AI 기초

2019-2020

강봉주

특별한 분포

특별한 분포

[베르누이 분포]

- 베르누이(Bernoulli) 시행: 시행의 결과가 2가지 중 하나만 나오는 경우. 보통 1 또는 0만 나오는 경우
- 베르누이 분포: 베르누이 시행의 결과 분포

$$P(X=1)=p, P(X=0)=1-p$$

특별한 분포

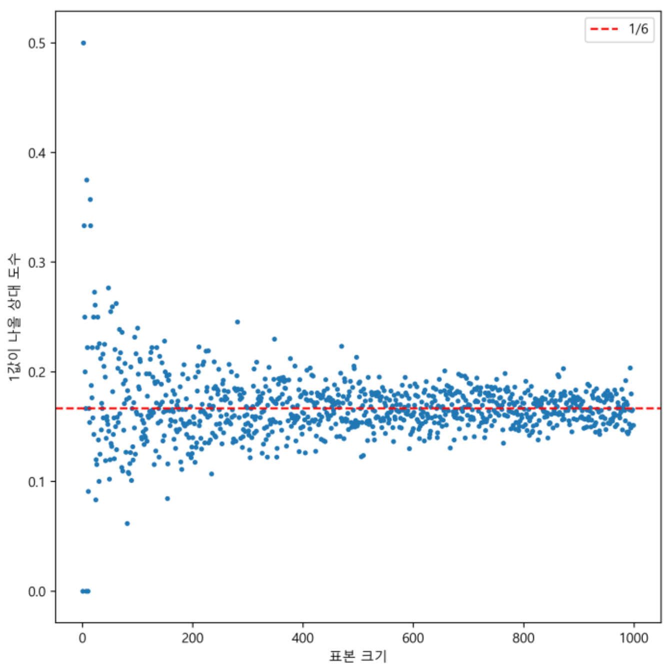
[베르누이 분포]

예제) $X \sim Brn(p)$ 일 때 평균과 분산을 구하세요.

특별한 분포

[베르누이 분포]

- $X \sim Brn(p)$ 인 분포에서 임의 표본의 추출
- *p=1/2* 인 경우에 베르누이 시행과 대수의 법칙



특별한 분포

[베르누이 분포]

- $X \sim Brn(p)$ 인 분포에서 임의 표본의 추출
- *p=1/2* 인 경우에 베르누이 시행과 대수의 법칙

In

```
# 표본의 개수: 10
trials = 1
event_prob = 1/2
size = 10
np.random.seed(123)
result = np.random.binomial(n=trials,p=event_prob, size
=size)
result
```

Out

```
array([1, 0, 0, 1, 1, 0, 1, 1, 0, 0])
```

특별한 분포

[베르누이 분포]

예제) $Y_1, \dots, Y_n \sim Brn(p)$ 이고 서로 독립일 때 결합 확률 밀도 함수는?

특별한 분포

[베르누이 분포]

예제) $Y_1, \dots, Y_n \sim Brn(p)$ 이고 서로 독립일 때 결합 확률 밀도 함수는?

$$f(y, \dots, y_n) = f_1(y_1) \cdots f_n(y_n) = \prod_i f_i(y_i) = \prod_i p^{y_i} (1-p)^{1-y_i}$$

특별한 분포

[이항 분포(binomial distribution)]

- $Y = \sum_{i=1}^{n} X_i \,, X_i \overset{iid}{\sim} Brn(p), \mathrm{n} = \#\ \text{of trials}$
- $Y \sim B(n, p)$: 성공 건수

In	# 이항 분포의 정의
	n = 10
	p = event_prob = 1/2
	size = sample_size = 3
	
	np.random.seed(123)
	np.random.binomial(n, p, size)

Out	array([6, 4, 4])

특별한 분포

[이항 분포(binomial distribution)]

- 확률 밀도 함수

$$\Pr(Y = y) = f(y) = \binom{n}{y} p^y (1-p)^{n-y}, \qquad y = 0, \dots, n$$

- 이항 계수(binomial coefficient)

$$\binom{n}{y} = \frac{\overbrace{n(n-1)\cdots(n-y+1)}^{\text{개수가 y개}}}{y(y-1)\cdots 1} = \frac{n!}{(n-y)!\,y!} = \binom{n}{n-y}$$

특별한 분포

[**이항 분포(binomial distribution)**]

- 이항 계수(binomial coefficient)

$$\binom{5}{2} = \frac{5 \cdot 4}{2 \cdot 1} = \frac{5 \cdot 4 \cdot (3 \cdot 2 \cdot 1)}{(3 \cdot 2 \cdot 1) \cdot 2 \cdot 1} = \frac{5!}{3!\,2!} = \binom{5}{3}$$

특별한 분포

[**이항 분포(binomial distribution)**]

예제) 이항계수 $\binom{5}{2}$, $\binom{5}{3}$의 값을 구하세요.

특별한 분포

[이항 분포(binomial distribution)]

예제) 이항계수 $\binom{5}{2}, \binom{5}{3}$의 값을 구하세요.

In	`# 예제의 해답`
	`A = [2, 3]`
	`trials = 5`
	`scipy.special.comb(trials, A)`

Out	`array([10., 10.])`

특별한 분포

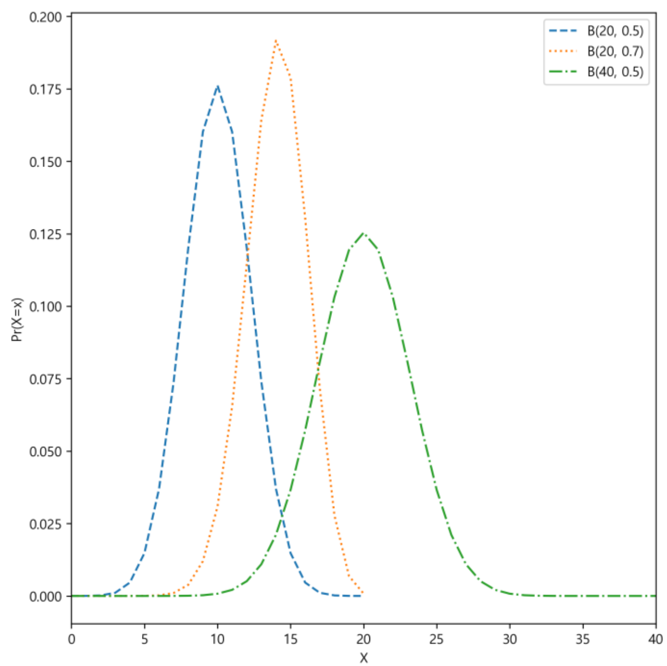
[**이항 분포(binomial distribution)**]

예제) $B(20, 0.5), B(20, 0.7), B(40, 0.5)$ 을 따르는 확률변수 각각에 대하여 확률 밀도 함수 그래프를 그리세요.

특별한 분포

[이항 분포(binomial distribution)]

예제) $B(20, 0.5), B(20, 0.7), B(40, 0.5)$ 을 따르는 확률변수 각각에 대하여 확률 밀도 함수 그래프를 그리세요.



특별한 분포

[이항 분포(binomial distribution)]

예제) 적률 생성함수를 이용하여 이항 분포의 평균과 분산을 구하세요.

$$\mu = E(Y) = \sum_{y} y \binom{n}{y} p^{y}(1-p)^{n-y} = ?$$

$$E(Y^2) = \sum_{y} y^2 \binom{n}{y} p^{y}(1-p)^{n-y} = ?$$

$$\sigma^2(Y) = E(Y^2) - \mu^2$$

$$M(t) = \sum\nolimits_{x} e^{tx} f(x) = \sum\nolimits_{x} \binom{n}{x} (pe^{t})^{x}(1-p)^{n-x} = [(1-p)+pe^{t}]^{n}$$

특별한 분포

[이항 분포(binomial distribution)]

예제) 적률 생성함수를 이용하여 이항 분포의 평균과 분산을 구하세요.

In	<pre># 적률 생성 함수를 이용한 이항 분포의 평균과 분산
	p, t, n = sympy.symbols('p, t, n')
	expr = ((1-p)+p*sympy.exp(t))**n
	
	# 1차 적률(기대값)
	M1 = sympy.Lambda(t, expr.diff(t).simplify())
	EX = M1(0)
	EX</pre>

Out	np

특별한 분포

[이항 분포(binomial distribution)]

예제) 적률 생성함수를 이용하여 이항 분포의 평균과 분산을 구하세요.

In	`# 2차 적률`
	`M2 = sympy.Lambda(t, expr.diff(t, t).simplify())`
	`EXX = M2(0)`
	`EXX`

Out	$np(np - p + 1)$

In	`# 분산` `(EXX- EX**2).simplify()`
Out	$np(1 - p)$

특별한 분포

[**정규 분포(normal distribution)**]

- 가우스 분포
- 천체의 거리 등에 대한 오차의 확률분포
- 갈톤(Galton), 피어슨(Pearson)
- 중심극한정리(central limit theorem): **표본 평균의 분포**는 데이터의 크기가 크다면 정규분포를 따름

Johann Carl Friedrich Gauss (1777–1855)

특별한 분포

[**정규 분포(normal distribution)**]

- $X \sim N(\mu, \sigma^2)$
- 확률 밀도 함수
- $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, -\infty < x < \infty$

- $Z \sim N(0, 1)$
- 표준 정규분포
- $\varphi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$

특별한 분포

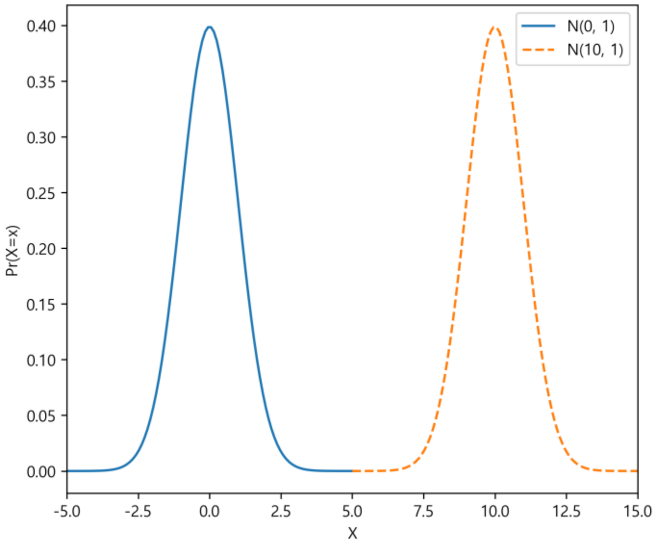
[**정규 분포(normal distribution)**]

예제) $N(0,1), N(10,1)$ 각각의 분포를 그리고 비교해보세요.

특별한 분포

[정규 분포(normal distribution)]

예제) $N(0,1), N(10,1)$ 각각의 분포를 그리고 비교해보세요.



특별한 분포

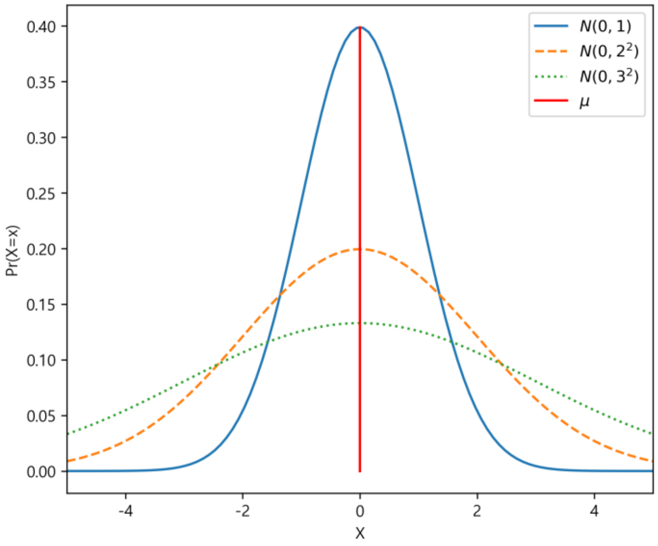
[**정규 분포(normal distribution)**]

예제) $N(0,1), N(0,2^2), N(0,3^2)$각각의 분포를 그리고 비교해보세요.

특별한 분포

[정규 분포(normal distribution)]

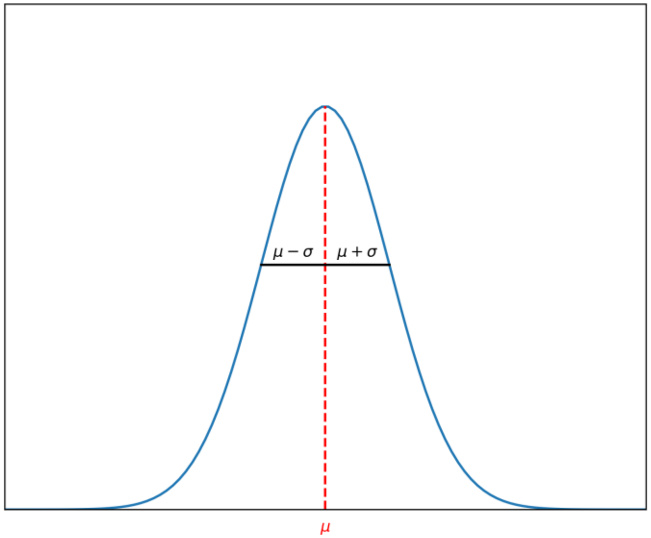
예제) $N(0,1), N(0,2^2), N(0,3^2)$각각의 분포를 그리고 비교해보세요.



특별한 분포

[**정규 분포(normal distribution)**]

- 표준편차의 의미
- 종 모양의 그래프에서의 변곡점



특별한 분포

[정규 분포(normal distribution)]

- 정규분포의 성질
- $X \sim N(\mu, \sigma^2)$

임의의 상수 a, b에 대하여 $aX + b \sim N(a\mu + b, a^2\ \sigma^2)$

$$\frac{X - \mu}{\sigma} \sim ?$$

특별한 분포

[정규 분포(normal distribution)]

- 정규분포의 성질
- $X \sim N(\mu, \sigma^2)$

임의의 상수 a, b에 대하여 $aX + b \sim N(a\mu + b, a^2\, \sigma^2)$

$$\frac{X - \mu}{\sigma} \sim N(0,1)$$

특별한 분포

[**정규 분포(normal distribution)**]

- 정규분포의 성질
- $Z \sim N(0, 1)$

$$\sigma Z + \mu \sim ?$$

특별한 분포

[**정규 분포(normal distribution)**]

- 정규분포의 성질
- $Z \sim N(0, 1)$

$$\sigma Z + \mu \sim N(\mu, \sigma^2)$$

특별한 분포

[정규 분포(normal distribution)]

- 정규분포의 성질
- $X \sim N(\mu, \sigma^2)$

$$\Pr(a \leq X \leq b) =?$$

특별한 분포

[**정규 분포(normal distribution)**]

- 정규분포의 성질
- $X \sim N(\mu, \sigma^2)$
- $\Phi(\cdot)$: 표준 정규분포의 분포함수

$$
\begin{aligned}
\Pr(a \le X \le b) &= \Pr\left(\frac{a-\mu}{\sigma} \le Z \le \frac{b-\mu}{\sigma}\right) \\
&= \Pr\left(Z \le \frac{b-\mu}{\sigma}\right) - \Pr\left(Z \le \frac{a-\mu}{\sigma}\right) \\
&= \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)
\end{aligned}
$$

특별한 분포

[**정규 분포(normal distribution)**]

예제) $\Phi(0), \Phi(1), \Phi(-1)$ 값을 구하세요

특별한 분포

[정규 분포(normal distribution)]

예제) $\Phi(0), \Phi(1), \Phi(-1)$ 값을 구하세요

In	`# 분포함수의 계산`
	`X = ss.norm()`
	`np.array([X.cdf(0), X.cdf(1), X.cdf(-1)]).round(3)`

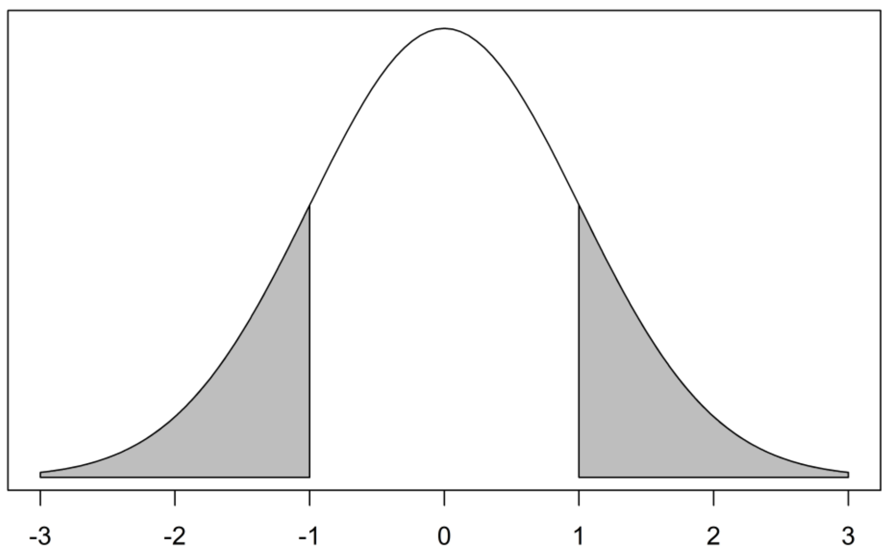
Out	$array([0.5\ \ , 0.841, 0.159])$

특별한 분포

[**정규 분포(normal distribution)**]

- 표준 정규분포의 성질
- $\Phi(\cdot)$: 표준 정규분포의 분포함수

$\Phi(-1)=1-\Phi(1)$



특별한 분포

[**정규 분포(normal distribution)**]

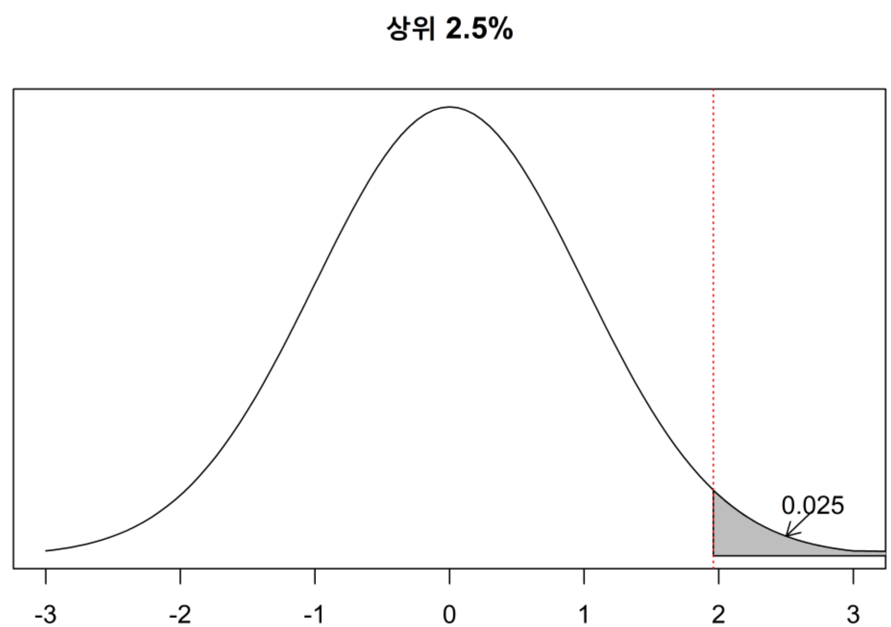
- 표준 정규분포의 성질
- $\Phi(\cdot)$: 표준 정규분포의 분포함수

$$\Pr(-1 \le Z \le 1) = \Phi(1) - \Phi(-1) = 1 - 2\Phi(-1) = 0.683$$
$$\Pr(-2 \le Z \le 2) = \Phi(2) - \Phi(-2) = 1 - 2\Phi(-2) = 0.954$$
$$\Pr(-3 \le Z \le 3) = \Phi(3) - \Phi(-3) = 1 - 2\Phi(-3) = 0.997$$

특별한 분포

[**정규 분포(normal distribution)**]

예제) 표준 정규분포에서 상위 2.5%에 해당하는 분위수는?



특별한 분포

[정규 분포(normal distribution)]

예제) 표준 정규분포에서 상위 2.5%에 해당하는 분위수는?

In	# 특정 분위수 계산
	X = ss.norm()
	alpha = 0.025
	X.ppf(1-alpha).round(3)

Out	1.96

특별한 분포

[**정규 분포(normal distribution)**]

[과제]

- 정규 분포의 적률생성함수를 이용하여 평균과 분산을 구하세요.
- 적률생성함수: $\mathrm{M(t)} = e^{\mu t+\frac{\sigma^2 t^2}{2}}$