1장. 확률분포

1.1 확률의 뜻과 성질

표본공간(標本空間 sample space), 사건(事件 event)(부록 I 참조)

확률(確率 probability)의 공리(p.11):
(1)(확률의 범위) 각 사건 A 에 대하여 $P(A) \geq 0$
(2)(전체의 확률) 표본공간 S에 대하여 P(S)=1
(3)(가산가법성(countable additivity)) 사건 $A_1, A_2, \cdots$에 대하여 $A_i \cap A_j = \varnothing \ (i \neq j)$이면

$$P(A_1 \cup A_2 \cup \cdots) = P(A_1) + P(A_2) + \cdots$$

합집합, 교집합과 여집합(부록 I)(p.487)

$$\bigcup_{n=1}^{\infty} A_n = A_1 \cup A_2 \cup \cdots = \{x : x \in A_n \text{ for } some\ n = 1, 2, \cdots\}$$

$$\bigcap_{n=1}^{\infty} A_n = A_1 \cap A_2 \cap \cdots = \{x : x \in A_n \text{ for } any\ n = 1, 2 \cdots\}$$

$$A^c = \{x : x \notin A\} = \{x : \text{A의 원소가 아닌 } x\}$$

드 모르간의 법칙

$$\left(\bigcup_{n=1}^{\infty} A_n\right)^c = \bigcap_{n=1}^{\infty} A_n^c, \quad \left(\bigcap_{n=1}^{\infty} A_n\right)^c = \bigcup_{n=1}^{\infty} A_n^c$$

(예)

$$\bigcup_{n=1}^{\infty} [\frac{1}{n}, 1] = (0, 1], \ \bigcap_{n=1}^{\infty} (-\frac{1}{n}, 1) = [0, 1), \ (\bigcup_{n=1}^{\infty} (-\infty, -\frac{1}{n}])^c = ((-\infty, 0))^c = [0, +\infty) = \bigcap_{n=1}^{\infty} (-\frac{1}{n}, +\infty)$$

정리 1.1.1: 확률의 기본 성질(p.12)
(a) 각 사건 A 에 대하여 $0 \leq P(A) \leq 1$ 이고 $P(\varnothing)=0$
(b)(여사건의 확률) $P(A^c)=1-P(A)$
(c)(단조성) $A \subseteq B$ 이면 $P(A) \leq P(B)$
[Key] 확률의 세 공리와 다음의 관계식

$$A \cap A^c = \varnothing, \ A \cup A^c = S, \ \varnothing = S^c$$

$$A \subseteq B \text{ 이면 } A \cap (B-A) = \varnothing, \ A \cup (B-A) = B$$

예 1.1.1: 베르누이 시행(Bernoulli trial)
(a) S={H,T}; P({H})=1/2, P({T})=1/2
(b) S={s, f}; P({s})=p, P({f})=1−p $(0 \leq p \leq 1)$

정리 1.1.2: 합사건의 확률(p.12)

(a) $P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2)$

(b) $P(A_1 \cup A_2 \cup A_3) = P(A_1) + P(A_2) + P(A_3) -$

$$-P(A_1 \cap A_2) - P(A_2 \cap A_3) - P(A_3 \cap A_1) + P(A_1 \cap A_2 \cap A_3)$$

(c) $P(A_1 \cup A_2 \cup \cdots \cup A_n) = P(A_1) + P(A_2) + \cdots + P(A_n) -$

$$-P(A_1 \cap A_2) - P(A_1 \cap A_3) - \cdots - P(A_{n-1} \cap A_n) +$$

$$+ \cdots\cdots$$

$$+ (-1)^{n-1} P(A_1 \cap A_2 \cap \cdots \cap A_n)$$

(d)(**가산반가법성 countable subadditivity**)

$$P(A_1 \cup A_2 \cup \cdots) \le P(A_1) + P(A_2) + \cdots$$

[Key]

(d) $B_1 = A_1,\ B_2 = A_2 - A_1,\ \cdots\cdots,\ B_n = A_n - (A_1 \cup A_2 \cup \cdots \cup A_{n-1}), \cdots$

$$B_i \cap B_j = \varnothing\ (i \ne j),\ B_1 \cup B_2 \cup \cdots = A_1 \cup A_2 \cup \cdots$$

(가산가법성) $P(A_1 \cup A_2 \cup \cdots) = P(B_1 \cup B_2 \cup \cdots) = P(B_1) + P(B_2) + \cdots$

(단조성) $B_1 \subseteq A_1,\ B_2 \subseteq A_2,\ \cdots,\ B_n \subseteq A_n,\ \cdots$

정리 1.1.3: 확률측도(確率測度 probability measure)**의 연속성(p.14)**

(a) $A_1 \subseteq A_2 \subseteq \cdots \subseteq A_n \subseteq \cdots$ 이면 $P(\bigcup_{n=1}^{\infty} A_n) = \lim_{n\to\infty} P(A_n)$

(b) $B_1 \supseteq B_2 \supseteq \cdots \supseteq B_n \supseteq \cdots$ 이면 $P(\bigcap_{n=1}^{\infty} B_n) = \lim_{n\to\infty} P(B_n)$

[Key] (a) $C_1 = A_1,\ C_2 = A_2 - A_1,\ \cdots\cdots,\ C_n = A_n - A_{n-1}, \cdots$

$$C_i \cap C_j = \varnothing\ (i \ne j),\ C_1 \cup C_2 \cup \cdots = A_1 \cup A_2 \cup \cdots$$

$$P(C_k) = P(A_k) - P(A_{k-1})\ \ (k = 2, 3, \cdots)$$

(가산가법성) $P(\bigcup_{n=1}^{\infty} A_n) = P(\bigcup_{n=1}^{\infty} C_n) = \sum_{n=1}^{\infty} P(C_n) = \lim_{n\to\infty} \sum_{k=1}^{n} P(C_k) = \lim_{n\to\infty} P(A_n)$

(b) $A_n = (B_n)^c;\ A_1 \subseteq A_2 \subseteq \cdots \subseteq A_n \subseteq \cdots,\quad \bigcup_{n=1}^{\infty} A_n = (\bigcap_{n=1}^{\infty} B_n)^c$

사건의 극한과 연속성(p.15)

(커지는 사건들) $A_1 \subseteq A_2 \subseteq \cdots \subseteq A_n \subseteq \cdots \subseteq \bigcup_{n=1}^{\infty} A_n \equiv \lim_{n\to\infty} A_n$

$$P(\lim_{n\to\infty} A_n) = \lim_{n\to\infty} P(A_n)$$

(작아지는 사건들) $B_1 \supseteq B_2 \supseteq \cdots \supseteq B_n \supseteq \cdots \bigcap_{n=1}^{\infty} B_n \equiv \lim_{n\to\infty} B_n$

$$P(\lim_{n\to\infty} B_n) = \lim_{n\to\infty} P(B_n)$$

예 1.1.3 $S = [0,\ 1],\ P((a,\ b]) = b - a\ (a < b)$:

$$P(\{b\}) = P(\bigcap_{n=1}^{\infty} (b - \frac{1}{n},\ b]) = \lim_{n\to\infty} P((b - \frac{1}{n},\ b]) = \lim_{n\to\infty} \frac{1}{n} = 0$$

1.2 조건부확률과 독립성

사건이 일어날 가능성은 그 사건에 관한 정보에 따라 다르다:
주사위를 던질 때 눈의 수가 3일 가능성은 1/6,
그 눈의 수가 홀수인 것을 알고 있다면 1, 3, 5 중의 하나로서 3일 가능성은 1/3

조건부확률(條件附確率 conditional probability):(p.15)
사건 A가 주어진 경우에 사건 B의 조건부확률

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

정리 1.2.1: 조건부확률의 성질(p.16)
(a)(곱셈공식) P(A)>0, P(B)>0 이면

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

(b)(전확률공식) $A_i\cap A_j=\varnothing\,(i\neq j),\, A_1\cup A_2\cup\cdots=S$ 일 때, $P(A_i)>0$ 이면

$$P(B)=P(B|A_1)P(A_1)+P(B|A_2)P(A_2)+\cdots$$

[Key] (b) 다음의 관계식에 가산가법성과 곱셈공식 적용:

$$B=B\cap S=B\cap(A_1\cup A_2\cup\cdots)=(B\cap A_1)\cup(B\cap A_2)\cup\cdots$$

$$(B\cap A_i)\cap(B\cap A_j)=\varnothing\ (i\neq j)$$

예 1.2.1: Laplace의 해 뜰 확률(생략)

정리 1.2.2: 베이즈(Bayes) 정리(p.17)
사건 $A_{1,}A_{2,}\cdots$이 표본공간 S를 공통부분이 없게 분할하고 $P(A_i)>0$ 일 때, $P(B)>0$이면

$$P(A_j|B)\propto P(B|A_j)P(A_j)\ (j=1,2,\cdots)$$

이고 <u>이 비례식에서 비례상수는 좌변의 합이 1임으로부터 결정된다.</u>
[Key]
곱셈공식; $P(A_j|B)P(B)=P(B\cap A_j)=P(B|A_j)P(A_j)\ (j=1,2,\cdots)$.
비례상수 c의 결정;

$$P(A_1|B)+P(A_2|B)+\cdots=c\{P(B|A_1)P(A_1)+P(B|A_2)P(A_2)+\cdots\}$$

$$P(A_1\cup A_2\cup\cdots|B)=c\{P(B|A_1)P(A_1)+P(B|A_2)P(A_2)+\cdots\}$$

$$1=P(S|B)=c\{P(B|A_1)P(A_1)+P(B|A_2)P(A_2)+\cdots\}$$

$$c=1/\{P(B|A_1)P(A_1)+P(B|A_2)P(A_2)+\cdots\}$$

예 1.2.2(p.18)

한 공장에서 전체 생산량의 20%, 30%, 50%를 세 기계 M_1, M_2, M_3로 생산하고 있고, 세 기계에서의 불량품 제조 비율은 각각 3%, 2%, 1%로 알려져 있다. 어느 날 이 공장에서 생산된 제품 중 임의로 1개를 택하여 검사하였더니 불량품이었다. 이 제품이 각 기계에서 생산되었을 확률을 구하여라.

[풀이] 검사 전 한 제품이 기계 M_1, M_2, M_3에서 생산되었을 확률:

$$P(M_1) = 0.2, P(M_2) = 0.3, P(M_3) = 0.5$$

한 제품이 불량품일 사건을 B라고 하면 각 기계에서의 불량품 제조 비율:

$$P(B|M_1) = 0.03, P(B|M_2) = 0.02, P(B|M_3) = 0.01$$

베이즈 정리:

$$P(M_1|B) \propto 0.2 \times 0.03, P(M_2|B) \propto 0.3 \times 0.02, P(M_3|B) \propto 0.5 \times 0.01$$

$$P(M_1|B) : P(M_2|B) : P(M_3|B) = 6 : 6 : 5$$

검사 후 불량품으로 밝혀진 제품이 각 기계에서 생산되었을 확률은

$$P(M_1|B) = 6/17, P(M_2|B) = 6/17, P(M_3|B) = 5/17$$

사건의 독립성(p.19~20):

(a)(두 사건이 서로 독립)

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

(b)(세 사건이 서로 독립)

$$P(A \cap B) = P(A)P(B),\ P(B \cap C) = P(B)P(C),\ P(C \cap A) = P(C)P(A)$$

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

(c)(n개 사건이 서로 독립)

$$P(A_i \cap A_j) = P(A_i)P(A_j)\ (1 \le i < j \le n)$$

$$P(A_i \cap A_j \cap A_k) = P(A_i)P(A_j)P(A_k)\ (1 \le i < j < k \le n)$$

$$\cdots\cdots\cdots$$

$$P(A_1 \cap A_2 \cap \cdots \cap A_n) = P(A_1)P(A_2)\cdots P(A_n)$$

예 1.2.3(p.19)

두 개의 주사위를 던지는 경우에 다음 사건들을 생각해보자

A={첫 번째 주사위의 눈이 짝수}

B={두 번째 주사위의 눈이 홀수}

C={두 주사위의 눈의 합이 홀수}

(a) A와 B, B와 C, C와 A는 서로 독립. (b) 사건 C와 $A \cap B$는 서로 독립이 아님.

[Key] 각각의 경우의 수를 구하여 전체 경우의 수로 나누면

$$P(A) = 18/36, P(B) = 18/36, P(C) = 18/36$$

$$P(A \cap B) = 9/36, P(B \cap C) = 9/36, P(C \cap A) = 9/36$$

$$\therefore P(A) = P(B) = P(C) = 1/2,\ P(A \cap B) = P(B \cap C) = P(C \cap A) = 1/4$$

(a) $P(A \cap B) = P(A)P(B)$, 같은 방법

(b) $A \cap B \subset C$, 즉 $A \cap B \cap C = A \cap B$

$$P(C|A \cap B) = P(A \cap B \cap C)/P(A \cap B) = 1, \quad \therefore P(C|A \cap B) \neq P(C)$$

1.3 확률변수와 확률분포

랜덤한 실험(random experiment) 확률변수(確率變數 random variable)(p.21~22)

$$\{X=1\}, \{X \le 1\}, \{1 < Y \le 3\}, \{Z > 0\}$$

예 1.3.1,예 1.3.2: 동전을 두 번 던지는 실험에서 앞면이 나오는 횟수(p.22)

표본공간 S={(T,T), (T,H), (H,T), (H,H)}, 앞면이 나오는 횟수 X

결과	(T,T)	(T,H)	(H,T)	(H,H)
X의 값	0	1	1	2
확률	1/4	1/4	1/4	1/4

확률변수 X의 확률분포표

X의 값	0	1	2
확률	1/4	1/2	1/4

예 1.3.3: 앞면이 나올 때까지 동전을 던지는 실험

앞면이 나올 때까지 동전을 던지는 실험에서 던지게 될 횟수를 X

S={ H, TH, TTH, TTTH,…}

P(X=1)=P(H)=1/2, P(X=2)=P(TH)=1/4, P(X=3)=P(TTH)=1/8, …

$$P(X=x) = (1/2)^x, \; x = 1, 2, \cdots$$

이산형(離散型 discrete type) 확률변수가 가질 수 있는 값들의 집합:

$$\{x_1, x_2, x_3, \cdots\}$$

확률밀도함수(確率密度函數 probability density function):

$$f(x_k) = P(X = x_k) \; (k = 1,2,3,\cdots)$$

이산형 확률변수의 확률밀도함수:(p.23)

(a)[1] $f(x) \ge 0 \;\; \forall x : -\infty < x < +\infty,$

$f(x) = 0 \;\; \forall x : x \ne x_k \; (k = 1,2,\cdots)$

(b)[2] $\sum_x f(x) = \sum_{k=1}^{\infty} f(x_k) = 1$

(c) $\sum_{x:a \le x \le b} f(x) = P(a \le X \le b)$

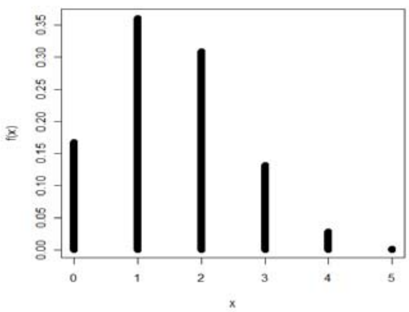


그림 1.3.1 이산형의 확률밀도함수

---

1) '$\forall x$' 는 '모든 x에 대하여'를 뜻하며, $\forall$ 는 any, all, arbitrary 에서의 A를 뒤집어 놓은 기호이다.

2) 앞으로 $\sum_x f(x)$ 의 의미는 $f(x) \ne 0$ 인 x에 대하여 $f(x)$ 를 더한다는 것을 뜻한다.

예 1.3.4 (p.24)

어느 병원에서 가까운 지하철역까지 **15분 간격으로 왕복 셔틀 버스가 운행**되고 있다. 아무런 정보가 없는 사람이 지하철을 타고 와서 셔틀 버스로 병원에 올 때 **버스를 기다릴 시간을 X(분)**이라고 하면 X 는 실수 구간의 값들을 가질 수 있고 그 확률은 다음과 같이 길이의 비로 주어진다.

$$\mathrm{P}(0 \le \mathrm{X} \le 15) = 1,\ \mathrm{P}(0 \le \mathrm{X} \le 5) = 5/15,\ \mathrm{P}(a \le \mathrm{X} \le b) = (b-a)/15\ (0 \le a < b \le 15)$$

연속형(連續型 continuous type) 확률변수:

실수 구간의 값들을 가질 수 있고 그에 관한 확률이 적분으로 주어진다.

연속형 확률변수의 확률밀도함수:(p.24)

(a) $f(x) \ge 0\ \forall x : -\infty < x < +\infty$

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

(c) $\int_a^b f(x)\,dx = \mathrm{P}(a \le \mathrm{X} \le b) \quad (-\infty < a < b < +\infty)$

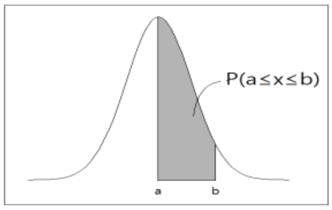


그림 1.3.2 연속형의 확률밀도 함수 **X의**

확률분포(確率分布 probability distribution) 또는 분포:

$$\mathrm{P}(a \le \mathrm{X} \le b) \quad (-\infty < a < b < +\infty), \quad \mathrm{X} \sim f \quad (\mathrm{pdf})$$

예 1.3.5

다음 함수가 확률밀도함수가 되기 위한 상수 c 의 값을 구하고, 이 함수가 확률변수 X 의 확률밀도 함수일 때 확률 $\mathrm{P}(1/2 \le \mathrm{X} \le 3/4)$을 구하여라.

$$f(x) = \begin{cases} cx(1-x) & 0 \le x \le 1 \\ 0 & x < 0, x > 1 \end{cases}$$

[Key] 전체 확률이 1 :

$$\int_{-\infty}^{+\infty} f(x)\,dx = \int_0^1 cx(1-x)\,dx = 1$$

$$\therefore\quad c[x^2/2 - x^3/3]_0^1 = c/6 = 1$$

$$\mathrm{P}(1/2 \le \mathrm{X} \le 3/4) = \int_{1/2}^{3/4} 6x(1-x)\,dx = 11/32$$

지표함수(指標函數 indicator function):(p.25)

$$\mathrm{I}_{\mathrm{A}}(x) = \begin{cases} 1 & x \in \mathrm{A} \\ 0 & x \notin \mathrm{A} \end{cases}, \qquad \mathrm{I}_{(x \in \mathrm{A})}$$

$$f(x) = 6x(1-x)\mathrm{I}_{[0\ 1]}(x) \quad \text{또는} \quad f(x) = 6x(1-x)\mathrm{I}_{(0 \le x \le 1)}$$

연속형 확률변수의 경우에

$$\mathrm{P}(\mathrm{X} = a) = \int_a^a f(x)\,dx = 0$$

$$\triangle x \fallingdotseq 0 \text{ 이면} \quad \mathrm{P}(a \le \mathrm{X} \le a + \triangle x) \fallingdotseq f(a)\triangle x$$

4. 정적분과 이상적분(p.497)

(1) (리이만)적분: 닫힌(closed) 구간 $[a,b]$에서 연속인 함수 f에 대하여

$$\int_{[a,b]} f(x)dx = \lim_{n\to\infty} \frac{b-a}{n} \sum_{i=1}^{n} f(a+\frac{b-a}{n}i)$$

(2) 음이 아닌 함수의 이상(improper)적분[3]

(a:열린 구간에서의 적분)

구간 $(a,b]$에서 연속이고 음이 아닌 함수 f에 대하여,

$$\int_{(a,b]} f(x)dx = \lim_{\substack{h\to 0 \\ h>0}} \int_{[a+h,b]} f(x)dx$$

(b:무한 구간에서의 적분)

구간 $(-\infty,b]$에서 연속이고 음이 아닌 함수 f에 대하여,

$$\int_{(-\infty,b]} f(x)dx = \lim_{a\to-\infty} \int_{[a,b]} f(x)dx$$

(c:수직선 전체에서의 적분)

수직선 전체인 $(-\infty,+\infty)$에서 연속이고 음이 아닌 함수 f에 대하여,

$$\int_{(-\infty,+\infty)} f(x)dx = \lim_{a\to-\infty} \int_{[a,0]} f(x)dx + \lim_{b\to+\infty} \int_{[0,b]} f(x)dx$$

(3) 적분가능(integrable) 함수(p.498)

구간 I (열린구간, 무한 구간, 수직선 전체 등)에서 연속인 함수 f에 대하여,

$$f^+(x) = \max(f(x),0), \quad f^-(x) = -\min(f(x),0)$$

$$f = f^+ - f^-$$

$$\int_I f(x)dx = \int_I f^+(x)dx - \int_I f^-(x)dx$$

이 적분 값이 실수이기 위하여는

$$\int_I f^+(x)dx < +\infty, \quad \int_I f^-(x)dx < +\infty$$

그런데 $|f| = f^+ + f^-$ 이므로, 이 조건은 아래 조건과 동치이다.

$$\int_I |f(x)|dx < +\infty$$

이 조건을 만족시키는 함수를 구간 I에서 **적분가능(integrable) 함수**라고 한다.

예 I.4.1

(a) $\int_1^\infty x^{-n}dx < +\infty \Leftrightarrow n > 1$ (b) $\int_0^1 x^{\alpha-1}dx < +\infty \Leftrightarrow \alpha > 0$

(c) $\int_{-\infty}^{+\infty} (\frac{x}{1+x^2})^+ dx = \int_0^{+\infty} \frac{x}{1+x^2}dx = +\infty$, $\int_{-\infty}^{+\infty} (\frac{x}{1+x^2})^- dx = \int_{-\infty}^{0} \frac{-x}{1+x^2}dx = +\infty$

$$\int_{-\infty}^{+\infty} (\frac{x}{1+x^2})^+ dx - \int_{-\infty}^{+\infty} (\frac{x}{1+x^2})^- dx = \infty - \infty \,(???)(\text{부정형})$$

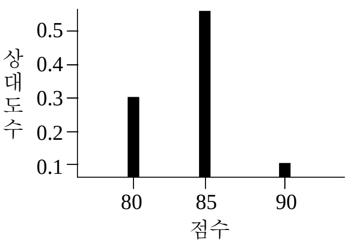
---

3) 음이 아닌 함수를 전제로 하므로 이 정의에서의 극한값들은 실수이거나 $+\infty$이다. 따라서, (c)의 이상 적분의정의에서 $\infty-\infty$꼴의 부정형은 나오지 않는다.

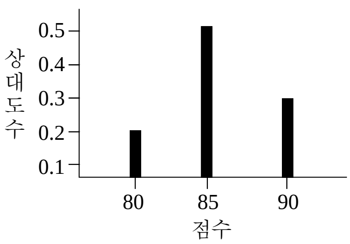
1.4 확률분포의 특성치

'평균'의 의미는 무엇이고, 무슨 목적으로 '평균'을 사용할까?(p.26)

1반



2반



1 반

점수	80	85	90	합계
도수	3	6	1	10
상대도수	3/10	6/10	1/10	1

2반

점수	80	85	90	합계
도수	2	5	3	10
상대도수	2/10	5/10	3/10	1

$(80\times 3+85\times 6+90\times 1)/10=84.0$ $\qquad$ $(80\times 2+85\times 5+90\times 3)/10=85.5$

그림 1.4.1 두 반의 성적 분포와 평균성적

평균은 분포의 위치를 나타내는 값

$$80\times(\frac{3}{10})+85\times(\frac{6}{10})+90\times(\frac{1}{10})=84.0,\qquad 80\times(\frac{2}{10})+85\times(\frac{5}{10})+90\times(\frac{3}{10})=85.5$$

확률분포의 평균:(p.27)

확률변수 X 의 분포의 평균[4)](平均 mean)은

$$\mu=\begin{cases}\sum_x x\,pdf(x) & (\text{X가 이산형일 때})\\ \int_{-\infty}^{+\infty} x\,pdf(x)\,dx & (\text{X가 연속형일 때})\end{cases}$$

예1.4.1

확률변수 X 의 확률밀도함수가 다음과 같을 때

X 의 확률분포의 평균 μ 를 구하여라.

$$f(x)=\begin{cases}(-x^2+2x)/2, & 0\le x\le 1\\ (-x^2+2x+3)/8, & 1\le x\le 3\\ 0, & x<0, x>3\end{cases}$$

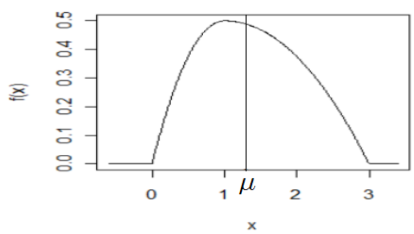


그림 1.4.2

[풀이] 평균의 정의에 따라

$$\mu=\int_{-\infty}^{+\infty} x f(x)\,dx=\int_0^1 x(-x^2+2x)/2\,dx+\int_1^3 x(-x^2+2x+3)/8\,dx$$

$$\therefore\ \mu=[-x^4/8+x^3/3]_0^1+[-x^4/32+x^3/12+3x^2/16]_1^3=11/8$$

4) 여기에서 무한 합이나 이상 적분이 절대 수렴하는 경우에만 평균이 실수로 정의되는 것에 유의하여야 한다.

기댓값(期待값 expected value)(p.28): $g(\mathrm{X})$ 의 기댓값[5] $\mathrm{E}[g(\mathrm{X})]$

$$\mathrm{E}[g(\mathrm{X})]=\begin{cases}\sum_x g(x)pdf(x) & (\mathrm{X}\text{가 이산형일 때})\\ \int_{-\infty}^{+\infty} g(x)pdf(x)\,dx & (\mathrm{X}\text{가 연속형일 때})\end{cases}$$

확률분포의 분산(分散 variance)과 표준편차(標準偏差 standard deviation)(p.29):
분포가 평균을 기준으로 어떻게 산포되어 있는가를 나타내는 특성치

$$\mathrm{Var}(\mathrm{X})=\mathrm{E}[(\mathrm{X}-\mu)^2]=\begin{cases}\sum_x (x-\mu)^2 pdf(x) & (\mathrm{X}\text{가 이산형일 때})\\ \int_{-\infty}^{+\infty} (x-\mu)^2 pdf(x)\,dx & (\mathrm{X}\text{가 연속형일 때})\end{cases}$$

$$\mathrm{Sd}(\mathrm{X})=\sqrt{\mathrm{Var}(\mathrm{X})}$$

예 1.4.3
다음 두 확률분포의 분산을 구하여라.

(a) $\mathrm{X}\sim f_1\,(\mathrm{pdf})$

$f_1(x)=\mathrm{I}_{[0\ \ 1]}(x)$

(b) $\mathrm{X}\sim f_2\,(\mathrm{pdf})$

$f_2(x)=(2-4|x-1/2|)\mathrm{I}_{[0\ \ 1]}(x)$

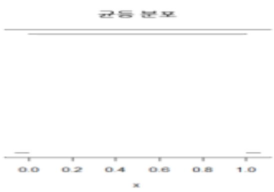


그림 1.4.3 균등분포

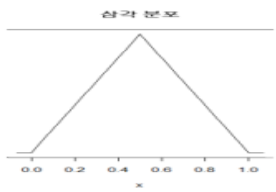


그림 1.4.4 삼각분포

[풀이] 두 확률분포의 분산을 각각 $\sigma_1^2,\ \sigma_2^2$ 이라고 하면

$$\sigma_1^2=\int_{-\infty}^{+\infty}(x-\mu_1)^2 f_1(x)\,dx=\int_0^1 (x-1/2)^2 dx=1/12,$$

$$\sigma_2^2=\int_0^1 (x-1/2)^2(2-4|x-1/2|)\,dx=\int_{-1/2}^{1/2} t^2\,(2-4|t|)dt=1/24$$

X 의 평균, 분산 : $\mu_{\mathrm{X}},\sigma_{\mathrm{X}}^2$ 으로 나타내거나, X를 생략하여도 혼동이 없을 때에는 μ,σ^2

정리 1.4.1: 기댓값의 성질[6](p.30)

(a)(선형성) $\mathrm{E}(a\mathrm{X}+b)=a\mathrm{E}(\mathrm{X})+b$ (a,b는 확률변수가 아닌 상수)

(b)(선형성) $\mathrm{E}[c_1g_1(\mathrm{X})+c_2g_2(\mathrm{X})]=c_1\mathrm{E}[g_1(\mathrm{X})]+c_2\mathrm{E}[g_2(\mathrm{X})]$ (c_1,c_2 는 상수)

(c)(단조성) $g_1(\mathrm{X})\le g_2(\mathrm{X})$ 이면 $\mathrm{E}[g_1(\mathrm{X})]\le \mathrm{E}[g_2(\mathrm{X})]$

[Key] 합, 적분의 성질, pdf의 합은 1

5) 여기에서 무한 합이나 이상적분이 실수로 정의되려면 합이나 적분이 절대 수렴하여야 한다.
6) 이러한 선형성과 단조성은 각각의 기댓값이 실수로 정의되는 것을 전제로 하는 것이다.

정리 1.4.2: 분산의 성질과 계산 공식[7](p.30)

(a) $\mathrm{Var}(aX+b)=a^2\mathrm{Var}(X)$ (a, b 는 상수)

(b) $\mathrm{Var}(X)=E(X^2)-\{E(X)\}^2$

[Key] 정의와 기댓값의 선형성:

$$E(aX+b)=aE(X)+b=a\mu+b$$

(a) $Y=aX+b$ 라고 하면

$$\mathrm{Var}(aX+b)=E[(aX+b-(a\mu+b))^2]=E[a^2(X-\mu)^2]=a^2E[(X-\mu)^2]$$

(b)

$$\mathrm{Var}(X)=E[(X-\mu)^2]=E[(X^2-2\mu X+\mu^2)]=E(X^2)-2\mu^2+\mu^2$$

예 1.4.4 (생략)

예 1.4.5: 확률변수의 표준화(標準化 standardize)

평균이 μ 이고 표준편차가 σ이고 $\sigma>0$ 일 때

$$Z=\frac{X-E(X)}{\sqrt{\mathrm{Var}(X)}}=\frac{X-\mu}{\sigma}$$

$$E(Z)=(E(X)-\mu)/\sigma=0, \qquad \mathrm{Var}(Z)=\mathrm{Var}(X)/\sigma^2=1$$

평균이나 분산이 실수로 정의될 수 없는 확률분포가 있음에 유의(p.31)

$$f(k)=P(X=k)=\frac{1}{k(k+1)},\ k=1,2,\cdots$$

7) 이러한 성질과 공식은 각 변수의 분산과 기댓값이 모두 실수라는 전제 하에 성립하는 것이다.

1.5 누적분포함수와 생성함수

누적분포함수(累積分布函數 cumulative distribution function):(p.32)

확률변수 X 의 누적 확률을 나타내는 함수

$$cdf_X(x) = P(X \le x)$$

(이산형 확률변수의 경우)

$$F(x) = P(X \le x) = \sum_{t:t \le x} f(t)$$

$x_1 < x_2 < \cdots < x_{n-1} < x_n < \cdots$ 인 경우에는

$$F(x_n) = \sum_{k=1}^{n} f(x_k), \quad f(x_n) = F(x_n) - F(x_{n-1})$$

(연속형 확률변수의 경우)

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

f 가 연속인 x 에 대하여는

$$\frac{d}{dx}F(x) = f(x)$$

예 1.5.1: 앞면이 나올 때까지 동전을 던지는 실험

$$F(n) = P(X \le n) = \sum_{k=1}^{n} pdf(k) = \sum_{k=1}^{n} (1/2)^k = 1 - (1/2)^n \ \ (n = 1, 2, \cdots)$$

$$F(x) = \sum_{t:t \le x} f(t) = \sum_{\substack{k=1 \\ k \le x}}^{\infty} f(k) = \begin{cases} 0, & x < 1 \\ 1 - (1/2)^n, & n \le x < n+1\,(n = 1, 2, \cdots) \end{cases}$$

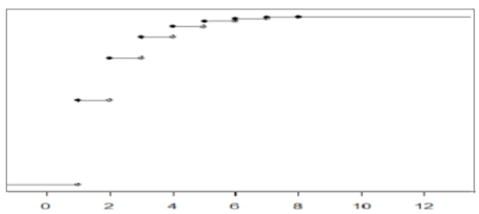
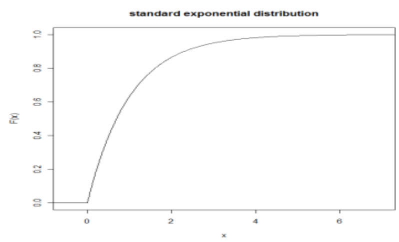


그림 1.5.1 이산형 확률변수의 누적분포함수의 예

예 1.5.2: 표준지수분포(標準指數分布 standard exponential distribution)(p.33)

$$pdf(x) = e^{-x} I_{(x \ge 0)}$$



$$cdf(x) = \int_{-\infty}^{x} e^{-t} I_{(t \ge 0)}\,dt = \begin{cases} 0, & x < 0 \\ \int_0^x e^{-t}\,dt = 1 - e^{-x}, & x \ge 0 \end{cases}$$

정리 1.5.1: 누적분포함수의 성질(p.33)

(a)(증가성) $x_1 < x_2$ 이면 $\mathrm{F}(x_1) \le \mathrm{F}(x_2)$

(b)(전체 변동) $\lim_{x \to -\infty} \mathrm{F}(x) = 0, \quad \lim_{x \to +\infty} \mathrm{F}(x) = 1$

(c)(오른쪽 연속성) $\lim_{\substack{h \to 0 \\ h > 0}} \mathrm{F}(x+h) = \mathrm{F}(x)$

[Key] (a) 확률의 단조성:

$$x_1 < x_2 \text{ 이면 } (\mathrm{X} \le x_1) \subseteq (\mathrm{X} \le x_2)$$

(b) 확률측도의 연속성과 (a)의 증가성:

$$\lim_{k \to +\infty} (\mathrm{X} \le -k) = \bigcap_{k=1}^{\infty} (\mathrm{X} \le -k) = \varnothing, \qquad \lim_{k \to +\infty} (\mathrm{X} \le k) = \bigcup_{k=1}^{\infty} (\mathrm{X} \le k) = \mathrm{S}$$

$$\lim_{k \to +\infty} \mathrm{P}(\mathrm{X} \le -k) = \mathrm{P}(\varnothing) = 0, \qquad \lim_{k \to +\infty} \mathrm{P}(\mathrm{X} \le k) = \mathrm{P}(\mathrm{S}) = 1$$

(c) 확률측도의 연속성과 (a)의 증가성:

$$\lim_{n \to +\infty} (\mathrm{X} \le x + 1/n) = \bigcap_{n=1}^{\infty} (\mathrm{X} \le x + 1/n) = (\mathrm{X} \le x)$$

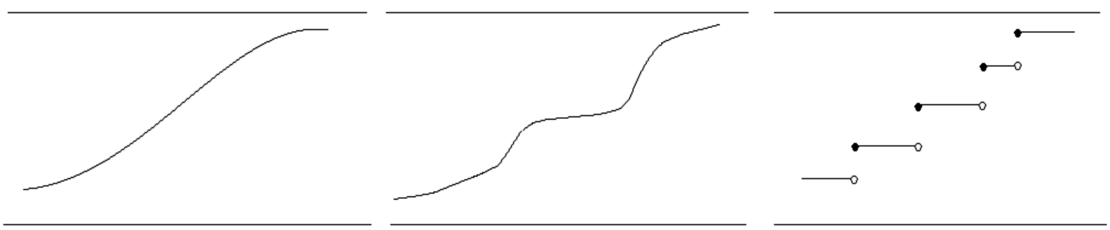


그림 1.5.3 누적분포함수의 여러 가지 형태

(오른쪽 연속성)

$$\mathrm{F}(a+) = \lim_{\substack{x \to a \\ x > a}} \mathrm{F}(x) = \lim_{n \to +\infty} \mathrm{F}(a + 1/n), \quad \mathrm{F}(a+) = \mathrm{F}(a) = \mathrm{P}(\mathrm{X} \le a)$$

(왼쪽 극한)

$$\mathrm{F}(a-) = \lim_{\substack{x \to a \\ x < a}} \mathrm{F}(x) = \lim_{n \to +\infty} \mathrm{F}(a - 1/n) = \lim_{n \to +\infty} \mathrm{P}(\mathrm{X} \le a - 1/n)$$

$$\lim_{n \to +\infty} (\mathrm{X} \le a - 1/n) = \bigcup_{n=1}^{\infty} (\mathrm{X} \le a - 1/n) = (\mathrm{X} < a)$$

$$\mathrm{F}(a-) = \lim_{n \to +\infty} \mathrm{P}(\mathrm{X} \le a - 1/n) = \mathrm{P}(\lim_{n \to +\infty} (\mathrm{X} \le a - 1/n)) = \mathrm{P}(\mathrm{X} < a)$$

($\because$ 확률측도의 연속성)

(왼쪽 방향에서의 불연속성의 크기)

$$\mathrm{F}(a) - \mathrm{F}(a-) = \mathrm{P}(\mathrm{X} \le a) - \mathrm{P}(\mathrm{X} < a) = \mathrm{P}(\mathrm{X} = a)$$

예 1.5.3(불연속인 누적분포함수의 예)(p.35)

어느 병원에서 가까운 지하철역까지 10분 간격으로 왕복 셔틀 버스가 운행되고 있고, 지하철역에서 병원까지 버스로는 10분, 도보로는 15분이 걸린다. 아무런 정보가 없는 사람이 지하철을 타고 와서 동전을 던져 앞면이 나오면 버스로 뒷면이 나오면 도보로 병원에 가려고 한다. 이 사람이 지하철역에서 병원에 도착할 때까지의 시간을 X(분)이라고 하면,

$$\mathrm{P}(\mathrm{X} \le x | \text{앞면}) = \begin{cases} 0, & x < 10 \\ (x-10)/10, & 10 \le x < 20 \\ 1, & x \ge 20 \end{cases} \qquad \mathrm{P}(\mathrm{X} \le x | \text{뒷면}) = \begin{cases} 0, & x < 15 \\ 1, & x \ge 15 \end{cases}$$

$$\mathrm{P}(\text{앞면}) = \mathrm{P}(\text{뒷면}) = 1/2$$

따라서

$$\mathrm{P}(\mathrm{X} \le x) = \mathrm{P}(\mathrm{X} \le x | \text{앞면})\mathrm{P}(\text{앞면}) + \mathrm{P}(\mathrm{X} \le x | \text{뒷면})\mathrm{P}(\text{뒷면})$$

으로부터 X 의 누적분포함수는

$$\mathrm{F}(x) = \begin{cases} 0, & x < 10 \\ (x-10)/20, & 10 \le x < 15 \\ x/20, & 15 \le x < 20 \\ 1, & x \ge 20 \end{cases}$$

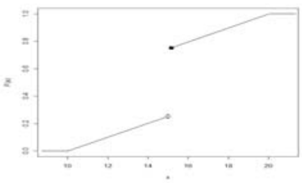


그림 1.5.4 불연속인 누적분포함수의 예

확률생성함수(確率生成函數 probability generating function):(p.36)

<u>음이 아닌 정수의 값</u>을 가질 수 있는 이산형 확률변수 X

확률밀도함수: $f(x) = \mathrm{P}(\mathrm{X} = x),\ x = 0, 1, \cdots$ 또는 $(f(0), f(1), f(2), \cdots)$

확률생성함수[8]: $f(0)s^0 + f(1)s^1 + f(2)s^2 + \cdots = \sum_{x=0}^{\infty} s^x f(x),\ -1 < s < 1$

$$\mathrm{G}(s) = \mathrm{E}(s^{\mathrm{X}}) = \sum_{x=0}^{\infty} s^x \mathrm{P}(\mathrm{X} = x),\ -1 < s < 1$$

확률 생성 관계: $f(k) = \mathrm{P}(\mathrm{X} = k) = \mathrm{G}^{(k)}(0)/k!,\ k = 0, 1, 2, \cdots$

$$f(k) = \mathrm{P}(\mathrm{X} = k) = \mathrm{G}^{(k)}(0)/k!,\ k = 0, 1, 2, \cdots$$

예 1.5.4: 앞면이 나올 때까지 동전을 던지는 실험

앞면이 나올 때까지 동전을 던지게 될 횟수 X 의 확률밀도함수는

$$f(k) = \mathrm{P}(\mathrm{X} = k) = (1/2)^k\ (k = 1, 2, \cdots), \quad f(x) = 0\ \ \forall x : x \ne k\ (k = 1, 2, \cdots)$$

이므로, X 의 확률생성함수는

$$\mathrm{G}(s) = \mathrm{E}(s^{\mathrm{X}}) = \sum_{k=1}^{\infty} s^k (1/2)^k = s/(2-s), \quad -2 < s < 2$$

확률생성함수 $\mathrm{G}(s)$ 의 정의에서 s 가 양수이면

$$\mathrm{G}(s) = \mathrm{E}(s^{\mathrm{X}}) = \mathrm{E}(e^{\mathrm{X}\log s}) = \mathrm{E}(e^{t\mathrm{X}}),\ t = \log s$$

와 같이 나타낼 수 있으므로 $\mathrm{E}(e^{t\mathrm{X}})$도 분포를 정할 수 있음을 알 수 있다.

---

8) 확률생성함수를 뜻하는 멱급수는 예 1.5.4에서와 같이 1보다 큰 s 값에 대해서도 절대 수렴할 수 있다.

적률생성함수[9](積率生成函數 moment generating function):(p.37)

0을 포함하는 열린구간의 t 값에 대하여 $\mathrm{E}(e^{t\mathrm{X}})$ 가 실수일 때, 함수

$$mgf_{\mathrm{X}}(t) = \mathrm{E}(e^{t\mathrm{X}}),\ \ -h < t < h\,(\exists\, h > 0)$$[10]

를 확률변수 X 의 적률생성함수라고 한다.

예 1.5.5: 앞면이 나올 때까지 동전을 던지는 실험

앞면이 나올 때까지 동전을 던지게 될 횟수 X 의 확률밀도함수는

$$f(k) = \mathrm{P}(\mathrm{X}=k) = (1/2)^k\ (k = 1, 2, \cdots),$$

$e^t/2 < 1$ 즉 $t < \log 2$ 일 때 $\mathrm{E}(e^{t\mathrm{X}}) = \sum_{k=1}^{\infty} e^{tk}(1/2)^k < +\infty$

$$\therefore mgf_{\mathrm{X}}(t) = \mathrm{E}(e^{t\mathrm{X}}) = \sum_{k=1}^{\infty} e^{tk}(1/2)^k = (e^t/2)/(1-e^t/2),\ \ t < \log 2$$

(지수함수의 멱급수 전개식)

$$e^{t\mathrm{X}} = 1 + \frac{t\mathrm{X}}{1!} + \frac{(t\mathrm{X})^2}{2!} + \cdots \frac{(t\mathrm{X})^k}{k!} + \cdots$$

(오른쪽 변의 기댓값을 항별로 취하여 더할 수 있다면 다음을 예상)

$$\mathrm{M}(t) = \mathrm{E}(e^{t\mathrm{X}}) = 1 + \frac{\mathrm{E}(\mathrm{X})}{1!}t + \frac{\mathrm{E}(\mathrm{X}^2)}{2!}t^2 + \cdots + \frac{\mathrm{E}(\mathrm{X}^k)}{k!}t^k + \cdots$$

적률[11]**: X 의 k차 적률**$(k = 1, 2, \cdots)$(p.38)

$\mathrm{E}(|\mathrm{X}|^k) < +\infty$ 일 때

$$m_k = \mathrm{E}(\mathrm{X}^k) = \begin{cases} \sum_x x^k f(x) & (\mathrm{X}\text{ 가 이산형일 때}) \\ \int_{-\infty}^{+\infty} x^k f(x)\,dx & (\mathrm{X}\text{ 가 연속형일 때}) \end{cases}$$

정리 1.5.2: 적률생성함수의 성질(p.38)

(a)(적률생성 성질) 확률변수 X 의 적률생성함수가 존재하면, 즉

$$\mathrm{M}(t) = \mathrm{E}(e^{t\mathrm{X}}) < +\infty\ \ \forall t : -h < t < h\,(\exists\, h > 0)$$

이면, X 의 모든 적률이 존재하고

$$\mathrm{E}(\mathrm{X}^k) = \mathrm{M}^{(k)}(0),\quad \mathrm{M}(t) = \sum_{k=0}^{\infty} \frac{\mathrm{E}(\mathrm{X}^{\mathrm{k}})}{k!} t^k,\ \ -\epsilon < t < \epsilon\,(\exists\, \epsilon > 0)$$

(b)(분포 결정성)[12] 두 확률변수 X, Y의 적률생성함수 $\mathrm{M_X}(t), \mathrm{M_Y}(t)$가 존재하고 0을 포함하는 열린구간에서 일치한다면, 즉

$$\mathrm{M_X}(t) = \mathrm{M_Y}(t)\ \ \forall t : -h < t < h\,(\exists\, h > 0)$$

이면, X 와 Y의 확률분포가 일치 즉 X 와 Y의 확률밀도함수, 누적분포함수가 일치한다.

[증명]

---

9) $\mathrm{E}(e^{t\mathrm{X}}) < +\infty\ \ \forall t : -h < t < h(h > 0)$ 가 성립하지 않으면 적률생성함수가 존재하지 않는다고 한다.

10) $(\exists\, h > 0)$ 의 표현에서 $\exists$ 는 exist 의 E를 뒤집어 표기한 것으로, '양수 h가 존재하여'라는 뜻이다.

11) 적률이 실수로 정의되기 위하여는 무한 합이나 이상적분이 절대 수렴, 즉 $\mathrm{E}(|\mathrm{X}|^k) < +\infty$ 이어야 한다.

12) 적률생성함수를 일반화한 특성함수(characteristic function)를 이용한 증명은 참고서적을 참조

(a)
$$e^{a} = 1 + \frac{a}{1!} + \frac{a^2}{2!} + \cdots \frac{a^k}{k!} + \cdots$$
$$(|tx|)^k / k! \le e^{|tx|} \le e^{tx} + e^{-tx}$$
$$\therefore |t|^k \mathrm{E}(|\mathrm{X}|^k)/k! \le \mathrm{M}(t) + \mathrm{M}(-t) < +\infty \quad (-h < t < h)$$
따라서 X 의 모든 적률이 존재한다.

(테일러 정리: 부록 정리 I.3.1) 지수함수에 적용하면
$$e^{tx} - \sum_{k=0}^{n-1} \frac{(tx)^k}{k!} = \frac{1}{(n-1)!} \int_0^1 (1-u)^{n-1} e^{utx}\, du\ (tx)^n$$
$$\therefore |e^{tx} - \sum_{k=0}^{n-1} \frac{(tx)^k}{k!}| \le \int_0^1 (1-u)^{n-1} e^{u|tx|}\, du\ \frac{(|tx|)^{n-1}}{(n-1)!} \frac{|tx|}{1!}$$
$$\le \int_0^1 (1-u)^{n-1} e^{|tx|}\, du\ e^{|tx|} e^{|tx|}$$
$$\le \frac{1}{n}(e^{3tx} + e^{-3tx})$$
이 부등식의 각 변에 확률밀도함수를 곱하여 합하거나 적분하면
$$|\mathrm{M}(t) - \sum_{k=0}^{n-1} \mathrm{E}(\mathrm{X}^k) t^k / k!| \le \frac{1}{n}\{\mathrm{M}(3t) + \mathrm{M}(-3t)\},\ -h < 3t < h$$
$$\therefore \mathrm{M}(t) = \sum_{k=0}^{\infty} \frac{\mathrm{E}(\mathrm{X}^k)}{k!} t^k,\quad -\epsilon < t < \epsilon\,(\epsilon = h/3 > 0)$$
(b) (생략) [P. Billingsley] Probability and Measure 3rd ed. p.346~p.347을 참조하시오.

로그정규분포(3장 문제 (3.19)): 모든 적률은 존재하지만 적률생성함수가 존재하지 않는 예

예 1.5.6: 표준지수분포의 적률생성함수와 적률(p.40)

확률밀도함수: $f(x) = e^{-x} \mathrm{I}_{(x>0)}$

적률생성함수: $\mathrm{M}(t) = \int_0^{+\infty} e^{tx} e^{-x} dx = (1-t)^{-1},\quad t < 1$

적률생성함수의 멱급수 전개식과 적률:
$$\mathrm{M}(t) = \sum_{k=0}^{\infty} t^k = \sum_{k=0}^{\infty} \frac{k!}{k!} t^k,\quad t < 1$$
$$\mathrm{E}(\mathrm{X}^k) = \mathrm{M}^{(k)}(0) = k! \quad (k = 1, 2, \cdots)$$
$$\int_0^{\infty} x^k e^{-x} dx = \int_0^{\infty} x^k d(-e^{-x}) = \int_0^{\infty} e^{-x} d(x^k) = k \int_0^{\infty} x^{k-1} e^{-x} dx = \cdots = k!$$

누율생성함수(累率生成函數 cumulant generating function)와 누율(累率 cumulant):(p.40)

적률생성함수 $\mathrm{M}(t)$가 존재할 때,

$$cgf(t) = \log\ mgf(t) = \log \mathrm{E}(e^{t\mathrm{X}}),\quad -h < t < h(\exists\, h > 0)$$

$$cgf_{\mathrm{X}}(t) = \sum_{r=1}^{\infty} \frac{c_r(\mathrm{X})}{r!} t^r,\quad -h < t < h(\exists\, h > 0)$$

$$c_r(\mathrm{X}) = \ cgf_{\mathrm{X}}^{(r)}(0)\ \ (r = 1, 2, \cdots)\ \ (r\text{차 누율})$$

누율 c_r과 적률 m_k의 관계:(p.40)

$$\mathrm{c}_1(\mathrm{X}) = \mathrm{E}(\mathrm{X}), c_2(\mathrm{X}) = \mathrm{Var}(\mathrm{X}), \cdots, (\text{연습문제 (1.19) 참조})$$

[Key]

$$\log(1 + \sum_{k=1}^{\infty} \frac{m_k}{k!} t^k) = \sum_{r=1}^{\infty} \frac{c_r}{r!} t^r,\quad -h < t < h(\exists\, h > 0)$$

$$\log(1 + \mathrm{A}) = \mathrm{A} - \mathrm{A}^2/2 + \mathrm{A}^3/3 - \cdots,\ \ (-1 < \mathrm{A} < 1)$$

$$\mathrm{A} = m_1 t + m_2 t^2/2! + m_3 t^3/3! + \cdots$$

결과를 t의 오름차순으로 정리하면

$$c_1 = m_1,\ c_2 = 2!\left\{m_2/2! - (m_1)^2/2\right\} = m_2 - (m_1)^2, \cdots$$

1.6 여러 가지 부등식

볼록(convex) 함수: 수직선 위의 구간 I에서 정의된 함수로서 다음을 만족시키는 함수 ϕ

$$\phi(\lambda x+(1-\lambda)y) \le \lambda\phi(x)+(1-\lambda)\phi(y) \quad \forall\,\lambda: 0 \le \lambda \le 1,\ \forall\, x\in \mathrm{I}, \forall\, y\in \mathrm{I}$$

두 번 미분가능한 함수의 경우에 볼록한 함수: $\phi''(x) \ge 0 \quad \forall\, x\in \mathrm{I}$

정리 1.6.1: 젠센(Jensen)의 부등식(p.41)

수직선 위의 구간 I 에서의 값을 갖는 확률변수 X 의 기댓값이 존재하면, 구간 I 에서 볼록한 함수 ϕ 에 대하여 다음 부등식이 성립한다.

$$\phi(\mathrm{E}(\mathrm{X})) \le \mathrm{E}(\phi(\mathrm{X}))$$

[증명] 두 번 미분 가능한 볼록함수 ϕ의 경우에

$$\phi''(x) \ge 0 \quad \forall\, x\in \mathrm{I}$$

$$\phi(x)=\phi(\mu)+\phi'(\mu)(x-\mu)+\phi''(\xi)(x-\mu)^2/2 \quad (\exists\,\xi\in \mathrm{I})$$

$$\phi(x) \ge \phi(\mu)+\phi'(\mu)(x-\mu) \quad \forall\, x\in \mathrm{I}$$

이 부등식의 양변에 확률밀도함수를 곱하여 합하거나 적분하면

$$\mathrm{E}(\phi(\mathrm{X})) \ge \phi(\mu)+\phi'(\mu)\mathrm{E}(\mathrm{X}-\mu)=\phi(\mu)$$

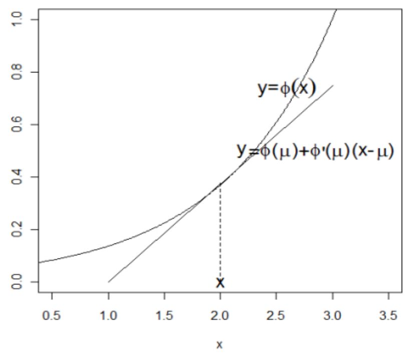


그림 1.6.1 볼록한 함수

예 1.6.1: $\phi(x)=x^2$:볼록함수

$$\mathrm{E}(\mathrm{X}^2)=\mathrm{E}(\phi(\mathrm{X})) \ge \phi(\mathrm{E}(\mathrm{X}))=(\mathrm{E}(\mathrm{X}))^2$$

정리 1.6.2: 리아푸노프(Liapounov) 부등식(p.42)

확률변수 X 에 대하여 $\mathrm{E}(|\mathrm{X}|^s)<\infty$이면 $0<r<s$인 r에 대하여

$$(\mathrm{E}(|\mathrm{X}|^r))^{1/r} \le (\mathrm{E}(|\mathrm{X}|^s))^{1/s}$$

[Key] $\phi(y)=y^p\mathrm{I}_{(y\ge 0)}(p=s/r>1)$:볼록함수, $\mathrm{Y}=|\mathrm{X}|^r$에 젠센의 부등식

$$(\mathrm{E}(|\mathrm{X}|^r))^{s/r}=(\mathrm{E}(\mathrm{Y}))^p \le \mathrm{E}(\mathrm{Y}^p)=\mathrm{E}((|\mathrm{X}|^r)^{s/r})=\mathrm{E}(|\mathrm{X}|^s)$$

(참고: $\mathrm{E}(\mathrm{Y})=\mathrm{E}(|\mathrm{X}|^r(\mathrm{I}_{(|\mathrm{X}|\le 1)}+\mathrm{I}_{(|\mathrm{X}|>1)})) \le \mathrm{E}(1+|\mathrm{X}|^s)<\infty$)

정리 1.6.3: 마코프(Markov) 부등식과 체비셰프(Chebyshev) 부등식(p.43)

(a)(마코프) 확률변수 Z 에 대하여 $\mathrm{E}(|\mathrm{Z}|^r)<\infty$ $(r>0)$이면, 임의의 양수 k에 대하여

$$\mathrm{P}(|\mathrm{Z}| \ge k) \le \mathrm{E}(|\mathrm{Z}|^r)/k^r$$

(b)(체비셰프) 확률변수 X 에 대하여 $\mathrm{Var}(\mathrm{X})<\infty$이면, 임의의 양수 k 에 대하여

$$\mathrm{P}(|\mathrm{X}-\mathrm{E}(\mathrm{X})| \ge k) \le \mathrm{Var}(\mathrm{X})/k^2$$

[Key]

(a)

$$\mathrm{P}(|\mathrm{Z}| \ge k)=\mathrm{E}(\mathrm{I}_{(|\mathrm{Z}|\ge k)})$$

$$\mathrm{I}_{(|\mathrm{Z}|\ge k)}=1 \bullet \mathrm{I}_{(|\mathrm{Z}|/k\ge 1)} \le (|\mathrm{Z}|/k)^r \bullet \mathrm{I}_{(|\mathrm{Z}|/k\ge 1)} \le |\mathrm{Z}|^r/k^r$$

$$\mathrm{P}(|\mathrm{Z}| \ge k)=\mathrm{E}(\mathrm{I}_{(|\mathrm{Z}|\ge k)}) \le \mathrm{E}(|\mathrm{Z}|^r)/k^r$$

(b) 마코프 부등식에서 $\mathrm{Z}=\mathrm{X}-\mathrm{E}(\mathrm{X}),\ r=2$

예 1.6.2: 6 시그마 밖의 확률

확률변수 X 의 평균, 표준편차를 각각 $\mu, \sigma\,(0 < \sigma < \infty)$라고 하면

$$\mathrm{P}(|\mathrm{X}-\mu| \geq 6\sigma) \leq 1/36 = 0.0277\cdots$$

즉 X 의 분포가 무엇이든 평균으로부터 여섯 배의 표준편차 밖의 확률은 2.8%이하이다.

정리 1.6.4: 분산 0의 의미(p.44)

확률변수 X 의 평균이 $\mu = \mathrm{E}(\mathrm{X})$ 이고 $\mathrm{Var}(\mathrm{X}) = 0$ 이면

$$\mathrm{P}(\mathrm{X} = \mu) = 1$$

[Key] $$(|\mathrm{X}-\mu| > 0) = \bigcup_{n=1}^{\infty}(|\mathrm{X}-\mu| \geq \frac{1}{n})$$

(가산반가법성) $$\mathrm{P}(|\mathrm{X}-\mu| > 0) = \mathrm{P}\left\{\bigcup_{n=1}^{\infty}(|\mathrm{X}-\mu| \geq \frac{1}{n})\right\} \leq \sum_{n=1}^{\infty}\mathrm{P}(|\mathrm{X}-\mu| \geq \frac{1}{n})$$

(체비셰프 부등식과 $\mathrm{Var}(\mathrm{X}) = 0$) $$0 \leq \mathrm{P}(|\mathrm{X}-\mu| \geq \frac{1}{n}) \leq \mathrm{Var}(\mathrm{X})/(1/n)^2 = 0$$

$$\therefore\ \mathrm{P}(|\mathrm{X}-\mu| \geq \frac{1}{n}) = 0 \quad (n = 1, 2, \cdots)$$

$$0 \leq \mathrm{P}(|\mathrm{X}-\mu| > 0) \leq \sum_{n=1}^{\infty}\mathrm{P}(|\mathrm{X}-\mu| \geq \frac{1}{n}) = 0$$

$$\therefore\ \mathrm{P}(|\mathrm{X}-\mu| > 0) = 0$$

$$\therefore\ \mathrm{P}(\mathrm{X} = \mu) = 1 - \mathrm{P}(|\mathrm{X}-\mu| > 0) = 1$$