

Mathematical Statistics_v1.0

이산확률분포

- 베르누이분포
- 이항분포
- 기하분포
- 음이항분포
- 포아송분포
- 초기하분포
- 다항분포
- 지수분포
- 감마분포

베르누이 분포 (Bernoulli Distribution)

성공 or 실패 = 1 or 0
=> 베르누이 시행

분포 = 함수
함수 = 일반화
일반화 = 규칙

공식 ex) $P(x) = 0.3^x \cdot 0.7^{(1-x)}$
$E(x) = \text{sigma} \cdot x^t \cdot p(x) = p$
$V(x) = E(x^2) - (E(x))^2 = p \cdot (1-p)$

베르누이 실행 => 기하 분포
이항 분포 => 음이항 분포

이항 분포 (Binomial distribution)

베르누이 시행에서 1의 사건이
"몇 번" 발생했나에 대한 분포

공식 ex) $P(x) = nCr \cdot p^x \cdot (1-p)^{(n-x)}$
주요표현법 $nCr = \binom{n}{x}$
$X \sim B(n, p)$

기하 분포 (Geometric distribution)

베르누이 시행에서 1의 사건이 "몇 번째"에 성공을 하였느냐에 대한 분포
정의1) 처음 성공이 나올때까지 시행한 #회수를 확률변수로 하는 분포
x번째 사건 이성과 결혼하게 될 확률 $P(x)$

$P(1) = 0.05$
 $P(2) = 0.95 \cdot 0.05$
 $P(3) = 0.95 \cdot 0.95 \cdot 0.05$
 $P(x) = 0.95^{(x-1)} \cdot 0.05$

공식
 $P(x) = (1-p)^{(x-1)} \cdot p$
 $X \sim \text{Geo}(P)$



성공확률이 높을수록 감소하는 속도가 빠르다

기하분포 이름의 유래
기하수열 = 등비수열
기하평균
기하 (Geometric) => 선, 곡선, 도형

기하수열
 $a_0, a_0r^1, a_0r^2, a_0r^3, \dots$
 $a_n = a_0 \cdot r^n (n-1)$

기하분포
 $p(x) = p \cdot (1-p)^{(x-1)}$
##기하수열 공식과의 유사성확인

음이항 분포 (Negative Binomial Distribution)

성공확률 p, x번째 시행에서 k번째에 성공이 나올 확률 $p(x)$

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

따라서, $k=1$ 일 때, 기하분포가 된다.
전체시행횟수: n
성공횟수: k
실패횟수: r
 $n = k + r$ (독립, 종속, 상수)

- 1) r이 정해지고, k가 변수인 경우: 마지막이 실패
- 2) r이 정해지고, n이 변수인 경우: 마지막이 실패
- 3) k가 정해지고, r이 변수인 경우: 마지막이 성공
- 4) k가 정해지고, n이 변수인 경우: 마지막이 성공
- 5) n이 정해지고, k or r이 변수인 경우: 이항분포가 된다.

정의

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

공식

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

이름이 음이항 분포인 이유 (negative binomial)

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

포아송 분포 (Poisson Distribution)

이항분포 => 푸아송 분포
 $n \Rightarrow$ 매우 큼
 $p \Rightarrow$ 매우 작음

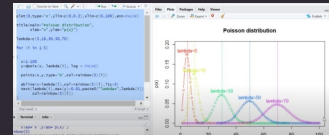
거리 길남이
하루(24)시간 거리 길남이수(x)
마주칠 확률 p -> 적다?
시행횟수 n -> 무슨단위로 정의? 분?시?모든순간?

n, p 두 가지 모두 다룰 수 없지만 np 는 다룰 수 있다.
np = 하루종일 돌아다니며 길남이를 마주치는 평균횟수
np = lambda

이항분포로부터 포아송분포 공식 유도 => 미분으로 증명하는 건 나중에

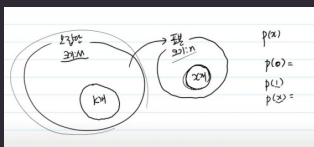
공식

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



그래프

초기하 분포 (Hypergeometric distribution)



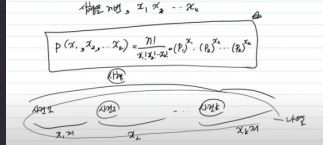
공식

$$P(X=x) = \frac{\binom{K}{x} \binom{N-K}{n-x}}{\binom{N}{n}}$$

다항 분포 (Multinomial distribution)

공식

$$P(X_1=x_1, X_2=x_2, \dots, X_k=x_k) = \frac{n!}{x_1! x_2! \dots x_k!} p_1^{x_1} p_2^{x_2} \dots p_k^{x_k}$$



지수분포 (Exponential distribution)

지수분포 = 포아송분포를 적분

지수분포 공식

$$f(t) = \lambda e^{-\lambda t}$$

지수분포 공식

$$F(t) = 1 - e^{-\lambda t}$$

지수분포 공식

$$P(X > t) = e^{-\lambda t}$$

지수분포 = from the question : 길남이를 1번 마주칠 때 까지 걸린 시간이 t 이하일 확률

감마분포 = from the question : 길남이를 n번 마주칠 때 까지 걸린 시간이 t 이하일 확률