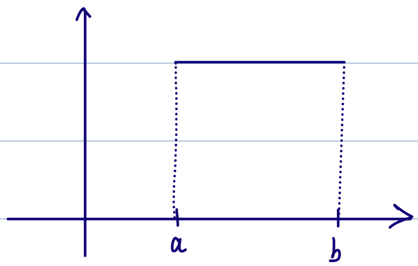
$P(X=x)=0$ 을 만족하면 연속형 확률변수.

균일분포 $X\sim U(a,b)$

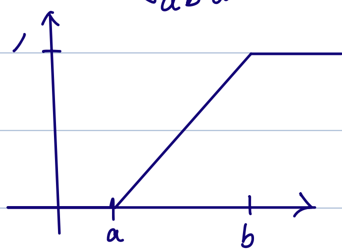
$f(x) = \begin{cases} \frac{1}{b-a} & a\le x\le b \text{ 인 경우} \\ 0 & \text{otherwise}\end{cases}$

$F(x) = P(X\le x) = \int_{-\infty}^{x} f(t)\,dt = \int_{-\infty}^{x}\frac{1}{b-a}I(a\le t\le b)\,dt$

$\int_a^x \frac{1}{b-a}\,dt = \begin{cases}0 & x<a\\ \frac{x-a}{b-a} & a\le x<b \\ 1 & x\ge b\end{cases}$



확률밀도함수(pdf)



확률분포함수

$E(X)=\int_a^b \frac{x}{b-a}\,dx = \frac{b+a}{2}$ $\quad Var(X)=\frac{(b-a)^2}{12}$

· 지수분포. $X\sim Exp(\theta)$

포아송분포를 고려했을 때, 특정한 사건 A가 일어나고 다음 사건이 일어날 때까지 걸리는 시간을 W라고 할 때

$F(w)=P(W\le w) = 1-P$[구간 $[0,w)$에서 사건 A가 일어나지 않음] $=1-e^{-\lambda w}$ 이다.

$F(w)$는 시간 w안에 적어도 한 번 사건이 발생할 확률이다.

즉 $1-P(W>w)$는 $1-P$[구간 $[0,w)$에서 사건 A가 일어나지 않음]으로 표현이 가능하다.

포아송분포에서 P[구간 $[0,w)$에서 사건 A가 일어나지 않음]은 $e^{-\lambda w}$로 계산된다.

왜? 구간을 n으로 나누었을 때, n개의 구간에서 사건 A가 발생하지 않을 확률은 $\prod_{i=1}^{n}(1-\lambda\cdot\frac{w}{n})$이다.

$\lim_{n\to\infty}(1-\lambda\frac{w}{n})^n$은 지수함수의 정의에 따라 $e^{-\lambda w}$이다.

즉 $F(w)=1-e^{-\lambda w}$이다.

(미분) /

$f_W(w)=\lambda e^{-\lambda w},\ w>0$ $\theta=\frac{1}{\lambda}$이라고 하면. $f(x)=\frac{1}{\theta}e^{-\frac{x}{\theta}}I(x>0)$ $F(x)=1-e^{-\frac{x}{\theta}}$

$$E(X)=\int_0^\infty x\cdot\frac{1}{\theta}\cdot e^{-\frac{x}{\theta}}dx=-xe^{-\frac{x}{\theta}}\Big|_0^\infty+\int_0^\infty e^{-\frac{x}{\theta}}dx=\theta$$

$$Var(X)=\int_0^\infty x^2\frac{1}{\theta}e^{-\frac{x}{\theta}}dx=\theta^2$$

포아송 확률변수에서 모수 λ가 커지면 특정 시간의 발생횟수가 증가한다는 의미

⇒ 즉 사건 사이 시간 간격은 줄어든다. 즉, $\lambda=\frac{1}{\theta}$가 성립.

포아송분포에서의 시간 간격의 확률분포를 지수분포라고 한다.

E. 2.37

α-입자가 10초당 평균 5개씩 포아송 과정을 따라 발생. $P(X\geq 5)$를 구하여라.

1초당 $\frac{1}{2}$개 $\Rightarrow \lambda=\frac{1}{2} \Rightarrow \theta=2$ $P(X\geq 5)=\int_5^\infty\frac{1}{2}e^{-\frac{x}{2}}dx=e^{-\frac{5}{2}}=0.08208$

· 무기억성 : $X\sim EXP(\theta)$ 이면 $P(X>a+t\mid X>a)=P(X>t)$

정규분포 $X\sim N(\mu,\sigma^2)$

$$f(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}I(-\infty<x<\infty)\quad E(X)=\mu,\ Var(x)=\sigma$$

정규분포는 선형변환을 해도 다시 정규분포를 따른다.

$X\sim N(\mu,\sigma^2)$일 때, $Y=aX+b$의 분포는 $Y\sim N(a\mu+b,a^2\sigma^2)$이다.

<증명>

$$F_Y(y)=P(Y\leq y)=P(aX+b\leq y)=P\left(X\leq\frac{y-b}{a}\right)=\int_{-\infty}^{\frac{y-b}{a}}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}dx\quad\left(\frac{t-b}{a}=x\text{로 치환}\right)$$

$$=\int_{-\infty}^{y}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{1}{2}\left(\frac{\frac{t-b}{a}-\mu}{\sigma}\right)^2}dx=\int_{-\infty}^{y}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{1}{2}\left(\frac{t-a\mu-b}{a\sigma}\right)^2}dx$$

$\int_{-\infty}^{\infty} \phi(z)dz = 1$ 이다.

$$E(X) = \int_{-\infty}^{\infty} x \cdot \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\cdot\left(\frac{x-\mu}{\sigma}\right)^2} dx = \int_{-\infty}^{\infty} (\mu + z\sigma)\cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz = \int_{-\infty}^{\infty} (\mu + z\sigma)\cdot \phi(z) dz$$

$$= \int_{-\infty}^{\infty} \mu \cdot \phi(z) dz + \int_{-\infty}^{\infty} z\sigma \cdot \phi(z) dz = \mu$$ ($z\phi(z)$는 적분가능한 기함수이므로 적분시 0이다.)

$$E(X^2) = \int_{-\infty}^{\infty} x^2 \cdot \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx = \int_{-\infty}^{\infty} (\mu + z\sigma)^2 \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz$$

$$= \mu^2 \int_{-\infty}^{\infty} \phi(z) dz + \underbrace{2\mu\sigma \int_{-\infty}^{\infty} z\phi(z) dz}_{=0} + \underline{\sigma^2 \int_{-\infty}^{\infty} z^2 \phi(z) dz} = \mu^2 + \sigma^2$$ 이므로

$$\hookrightarrow \int_{-\infty}^{\infty} z^2 \phi(z) dz = \int_{-\infty}^{\infty} (\phi''(z) + \phi(z)) dz = 1$$

$$Var(X) = E(X^2) - E(X)^2 = \sigma^2$$

$\Phi(x) = P(Z \le z) = \int_{-\infty}^{x} \phi(t) dt$, $\Phi(x) = 1 - \Phi(-x)$, $\underline{\Phi(x) = F\left(\frac{x-\mu}{\sigma}\right)}$

$$\hookrightarrow F(x) = P(X \le x) = P\left(\frac{X-\mu}{\sigma} \le \frac{x-\mu}{\sigma}\right)$$

$$= P\left(Z \le \frac{x-\mu}{\sigma}\right) = \Phi\left(\frac{x-\mu}{\sigma}\right).$$

E. 2.38

$X \sim N(25, 16)$ $\quad P(20 \le X \le 35) = P\left[\frac{20-25}{4} \le Z \le \frac{35-25}{4}\right] = P[-1.25 \le Z \le 2.5]$

$$= \Phi(2.5) - \Phi(-1.25) = 0.8882$$

$P(|X-25| \le C) = 0.9$ $\quad P(|X-25| \le c) = P(-c \le X-25 \le c) = P\left(-\frac{c}{4} \le Z \le \frac{c}{4}\right)$

$$= \Phi\left(\frac{c}{4}\right) - \Phi\left(-\frac{c}{4}\right) = 0.90 \quad \left(\Phi\left(\frac{c}{4}\right) + \Phi\left(-\frac{c}{4}\right) = 1\right)$$

$$= \Phi\left(\frac{c}{4}\right) - 1 + \Phi\left(\frac{c}{4}\right) = 0.9 \Rightarrow 2\Phi\left(\frac{c}{4}\right) = 1.9 \Rightarrow \Phi\left(\frac{c}{4}\right) = 0.95.$$

$$\frac{c}{4} = 1.645 \Rightarrow C = 6.58$$

이변량 정규분포 $(X,Y) \sim BVN(\mu_X, \mu_Y, \sigma_X, \sigma_Y, \rho)$

$$f_{X,Y}(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_X}{\sigma_X}\right)^2 - 2\rho\frac{x-\mu_X}{\sigma_X}\cdot\frac{x-\mu_Y}{\sigma_Y} + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2\right]}$$

$-\infty < x, y < \infty, \ -\infty < \mu_X, \mu_Y < \infty, \ 0 < \sigma_X, \sigma_Y < \infty, \ -1 < \rho < 1$

$\mu_X = E(X), \ \mu_Y = E(Y), \ Var(X) = \sigma_X^2, \ Var(Y) = \sigma_Y^2, \ Cov(X,Y) = \rho\sigma_X\sigma_Y$

· 이변량 확률벡터 (X,Y)가 $BVN(\mu_X, \mu_Y, \sigma_X, \sigma_Y, \rho)$를 따르면

~~주변확률밀도함수~~ X는 $N(\mu_X, \sigma_X)$를 따르고 Y는 $N(\mu_Y, \sigma_Y)$를 따른다.

감마분포 $X \sim GAM(k, \theta)$

지수분포들 여러개의 분포를 포함하는 분포족(family of distributions)이다.

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

$\Gamma(1) = \int_0^\infty e^{-t} dt = 1.$ (i) $\Gamma(k) = (k-1)\Gamma(k-1)$ (ii) $\Gamma(n) = (n-1)!$ (iii) $\Gamma(\frac{1}{2}) = \sqrt{\pi}$

$$f(x; k, \theta) = \frac{1}{\theta^k \Gamma(k)} x^{k-1} \cdot e^{-\frac{x}{\theta}}, \ x>0$$

$k=1$이라면 모수가 θ인 지수분포를 따른다.

$$E(X) = \int_{-\infty}^{\infty} x \cdot \frac{1}{\theta^k \cdot \Gamma(k)} x^{k-1} \cdot e^{-\frac{x}{\theta}} I(x>0)\, dx = \int_0^\infty \frac{1}{\theta^k \cdot \Gamma(k)} \cdot x^{(1+k)-1} \cdot e^{-\frac{x}{\theta}} dx$$

$$= \int_{-\infty}^{\infty} \frac{1}{\theta^k \Gamma(k)} \cdot \theta^{k+1} \cdot \Gamma(k+1) \cdot \underbrace{\frac{1}{\theta^{k+1}\Gamma(k+1)} x^{(1+k)-1} \cdot e^{-\frac{x}{\theta}}}_{GAM(k+1,\theta)} dx = \frac{\theta^{k+1} \cdot \Gamma(k+1)}{\theta^k \Gamma(k)} = \frac{\cancel{\theta^k} \cdot \theta \cdot k\cancel{\Gamma(k)}}{\cancel{\theta^k}\ \cancel{\Gamma(k)}} = k\theta$$

$$E(X^2) = \int_0^\infty x^2 \frac{1}{\theta^k \Gamma(k)} x^{k-1} \cdot e^{-\frac{x}{\theta}} dx = \int_0^\infty \frac{1}{\theta^k \Gamma(k)} \underbrace{x^{(k+2)-1} \cdot e^{-\frac{x}{\theta}} \cdot \frac{1}{\theta^{k+2}\Gamma(k+2)}}_{GAM(k+2,\theta)} \cdot \theta^{k+2} \cdot \Gamma(k+2)\, dx$$

$$= \frac{\theta^{k+2}\Gamma(k+2)}{\theta^k \Gamma(k)} = \theta^2 \cdot \frac{(k+1) \cdot k \cdot \cancel{\Gamma(k)}}{\cancel{\Gamma(k)}} = (k^2+k)\theta^2$$

$Var(X) = E(X^2) - (E(X))^2 = k^2\theta^2 + k\theta^2 - k^2\theta^2 = k\theta^2 \qquad \therefore E(X) = k\theta, \ Var(X) = k\theta^2$

베타분포 $X \sim BETA(a,b)$

$$B(a,b) = \int_0^1 x^{a-1} \cdot (1-x)^{b-1}\, dx \qquad a>0,\ b>0$$

$$f(x;a,b) = \frac{1}{B(a,b)} x^{a-1}(1-x)^{b-1} \quad (0<x<1)$$

$a=1, b=1$ 인 경우에는 $U(0,1)$이 된다.

$$E(X^k) = \int_0^1 \frac{1}{B(a,b)} x^{a-1} \cdot (1-x)^{b-1} \cdot x^k\, dx = \frac{1}{B(a,b)} \int_0^1 x^{k+a-1}(1-x)^{b-1}\, dx \qquad k>-a \text{에 대하여,}$$

$$= \frac{B(a+k,b)}{B(a,b)} = \frac{\Gamma(a+k)\Gamma(b)}{\Gamma(a+k+b)} \cdot \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} = \frac{\Gamma(a+k)\Gamma(a+b)}{\Gamma(a+b+k)\Gamma(a)}$$

$$E(X) = \frac{\Gamma(a+1)\Gamma(a+b)}{\Gamma(a+b+1)\Gamma(a)} = \frac{a\cancel{\Gamma(a)}\cancel{\Gamma(a+b)}}{(a+b)\cancel{\Gamma(a+b)}\cancel{\Gamma(a)}} = \frac{a}{a+b}$$

$$E(X^2) = \frac{\Gamma(a+2)\Gamma(a+b)}{\Gamma(a+b+2)\Gamma(a)} = \frac{(a+1)\cdot a\cancel{\Gamma(a)}\cancel{\Gamma(a+b)}}{(a+b+1)(a+b)\cancel{\Gamma(a+b)}\cancel{\Gamma(a)}} = \frac{a(a+1)}{(a+b+1)(a+b)}$$

$$Var(X) = E(X^2) - (E(X))^2 = \frac{a(a+1)}{(a+b+1)(a+b)} - \frac{a \times a}{(a+b)(a+b)} = \frac{a}{(a+b)}\left(\frac{a+1}{a+b+1} - \frac{a}{a+b}\right) = \frac{a}{a+b} \cdot \frac{b}{(a+b+1)(a+b)}$$

$$= \frac{ab}{(a+b+1)(a+b)^2} \qquad \therefore E(X) = \frac{a}{a+b} \quad Var(X) = \frac{ab}{(a+b+1)(a+b)^2}$$

베타분포는 모수 a,b에 따라서 구간 $(0,1)$ 상에 여러가지 분포형태를 가진다.
적절한 변환을 사용하면 베타분포는 임의의 유한구간 내의 값을 갖는 변수에 대한 확률모형으로 활용가능하다.

정규분포의 적분.

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\cdot\frac{(x-\mu)^2}{\sigma^2}}\, dx = 1 \qquad \frac{x-\mu}{\sqrt{2}\sigma} = t \qquad \frac{1}{\sqrt{2}\sigma} dx = dt$$

$$= \int_{-\infty}^{\infty} \frac{1}{\cancel{\sqrt{2\pi}\sigma}} e^{-t^2} \cdot \cancel{\sqrt{2}\sigma}\, dt = \int_{-\infty}^{\infty} \frac{1}{\sqrt{\pi}} e^{-t^2}\, dt = 1 \text{ 이므로 } \int_{-\infty}^{\infty} e^{-t^2}\, dt = \sqrt{\pi} \text{ 이다.}$$

$$\int_{-\infty}^{\infty} e^{-t^2}\, dt = I \text{ 라고 하고 } I^2 = \int_{-\infty}^{\infty} e^{-x^2}\, dx \int_{-\infty}^{\infty} e^{-y^2}\, dy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-(x^2+y^2)}\, dx\, dy \quad \begin{matrix} x = r\cos\theta \\ y = r\sin\theta \end{matrix} \text{ 로 치환}$$

$$dxdy = \left|\frac{\partial(x,y)}{\partial(r,\theta)}\right| drd\theta = \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{vmatrix} drd\theta = \begin{vmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix} drd\theta = (r\cos^2\theta + r\sin^2\theta)drd\theta = rdrd\theta$$

$$= \int_0^{2\pi}\int_0^{\infty} e^{-r^2}\cdot rdrd\theta = \int_0^{2\pi}\int_0^{\infty} \frac{1}{2}e^{-t}\cdot 2rdrd\theta = \int_0^{2\pi}\int_0^{\infty}\frac{1}{2}e^{-t}dt\,d\theta = \int_0^{2\pi}\frac{1}{2}d\theta = \pi$$

$r^2 = t \quad 2rdr = t$ $\qquad \therefore I^2 = \pi \quad I = \sqrt{\pi}$

즉 $\int_{-\infty}^{\infty}\frac{1}{\sqrt{\pi}}e^{-t^2}dt = 1.$

$X \sim N(\mu, \sigma^2)$ mgf

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

지수부분 정리 : $-\frac{x^2}{2\sigma^2} + \frac{2x\mu}{2\sigma^2} - \frac{\mu^2}{2\sigma^2} + tx = -\frac{x^2}{2\sigma^2} + \left(t+\frac{\mu}{\sigma^2}\right)x - \frac{\mu^2}{2\sigma^2} = -\frac{1}{2\sigma^2}\left(x^2 - 2\sigma^2 x\left(t+\frac{\mu}{\sigma^2}\right)\right) - \frac{\mu^2}{2\sigma^2}$

$$= -\frac{1}{2\sigma^2}\left(x^2 - 2\sigma^2 x\left(t+\frac{\mu}{\sigma^2}\right) + \left(\sigma^2\left(t+\frac{\mu}{\sigma^2}\right)\right)^2\right) + \frac{1}{2\sigma^2}\left(\sigma^2\left(t+\frac{\mu}{\sigma^2}\right)\right)^2 - \frac{\mu^2}{2\sigma^2}$$

$$= -\frac{\left(x-\sigma^2\left(t+\frac{\mu}{\sigma^2}\right)\right)^2}{2\sigma^2} + \frac{\sigma^4\left(t+\frac{\mu}{\sigma^2}\right)^2}{2\sigma^2} - \frac{\mu^2}{2\sigma^2}$$

$$E(e^{tX}) = \frac{1}{\sqrt{2\pi}\sigma}\underbrace{\int_{-\infty}^{\infty} e^{-\frac{\left(x-\left(\sigma^2\left(t+\frac{\mu}{\sigma^2}\right)\right)\right)^2}{2\sigma^2}}}_{\sqrt{2\pi}\sigma} e^{\frac{\sigma^4\left(t+\frac{\mu}{\sigma^2}\right)^2}{2\sigma^2}} e^{-\frac{\mu^2}{2\sigma^2}}dx$$

$$= e^{\frac{\sigma^4\left(t+\frac{\mu}{\sigma^2}\right)^2 - \mu^2}{2\sigma^2}} = e^{\frac{\sigma^4\left(t^2+\frac{2t\mu}{\sigma^2}+\frac{\mu^2}{\sigma^4}\right)-\mu^2}{2\sigma^2}} = e^{\frac{\sigma^4t^2+2t\mu\sigma^2+\cancel{\mu^2}-\cancel{\mu^2}}{2\sigma^2}} = e^{\frac{\sigma^2t^2}{2}+\mu t}$$

$$\therefore E(e^{tX}) = e^{\frac{\sigma^2t^2}{2}+\mu t}$$

$Z \sim N(0,1)$ mgf

$$E(e^{tZ}) = \int_{-\infty}^{\infty} e^{tz}\cdot\frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}dz = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{z^2}{2}+tz}dz = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{1}{2}(z^2-2tz+t^2-t^2)}dz$$

$$= \underbrace{\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{1}{2}(z-t)^2+\frac{t^2}{2}}dz}_{N(t,1)} = e^{\frac{t^2}{2}} \qquad \therefore E(e^{tZ}) = e^{\frac{t^2}{2}}$$
