# 5장. 연속확률분포







## 연속균등분포(continuous uniform distribution)

확률변수 $X$의 확률밀도함수가

$$f_X(x) = \begin{cases} \dfrac{1}{b-a}, & a < x < b \\ 0, & \text{otherwise} \end{cases}, \quad -\infty < a < b < \infty$$

이면, $X$는 연속균등분포를 따른다고 하고 $X \sim U_{(a,b)}$로 나타낸다.

### Theorem

$X \sim U_{(a,b)}$ 이면,

(1) $E[X] = \frac{a+b}{2}$

(2) $\mathrm{Var}[X] = \frac{(b-a)^2}{12}$





## Theorem

$X \sim U_{(a,b)}$ 이면, $m_X(t) = \begin{cases} \dfrac{e^{bt} - e^{at}}{(b-a)t}, & t \neq 0 \\ 1, & t = 0 \end{cases}$

**Proof.** $m_X(t) = E\left[e^{tX}\right]$

(1) $t = 0$이면,

$$m_X(0) = E\left[1\right] = 1$$

(2) $t \neq 0$이면,

$$m_X(t) = E\left[e^{tX}\right] = \int_{-\infty}^{\infty} e^{tx} f_X(x)\, dx$$

$$= \int_a^b e^{tx} \cdot \frac{1}{b-a}\, dx = \left[\frac{1}{b-a} \cdot \frac{1}{t} e^{tx}\right]_a^b = \frac{e^{bt} - e^{at}}{(b-a)t}$$





## Example

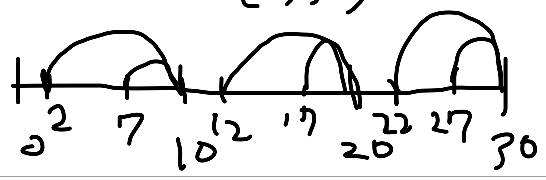
지하철이 정해진 정차역에 오전 $6:00$부터 $10$분 간격으로 도착한다고 한다. 지하철을 타려는 승객이 이 정차역에 도착하는 시각은 오전 $6:00$에서 오전 $6:30$ 사이에 연속균등분포를 따른다고 하다.

(1) 이 승객이 $3$ 분 이내에 승차할 수 있는 확률을 구하시오.

(2) 이 승객이 기다리는 시간이 적어도 $8$분일 확률을 구하시오.

(3) 이 승객이 $6:11$이 넘어서 도착 했을 때, 앞으로 적어도 $8$분을 기다릴 확률을 구하시오.

$X \sim U(0,30)$ $\quad f_X(x) = \begin{cases} \frac{1}{30}, & 0<x<30 \\ 0, & \text{o/w} \end{cases}$

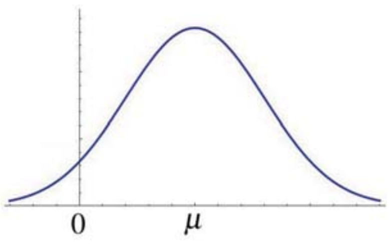


## 정규분포(Normal distribution)

확률변수 $X$의 확률밀도함수가

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} \quad (\sigma > 0)$$

이면, 확률변수 $X$는 정규분포를 따른다고 하고 $X \sim N(\mu, \sigma^2)$로 나타낸다.








## Theorem

$X \sim N(\mu, \sigma^2)$이면,

(1) $E[X] = \mu$

(2) $\mathrm{Var}[X] = \sigma^2$

(3) $m_X(t) = \exp\left\{\mu t + \frac{1}{2}\sigma^2 t^2\right\}$

**Proof.**

$$\begin{aligned}
(3)\quad m_X(t) = E\left[e^{tX}\right] &= \int_{-\infty}^{\infty} e^{tx} f_X(x)\, dx = \int_{-\infty}^{\infty} e^{tx} \cdot \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\, dx \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{x^2 - 2(\mu+\sigma^2 t)x + \mu^2}{2\sigma^2}\right\} dx \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{\left(x-(\mu+\sigma^2 t)\right)^2}{2\sigma^2}\right\} \cdot \exp\left\{-\frac{\mu^2 - (\mu+\sigma^2 t)^2}{2\sigma^2}\right\} dx \\
&= \exp\left\{\mu t + \frac{\sigma^2 t^2}{2}\right\} \cdot \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{\left(x-(\mu+\sigma^2 t)\right)^2}{2\sigma^2}\right\} dx \\
&= \exp\left\{\mu t + \frac{\sigma^2 t^2}{2}\right\}
\end{aligned}$$





## Theorem

(1) $X \sim N(\mu, \sigma^2)$이고 $Y = aX + b,\ a \neq 0$이면
$Y \sim N(a\mu + b, (a\sigma)^2)$이다.

(2) $X \sim N(\mu_1, \sigma_1^2),\ Y \sim N(\mu_2, \sigma_2^2)$이고 $X$와 $Y$가 독립이면,
$W = X + Y \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$이다.

(3) $X \sim N(\mu, \sigma^2)$이고 $Z = \dfrac{X-\mu}{\sigma}$이면 $Z \sim N(0,1)$이다.

## 표준정규분포(Standard Normal distribution)

$Z \sim N(0,1)$일 때, 확률변수 $Z$는 표준정규분포를 따른다고 한다.





## Example

$X \sim N(175, 5^2)$

표준정규분포

1. 어느 직장의 신입사원들의 키($X$)는 근사적으로 평균이 $175cm$, 표준편차가 $5cm$인 정규분포를 따른다고 한다. 임의로 선택한 신입사원의 키가 $170cm$ 보다 크고 $180cm$ 보다 작은 확률을 구하시오.

   $P[170 < X < 180]$
   $= P[170 < X' < 180] = P\left[\frac{170-175}{5} < \frac{X'-175}{5} < \frac{180-175}{5}\right]$
   $= P[-1 < Z < 1]$

2. 수학과 학생 40명의 통계학 성적분포는 평균 70점, 표준편차 10점인 정규분포를 따른다고 한다. 통계학 성적이 60점 미만인 학생은 약 몇 명인지 구하시오.

   $N(70, 10^2)$ $\quad P[X < 60] = P\left[Z < \frac{60-70}{10}\right]$
   $= P[Z < -1]$

## 이항분포의 정규근사식

$X \sim B(n,p)$이라 하고 $np > 5$, $n(1-p) > 5$이라 하자. $X' \sim N(np, np(1-p))$ 일 때, 중심극한의 정리에 의해 $X \approx X'$ 이다. 그러면 임의의 $x \in R_X = \{0,1,2,\cdots,n\}$에 대해

$$P[X \leqslant x] = P[X < x+1]$$

이고, $P[X \leqslant x]$는 $P[X' \leqslant x]$ 보다 $P[X' \leqslant x+0.5]$로 더 잘 근사된다. 임의의 $a,b \in R_X = \{0,1,2,\cdots,n\}$에 대하여 다음과 같이 근사시키고, 이를 근삿값의 연속성 보정(continuous correction)이라 한다.

$$P[a \leqslant X \leqslant b] = \sum_{x=a}^{b} f_X(x) = \sum_{x=a}^{b} \binom{n}{x} p^x (1-p)^{n-x}$$
$$\approx P\left[a - 0.5 \leqslant X' \leqslant b + 0.5\right]$$





## Example

$X \sim B\left(1200, \frac{1}{4}\right)$

100원짜리 동전을 2개 던지는 실험을 1200회 시행하여 2개 모두 뒷면이 나올 횟수를 $X$라 할 때, 다음 확률을 구하시오.

(1) $X$가 280 이상 310 이하일 확률

(2) $X$가 260 이상일 확률

$X' \sim N(300, 225)$

$P[280 \leq X \leq 310]$

$\fallingdotseq P[280 \leq X' \leq 310]$

$\fallingdotseq P[280-0.5 \leq X' \leq 310+0.5]$

$= P\left[\bigcirc \leq \frac{X'-300}{\sqrt{225}} \leq \bigcirc\right]$





## 지수분포(Exponential distribution)

확률변수 $X$의 확률밀도함수가

$$f_X(x) = \begin{cases} \frac{1}{\beta} e^{-x/\beta}, & x > 0 \\ 0, & \text{otherwise} \end{cases}, \quad \beta > 0$$

$X \sim exp(\lambda)$

$$\begin{cases} \lambda e^{-x\lambda}, & x > 0 \\ 0, & o.w \end{cases}$$

이면, $X$는 모수 $\beta$를 갖는 지수분포를 따른다고 하고 $X \sim \exp(\beta)$로 나타낸다.

### Remark

$$\int_{-\infty}^{\infty} f_X(x)\,dx = \int_0^{\infty} \frac{1}{\beta} e^{-x/\beta}\,dx = \left[-e^{-x/\beta}\right]_0^{\infty} = 1$$





## Theorem

$X \sim \exp(\beta)$이면,

(1) $E[X] = \beta, \quad \mathrm{Var}[X] = \beta^2$

(2) $m_X(t) = \frac{1}{1-\beta t}, \quad t < \frac{1}{\beta}$

$X \sim exp(\lambda)$

$E[X] = \frac{1}{\lambda}$

**Proof.**

$$m_X(t) = E\left[e^{tX}\right] = \int_{-\infty}^{\infty} e^{tx} f_X(x)\,dx$$
$$= \int_0^{\infty} e^{tx} \cdot \frac{e^{-x/\beta}}{\beta}\,dx = \int_0^{\infty} \frac{e^{-\frac{1-\beta t}{\beta}x}}{\beta}\,dx$$
$$= \left[-\frac{1}{1-\beta t} e^{-\frac{1-\beta t}{\beta}x}\right]_0^{\infty} = \frac{1}{1-\beta t}, \quad 1 - \beta t > 0$$

## Theorem(무기억성)

$X \sim \exp(\beta)$이면, 임의의 $a, b > 0$에 대하여,

$$P\Big[X > a+b \;\Big|\; X > a\Big] = P[X > b].$$

**Proof.**

$$P[X > b] = \int_b^\infty \frac{1}{\beta} e^{-x/\beta}\, dx = \Big[-e^{-x/\beta}\Big]_b^\infty = e^{-b/\beta}$$

$$P\Big[X > a+b \;\Big|\; X > a\Big] = \frac{P[X > a+b]}{P[X > a]} = \frac{e^{-(a+b)/\beta}}{e^{-a/\beta}} = e^{-b/\beta}$$





$X \sim \exp(3)$

$f_X(x) = \begin{cases} \frac{1}{3} e^{-\frac{x}{3}}, & x > 0 \\ 0, & o/w \end{cases}$

## Example

1. 퓨즈의 수명 $X$는 평균이 3인 지수분포를 따른다고 하자. 즉, $X \sim \exp(3)$.
   (1) 퓨즈의 수명이 10년이 넘을 확률을 구하시오. $p$
   (2) 퓨즈를 동시에 5개 사용한 후에 10년 후에도 절반 이상이 작동할 확률을 구하시오. $Y$

   (1) $P[X > 10] = \int_{10}^\infty \frac{1}{3} e^{-\frac{x}{3}} dx = p$

   (2) $Y \sim B(5, p)$ $\quad P[Y \geq 3] = \sum_{x=3}^{5} \binom{5}{x} p^x (1-p)^{5-x}$

2. 어떤 스마트폰 건전지의 충전 후 사용시간은 평균 48 시간인 지수분포를 따른다고 한다. 만약 이 건전지를 충전 후 12 시간 동안 사용했다면 앞으로 40 시간 더 사용할 확률을 구하시오.

$P[X \geq 12+40 \mid X > 12] = P[X > 40]$

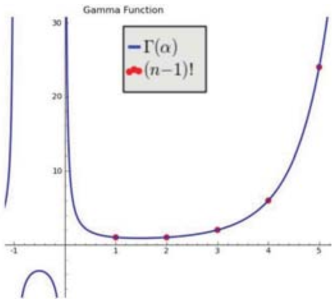




## 감마함수(Gamma function)

감마함수는 다음과 같이 정의된다.

$$\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t}\, dt, \quad \alpha > 0$$




### Theorem

(1) $\Gamma(1) = 1$

(2) $\forall \alpha > 0,\ \Gamma(\alpha+1) = \alpha\,\Gamma(\alpha)$.
특히, $\forall n \in \mathbb{N},\ \Gamma(n+1) = n!$

(3) $\Gamma(\frac{1}{2}) = \sqrt{\pi}$





## 감마분포(Gamma distribution)

확률변수 $X$의 확률밀도함수가

$$f_X(x) = \begin{cases} \dfrac{x^{\alpha-1} e^{-x/\beta}}{\Gamma(\alpha)\beta^\alpha}, & x > 0 \\ 0, & \text{otherwise} \end{cases}, \quad \alpha > 0,\ \beta > 0$$

이면, $X$는 모수 $\alpha$와 $\beta$를 갖는 감마분포를 따른다고 하고 다음과 같이 나타낸다. $X \sim \mathrm{Ga}(\alpha, \beta)$로 나타낸다.

$X \sim Ga(1, \beta) = X \sim \exp(\beta)$

### Remark

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

$$= \frac{1}{\Gamma(\alpha)} \int_0^\infty t^{\alpha-1} e^{-t}\, dt = \frac{1}{\Gamma(\alpha)} \cdot \Gamma(\alpha) = 1$$

$$(\ t = x/\beta, \quad dt = \tfrac{1}{\beta}\, dx\ )$$

## Theorem

$X \sim \text{Ga}(\alpha, \beta)$이면,

(1) $E[X] = \alpha\beta$ (3) $m_X(t) = \left(\frac{1}{1-\beta t}\right)^{\alpha}, \quad t < \frac{1}{\beta}$

(2) $\text{Var}[X] = \alpha\beta^2$ (4) $\text{Ga}(1, \beta) = \exp(\beta)$

**Proof.**

$$m_X(t) = E\left[e^{tX}\right] = \int_{-\infty}^{\infty} e^{tx} \cdot f_X(x)\,dx = \int_0^{\infty} e^{tx} \cdot \frac{x^{\alpha-1}e^{-x/\beta}}{\Gamma(\alpha)\beta^{\alpha}}\,dx$$

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

$$= \left(\frac{1}{1-\beta t}\right)^{\alpha}, \quad 1-\beta t > 0$$

$\text{Ga}\left(\alpha, \frac{\beta}{1-\beta t}\right)$





## 감마분포와 푸아송분포의 관계

$X \sim \text{Pois}(m)$이라고 하고 $X_{(0,t)}$를 시간 $(0,t)$에서 일어나는 사건의 출현수라고 하면 $X_{(0,t)} \sim \text{Pois}(mt)$ 이므로,

$$f_{X_{(0,t)}}(x) = P[X_{(0,t)} = x] = \begin{cases} \frac{(mt)^x e^{-mt}}{x!}, & x \in \{0, 1, 2, \cdots\} \\ 0, & \text{otherwise} \end{cases}$$

이다. $T$를 $\alpha$ 번째 사건이 일어날 때까지의 시간의 길이라고 하면,

$$F_T(t) = P[T \leqslant t] = P\left[X_{(0,t)} \geqslant \alpha\right] = \begin{cases} 1 - P\left[X_{(0,t)} < \alpha\right], & t > 0 \\ 0, & \text{o/w} \end{cases}$$

이고, $t > 0$이면

$$1 - P\left[X_{(0,t)} < \alpha\right] = 1 - \sum_{k=0}^{\alpha-1} P\left[X_{(0,t)} = k\right] = 1 - \sum_{k=0}^{\alpha-1} \frac{(mt)^k e^{-mt}}{k!}$$

이다.





$$\frac{d}{dt}\left[1 - \sum_{k=0}^{\alpha-1}\frac{(mt)^k e^{-mt}}{k!}\right] = \frac{d}{dt}\left[1 - e^{-mt} - \sum_{k=1}^{\alpha-1}\frac{(mt)^k e^{-mt}}{k!}\right]$$

$$= me^{-mt} - e^{-mt}\sum_{k=1}^{\alpha-1}\left[\frac{-m(mt)^k}{k!} + \frac{m(mt)^{k-1}}{(k-1)!}\right]$$

$$= \frac{m(mt)^{\alpha-1}}{(\alpha-1)!}e^{-mt}$$

$$= \frac{t^{\alpha-1}}{\Gamma(\alpha)\beta^{\alpha}}e^{-t/\beta}, \qquad \beta = \frac{1}{m}$$

가 되어 $T$는 모수 $\alpha$, $\beta = \frac{1}{m}$인 감마분포가 된다. 즉,

$$T \sim \text{Ga}(\alpha, \beta) = \text{Ga}\left(\alpha, \frac{1}{m}\right)$$





## Example

$X_{시간} \sim \text{Pois}(15)$
$X_{1분} \sim \text{Pois}\left(\frac{15}{60}\right)$
$T \sim \text{Ga}(1, 4)$
$= T \sim \exp(4)$

1. 어떤 가게에서 만 원 이상 구매하는 고객들의 수는 시간당 15명의 비율로 도착하는 푸아송분포를 따른다고 하자. 만 원 이상 구매하는 첫 고객이 생길 때 까지 가게 종업원이 기다려야 하는 시간이 4분 이상일 확률을 구하시오.

   $P[T \geq 4] = \int_4^{\infty} \frac{1}{4}e^{-t/4}\,dt$

   $f_T(t) = \begin{cases} \frac{1}{4}e^{-t/4}, & t>0 \\ 0 & \text{o/w} \end{cases}$

2. 어떤 햄버거 가게에는 시간당 평균 30명의 손님이 오는 비율로 푸아송분포를 따른다고 한다. $X_{시간} \sim \text{Pois}(30) = X_{분} \sim \text{Pois}\left(\frac{1}{2}\right)$
   (1) 처음 두 명의 손님이 방문하기까지 가게 주인이 기다리는 시간이 7분 이상일 확률을 구하시오. $T \sim \text{Ga}(2, 2)$
   (2) 처음 두 명의 손님이 방문하기까지 가게 주인이 기다리는 시간의 평균과 분산을 구하시오.

   $E(X) = \alpha\beta = 4$

   $V[X] = \alpha\beta^2 = 8$

$P[T \geq 7] = \int_7^{\infty} \frac{x^{\alpha-1}e^{-x/\beta}}{\Gamma(\alpha)\beta^{\alpha}}\,dx = \frac{9}{2}e^{-\frac{7}{2}}$