통계 분석

2019년 2학기

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]

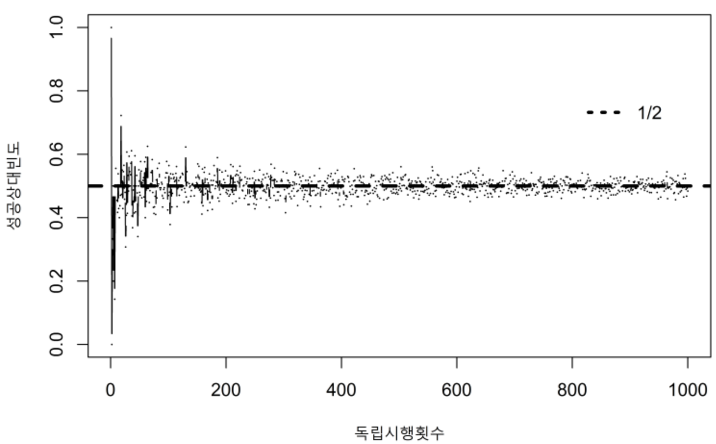
예제) $X \sim Ber(p)$ 일 때 평균과 분산을 구하세요.

특별한 분포

[베르누이 분포]

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

```
# 시행 횟수:1, 성공확률:1/2, 표본크기: 1000
# 결과 값: 성공 건수
rbinom(1000, 1, 1/2)
```



특별한 분포

[베르누이 분포]

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

특별한 분포

[베르누이 분포]

예제) $Y_1, \dots, Y_n \sim Ber(p)$ 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)]

- $Y = \sum_{i=1}^{size} X_i, X_i \overset{\text{서로 독립}}{\sim} Ber(p), \text{size} = \#\text{ of trials}$
- $Y \sim B(size, p)$: 성공 건수

```
> rbinom(10, 4, 1/2)
 [1] 3 2 3 2 3 2 0 3 4 3
```

size가 4이면 성공 건수는 0, 1, 2, 3, 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) factorial 함수를 이용하여 이항 계수 값을 구하는 함수를 작성하고 검증하세요.

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.

dbinom() 함수 이용

특별한 분포

[**이항 분포(binomial distribution)**]

예제) 이항 분포의 평균과 분산을 구하세요.

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

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

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

특별한 분포

[이항 분포(binomial distribution)]

예제) 이항 분포의 평균과 분산을 구하세요.

$$E(Y) = \sum_{y=0}^{n} y \binom{n}{y} p^{y} (1-p)^{n-y}$$

$$= \sum_{y=0}^{n} y \frac{n!}{(n-y)!\,y!} p^{y} (1-p)^{n-y}$$

$$= np \sum_{y=1}^{n} \frac{(n-1)!}{(n-y)!\,(y-1)!} p^{y-1} (1-p)^{n-y}$$ let x=y-1

$$= np \sum_{x=0}^{n-1} \frac{(n-1)!}{(n-1-x)!\,x!} p^{x} (1-p)^{n-1-x}$$

$$= np\big(p + (1-p)\big)^{n-1} = np$$

특별한 분포

[이항 분포(binomial distribution)]

예제) 이항 분포의 평균과 분산을 구하세요.

$$E(Y^2) = \sum_{y=0}^{n} y^2 \binom{n}{y} p^y (1-p)^{n-y}$$

$$= \sum_{y=0}^{n} y^2 \frac{n!}{(n-y)!\,y!} p^y (1-p)^{n-y}$$

$$= np \sum_{y=1}^{n} y \frac{(n-1)!}{(n-y)!\,(y-1)!} p^{y-1} (1-p)^{n-y} \quad \text{let x=y-1}$$

$$= np \sum_{x=0}^{n-1} (x+1) \frac{(n-1)!}{(n-1-x)!\,x!} p^x (1-p)^{n-1-x}$$

$$= np\big((n-1)p+1\big) = np(np+1-p)$$

특별한 분포

[**이항 분포(binomial distribution)**]

예제) 이항 분포의 평균과 분산을 구하세요.

$$\sigma^2(Y) = E(Y^2) - \mu^2$$
$$= np(np + 1 - p) - (np)^2$$
$$= np(1 - p) = npq$$

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)**]

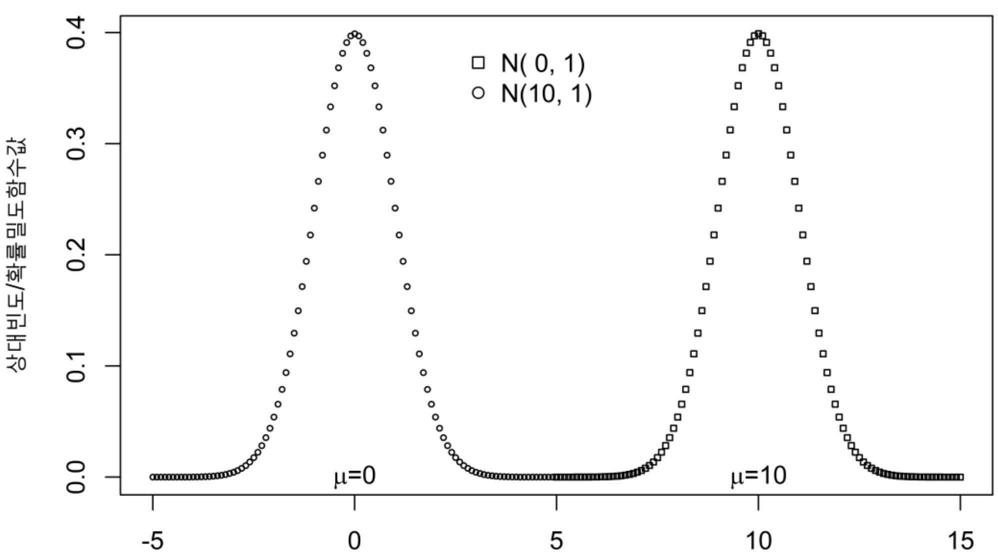
예제) $N(0,1), N(10,1)$ 각각의 분포를 그리고 비교해보세요.

dnorm() 함수를 이용

특별한 분포

[정규 분포(normal distribution)]

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



특별한 분포

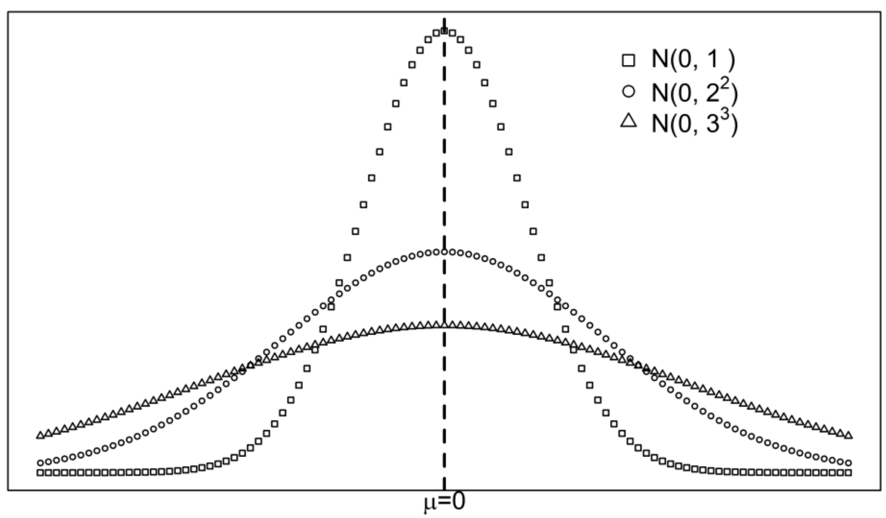
[**정규 분포(normal distribution)**]

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

특별한 분포

[**정규 분포(normal distribution)**]

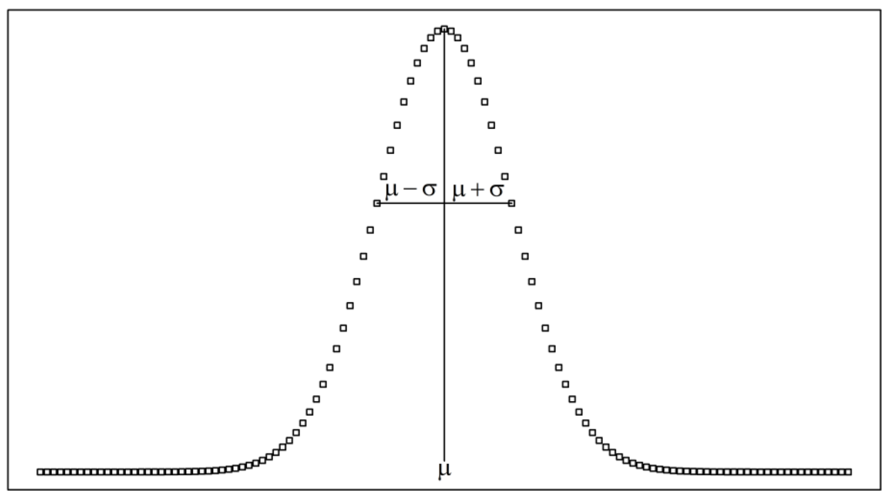
예제) $N(0,1), N(0,2^2), N(0,3^2)$각각의 분포를 그리고 비교해보세요.



특별한 분포

[**정규 분포(normal distribution)**]

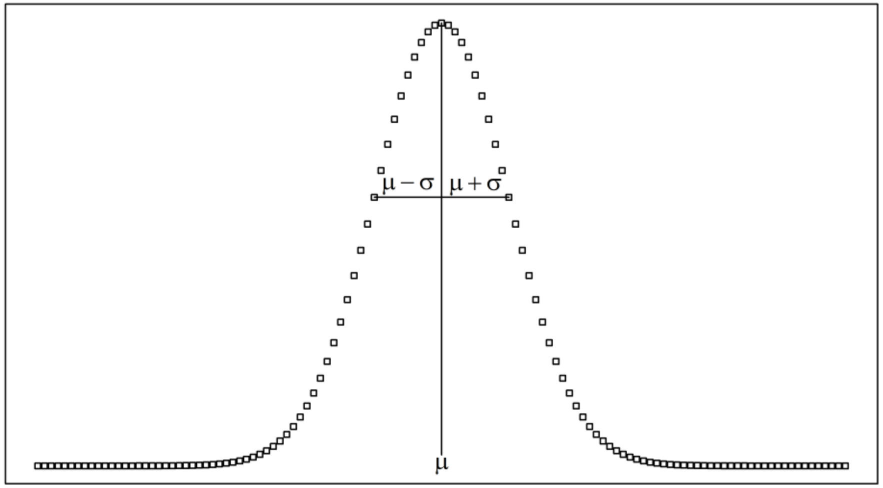
- 표준편차의 의미
- 종 모양의 그래프에서의 변곡점



특별한 분포

[정규 분포(normal distribution)]

과제1) 아래 그림을 생성하는 R 프로그램을 작성하세요.



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

$$\Pr(a \le X \le b) = \Pr\left(\frac{a-\mu}{\sigma} \le Z \le \frac{b-\mu}{\sigma}\right)$$

$$= \Pr\left(Z \le \frac{b-\mu}{\sigma}\right) - \Pr\left(Z \le \frac{a-\mu}{\sigma}\right)$$

$$= \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$$

특별한 분포

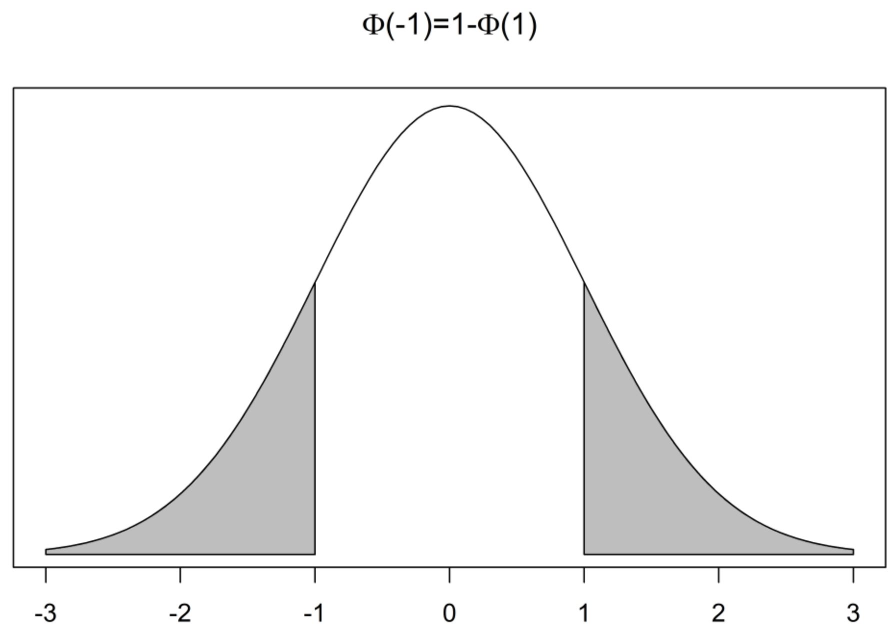
[**정규 분포(normal distribution)**]

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

특별한 분포

[**정규 분포(normal distribution)**]

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



특별한 분포

[정규 분포(normal distribution)]

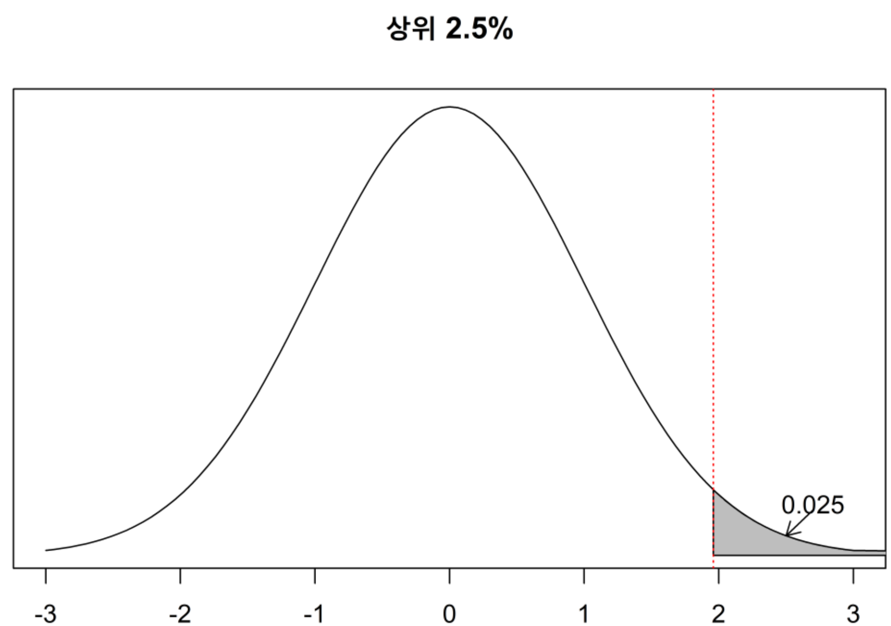
- 표준 정규분포의 성질
- $\Phi(\cdot)$: 표준 정규분포의 분포함수

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

특별한 분포

[정규 분포(normal distribution)]

예제) 표준 정규분포에서 상위 2.5%에 해당하는 분위수는?



특별한 분포

[정규 분포(normal distribution)]

과제2) 아래 그림을 생성하는 R 프로그램을 작성하세요.