제8주 연속확률분포모형

hylee@silla.ac.kr

확률및통계(2)

제 8장 연속확률분포모형

- 연속확률분포모형
- 연속확률밀도함수
- 연소확률변수의 합의 분포

제 1절 연속균일분포

[정의] 연속균일분포

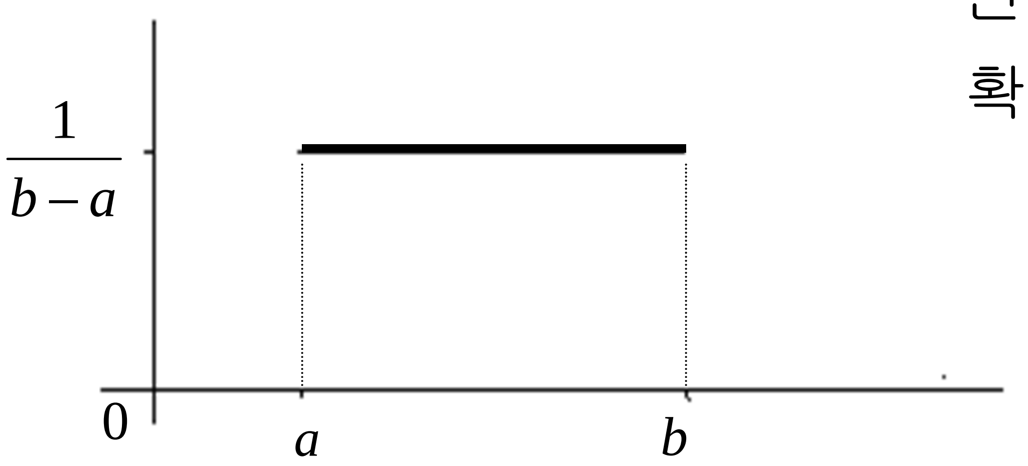
확률변수 X 가 두 점 $a, b(a < b)$ 사이에서 균일하게 분포되어 있을 때, X 의 확률밀도함수 $f(x)$ 는

$$f(x) = \begin{cases} \dfrac{1}{b-a} & (a < x < b) \\ 0 & (x \le a \text{ 또는 } x \ge b) \end{cases}$$

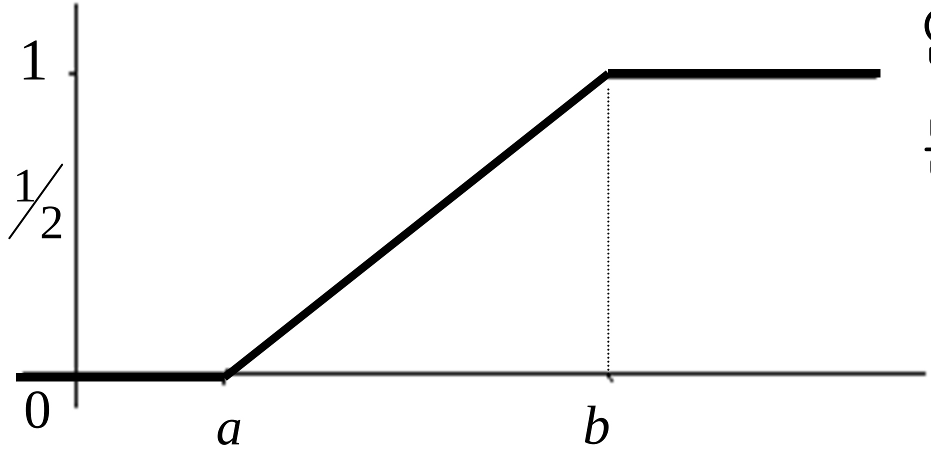
이고 , 분포함수 $F(x)$는

$$F(x)=\begin{cases} 0 & (x \le a) \\ \dfrac{x-a}{b-a} & (a < x < b) \\ 1 & (x \ge b) \end{cases}$$

이 된다. 이 확률분포를 **연속균일분포** (continuous uniform distribution) 또는 **일양분포**라 하고, $U(a,b)$로 표기한다.



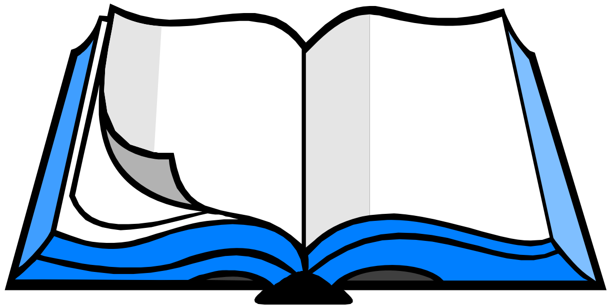
연속균등분포의
확률밀도함수



연속균등분포의
분포함수

예제 1

연속균일분포가 확률밀도함수임을 보이시오



[풀이]

$$f(x) = \frac{1}{b-a} \geq 0 \qquad (a < x < b)$$

이고,

$$\int_a^b f(x)dx = \int_a^b \frac{1}{b-a}dx = \frac{1}{b-a}\left[x\right]_a^b$$

$$= 1$$

이므로, 연속균등분포는 확률밀도 함수이다.

제 2절 정규분포

[정의] 정규분포

확률변수 X 의 확률밀도함수 $f(x)$가

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, (-\infty < x < \infty)$$

일 때, X 는 모수 μ 와 σ^2 을 갖는 정규분포 (normal distribution)를 따른다고 하고, $N(\mu, \sigma^2)$으로 표기한다.

예제 4

정규분포가 확률밀도함수임을 보이시오.

[풀이]

$$f(x)=\frac{1}{\sqrt{2\pi\sigma}}e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (-\infty < x < \infty)$$

가 다음 두 조건을 만족함을 보이자.

(1) $f(x) \geq 0$

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

지수함수는 언제나 양수값을 함수값으로 갖는다. 따라서 (1)은쉽게 보여진다. 다음으로

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

에서

$I = \int_{-\infty}^{\infty} e^{-x^2} dx$ 라 두자.

$$I^2 = \int_{-\infty}^{\infty} e^{-x^2} dx \int_{-\infty}^{\infty} e^{-y^2} dy$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-(x^2+y^2)} dx dy$$

이다. 이 때, $x = r\cos\theta,\ \ y = r\sin\theta$ 의 극좌표로 변수변환하면

$$I^2 = \int_0^{2\pi} d\theta \int_0^{\infty} r e^{-r^2} dr = \pi$$

이므로

$$I = \int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}$$

이다. 따라서, $y = \dfrac{x-\mu}{\sqrt{2}\sigma}$ 로 치환하여

적분하면 다음을 얻는다.

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \int_{-\infty}^{\infty} \frac{1}{\sqrt{\pi}} e^{-y^2} dy = 1$$

그러므로, 정규분포는 확률함수이다.

예제 5

X 가 $N(\mu,\sigma^2)$ 를 따를 때, $Y = aX + b$

$(a > 0)$ 의 확률밀도함수를 구하고

$Z = \dfrac{X - \mu}{\sigma}$ 의 확률밀도함수를구하시오.

[풀이] X 의 확률밀도함수는

$$f(x)=\frac{1}{\sqrt{2\pi\sigma}}e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (-\infty < x < \infty)$$

이다. $y=ax+b$ 의 역함수는 $x=\dfrac{y-b}{a}$

이므로, Y 의 확률밀도함수는

다음과 같다.

$$g(y)=\frac{1}{a}f\left(\frac{y-b}{a}\right)$$

$$=\frac{1}{a\sqrt{2\pi}\sigma}\exp\left\{-\frac{\left(\frac{y-b}{a}-\mu\right)^2}{2\sigma^2}\right\}$$

$$=\frac{1}{a\sqrt{2\pi}(a\sigma)}\exp\left\{-\frac{[y-(a\mu+b)]^2}{2(a\sigma)^2}\right\}$$

따라서, Y는 모수가 $a\mu+b$ 와 $(a\sigma)^2$ 인 정규분포를 따른다. 특히, $a = {}^{1}\!/_{\sigma} > 0$, $b = -{}^{\mu}\!/_{\sigma}$ 이라 놓으면, $Z = \dfrac{X-\mu}{\sigma}$의 확률밀도함수 $\varphi(z)$는 y의 확률밀도함수에서 $a\sigma = 1,\ a\mu + b = 0$ 이므로

$$\varphi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

이 된다. 이 경우 Z는 $N(0,1)$을 따른다.

위에서 구한 확률밀도함수

$$\varphi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

을 표준정규확률밀도함수

(standard normal density function)

이라 한다.

표준정규밀도함수 (STANDARD NORMAL FUNCTION)

정규분포 $N(\mu, \sigma^2)$에서 $\mu = 1, \sigma^2 = 1$일 때, 이를 표준정규분포라 한다.

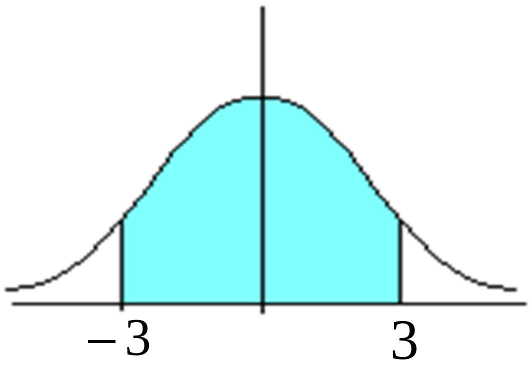
예제 6

(1) X 가 $N(0,1)$ 를 따를 때,

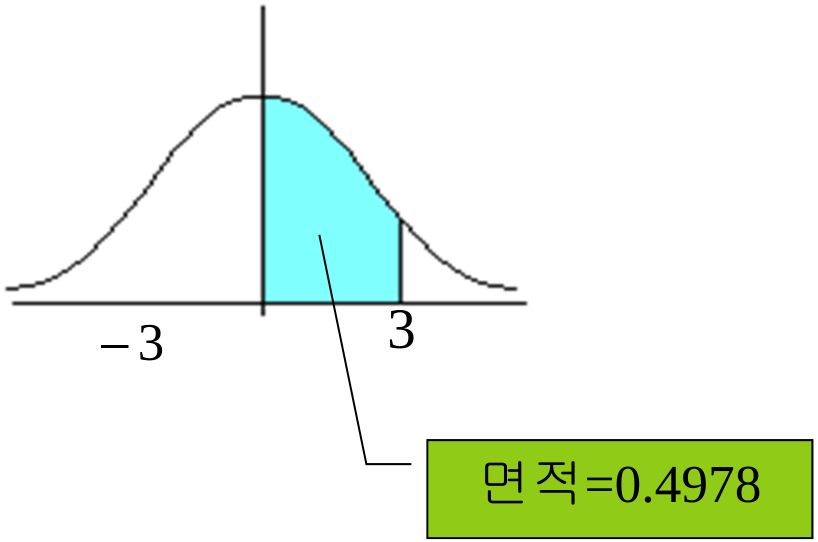
$P(-3 \le X \le 3)$ 을 구하시오.

(2) X 가 $N(2,25)$ 를 따를 때,

$P(-3 \le X \le 8)$ 을 구하시오.



$= 2\times$



$$= 2\times(0.4978)$$

$$= 0.9974$$

[풀이]

(1) $P(-3 \le X \le 3) = 2P(0 \le Z \le 3)$

$= 2 \cdot (0.4978)$

$= 0.9974$

(2) $P(-3 \le X \le 8) = P(\frac{-3-2}{5} \le \frac{X-2}{5} \le \frac{8-2}{5})$

$= P(-1 \le Z \le 1.2)$

$= P(0 \le Z \le 1) + P(0 \le Z \le 1.2)$

$= 0.3413 + 0.3849$

$= 0.7262$

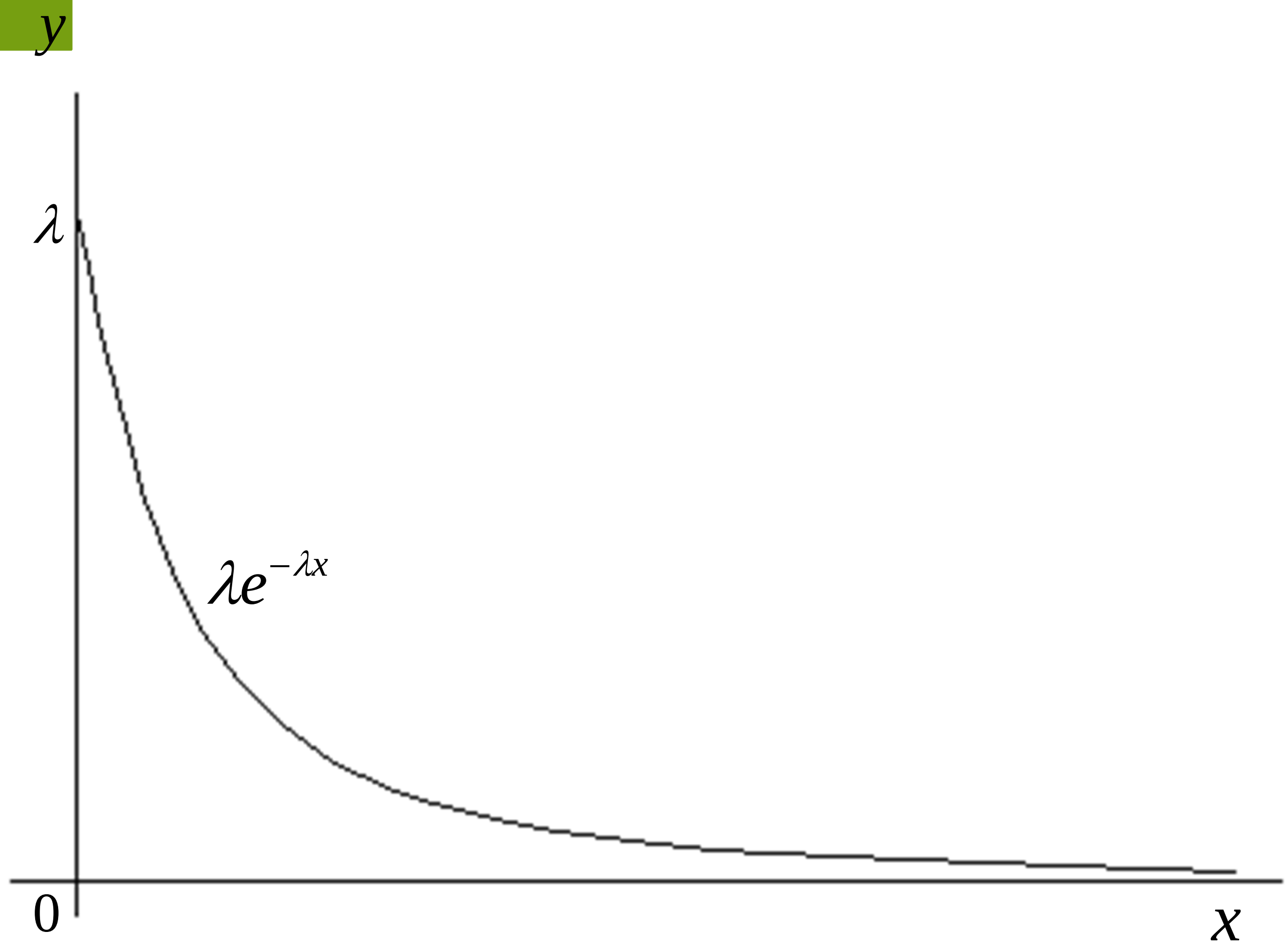
제 3절 지수분포

[정의] 지수분포

확률변수 X 의 확률밀도함수 $f(x)$ 가

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

로 주어질 때, X 는 모수가 λ 인 **지수분포** (exponential distribution) 를 따른다고 하고 $E(\lambda)$ 로 표기한다.



지수분포의 확률밀도함수의 그래프

예제 7

지수분포가 확률밀도함수임을 보이시오.

[풀이]

지수함수는 항상 양수값을 갖는다.

그러므로 $f(x) = \lambda e^{-\lambda x} \geq 0, \quad (x \geq 0)$

한편,

$$\int_0^\infty \lambda e^{-\lambda x} dx = \lim_{T\to\infty} \left[e^{-\lambda x} \right]_0^T = 1$$

이므로, 지수분포는 확률밀도함수이다.

지수분포의 분포함수를 $F(x)$ 라하면

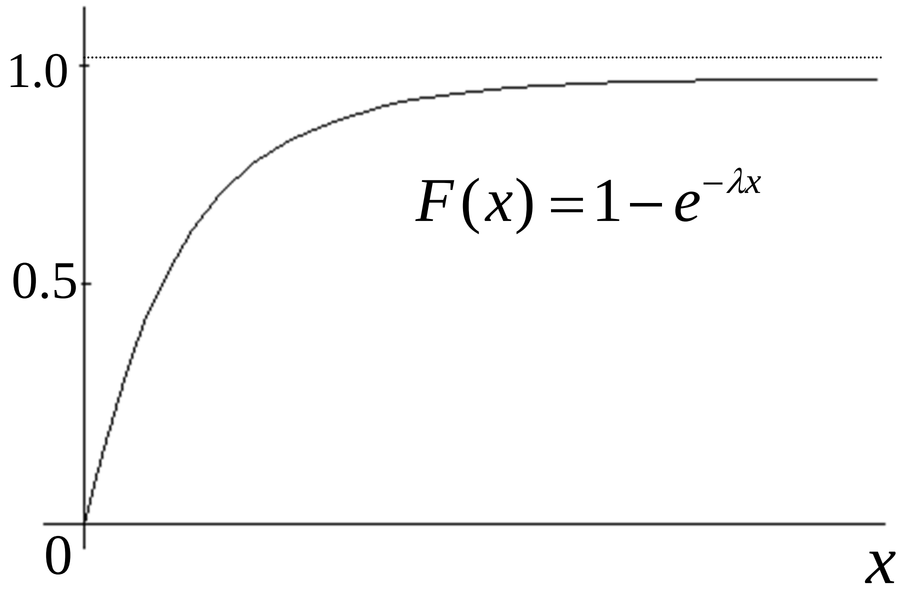
$$F(x) = P(X \le x)$$

$$= \int_0^x \lambda e^{-\lambda t} dt = [-e^{-\lambda t}]_0^x = 1 - e^{-\lambda x}$$

이므로, 다음과 같이 표현할 수있다.

$$F(x)=\begin{cases}1-e^{-\lambda x} & (x\geq 0)\\ 0 & (x<0)\end{cases}$$ 이고 그래프는

$\lambda=1$ 인 지수분포의 분포함수 그래프



예제 8

전화통화시간(분)은 $\lambda = 1/10$ 인 지수분포를 한다.

(1) 10분이상 통화할 확률을 구하시오.

(2) 10분이상 20분 이하 통화할확률을 구하시오.

[풀이]

$$(1)\ \ P(X \ge 10) = \int_{10}^{\infty} \frac{1}{10} e^{-\frac{x}{10}} dx = \left[-e^{-\frac{x}{10}} \right]_{10}^{\infty}$$

$$= e^{-1} = 0.368$$

$$(2)\ \ P(10 \le X \le 20) = \int_{10}^{20} \frac{1}{10} e^{-\frac{x}{10}} dx$$

$$= \left[-e^{-\frac{x}{10}} \right]_{10}^{20}$$

$$= e^{-1} - e^{-2} = 0.233$$

예제 9

X 가 지수분포를 할때

$P(X > x + y \mid X > x) = P(X > y)$

임을 보이시오. 이 때, 위의 성질을 만족하는 음이아닌 확률변수 X를 비기억(memoryless)이라고한다.

[풀이]

$$P(X > x+y \mid X > x) = \frac{P(X > x+y, X > x)}{P(X > x)}$$

$$= \frac{P(X > x+y)}{P(X > x)} = \frac{\int_{x+y}^{\infty} \lambda e^{-\lambda t} dt}{\int_{x}^{\infty} \lambda e^{-\lambda t} dt} = \frac{e^{-\lambda(x+y)}}{e^{-\lambda x}}$$

$$= e^{-\lambda y} = P(X > y)$$

쉬어가기

제 4절 감마분포

[정의] 감마분포

확률변수 X 가 모수 $(\alpha, \beta), \alpha > 0, \beta > 0$

를 가지며 다음 확률밀도함수

$$f(x) = \begin{cases} \dfrac{1}{\beta^{\alpha}\Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\beta}} & (x \geq 0) \\ 0 & (\mathrm{x} < 0) \end{cases}$$

을 가질 때, X 는 감마분포(gamma distribution)

를 따른다고하고 $\gamma(\alpha, \beta)$ 로 표기한다.

여기서, $\Gamma(\alpha) = \int_0^\infty e^{-x} x^{\alpha-1} dx$

는 감마함수(gamma function)이며, 부분적분

에 의해서

$$\Gamma(\alpha) = [-e^{-x} x^{\alpha-1}]_0^\infty + \int_0^\infty e^{-x} (\alpha-1) x^{\alpha-2} dx$$

$$= (\alpha-1) \int_0^\infty e^{-x} x^{\alpha-2} dx$$

$$= (\alpha-1)\Gamma(\alpha-1)$$

α 가 정수값일 때, 예를 들면 $\alpha = n$ 이고,

위의 식에서 $\Gamma(1) = \int_0^\infty e^{-x}dx = 1$ 이므로

$$
\begin{aligned}
\Gamma(n) &= (n-1)\Gamma(n-1) \\
&= (n-1)(n-2)\Gamma(n-2) \\
&\quad \cdots\cdots \\
&= (n-1)(n-2)\cdots 3\cdot 2\cdot 1\cdot \Gamma(1) \\
&= (n-1)!
\end{aligned}
$$

예제 10

감마분포가 확률밀도함수임을 보이시오.

[풀이]

$$f(x) = \frac{1}{\beta^{\alpha}\Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\beta}} \geq 0 \quad (x \geq 0)$$

이고,

$$\int_0^{\infty} \frac{1}{\beta^{\alpha}\Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\beta}} dx$$

$$= \frac{1}{\Gamma(\alpha)} \int_0^{\infty} \frac{x^{\alpha-1}}{\beta^{\alpha}} e^{-\frac{x}{\beta}} dx$$

이므로

$$\Gamma(\alpha) = \int_0^\infty e^{-y} y^{\alpha-1} dy$$

에서 $y = x/\beta$, $\beta > 0$ 라 놓으면

$$\Gamma(\alpha) = \int_0^\infty \frac{x^{\alpha-1}}{\beta^{\alpha-1}} e^{-x/\beta} \frac{1}{\beta} dx$$

$$= \int_0^\infty \frac{x^{\alpha-1}}{\beta^{\alpha}} e^{-x/\beta} dx$$

따라서

$$\frac{1}{\Gamma(\alpha)}\int_0^\infty \frac{x^{\alpha-1}}{\beta^\alpha} e^{-\frac{x}{\beta}} dx = \frac{\Gamma(\alpha)}{\Gamma(\alpha)} = 1$$

그러므로 감마분포는 확률밀도함수이다.

***

감마분포의 확률밀도함수 식에서

$\alpha = 1,\ \beta = \dfrac{1}{\lambda}$ 이면 지수분포의 확률

밀도함수와 일치한다. 즉, 지수분포는

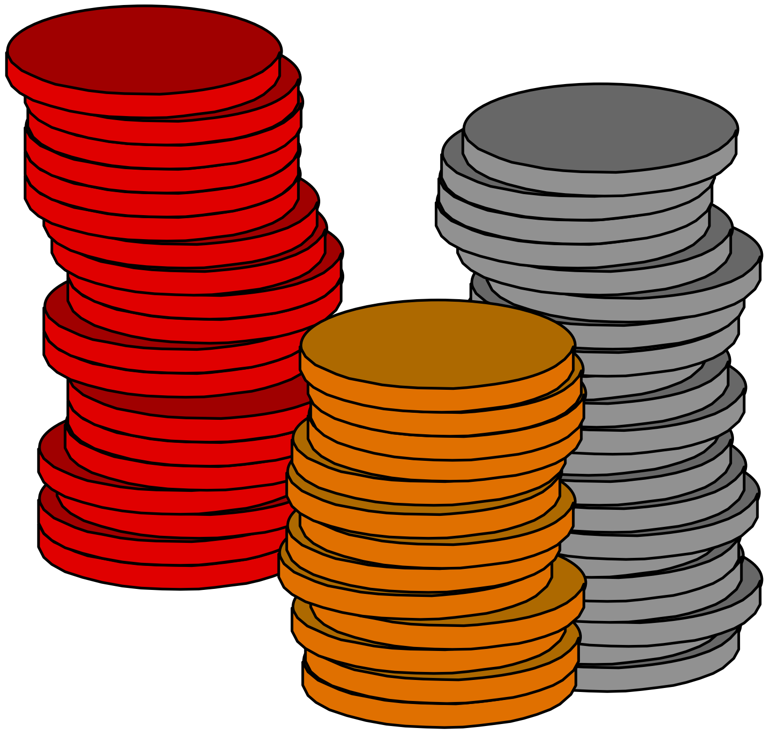
감마분포의 특수한 경우이다.

제1절 표본공간과 사상

- 표본공간과 사상을 설명한다
- 사상의 연산과 연산법칙

동전 2번 던지는 실험

- 동전을 두 번 던졌을 때 나타나는 현상들을 모두 나열해 봅시다



앞면을 H 로 뒷면을 T 로 표시하면 다음의 표본공간을 얻는다

head

tail

$$\Omega = \{HH, HT, TH, TT\}$$

- 여기서 얻어진 표본공간은 4개의 원소로 이루어진 집합이다.
- 표본공간의 한 부분집합을 **사상**(event) 이라고 정의한다.

Ω 의 부분집합의 집합 $P(\Omega)$ 를 생각하면

$P(\Omega) = \Omega$ 의 멱집합

사상

$$
= \left\{ \begin{array}{l}
\phi, \\
\{HH\},\{HT\},\{TH\},\{TT\}, \\
\{HH,HT\},\{HH,TH\},\{HH,TT\},\{HT,TH\}, \\
\{HT,TT\},\{TH,TT\}, \\
\{HH,HT,TH\},\{HH,HT,TT\},\{HH,TH,TT\}, \\
\{HT,TH,TT\}, \\
\Omega
\end{array} \right\}
$$

그러므로 예를 들면

$\{HH, HT\} =$ 첫번째 동전이 앞면인 사상

$\{HH, TT\} =$ 첫번째와 두번째 동전표면이 같은경우

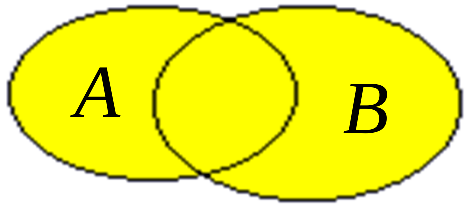
등 이다.

사상의 연산

- 1. 합사상
- 2. 곱사상
- 3. 여사상
- 4. 공사상
- 5. 배반사상

1. 합사상

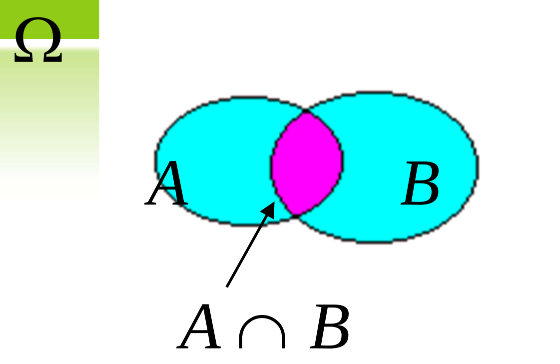
$A \cup B$



- 표본공간 Ω 의 임의의 두 사상 A 와 B 의 적어도 한쪽이 일어나는 사상을 **합사상** (union event)이라 한다.

$$A \cup B = \{\omega \mid \omega \in A \text{ 또는 } \omega \in B\}$$

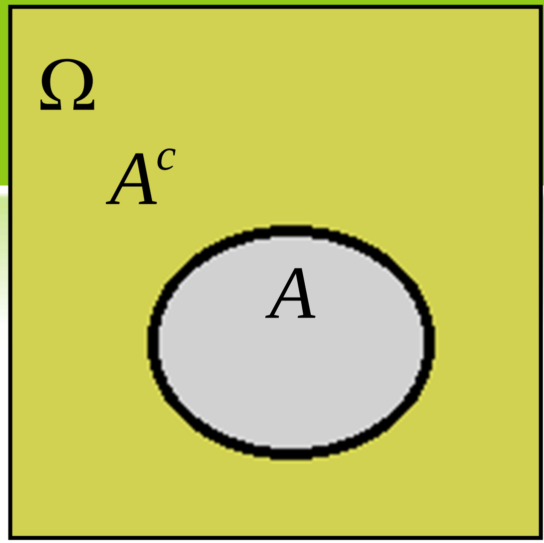
2. 곱사상



- 표본공간 Ω의 임의의 두 사상 A 와 B 가 함께 일어나는 사상을 곱사상(intersection event)이라 하며, 다음과 같이 표현한다.

$$A \cap B = \{\omega \mid \omega \in A \text{ 그리고 } \omega \in B\}$$

3. 여사상



- 사상 A가 일어나지 않는 사상을 **여사상** (complement event)이라 하며 다음과 같이 표현한다.

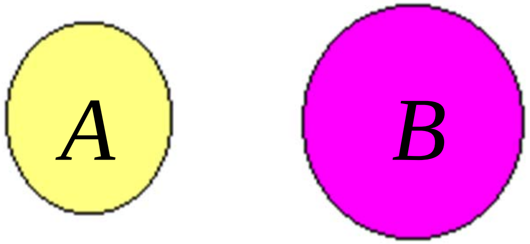
$$A^c = \{\omega \mid \omega \notin A\}$$

4. 공사상

- 사건 A 가 어떠한 결과도 포함하지 않는 경우, 이러한 사상을 공사상(null event or empty event)이라 하고 라고 표현한다.

ϕ

5. 배반사상



$$A \cap B = \phi$$

A 와 B 는

교집합이 없다.

사상의 연산법칙

교환법칙

$$A \cup B = B \cup A, \quad A \cap B = B \cap A$$

결합법칙

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

사상의 연산법칙

배분법칙

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

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

드모르강$(De\,Morgan)$의법칙

$$(A \cup B)^c = A^c \cap B^c$$

$$(A \cap B)^c = A^c \cup B^c$$

예제 4

동전을 두번 던지는 시행에서

$$A_1 = \{HH\}, \qquad A_2 = \{HT\},$$

$$A_3 = \{TH\}, \qquad A_4 = \{TT\}$$

라 할 때 다음사상

$$E_1 = A_1 \cup A_2 \cup A_3,$$

$$E_2 = A_3 \cup A_4$$

의 확률 $P(E_1), P(E_2)$ 를 구하시오.

여기서

$$P(A_1) = \frac{1}{16}, \quad P(A_2) = \frac{3}{16}$$

$$P(A_3) = \frac{3}{16}, \quad P(A_4) = \frac{9}{16}$$

라 가정한다

풀이

$$P(E_1) = P(A_1 \cup A_2 \cup A_3)$$

$$= P(A_1) + P(A_2) + P(A_3)$$

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

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

$$= \frac{1}{16} + \frac{3}{16} + \frac{3}{16} - \{0 + 0 + 0\} + 0 = \frac{7}{16}$$

(2) $A_1 \cap A_2 = 1$이 첫번째, 2가 두번째 자리에 위치하는 사상이므로,

$$A_1 \cap A_2$$
$$= \{(1,2,\boxed{k_1,\cdots,k_{n-2}}) \mid k_1,\cdots,k_{n-2} = 3,\cdots,n\}$$

$|A_1 \cap A_2| = (n-2)!$ 이므로

구하는 확률은

$$P(A_1 \cap A_2) = \frac{(n-2)!}{n!} = \frac{1}{n(n-1)}$$

이다.

(3) $P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2)$
을 이용한다.

따라서

$$P(A_1 \cup A_2) = \frac{1}{n} + \frac{1}{n} - \frac{1}{n(n-1)}$$

$$= \frac{2n-3}{n(n-1)}$$

(4) 어느 숫자도 제 번호에 있지 않으려면 사상은

첫번째 자리도 1 이 아니고 (즉, A_1^c),

두번째 자리도 2 가 아니고 (A_2^c),

$\cdots,$

n 번 째 자리도 n 이 아니어야 한다.

즉,

$$A_1^c \cap A_2^c \cap \cdots \cap A_n^c$$

의확률을 구하는 것이다.

그런데

$$A_1^c \cap A_2^c \cap \cdots \cap A_n^c$$

$$= (A_1 \cup A_2 \cup \cdots \cup A_n)^c$$

의확률을 구하는 것이다.

즉,

$A_1 \cup A_2 \cup \cdots \cup A_n$ 의 여사상을 의미한다.

그러므로 구하고자 하는 확률은

$$1 - P(A_1 \cup A_2 \cup \cdots \cup A_n)$$

이고,

위에서 배운 확률계산법을 이용하면,

$$1-P(A_1 \cup \cdots \cup A_n)$$

$$=1-\{\sum_{i} P(A_i)-\sum_{i<j} P(A_i \cap A_j)$$

$$+\sum_{i<j<k} P(A_i \cap A_j \cap A_k)-\cdots$$

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

이다.