

▼ 기하분포, 초기하분포, 음이항분포

01. 기하분포

베르누이 시행에서 처음 성공할 때까지 반복한 시행횟수

즉, 베르누이 시행에서 첫 성공까지의 시행횟수(한번의 성공만 포함)

- 기하분포의 확률질량함수(PMF)

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

$$P(X \leq x) = 1 - (1 - p)^x$$

- 예상치(Expectation)와 분산(Variance)

$$E(X) = \frac{1}{p}$$

$$Var(X) = \frac{1 - p}{p^2}$$

02. 음이항분포

x번의 독립적인 베르누이 시행 중 r 번의 성공을 포함하는 이항분포

- 이항분포의 확률질량함수(PMF)

$$P(X = x) = \binom{x-1}{r-1} (1-p)^{x-r} p^r = {}_{x-1}C_{r-1} (1-p)^{x-r} p^r$$

- 예상치(Expectation)와 분산(Variance)

$$E(X) = \frac{r}{p}$$

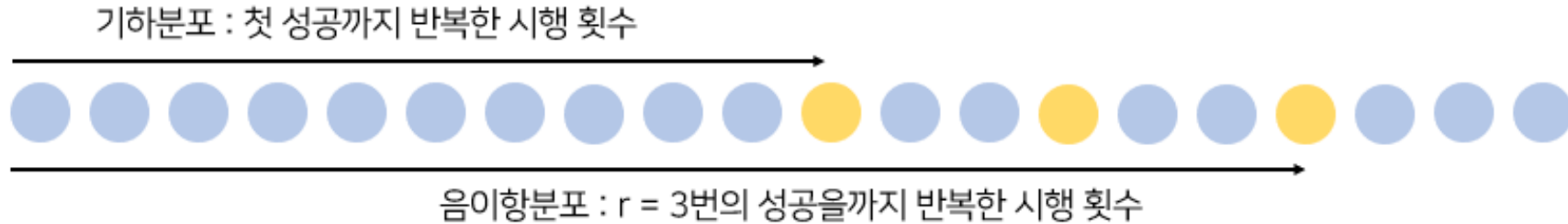
$$Var(X) = \frac{r(1-p)}{p^2}$$

- r = 1인 경우 = 기하분포

- 이항분포? 음이항분포? 이항분포의 경우 전체 시행 횟수 n, 성공횟수를 r이라고 할 때, $X \sim NB(r, p)$, $Y \sim B(n, p)$ 의 관계는

$$P(X \leq n) = P(Y \geq r)$$

즉, 이항분포의 경우 시행 횟수 n 이 고정되어 있고, 성공 횟수 r 이 확률변수이지만, 음이항분포의 경우 시행횟수 n 이 확률변수이고 성공횟수 r 이 고정되어있다



▼ 03. 초기하분포

N 개 중 원하는 결과가 x 개 있고, 이 중 n 개를 **비복원추출**로 뽑았을 때, 원하는 x 개가 뽑히는 경우의 분포

- 이항분포(Binomial) : 초기하 분포를 복원추출로 진행
- 초기하분포의 확률질량함수(PMF)

$$P(X = x) = \frac{\binom{r}{x} * \binom{N-r}{n-x}}{\binom{N}{n}} = \frac{rC_x * {}_{N-r}C_{n-x}}{NC_n}$$

- 예상치(Expectation)와 분산(Variance)

$$E(X) = n \frac{r}{N} \quad \text{Var}(X) = n \frac{r}{N} * \frac{N-r}{N-1} * \left(1 - \frac{r}{N}\right)$$

- 초기하분포의 N이 무한대로 간다면, 초기하분포는 이항분포와 동일해진다.

