02 강

통계적 추론

확률분포 1

한국방송통신대학교 통계·데이터과학과
이 긍희 교수

학습내용

1. 확률분포함수를 이해한다.
2. 기댓값과 분산을 이해한다.
3. 마코프 부등식을 이해한다.
4. 적률생성함수를 이해한다.

01

확률분포함수

1 확률변수

이산형 확률변수

- **확률변수의 가능한 값이 유한 개이거나 셀 수 있는 경우**

연속형 확률변수

- **확률변수의 가능한 값이 무한 개이며 셀 수 없는 경우**

1 확률변수

이산형 확률변수

- **확률질량 함수 :** $f(x) = P(X = x)$

확률질량함수의 성질

- $0 \leq f(x) \leq 1$
- $\sum_{x=1}^{n} f(x) = 1$
- $P(a \leq X \leq b) = \sum_{a}^{b} f(x)$

1 확률변수

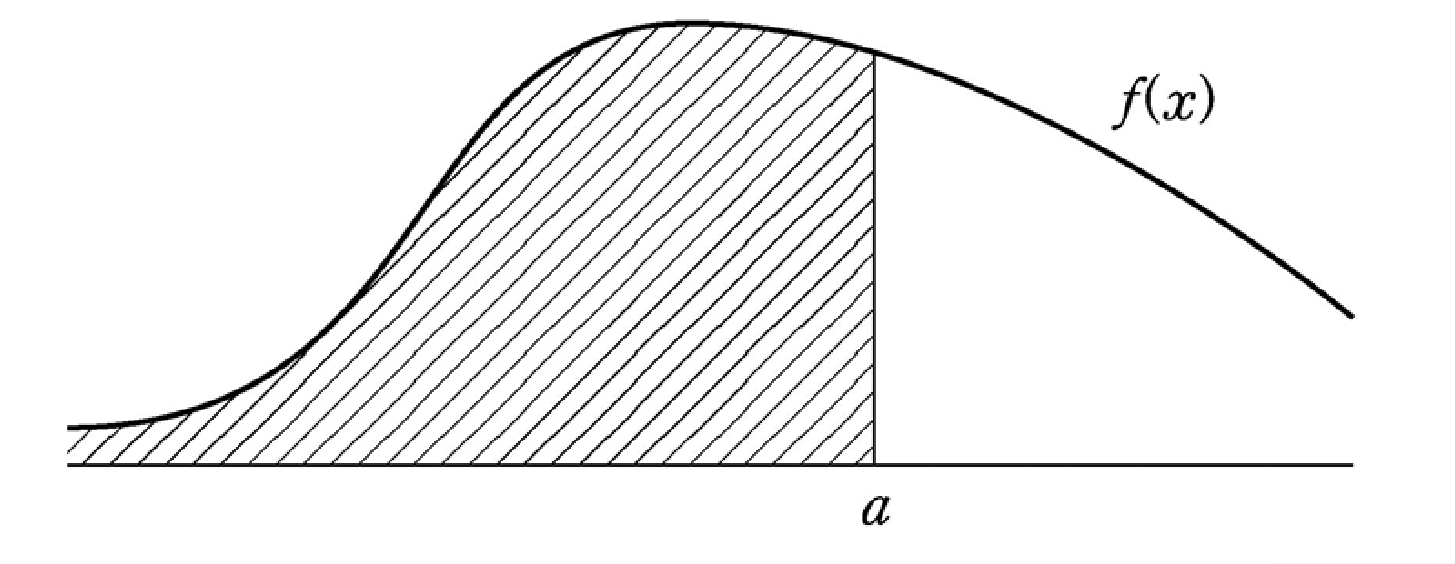
예 불량품 3개, 양품 2개인 상자에서 2개의 제품을 꺼낼 때 불량품 수의 확률질량함수는?

1 확률변수

연속형 확률변수

- **연속형 확률변수의 분포를 결정하는 함수 :**

$$F(x) = P(X \le x) = \int_{-\infty}^{x} f(t)\,dt, \qquad f(x) = \frac{d}{dx}F(x)$$



1 확률변수

예 **10분 간격으로 출발하는 버스를 기다리는 시간의 확률밀도함수는?**

1 확률변수

확률밀도함수의 성질

- $0 \leq f(x)$
- $\int_{-\infty}^{\infty} f(x)dx = 1$
- $P(a \leq X \leq b) = \int_{a}^{b} f(x)dx$

1 확률변수

예 다음 확률밀도함수에서 C와 $P(1 < X < 2)$를 구하시오.

$$f(x) = \begin{cases} Ce^{-x} & , x \geq 0 \\ 0 & , x < 0 \end{cases}$$

1 확률변수

예 다음 확률밀도함수에서 C와 $P(1 < X < 2)$를 구하시오.

02

기댓값

1 확률분포

- 확률변수 모든 값에 대응하는 확률을 나타내는 분포
  - **이산형 확률변수 : 확률질량함수**
  - **연속형 확률변수 : 확률밀도함수**

2 기댓값

분포의 무게중심(균형점)

- **이산형 확률변수**: $E(X) = \sum_{x=0}^{\infty} xf(x)$
- **연속형 확률변수**: $E(X) = \int_{-\infty}^{\infty} xf(x)dx$

2 기댓값

예 동전을 2번 던져서 나온 앞면의 수의 기댓값은?

2 기댓값

예 **다음 확률밀도함수를 가지는 X의 기댓값은?**

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

2 기댓값

예 **다음 확률밀도함수를 가지는 X의 기댓값은?**

3 기댓값의 성질

기댓값의 성질

- $E(aX + b) = aE(X) + b$
- $E\big(g(X) + h(X)\big) = E\big(g(X)\big) + E\big(h(X)\big)$

03

분산과 표준편차

1 분산

- 분산(variance) : 확률변수의 흩어져 있는 정도

  - $$Var(X) = E\left(\left(X - E(X)\right)^2\right)$$
    $$= E(X^2) - (E(X))^2$$

1 분산

분산의 성질

- $Var(aX + b) = a^2 Var(X)$

2 표준편차

표준편차(standard deviation)

- $Sd(X) = \sqrt{Var(X)}$

3 분산의 계산

예 **동전을 2번 던져서 나온 앞면의 수의 분산과 표준편차는?**

3 분산의 계산

예 **다음 확률밀도함수를 가지는 X의 분산과 표준편차는?**

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

3 분산의 계산

예 다음 확률밀도함수를 가지는 X의 분산과 표준편차는?

04

마코프 부등식

1 체비세프 부등식

- 체비세프(Chebyshev) 부등식
  - **확률변수** $X \sim (\mu, \sigma^2), \sigma^2 < \infty,\ k$: **양의 상수**

$$\Rightarrow P(|X-\mu| \geq k) \leq \frac{\sigma^2}{k^2}$$

2 마코프 부등식

마코프(Markov) 부등식

- **확률변수** X, $E(|X|^r) < \infty,\ r, k > 0$

$$\Rightarrow \mathrm{P}(|X| \geq k) \leq \frac{E(|X|^r)}{k^r}$$

3 분산의 계산

예 **X 가 다음 확률밀도함수를 가질 때** $P(|X-\mu_X| \geq \frac{3}{2}\sigma_X)$ **의 상한 값은?**

$$f(x) = \frac{1}{2\sqrt{3}}, \qquad -\sqrt{3} < x < \sqrt{3}$$

3 분산의 계산

예 **X 가 다음 확률밀도함수를 가질 때** $P(|X-\mu_X| \geq \frac{3}{2}\sigma_X)$ **의 상한 값은?**

05

적률생성함수

1 적률생성함수의 정의

k차 적률(moment)

- $E(X^k) = \begin{cases} \sum_x x^k f(x), & \text{이산형} \\ \int_{-\infty}^{\infty} x^k f(x) dx, & \text{연속형} \end{cases}$

1 적률생성함수의 정의

- 적률생성함수(moment generating function, mgf)
  - **0을 포함한 열린구간의 t에 대해서** $E(e^{tX}) < \infty,$

$$M(t) = E(e^{tX}) = \begin{cases} \sum_x e^{tx} f(x), & \text{이산형} \\ \int_{-\infty}^{\infty} e^{tx} f(x) dx, & \text{연속형} \end{cases}$$

2 적률생성함수의 성질

- 적률생성함수가 존재하면 모든 적률이 존재
  - $\mathrm{M}(0) = 1$
  - $M'(0) = E(X)$
  - $M^{(k)}(0) = E(X^k)$

2 적률생성함수의 성질

- 두 확률변수 X, Y의 적률생성함수가 존재하고, 0을 포함한 열린구간에서 일치하며 두 확률변수 X, Y의 확률분포가 일치
  - $M_X(\mathrm{t}) = M_Y(\mathrm{t}), \forall t : -h < t < t\ (h > 0)$
    $\Leftrightarrow f_X(x) = f_Y(y)$

3 적률생성함수의 예

예 **확률변수 X의 확률분포가 다음과 같을 때 X의 적률생성함수와 기댓값은?**

	0	1	2	3
$f(x)$	1/6	1/3	1/3	1/6

3 적률생성함수의 예

예 **확률변수 X의 확률분포가 다음과 같을 때 X의 적률생성함수와 기댓값은?**

3 적률생성함수의 예

예 **확률변수 X의 확률분포가 다음과 같을 때 X의 적률생성함수와 기댓값은?**

$$f(x) = \begin{cases} e^{-x} & , x \geq 0 \\ 0 & , x < 0 \end{cases}$$

3 적률생성함수의 예

예 **확률변수 X의 확률분포가 다음과 같을 때 X의 적률생성함수와 기댓값은?**

정리하기

- 확률분포는 이산형 확률변수의 경우 확률질량함수로, 연속형 확률변수의 경우 확률밀도함수로 결정된다.
- 기댓값은 확률분포의 무게중심(균형점)이다.
- 분산으로 확률분포가 흩어져 있는 정도를 측정한다.
- 마코프 부등식은 확률분포의 상한을 제시한다.
- 적률생성함수가 존재하면 확률변수의 적률을 구할 수 있고, 확률분포와 대응한다.

03 강

다음시간 안내

확률분포 2