고급수리통계학
Note.

Oh Jae Kwon

충북대학교 통계학과

2021년 3월 27일

차 례

제 1 장

Some Special Distributions

1 확률변수, 분포

확률변수Random variable

- 확률변수는 확률실험random experiment을 했을 때 발생할 수 있는 결과들을 숫자로 요약 해주는 **대응규칙(함수)**

$$X : s \to \mathbb{R}, \quad s \in S$$

- s : 표본공간의 하나의 원소, S : 표본공간

- 확률질량함수probability mass function는 이산형 확률변수가 가질 수 있는 값 $x_1, x_2, \ldots$을 취할 확률을 계산해주는 함수

$$p_X(x) = P(X = x), \quad x = x_1, x_2, \ldots$$

- p_X 확률변수 x에 대한 pmf

- $P : A \to [0, 1]$, P(사건, 집합)

$\star\star\star$ $X = x$의 의미는 사건(event)이다. : $\{s : X(s) = x, s \in S\}$

- 누적확률함수cumulative distribution function는 확률변수가 특정값 또는 그 이하의 값을 취할 확률을 계산해주는 함수

$$F_X(x) = P(X \le x), \quad x \in \mathbb{R}$$

  < 특징 >

  1. 비감소함수
  2. $\lim_{x \to -\infty} F_X(x) = 0, \quad \lim_{x \to \infty} F_X(x) = 1$
  3. $\lim_{x \to x_0^+} F_X(x) = x_0 \qquad$ (right continuous)

분포distribution

- 확률변수가 가질 수 있는 값들이 흩어져 있는 정도

- 확률변수 X의 분포를 구하기

  - 확률변수 X의 의미

- 확률변수 X가 가질수 있는 값
- probability mass function, **check!** sum to 1 & cumulative distribution function
- $E(X)$, $Var(X)$
- moment generating function

2 베르누이 분포, 이항 분포

베르누이 시행 Bernoulli trial

- 표본공간이 **오직 두 개의 상호배타적인 원소로 구성된 실험**에서 시행

참고

임의로 표본공간을 두개의 원소만을 가질 수 있게 나눠서 확률변수를 정의할 경우

- 확률변수 X는 Bernoulli 확률변수이다.
- But, Bernoulli trial 은 아니다!

ex) 주사위 1개를 던지는 실험 $S = 1, 2, 3, 4, 5, 6$

$$X = \begin{cases} 1, & \text{홀수} \\ 0, & \text{짝수} \end{cases}$$

X는 Bernoulli 확률변수 이지만 Bernoulli trial 은 아니다.

베르누이 확률변수 $X \sim Bern(p)$

- 확률변수 X가 취할 수 있는 값이 오직 1 또는 0
- $P(X = 1) = p, \quad P(X = 0) = 1 - p, \quad 0 \leq p \leq 1$

베르누이 확률변수의 확률질량함수

1.
$$p_X(x) = \begin{cases} p, & x = 1 \\ 1 - p, & x = 0 \\ 0, & o.w \end{cases}$$

2.
$$p^x(1-p)^{1-x}, \quad x = 0, 1$$

- $x = 0, 1$: support, x가 가질 수 있느느 값 이외에는 0값을 갖는다.

독립인 두 베르누이 확률변수의 합

$$X_1, X_2 \overset{iid}{\sim} Bern(p)$$

iid : 동일하며identically 서로 독립인independent 분포distributed를 따른다.

독립성^{independence}

$$P_{X_1,X_2}(x_1,x_2) = P_{X_1}(x_1)P_{X_2}(x_2)$$

Q. 확률변수 X를 $X = X_1 + X_2$라 하자.

- X가 가질 수 있는 값은? $0, 1, 2$
- X의 확률질량함수는?

$$\begin{aligned}
P(X=0) &= P(X_1=0, X_2=0) = P(X_1=0)P(X_2=0) = (1-p)^2 \\
P(X=1) &= P(X_1=1, X_2=0) + P(X_1=0, X_2=1) \\
&= P(X_1=1)P(X_2=0) + P(X_1=0)P(X_2=1) \\
&= p(1-p) + p(1-p) = 2p(1-p) \\
P(X=2) &= P(X_1=1, X_2=1) = P(X_1=1)P(X_2=1) = p^2
\end{aligned}$$

- X의 의미 : 성공 횟수

이항 분포$^{\text{binomial distribution}}$

성공확률 p인 베르누이 시행을 n번 독립적으로 반복시행하는 경우, 확률변수 X를 '성공 횟수'라고 하자. X는 시행 횟수 n, 성공확률p를 모수$^{\text{parameter}}$로 갖는 이항 분포를 따른다.

$$X \sim Bin(n,p)$$

$$X_1, X_2, \ldots, X_n \overset{iid}{\sim} Bern(p), \quad X = X_1 + \cdots + X_n \sim Bin(n,p)$$

- X가 가질 수 있는 값 : $0, \cdots, n$
- 베르누이 확률변수와 이항 확률변수와의 관계

$$Bin(1,p) \equiv Bern(p)$$

이항 확률변수의 확률질량함수

- $p_X(k) = P(X=k) = \binom{n}{k} p^x (1-p)^{n-k}, \quad k = 0, 1, \ldots, n$
- 확률질량함수는

$$\sum_{k=0}^{n} p_X(k) = 1$$

  을 반드시 만족해야한다.

이항 정리*binomial theorem*

$$(a+b)^n = \sum_{j=0}^{n} \binom{n}{j} a^j b^{n-j}$$

독립인 이항확률변수의 합의 분포

$$X \sim Bin(n,p), \quad Y \sim Bin(m,p), \quad (n,m\text{은 자연수})$$

$$Z = X+Y \sim Bin(n+m,p)$$

Z의 확률질량함수

$$\begin{aligned} P_Z(k) = P(Z=k) = P(X+Y=k) &= \sum_{j=0}^{m} P(X=k-j|Y=j)P(Y=j) \\ &= \sum_{j=0}^{m} P(X=k-j)P(Y=j) = \sum_{j=0}^{m} \binom{n}{k-j} p^{k-j}(1-p)^{n-k+j} \binom{m}{j} p^j (1-p)^{m-j} \\ &= \sum_{j=0}^{m} \binom{n}{k-j}\binom{m}{j} p^k(1-p)^{n+m-k} = p^k(1-p)^{n+m-k} \sum_{j=0}^{m} \binom{n}{k-j}\binom{m}{j} \\ &= \binom{n+m}{k} p^k (1-p)^{n+m-k} \end{aligned}$$

전확률 공식*Law of Total Probability*

$$\begin{aligned} P(A) &= P((A\cap B_1)\cup\cdots\cup(A\cap B_J)) \\ &= \sum_{j=1}^{J} P(A\cap B_j) = \sum_{j=1}^{J} P(A|B_j)P(B_j) \end{aligned}$$

Vandermonde convolution

$$\binom{n+m}{k} = \sum_{j=0}^{k} \binom{n}{j}\binom{m}{k-j}$$

3 기대값, 지시확률변수, 초기하 분포

예제 : 52장의 트럼프 카드에서 임의로 5장을 뽑는 실험

- 확률변수 X : Ace의 개수
- 확률변수 X가 갖을 수 있는 값 : 0, 1, 2, 3, 4

- X의 확률질량함수 - 비복원

$$P(X=k) = \frac{\binom{4}{k}\binom{48}{5-k}}{\binom{52}{5}}$$

- ex) 초기하 (Hypergeometry) - 복원

$$Y_1, Y_2, Y_3, Y_4, Y_5 \sim Bern(\frac{4}{52})$$
$$Y_j \triangleq \begin{cases} 1 & , Ace \\ 0 & , NotAce \end{cases}$$
$$Y = \sum_{j=1}^{5} Y_i \sim Bin(5, \frac{4}{52})$$
$$P_Y(k) = \binom{5}{k}\left(\frac{4}{52}\right)^k \left(\frac{48}{52}\right)^{5-k}$$

기대값 : 이산형 확률변수 X의 기대값은 X가 가질 수 있는 모든 값들에 각각의 확률을 가중치로 하여 계산된 산술평균

$$E(X) = \sum_x xP(X=x) = \sum_x p_x(x)$$

- 이산형 확률변수의 기대값을 계산하려면 pmf가 필요
- 확률변수의 기대값 $E(X)$은 고정된 상수[constant]이다. (randomness가 없음)

베르누이, 이항분포의 기대값

- **Bernoulli** : $X \sim Bern(p)$일때, $E(X)$?

$$0 \cdot P(X=0) + 1 \cdot P(X=1) = p$$

- **Binomial** : $X \sim Bin(n,p)$일때, $E(X)$?

$$\begin{aligned}\sum_{x=0}^{n} x\binom{n}{x}p^x(1-p)^{n-x} &= \sum_{x=1}^{n} n\binom{n-1}{x-1}p^x(1-p)^{n-x} \\ &= n\sum_{x=1}^{n}\binom{n-1}{x-1}p^x(1-p)^{n-x} \\ &= n\sum_{t=0} n-1\binom{n-1}{t}p^{t+1}(1-p)^{(n-1)-t} \\ &= np\sum_{t=0} n-1\binom{n-1}{t}p^t(1-p)^{(n-1)-t} = np\end{aligned}$$

공식

n개 중 k개를 선택하여 대표를 뽑는 경우의 수와 대표 한 명을 먼저 뽑고 나머지 $(n-1)$명 중에서 $(k-1)$명을 뽑는 경우의 수

$$k\binom{n}{k} = n\binom{n-1}{k-1}$$

기댓값의 선형성linearity

기대값의 성질 : 임의의 상수 a, b와 확률변수 X, Y에 대하여

1. $E(a) = a$
2. $E(bX) = bE(X)$
3. $E(a + bX) = a + bE(X)$
4. $E(X + Y) = E(X) + E(Y)$, ($\star$ X, Y의 독립성과 무관)
5. $E(aX + bY) = aE(X) + bE(Y)$, ($\star$ X, Y의 독립성과 무관)

지시확률변수indicator random variable

- 1 또는 0을 갖는 확률변수(베르누이 확률변수와 매우 유사)
- 복잡한 문제를 우리가 관심있는 사건과 그 나머지에만 관심을 둔다면, 어떤 문제든지 지시확률변수를 이용해 표현이 가능

$$X = \begin{cases} 1, & if\ A\text{라는 사건이 발생} \\ 0, & o.w \end{cases}$$

- 지시확률변수를 잘 이용하면 관심있는 사건에 대한 확률을 기대값으로 표현이 가능 $\Rightarrow$ **fundamental bridge**

$$E(X) = P(A)$$

초기하분포의 기대값

- $X \sim Hypergeom(n, K, N)$
- 단, $n < N,\ K < N,\ K < n$
- pmf

$$P(X = k) = \frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}}, \quad k = 0, \cdots, K$$

- X의 기대값 $E(X)$?

1. pmf를 이용

$$
\begin{aligned}
E(X) = \sum_{k=0}^{K} kP(X=k) &= \sum_{k=1}^{K} k\frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}} \\
&= \sum_{k=1}^{K} K\frac{\binom{K-1}{k-1}\binom{N-K}{n-k}}{\binom{N}{n}} \\
&= \sum_{k=1}^{K} \frac{Kn}{N}\frac{\binom{K-1}{k-1}\binom{(N-1)-(K-1)}{(n-1)-(k-1)}}{\binom{N-1}{n-1}} \\
&= n\cdot\frac{K}{N}
\end{aligned}
$$

공식

$$
\binom{n}{r} = \frac{n!}{r!(n-r)!} = \frac{n}{r}\cdot\frac{(n-1)!}{(r-1)!((n-1)-(r-1))!} = \frac{n}{r}\cdot\binom{n-1}{r-1}
$$

2. 선형성, 지시확률변수를 이용

$$
\begin{aligned}
& Y_1, Y_2, Y_3, Y_4, Y_5 \sim Bern(\frac{4}{52}) \\
& Y_j \triangleq \begin{cases} 1 & , Ace \\ 0 & , NotAce \end{cases} \\
& Y = \sum_{j=1}^{5} Y_i \sim Bin(5, \frac{4}{52}) \\
E(Y) &= E(Y_1 + \cdots + Y_5) = E(Y_1) + \cdots + E(Y_5) \\
&= \sum_{j=1}^{5} P(j\text{th card Ace}) = 5\cdot\frac{4}{52}
\end{aligned}
$$

4 기하 분포

기하 분포geometric distribution

성공확률 p인 베르누이 시행을 독립적으로 반복시행하는 경우, 확률변수 X를 'Ḣ한번 성공할때까지 시

행한 횟수' 라고 하자. 이 때 X는 성공확률 p를 모수로 갖는 기하 분포를 따른다고 한다.

$$X \sim Geom(p)$$

- X가 가질 수 있는 값은? $1, 2, 3, \cdots, \infty$
- X의 pmf? $P_X(k) = P(X = k) = (1-p)^{k-1}p, \quad k = 1, \cdots, \infty$

  check!

$$\sum_{k=1}^{\infty}(1-p)^{k-1}p = p\sum_{k=1}^{\infty}(1-p)^{k-1} = p \cdot \frac{1}{1-(1-p)} = 1$$

무한등비급수*infinite geometric series*

$a \neq 0$일때 무한등비급수

$$a + ar + ar^2 + \cdots = \sum_{n=1}^{\infty} ar^{n-1} = \frac{a}{1-r}$$

- $X \sim Geom(p)$ 일 때, $E(X)$?

$$\begin{aligned} E(X) &= \sum_{k=1}^{\infty} kP(X = k) = \sum_{k=1}^{\infty} k(1-p)^{k-1}p \\ &= p\sum_{k=1}^{\infty} k(1-p)^{k-1} = p\frac{1}{p^2} = \frac{1}{p} \end{aligned}$$

무한등비급수 변형

$$\begin{aligned} & \sum_{k=1}^{\infty}(1-p)^{k-1} = \frac{1}{p} \\ \Leftrightarrow \quad & \frac{1}{1-p}\sum_{k=1}^{\infty}(1-p)^{k} = \frac{1}{p} \\ \Leftrightarrow \quad & \sum_{k=1}^{\infty}(1-p)^{k} = \frac{1-p}{p} \quad \text{양변 } p\text{에 대하여 미분} \\ \Leftrightarrow \quad & -\sum_{k=1}^{\infty}k(1-p)^{k-1} = -\frac{1}{p^2} \\ \Leftrightarrow \quad & \sum_{k=1}^{\infty}k(1-p)^{k-1} = \frac{1}{p^2} \end{aligned}$$

- **Story proof** $p = 0.1$이라는 것은 한번 성공할 확률이 0.1이다. 다시 말하면 한 번 성공하려면 10번은 시행해야 한다.

5 음이항 분포

음이항 분포 **negative binomial distribution**

성공확률 p인 베르누이 시행을 독립적으로 반복시행하는 경우, 확률변수 X를 'r번 성공할때까지 시행한 횟수'

$$X \sim NB(r, p)$$

- X가 가질 수 있는 값 : $r, r+1, r+2, \ldots$
- X의 pmf ?

$$P_X(k) = P(X = k) = \binom{k-1}{r-1} p^r (1-p)^{k-r}$$

음이항 정리 *(음이항 급수)*

$$(a+b)^{-n} = \sum_{k=0}^{\infty} \binom{-n}{k} a^k b^{-n-k}$$

Tayler expansion

$$f(x) = (1-x)^{-n}$$

$$\begin{aligned} f(x) =& f(0) + f'(0)x + \frac{1}{2}f''(0)x^2 + \frac{1}{3!}f'''(0)x^3 + \cdots \\ =& 1 + (-n)(-1)x + \frac{1}{2}(-n)(-n-1)(-1)(-1)x^2 \\ &+ \frac{1}{3!}(-n)(-n-1)(-n-2)(-1)(-1)(-1)x^3 + \cdots \\ =& \sum_{k=0}^{\infty} \binom{-n}{k} (-1)^k x^k = \sum_{k=0}^{\infty} \binom{n+k-1}{k} x^k \end{aligned}$$

$$\begin{aligned} \sum_{k=0}^{\infty} \binom{k-1}{r-1} p^r (1-p)^{k-r} =& p^r \sum_{k=0}^{\infty} \binom{k-1}{r-1} (1-p)^{k-r} \\ =& p^r \sum_{r=0}^{\infty} \binom{r+t-1}{r-1} (1-p)^t \\ =& p^r \sum_{r=0}^{\infty} \binom{r+t-1}{t} (1-p)^t \\ =& p^r (1-(1-p))^- r = p^r p^{-r} = 1 \end{aligned}$$

- 기하분포와의 관계

$$X_1, \ldots, X_r \sim Geom(p), \quad X = \sum_{i=1}^{r} X_i \sim NB(r, p)$$

6 포아송 분포

자연상수 $e = 2.71828$

자연상수를 정의하는 방법

1. $x = 1$부터 $x = a$까지 $y = \frac{1}{x}$ 의 적분값이 1이 되게하는 a 값

$$\int_1^e \frac{1}{x} dx = 1$$

2.

$$e = \sum_{n=0}^{\infty} \frac{1}{n!} = 1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \cdots$$

3.

$$e = \lim_{t \to 0} (1+t)^{\frac{1}{t}} \quad or \quad e = \lim_{n \to \infty} \left(1 + \frac{1}{n}\right)^n$$

4. 성질

$$e^a = \lim_{n \to \infty} \left(1 + \frac{a}{n}\right)^n$$

포아송 분포 **poisson**

단위시간(공간) 에서 발생하는 사건의 발생횟수

- 시행횟수 n 에는 제약을 두지 않음 (관심이 없음) : $n \to \infty$
- (대체적으로) 사건이 발생할 확률 p가 아주 작은 경우에 사용됨 : $p \to 0$

$\lambda = np$ 를 단위시간(공간) 평균 발생횟수라고 하면, 확률변수 X 는 모수가 λ인 포아송 분포를 따른다고 한다.

$$X \sim P(\lambda)$$

이항분포의 포아송 근사 **approximation**

$X \sim Bin(n, p)$ 이면

$$P_X(x) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0, 1, 2, 3, \ldots, n$$

이때 $n \to \infty, p \to 0, \lambda = np(\lambda \text{ is fixed})$ 인 경우?

$$\begin{aligned} p_X(x) &= \frac{n(n-1)\cdots(n-x+1)}{x!} \left(\frac{\lambda}{n}\right)^x \left(1 - \frac{\lambda}{n}\right)^{n-x} \\ &= \frac{\lambda^x}{x!} e^{-\lambda} = \frac{\lambda^x e^{-\lambda}}{x!}, \quad x = 0, 1, 2, \ldots \end{aligned}$$

기대값

$$E(x) = \sum_{x=0}^{\infty} x \frac{\lambda^x e^{-\lambda}}{x!} = \lambda \sum_{x=1}^{\infty} \frac{\lambda^{x-1} e^{-\lambda}}{(x-1)!} = \lambda \sum_{t=0}^{\infty} \frac{\lambda^t e^{-\lambda}}{(t)!} = \lambda$$

e의 성질

$$\sum_{t=0}^{\infty} \frac{\lambda^t}{t!} = e^{\lambda}$$

증명 : Linearity $E(X+Y) = E(X) + E(Y)$

$$E(X) = \sum_x xP(X=x) \Rightarrow \sum_s X_{(s)} P(\{s\})$$

$$\begin{aligned} E(X+Y) &= \sum_s (X+Y)_{(s)} P(\{s\}) = \sum_s (X_{(s)} + Y_{(s)}) P(\{s\}) \\ &= \sum_s X_{(s)} P(\{s\}) + \sum_s Y_{(s)} P(\{s\}) = E(X) + E(Y) \end{aligned}$$

7 분산, LOTUS

$$Var(X) = E[(X - E(X))^2] = E(X^2) - \{E(X)\}^2 \geq 0 \quad \leftrightarrow \quad E(X^2) \geq \{E(X)\}^2$$

Law Of The Unconscious Statistician (LOTUS)
이산형 확률변수 X의 함수꼴인 확률변수 $g(X)$의 기대값 $E[g(X)]$

$$E[g(X)] = \sum_x g(x) p_X(x)$$

$Y = g(x), X = g^{-1}(y)$

$$\begin{aligned} E(Y) &= \sum_y yP(g(x) = y) = \sum_y y \sum_{x \in g^{-1}(y)} f(x) \\ &= \sum_y \sum_{x \in g^{-1}(y)} g(x) f(x) = \sum_x g(x) f(x) \end{aligned}$$

이산형분포들의 분산 계산

$$Var(X) = E(X^2) - \{E(X)\}^2$$

- 베르누이 분포 $X \sim Bern(p)$
- 이항 분포 $X \sim Bin(n,p)$
- 초기하 분포 $X \sim Hypergeom(n, K, N)$
- 기하 분포 $X \sim Geom(p)$
- 음이항 분포 $X \sim NB(r,p)$
- 포아송 분포 $X \sim P(\lambda)$

8 연속형 continuous 확률변수

주어진 실수구간 내에 속하는 어떠한 실수값도 가질 수 있는 확률변수

연속확률변수는 취할 수 있는 값이 무한히 많기 때문에, 이 중 어떤 한 실수값을 정확히 취할 확률은 0이다. (측정이 불가!)

확률밀도함수 probability density function, pdf

확률변수 X가 구간 interval $[L, U]$에서 값을 가지는 연속형 확률변수인 경우,

- 모든 $x \in [L, U]$에 대하여 $f_X(x) \geq 0$
- 모든 $xn \in [L, U]$에 대하여 $f_X(x) = 0$
- 값을 가지는 모든 구간에서 f_X를 적분 intergration 하면 1

$$\int_L^U f_X(x)dx = 1$$

- $L \leq a \leq b \leq U$에 대하여

$$P(a \leq X \leq b) = \int_a^b f_X(x)dx$$

위 조건들을 만족하면 $f_X(x)$를 X의 확률밀도함수라 한다.

확률밀도함수의 성질

1. $f_X(x) \geq 0$, $\int_{-\infty}^{+\infty} f_X(x)dx = 1$
2. $P(x_1 \leq X \leq x_2) = \int_{x_1}^{x_2} f_X(t)dt$
3. 함수값 $f_X(x)$가 갖는 의미는?
4. 연속형 확률변수의 누적밀도함수

$$F_X(x) = P(X \leq x) = \int_{-\infty}^{x} f_X(t)dt$$

9 균등분포 Uniform distribution

- 연속형 확률변수 X 는 실구간 $[a, b]$에서 값을 가지는 연속형 확률변수

$$f_X(x) = \begin{cases} \frac{1}{b-a} & \text{if} \quad a \leq x \leq b, \\ 0 & o.w \end{cases}$$

$$\int_a^b c dx = cb - ca = 1,$$

$$c = \frac{1}{b-a}$$

$$E(X) = \int_a^b x \frac{1}{b-a} dx = \frac{b^2 - a^2}{2(b-a)} = \frac{b+a}{2}$$

확률적분변환[**probability integral transformation**]

- 목적 : 확률변수 X를 생성[generation]하고 싶음
- 주어진 것
  - 균등분포 $U \sim Unif(0,1)$에서 관측된 값 $u_1, \ldots, u_n$
  - 확률변수 X 의 cdf F_X
  - 단, F_X는 연속이며 순증가[strictly increasing] 함수 → 역함수 존재

$$F_X^{-1}(U) \overset{d}{\equiv} X$$

  - 즉, 확률변수 X는 $F_X(u_1), \ldots, F_X(u_n)$으로 생성가능

$$F_X^{-1}(U) \overset{d}{\equiv} X$$

$$F_X(x) = P(X \leq x) = P(F^{-1}(U) \leq x) = P(U \leq F_X(x)) = F_X(x)$$

10 정규분포[normal distribution]

- 표준정규분포 $Z \sim N(0,1)$의 pdf

$$f_z(z) = \frac{1}{\sqrt{2\pi}} exp\left(-\frac{z^2}{2}\right)$$

- Z의 cdf를 Φ라 하면

$$\Phi(z) = P(Z \leq z),\ \Phi'(z) = f_Z(z)$$

- 평균 μ, 표준편차 σ인 정규분포 $X \sim N(\mu, \sigma^2)$와 Z의 관계

$$X = \mu + \sigma Z$$

- X의 cdf

$$P(X \leq x) = P(\mu + \sigma Z \leq x) = P(Z \leq \frac{x-\mu}{\sigma} = \Phi\left(\frac{x-\mu}{\sigma}\right)$$

- X의 pdf 는 cdf를 미분하면

$$f_X(x) = \Phi'\left(\frac{x-\mu}{\sigma}\right)\frac{1}{\sigma} = f_z\left(\frac{x-\mu}{\sigma}\right)\frac{1}{\sigma} = \frac{1}{\sqrt{2\pi}\sigma} exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

11 지수분포exponential distribution

- 단위시간당 사건의 편균 발생횟수를 λ라 할 때, 함번 사건이 일어날때까지 걸리는 시간을 X라 하면
$$X \sim Exp(\lambda)$$

포아송과정poisson process

- λ : 단위시간당 사건의 평균 발생횟수 $[0, 1]$
- λx : 어떤 특정 시점 $x > 0$까지 평균 사건 발생횟수 $[0, x]$
- X의 cdf

$$\begin{aligned}
F_X(x) = P(X \leq x) &= 1 - P(X \geq x) \\
&= 1 - P([0, x]\text{동안 사건이 일어나지 않음}) \\
&= 1 - P(Y = 0) \quad , Y \sim Poisson(\lambda x) \\
&= 1 - \frac{e^{-\lambda x}(\lambda x)^0}{0!} \\
&= 1 - e^{-\lambda x}
\end{aligned}$$

- X의 pdf
$$f_X(x) = F'_X(x) = \lambda e^{-\lambda x} = \frac{1}{\theta} e^{-\frac{x}{\theta}}$$
- $\theta = \frac{1}{\lambda}$: 사건이 한번 발생할 때까지 걸리는 평균 시간
- $Y \sim Exp(1)$의 기대값과 분산

$$\begin{aligned}
E(Y) &= \int_0^\infty y e^{-y} dy = y - \frac{1}{y} e^{-y}|_0^\infty = 1 \\
E(Y^2) &= \int_0^\infty y^2 e^{-y} dy = 2 \\
Var(Y) &= 1
\end{aligned}$$

- $Y = \lambda X,\ X = \frac{Y}{\lambda}$
$$E(X) = \frac{1}{\lambda} = \theta, \quad Var(X) = \frac{1}{\lambda^2} = \theta^2$$
- $X \sim Exp(\lambda)$

무기억성$^{memoryless\ property}$

지수분포, 기하분포에서 나타나는 성질

$$
\begin{aligned}
P(X \geq s+t | X \geq s) &= \frac{P(X \geq s+t, X \geq s)}{P(X \geq s)} = \frac{P(X \geq s+t)}{P(X \geq s)} \\
&= \frac{1-F_X(s+t)}{1-F_X(s)} = \frac{1-(1-e^{-\lambda(s+t)})}{1-(1-e^{-\lambda s})} \\
&= \frac{e^{-\lambda(s+t)}}{e^{-\lambda s}} = e^{-\lambda t} = 1-(1-e^{-\lambda t} \\
&= 1-F_X(t) = P(X > t)
\end{aligned}
$$

12 적률생성함수$^{\text{moment generating function, MGF}}$

- 통계학에서 적률$^{\text{moment}}$: 양수 n에 대하여 $E[X^n]$을 확률변수 X의 n차 적률$^{\text{n - th moment}}$이라 한다.
- 기대값 : 1차 적률, 분산 : 2차 적률과 관계
- 확률변수 X의 적률생성함수 M_X는 다음으로 정의

$$M_X(t) = E[e^{tX}]$$

- M_X가 왜 적률생성함수일까?

$$M_X(t) = E[e^{tX}] = E\left(\sum_{n=0}^{\infty} \frac{X^n t^n}{n!}\right) = \sum_{n=0}^{\infty} \frac{t^n}{n!} E(X^n)$$

기하 급수$^{geometric\ series}$

$$e = \sum_{n=0}^{\infty} \frac{1}{n!} \quad \Rightarrow \quad e^{ax} = \sum_{n=0}^{\infty} \frac{a^n x^n}{n!}$$

$$
\begin{aligned}
M_X(t) &= \sum_{n=0}^{\infty} \frac{t^n}{n!} E(X^n) \\
&= \frac{t^0 E(X^0)}{0!} + \frac{t^1 E(X^1)}{1!} + \frac{t^2 E(X^2)}{2!} + \cdots \\
M'_X(t) &= E(X) + t \times \star + t^2 \times \star \\
M'_X(0) &= E(X) \\
M''_X(t) &= 0 + E(X^2) + t \times \star \\
M''_X(0) &= E(X^2) \\
&\vdots \\
M_X^{(n)}(0) &= E(X^n)
\end{aligned}
$$

- M_X를 $t = 0$을 포함하는 구간에서만 정의하면 적률을 구할수 있음

적률생성함수 중요성

- MGF M_X는 모든 실수에서 정의될 필요까지 없이, 0을 포함하는 정당한 구간 $(-a, a), a > 0$에서만 정의가 되면 된다.
- 확률변수 X의 n차 적률 $E[X^n]$은 적률생성함수를 n번 미분하여 얻을 수 있음

$$E[X^n] = M_X^{(n)}(0)$$

- 적률생성함수는 확률변수의 분포를 결정한다. 즉, 두 확률변수의 MGF가 같으면 분포가 같음을 의미
- 두 확률변수 X, Y의 MGF가 각각 M_X, M_Y 이고 X와 Y 가 독립이면, 확률변수 $Z = X + Y$의 MGF M_Z는 다음과 같이 얻어짐

$$M_Z(t) = E[e^{t(X+Y)}] = E[e^{tX} e^{tY}] = E[e^{tX}]E[e^{tY}] = M_X(t) M_Y(t)$$

예제 : 적률생성함수

- $X \sim Bern(p),\ X = 0, 1$

$$
\begin{aligned}
M_X(t) &= E[e^{tx}] = \sum_x e^{tx} P(X = x) = e^0(1-p) + e^t p = (1-p) + e^t p \\
M'_X(t) &= pe^t, \quad M'_X(0) = p \\
M''_X(t) &= pe^t, \quad M''_X(0) = p \\
Var(X) &= E[X^2] - \{E[X]\}^2 = p - p^2 = (1-p)p
\end{aligned}
$$

- $X \sim Bin(n,p),\ X = 0,1,2,\dots,n$

$$\begin{aligned}
M_X(t) &= E[e^{tx}] = \sum_x e^{tx} P(X=x) = \sum_{x=0}^{n} e^{tx} \binom{n}{x} p^x (1-p)^{n-x} \\
&= \sum_{x=0}^{n} \binom{n}{x} (pe^t)^x (1-p)^{n-x} = (pe^t + (1-p))^n, \quad \text{이항정리} \\
M'_X(t) &= npe^t(pe^t + (1-p))^{n-1}, \quad M'_X(0) = np \\
M''_X(t) &= n(n-1)p^2 e^{2t}(pe^t + (1-p))^{n-2} + npe^t(pe^t + (1-p))^{n-1}, \\
M''_X(0) &= n^2p^2 - np^2 + np \\
Var(X) &= E[X^2] - \{E[X]\}^2 = np(1-p)
\end{aligned}$$

- $X \sim Exp(1)$

$$\begin{aligned}
M_X(t) &= E[e^{tx}] = \int_0^\infty e^{tx} e^{-x} dx = \int_0^\infty e^{-x(1-t)} dx = (1-t)^{-1}, \quad |t| \le 1 \\
M_X^{(n)}(t) &= n!(1-t)^{-n-1} = E(X^n) \\
M_X(t) &= \frac{1}{1-t} = \sum_{n=0}^\infty t^n = \sum_{n=0}^\infty \frac{n! t^n}{n!} = \sum_{n=0}^\infty \frac{E(X^n) t^n}{n!} \\
E(X^n) &= n!
\end{aligned}$$

- $Y \sim Exp(\lambda),\ X \sim Exp(1)$

$$E(Y^n) = E\left(\frac{X^n}{\lambda^n}\right) = \frac{1}{\lambda^n} E(X^n) = \frac{n!}{\lambda^n} = \theta^n n!$$

- $Z \sim N(0,1)$

$$\begin{aligned}
M_X(t) = E[e^{tx}] &= \int_{-\infty}^\infty e^{tz} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^\infty e^{-\frac{z^2}{2} + tz} dz = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^\infty e^{-\frac{(z-t)^2}{2}} e^{\frac{t^2}{2}} dz \\
&= e^{\frac{t^2}{2}} \int_{-\infty}^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{(z-t)^2}{2}} dz = e^{\frac{t^2}{2}}
\end{aligned}$$

  - $E(Z^n)$
  - 홀수

$$E(Z^n) = \int_{-\infty}^\infty z^n \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz = 0$$

  Z_n 기함수로 전구간에 적분을 하면 0 이된다.

  $\frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$ 우함수 : y축 대칭

  기함수 $\times$ 우함수 $=$ 기함수

– 짝수

$$M_Z(t) = e^{\frac{t^2}{2}} = \sum_{n=0}^{\infty} \frac{(\frac{t^2}{2})^n}{n!} = \sum_{n=0}^{\infty} \frac{\frac{t^{2n}}{2^n}(2n)!}{n!(2n)!} = \sum_{n=0}^{\infty} \frac{\frac{(2n)!}{2^n n!} t^{2n}}{(2n)!}$$

$$E(Z^{2n}) = \frac{(2n)!}{2^n n!}$$
$$E(Z^2) = 1$$
$$E(Z^4) = 1 \times 3$$
$$E(Z^6) = 1 \times 3 \times 5$$

- $X \sim Poisson(\lambda)$

$$M_X(t) = E(e^{tx}) = \sum_{x=0}^{\infty} e^{tx} \frac{e^{-\lambda}\lambda^x}{x!} = e^{-\lambda} \sum_{x=0}^{\infty} \frac{(\lambda e^t)^x}{x!} = e^{-\lambda} e^{\lambda e^t} = e^{-\lambda + \lambda e^t}$$

- $X \sim Poisson(\lambda_1),\ Y \sim Poisson(\lambda_2),\ X \perp Y,\ Z = X + Y$

$$M_Z(t) = M_X(t) M_Y(t) = e^{-\lambda_1 + \lambda_1 e^t} e^{-\lambda_2 + \lambda_2 e^t} = e^{-(\lambda_1 + \lambda_2) + e^t(\lambda_1 + \lambda_2)}$$