■ 기하분포(Geometric Distribution)

- 성공할 확률이 p인 베르누이 시행을 성공할 때 까지 시행하는 경우 실패(시행) 횟수의 분포
  - 표본공간: $\Omega = \{S,\ FS,\ FFS,\ FFFS,\ \ldots\}$
  - 확률질량함수: $f(x) = (1-p)^x p, \quad x = 0, 1, 2, \ldots$
  - $X \sim \text{Geo}(p)$
  - 제1항이 p이고 공비가 $1-p$인 등비급수 형태
  - $Y = X+1$: 시행 횟수

$$f_Y(y) = (1-p)^{y-1} p, \quad y = 1, 2, \ldots$$

○ 등비급수의 합:

$$
\begin{aligned}
S &= a + ar + ar^2 + \cdots + ar^x \\
rS &= \quad\;\; ar + ar^2 + \cdots + ar^x + ar^{x+1}
\end{aligned}
$$

$$\Rightarrow (1-r)S = a - ar^{x+1} \quad \Rightarrow S = \frac{a - ar^{x+1}}{1-r}$$

○ $P(Y \le x) = P(X \le x-1)$: x번째 실험 이전에 성공할 확률

$$P(X \le x) = \sum_{k=0}^{x} p(1-p)^k = \frac{p - p(1-p)^{x+1}}{1-(1-p)} = 1-(1-p)^{x+1}$$

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

◉ 동전던지기

- $P(X \geq x)$: $x+1$번째 시행 이후에 성공(x번째까지 실패)할 확률

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

○ x번째까지 실패했다고 할 때, 다음($x+1$번째) 시행에서의 성공 확률

$$\frac{f_X(x)}{P(X \geq x)} = \frac{(1-p)^x p}{(1-p)^x} = p$$

• 무기억성(memoryless)

○ 5번 연속 뒷면이 나왔다고 하더라도 6번째가 앞면이 확률은 0.5

○ $P(Y \le x) = P(X \le x-1)$: x 번째 실험 이전에 성공할 확률

- $P(Y \le y) = P(X+1 \le y) = P(X \le y-1) = 1-(1-p)^y$
- $P(Y > y) = 1 - P(Y \le y) = (1-p)^y$

◉ 동전던지기

- 앞면($p=1/2$)이 나올 때 까지 동전던지기
- $P(Y \leq 2) = \frac{1}{2} + \left(\frac{1}{2}\right)^2 = 0.75 \Rightarrow$ 2번 이내에 끝날 확률은 0.75
- $P(Y \leq y) \geq 0.9$을 만족하는 최소 y는?

$$P(Y \leq y) = 1 - (1-p)^y = 1 - \frac{1}{2^y} \geq 0.9$$

$$\Rightarrow \frac{1}{2^y} \leq 0.1 \Rightarrow y \geq 4$$

- 기대값

$$\begin{aligned} S &= ar + 2ar^2 + 3ar^3 + 4ar^4 + \cdots \\ rS &= \quad\;\; ar^2 + 2ar^3 + 3ar^4 + \cdots \end{aligned}$$

$$(1-r)S = ar + ar^2 + ar^3 + \cdots = \frac{ar}{1-r} \Rightarrow S = \frac{ar}{(1-r)^2}$$

- ○ $E(X) = \sum_{x=0}^{\infty} xp(1-p)^x = p(1-p)/p^2 = (1-p)/p$
- ○ $E(Y) = E(X+1) = 1/p$
- ○ 동전던지기: $E(Y) = 2 \Rightarrow$

  게임을 끝내려면 평균 2번을 던져야 함

■ 음이항분포(Negative Binomial Distribution)

- 성공할 확률이 p인 베르누이 시행을 r번 성공할 때까지 시행하는 경우 실패(시행)횟수의 분포
  - X: 실패횟수, Y 시행횟수 ($Y = X + r$)
  - $Y = y$라고 하면, y번째는 S
    - · $y-1$번째까지 결과: $r-1$개 S, $y-r$개 F

$$f_Y(y) = \binom{y-1}{r-1} p^{r-1}(1-p)^{y-r} p, \quad y = r, r+1, \ldots$$

$$\Rightarrow f_Y(y) = \binom{y-1}{r-1} p^r (1-p)^{y-r}$$

- $Y \sim NB(r, p)$

○ $X = x$ 라고 하면, $x + r$ 번째 S,

- $x + r - 1$ 번째까지 결과: $r - 1$ 개 S, x 개 F

$$f(x) = \binom{x + r - 1}{r - 1} p^r (1 - p)^x, \quad x = 0, 1, 2, \ldots$$

◉ 가위바위보

- 5명과 차례로 가위바위보 게임
- 비기거나 지면 계속 게임을 진행하고 이기면 다음 사람과 게임

  $\Rightarrow\ p = 1/3$

- 게임이 완료될 때까지 10회 이하로 가위바위보 할 확률

$$P(Y \le 10) = \sum_{y=5}^{10} \binom{y-1}{5-1} \left(\frac{1}{3}\right)^{5} \left(\frac{2}{3}\right)^{y-5} = 0.213$$

- 기댓값
  - $X_i \sim$ **Geo**(p), X_i들은 서로 독립

$$X = X_1 + X_2 + \cdots + X_r$$
$$Y = Y_1 + Y_2 + \cdots + Y_r$$

  - $E(X) = E(X_1 + \cdots + X_r) = r(1-p)/p$
  - $E(Y) = r/p$

- 계수자료 분석에서 포아송분포의 대안으로 사용가능
  - 포아송분포: $E(X) = \lambda = Var(X)$

$\Rightarrow$ $\overline{x}$와 s^2의 차이가 심하면 ($\overline{x} \ll s^2$), 포아송분포?

● 요약

○ 기하분포

- 베르누이 시행을 성공할 때까지의 실패(시행) 횟수의 분포
- 무기억성

○ 음이항분포

- 베르누이 시행을 r 번 성공할 때까지의 실패(시행) 횟수의 분포
- $X_i \sim \text{Geo}(p)$, X_i 들은 서로 독립

$$X = X_1 + X_2 + \cdots + X_r$$