5. 연속확률변수와 분포

정규분포 1문제
중간고사에 나옴

담당교수 : 김 덕 기

toby123@cbnu.ac.kr

연속확률분포(continuous distribution)

- 연속확률변수(continuous random variable): 어느 구간에 속하는 모든 값을 가질 수 있는 확률변수
  - 예) 버스를 기다리는 시간, 제품의 수명 등

- 예) $X =$ 구간 $[0, 1]$에서 임의로 선택된 수
  - $P\left(X \le \frac{1}{2}\right) = \frac{1}{2}, P\left(X \le \frac{1}{3}\right) = \frac{1}{3}$

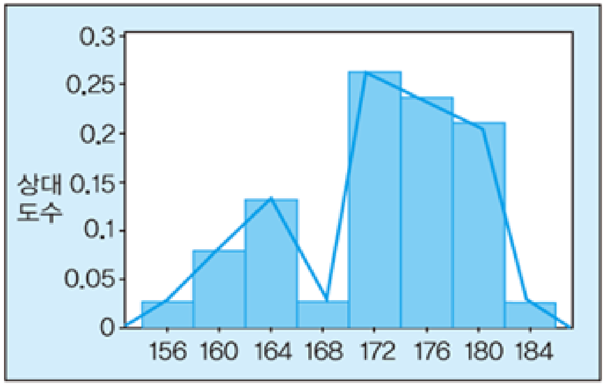
- 위 예에서 $P(X = 0.3)$ 일 확률은?
  - 이산확률변수 경우는 각 점(값)에서의 point mass (확률) 로 구함
  - 연속확률변수 경우는 어떻게 확률을 계산하는가?
    - 이산처럼 구간내의 가능한 실수 값이 셀 수 없이(uncountable) 많으므로 모든 값에서 point mass를 부여할 수 없다.
    - 가능한 방법은 임의의 실수 구간에 대한 확률을 부여하는 규칙이 필요. 위의 예처럼 구간 [0, 0.5]에 포함될 확률 $P(0 \le X \le 0.5)$을 구하는 방법

연속형 자료의 도수분포표와 확률분포

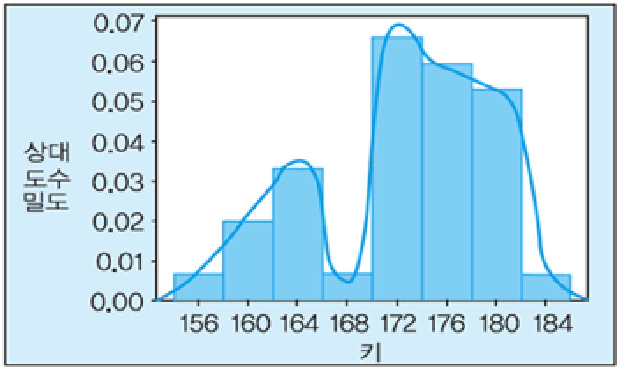
연속확률분포

학생들 키의 도수분포표

학생들의 키(cm) X	도수 f	상대도수 f/n	상대도수밀도 $f/(4n)$
154 이상 158 미만	1	0.026	0.00650
158 이상 162 미만	3	0.079	0.01975
162 이상 166 미만	5	0.132	0.03300
166 이상 170 미만	1	0.026	0.00650
170 이상 174 미만	10	0.263	0.06575
174 이상 178 미만	9	0.237	0.05925
178 이상 182 미만	8	0.211	0.05275
182 이상 186 미만	1	0.026	0.00650
합계	n=38	1.000	



[표 6.1]의 상대도수분포도와
대도수분포다각형



[표 6.1]의 상대도수밀도분포도와
확률분포곡선

연속확률변수의 확률 분포의 조건

➔ 연속확률변수 : 셀 수 없는 어떤 구간 사이에 나타나는 임의의 값들을 취하는 확률변수.

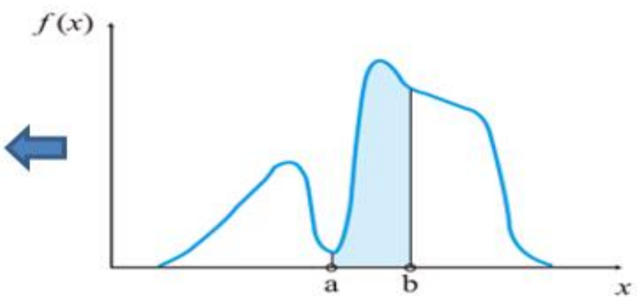
ex) 사람들의 체중 , 전화 통화 시간, 성인의 혈압

연속확률변수의 확률분포 $f(x)$가 만족해야 하는 조건

1. x의 모든 값에 대하여 $f(x)$는 0과 같거나 커야 한다. 즉, $f(x) \geq 0$.
2. $f(x)$ 곡선 아래의 총면적은 반드시 1이 되어야 한다.

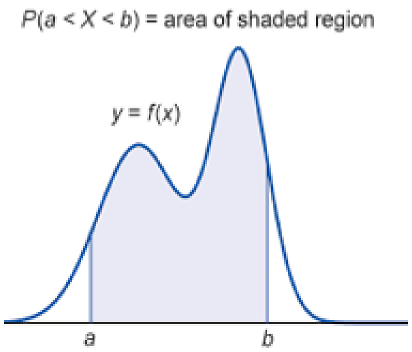
즉, $P(-\infty < X < \infty) = \int_{-\infty}^{\infty} f(x)\,dx = 1.$

$$P(a < X < b) = \int_a^b f(x)\,dx$$



$$P(a < X < b) = P(a \leq X < b) = P(a < X \leq b) = P(a \leq X \leq b)$$

확률밀도함수(pdf), 누적분포함수(cdf)



- 연속확률변수 X 의 확률밀도함수 : $f(x)$
- $P(a \le X \le b) = \int_a^b f(x)dx$: 면적
- $P(a \le X \le a) = P(X = a) = \int_a^a f(x)dx = 0$
- $P(a < X < b) = P(a < X \le b) = P(a \le X < b) = P(a \le X \le b)$
- 누적분포함수 : $F(x)$
  - $F(x) = P(X \le x) = \int_{-\infty}^x f(y)dy$
  - $F(b) - F(a) = P(X \le b) - P(X \le a) = \int_a^b f(x)dx$
  - $F'(x) = \frac{d}{dx}F(x) = f(x)$, 미분가능한 점 x

확률밀도함수와 누적분포함수의 관계

- $F(x) = \int_{-\infty}^{x} f(t)\, dt$ and $f(x) = \frac{d}{dx} F(x) = F'(x)$
- $\lim_{\Delta x \to 0} \frac{1}{\Delta x} P(x \le X \le x + \Delta x) = \lim_{\Delta x \to 0} \frac{\{F(x+\Delta x) - F(x)\}}{\Delta x} = F'(x) = f(x)$
- $P(x \le X \le x + \Delta x) \approx f(x)\, \Delta x$

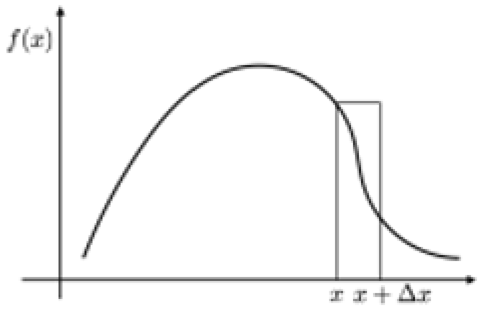
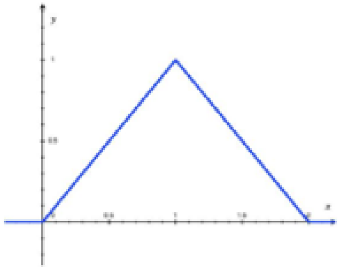


Figure: 구간 $[x, x + \Delta x]$에 속할 확률

누적분포함수 : 예제

[예제6] 연속확률변수 X 의 확률밀도함수가 다음과 같을 때

$$f(x) = \begin{cases} x & 0 \le x \le 1 \\ 2 - x & 1 < x \le 2 \\ 0 & otherwise \end{cases}$$



누적분포함수 $F(x)$ 를 구하라.

[풀이] $x < 0$ 이면 $f(x) = 0$ 이므로 $F(x) = \int_{-\infty}^{x} f(y)dy = 0$.

$0 \le x \le 1,\ F(x) = \int_{-\infty}^{x} f(y)dy = \int_{0}^{x} y dy = \frac{x^2}{2}$.

$1 < x \le 2,\ F(x) = \int_{-\infty}^{x} f(y)dy = \int_{0}^{1} y dy + \int_{1}^{x} (2 - y)dy = 2x - \frac{x^2}{2} - 1$

$x > 2,\ F(x) = \int_{-\infty}^{x} f(y)dy = \int_{0}^{2} f(y)dy + \int_{2}^{x} 0dy = 1.$

연속확률변수의 기댓값, 분산

연속확률변수의 기댓값

확률밀도함수 $f(x)$를 갖는 연속확률변수 X의 **기댓값(expected value)** 또는 **평균값(mean value)**은

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

으로 정의한다.

분산, 표준편차

X를 확률밀도함수 $f(x)$를 갖는 연속확률변수라고 하면 X의 **분산(variance)**은

$$\sigma_X^2 = \mathrm{Var}(X) = E[(X-\mu)^2] = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\, dx$$
$$= E(X^2) - [E(X)]^2$$

이 성립하고, X의 **표준편차(standard deviation)**는

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

로 정의된다.

또한 임의의 상수 a와 b에 대하여

$$E(aX+b) = aE(X) + b$$

이고

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

가 성립한다.

연속확률변수의 기댓값, 분산 : 예제

예제 다음 확률밀도함수를 갖는 연속확률변수를 X 라고 하자. X 의 분산 $\mathrm{Var}(X)$ 을 구하여라.

$$f(x) = \begin{cases} \frac{3}{2}\left(x - \frac{1}{2}x^2\right), & 0 < x < 2 \\ 0 \quad , & \text{기타} \end{cases}$$

풀이 X 의 분산을 구하기 위하여 X 의 기댓값 $E(X)$ 와 $E(X^2)$ 을 구하면

$$\begin{aligned} E(X) &= \int_{-\infty}^{\infty} x f(x)\,dx = \int_0^2 x\,\frac{3}{2}\left(x - \frac{1}{2}x^2\right)dx \\ &= \frac{3}{2}\int_0^2 \left(x^2 - \frac{1}{2}x^3\right)dx = 1 \end{aligned}$$

$$\begin{aligned} E(X^2) &= \int_{-\infty}^{\infty} x^2 f(x)\,dx = \int_0^2 x^2\,\frac{3}{2}\left(x - \frac{1}{2}x^2\right)dx \\ &= \frac{3}{2}\int_0^2 \left(x^3 - \frac{1}{2}x^4\right)dx = \frac{6}{5} \end{aligned}$$

이다. 이 결과를 이용하여 분산을 구하면

$$\begin{aligned} \mathrm{Var}(X) &= E(X^2) - [E(X)]^2 \\ &= \frac{6}{5} - 1^2 = \frac{1}{5} \end{aligned}$$

정규분포(Normal distribution) ☆

- C. Gauss 의 정규분포 (Normal distribution), bell distribution

C. Gauss(1777~1855) 는 각종의 물리학 실험을 행할 때 수반되는 계측오차에 대한 확률분포 로서 , 오늘날 가우스분포 (Gauss distribution) 라 알려져 있는 연속확률분포 제시 . 그 후로 이 분포는 물리학 뿐만 아니라 다른 모든 학문 분야에서도 널리 사용

▶ 정규분포: $\boldsymbol{X \sim N(\mu, \sigma^2)}$

X의 확률밀도 함수가

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

- $E(X)=\int_{-\infty}^{\infty} x\,f(x)dx=\mu$
- $Var(X)=E(X-\mu)^2=\int_{-\infty}^{\infty}(x-\mu)^2\,f(x)dx=\sigma^2$

이면, 연속확률변수 X가 모수 $\mu(-\infty<x<\infty)$와 $\sigma^2(0<\sigma^2<\infty)$를 갖는 정규분포라 한다.

— 정규분포의 성질

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

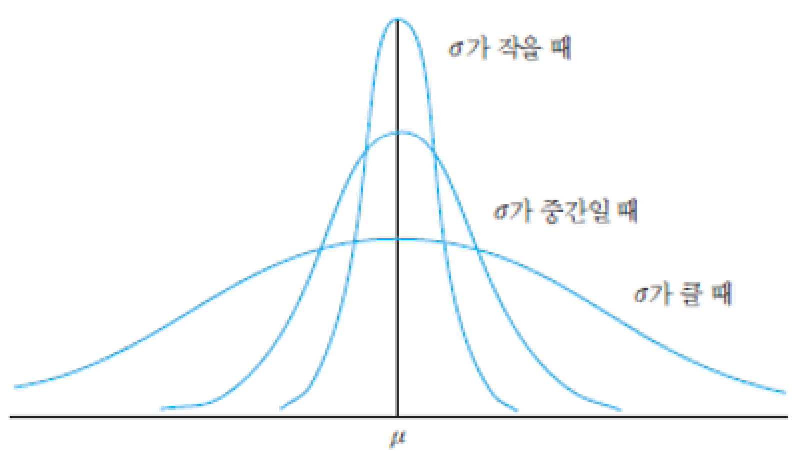
② $E[X]=\mu$

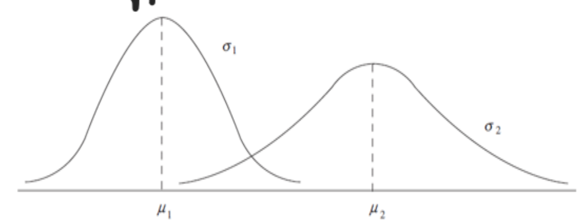
③ $Var(X)=\sigma^2$

정규분포의 특징1

- 정규분포의 특징
  1. 평균 μ를 중심으로 좌우 대칭이며, μ에서 확률밀도함수의 값이 가장 크다.
  2. 평균 μ에서 멀어질수록 확률밀도함수의 값은 점차 작아진다.
  3. 분산 σ^2이 클수록 확률밀도함수의 꼬리가 두꺼워진다 (μ에서 멀리 떨어져 있을 확률이 커진다.)

$\sigma_A^2 < \sigma_B^2$



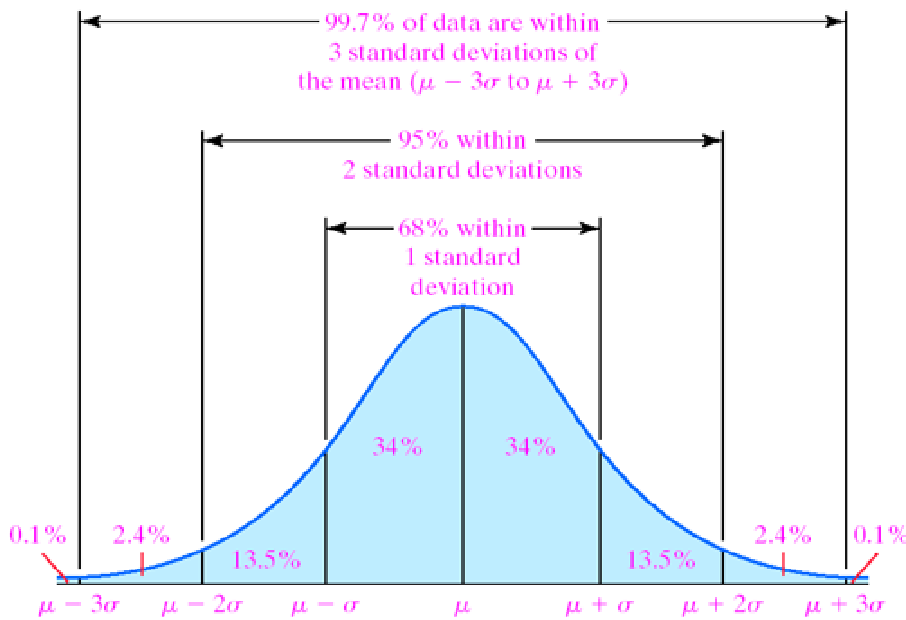


$\mu_A < \mu_B$

| 그림 4 | 평균은 같고 분산이 다른 정규분포의 확률밀도함수 비교

정규분포의 특징2

- 평균 μ를 중심으로 $\pm\sigma$ 안에 들어갈 확률이 0.6827이다.
- 평균 μ를 중심으로 $\pm 2\sigma$ 안에 들어갈 확률이 0.9545이다.
- 평균 μ를 중심으로 $\pm 3\sigma$ 안에 들어갈 확률이 0.9973이다.



$\boxed{6\sigma-\text{운동}}$

1표준편차

$$P(\mu-\sigma \le X \le \mu+\sigma) = 0.6827$$

$$P(\mu-2\sigma \le X \le \mu+2\sigma) = 0.9545$$

$$P(\mu-3\sigma \le X \le \mu+3\sigma) = 0.9973$$

대칭성 이용

Chebyshev 부등식

$$\boxed{P[\mu-k\sigma \le X \le \mu+k\sigma] \ge \left[1-\frac{1}{k^2}\right]\times 100\%}$$

- 정규분포의 확률밀도함수를 직접 적분하여 확률을 구하는 것은 쉽지 않다.

표준화(standardization)

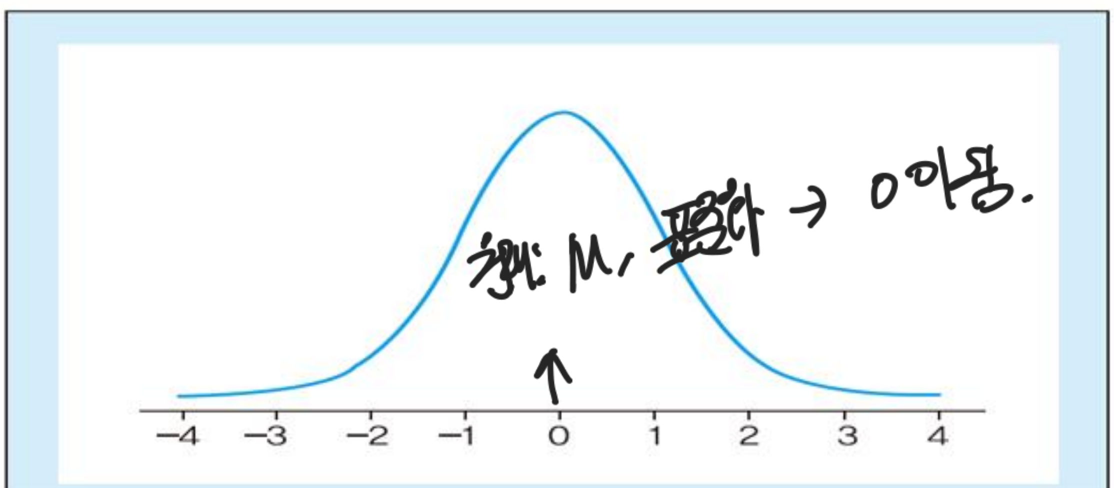
표준정규확률변수 Z와 정규확률변수 X와의 관련성

정규확률변수 X가 평균 μ와 표준편차 σ를 가질 때, 다음의 표준정규확률변수 Z는

$$Z = \frac{X-\mu}{\sigma} \qquad \sim N(\mu, \sigma^2) = N(0, 1)$$

로 평균 0, 분산 1인 정규분포를 갖는다. 그리고 이렇게 X에서 Z로 변환된 값을 z값이라고 하고, z값은 표준편차의 단위로 표시된 평균과의 거리를 의미한다. 이렇게 X에서 Z로 변환하는 과정을 표준화라고 한다.

단위 없애기



중심: μ, 표준편차 → 0이 됨.

표준정규분포

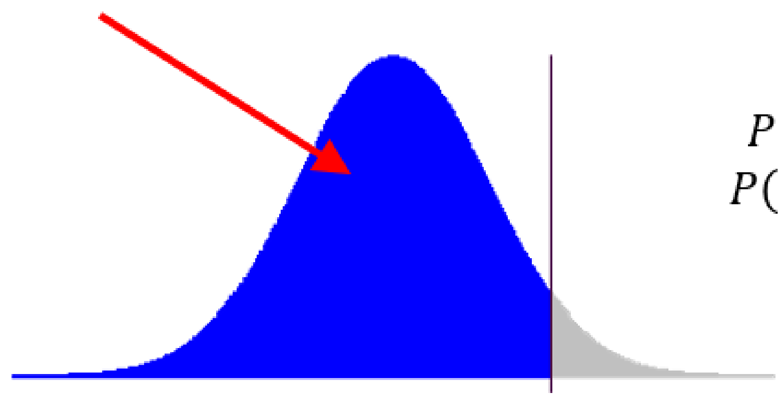
- 평균이 0이고 분산이 1인 정규분포 $N(0,1)$를 표준정규분포라고 한다.
- 표준정규분포를 따르는 확률변수를 흔히 Z로 나타낸다.
- 표준정규분포의 pdf : $\varphi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}, \quad -\infty < z < \infty$
- <부록-표3> 표준정규분포표: $Z \sim N(0,1)$일 때

cdf = $\Phi(z) = P(Z \le z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$ 의 값을 계산한 표

$P(Z \le z)$: 누적확률(cumulative distribution)

$P(Z_1 \le Z \le Z_2)$

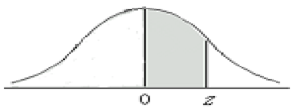
$= \Phi(Z_2) - \Phi(Z_1)$



$P(Z \le 1.98) = 0.9761$
$P(Z \le -0.03) = 0.4880$

표준 정규분포 표 Φ(z)

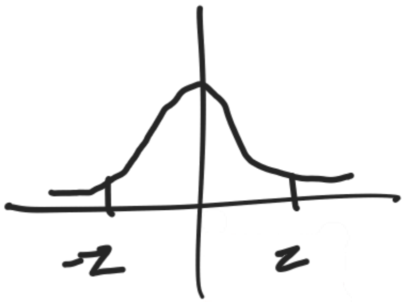
[표준정규분포표] $P(0 \le Z \le z) = \int_0^z \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt$



Z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0398	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0793	0.0871	0.091	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1179	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.148	0.1517
0.4	0.1554	0.1554	0.1628	0.1664	0.17	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1915	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.219	0.2224
0.6	0.2257	0.2257	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.258	0.258	0.2642	0.2673	0.2703	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2881	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3159	0.3212	0.3238	0.3264	0.3289	0.3315	0.334	0.3365	0.3389
1.0	0.3413	0.3413	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3643	0.3686	0.3708	0.3729	0.3749	0.377	0.379	0.381	0.383
1.2	0.3849	0.3849	0.3888	0.3907	0.3925	0.3944	0.3962	0.398	0.3997	0.4015
1.3	0.4032	0.4032	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4192	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4332	0.4357	0.437	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4452	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4554	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4641	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4713	0.4726	0.4732	0.4738	0.4744	0.475	0.4756	0.4761	0.4767
2.0	0.4772	0.4772	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4821	0.483	0.4834	0.4838	0.4842	0.4846	0.485	0.4854	0.4857
2.2	0.4861	0.4861	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.489
2.3	0.4893	0.4893	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4918	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4938	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4953	0.4956	0.4957	0.4959	0.496	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4965	0.4967	0.4968	0.4969	0.497	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4974	0.4976	0.4977	0.4977	0.4978	0.4979	0.4979	0.498	0.4981
2.9	0.4981	0.4981	0.4982	0.4983	0.4984	0.4984	0.4985	0.4985	0.4986	0.4986
3.0	0.4987	0.4987	0.4987	0.4988	0.4988	0.4989	0.4989	0.4989	0.499	0.499
3.1	0.499	0.499	0.4991	0.4991	0.4992	0.4992	0.4992	0.4992	0.4993	0.4993
3.2	0.4993	0.4993	0.4994	0.4994	0.4994	0.4994	0.4994	0.4995	0.4995	0.4995
3.3	0.4995	0.4995	0.4995	0.4996	0.4996	0.4996	0.4996	0.4996	0.4996	0.4997
3.4	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4998
3.5	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998

표준정규분포 특징- 예제

- Z의 확률밀도함수는 0에 대해 대칭이므로
  - $P(Z \le -z) = P(Z \ge z)$
  - $P(Z \le 0) = P(Z \ge 0) = \frac{1}{2}$



[예제10]

- $P(Z \le 1.37) = 0.9147$
- $P(Z \ge 1.37) = 1 - P(Z < 1.37) = 1 - 0.9147 = 0.0853$
- 또는 $P(Z \ge 1.37) = P(Z \le -1.37) = 0.0853$

[예제12] $P(-z \le Z \le z) = 0.90$인 z의 값은?

- $P(-z \le Z \le z) = 0.90$이 성립하기 위해서는 $P(Z \le -z) = 0.05$이어야 한다.
- $P(Z \le -1.65) = 0.0495, P(Z \le -1.64) = 0.0505$이므로 $P(Z \le -z) = 0.05$인 z는 근사적으로 1.645이다.

표준정규분포의 확률 계산 예제

예제 평균이 18이고 표준편차가 3인 정규분포를 따르는 확률변수 X에 대하여 다음을 구하여라.

(1) $P(X<15)$

(2) $P(12<X<21)$

(3) $P(X<k) = 0.3050$을 만족하는 k값

풀이 $X \sim N(18, 3^2)$이므로 $\mu = 18$, $\sigma = 3$이다.

$$(1)\ P(X<15) = P\left(Z<\frac{15-18}{3}\right) = P(Z<-1)$$
$$= \Phi(-1) = 0.1587$$

$$(2)\ P(12<X<21) = P\left(\frac{12-18}{3}<Z<\frac{21-18}{3}\right) = P(-2<Z<1)$$
$$= \Phi(1) - \Phi(-2) = 0.8413 - 0.0228 = 0.8185$$

(3) $P(X<k) = P\left(Z<\frac{k-18}{3}\right) = 0.3050$이므로 [부록 표 3]에서 확률값 0.3050값을 갖는 z값은 $z = -0.51$이다. 즉, $\frac{k-18}{3} = -0.51$이므로 $k = 16.47$이다.

중요

기말

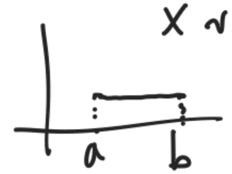
균등분포(Uniform distribution)

[정의] 연속확률변수 X 의 확률밀도함수가 다음과 같으면 (단 $b > a$)

$$f(x) = \begin{cases} \dfrac{1}{b-a}, & a \le x \le b \\ 0, & otherwise \end{cases}$$

균등분포를 따르며 $X \sim U[a,b]$ 로 표기한다.

$X \sim U[a, b]$



- 기댓값과 분산 $X \sim U[a,b]$
  - $E(X) = \int_a^b x \frac{1}{b-a} dx = \frac{a+b}{2}$
  - $Var(X) = E(X-\mu)^2 = \int_{-\infty}^{\infty} (x-\mu)^2 f(x) dx = \frac{(b-a)^2}{12}$

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

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

- $X \sim U[0,1]$ 이면 $E(X) = \frac{1}{2}, Var(X) = \frac{1}{12}$
  - 누적분포함수 $F(x) = \int_0^x 1\, dy = x,\ 0 \le x \le 1,\ F(x) = 0, x < 0, F(x) = 1, x > 1.$
  - $Y = aX + b$ (단, $b > a$) 의 평균, 분산을 구하면

    E(Y)=E(aX+b)=a/2+b, V(Y)=V(aX+b)=$a^2/12$

$E(Y) = \frac{a}{2} + b$

$V(Y) = \frac{a^2}{12}$

$Y = aX + b$

$f(x) = \lambda e^{-\lambda x}$

$X \sim Exp(\lambda)$

$\lambda \Rightarrow$ 평균 시간

지수분포(Exponentional distribution)

➔ 정규분포와 더불어 대표적인 연속확률분포이며, 대칭이 아닌 비대칭(skewed dist.)분포이며, 제품의 수명 분포로 자주 사용

[정의] 연속확률변수 X 의 확률밀도함수가 다음과 같으면 (단 $b > a$)

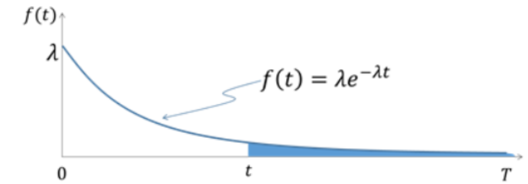
$$f(x) = \lambda e^{-\lambda x}, \qquad x \geq 0$$

지수분포를 따르며 $X \sim Exp(\lambda)$ 로 표기한다.

- 평균과 분산 $X \sim Exp(\lambda)$
  - $E(X) = \int_0^\infty x\, f(x)\, dx = \frac{1}{\lambda}$
  - $Var(X) = E(X-\mu)^2 = \int_{-\infty}^{\infty} (x-\mu)^2 f(x) dx = \frac{1}{\lambda^2}$

$E(X) = \frac{1}{\lambda}$

$V(X) = \frac{1}{\lambda^2}$



지수분포의 t시점 초과확률

- 누적분포함수 $F(x) = \int_0^x \lambda e^{-\lambda y} dy = -e^{-\lambda y}|_0^x = 1 - e^{-\lambda x},\ x \geq 0.$

지수분포 : 예제

[예제17] 신경세포의 나트륨 채널 오픈 시간이 평균 10(ms)인 지수분포를 따른다고 하면

- 채널 오픈 시간이 10ms 미만일 확률과
- 5ms와 15ms 사이에 오픈 될 확률과 5ms 과 15ms 사이에 오픈 될 확률을 구하라.

지수분포에서 평균의 역수 = λ.

[풀이] 채널 오픈 시간을 확률변수 X 라면 $X \sim Exp(\frac{1}{10})$ 즉 $\lambda = 1/10$.

$$P(X < 10) = P(X \leq 10) = F(10) = 1 - e^{-\left(\frac{1}{10}\right)10} = 1 - e^{-1} = 0.6321$$

$$P(5 < X < 15) = F(15) - F(5) = e^{-5\left(\frac{1}{10}\right)} - e^{-15\left(\frac{1}{10}\right)} = 0.3834.$$

$f(x) = \lambda e^{-\lambda x}$ $P(X) = \frac{e^{-\lambda}\lambda^{x}}{x!}$ ☆

지수분포 vs 포아송분포의 관계

[지수분포와 포아송분포]

[예제18] 사건이 시간당 평균 $\lambda = 3$ 회 발생할 때 확률변수 N_t를 t시간 동안 발생하는 사건의 수 라 하면 $N_t \sim Poisson(\lambda t)$ 가 된다. 여기서 확률변수 X를 첫 사건까지 기다리는 시간이라 하면

$X \sim Exp(\lambda)$

$$\{X > t\} = \{N_t = 0\}$$ ⇒ t번째 사건까지는 요랑 X

이 된다. 즉 첫 사건 발생까지 기다리는 시간이 t 보다 큰 사건은 t 시간동안 사건이 발생하지 않은 사건과 동일하게 된다. 따라서

$$P(X > t) = P(N_t = 0) = \frac{e^{-\lambda t}(\lambda t)^0}{0!} = e^{-\lambda t}$$

가 된다. 즉 $P(X \le t) = 1 - P(X > t) = 1 - e^{-\lambda t}$ 로 지수분포의 cdf.

누적 분포 함수

P.184 ~ 예제 18 참고

무기억성

지수분포의 특징(memoryless property)

$X \sim Exp(\lambda)$ 일 때 임의의 양의 실수 $t, s > 0$ 에 대하여

$$P(X > t + s | X > t) = P(X > s)$$

가 항상 성립한다.

[해설]

X 를 제품의 수명 또는 고장이 발생할 때까지의 대기시간이라 하면 위 등식은 처음부터 t 시간 까지 사건이 발생하지 않았을 때 이후 추가로 s 시간 동안에도 고장이 발생하지 않을 확률이 새 제품이 처음부터 s시간동안 고장이 없을 확률과 같음을 의미한다. 즉 t 시간의 작동기억이 상실된 것과 같음을 의미한다.

지수분포의 누적분포함수를 이용하면 $P(X > s) = e^{-\lambda s}$가 되고 따라서

$$P(X > t + s | X > t) = \frac{P(X > s + t)}{P(X > t)} = \frac{e^{-\lambda(s+t)}}{e^{-\lambda t}} = e^{-\lambda s} = P(X > s)$$