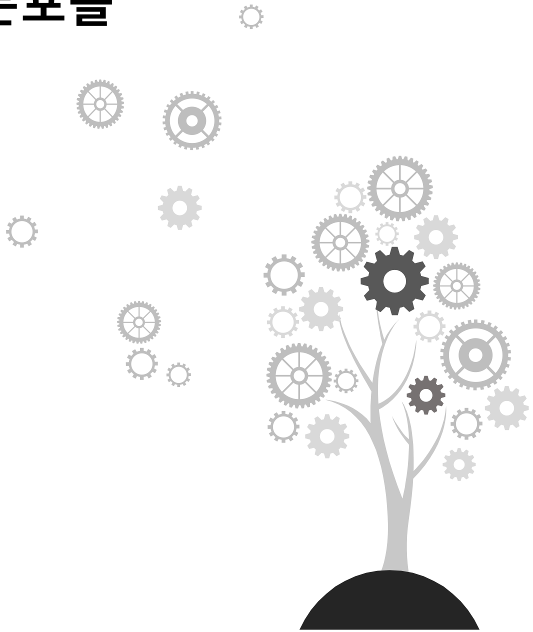
Contents



이항분포와 그에 관련된 분포들

01 베르누이 시행

- 베르누이 시행 (Bernoulli trial) : 성공과 실패의 두 가지 중에서 하나가 나타나는 실험

  ① 베르누이 시행의 각 시행은 성공과 실패의 두 결과만을 갖는다.
  ② 각 시행에서 성공할 확률은 p, 실패할 확률은 $1-p$로 일정하다.
  ③ 각 시행은 서로 독립으로 서로 영향을 미치지 않는다.

- 예 : 동전을 1번 던져서 앞면이 나오는 실험
  앞면이 나오면 성공(1), 뒷면이 나오면 실패(0)

$$\Omega = \{1, 0\}$$

- 예제 1, 예제 2

Chapter 7 이항분포와 그에 관련된 분포들

- 베르누이 분포 (Bernoulli distribution) :
  X : 성공의 확률이 p인 베르누이 시행의 결과

  $X \sim \text{Bernoulli}(p)$:
  확률변수 X는 모수가 p인 베르누이 분포를 따른다

- X의 확률분포

X	$P(X = x)$
0	$1 - p$
1	p
합계	1

$$P(X = x) = p^x \cdot (1-p)^{1-x}, \quad x = 0, 1$$

$$E(X) = p, \quad E(X^2) = p^2$$
$$Var(X) = p(1-p)$$

- 모수 p(베르누이 시행의 성공 확률)에 대해서 표본으로부터 추론

02 이항분포

$\langle$ m : 총시행 횟수
p : 성공 확률 $\rangle$

- 이항 분포 (Binomial distribution) :
  X : m번 베르누이 시행 중 성공의 수
  (각 베르누이 시행은 서로 독립이며 동일한 성공 확률 p를 가짐)

  $X \sim \mathrm{Binomial}(m, p)$ or $X \sim \mathrm{Bin}(m, p)$:
  확률변수 X는 모수가 (m, p)인 이항분포를 따른다

- X의 확률분포

$$P(X = x) = \binom{m}{x} p^x \cdot (1-p)^{m-x}, \quad x = 0, 1, \cdots, m$$

$$E(X) = mp, \quad Var(X) = mp(1-p)$$

- 모수 p(베르누이 시행의 성공 확률)에 대해서 표본을 이용하여 추론
- 모수 m(베르누이 시행의 시행 회수)는 처음부터 주어지는 정보

Chapter 7 이항분포와 그에 관련된 분포들

예) $X \sim \text{Bin}(m = 5, p = 0.7)$

$$P(X = 1) = \binom{5}{1} \times 0.7^{1} \times 0.3^{5-1}$$

- 이항분포는 서로 독립이고 같은 분포를 갖는 베르누이 분포들의 합

$$X = \sum_{i=1}^{m} X_i \sim \text{Bin}(m, p)$$

$$X_i \sim \text{Bernoulli}(p), \quad i = 1, \cdots, m, \quad E(X_i) = p, \quad Var(X_i) = p(1-p)$$

$$P(X_i = x_i) = \binom{1}{x_i} p^{x_i} \cdot (1-p)^{1-x_i} = p^{x_i} \cdot (1-p)^{1-x_i}, \quad x_i = 0, 1$$

$$P(X = x) = \binom{m}{x} p^{x} \cdot (1-p)^{m-x}, \quad x = 0, 1, \cdots, m$$

$$E(X) = mp, \quad Var(X) = mp(1-p)$$

이항분포와 그에 관련된 분포들

- 부록 표-1

$$P(X \le c) = \sum_{x=0}^{c} {}_{m}C_{x} \cdot p^{x}(1-p)^{m-x}, \quad {}_{m}C_{x} = \binom{m}{x}$$

$X \sim \mathrm{Bin}(m = 3, p = 0.3), \quad x = 0, 1, \cdots, m = 3$

		$\cdots$	$p = 0.3$	$\cdots$
$m = 3$	$c = 0$		$P(X \le 0)$	
	$c = 1$		$P(X \le 1)$	
	$c = 2$		$P(X \le 2)$	
	$c = 3$		$P(X \le 3)$	

Chapter 7 이항분포와 그에 관련된 분포들

- 이항분포에서의 확률계산

부록 표-1의 확률표를 이용하여 아래의 확률들을 계산 가능

$$P(a \le X \le b) = P(X \le b) - P(X \le a-1)$$
$$P(a < X < b) = P(X \le b-1) - P(X \le a)$$
$$P(a \le X < b) = P(X \le b-1) - P(X \le a-1)$$
$$P(a < X \le b) = P(X \le b) - P(X \le a)$$

$$P(X = a) = P(X \le a) - P(X \le a-1)$$
$$P(X \ge a) = 1 - P(X \le a-1)$$

확률 계산 시, 부등호와 등호에 따른 차이에 주의

(Bin
Hypergeometric
poisson)

03 초기하분포

- 초기하분포 (Hypergeometric distribution) :
  X : D개의 원소로 이루어진 그룹 1과 $N-D$개의 원소로 이루어진 그룹 2에서 비복원추출한 m개의 표본 내 그룹 1의 원소의 수

  $X \sim \text{Hypergeometric}(N, m, D)$:
  확률변수 X는 모수가 (N, m, D)인 초기하분포를 따른다

- X의 확률분포

$$P(X=x) = \frac{\binom{D}{x} \times \binom{N-D}{m-x}}{\binom{N}{m}}, \quad x = 0, 1, \cdots, m$$

N : 총 원소의 개수, D : 그룹 1의 원소의 수, $N-D$: 그룹 2의 원소의 수
m : 비복원추출한 표본의 개수, X : 표본 내 그룹 1의 원소의 수

이항분포와 그에 관련된 분포들

- 초기하분포의 평균과 분산, 초기하분포의 이항분포 근사

$$\text{Let } p = \frac{D}{N},$$

$$E(X) = m \cdot \frac{D}{N} = m \cdot p, \quad Var(X) = mp(1-p) \cdot \frac{N-m}{N-1} \xrightarrow{m < 0.05\,N} mp(1-p)$$

$$\therefore X \sim \text{Hypergeometric}(N, m, D) \xrightarrow{m < 0.05\,N} \text{Bin}\left(m, p = \frac{D}{N}\right)$$

N(총 원소의 개수)이 m(비복원추출한 표본의 개수) 보다 월등히 크다면,
초기하분포에서의 각각의 비복원추출을 복원추출(베르누이 시행)로 볼 수 있다. 그러면 X를 m번 베르누이 시행 중 성공의 수로 취급할 수 있다.

- 예제 5, 4.2, 4.6

'특정 단위시간에 몇 명의 손님이 들어올 것인가?
그리고 이에 대한 확률은 어떻게 되는가?'

04 포아송분포

- 포아송분포 (Poisson distribution) :
  X : 특정 시간(구간) 동안의 사건의 발생횟수 (p.216)

  $X \sim \text{Poisson}(\lambda)$:
  확률변수 X는 평균이 λ 인 포아송분포를 따른다
  특정 시간(구간) 동안 사건이 평균적으로 λ번 발생한다
  예는 주어질 것!

- X의 확률분포

자연상수 e.

$$\boxed{P(X = x) = \frac{e^{-\lambda} \cdot \lambda^x}{x!}, \quad x = 0, 1, \cdots}$$

lambda

$$E(X) = \lambda, \quad Var(X) = \lambda$$

- 모수 λ(포아송 분포의 평균)에 대해서 표본으로부터 추론

X 포아송분포의 조건
① 한 순간(즉, 특정 시각)에 2회 이상의 사건이 동시에 발생하지 않는다.
② 특정 시간에 발생한 사건의 횟수는 다른 시간에 발생하는 사건의 횟수에 영향을 받지 않는다.
③ 사건이 발생할 확률은

이항분포와 그에 관련된 분포들

주간의 길이에 비례한다.

"단위 시간의 길이"

- 부록 표-2

$$P(X \le c) = \sum_{x=0}^{c} \frac{e^{-\lambda} \cdot \lambda^{x}}{x!}$$

행

열

이 것들이 누적분포이다.

$X \sim \mathrm{Pois}(\lambda = 3.0), \quad x = 0, 1, \cdots$

	$\cdots$	$\lambda = 3.0$	$\cdots$
$c = 0$		$P(X \le 0)$	
$c = 1$		$P(X \le 1)$	
$c = 2$		$P(X \le 2)$	
$\vdots$		$\vdots$	

Chapter 7 이항분포와 그에 관련된 분포들

- 포아송분포에서의 확률계산

  부록 표-2의 확률표를 이용하여 아래의 확률들을 계산 가능

$$
\begin{aligned}
P(a \le X \le b) &= P(X \le b) - P(X \le a-1) \\
P(a < X < b) &= P(X \le b-1) - P(X \le a) \\
P(a \le X < b) &= P(X \le b-1) - P(X \le a-1) \\
P(a < X \le b) &= P(X \le b) - P(X \le a)
\end{aligned}
$$

$$
\begin{aligned}
P(X = a) &= P(X \le a) - P(X \le a-1) \\
P(X \ge a) &= 1 - P(X \le a-1)
\end{aligned}
$$

  확률 계산 시, 부등호와 등호에 따른 차이에 주의하자.

- 예제 6, 예제 7, 5.6

ex) $P(X=2) = P(X \le 2) - P(X \le 1)$

이산확률변수라 소수를 생각하지 않는다.