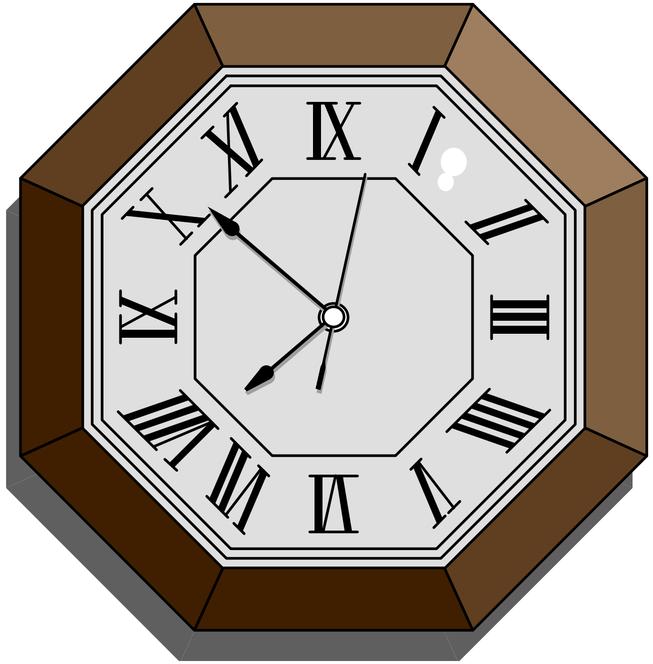
확률및통계(2)

제6주 이산형 확률 분포모형 (2)

hylee@silla.ac.kr

제 5절 초기하분포

- 초기하분포의 정의
- 초기하분포의 이항분포수렴



[정의] 초기하분포

크기 N인 유한모집단에서 임의의 사상 A에 속하는원소가 k개 있다고 하자. 이로부터 n개의 표본을 비복원임의추출하였을때 x개가임의의 사상 A에 속한다면 X의 확률분포는 다음과 같다.

$$P(X=x)=\frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}}, \quad x=0,1,2,\cdots,\min(k,n)$$

이 확률분포를 초기하분포(hypergeometric distribution) 하고, $H(N,k,n)$ 로표기한다.

예제 12

초기하분포가 확률함수임을 보이시오.

[풀이]

$$P(X = x) = \frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}} \geq 0,$$

$$x = 0,1,2,\cdots,\min(k,n)$$

이고,

$$\sum_{x=0}^{\min(k,n)} \binom{k}{x}\binom{N-k}{n-x} = \binom{N}{n}$$ 이므로

이것이 성립함을 증명하시오(숙제)

$$\sum_{x=0}^{\min(k,n)} \frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}} = 1$$ 이므로 초기하분포는

확률함수이다.

[숙제힌트]

$$(1+t)^k = \sum_{i=0}^{k} \binom{k}{i} t^i$$ 이고 $i = x$ 일 때,

$$(1+t)^{N-k} = \sum_{j=0}^{N-k} \binom{N-k}{j} t^j$$ 이고 $j = n - x$

일 때를 고려하고 $(1+t)^k (1+t)^{N-k} = (1+t)^N$

을 생각해서 $t^x t^{n-x} = t^n$ 의 계수를 계산해볼것

제 6절 포아송분포

- 어느 하루 동안 잘못 걸려온 전화의 회수
- 어떤 지역에서 1일 교통사고 사망자 수
- 인쇄된 책자의 페이지당 오자 수
- 일정한 혈액 속에 있는 적혈구의 수

- 주어진 시간 또는 영역에서 성공의 출현횟수를 X 라 할 때 X는 포아송분포를 따른다.

[정의] 포아송분포

확률변수 X 가 취하는 값이 0,1,2,...이고,

X 의 확률분포가

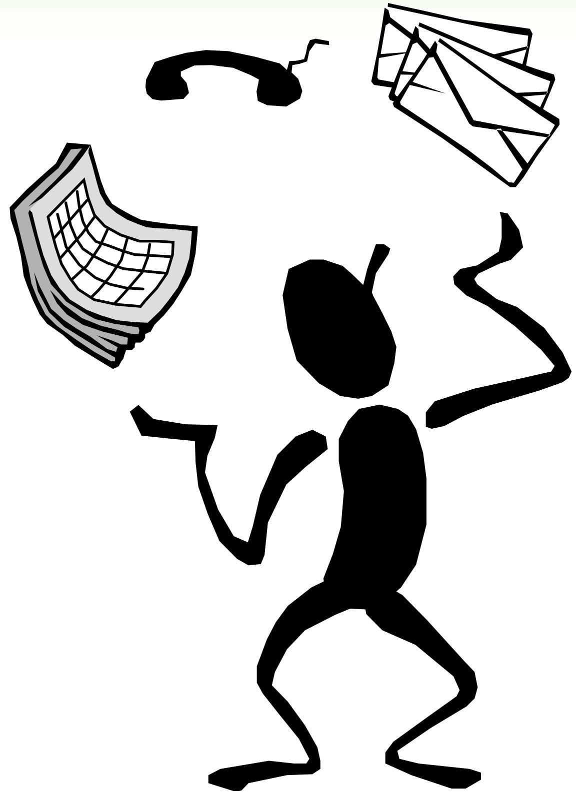
$$P(X = x) = \frac{\lambda^x e^{-\lambda}}{x!}, \quad x = 0,1,2,\cdots;(\lambda > 0)$$

일 때, X 는 포아송분포(Poisson distribution)

를 한다고 하며 $P(\lambda)$ 로 표기한다.

예제 14

포아송 분포가 확률함수임을 보이시오.



[풀이]

$$P(X=x)=\frac{\lambda e^{-\lambda}}{x!}\geq 0 \quad (x=0,1,2,\cdots)$$

이고

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

$$=e^{-\lambda}e^{\lambda}=1$$

이므로, 포아송분포는 확률함수이다.

제 7절 기하분포

- 기하분포의 정의
- 확률변수 X를 첫 번째 성공이 나올 때까지의 시행 횟수라고 정의하고, 시행 횟수 X에 대한 확률을 생각한다.

[정의] 기하분포

확률실험에서 성공의 확률이 $P(S)=p$ 이고, 실패의 확률이 $P(F)=q=1-p$ 인 반복독립시행에서 확률변수 X 를 처음으로 성공이 나타날 때까지의 시행횟수라고하면 X 의 확률분포는 다음과같다.

$$P(X=x)=q^{x-1}p \quad (x=1,2,\cdots)$$

이 확률분포를 기하분포(geometric distribution)라하고, $G(p)$로 표기한다.

예제 18

기하분포 $P(X = x) = q^{x-1} p \; (x = 1, 2, \cdots)$

가 확률함수임을 보이시오.

[풀이]

$$P(X=x)=q^{x-1}p\geq 0 \quad (x=1,2,3,\cdots)$$

이고,

$$\sum_{x=1}^{\infty}P(X=x)=p\sum_{x=1}^{\infty}q^{x-1}=\frac{p}{1-q}=1$$

이므로, 기하분포는 확률함수이다.

제 8절 음이항분포

- 시행횟수 n 을 미리 정하지 않고 어떤 주어진 성공의 횟수가 나올 때까지 베르누이 시행을 반복하는 경우를 생각해 보자.
- 즉, k 번째 성공이 나올 때까지의 시행횟수를 X 라 두면 X=x 가 될 확률은 실험을(x-1) 번 시행한 중에 k-1회 성공하고 x-k 번 실패한 후 맨 마지막에 성공하는 확률을 의미한다.

[예] 3회의 성공을 할때가지의 횟수를

X 라 하자. 그러면 표본공간은

$$\Omega = \{(S,S,S),(F,S,S,S),(S,F,S,S),(S,S,F,S)$$

$$(F,F,S,S,S),(F,S,F,S,S),(F,S,S,F,S),$$

$$(S,F,F,S,S),(S,F,S,F,S),$$

$$(S,S,F,F,S,),\cdots\}$$

이다.

여기서 각 샘플들의 마지막은 항상 S(성공)

이다.

그러므로 $X = x$ 가되는 경우는 x 번째는 S(성공)이고 그앞의 첫번째부터 $x-1$ 번째까지는 성공의 횟수가 $k-1$ 번 나오는 경우이다. 즉,

$$\overbrace{\underbrace{O\quad O\quad O\cdots O\quad O}_{k-1\text{ 번성공}}}^{\text{x-1 회}}\qquad \overset{\boxed{\text{k 번째 성공}}}{S}$$

총 횟수는 x 회

그러므로 이런 경우의 수는

$$\binom{x-1}{k-1} = \frac{(x-1)!}{(k-1)!(x-k)!}$$

이고, 각각의 경우에 확률은 모두

$$\boxed{p^{k-1}q^{(x-1)-(k-1)}}\boxed{p}$$

가 된다.

x-1 번 중
k-1 번 성공
할 확률

x 번째가 성공할
확률

그러므로

$$P(X = x) = \binom{x-1}{k-1} p^{k-1} q^{x-k} p$$

$$= \binom{x-1}{k-1} p^{k} q^{x-k}$$

[정의] 음이항분포

성공할확률이 p인 독립시행에서 k회성공할 때까지 시행을 반복할 때, X를 시행이 요구되는 횟수라 하면 X의 확률분포는 다음과같다.

$$P(X=x)=\binom{x-1}{k-1}p^k(1-p)^{x-k},\quad x=k,k+1,\cdots$$

이 확률분포를 음이항분포(negative binomial distribution)라하고, $NB(k,p)$로표기한다.

예제 20

음이항분포

$$P(X = x) = \binom{x-1}{k-1} p^k (1-p)^{x-k},$$

$$x = k, k+1, \cdots$$

가 확률함수임을 보이시오.

[풀이]

$$P(X=x)=\binom{x-1}{k-1}p^{x}(1-p)^{x-k}\geq,\ \ x=k,k+1,\cdots$$

이고,

$$\sum_{x=k}^{\infty}P(X=x)=\sum_{x=k}^{\infty}\binom{x-1}{k-1}p^{k}(1-p)^{x-k}$$

에서, $i=x-k$ 라 놓으면

$$\sum_{x=k}^{\infty} P(X = x) = p^k \sum_{i=0}^{\infty} \binom{i+k-1}{k-1} (1-p)^i$$

$$= p^k \{1-(1-p)\}^{-k}$$

$$= 1$$

이 되므로 음이항분포는 확률함수이다.