라.

Definition

X가 확률공간 $(\Omega, \mathfrak{F}, P)$ 위에서 정의된 이산 확률 변수일 때, 다음과 같이 정의된 함수 $f_X : \mathbb{R} \to \mathbb{R}$ 를 확률변수 X의 **확률질량함수**(probability mass function)라 한다.

$$f_X(x)=\begin{cases} p_X(x), & x \in R_X \\ 0, & x \notin R_X \end{cases}$$

참고 (1) 모든 $x \in \mathbb{R}$ 에 대하여, $f_X(x) \geq 0$이다.

(2) 모든 $x \in \mathbb{R}$ 에 대하여,

$$P(X \leq x)=\sum_{\substack{x_i \leq x \\ (x_i \in R_X)}} f_X(x_i)$$

이다.

(3) 임의의 $B \subset \mathbb{R}$ 에 대하여

$$X^{-1}(B)=\{\omega \in \Omega \mid f_X(\omega) \in B\}$$

이면, $X \in B$인 사건의 확률은

$$P(B)=P(X^{-1}(B))$$

로 정의한다.

Example

주사위를 두 번 반복하여 던지는 게임에서 나온 두 눈의 합을 확률변수 X라 할 때 다음을 구하여라.

(1) 확률변수 X의 상태공간

(2) X의 확률분포표와 확률질량함수

(3) 두 눈의 합이 7이상 10이하일 확률

Definition

X가 확률공간 $(\Omega, \mathfrak{F}, P)$에서 정의된 이산확률변수일 때,

$$F_X(x)=P(X \leq x)=\sum_{\substack{x_i \leq x \\ (x_i \in R_X)}} f_X(x_i)$$

로 정의되는 함수 $F_X : \mathbb{R} \to \mathbb{R}$ 을 확률변수 X의 **누적분포함수**(cumulative distribution function)라 한다.

Example

이산확률변수 X의 확률질량함수 f_X가 다음과 같을 때, X의 누적분포함수 F_X를 구하여라.

$$f_X(x)=\begin{cases} \frac{1}{4}, & x=0,2 \\ \frac{1}{2}, & x=1 \\ 0, & \text{otherwise} \end{cases}$$

Example

동전을 세 번 던져서 '앞면'이 나온 횟수 X에 대한 누적분포함수 F_X를 구하여라.

Theorem

이산확률변수 X의 누적분포함수 F_X는 다음 성질을 갖는다.

(1) 단조증가 :

$$x_1 < x_2 \Rightarrow F_X(x_1) \le F_X(x_2)$$

(2) 오른쪽 방향으로의 연속성 :

$$\lim_{h \to 0^+} F_X(x+h) = F_X(x)$$

(3) ① $F_X(-\infty) = \lim_{x \to -\infty} F_X(x) = 0$,

② $F_X(\infty) = \lim_{x \to \infty} F_X(x) = 1$

참고 $F_X(x^+) = \lim_{h \to 0^+} F_X(x+h)$, $F_X(x^-) = \lim_{h \to 0^+} F_X(x-h)$

Theorem

임의의 실수 a, b에 대하여 다음이 성립한다.

(1) $P(a < X \le b) = P(X \le b) - P(X \le a)$

$$= F_X(b) - F_X(a)$$

(2) $P(X = b) = \lim_{h \to 0^+} P(b-h < X \le b) = F_X(b) - F_X(b^-)$

(3) $P(X < b) = P(X \le b) - P(X = b) = F_X(b^-)$

(4) $P(a \le X \le b) = P(a < X \le b) + P(X = a)$

$$= F_X(b) - F_X(a) + P(X = a)$$

(5) $P(X \ge a) = 1 - P(X < a) = 1 - F_X(a^-)$

Corollary

이산확률변수 X의 누적분포함수가 F_X일 때,

(1) 만약 $a \not\in R_X$이면,

$$\begin{aligned} F_X(a) - F_X(a^-) &= P(X = a) \\ &= P(X \le a) - P(X < a) \\ &= \sum_{\substack{x_i \le a \\ (x_i \in R_X)}} f_X(x_i) - \sum_{\substack{x_i < a \\ (x_i \in R_X)}} f_X(x_i) = 0 \end{aligned}$$

이다.

(2) F_X가 $x = a$에서 점프불연속점을 가지면, $a \in R_X$이다.

Example

이산확률변수 X의 누적분포함수 F_X가 다음과 같을 때,

$$F_X(x) = \begin{cases} 0 & , \ x < 0 \\ 0.2 & , \ 0 \le x < 5 \\ 0.45 & , \ 5 \le x < 10 \\ 0.85 & , \ 10 \le x < 15 \\ 1 & , \ x \ge 15 \end{cases}$$

(1) X의 확률질량함수 f_X를 구하여라.

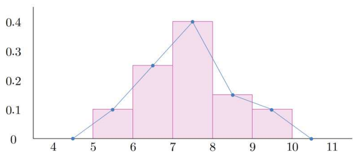
(2) $P(3 < X \le 10)$, $P(5 \le X \le 12)$를 구하여라.

§2.2. 연속확률변수

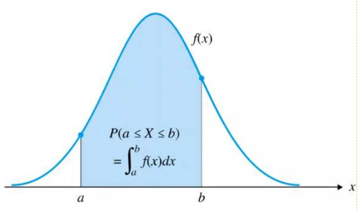
Example

다음은 큐브 동호회 회원 20명의 기록을 조사한 표와 $\dfrac{\text{상대도수}}{\text{계급의 크기}}$에 대한 히스토그램과 분포다각형이다.

기록(초)	도수(명)	상대도수
5이상 ~ 6미만	2	0.1
6 ~ 7	5	0.25
7 ~ 8	8	0.4
8 ~ 9	3	0.15
9 ~ 10	2	0.1
합계	20	1



$$\sum\left(\frac{\text{상대도수}}{\text{계급의 크기}}\times \text{계급의 크기}\right)=\sum \text{상대도수}=1$$



$$\int_{R_X} f_X(x)\,dx=1$$

Definition

확률공간 $(\Omega, \mathfrak{F}, P)$에서 정의된 연속확률변수 X에 대하여, 함수 $f_X : \mathbb{R}\to\mathbb{R}$ 가 다음 두 조건

(1) $f_X(x)\ge 0$

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

을 만족하면, f_X를 연속확률변수 X의 **확률 밀도 함수**(probability density function)라고 한다.

참고 (1) $P(a\le X\le b)=\displaystyle\int_a^b f_X(x)\,dx$

(2) $P(X=a)=P(a\le X\le a)=\displaystyle\int_a^a f_X(x)\,dx=0$

Theorem

(1) $P(a\le X\le b)=P(X=a)+P(a<X<b)+P(X=b)$
$\qquad =P(a<X<b)$

(2) $P(a\le X\le b)=P(a<X<b)$
$\qquad =P(a<X\le b)=P(a\le X<b)$

Example 1

연속확률변수 X의 확률밀도함수가 다음과 같다.

$$f_X(x)=\begin{cases} ke^{-2x}, & x\ge 0 \\ 0, & x<0 \end{cases}$$

(1) 상수 k를 구하여라.

(2) 확률 $P(1<X\le 2)$를 구하여라.

(3) 확률 $P(X\ge 1.5)$를 구하여라.

(sol)

Definition

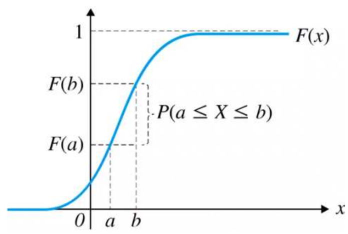
확률공간 $(\Omega, \mathfrak{F}, P)$에서 정의된 연속확률변수 X에 대하여,

$$F_X(x)=P(X\le x)=\int_{-\infty}^{x} f_X(t)\,dt$$

를 확률변수 X의 **누적분포함수**(cumulative distribution function)라 한다.

참고 (1) $F_X(b)-F_X(a)=P(a<X\le b)$
$\qquad =P(a<X<b)$
$\qquad =P(a\le X<b)=P(a\le X\le b)$

(2) x가 $\dfrac{d}{dx}F_X(x)=\dfrac{d}{dx}\displaystyle\int_{-\infty}^{x} f_X(t)\,dt=f_X(x)$



Theorem

연속확률변수 X의 누적분포함수 F_X는 다음 성질을 갖는다.

(1) F_X는 단조증가한다.

$$x_1<x_2 \Rightarrow F_X(x_1)\le F_X(x_2)$$

(2) F_X는 모든 실수범위에서 연속이다.

(3) ① $F_X(-\infty)=\displaystyle\lim_{x\to-\infty} F_X(x)=0$

② $F_X(\infty)=\displaystyle\lim_{x\to\infty} F_X(x)=1$

(4) 모든 실수 x에 대하여 $0\le F_X(x)\le 1$이다.

Example 2

[예제 1]의 확률밀도함수를 갖는 연속확률변수 X에 대하여 다음을 구하여라.

$$f_X(x)=\begin{cases} 2e^{-2x}, & x\geq 0 \\ 0, & x<0 \end{cases}$$

(1) 누적분포함수 $F_X(x)$

(2) F_X를 이용한 확률 $P(1<X\leq 2)$

(3) F_X를 이용한 확률 $P(X\geq 1.5)$

Example

연속확률변수 X의 누적분포함수 $F_X(x)$가 다음과 같다.

$$F_X(x)=\begin{cases} 0, & x<0 \\ A+Be^{-x}, & x\geq 0 \end{cases}$$

(1) 상수 A, B를 구하여라.

(2) $P(2<X\leq 5)$

(3) 확률밀도함수를 구하여라.

Definition

확률공간 $(\Omega, \mathfrak{F}, P)$에서 정의된 확률변수 X의 치역공간이

$$R_X=R_{X_d}\cup R_{X_c},\quad R_{X_d}\cap R_{X_c}=\varnothing$$

이고, X_d가 이산확률변수, X_c가 연속확률변수일 때, 각각의 누적분포함수를 F_{X_d}, F_{X_c}라 하자.

임의의 $0<\alpha<1$에 대하여

$$F_X(x)=\alpha F_{X_d}(x)+(1-\alpha)F_{X_c}(x)$$

를 누적분포함수로 갖는 확률변수 X를 **혼합 확률 변수** (mixed random variable)라 한다.

Example

대흥역에 지하철이 8분 간격으로 도착하여 1분 동안 정차한다고 한다. 어떤 사람이 역에 도착하여 지하철을 기다리는 시간을 X라 할 때, X의 누적분포함수를 구하여라.

Example

확률변수 X의 누적분포함수가 다음과 같을 때, 다음을 구하여라.

$$F_X(x)=\begin{cases} 0, & x<0 \\ \frac{1}{2}x, & 0\leq x<1 \\ 1-\frac{1}{4x^2}, & x\geq 1 \end{cases}$$

(1) $P(X=0)$
(2) $P(X=1)$
(3) $P(X\leq 1)$
(4) $P(X<1)$
(5) $P(X\geq 2)$

§2.3. 기댓값 (Expected Value)

Example

주사위를 던지는 확률실험에서 나온 눈의 수를 X라 할 때, X의 평균 μ_X을 구하여라.

Definition

(1) 이산확률변수 X가 확률질량함수가 $f_X(x)$일 때, 이산확률변수 X의 **기댓값** $E[X]$ (또는 **평균** μ_X)은

$$\mu_X = E[X] = \sum_{x \in R_X} x f_X(x)$$

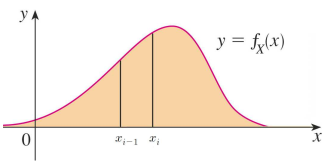
로 정의한다.

(2) 연속확률변수 X의 확률밀도함수가 $f_X(x)$일 때, 연속확률변수 X의 **기댓값** $E[X]$ (또는 **평균** μ_X)은

$$\mu_X = E[X] = \int_{-\infty}^{\infty} x f_X(x)\, dx$$

로 정의된다.

참고



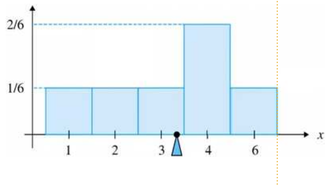
$$\begin{aligned}\mu_X &= \lim_{n\to\infty}\sum_{i=1}^{n} x_i P(x_{i-1} \le X \le x_i) \\ &= \lim_{n\to\infty}\sum_{i=1}^{n} x_i \frac{F_X(x_i) - F_X(x_{i-1})}{x_i - x_{i-1}} \Delta x \\ &= \int_{R_X} x f_X(x)\, dx\end{aligned}$$

Example

동전을 두 번 던지는 게임에서 앞면이 나온 횟수를 X라 하자. X의 기댓값 $E[X]$을 구하여라.

Example

주사위 눈이 1, 2, 3, 4, 4, 6으로 된 주사위를 던질 때, 나온 눈의 수를 X라 하자. X의 평균 μ_X을 구하여라.



Example

흰 공이 2개, 검은 공이 3개로 모두 5개의 공이 들어 있는 주머니에서 동시에 임의로 3개의 공을 꺼낼 때, 그 속에 포함된 흰 공의 개수를 X라 하자. X의 기댓값 $E[X]$를 구하여라.

Example 1

500원짜리 동전과 100원짜리 동전이 각각 5개씩 들어있는 주머니에서 동시에 임의로 동전 4개를 꺼낼 때, 꺼낸 동전 안에 포함된 100원짜리 동전의 개수를 확률변수 X라 하자. X의 평균 μ_X을 구하여라.

Example 2

다음 확률밀도함수를 갖는 연속확률변수 X의 기댓값 $E[X]$를 구하여라.

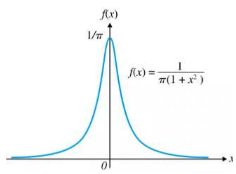
$$f_X(x) = \begin{cases} 2e^{-2x}, & x \ge 0 \\ 0, & x < 0 \end{cases}$$

Example

연속확률변수 X가 확률밀도함수

$$f_X(x) = \frac{1}{\pi}\frac{1}{1+x^2}, \quad -\infty < x < \infty$$

를 갖는다고 하자. X의 기댓값 $E[X]$을 구하여라.



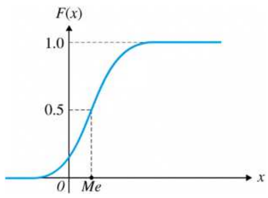
Definition

확률변수 X의 누적분포함수 $F_X(x)$에 대하여,

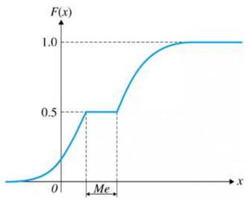
$$F_X(x_0)=0.5$$

를 만족하는 상수 x_0를 확률변수 X의 **중앙값**(median)이라 하고, 이 값을 기호로 $M_e=x_0$로 나타낸다.

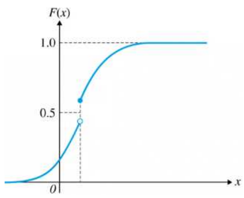
참고 $P(X\le M_e)=P(X>M_e)=0.5$



중앙값의 의미



중앙값이 무수히 많은 경우



중앙값이 없는 경우

Definition

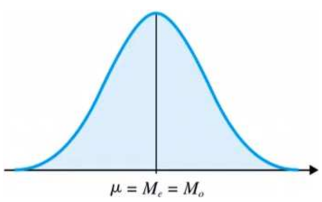
확률변수 X의 확률질량(밀도)함수가 $f_X(x)$일 때,

$$f_X(x_0)=\max\{f_X(x)\mid x\in R_X\}$$

를 만족하는 상수 x_0를 확률변수 X의 **최빈값**(mode)이라 하고, 이 값을 기호로

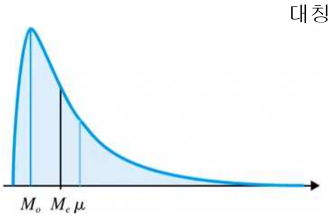
$$M_o=x_0$$

로 나타낸다.



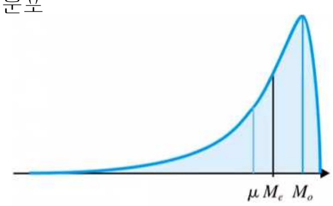
$M_o=M_e=\mu_X$

대칭분포



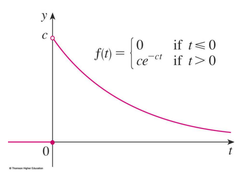
$M_o<M_e<\mu_X$

양의 비대칭분포

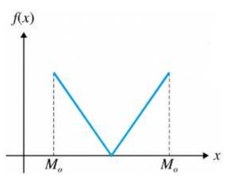


$\mu_X<M_e<M_o$

음의 비대칭분포



최빈값이 없는 경우



최빈값이 여러 개인 경우

Definition

(1) 확률변수 X의 누적분포함수 $F_X(x)$에 대하여,

$$F_X(Q_i)=\frac{i}{4},\quad i=1,2,3$$

를 만족하는 값 Q_1, Q_2, Q_3을 확률변수 X의 **사분위수**(quartiles)라 한다.

(2) 확률변수 X의 누적분포함수 $F_X(x)$에 대하여,

$$F_X(x_p)=p\quad(0<p<1)$$

를 만족하는 값 x_p를 확률변수 X의 $100p-$**백분위수**(percentiles) 라 한다.

Example 3

다음 확률밀도함수를 갖는 연속확률변수 X에 대하여,

$$f_X(x)=\begin{cases}2e^{-2x}, & x\ge 0\\ 0, & x<0\end{cases}$$

(1) X의 중앙값과 최빈값을 구하여라.

(2) 제 1사분위수 Q_1과 제 3사분위수 Q_3을 구하여라.

▣ 확률변수의 함수 $Y=u(X)$의 기댓값

Theorem

(1) 이산확률변수 X의 확률질량함수가 $f_X(x)$라 하자.

이산확률변수 $Y=u(X)$의 기댓값은

$$E[Y]=\sum_{x\in R_X}u(x)f_X(x)$$

이다.

(2) 연속확률변수 X의 확률밀도함수가 $f_X(x)$라 하자.

확률변수 $Y=u(X)$의 기댓값은

$$E[Y]=\int_{-\infty}^{\infty}u(x)f_X(x)\,dx$$

이다.

(proof)

Corollary

임의의 상수 a, b와 두 함수 $u(X)$, $v(X)$에 대하여, 다음이 성립한다.

(1) $E[aX+b]=aE[X]+b$

(2) $E[au(X)+bv(X)]=aE[u(X)]+bE[v(X)]$

Example

확률변수 X의 확률질량함수가 다음과 같다고 하자.

$$f_X(x)=\begin{cases} \frac{1}{4} & ,\ x=-1,1 \\ \frac{1}{2} & ,\ x=0 \\ 0 & ,\ \text{others} \end{cases}$$

(1) $Y=X^2$일 때, 확률변수 Y의 치역공간 R_Y를 구하여라.

(2) 확률변수 Y의 확률질량함수 $f_Y(y)$를 구하여라.

(3) 확률변수 Y의 기댓값을 구하여라.

Example 4

동전을 세 번 반복하여 던지는 실험에서 앞면이 나온 횟수를 X라 할 때, X의 기댓값 $E[X]$와 X^2의 기댓값 $E[X^2]$을 구하여라.

Example 5

연속확률변수 X의 확률밀도함수가

$$f_X(x)=\begin{cases} \frac{x}{8} & ,\ 0\le x\le 4 \\ 0 & ,\ x<0 \text{ 또는 } x>4 \end{cases}$$

일 때, X의 기댓값 $E[X]$와 X^2의 기댓값 $E[X^2]$을 구하여라.

Definition

(1) 이산확률변수 X의 확률질량함수가 $f_X(x)$이고 기댓값이 μ이면, X의 **분산**(variance)은

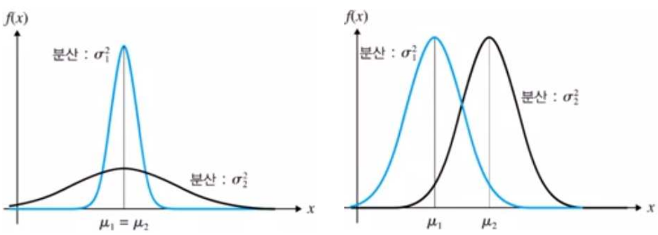
$$\sigma_X{}^2=Var[X]=\sum_{x\in R_X}(x-\mu_X)^2 f_X(x)$$

로 정의한다.

(2) 연속확률변수 X의 확률밀도함수가 $f_X(x)$이고 기댓값이 μ이면, X의 **분산**(variance)은

$$\sigma_X{}^2=Var[X]=\int_{-\infty}^{\infty}(x-\mu_X)^2 f_X(x)\,dx$$

로 정의한다.



분산에 따른 확률분포 비교

Definition

확률변수 X의 분산 $Var[X]$의 양의 제곱근

$$\sigma_X=\sqrt{Var[X]}$$

를 X의 **표준편차**(standard deviation)라 한다.

Example

[예제 4]에서 주어진 이산확률변수 X의 분산과 표준편차를 구하여라.

Example

[예제 5]에서 주어진 연속확률변수 X의 분산과 표준편차를 구하여라.

Theorem

$$\sigma_X{}^2 = Var[X] = E[X^2] - (E[X])^2$$

참고 $E[X^2] \neq (E[X])^2$

$(proof)$

Example 6

[예제 4]에서 주어진 이산확률변수 X의 분산과 표준편차를 구하여라.

Example 7

[예제 5]에서 주어진 연속확률변수 X의 분산과 표준편차를 구하여라.

Theorem

임의의 상수 a, b에 대하여 다음이 성립한다.

$$Var[aX+b] = a^2\, Var[X]$$

$$\sigma[aX+b] = |a| \cdot \sigma[X]$$

$(proof)$

Example 8

확률변수 X의 확률밀도함수가 $f_X(x) = e^{-x}$, $x>0$일 때, $Y=3X+1$의 평균과 분산을 구하여라.

Definition

확률변수 X에 대하여, $Z = \dfrac{X-\mu_X}{\sigma_X}$로 정의되는 확률변수 Z를 확률변수 X의 **표준화확률변수**(standardized random variable)라 한다. 이때,

$$\mu_Z = 0, \quad \sigma_Z{}^2 = 1$$

이 성립한다.

Theorem Markov Inequality

확률변수 X가 음이 아닌 값을 취할 때, $a>0$에 대하여

$$P(X \geq a) \leq \frac{E[X]}{a}$$

이 성립한다.

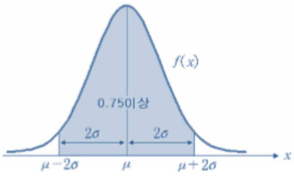
$(proof)$

Theorem Chebyshev's Inequality

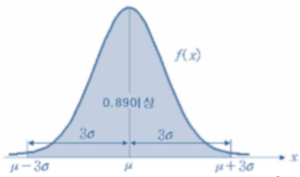
확률변수 X의 평균이 μ, 분산이 σ^2이면, $k>1$에 대하여

$$P(|X-\mu| \leq k\sigma) \geq 1 - \frac{1}{k^2}$$

이 성립한다.



$k=2$: $P(|X-\mu| \leq 2\sigma) \geq \frac{3}{4}$



$k=3$: $P(|X-\mu| \leq 3\sigma) \geq \frac{8}{9}$

$(proof)$

Example 9

확률변수 X의 평균이 $\mu_X = 9$이고 분산이 $\sigma_X{}^2 = 4$일 때, $P(6 \leq X \leq 12)$의 하한값을 구하여라.

▣ 적률과 적률생성함수

Definition

음이 아닌 정수 n에 대하여,

(1) 확률변수 X의 n차 **적률**(moment)은

$$E[X^n]=\begin{cases} \sum_{x \in R_X} x^n f_X(x) & , X : \text{이산} \\ \int_{-\infty}^{\infty} x^n f_X(x)dx & , X : \text{연속} \end{cases}$$

이다.

(2) 확률변수 X의 n차 **중심적률**(central moment)은

$$E[(X-\mu_X)^n]=\begin{cases} \sum_{x \in R_X} (x-\mu_X)^n f_X(x) & , X : \text{이산} \\ \int_{-\infty}^{\infty} (x-\mu_X)^n f_X(x)dx & , X : \text{연속} \end{cases}$$

이다.

Definition

확률변수 X의 **적률생성함수**(moment generating function)은

$$m_X(t)=E[e^{tX}]=\begin{cases} \sum_{x \in R_X} e^{tx} f_X(x) & , X : \text{이산} \\ \int_{-\infty}^{\infty} e^{tx} f_X(x)dx & , X : \text{연속} \end{cases}$$

이다.

참고 $e^{tX}=\sum_{n=0}^{\infty}\frac{1}{n!}(tX)^n$

$$=1+tX+\frac{t^2}{2!}X^2+\frac{t^3}{3!}X^3+\cdots+\frac{t^n}{n!}X^n+\cdots$$

$$E[e^{tX}]=E[1+tX+\frac{t^2}{2!}X^2+\frac{t^3}{3!}X^3+\cdots+\frac{t^n}{n!}X^n+\cdots]$$

$$=E[1]+E[tX]+E[\frac{t^2}{2!}X^2]+\cdots+E[\frac{t^n}{n!}X^n]+\cdots$$

$$=1+tE[X]+\frac{t^2}{2!}E[X^2]+\cdots+\frac{t^n}{n!}E[X^n]+\cdots$$

$$=\sum_{n=0}^{\infty}\frac{t^n}{n!}E[X^n]$$

Theorem

적률생성함수 $m_X(t)$의 n계 도함수

$$m_X{}^{(n)}(0)=\frac{d^n}{dt^n}m_X(t)\Big|_{t=0}=E[X^n]$$

(proof)

Theorem

(1) $E[X]=m_X{}'(0)$

(2) $Var[X]=E[X^2]-(E[X])^2=m_X{}''(0)-(m_X{}'(0))^2$

Example

동전을 두 번 던지는 게임에서 앞면이 나온 횟수를 X라 할 때,

(1) X의 적률생성함수 $m_X(t)$를 구하여라.

(2) $m_X(t)$를 이용하여, X의 기댓값과 분산을 구하여라.

Example

연속확률변수 X의 확률밀도함수가

$$f_X(x)=\begin{cases} e^{-x} & , x \geq 0 \\ 0 & , x<0 \end{cases}$$

일 때, X의 적률생성함수 $m_X(t)$를 구하고, $m_X(t)$를 이용하여, X의 기댓값과 분산을 구하여라.