

1장. 확률분포

Jiho KWAK, Department of Geography, SNU

April 6, 2024

1 확률의 정의

1.1 용어들

- 랜덤실험: 결과를 확실하게 예측할 수 없는 실험
- 표본공간(sample space): 가능한 모든 실험 결과들의 집합
- 사건(event): 표본공간의 부분집합
- 사건 C 의 확률: 사건 C 가 일어날 그럴듯한 정도

Example. (동전 던지기 실험)

표본공간 $e = \{H, T\}$

사건 $C = \{H\}, \{T\}, \emptyset, \{H, T\}$

Definition. (시그마 필드) $\mathcal{B}$ 를 표본공간 e 의 부분집합들의 모임(즉, $B \subseteq \mathcal{P}(e)$)이라 하자. $\mathcal{B}$ 가 다음의 성질을 만족하면 **시그마 필드(σ -field)**라 한다.

- (i) $\emptyset \in \mathcal{B}$
- (ii) (여집합에 닫힘) $B \in \mathcal{B} \implies B^c \in \mathcal{B}$
- (iii) (가산 합집합에 닫힘) $C_1, C_2, \dots \in \mathcal{B} \implies \bigcup_{i=1}^{\infty} C_i \in \mathcal{B}$

Theorem. $\mathcal{D}$ 를 공집합이 아닌 e 의 부분집합들의 모임이라 하자. $\mathcal{D}$ 를 포함하는 모든 시그마 필드의 교집합은 시그마 필드이다. 즉, $\mathcal{B} := \bigcap \{\epsilon : \mathcal{D} \subseteq \epsilon, \epsilon \text{은 시그마 필드}\}$. 이 때, $\mathcal{B}$ 를 $\sigma(\mathcal{D})$ 라 표시하며, $\mathcal{B}$ 를 $\mathcal{D}$ 가 생성하는 시그마 필드라 한다.

Proof. (i) ($\emptyset \in \mathcal{B}$) $\mathcal{D}$ 를 포함하는 모든 시그마 필드 ϵ 는 $\emptyset$ 을 포함한다. $\therefore \emptyset \in \mathcal{B}$.

(ii), (iii) can be proved similarly. ■

Theorem. $\mathcal{B}$ 를 시그마 필드, e 를 표본공간이라 하자.

- (i) $e \in \mathcal{B}$
- (ii) $\mathcal{B}$ 는 가산 교집합에 대해 닫혀 있다. 즉 $C_1, C_2, \dots \in \mathcal{B} \implies \bigcap_{i=1}^{\infty} C_i \in \mathcal{B}$

Proof. 시그마 필드는 여집합에 대해 닫혀 있음을 이용한다. ■

-보렐 시그마 필드: 거리 공간에서 열린 집합들로 생성되는 시그마 필드

1.2 확률

Definition. (확률) e 를 표본공간, $\mathcal{B}$ 를 e 상의 시그마 필드라 하자. 함수 $P : \mathcal{B} \rightarrow \mathbb{R}$ 가 다음을 만족하면 **확률 (probability)** 혹은 **확률 측도 (probability measure)**라 한다.

- (i) $P(C) \geq 0, \forall C \in \mathcal{B}$
- (ii) $P(e) = 1$
- (iii) (가산가법성, countable additivity) $C_1, C_2, \dots \in \mathcal{B}, C_i \cap C_j = \emptyset, \forall i \neq j \implies P(\bigcup_{i=1}^{\infty} C_i) = \sum_{i=1}^{\infty} P(C_i)$

Definition. (측도) $\mu : \mathcal{B} \rightarrow \mathbb{R}$ 이 다음의 조건을 만족하면 **측도 (measure)**라 한다.

- (i) $\mu(C) \geq 0, \forall C \in \mathcal{B}$
- (ii) $\mu(\emptyset) = 0$
- (iii) (가산가법성) $C_1, C_2, \dots \in \mathcal{B}, C_i \cap C_j = \emptyset, \forall i \neq j \implies \mu(\bigcup_{i=1}^{\infty} C_i) = \sum_{i=1}^{\infty} \mu(C_i)$

Note. 확률은 $P(e) = 1$ 인 측도이다.

Theorem. $\mathcal{B}$ 는 e 상의 시그마 필드, $C, C_1, C_2, \dots \in \mathcal{B}$, P 는 $\mathcal{B}$ 상의 확률이다.

- (i) $P(\emptyset) = 0$
- (ii) (유한 가법성, finite additivity) $C_1, C_2 \in \mathcal{B}$ 서로소 $\implies P(C_1 \cup C_2) = P(C_1) + P(C_2)$
- (iii) $P(C^c) = 1 - P(C)$
- (iv) $0 \leq P(C) \leq 1, \forall C \in \mathcal{B}$
- (v) (단조성, monotonicity) $C_1 \subseteq C_2 \implies P(C_1) \leq P(C_2)$
- (vi) (가산 준가법성, countable subadditivity) $P(\sum_{i=1}^{\infty} C_i) \leq \sum_{i=1}^{\infty} P(C_i)$
- (vii) (a) (아래로부터 연속성, continuity from below) $C_1 \subseteq C_2 \subseteq \dots$ 이면 $P(\lim_{n \rightarrow \infty} C_n) = \lim_{n \rightarrow \infty} P(C_n), \lim_{n \rightarrow \infty} C_n = \bigcup_n C_n$
 (b) (위로부터 연속성, continuity from above) $C_1 \supseteq C_2 \supseteq \dots$ 이면 $P(\lim_{n \rightarrow \infty} C_n) = \lim_{n \rightarrow \infty} P(C_n), \lim_{n \rightarrow \infty} C_n = \bigcap_n C_n$

Proof.

- (i) Letting $C_1 = C_2 = \dots = \emptyset, C_i \cap C_j = \emptyset, \forall i \neq j$.

가산가법성에 의해 $P(\emptyset) = P(\bigcup_{i=1}^{\infty} C_i) = \sum_{i=1}^{\infty} P(C_i) = \sum_{i=1}^{\infty} P(\emptyset) \therefore P(\emptyset) = 0$

- (ii) $C_1, C_2 \in \mathcal{B}, C_1 \cap C_2 = \emptyset$ 이라 하자. $P(C_1 \cup C_2) = P(C_1) + P(C_2)$ 를 보이므로 $C_3 = C_4 = \dots = \emptyset$ 라 두자. 가산가법성에 의해 다음이 성립한다.

$$\begin{aligned} P(C_1 \cup C_2) &= P(\bigcup_{i=1}^{\infty} C_i) = \sum_{i=1}^{\infty} P(C_i) = P(C_1) + P(C_2) + 0 + 0 + \dots \\ &= P(C_1) + P(C_2) \end{aligned}$$

- (iii) $e = C \cup C^c, C \cap C^c = \emptyset$ 이므로 다음이 성립한다.

$$1 = P(e) = P(C \cup C^c) = P(C) + P(C^c)$$

(v) $C_1 \subseteq C_2, C_1, C_2 \in \mathcal{B}$ 라 하자. $P(C_1) \leq P(C_2)$ 를 보이므로 $C_1 \subseteq C_2$ 이므로 $C_2 = C_1 \cup (C_2 \setminus C_1)$ 이고 $C_1 \cap (C_2 \setminus C_1) = \emptyset$. 유한가법성에 의해 다음이 성립한다.

$$\begin{aligned} P(C_2) &= P(C_1) + P(C_2 \setminus C_1) \\ &\geq P(C_1) \quad (\because P(C_2 \setminus C_1) \geq 0) \end{aligned}$$

(iv)는 (v)에 의해 성립한다. ($\because P(\emptyset) = 0, P(e) = 1$)

(vi) $C_1, C_2, \dots \in \mathcal{B}$. $A_1 = C_1, A_2 = C_2 \setminus C_1, A_3 = C_3 \setminus (C_1 \cup C_2), \dots, A_n = C_n \setminus (C_1 \cup \dots \cup C_{n-1}), \dots$ 라 하자. $\bigcup C_i = \bigcup A_i$, A_i 들은 서로소이다. 따라서

$$\begin{aligned} P\left(\bigcup C_i\right) &= P\left(\bigcup A_i\right) \\ &= \sum_{i=1}^{\infty} P(A_i) \\ &\leq \sum_{i=1}^{\infty} P(C_i) \end{aligned}$$

(vii) (a) $C_1 \subseteq C_2 \subseteq \dots$. $A_1 = C_1, A_2 = C_2 \setminus C_1, A_3 = C_3 \setminus C_2, \dots, A_i = C_i \setminus C_{i-1}, \dots$ 라 하자. $\bigcup C_i = \bigcup A_i$, A_i 들은 서로소이다. 따라서

$$\begin{aligned} P\left(\lim_{n \rightarrow \infty} C_n\right) &= P\left(\bigcup_n C_n\right) = P\left(\bigcup_n A_n\right) \\ &= \sum_{i=1}^{\infty} P(A_i) \\ &= \lim_{n \rightarrow \infty} \sum_{i=1}^n P(A_i) \\ &= \lim_{n \rightarrow \infty} P\left(\bigcup_{i=1}^n A_i\right) \\ &= \lim_{n \rightarrow \infty} P(C_n) \end{aligned}$$

(b)는 (a)에서 여집합을 취하면 된다. ■

2 조건부 확률과 독립성

2.1 조건부 확률

Definition. $P(C) > 0$ 을 가정하자. C 가 주어졌을 때, A 의 **조건부 확률**은 $P(A|C) = \frac{P(A \cap C)}{P(C)}$ 로 정의한다.

Theorem. $P(C) > 0$.

(1) $P(C_1|C) \geq 0$

(2) $C_1, C_2, \dots$ 서로소이면 $P\left(\bigcup_{i=1}^{\infty} C_i|C\right) = \sum_{i=1}^{\infty} P(C_i|C)$

(3) $P(C|C) = 1$

(4) (곱의 법칙) $P(C_1) > 0$.

$$P(C_1 \cap C_2) = P(C_1) \cdot P(C_2|C_1)$$

$$P(C_1 \cap \cdots \cap C_n) = P(C_1) \cdot P(C_2|C_1) \cdot P(C_3|C_2 \cap C_1) \cdots P(C_n|C_1 \cap \cdots \cap C_{n-1})$$

(5) (전확률공식) $P(C_i) > 0, i = 1, 2, \dots, n, e = \bigcup_{i=1}^{\infty} C_i, C_i$ 들은 서로소.

$$P(C) = \sum_{i=1}^{\infty} P(C_i) \cdot P(C|C_i)$$

(6) (베이즈 정리) $P(C_i) > 0, i = 1, 2, \dots, n, e = \bigcup_{i=1}^n C_i, C_i$ 들은 서로소.

$$P(C_j|C) = \frac{P(C_j)P(C|C_j)}{\sum_{i=1}^n P(C_i)P(C|C_i)}$$

Proof.

(2) 정의를 이용한다.

$$P\left(\bigcup_{i=1}^{\infty} C_i|C\right) = \frac{P\left(\left(\bigcup_{i=1}^{\infty} C_i\right) \cap C\right)}{P(C)} = \frac{P\left(\bigcup_{i=1}^{\infty} (C_i \cap C)\right)}{P(C)}$$

$$= \sum_{i=1}^{\infty} \frac{P(C_i \cap C)}{P(C)} = \sum_{i=1}^{\infty} P(C_i|C)$$

(4) $P(C_2|C_1) = \frac{P(C_1 \cap C_2)}{P(C_1)}.$

(5)

$$P(C) = P(C \cap e) = P\left(C \cap \bigcup_{i=1}^n C_i\right)$$

$$= P\left(\bigcup_{i=1}^n (C \cap C_i)\right) = \sum_{i=1}^n P(C \cap C_i) \quad (\because \text{서로소})$$

$$= \sum_{i=1}^n P(C_i) \cdot P(C|C_i)$$

(6)

$$P(C_j|C) = \frac{P(C_j \cap C)}{P(C)}$$

$$= \frac{P(C_j)P(C|C_j)}{\sum_{i=1}^n P(C_i)P(C|C_i)}$$

Note. $P(C_j|C) \propto P(C_j) \cdot P(C|C_j), j = 1, \dots, n$

■

2.2 독립성

Definition.

(a) (두 개의 사건)

$$C_1, C_2 : \text{독립} \iff P(C_1 \cap C_2) = P(C_1)P(C_2)$$

(b) (세 개 이상 사건들)

$$\begin{aligned} C_1, \dots, C_n : \text{독립} &\iff P(C_i \cap C_j) = P(C_i)P(C_j), \forall i \neq j \\ &\quad P(C_i \cap C_j \cap C_k) = P(C_i)P(C_j)P(C_k), \forall i \neq j \neq k \\ &\quad \vdots \\ &\quad P(C_1 \cap \dots \cap C_n) = P(C_1) \cdots P(C_n) \end{aligned}$$

(c) (두 개의 실험)

$$\epsilon_1 = \{C_1, \dots, C_m\}, \epsilon_2 = \{A_1, \dots, A_n\} : \text{독립} \iff C_i \text{와 } A_j \text{가 독립 } \forall i = 1, \dots, m, \forall j = 1, \dots, n$$

Note. $P(C_1) > 0$ 이라고 하자. $C_1 \perp\!\!\!\perp C_2$ 이면 $P(C_2 \cap C_1) = P(C_1) \cdot P(C_2|C_1) = P(C_1)P(C_2)$ 이므로

$$P(C_2|C_1) = P(C_2).$$

Example. (독립은 아니지만 짝으로는 독립일 때) 1, 2, 3, 123이 적힌 네 개의 공이 있다고 하자.

$$A_i = \{i \text{가 쓰인 공을 추출}\}, i = 1, 2, 3$$

$$P(A_1) = \frac{1}{2} = P(A_2) = P(A_3)$$

$$P(A_1 \cap A_2) = \frac{1}{4} = P(A_1 \cap A_3) = P(A_2 \cap A_3)$$

$$P(A_1 \cap A_2) = P(A_1)P(A_2) \text{이므로 } A_1, A_2, A_3 : \text{짝으로 독립 (pairwise independent)}$$

하지만 $P(A_1 \cap A_2 \cap A_3) = \frac{1}{4} \neq P(A_1)P(A_2)P(A_3) = \frac{1}{8}$ 이므로 A_1, A_2, A_3 은 독립이 아니다.

3 확률변수, 누적분포함수, 확률밀도함수

3.1 확률변수

Definition.

(1) (확률변수)

$$X \text{는 } e \text{ 상에 정의된 확률변수} \iff X \text{는 } e \text{ 상에 정의된 실수값을 갖는 함수. 즉, } X : e \rightarrow \mathbb{R}$$

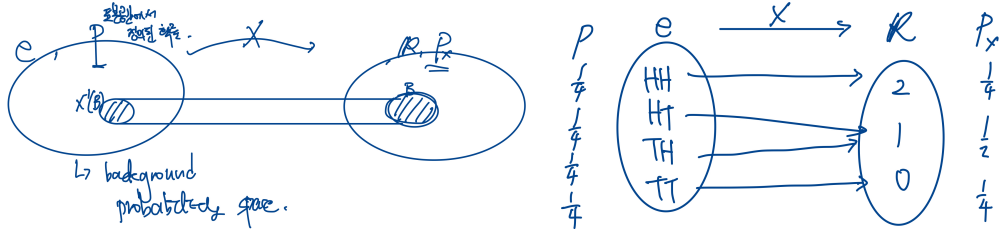
(2) (X 의 공간)

$$\begin{aligned} X \text{의 치역} &= X \text{의 공간} \\ &= \{X(w) : w \in e\} \end{aligned}$$

(3) (X 의 분포) $\mathbb{R}$ 상에 정의된 확률 P_X

$$P_X(B) = P(X^{-1}(B)) = P(\{w \in e : X(w) \in B\}) = P(X \in B), \forall B \subseteq \mathbb{R}$$

Example. X : 동전을 두 번 던져서 앞면이 나온 횟수



3.2 누적분포함수

Definition. X 의 누적분포함수(cumulative distribution function, cdf)는 $F_X : \mathbb{R} \rightarrow [0, 1]$ 인 함수로

$$F_X(x) := \mathbb{P}(X \leq x), \forall x \in \mathbb{R}$$

와 같이 정의된다.

Theorem. (특성화 정리) $F : \mathbb{R} \rightarrow \mathbb{R}$ 이 어떤 확률변수의 cdf

- $\iff$ (i) F 는 증가(nondecreasing) 함수
(ii) $\lim_{x \rightarrow -\infty} F(x) = 0; \lim_{x \rightarrow \infty} F(x) = 1$
(iii) F 는 우측연속. 즉, $\lim_{x \searrow x_0} F(x) = F(x_0)$

Proof. ($\Rightarrow$) $X \sim F$ 라 하자.

(i) $x_1 < x_2$ 일 때 $F(x_1) \leq F(x_2)$ 를 보이하고자

$$F(x_2) = \mathbb{P}(X \leq x_2) \geq \mathbb{P}(X \leq x_1) = F(x_1) \quad (\because [X \leq x_1] \subseteq [X \leq x_2])$$

Note. $[X \leq x_1] = \{w \in e : X(w) \leq x_1\}$

(ii) $x_n \nearrow \infty$ 일 때, $F(x_n) \rightarrow 1$ 을 보이하고자

$$\begin{aligned} \lim_{n \rightarrow \infty} F(x_n) &= \lim_{n \rightarrow \infty} \mathbb{P}(X \leq x_n) \\ &= \mathbb{P}(X \in \mathbb{R}) \quad (\because [X \leq x_n] \text{는 증가, 확률측도의 연속성에 의해 } \bigcup_n [X \leq x_n] = [X \in \mathbb{R}]) \end{aligned}$$

(iii) $x_n \searrow x_0$ 일 때, $F(x_n) \rightarrow F(x_0)$ 를 보이하고자

$$\begin{aligned} \lim_{n \rightarrow \infty} F(x_n) &= \lim_{n \rightarrow \infty} \mathbb{P}(X \leq x_n) \\ &= \mathbb{P}(X \leq x_0) \quad (\because [X \leq x_n] \text{은 감소, } \bigcap_n [X \leq x_n] = [X \leq x_0]) \end{aligned}$$

($\Leftarrow$) F 가 3가지 조건을 만족하면 F 를 cdf로 갖는 확률변수를 만들어낼 수 있음을 보여야 한다.

(sketch) 간단한 확률 공간 ($e = [0, 1], P$)를 생각하자.

$P(A) := \int_A dx, A \subseteq [0, 1]$ 로 정의하면 $P(A)$ 는 확률이 된다.

이제 확률변수를 다음과 같이 정의하자. $X(w) = \inf\{x \in \mathbb{R} : F(x) \geq w\}$

F 가 연속이고 순증가임을 가정하면,

$$\mathbb{P}(X \leq x) = \mathbb{P}(w \in [0, 1] : X(w) \leq x) = \mathbb{P}(w \in [0, 1] : w \leq F(x)) = F(x)$$

■

Theorem. 두 개의 확률변수 X_1 과 X_2 가 동일한 누적분포함수를 갖는다면, X_1 과 X_2 의 분포는 동일하다.

Proof. 증명은 생략한다.

■

* X_1 과 X_2 가 동일한 누적분포함수를 갖는다:

$$\begin{aligned} F_{X_1} = F_{X_2} &\Leftrightarrow F_{X_1}(x) = F_{X_2}(x), \forall x \in \mathbb{R} \\ &\Leftrightarrow \mathbb{P}(X_1 \leq x) = \mathbb{P}(X_2 \leq x), \forall x \\ &\Leftrightarrow P_{X_1} \text{과 } P_{X_2} \text{가 } (-\infty, x], x \in \mathbb{R} \text{ 상에서 같다.} \\ &\Leftrightarrow P_{X_1}(-\infty, x] = P_{X_2}(-\infty, x], \forall x \end{aligned}$$

* X_1 과 X_2 의 분포가 같다:

$$P_{X_1} = P_{X_2} \Leftrightarrow P_{X_1}(B) = P_{X_2}(B), \forall B \subseteq \mathbb{R}$$

Note. $P_{X_1}(B) = \mathbb{P}(X \in B), B \subseteq \mathbb{R}$

Theorem. F 는 확률변수 X 의 cdf.

- (a) $0 \leq F(x) \leq 1, \forall x \in \mathbb{R}$
- (b) $\mathbb{P}(a < X \leq b) = F(b) - F(a), a < b$
- (c) $\forall x, F$ 의 좌극한이 존재한다. 즉, $F(x_0^-) = \lim_{x \nearrow x_0} F(x) = \mathbb{P}(X < x_0)$
- (d) $\mathbb{P}(X = x) = F(x) - F(x^-)$

Proof.

- (a) $F(x)$ 는 어떤 사건의 확률이므로 성립한다.
- (b) $\mathbb{P}(a < x \leq b) = \mathbb{P}(X \leq b) - \mathbb{P}(X \leq a) = F(b) - F(a)$
- (c) $x_n \nearrow x_0$ 라 하자.

$$\begin{aligned} F(x_0^-) &= \lim_{n \rightarrow \infty} F(x_n) = \lim_{n \rightarrow \infty} \mathbb{P}(X \leq x_n) \\ &= \mathbb{P}(X < x_0) \quad (\because \bigcup (-\infty, x_n] = (-\infty, x_0)) \end{aligned}$$

Note. $F(x_0^+) = F(x_0)$

■

3.3 이산형 확률변수

Definition.

(i) 확률변수 X 가 $X : (e, \mathbb{P}) \rightarrow (\mathcal{D}, P_X)$ 인 함수이고, $\mathcal{D}$ 가 가산(countable)이면 X 를 **이산형(discrete)** 확률변수라 한다.

(ii) $\mathcal{D} = \{d_1, d_2, \dots\}$ 일 때, X 의 **확률밀도함수**(밀도함수, 확률질량함수, probability density/mass function)는 다음과 같이 정의된다.

$$P_X(x) := \mathbb{P}(X = x), x \in \mathcal{D}$$

Theorem.

(i) $\forall B \subseteq \mathbb{R}$ 에 대해, $\mathbb{P}(X \in B) = \sum_{x \in B} P_X(x)$

(ii) X_1 과 X_2 가 이산형 확률변수들이고, 밀도함수들이 같다면, 즉, $p_{X_1} = p_{X_2}$ 이라면 X_1 과 X_2 의 분포는 같다. 즉, $P_{X_1} = P_{X_2}$ 이다.

3.4 연속형 확률변수

Definition. X : 확률변수, $\mathbb{P}(X \in B) = \int_B f(x)dx, \forall B \in \mathbb{R}$ 을 만족하는 f 를 X 의 **밀도함수**라 한다. 이 때, X 를 **연속형 확률변수**라 한다.

Theorem. X : 연속형 확률변수, f : X 의 밀도함수(pdf), F : X 의 cdf라 하자.

(i) $F(x) = \int_{-\infty}^x f(t)dt$

(ii) $F'(x) = f(x), \forall x \in C_f$ (f 의 연속점들의 집합)

Proof.

(i) $F(x) = \mathbb{P}(X \leq x) = \int_{-\infty}^x f(t)dt$

(ii) 미적분의 기본정리에 의해 성립한다. 즉, pdf가 미분가능한 점에서 cdf를 구할 수 있다. ■

Example. (이산형도 연속형도 아닌 확률변수)

$u = [0, 1]$ 에서 랜덤하게 뽑은 수, $X := \begin{cases} 1 & \text{동전이 앞면} \\ u, & \text{동전이 뒷면} \end{cases}$ 이라고 하면,

$$\begin{aligned} F_X(x) &= \mathbb{P}(X \leq x) = \mathbb{P}(X \leq x|H) \cdot \mathbb{P}(H) + \mathbb{P}(X \leq x|T) \cdot \mathbb{P}(T) \\ &= \begin{cases} 0, & x < 0 \\ \frac{x}{2}, & 0 < x < 1 \\ 1, & x \geq 1 \end{cases} \end{aligned}$$

이므로 X 는 연속형도($\because$ 적분 불가능), 이산형도 아니다.

4 확률변수의 기대값

4.1 기대값의 정의

Definition. $X \sim f$ (pdf), $\sum |x|f(x) < \infty$ (이산형), $\int |x|f(x)dx < \infty$ (연속형)

$$\mathbb{E}X := \begin{cases} \sum_x xf(x), & X : \text{이산형} \\ \int xf(x)dx, & X : \text{연속형} \end{cases}$$

Note. 기호상 편의를 위해 $\mathbb{E}X = \int x dF(x) = \int x dP(x)$ 혹은 $\mathbb{E}X = \int x F(dx) = \int x P(dx)$ 라고 쓰며, pdf가 존재하는 경우 $F(dx) = f(x)dx$ 이다.

Example. (기대값의 정의에서 피적분함수의 절대값이 적분가능하기를 요구하는 이유)

적분 $\int_{-\infty}^{\infty} h(x)dx$ 이 혼란 없이 정의되기 위해서는 $\int |h(x)|dx < \infty$ 의 조건이 필요하다.

$\int_{-\infty}^{\infty} \frac{x}{1+x^2} dx$ 를 고려해 보자.

$$\begin{aligned} \int_0^{\infty} \frac{x}{1+x^2} dx &= \frac{1}{2} \int_0^{\infty} \frac{1}{1+t} dt \quad (x^2 = t) \\ &= \frac{1}{2} \left| \log(1+t) \right|_0^{\infty} = \infty \end{aligned}$$

따라서 $\int_{-\infty}^{\infty} \left| \frac{x}{1+x^2} \right| dx = \infty$ 이다.

한편, $\int_{-\infty}^{\infty} \frac{x}{1+x^2} dx = \lim_{M \rightarrow \infty} \int_{-M}^M \frac{x}{1+x^2} dx = \lim_{M \rightarrow \infty} 0 = 0$ 이고,

$\int_{-\infty}^{\infty} \frac{x}{1+x^2} = \int_0^{\infty} \frac{x}{1+x^2} + \int_{-\infty}^0 \frac{x}{1+x^2} dx = \infty - \infty$ 이므로 적분이 정의되지 않는다.

Definition. $\int |g(x)|dF(x) < \infty$ 일 때,

$$\begin{aligned} \mathbb{E}g(x) &= \int g(x)dF(x) \\ &= \begin{cases} \sum g(x)f(x), & X : \text{이산형} \\ \int g(x)f(x)dx, & X : \text{연속형} \end{cases} \end{aligned}$$

Proof. (이산형인 경우)

$Y = g(X)$ 라 하자.

$$\begin{aligned} \mathbb{E}Y &= \sum_y y f_Y(y), \quad (f_Y : y \text{의 pdf}) \\ &= \sum_y \sum_{x:g(x)=y} g(x) \mathbb{P}(g(x)=y) = \sum_y \sum_{x:g(x)=y} g(x) \mathbb{P}(X : g(x)=y) \\ &= \sum_x g(x) \mathbb{P}(X=x) \end{aligned}$$

■

Theorem. $a, b \in \mathbb{R}, X : \text{확률변수}, k_1, k_2 \text{는 상수. 기대값들이 존재할 때,}$

- (a) $\mathbb{E}(aX + b) = a\mathbb{E}X + b$
- (b) (선형성) $\mathbb{E}(k_1 g_1(X) + k_2 g_2(X)) = k_1 \mathbb{E}g_1(X) + k_2 \mathbb{E}g_2(X)$
- (c) (양부호, positivity) $g(x) \geq 0, \forall x \implies \mathbb{E}g(X) \geq 0$

Proof. 적분의 성질에 의해 성립한다.

■

4.2 확률분포의 특성치

Definition.

(a) 평균: $\mathbb{E}X$ (b) 분산: $Var(X) = \mathbb{E}(X - \mathbb{E}X)^2$ (c) 표준편차: $Sd(X) = \sqrt{Var(X)}$ **Theorem.** $Var(X) < \infty, \mu = \mathbb{E}X, a \in \mathbb{R}$.(1) $Var(X) \geq 0$ (2) $Var(X) = 0 \Leftrightarrow \mathbb{P}(X = \mu) = 1$ (3) $Var(X) = \mathbb{E}X^2 - (\mathbb{E}X)^2$ (4) $Var(aX + b) = a^2 Var(X), Sd(aX + b) = |a| Sd(X)$ *Proof.* (1), (4)는 분산의 정의에 의해 성립하고, (3)은 기대값의 linearity에 의해 성립.

(2)의 왼쪽 방향은 당연하다. 오른쪽 방향을 보이겠다

$$\begin{aligned}
0 = Var(X) &= \mathbb{E}(X - \mathbb{E}X)^2 \\
&\geq \mathbb{E}[(X - \mu)^2 I(|X - \mu| \geq \frac{1}{n})] \quad n = 1, 2, \dots \\
&\geq \mathbb{E}(\frac{1}{n^2} I(|X - \mu| \geq \frac{1}{n})) \\
&= \frac{1}{n^2} \mathbb{P}(|X - \mu| \geq \frac{1}{n})
\end{aligned}$$

따라서 $\mathbb{P}(|X - \mu| \geq \frac{1}{n}) = 0 \quad \forall n = 1, 2, \dots$ 이다.

$$\begin{aligned}
\mathbb{P}(X = \mu) &= \mathbb{P}(|X - \mu| = 0) \\
&= \mathbb{P}(\bigcap_{n=1}^{\infty} [|X - \mu| < \frac{1}{n}]) \\
&= \lim_{n \rightarrow \infty} \mathbb{P}(|X - \mu| < \frac{1}{n}) \\
&= \lim_{n \rightarrow \infty} 1 = 1
\end{aligned}$$

■

Example. (기대값이 존재하지 않는 경우) $X \sim Ca(0, 1)$ (코쉬 분포)

$$\begin{aligned}
\Leftrightarrow f(x) &= \frac{1}{\pi} \frac{1}{1 + x^2}, \quad x \in \mathbb{R} \\
\mathbb{E}|X| &= \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{|x|}{1 + x^2} dx = \infty
\end{aligned}$$

Definition. (비대칭도, 뽀족한 정도) $\mu = \mathbb{E}X, \sigma^2 = Var(X)$ (1) (비대칭도, skewness, 왜도) $\mathbb{E}(\frac{X - \mu}{\sigma})^3$ (2) (뽀족한 정도, kurtosis, 첨도) $\mathbb{E}(\frac{X - \mu}{\sigma})^4$ 혹은 $\mathbb{E}(\frac{X - \mu}{\sigma})^4 - 3$

Note. skewed to the left = negative skewness, skewed to the right = positive skewness: 꼬리가 치우쳐진 쪽을 기준으로 한다.

4.3 적률, 적률생성함수

Definition.

- (a) (r 번째 적률, moment) $m_r = \mathbb{E}X^r$ ($\mathbb{E}|X|^r < \infty$), $r = 1, 2, \dots$
- (b) (적률생성함수, moment generating function, mgf) 어떤 $h < \infty$ 에 대해, $\mathbb{E}e^{tX} < \infty$, $\forall t \in (-h, h)$ 일 때, $M_X(t) := \mathbb{E}e^{tX}$ 를 X 의 **적률생성함수**라 한다.

Theorem. (mgf의 성질들)

- (1) (적률생성함수는 적률을 생성한다.) 확률변수 X 의 mgf $M_X(t)$ 가 존재하면,
- (a) 모든 차수의 적률이 존재한다. 즉, $\mathbb{E}|X|^r < \infty$, $r = 1, 2, \dots$
 - (b) $\mathbb{E}X^r = M_X^{(r)}(0)$, $r = 1, 2, \dots$
 - (c) 어떤 $\epsilon > 0$ 에 대해 $M_X(t) = \sum_{r=0}^{\infty} \frac{\mathbb{E}X^r}{r!} t^r$, $|t| < \epsilon$
- (2) (적률생성함수의 유일성) 만약 0 근방에서 X 와 Y 의 mgf가 같으면, $X \stackrel{d}{=} Y$ (X 와 Y 의 분포가 같다.)
 $\iff \exists h > 0$ s.t. $M_X(t) = M_Y(t), \forall t \in (-h, h) \implies P(X \in B) = P(Y \in B), \forall B \subseteq \mathbb{R}$

Proof. (2)는 받아들이자. (특성함수를 이용해 증명한다.)(1-a) $r = 1, 2, \dots$, $\mathbb{E}|X|^r < \infty$ 을 보이려고자

$$\begin{aligned}
 e^{|tx|} &= \sum_{k=0}^{\infty} \frac{|tx|^k}{k!} \\
 \frac{|tx|^r}{r!} &\leq e^{|tx|} \leq e^{tx} + e^{-tx} \\
 |x|^r &\leq \frac{r!}{|t|^r} (e^{tx} + e^{-tx}) \\
 \mathbb{E}|X|^r &\leq \frac{r!}{|t|^r} (\mathbb{E}e^{tx} + \mathbb{E}e^{-tx}) = M(t) + M(-t) < \infty \quad t \in (-h, h)
 \end{aligned}$$

(1-b) $M_X(t) = \sum_{r=0}^{\infty} \frac{\mathbb{E}X^r}{r!} t^r$ 을 보이려고자 $\lim_{n \rightarrow \infty} \sum_{r=0}^{n-1} \frac{\mathbb{E}X^r}{r!} t^r = M_X(t)$ 을 보이면 된다.

즉, $|M_X(t) - \sum_{r=0}^{n-1} \frac{\mathbb{E}X^r}{r!} t^r| \rightarrow 0$ 을 보이면 된다.

Note. (테일러 정리) 어떤 $h > 0$ 에 대해, f 가 $(a-h, a+h)$ 에서 n 번 연속적으로 미분가능하면

$$\begin{aligned}
 f(x) &= \sum_{r=0}^{n-1} \frac{f^{(r)}(a)}{r!} (x-a)^r + R_n(x, a) \\
 \text{where } R_n(x, a) &= \frac{(x-a)^n}{(n-1)!} \int_0^1 (1-u)^{n-1} f^{(n)}(a+u(x-a)) du
 \end{aligned}$$

테일러 정리를 지수함수에 적용하면, $e^{tx} = \sum_{r=0}^{n-1} \frac{(tx)^r}{r!} + \frac{(tx)^n}{(n-1)!} \int_0^1 (1-u)^{n-1} e^{utx} du$ 이므로

$$|M_X(t) - \sum_{r=0}^{n-1} \frac{\mathbb{E}(tX)^r}{r!}| = |\mathbb{E}(e^{tX} - \sum_{r=0}^{n-1} \frac{(tX)^r}{r!})| \text{ 이고,}$$

$$\begin{aligned}
|e^{tx} - \sum_{r=0}^{n-1} \frac{(tX)^r}{r!}| &= \left| \frac{(tx)^n}{(n-1)!} \int_0^1 (1-u)^{n-1} e^{utx} du \right| \\
&\leq \frac{|tx|^n}{(n-1)!} \int_0^1 (1-u)^{n-1} e^{|tx|} du \\
&\leq \frac{|tx|^n}{(n-1)!} |tx| e^{|tx|} \int_0^1 (1-u)^{n-1} du \\
&\leq e^{|tx|} \cdot e^{|tx|} \cdot e^{|tx|} \cdot \frac{1}{n} = \frac{1}{n} e^{3|tx|}
\end{aligned}$$

(1-c) 해석함수 $M_X(t) = \sum_{r=0}^{\infty} \frac{M_X^{(r)}(0)}{r!} t^r = \sum_{r=0}^{\infty} \frac{\mathbb{E}X^r}{r!} t^r$ 이므로 $M_X^{(r)}(0) = \mathbb{E}X^r, \forall r = 1, 2, \dots$

Theorem. $M_{aX+b}(t) = e^{bt} M_X(at), \quad a, b \in \mathbb{R}$

Proof. $M_{aX+b}(t) = \mathbb{E}e^{t(aX+b)} = e^{tb} M_X(at)$

Example. (정규분포) $X \sim N(0, 1), f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$ 1 $x \in \mathbb{R}$

$$\begin{aligned}
M_X(t) &= \mathbb{E}e^{tX} = \int e^{tx} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx \\
&= \frac{1}{\sqrt{2\pi}} \int e^{-\frac{1}{2}x^2+tx} dx = \frac{1}{\sqrt{2\pi}} \int e^{-\frac{1}{2}(x^2-2tx+t^2)+\frac{1}{2}t^2} dx \\
&= \frac{1}{\sqrt{2\pi}} \int e^{-\frac{1}{2}(x-t)^2} dx \cdot e^{\frac{1}{2}t^2} = e^{\frac{1}{2}t^2}, \quad t \in \mathbb{R}
\end{aligned}$$

$$e^{\frac{1}{2}t^2} = \sum_{r=0}^{\infty} \frac{(\frac{1}{2}t^2)^r}{r!} = \sum_{r=0}^{\infty} \frac{t^{2r}}{2^r r!} = \sum_{r=0}^{\infty} \frac{(2r)!}{(2r)!} t^{2r} \text{ 이므로 } \begin{cases} \mathbb{E}X^{2r} = \frac{(2r)!}{2^r r!}, & r = 1, 2, \dots \\ \mathbb{E}X^{2r+1} = 0, & r = 0, 1, 2, \dots \end{cases}$$

참고: mgf가 존재하면 모든 차수의 적률이 존재한다(1-a). 역은 성립하지 않는다.

Example. (로그정규분포) $\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(\log x)^2} \frac{1}{x}, \quad x > 0$

로그정규분포의 적률생성함수는 존재하지 않지만, 모든 적률은 존재한다.

4.4 누울생성함수

Definition. (누울, 누울생성함수)

(a) 확률변수 X 의 mgf가 $|t| < h$ 에서 존재할 때 ($h < 0$),

$$C(t) := \log M_X(t), \forall t : |t| < h$$

를 X 의 누울생성함수(cumulant generating function, cgf)라 한다.

¹ $f(x)$ 가 확률밀도함수가 되는가?: 적분해서 1이 되는지를 확인한다

(b) $C(t) = \sum_{r=0}^{\infty} \frac{C^{(r)}(0)}{r!} t^r = \sum_{r=0}^{\infty} \frac{C_r(X)}{r!} t^r$, $|t| < h$ 와 같이 표현될 때, $C_r(X)$ 를 X 의 r 번째 누울이라 한다.

Example. $X \sim N(0, 1)$ 일 때, $M_X(t) = e^{\frac{1}{2}t^2}$ 이고 $C_X(t) = \log M_X(t) = \frac{1}{2}t^2$
따라서 $C_0(X) = 0, C_1(X) = 0, C_2(X) = 1, C_3(X) = C_4(X) = \dots = 0$

Theorem. (적률과 누울의 관계) X 의 mgf가 존재한다고 하자. m_r : X 의 r 차 적률, c_r : X 의 r 차 누울

$$m_0 = 1, c_0 = 0$$

$$m_k = \sum_{j=0}^{k-1} \binom{k-1}{j} c_{k-j} m_j, \quad k \geq 1$$

$$c_k = m_k - \sum_{j=0}^{k-1} \binom{k-1}{j} c_{k-j} m_j, \quad k \geq 1$$

Proof. 해석함수는 멱급수로 표현가능하다.

(Apostol, 정리 9.25.) $f(z) = \sum_{n=0}^{\infty} a_n z^n (|z| < r)$, $g(z) = \sum_{n=0}^{\infty} b_n z^n (|z| < R)$ 이라고 하자.

$|z| < R$ 이고 $\sum_{n=0}^{\infty} |b_n z^n| < r$ 일 때, $f(g(z)) = \sum_{k=0}^{\infty} c_k z^k$

여기서 $c_k = \sum_{n=0}^{\infty} a_n b_k(n)$, $g(z)^n = \left(\sum_{k=0}^{\infty} b_k z^k\right)^n = \sum_{k=0}^{\infty} b_k(n) z^k$

$$P(s) = \sum_{k=0}^{\infty} p_k s^k$$

$$Q(s) = e^{P(s)} = \sum_{k=0}^{\infty} q_k s^k \text{일 때,}$$

$$q_0 = e^{p_0}$$

$$q_k = \frac{1}{k} \sum_{j=0}^{k-1} (k-j) p_{k-j} q_j, \quad k \geq 1$$

■

Example.

$$m_0 = 1$$

$$m_1 = \sum_{j=0}^0 \binom{0}{j} c_{1-j} m_j = c_1 m_0 = c_1$$

$$m_2 = \sum_{j=0}^1 \binom{1}{j} c_{2-j} m_j = c_2 m_0 + c_1 m_1 = c_2 \cdot 1 + c_1 \cdot c_1 = c_2 + c_1^2$$

$$m_3 = \sum_{j=0}^2 \binom{2}{j} c_{3-j} m_j = c_3 m_0 + 2c_2 m_1 + c_1 m_2 = c_3 + 2c_2 c_1 + c_1(c_2 + c_1^2) = c_3 + 3c_1 c_2 + c_1^3$$

Definition.

(a) (확률생성함수) X : 음이 아닌 정수 값을 갖는 이산형 확률변수라고 하자. 이 때 확률생성함수

$$G(s) = \mathbb{E}s^X, s \in (-1, 1)$$

$$= \sum_{x=0}^{\infty} s^x \mathbb{P}(X = x)$$

(b) (라플라스 변환) $X \geq 0$ w.p.

$$L(t) = \mathbb{E}e^{-tX}, \quad t \geq 0$$

(c) (특성함수)

$$\phi_X(t) = \mathbb{E}e^{itX}, \quad t \in \mathbb{R}$$

Note. $L(t)$ 와 $\phi_X(t)$ 는 항상 존재한다. ($\because e^{-t}, e^{it} \leq 1$)

5 부등식

5.1 볼록함수

Definition.

(a) 함수 $\phi: (a, b) \rightarrow \mathbb{R}$ 가

$$\phi(\alpha x + (1 - \alpha)y) \leq \alpha\phi(x) + (1 - \alpha)\phi(y) \quad \forall a < x < y < b, \alpha \in (0, 1)$$

를 만족하면 ϕ 를 **볼록함수(convex function)**라 한다. 위의 부등식이 등호 없는 부등식을 만족하면 ϕ 를 **순볼록(strictly convex)**라 한다.

(b) $-\phi$ 가 볼록(순볼록)이면 ϕ 를 오목(순오목)이라 한다.

Theorem. $t = \alpha x + (1 - \alpha)y$ 라고 하면,

$$\begin{aligned} \phi(t) \leq \alpha\phi(x) + (1 - \alpha)\phi(y) &\iff \frac{\phi(t) - \phi(x)}{t - x} \leq \frac{\phi(y) - \phi(x)}{y - x} \\ &\iff \frac{\phi(t) - \phi(x)}{t - x} \leq \frac{\phi(y) - \phi(t)}{y - t} \\ &\iff \frac{\phi(y) - \phi(x)}{y - x} \leq \frac{\phi(y) - \phi(t)}{y - t} \end{aligned}$$

Theorem.

1. ϕ 가 (a, b) 에서 미분가능

$$(a) \phi \text{가 볼록} \iff \phi'(x) \leq \phi'(y) \quad \forall a < x < y \leq b$$

$$(b) \phi \text{가 순볼록} \iff \phi'(x) < \phi'(y) \quad \forall a < x < y \leq b$$

2. ϕ 가 (a, b) 에서 두 번 미분가능

$$(a) \phi \text{가 볼록} \iff \phi''(x) \geq 0 \quad \forall x \in (a, b)$$

$$(b) \phi \text{가 순볼록} \iff \phi''(x) > 0 \quad \forall x \in (a, b)$$

Proof. (1-a)

■

Example.

(1) $(e^x)' = e^x$: 증가, $\forall x > 0 \implies e^x$: 순볼록

(2) $(\log x)' = \frac{1}{x}$: 감소, $\forall x > 0 \implies \log x$: 순오목

Definition. (산술평균과 기하평균)

(a) $a^{\frac{1}{2}}b^{\frac{1}{2}} \leq \frac{1}{2}a + \frac{1}{2}b, \quad a, b > 0$

(b) $a^\alpha b^\beta \leq \alpha a + \beta b, \quad a, b > 0, \alpha + \beta = 1, \alpha, \beta > 0$

(c) $cd \leq \frac{1}{p}c^p + \frac{1}{q}d^q, \quad \frac{1}{p} + \frac{1}{q} = 1, p, q > 1, c, d > 0$

Proof. (a)는 (b)의 특별한 경우이다.

(b) (b의 부등식)

$$\begin{aligned} &\iff \left(\frac{a}{a+b}\right)^\alpha \left(\frac{b}{a+b}\right)^\beta \leq \alpha \frac{a}{a+b} + \beta \frac{b}{a+b} \\ &\iff \alpha \log \frac{a}{a+b} + \beta \log \frac{b}{a+b} \leq \log\left(\alpha \frac{a}{a+b} + \beta \frac{b}{a+b}\right) \quad (\because \log: \text{순오목}) \end{aligned}$$

$$(c) \begin{cases} a^\alpha \rightarrow c \\ b^\beta \rightarrow d \\ \alpha \rightarrow \frac{1}{p} \\ \beta \rightarrow \frac{1}{q} \end{cases}, \text{ (b)를 이용하면 } cd \leq \frac{1}{p}c^p + \frac{1}{q}d^q$$

■

Theorem. (리아푸노프 부등식, Liapounov inequality)

$$\mathbb{E}|X|^s \implies (\mathbb{E}|X|^r)^{\frac{1}{r}} \leq (\mathbb{E}|X|^s)^{\frac{1}{s}}, \quad 0 < r < s$$

따라서 $\mathbb{E}|X|^r < \infty$.

Note 1. $\|X\|_r = (\mathbb{E}|X|^r)^{\frac{1}{r}}$: X 의 Lr -노름. 즉, $\|X\|_r \leq \|X\|_s$

Note 2. 위 부등식에 의하면, s 번째 적률이 존재하면 $r(\leq s)$ 번째 적률은 존재한다.

Proof. $c = \frac{|X|^r}{(\mathbb{E}|X|^s)^{r/s}}, d = 1, p = \frac{s}{r}, \frac{1}{q} = 1 - \frac{1}{p} = \frac{s-r}{s}$ 이라 두고 위 Definition의 (c)에 대입하면

$$\begin{aligned} \frac{|X|^r}{(\mathbb{E}|X|^s)^{r/s}} \cdot 1 &\leq \frac{r}{s} \left(\frac{|X|^r}{(\mathbb{E}|X|^s)^{r/s}} \right)^{s/r} + \frac{s}{s-r} \cdot 1^{\frac{s-r}{s}} \\ \frac{|X|^r}{(\mathbb{E}|X|^s)^{r/s}} &\leq \frac{r}{s} \frac{|X|^s}{\mathbb{E}|X|^s} + \frac{s}{s-r} \\ \frac{\mathbb{E}|X|^r}{(\mathbb{E}|X|^s)^{r/s}} &\leq 1 \\ \mathbb{E}(|X|^r)^{1/r} &\leq \mathbb{E}(|X|^s)^{1/s} \end{aligned}$$

■

Theorem. (마르코프 부등식, 체비셰프 부등식)

(a) (Markov inequality) $X \geq 0$ w.p.1., $c > 0$

$$\mathbb{P}(X \geq c) \leq \frac{\mathbb{E}X}{c}$$

(b) $\mathbb{P}(|Z| \geq c) \leq \frac{\mathbb{E}|Z|^r}{c^r}, \quad c, r > 0$

(c) (Chebyshev inequality)

$$\mathbb{P}(|X - \mathbb{E}X| \geq k\sigma) \leq \frac{1}{k^2}, \quad k > 0, \quad \sigma^2 = \text{Var}(X)$$

Proof.

(a) X 의 밀도함수 f 가 존재한다고 가정하자.

$$\begin{aligned} \mathbb{E}X &= \int_0^\infty xf(x)dx \\ &\geq \int_c^\infty xf(x)dx \quad (\because x, f(x) \geq 0) \\ &\geq \int_c^\infty cf(x)dx \quad (\because c < x) \\ &= c\mathbb{P}(X \geq c) \end{aligned}$$

밀도함수의 존재를 가정하지 않는 경우,

$$\begin{aligned} \mathbb{E}X &\geq \mathbb{E}[XI_{(X \geq c)}] \\ &\geq \mathbb{E}[cI_{(X \geq c)}] \\ &= c\mathbb{P}(X \geq c) \end{aligned}$$

(b) $\mathbb{P}(|Z| \geq c) = \mathbb{P}(|Z|^r \geq c^r) \leq \frac{\mathbb{E}|Z|^r}{c^r} \quad (\because (a))$

(c) $\mathbb{P}(|X - \mathbb{E}X| \geq k\sigma) \leq \frac{\mathbb{E}|X - \mathbb{E}X|^2}{k^2\sigma^2} = \frac{1}{k^2}$

■

Note. 체비셰프의 부등식에 의하면, 1σ 밖의 확률 ≤ 1 , 2σ 밖의 확률 $\leq \frac{1}{4}$, ... (X 의 분포가 무엇이든)

Theorem. (젠센의 부등식, Jensen's inequality) ϕ 는 I 상에서 볼록, I 는 열린 구간이라고 하자.

$$\mathbb{E}|X| < \infty, \quad X \in I \text{ w.p.1.} \implies \phi(\mathbb{E}X) \leq \mathbb{E}\phi(X)$$

Proof. ϕ'' 가 존재한다고 가정하자(두 번 미분가능). $\mu = \mathbb{E}X$ 라 하면,

$$\begin{aligned} \phi(x) &= \phi(\mu) + \phi'(\mu)(x - \mu) + \frac{\phi''(x^*)}{2}(x - \mu)^2, x^* \text{는 } x \text{와 } \mu \text{ 사이값} \\ &\geq \phi(\mu) + \phi'(\mu)(x - \mu), \quad \forall x \in I \\ \mathbb{E}\phi(x) &\geq \phi(\mu) + \phi'(\mu)\mathbb{E}(X - \mu) = \phi(\mu) \end{aligned}$$

Note. 일반적인 증명은 임의의 점에서 그은 접선이 $f(x)$ 보다 아래에 있음을 증명한다.

■

Example.

- $e^{\mathbb{E}X} \leq \mathbb{E}e^X$
- $\log(\mathbb{E}X) \geq \mathbb{E}(\log X)$

2장. 다차원 확률변수의 확률분포

Jiho KWAK, Department of Geography, SNU

April 12, 2024

1 두 확률변수의 분포

1.1 용어들

- 함수 $(X_1, X_2) : e \rightarrow \mathbb{R}^2$ 를 확률벡터라 한다.
- (X_1, X_2) 의 공간 = (X_1, X_2) 의 치역
- (X_1, X_2) 의 확률분포: $\mathbb{R}^2$ 상의 확률
- $P_{(X_1, X_2)}(B) = \mathbb{P}((X_1, X_2) \in B) = \mathbb{P}(w \in e : (X_1(w), X_2(w)) \in B), B \subseteq \mathbb{R}^2$

1.2 결합확률분포, 주변확률분포

Definition. (결합누적분포함수) (X_1, X_2) 의 **누적분포함수**는 $\mathbb{R}^2$ 상의 함수

$$F_{X_1, X_2}(x_1, x_2) = \mathbb{P}(X_1 \leq x_1, X_2 \leq x_2), \quad \forall x_1, x_2 \in \mathbb{R}$$

Theorem. F : X_1, X_2 의 cdf

(a) $\tilde{a} = (a_1, a_2), \tilde{b} = (b_1, b_2)$

$$\triangle_{(\tilde{a}, \tilde{b}]} F := F(b_1, b_2) - F(a_1, b_2) - F(a_2, b_1) + F(a_1, a_2) \geq 0, \quad \forall a_1 \leq a_2, b_1 \leq b_2$$

(b) (각 인수에 대해 증가함수)

$$F(x_1, x_2) \leq F(y_1, x_2), \quad \forall x_1 < y_1, x_2$$

$$F(x_1, x_2) \leq F(x_1, y_2), \quad \forall x_2 < y_2, x_1$$

(c) (i) $0 \leq F(x) \leq 1 \quad \forall x_1, x_2$

(ii) $\lim_{x_1, x_2 \rightarrow \infty} F(x_1, x_2) = 1$

(iii) $\lim_{x_1 \rightarrow \infty} F(x_1, x_2) = 0, \quad \forall x_2$

$\lim_{x_2 \rightarrow \infty} F(x_1, x_2) = 0, \quad \forall x_1$

(d) (우측연속성) $x_n = (x_{n1}, x_{n2}) \downarrow x_0 = (x_{01}, x_{02}), \text{ 즉 } x_{n1} \downarrow x_{01}, x_{n2} \downarrow x_{02} \implies F(x_n) \downarrow F(x_0)$

(e) $\lim_{x_2 \rightarrow \infty} F(x_1, x_2) = \mathbb{P}(X_1 \leq x_1), \quad \forall x_1$

$\lim_{x_1 \rightarrow \infty} F(x_1, x_2) = \mathbb{P}(X_2 \leq x_2), \quad \forall x_2$

(f) $F : \mathbb{R}^2 \rightarrow \mathbb{R}$ 가 (a), (b), (c), (d)를 만족하면 F 는 어떤 확률변수 (X_1, X_2) 의 cdf.

Proof. (a) $\Delta_{(a,b]} = P(a_1 < x_1 \leq b_1, a_2 < x_2 \leq b_2)$

(b) $F(x_1, x_2) = P(X_1 \leq x_1, X_2 \leq x_2) \leq P(X_1 \leq y_1, X_2 \leq y_2) = F(y_1, y_2)$

(c) (ii) $x_n = (x_{n1}, x_{n2})$ 라고 두면 $x_{n1} \uparrow \infty, x_{n2} \uparrow \infty$

$\implies F(x_{n1}, x_{n2}) = P(X_1 \leq x_{n1}, X_2 \leq x_{n2}) \uparrow P(X_1 \in \mathbb{R}, X_2 \in \mathbb{R})$ ($\because$ 확률의 연속성)

■

Note. (f)의 의미: cdf를 구하면 확률분포와 일대일 대응 가능

Definition.

(i) $P((X_1, X_2) \in B) = \iint_B f(x_1, x_2) dx_1 dx_2, \quad \forall B \subseteq \mathbb{R}^2$ 이면 f 를 (X_1, X_2) 의 **결합확률밀도함수**라 한다.

이때 (X_1, X_2) 는 연속형.

(ii) $P((X_1, X_2) \in B) = \sum_{(x_1, x_2) \in \mathbb{R}^2} f(x_1, x_2), \quad \forall B \subseteq \mathbb{R}^2$ 이면 f 를 (X_1, X_2) 의 **결합확률질량함수**라 한다

이때 (X_1, X_2) 는 이산형.

Theorem.

(a) (X_1, X_2) 가 연속형일 때, $\iint_{\mathbb{R}^2} f(x_1, x_2) dx_1 dx_2 = 1$

(X_1, X_2) 가 이산형일 때, $\sum_{(x_1, x_2) \in \mathbb{R}^2} f(x_1, x_2) = 1$

(b) $f : (X_1, X_2)$ 의 pdf. f 가 (x_1, x_2) 에서 연속일 때,

$$f(x_1, x_2) = \frac{\partial^2}{\partial x_1 \partial x_2} F(x_1, x_2)$$

Proof. (a)는 당연하다.

(b)

$$\begin{aligned} \frac{\partial^2}{\partial x_1 \partial x_2} F(x_1, x_2) &= \lim_{h_1 \downarrow 0} \frac{1}{h_1} \left[\frac{\partial}{\partial x_2} F(x_1 + h_1, x_2) - \frac{\partial}{\partial x_2} F(x_1, x_2) \right] \\ &= \lim_{h_1 \downarrow 0} \frac{1}{h_1} \left[\lim_{h_2 \downarrow 0} \frac{1}{h_2} (F(x_1 + h_1, x_2 + h_2) - F(x_1 + h_1, x_2)) - \lim_{h_2 \downarrow 0} \frac{1}{h_2} (F(x_1, x_2 + h_2) - F(x_1, x_2)) \right] \\ &= \lim_{h_1 \downarrow 0} \lim_{h_2 \downarrow 0} \frac{1}{h_1 h_2} \left[F(x_1 + h_1, x_2 + h_2) - F(x_1 + h_1, x_2) - F(x_1, x_2 + h_2) + F(x_1, x_2) \right] \\ &= \lim_{h_1, h_2 \downarrow 0} \frac{1}{h_1 h_2} \Delta_{((x_1, x_2), (x_1 + h_1, x_2 + h_2))} F \\ &= \lim_{h_1, h_2 \downarrow 0} \frac{1}{h_1 h_2} P(x_1 < X_1 \leq x_1 + h_1, x_2 < X_2 \leq x_2 + h_2) \\ &= \lim_{h_1, h_2 \downarrow 0} \frac{1}{h_1 h_2} \int_{x_1}^{x_1 + h_1} \int_{x_2}^{x_2 + h_2} f(t_1, t_2) dt_1 dt_2 \end{aligned}$$

$$\frac{1}{h_1 h_2} \cdot h_1 h_2 \inf_{\substack{t_1 \in (x_1, x_1 + h_1) \\ t_2 \in (x_2, x_2 + h_2)}} f(t_1, t_2) \leq \frac{1}{h_1 h_2} \int_{x_1}^{x_1 + h_1} \int_{x_2}^{x_2 + h_2} f(t_1, t_2) dt_1 dt_2 \leq \frac{1}{h_1 h_2} \cdot h_1 h_2 \sup_{\substack{t_1 \in (x_1, x_1 + h_1) \\ t_2 \in (x_2, x_2 + h_2)}} f(t_1, t_2)$$

■

Definition. (주변분포, marginal distribution)

$$(a) \text{ (주변밀도함수)} \begin{cases} f_{X_1}(x_1) = \int_{x_2} f_{X_1, X_2}(x_1, x_2) dx_2 & (\text{연속형}) \\ f_{X_1}(x_1) = \sum_{x_2 \in \mathbb{R}} f_{X_1, X_2}(x_1, x_2) & (\text{이산형}) \end{cases}$$

$$(b) \text{ (주변누적분포함수)} \begin{cases} F_{X_1}(x_1) = \lim_{x_2 \uparrow \infty} F_{X_1, X_2}(x_1, x_2) \\ F_{X_2}(x_2) = \lim_{x_1 \uparrow \infty} F_{X_1, X_2}(x_1, x_2) \end{cases}$$

Theorem. f_{X_1, X_2} 는 (x_1, x_2) 의 pdf, F_{X_1, X_2} 는 (x_1, x_2) 의 cdf.

(a) f_{X_i} 는 X_i 의 pdf

(b) F_{X_i} 는 X_i 의 cdf

Proof. (a) f_{X_i} 이 X_i 의 pdf $\iff \mathbb{P}(X_1 \in B) = \int_B f_{X_1}(x_1) dx$, $\forall B \subseteq \mathbb{R}$ 를 보이므로

$$\begin{aligned} \mathbb{P}(X_1 \in B) &= \mathbb{P}((X_1, X_2) \in B \times \mathbb{R}) = \iint_{x_1 \in B, x_2 \in \mathbb{R}} f_{X_1, X_2}(x_1, x_2) dx_1 dx_2 \\ &= \int_{X_1 \in B} \left[\int_{X_2 \in \mathbb{R}} f_{X_1, X_2}(x_1, x_2) dx_2 \right] dx_1 \\ &= \int_{X_1 \in B} f_{X_1}(x_1) dx_1 \end{aligned}$$

(b) $\lim_{x_2 \rightarrow \infty} F(x_1, x_2) = \mathbb{P}(X_1 \leq x_1) = F_{X_1}(x_1)$ ($\because$ cdf의 성질에 의해 성립)

■

Example. $(X, Y) \sim f(x, y) = e^{-y} I(0 < x < y < \infty)$

(a) $\mathbb{P}(X + Y > 1)$

$$\iint_{x+y>1} f(x, y) dx dy = \int_0^{\frac{1}{2}} \int_{1-x}^{\infty} e^{-y} dy dx + \int_{\frac{1}{2}}^{\infty} \int_x^{\infty} e^{-y} dy dx = 2e^{-\frac{1}{2}} - e^{-1}$$

(b) $f_X(x) = e^{-x} I_{(x>0)}$, $f_Y(y) = y \cdot e^{-y} I_{y>0}$

2 결합확률분포의 특성치

2.1 기대값, 공분산, 상관계수

Definition. (X_1, X_2) 의 cdf $F_{1,2}$, 분포 $P_{1,2}$, 밀도함수 $f_{1,2}$

$$\begin{aligned} \mathbb{E}g(X_1, X_2) &:= \begin{cases} \sum_{x_1} \sum_{x_2} g(x_1, x_2) f(x_1, x_2) & (\text{이산형}) \\ \iint g(x_1, x_2) f(x_1, x_2) dx_1 dx_2 & (\text{연속형}) \end{cases} \\ &= (\gamma | \tilde{\mathcal{H}}) \begin{cases} \int g(x_1, x_2) F_{1,2}(dx_1, dx_2) = \int g(x_1, x_2) dF_{1,2}(x_1, x_2) \\ \int g(x_1, x_2) P_{1,2}(dx_1, dx_2) = \int g(x_1, x_2) dP_{1,2}(x_1, x_2) \end{cases} \end{aligned}$$

Note. 2차원 기대값과 1차원 기대값의 정의는 일치한다. $Y = g(X_1, X_2)$ 라 놓으면

$$\mathbb{E}Y = \sum_y yP(Y = y) = \sum_{x_1, x_2} g(x_1, x_2)f_{1,2}(x_1, x_2)$$

Theorem. (선형성) $\mathbb{E}|g_1(X_1, X_2)| < \infty$, $\mathbb{E}|g_2(X_1, X_2)| < \infty$ 이면, $\forall k_1, k_2 \in \mathbb{R}$ 에 대해 다음이 성립한다.

$$\mathbb{E}(k_1 g_1(X_1, X_2) + k_2 g_2(X_1, X_2)) = k_1 \mathbb{E}g_1(X_1, X_2) + k_2 \mathbb{E}g_2(X_1, X_2)$$

Definition. (공분산과 상관계수) $\mathbb{E}X_1^2, \mathbb{E}X_2^2 < \infty$.

$$\begin{aligned} \text{Cov}(X_1, X_2) &:= \mathbb{E}(X_1 - \mu_1)(X_2 - \mu_2) \\ \text{Corr}(X_1, X_2) &:= \frac{\text{Cov}(X_1, X_2)}{\sqrt{\text{Var}(X_1)\text{Var}(X_2)}} \end{aligned}$$

Theorem. (공분산의 성질) $\mathbb{E}X_1^2, \mathbb{E}X_2^2 < \infty$

$$\begin{aligned} \text{(a)} \quad & \begin{cases} \text{Cov}(X_1, X_2) = \mathbb{E}X_1X_2 - \mathbb{E}X_1\mathbb{E}X_2 = \text{Cov}(X_2, X_1) \\ \text{Cov}(X_1, X_1) = \text{Var}(X_1) \\ \text{Cov}(aX_1 + b, X_2) = a\text{Cov}(X_1, X_2) \end{cases} \\ \text{(b)} \quad & \begin{cases} \text{Var}(X_1 + X_2) = \text{Var}(X_1) + \text{Var}(X_2) + 2\text{Cov}(X_1, X_2) \\ \quad \quad \quad = \text{Var}(X_1) + \text{Var}(X_2) \quad (\text{Cov}(X_1, X_2) = 0 \text{ 일 때}) \end{cases} \\ \text{(c)} \quad & |\text{Cov}(X_1, X_2)| \leq \sqrt{\text{Var}(X_1)\text{Var}(X_2)} \end{aligned}$$

Lemma.

(a) (코쉬-슈바르츠 부등식)

$$\mathbb{E}|XY| \leq (\mathbb{E}|X|^2)^{\frac{1}{2}}(\mathbb{E}|Y|^2)^{\frac{1}{2}}$$

(b) (홀더 부등식) $\mathbb{E}|X|^p, \mathbb{E}|X|^q < \infty$, $p + q \geq 1$, $\frac{1}{p} + \frac{1}{q} = 1$ 일 때,

$$\mathbb{E}|XY| \leq (\mathbb{E}|X|^p)^{\frac{1}{p}}(\mathbb{E}|Y|^q)^{\frac{1}{q}}$$

Proof.

(a) Recall that $c \cdot d \leq \frac{1}{2}c^2 + \frac{1}{2}d^2$. $c = \frac{|X|}{(\mathbb{E}|X|^2)^{\frac{1}{2}}}$, $d = \frac{|Y|}{(\mathbb{E}|Y|^2)^{\frac{1}{2}}}$ 라고 두면,

$$\begin{aligned} \frac{|XY|}{\sqrt{\mathbb{E}|X|^2 \cdot \mathbb{E}|Y|^2}} &\leq \frac{1}{2} \frac{X^2}{\mathbb{E}|X|^2} + \frac{1}{2} \frac{|Y|^2}{\mathbb{E}|Y|^2} \\ \mathbb{E}|XY| &\leq \sqrt{\mathbb{E}|X|^2 \mathbb{E}|Y|^2} \end{aligned}$$

(b) Recall that $c \cdot d \leq \frac{1}{p}c^p + \frac{1}{q}d^q$

■

Proof. of the theorem (c):

$$\begin{aligned}
 |Cov(X_1, X_2)| &\leq |\mathbb{E}(X_1 - \mu_1)(X_2 - \mu_2)| \\
 &\leq \mathbb{E}|(X_1 - \mu_1)(X_2 - \mu_2)| \\
 &\leq \sqrt{\mathbb{E}(X_1 - \mu_1)^2 \mathbb{E}(X_2 - \mu_2)^2} \\
 &= \sqrt{Var(X_1)Var(X_2)}
 \end{aligned}$$

공분산의 기하학적 의미

- $X_1, X_2 \in L_2 = \{X : \mathbb{E}X^2 < \infty\}$. 이 공간의 내적 $\langle X_1, X_2 \rangle := \mathbb{E}(X_1 X_2)$.

$$\begin{aligned}
 Cov(X_1, X_2) &= \langle X_1 - \mu_1, X_2 - \mu_2 \rangle \\
 Var(X_1) &= \|X_1 - \mu_1\|_2^2 := \mathbb{E}(X_1 - \mu_1)^2 \\
 Corr(X_1, X_2) &= \frac{Cov(X_1, X_2)}{\sqrt{Var(X_1)Var(X_2)}} = \frac{\langle X_1 - \mu_1, X_2 - \mu_2 \rangle}{\|X_1 - \mu_1\|_2 \|X_2 - \mu_2\|_2}
 \end{aligned}$$

- 유클리드 공간이라면 $\langle X_1 - \mu_1, X_2 - \mu_2 \rangle = \|X_1 - \mu_1\| \|X_2 - \mu_2\| \cdot \cos(\theta)$

Theorem. $\mathbb{E}X_1^2, \mathbb{E}X_2^2 < \infty$, $\mu_i = \mathbb{E}X_i$, $\sigma_i^2 = Var(X_i)$, $i = 1, 2$

- (a) $Var(\frac{X_2 - \mu_2}{\sigma_2} - \rho \frac{X_1 - \mu_1}{\sigma_1}) = 1 - \rho^2$
- (b) $-1 \leq \rho^2 \leq 1$
- (c) $\rho = 1 \iff \frac{X_2 - \mu_2}{\sigma_2} = \frac{X_1 - \mu_1}{\sigma_1} \text{ w.p.1. } (\iff P(\frac{X_2 - \mu_2}{\sigma_2} - \frac{X_1 - \mu_1}{\sigma_1} = 1) = 1)$
- $\rho = -1 \iff \frac{X_2 - \mu_2}{\sigma_2} = -\frac{X_1 - \mu_1}{\sigma_1} \text{ w.p.1.}$

Proof. (a)

$$\begin{aligned}
 Var(\frac{X_2 - \mu_2}{\sigma_2} - \rho \frac{X_1 - \mu_1}{\sigma_1}) &= Var(\frac{X_2 - \mu_2}{\sigma_2}) + Var(\rho \frac{X_1 - \mu_1}{\sigma_1}) - 2Cov(\frac{X_2 - \mu_2}{\sigma_2}, \rho \frac{X_1 - \mu_1}{\sigma_1}) \\
 &= 1 + \rho^2 - 2\rho^2 = 1 - \rho^2
 \end{aligned}$$

- (b) $1 - \rho^2 \geq 0 \iff -1 \leq \rho \leq 1$

- (c) ($\Rightarrow$) $\rho = 1$ 이라면,

$$Var(\frac{X_2 - \mu_2}{\sigma_2} - \rho \frac{X_1 - \mu_1}{\sigma_1}) = 0 \implies \frac{X_2 - \mu_2}{\sigma_2} = \frac{X_1 - \mu_1}{\sigma_1} \text{ w.p.1.}$$

Definition.

- (a) (결합적률) $\mathbb{E}X_1^k X_2^m$: (X_1, X_2) 의 (k, m) 번째 적률
- (b) (결합적률생성함수) 아래 값이 $(0, 0)$ 의 열린 근방에서 정의될 때, (X_1, X_2) 의 mgf가 존재한다고 한다.

$$M_{X_1, X_2}(t_1, t_2) := \mathbb{E}e^{t_1 X_1 + t_2 X_2}, \quad t_1, t_2 \in \mathbb{R}$$

Theorem.

(1) $M_{X_1, X_2}(t_1, t_2)$ 가 존재할 때, 다음이 성립한다.

(a) (X_1, X_2) 의 모든 적률이 존재한다. 즉,

$$\mathbb{E}|X_1|^k |X_2|^m < \infty, \quad \forall k, m > 0.$$

$$(b) \mathbb{E}(X^k Y^m) = \frac{\partial^{k+m}}{\partial t_1^k \partial t_2^m} M_{X_1, X_2}(t_1, t_2) \Big|_{t_1=t_2=0}$$

$$(c) M_{X_1, X_2}(t_1, t_2) = \sum_{r_1=0}^{\infty} \sum_{r_2=0}^{\infty} \frac{\mathbb{E}(X_1^{r_1} X_2^{r_2})}{r_1! r_2!} t_1^{r_1} t_2^{r_2}$$

(2) (mgf의 유일성)

$$(0, 0) \text{ 근방에서 } M_{X_1, X_2}(t_1, t_2) = M_{Y_1, Y_2}(t_1, t_2) \iff (X_1, X_2) \stackrel{d}{=} (Y_1, Y_2)$$

3 조건부분포와 기대값

3.1 조건부분포

Definition. (조건부밀도함수) $X_1, X_2 \sim f_{1,2}$, f_1, f_2 : X_1, X_2 의 주변밀도함수. $X_2 = x_2$ 가 주어졌을 때, X_1 의 조건부밀도함수는

$$f_{1|2}(x_1|x_2) := \frac{f_{1,2}(x_1, x_2)}{f_2(x_2)}, \quad f_2(x_2) > 0$$

Note. 의미:

(이산형일 때)

$$f_{1|2}(x_1|x_2) = \frac{P(X_1 = x_1, X_2 = x_2)}{P(X_2 = x_2)} = P(X_1 = x_1 | X_2 = x_2)$$

(연속형일 때) $X_2 = x_2$ 일 때 X_1 의 분포 = $X_2 \in [x_2, x_2 + dx_2]$ 일 때 X_1 의 분포

$$\begin{aligned} \mathbb{P}(X_1 \in A | x_2 \leq X_2 \leq x_2 + \Delta x_2) &= \frac{\mathbb{P}(X_1 \in A, x_2 \leq X_2 \leq x_2 + \Delta x_2)}{\mathbb{P}(x_2 \leq X_2 \leq x_2 + \Delta x_2)} \\ &= \frac{\int_A \int_{x_2}^{x_2 + \Delta x_2} f_{1,2}(x_1, t) dx_1 dt}{\int_{x_2}^{x_2 + \Delta x_2} f_2(t) dt} \\ &= \frac{\int_A \frac{1}{\Delta x_2} \int_{x_2}^{x_2 + \Delta x_2} f_{1,2}(x_1, t) dt dx_1}{\frac{1}{\Delta x_2} \int_{x_2}^{x_2 + \Delta x_2} f_{1,2}(x_1, t) dt} \\ &\simeq \frac{\int_A f_{1,2}(x_1, x_2) dx_1}{f_2(x_2)} \quad (\because \text{중간값 정리}) \\ &= \int_A f_{1|2}(x_1|x_2) dx_1 \end{aligned}$$

Note. $G(x) := \mathbb{P}(X_1 \in A, X_2 \leq x)$, $F(x) := \mathbb{P}(X_2 \leq x)$

$$\mathbb{P}(X_1 \in A | X_2 \in [x_2, x_2 + \Delta x_2]) = \frac{G(x_2 + \Delta x_2) - G(x_2)}{F(x_2 + \Delta x_2) - F(x_2)} \simeq \frac{dG(x_2)}{dF(x_2)} \implies \text{'G에 대한 F의 미분'}$$

Theorem.

$$(i) f_{1|2}(x_1|x_2) \leq 0, \quad \forall x_1, x_2$$

$$(ii) \begin{cases} \sum f_{1|2}(x_1|x_2) = 1 & (\text{이산형}) \\ \int_{x_1} f_{1|2}(x_1|x_2) dx_1 = 1, f_2(x_2) > 0 & (\text{연속형}) \end{cases}$$

Definition. (조건부누적분포함수)

$$F_{1|2}(x_1|x_2) := \mathbb{P}(X_1 \leq x_1 | X_2 = x_2)$$

$$:= \begin{cases} \sum_{x \leq x_1} f_{1|2}(x_1|x_2) & (\text{이산형}) \\ \int_{x \leq x_1} f_{1|2}(x_1|x_2) dx & (\text{연속형}) \end{cases}$$

3.2 조건부기대값

Definition. (조건부기대값)

$$\mathbb{E}[u(X_1, X_2) | X_1 = x_1] = \begin{cases} \sum u(x_1, x_2) f_{1|2}(x_1|x_2) \\ \int_{x_1} u(x_1, x_2) f_{1|2}(x_1|x_2) dx_1 \end{cases}$$

Theorem.

$$(a) \mathbb{E}(aX_2 + b | X_1 = x_1) = a\mathbb{E}(X_2 | X_1 = x_1) + b, \quad a, b \text{는 양수}$$

$$\mathbb{E}(c(X_1)X_2 | X_1 = x_1) = c(x_1) \cdot \mathbb{E}(x_2 | X_1 = x_1)$$

$$(b) \mathbb{E}(u(X_2) + v(X_2) | X_1 = x_1) = \mathbb{E}(u(X_2) | X_1 = x_1) + \mathbb{E}(v(X_2) | X_1 = x_1)$$

$$(c) u(x_2) \geq 0, \forall x_2 \implies \mathbb{E}(u(X_2) | X_1 = x_1) \geq 0 \quad (\text{기대값이 존재할 때})$$

Proof. (a)의 두번째 식

$$\begin{aligned} \mathbb{E}(c(X_1)X_2 | X_1 = x_1) &= \int c(x_1)x_2 \cdot f_{2|1}(x_2|x_1) dx_2 \\ &= c(x_1) \int x_2 f_{2|1}(x_2|x_1) dx_2 \\ &= c(x_1) \mathbb{E}(X_2 | X_1 = x_1) \end{aligned}$$

■

3.3 조건부평균과 분산

Definition.

$$(\text{평균}) \mathbb{E}(X_2 | X_1 = x_1)$$

$$(\text{분산}) \text{Var}(X_2 | X_1 = x_1) = \mathbb{E}((X_2 - \mathbb{E}(X_2 | X_1 = x_1))^2 | X_1 = x_1)$$

Note. $\mathbb{E}(X_2|X_1) = h(x)$ 라는 함수로 생각할 수 있다. 즉, $h(x_1) = \mathbb{E}(X_2|X_1 = x_1)$

Theorem.

(a) $Var(X_2|X_1 = x) = \mathbb{E}(X_2^2|X_1 = x_1) - (\mathbb{E}(X_2|X_1 = x_1))^2$

(b) (반복적 기대값) $\mathbb{E}|X_2| < \infty$ 일 때,

$$\mathbb{E}(\mathbb{E}(X_2|X_1)) = \mathbb{E}(X_2)$$

(c) (최적예측값)

(i) $\arg \min_{u(X_1)} \mathbb{E}(X_2 - u(X_1))^2 = \mathbb{E}(X_2|X_1)$

(ii) $Cov(X_2 - \mathbb{E}(X_2|X_1), v(X_1)) = 0, \forall v : \mathbb{E}v(X_1) < \infty$

(d) (Eve의 공식) $Var(X_2) = \mathbb{E}[Var(X_2|X_1)] + Var(\mathbb{E}(X_2|X_1))$

Proof. (b)

$$\begin{aligned} \mathbb{E}[\mathbb{E}(X_2|X_1)] &= \int \mathbb{E}(X_2|X_1 = x) f_1(x_1) dx_1 \\ &= \iint x_2 f_{2|1}(x_2|x_1) dx_2 \cdot f_1(x_1) dx_1 \\ &= \iint x_2 \frac{f_{1,2}(x_1, x_2)}{f_1(x_1)} dx_2 \cdot f_1(x_1) dx_1 \\ &= \iint x_2 f_{1,2}(x_1, x_2) dx_1 dx_2 \\ &= \mathbb{E}(X_2) \end{aligned}$$

(c) (ii)

$$\begin{aligned} Cov(X_2 - \mathbb{E}(X_2|X_1), v(X_1)) &= \mathbb{E}((X_2 - \mathbb{E}(X_2|X_1)) \cdot v(X_1)) \\ &= \mathbb{E}(\mathbb{E}[(X_2 - \mathbb{E}(X_2|X_1)) \cdot v(X_1)|X_1]) \\ &= \mathbb{E}(v(X_1) \cdot \mathbb{E}(X_2 - \mathbb{E}(X_2|X_1)|X_1)) \\ &= 0 \end{aligned}$$

(c) (i)

$$\begin{aligned} \mathbb{E}(X_2 - u(X_1))^2 &= \mathbb{E}(X_2 - \mathbb{E}(X_2|X_1) + \mathbb{E}(X_2|X_1) - u(X_1))^2 \\ &= \mathbb{E}(X_2 - \mathbb{E}(X_2|X_1))^2 + \mathbb{E}(\mathbb{E}(X_2|X_1) - u(X_1))^2 + 2(AB) \\ &= \mathbb{E}(X_2 - \mathbb{E}(X_2|X_1))^2 + \mathbb{E}(\mathbb{E}(X_2|X_1) - u(X_1))^2 \\ &\geq \mathbb{E}(X_2 - \mathbb{E}(X_2|X_1))^2 \end{aligned}$$

where $AB = Cov(X_2 - \mathbb{E}(X_2|X_1), \mathbb{E}(X_2|X_1) - u(X_1)) = 0$ ($\cdot \cdot$ (ii))

■

Theorem. (선형예측) $0 < \text{Var}(X_1), \text{Var}(X_2) < \infty$

$$(1) \arg \min_{a+b < 1} \mathbb{E}(X_2 - (a + bX_1))^2 = \mu_2 + \sigma_{21}\sigma_{11}^{-1}(X_1 - \mu_1)$$

$$\begin{aligned} &= \mu_2 + \rho\sigma_2 \frac{X_1 - \mu_1}{\sigma_1} \\ \min_{a,b} \mathbb{E}(X_2 - (a + bX_1))^2 &= \sigma_{22} - \sigma_{21}\sigma_{11}^{-1}\sigma_{12} = \sigma_2^2(1 - \rho)^2 \end{aligned}$$

$$(2) \text{어떤 상수 } a, b \text{에 대해, } \mathbb{E}(X_2|X_1) = a + bX_1 \text{이라면, } \mathbb{E}(X_2|X_1) = \mu_2 + \rho\sigma_2 \frac{X_1 - \mu_1}{\sigma_1}$$

Note. (1)-1의 의미:

$$X_2 \simeq \mu_2 + \rho\sigma_2 \frac{X_1 - \mu_1}{\sigma_1} \iff \frac{X_2 - \mu_2}{\sigma_2} \simeq \rho \frac{X_1 - \mu_1}{\sigma_1}$$

즉, 표준화 한 뒤 예측하면 상관계수가 가장 좋은 선형계수가 된다. (구하기 어려운 조건부기대값 대신 X 의 선형함수로 예측 - X_1, X_2 의 1,2차 적률은 추정하기 쉬움)

4 확률변수의 독립성

Definition. (X_1, X_2) 가 pdf $f_{1,2}$ 를 갖는다. f_1, f_2 : X_1, X_2 의 주변확률밀도함수

$$X_1, X_2 \text{ 서로 독립} \iff f_{1,2}(x_1, x_2) = f_1(x_1)f_2(x_2), \forall x_1, x_2$$

Theorem. $f_{1,2}$: (X_1, X_2) 의 pdf. 다음은 동치이다.

- (a) $X_1 \perp\!\!\!\perp X_2$
- (b) $f_{1,2}(x_1, x_2) = f_1(x_1)f_2(x_2), \forall x_1, x_2$
- (c) $F_{1,2}(x_1, x_2) = F_1(x_1)F_2(x_2), \forall x_1, x_2$
- (d) $P(X_1 \in A, X_2 \in B) = P(X_1 \in A)P(X_2 \in B), \forall A, B$
- (e) $\mu_{1,2}(t_1, t_2) = M_1(t_1)M_2(t_2), \forall t_1, t_2 : |t_1| < h_1, |t_2| < h_2 (\exists h_1, h_2 > 0)$
- (f) $f_{2|1}(x_2|x_1) = x_2$ 의 함수 (x_1 에 의존하지 않음)
- (g) $f_{1,2}(x_1, x_2) = (x_1 \text{의 함수}) \times (x_2 \text{의 함수})$
- (h) $F_{1,2}(x_1, x_2) = (x_1 \text{의 함수}) \times (x_2 \text{의 함수})$
- (i) $M_{1,2}(t_1, t_2) = (t_1 \text{의 함수}) \times (t_2 \text{의 함수})$

5 다차원 확률변수의 분포

Definition. (기대값)

(i) $u : \mathbb{R}^k \rightarrow \mathbb{R}, X = (X_1, \dots, X_k)^T$: k 차원 확률변수, $f = X$ 의 밀도함수

$$\mathbb{E}u(X) = \begin{cases} \int_{x \in \mathbb{R}^k} u(x)f(x)dx & (\text{연속형}) \\ \sum_{x \in \mathbb{R}^k} u(x)f(x) & (\text{이산형}) \end{cases}$$

$$(ii) \ u : \mathbb{R}^k \rightarrow \mathbb{R}^l, u = \begin{pmatrix} u_1 \\ \vdots \\ u_l \end{pmatrix} \text{ (각각은 } u_i = \mathbb{R}^k \rightarrow \mathbb{R}^1 \text{)}$$

$$\mathbb{E}u(X) = \begin{pmatrix} \mathbb{E}u_1(X) \\ \vdots \\ \mathbb{E}u_l(X) \end{pmatrix}$$

$$(iii) \ u : \mathbb{R}^k \rightarrow \mathbb{R}^{m \times n}, u = \begin{pmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ u_{21} & u_{22} & \cdots & u_{2n} \\ \vdots & & \ddots & \\ u_{m1} & u_{m2} & \cdots & u_{mn} \end{pmatrix}, \text{ (각각은 } u_{ij} : \mathbb{R}^k \rightarrow \mathbb{R}^1 \text{)}$$

$$\mathbb{E}u(X) = \begin{pmatrix} \mathbb{E}u_{11}(X) & \mathbb{E}u_{12}(X) & \cdots & \mathbb{E}u_{1n}(X) \\ \mathbb{E}u_{21}(X) & \mathbb{E}u_{22}(X) & \cdots & \mathbb{E}u_{2n}(X) \\ \vdots & & \ddots & \\ \mathbb{E}u_{m1}(X) & \mathbb{E}u_{m2}(X) & \cdots & \mathbb{E}u_{mn}(X) \end{pmatrix}$$

Definition. (공분산행렬) X : k 차원 랜덤벡터, $X = (X_1, \dots, X_k)^T$, $\mathbb{E}X = \begin{pmatrix} \mathbb{E}X_1 \\ \vdots \\ \mathbb{E}X_k \end{pmatrix} := \mu$,

$$\begin{aligned} Var(X) &:= \mathbb{E}[(X - \mu)(X - \mu)^T] \\ &= \left(\sigma_{ij} = Cov(X_i, X_j), 1 \leq i, j \leq k \right) \end{aligned}$$

Theorem. X, X_1, X_2 : 랜덤벡터, $\mu = \mathbb{E}X$

$$(i) \ Var(X) = \mathbb{E}XX^T - \mu\mu^T$$

$$Cov(X_1, X_2) = \mathbb{E}[(X_1 - \mathbb{E}X_1)(X_2 - \mathbb{E}X_2)^T] = \mathbb{E}X_1X_2^T - (\mathbb{E}X_1)(\mathbb{E}X_2)^T$$

$$(ii) \ Var(AX + b) = AVar(X)A^T$$

$$Cov(A_1X_1 + b_1, A_2X_2 + b_2) = A_1Cov(X_1, X_2)A_2^T$$

(iii) (이중선형성)

$$Cov(X_1 + X_2, Y) = Cov(X_1, Y) + Cov(X_2, Y)$$

$$Cov(Y, X) = Cov(X, Y)^T$$

$$Var\left(\sum_{i=1}^k X_i\right) = \sum_{i=1}^k Var(X_i) + 2 \sum_{i < j} Cov(X_i, X_j) \quad (X_i \text{들이 1차원})$$

$$Var\left(\sum_{i=1}^k X_i\right) = \sum_{i=1}^k Var(X_i) + \sum_{i \neq j} Cov(X_i, X_j) \quad (X_i \text{들의 차원} \geq 2)$$

$$(iv) \text{ (음이 아닌 정부호, nonnegative definite) } a^T Var(X)a = Var(a^T X) \geq 0, \forall a \in \mathbb{R}^k$$

6 확률변수의 변환

6.1 이산형 확률변수일 때

$P(Y = y)$ 를 구하면 된다: $P(Y = y) = \sum_{x:u(x)=y} f(x) = f_Y(y)$

Example. $X_1 \sim Poi(\lambda_1) \perp\!\!\!\perp X_2 \sim Poi(\lambda_2)$, $\lambda_1, \lambda_2 > 0$.

$\Rightarrow Y = X_1 + X_2$ 의 밀도함수? (Note. $X \sim Poi(\lambda_i) \iff f(x) = \frac{e^{-\lambda} \lambda^x}{x!}$, $x \geq 0$)

$$\begin{aligned} P(Y = y) &= P(X_1 + X_2 = y), \quad y \geq 0 \\ &= \sum_{x=0}^y P(X_1 = x, X_2 = y - x) = \sum_{x=0}^y \frac{e^{-\lambda_1} \lambda_1^x}{x!} \cdot \frac{e^{-\lambda_2} \lambda_2^{(y-x)}}{(y-x)!} \\ &= \left(\sum_{x=0}^y \frac{y!}{x!(y-x)!} \left(\frac{\lambda_1}{\lambda_1 + \lambda_2} \right)^x \left(\frac{\lambda_2}{\lambda_1 + \lambda_2} \right)^{y-x} \right) \cdot \frac{1}{y!} (\lambda_1 + \lambda_2)^y \cdot e^{-(\lambda_1 + \lambda_2)} \\ &= \frac{1}{y!} e^{-(\lambda_1 + \lambda_2)} \cdot (\lambda_1 + \lambda_2)^y, \quad y \geq 0 \end{aligned}$$

$$\therefore X_1 + X_2 \sim Poi(\lambda_1 + \lambda_2)$$

6.2 연속형 확률변수일 때

Theorem. $Y = u(X)$, u : 1-1, 부드러운 경우

$$f_Y(y) = f_X(u^{-1}(y)) \left| \det \left(\frac{\partial u^{-1}(y)}{\partial y} \right) \right|$$

Note. 공식을 기억하는 법: $(* \ x \xrightarrow{u} y = u(x))$

$$\begin{aligned} f_X(x)dx &= f_X(u^{-1}(y)) \left\| \frac{dx}{dy} \right\| dy \\ &= f_Y(y)dy \end{aligned}$$

Example. $X \sim f_X(x) = 2xI_{(0 < x < 1)}$ 일 때, $Y = x^2$ 의 밀도함수는?

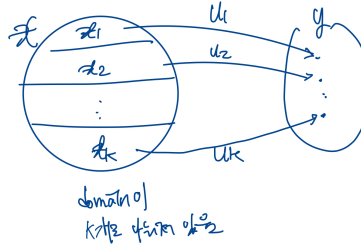
$$\begin{aligned} f_X(x)dx &= 2xI_{(0 < x < 1)}dx = 2\sqrt{y}I_{(0 < \sqrt{y} < 1)} \cdot \frac{1}{2}y^{-\frac{1}{2}}dy \\ &= 1 \cdot I_{(0 < y < 1)}dy \\ \Rightarrow f_Y(y) &= 1 \cdot I_{(0 < y < 1)} \end{aligned}$$

(다른 방법) Y 의 cdf 구한 다음 미분

$$\begin{aligned} F_Y(y) &= P(Y \leq y), \quad y \in (0, 1) \\ &= P(X^2 \leq y) = P(X \leq \sqrt{y}) = \int_0^{\sqrt{y}} 2xdx = y \\ \Rightarrow f_Y(y) &= 1, \quad y \in (0, 1) \end{aligned}$$

Theorem. $Y = u(X), u : \mathfrak{X} \rightarrow y$, $k-1$ 이고 연속적으로 미분가능, $u_i := u|_{x_i} : 1-1$

$$f_Y(y) = \sum_{i=1}^k f_X(u_i^{-1}(y)) \left| \det \frac{\partial u_i^{-1}(y)}{\partial y} \right|$$



Note. 기억하는 법:

$$\begin{aligned} f_X(x)dx &= \sum_{i=1}^k f_X(x)I(x \in x_i)dx = \sum_{i=1}^k f_X(u_i^{-1}(y))I_{(u_i^{-1}(y) \in x_i)} \left| \frac{\partial u_i^{-1}(y)}{\partial y} \right| dy \\ &= \left(\sum_{i=1}^k f_X(u_i^{-1}(y))I_{(u_i^{-1}(y) \in x_i)} \left| \frac{\partial u_i^{-1}(y)}{\partial y} \right| \right) dy \\ &= f_Y(y)dy \end{aligned}$$

Example. $X \sim f_X(x) = \frac{1}{2}I_{(-1 < x < 1)}$ 일 때, $Y = x^2$ 의 밀도함수는?

$$\begin{aligned} f_X(x)dx &= \frac{1}{2}I_{(-1 < x < 1)}dx \\ &= \frac{1}{2}I_{(-1 < x \leq 0)}dx + \frac{1}{2}I_{(0 < x < 1)}dx \\ &= \frac{1}{2}I_{(-1 < -\sqrt{y} \leq 0)}\frac{1}{2}y^{-\frac{1}{2}}dy + \frac{1}{2}I_{(0 < \sqrt{y} < 1)}\frac{1}{2}y^{-\frac{1}{2}}dy(*) \\ &= \frac{1}{2}I_{(0 \leq y < 1)}\frac{1}{2}y^{-\frac{1}{2}}dy + \frac{1}{2}I_{(0 < y < 1)}\frac{1}{2}y^{-\frac{1}{2}}dy \\ &= \frac{1}{2}y^{-\frac{1}{2}}I_{(0 < y < 1)}dy \end{aligned}$$

* $x \rightarrow y = x^2$ 이면 $(-1, 0] \rightarrow [0, 1]$ 에서 $y = x^2 \iff x = -\sqrt{y} \iff dx = \left| \frac{1}{2}y^{-\frac{1}{2}} \right| dy$ (적분 방향에 유의)

3장. 여러 가지 확률분포

Jiho KWAK, Department of Geography, SNU

May 27, 2024

1 이항분포와 관련된 분포들

1.1 베르누이 분포

Notation. $X \sim \text{Ber}(p)$, $0 \leq p \leq 1$

Properties.

1. (정의, 밀도함수) $f(x) = p^x(1-p)^{1-x} = \begin{cases} p & x = 1 \\ 1-p = q & x = 0 \end{cases}$

2. (적률생성함수) $M_X(t) = pe^t + q$, $t \in \mathbb{R}$, $q = 1-p$

3. $\mathbb{E}X = p$

$$\text{Var}(X) = p(1-p)$$

$$\mathbb{E}X^k = p, k \geq 1$$

Proof.

2. $M_X(t) = \mathbb{E}e^{tX} = pe^{t \cdot 1} + qe^{t \cdot 0} = pe^t + q$

3. $\mathbb{E}X = p \cdot 1 + q \cdot 0 = p$

$$\mathbb{E}X^k = p \cdot 1^k + q \cdot 0^k = p, \quad k \geq 1$$

$$\text{Var}(X) = \mathbb{E}X^2 - (\mathbb{E}X)^2 = p - p^2 = p(1-p)$$

Note. $M(t) = pe^t + q = p \sum_{r=1}^{\infty} \frac{1}{r!} t^r + q = p + q + \sum_{r=1}^{\infty} \frac{p}{r!} t^r = 1 + \sum_{r=1}^{\infty} \frac{p}{r!} t^r \implies \mathbb{E}X^r = p$ ■

1.2 이항전개

Definition. (이항계수, binomial coefficient) $\alpha \in \mathbb{C}, k \geq 0$ 정수

$$\binom{\alpha}{k} := \begin{cases} \frac{\alpha \cdot (\alpha-1) \cdots (\alpha-k+1)}{k!} & k \geq 1 \\ 1 & k = 0 \end{cases}$$

Note.

(1) 내림 팩토리얼 (descending factorial): $(\alpha)_k := \alpha(\alpha-1) \cdots (\alpha-k+1) \implies \binom{\alpha}{k} = \frac{(\alpha)_k}{(k)_k}$

(2) 오름 팩토리얼 (ascending factorial): $[\alpha]_k := \alpha(\alpha+1) \cdots (\alpha+k-1)$

Properties.

- (i) $(1+x)^\alpha = \sum_{k=0}^{\infty} \binom{\alpha}{k} x^k, \quad \alpha, x \in \mathbb{C}, |x| < 1$
- (ii) $\sum_{k=0}^{\infty} \binom{\alpha}{k} x^k$ 는 $|x| < 1$ 일 때 절대수렴(absolutely convergent)
- (iii) $n \in \mathbb{N}$ 이면, $\forall a, b \in \mathbb{R}, \quad (a+b)^n = \sum_{k=0}^n \binom{n}{k} a^k b^{n-k}$
- (iv) $\binom{\alpha}{k-1} + \binom{\alpha}{k} = \binom{\alpha+1}{k}$
- (v) $r > 0, k = 0, 1, 2, \dots, \quad \binom{-r}{k} = (-1)^k \binom{r+k-1}{k}$

1.3 이항분포

Definition. (베르누이 과정, Bernoulli process) $X_1, X_2, \dots \stackrel{\text{iid}}{\sim} \text{Ber}(p), p \in [0, 1]$ 일 때, 확률변수들의 모임 $(X_n)_{n \geq 1}$ 을 **베르누이 과정**이라 한다.

Notation. $X \sim \text{Bin}(n, p), 0 \leq p \leq 1, n = 1, 2, 3, \dots$

Properties.

1. (정의, 밀도함수) $f(x) = \binom{n}{x} p^x (1-p)^{n-x}$
2. (적률생성함수) $M(t) = (pe^t + q)^k, \forall t \in \mathbb{R}$
3. $\mathbb{E}X = np, \text{Var}(X) = npq$
4. $X \sim \text{Bin}(n, p) \iff X \stackrel{d}{=} X_1 + \dots + X_n, X_i \stackrel{\text{iid}}{\sim} \text{Ber}(p)$

1.4 기하분포

Notation. $X \sim \text{Geo}(p), 0 \leq p \leq 1. (p = 0 \text{일 때 } X = \infty \text{ w.p.1.})$

*의미: p -동전을 무한 번 던질 때, 첫 앞면이 나올 때까지 나오는 뒷면의 개수 $=: X \sim \text{Geo}(p)$

Properties.

1. (정의, 밀도함수) $f(x) = p(1-p)^x, x = 0, 1, 2, \dots$
2. (적률생성함수) $M(t) = \mathbb{E}e^{tX} = p(1-qe^t)^{-1}, t < -\log q, q = 1-p$
3. $\mathbb{E}X = \frac{q}{p}, \text{Var}(X) = \frac{q}{p^2}$
4. $X \sim \text{Geo}(p) \iff X \stackrel{d}{=} w_1 - 1, w_1 = \min\{n : X_1 + \dots + X_n \geq 1\}$

1.5 음이항분포(Negative Binomial Distribution)

Notation. $X \sim \text{Nbin}(r, p), r = 1, 2, \dots, 0 < p < 1$

*의미: 성공 확률이 p 인 베르누이 과정에서 r 번의 성공이 나올 때까지 나오는 실패의 횟수 $=: X \sim \text{Nbin}(r, p)$

Properties.

1. (밀도함수) $\mathbb{P}(X = x) = \binom{x+r-1}{x} p^r q^x, \quad x = 0, 1, 2, \dots$
2. (적률생성함수) $M(t) = \frac{p^r}{(1 - qe^t)^r}, \quad t < -\log q$
3. $X \sim Nbin(r, p) \iff X \stackrel{d}{=} W_r - r$
where $W_r := \min\{n : X_1 + \dots + X_n \geq r\}, \quad X_1, \dots, X_n \stackrel{i.i.d.}{\sim} Ber(p)$ (r 번째 성공까지 시행횟수)
4. $X \sim Nbin(r, p) \iff X \stackrel{d}{=} X_1 + \dots + X_r, \quad X_1, \dots, X_r \stackrel{iid}{\sim} Geo(p)$
5. $\mathbb{E}X = r \cdot \frac{q}{p}, \quad Var(X) = r \cdot \frac{q}{p^2}$

Proof.

4. 베르누이 과정에서 $W_1 - 1, W_2 - W_1 - 1, \dots, W_r - W_{r-1} - 1 \stackrel{iid}{\sim} Geo(p)$ 임을 보이므로

$$\begin{aligned}
& \mathbb{P}(W_1 - 1 = y_1, W_2 - W_1 - 1 = y_2, \dots, W_r - W_{r-1} - 1 = y_r) \\
&= \mathbb{P}(X_1 = \dots = X_{y_1} = 0, X_{y_1+1} = 1, X_{y_1+2} = \dots = X_{y_1+y_2+1} = 0, \dots, \\
&\quad X_{y_1+\dots+y_{r-1}+r} = \dots = X_{y_1+\dots+y_{r-1}+y_r+r-1} = 0, X_{y_1+y_r+r} = 1) \\
&= (q^{y_1} \cdot p)(q^{y_2} \cdot p) \dots (q^{y_r} \cdot p)
\end{aligned}$$

$W_r - r = (W_1 - 1) + (W_2 - W_1 - 1) + \dots + (W_r - W_{r-1} - 1)$ 이고 각각이 i.i.d. $Geo(p)$ 이므로
 $X \stackrel{d}{=} X_1 + \dots + X_r$

■

Properties.

1. $X_1 \sim Bin(n_1, p)$ 과 $X_2 \sim Bin(n_2, p)$ 이 독립 $\implies X_1 + X_2 \sim Bin(n_1 + n_2, p)$
2. $X_1 \sim Nbin(r_1, p)$ 과 $X_2 \sim Nbin(r_2, p)$ 이 독립 $\implies X_1 + X_2 \sim Nbin(r_1 + r_2, p)$

1.6 초기하분포(Hypereometric Distribution)

Definition. N 개의 공(빨간색 공 M 개, 검은색 공 $N - M$ 개)이 담겨 있는 상자로부터 n 개의 공을 추출하는 시행에서 $X = n$ 개의 공 중 빨간색 공의 수. 이때 X 의 분포를 초기하분포라 하고, $X \sim Hypergeo(N, M, n)$ 으로 표시한다.

Properties.

1. (밀도함수) $\mathbb{P}(X = x) = \frac{\binom{M}{x} \binom{N-M}{n-x}}{\binom{N}{n}}, \quad \max(0, M+n-N) \leq x \leq \min(M, n)$
2. (평균과 분산) $X \sim Hypergeo(N, M, n) \simeq Bin(n, p = \frac{M}{N})$
 $\mathbb{E}X = \frac{M}{N} \cdot n, \quad Var(X) = \frac{N-n}{N-1} \cdot n \cdot \frac{M}{N} \cdot (1 - \frac{M}{N})$
3. (초기하분포의 이항근사) $0 < p < 1$
 $\lim_{N \rightarrow \infty} \mathbb{P}(Hypergeo(N, M = N \cdot p, n) = x) = \mathbb{P}(Bin(n, p) = x)$

2 포아송 분포

2.1 포아송 분포

Theorem. (이항분포의 포아송 근사)

$$\lim_{\substack{n \rightarrow \infty \\ np_n \rightarrow \lambda > 0}} \binom{n}{x} p_n^x (1 - p_n)^{n-x} = \frac{e^{-\lambda} \lambda^x}{x!}, \quad x \geq 0$$

$$\iff \lim_{n \rightarrow \infty} \mathbb{P}(X_n = x) = \mathbb{P}(Y = x) \text{ where } X_n \sim \text{Bin}(n, p_n) \text{ and } Y \sim \text{Poi}(\lambda)$$

*증명을 위해 필요한 사실들

1. 스티링의 공식(Stirling's formula) $n! \simeq \sqrt{2\pi n} \left(\frac{n}{e}\right)^n$ i.e. $\lim_{n \rightarrow \infty} \frac{\sqrt{2\pi n} \left(\frac{n}{e}\right)^n}{n!} = 1$
2. 큰 O와 작은 o: as $n \rightarrow \infty$,
 $x_n = O(a_n) \stackrel{\text{def}}{\iff} \sup_n \left| \frac{x_n}{a_n} \right| < \infty \iff \frac{x_n}{a_n} = O(1)$
 $x_n = o(a_n) \stackrel{\text{def}}{\iff} \frac{x_n}{a_n} \rightarrow 0 \iff \frac{x_n}{a_n} = o(1)$
3. $(1 + \frac{a}{n})^n \xrightarrow{n \rightarrow \infty} e^a$
 $(1 + \frac{a + o(1)}{n})^n \xrightarrow{n \rightarrow \infty} e^a$
4. $a_n \simeq b_n \iff \lim_{n \rightarrow \infty} \frac{a_n}{b_n} = 1$

Proof.

$$\begin{aligned} \binom{n}{x} p_n^x (1 - p_n)^{n-x} &= \frac{n!}{x!(n-x)!} p_n^x (1 - p_n)^{n-x} \\ &\simeq \frac{\sqrt{2\pi n} \left(\frac{n}{e}\right)^n}{x! \sqrt{2\pi(n-x)} \left(\frac{n-x}{e}\right)^{n-x}} p_n^x (1 - p_n)^{n-x} \\ &= \frac{e^{-x}}{x!} \cdot \sqrt{\frac{n}{n-x}} \cdot n^x \cdot \left(\frac{n}{n-x}\right)^{n-x} \cdot p_n^x (1 - p_n)^{n-x} \\ &\xrightarrow{n \rightarrow \infty} \frac{e^{-x}}{x!} \cdot \lambda^x \cdot e^x \cdot e^{-\lambda} = \frac{e^{-\lambda} \lambda^x}{x!}, \quad x \geq 0 \end{aligned}$$

Note that $\begin{cases} n^x \cdot p_n^x = \lambda^x \\ \left(\frac{n}{n-x}\right)^{n-x} = \left(1 + \frac{x}{n-x}\right)^{n-x} \rightarrow e^x \\ (1 - p_n)^{n-x} = \left(1 - \frac{np_n}{n}\right)^{n-x} = \left(1 - \frac{\lambda + o(1)}{n}\right)^{n-x} = \left(1 - \frac{\lambda + o(1)}{n}\right)^n \left(1 - \frac{\lambda + o(1)}{n}\right)^{-x} \rightarrow e^{-\lambda} \end{cases}$ ■

Properties.

1. (밀도함수, 정의) $f(x) = \frac{e^{-\lambda} \lambda^x}{x!}, \quad x \geq 0$
2. (적률생성함수) $M_X(t) = e^{\lambda(e^t - 1)}, \quad t \in \mathbb{R}$
3. $\mathbb{E}X = \lambda, \quad \text{Var}(X) = \lambda$
4. (누율생성함수) $C_X(t) = \lambda(e^t - 1) \implies c_r = \lambda, \quad \forall r \geq 1$

5. $X_1, \dots, X_n$: 서로 독립, $X_i \sim Poi(\lambda)$, $i = 1, \dots, n$
 $\implies \sum_{i=1}^n X_i \sim Poi(\sum_{i=1}^n \lambda_i)$
6. $X_1 \sim Poi(\lambda_1) \perp\!\!\!\perp X_2 \sim Poi(\lambda_2)$
 $\implies X_1 | X_1 + X_2 = n \sim Bin(n, \frac{\lambda_1}{\lambda_1 + \lambda_2})$
7. (Thinning Property) $X \sim Poi(\lambda)$, $Y | X = x \sim Bin(x, p)$
 $\implies Y \sim Poi(\lambda p)$

2.2 포아송과정

Definition. $(N_t, t \geq 0)$ 은 확률과정으로 $N_t = 0, 1, 2, \dots$, $N_0 \equiv 0$ 이고, N_t 는 t 에 관해 증가(nondecreasing)하며, 다음의 조건을 만족하면 도착률(arrival rate) $\lambda > 0$ 인 포아송과정이라 하고, $(N_t, t \geq 0) \sim PP(\lambda)$ 라 표시한다.

- (1) (정상성, stationarity) $N_{t+s} - N_s \stackrel{d}{=} N_t$, $\forall t, s \geq 0$
- (2) (독립증분, independent increment) $N_{t_1}, N_{t_2} - N_{t_1}, \dots, N_{t_k} - N_{t_{k-1}}$ 는 서로 독립 $\forall 0 < t_1 < t_2 < \dots < t_k$
- (3) (희소성, rareness) $h \rightarrow 0$ 일 때, $\mathbb{P}(N_h \geq 2) = o(h)$
- (4) (비례성, proportionality) $h \rightarrow 0$ 일 때, $\mathbb{P}(N_h = 1) = \lambda h + o(h)$

Note. 확률과정(stochastic process): 확률변수들의 모임 (여기서 $N_t : t$ 시점까지 어떤 사건의 발생 횟수)

Theorem. $N_t \sim Poi(\lambda t)$, $t \geq 0$

Lemma. 미분방정식 $y'(x) = ay(x) + b(x)$ 의 해는 $y(x) = y(0)e^{ax} + \int_0^x e^{a(x-s)}b(s)ds$ 이다. 여기서 $a \in \mathbb{R}$, b 는 x 의 함수

Proof. $a = 0$ 일 때 성립. $a \neq 0$ 이라 하자.

$$\begin{aligned}
 y'(x)e^{-ax} &= ae^{-ax}y(x) + e^{-ax}b(x) \\
 (y'(x)e^{-ax})' &= e^{-ax}b(x) \\
 \int_0^x (y'(s)e^{-as})'ds &= \int_0^x e^{-as}b(s)ds \\
 y(x)e^{-ax} + y(0) &= \int_0^x e^{-as}b(s)ds
 \end{aligned}$$

■

Proof. (정리의 증명) $g(n, t) := \mathbb{P}(N_t = n) = \frac{e^{-\lambda t}(\lambda t)^n}{n!}$ 임을 보이자.

(1) $g(0, t) = \mathbb{P}(N_t = 0) = e^{-\lambda t}$ 를 보이자. $\iff$ (i) $g(0, 0) = 1$, (ii) $\frac{\partial}{\partial t}g(0, t) = -\lambda g(0, t)$

$$\begin{aligned}
 \text{(i) } g(0, 0) &= \mathbb{P}(N_0 = 0) = 1 \\
 \text{(ii) } \frac{\partial}{\partial t}g(0, t) &= \lim_{h \rightarrow 0} \frac{1}{h} [g(0, t+h) - g(0, t)] = \lim_{h \rightarrow 0} \frac{1}{h} [\mathbb{P}(N_{t+h} = 0) - \mathbb{P}(N_t = 0)] \\
 &= \lim_{h \rightarrow 0} \frac{1}{h} [\mathbb{P}(N_t = 0, N_{t+h} - N_t = 0) - \mathbb{P}(N_t = 0)] \\
 &\stackrel{(2)}{=} \lim_{h \rightarrow 0} \frac{1}{h} [\mathbb{P}(N_t = 0)\mathbb{P}(N_{t+h} - N_t = 0) - \mathbb{P}(N_t = 0)]
 \end{aligned}$$

$$\begin{aligned}
&\stackrel{(1)}{=} \lim_{h \rightarrow 0} \frac{1}{h} [\mathbb{P}(N_t = 0) \mathbb{P}(N_h = 0) - \mathbb{P}(N_t = 0)] \\
&= \lim_{h \rightarrow 0} \mathbb{P}(N_t = 0) \cdot \frac{1}{h} (1 - \lambda h + o(h) - 1) = \mathbb{P}(N_t = 0) \lim_{h \rightarrow 0} (-\lambda + o(h)) \\
&= g(0, t)(-\lambda)
\end{aligned}$$

By the lemma, $g(0, t) = g(0, 0)e^{-\lambda t} = e^{-\lambda t}$

(2) (i) $g(n, 0) = \mathbb{P}(N_0 = n) = 0$, $n \geq 1$ 이다. (ii) $g'(n, t) = -\lambda g(n, t) + \lambda g(n-1, t)$ 을 보이고자

$$\begin{aligned}
\frac{1}{h}(g(n, t+h) - g(n, t)) &= \frac{1}{h}(\mathbb{P}(N_{t+h} = n) - \mathbb{P}(N_t = n)) \\
&= \frac{1}{h}(\mathbb{P}(N_t = n, N_{t+h} - N_t = 0 \text{ or } N_t = n-1, N_{t+h} - N_t = 1 \\
&\quad \text{or } N_t \leq n-2, N_{t+h} - N_t = \dots) - \mathbb{P}(N_t = n)) \\
&= \frac{1}{h}(\mathbb{P}(N_t = n)\mathbb{P}(N_h = 0) + \mathbb{P}(N_t = n-1)\mathbb{P}(N_h = 1) + \mathbb{P}(N_t \leq n-2)o(h) - \mathbb{P}(N_t = n)) \\
&= \frac{1}{h}(g(n, t) \cdot (1 - \lambda h + o(h)) + g(n-1, t)(\lambda h + o(h)) + o(h) - g(n, t)) \\
&\rightarrow -\lambda g(n, t) + \lambda g(n-1, t) \text{ as } h \rightarrow 0
\end{aligned}$$

By the lemma,

$$\begin{aligned}
g(n, t) &= g(n, 0)e^{-\lambda t} + \int_0^t e^{-\lambda(t-s)} \lambda g(n-1, s) ds \\
&= \int_0^t \lambda e^{-\lambda(t-s)} g(n-1, s) ds
\end{aligned}$$

$$\begin{aligned}
n=1 \text{ 일 때: } g(1, t) &= \int_0^t \lambda e^{-\lambda(t-s)} g(0, s) ds = \int_0^t \lambda e^{-\lambda(t-s)} e^{-\lambda s} ds = \int_0^t \lambda e^{-\lambda t} ds = \lambda t e^{-\lambda t} \\
n=2, \dots \text{ 차례로 풀면 된다.}
\end{aligned}$$

■

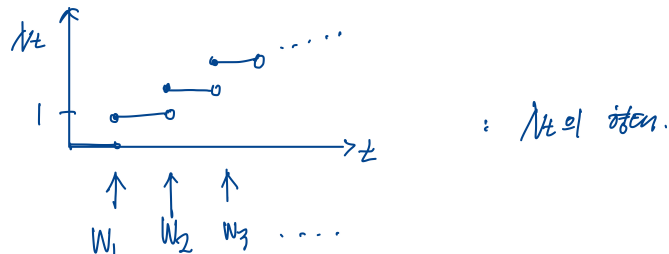
3 감마분포와 관련된 분포

3.1 포아송 과정의 대기시간

Theorem. $(N_t, t \geq 0) \sim PP(\lambda)$, $\lambda > 0$ 라 하자. $W_r := \min\{t : N_t \geq r\}$: r 번째 대기시간

(i) $W_r \sim Ga(r, \frac{1}{\lambda})$

(ii) $W_1, W_2 - W_1, \dots \stackrel{\text{iid}}{\sim} Ga(1, \frac{1}{\lambda}) = Exp(\frac{1}{\lambda})$



Proof.

(i)

$$\begin{aligned}
 \mathbb{P}(W_r > T) &= \mathbb{P}(N_t \leq r-1) \\
 &= \sum_{k=0}^{r-1} \frac{e^{-\lambda t} (\lambda t)^k}{k!} \\
 \frac{d}{dt} \mathbb{P}(W_r \leq t) &= \frac{d}{dt} (1 - \mathbb{P}(W_r > t)) \\
 &= \frac{d}{dt} \left(1 - \frac{e^{-\lambda t} (\lambda t)^k}{k!}\right) \\
 &= \frac{\lambda^r}{(r-1)!} \cdot t^{r-1} e^{-\lambda t} : X \sim Ga(r, \frac{1}{\lambda}) \text{의 밀도함수}
 \end{aligned}$$

■

3.2 감마함수

Definition. $\Gamma(x) := \int_0^\infty t^{x-1} e^{-t} dt, \forall x > 0$ ($\Gamma(0) := 1$)

Lemma. $\Gamma(x) < \infty, \forall x \geq 0$

Proof. $x \in (0, 1), x \geq 1$ 두 가지 case로 나누어 $\Gamma(x) = \int_0^1 + \int_1^\infty \dots dt < \infty$ 를 검토한다.

■

Properties.

- (1) $\Gamma(x+1) = x\Gamma(x), x > 0$
- (2) $\Gamma(n+1) = n!, n = 1, 2, \dots$
- (3) $\log \Gamma(x)$ 는 $(0, \infty)$ 상에서 볼록

Proof.

(1)

$$\begin{aligned}
 \Gamma(x+1) &= \int_0^\infty t^x e^{-t} dt \\
 &= e^{-t} t^x \Big|_0^\infty + \int_0^\infty x t^{x-1} e^{-t} dt \\
 &= x \int_0^\infty t^{x-1} e^{-t} dt
 \end{aligned}$$

(2) $\Gamma(n+1) = n\Gamma(n) = n(n-1)\Gamma(n-1) = \dots = n!$

(3) $0 < x < y < \infty, \alpha \in (0, 1)$

$$\log \Gamma(\alpha x + (1-\alpha)y) \leq \alpha \log \Gamma(x) + (1-\alpha) \log \Gamma(y)$$

$$\Gamma(\alpha x + (1-\alpha)y) \leq \Gamma(x)^\alpha \Gamma(y)^{1-\alpha} \text{를 보이하고자}$$

$$\begin{aligned}
\Gamma(\alpha x + (1-\alpha)y) &= \int_0^\infty t^{\alpha x + (1-\alpha)y} \frac{1}{t} e^{-t} dt \\
&\geq \left(\int_0^\infty (t^{\alpha x})^{\frac{1}{\alpha}} \frac{1}{t} e^{-t} dt \right)^\alpha \left(\int_0^\infty (t^{(1-\alpha)y})^{\frac{1}{1-\alpha}} \frac{1}{t} e^{-t} dt \right)^{1-\alpha} (\because \text{홀더 부등식}) \\
&= \left(\int_0^\infty t^{x-1} e^{-t} dt \right)^\alpha \left(\int_0^\infty t^{y-1} e^{-t} dt \right)^{1-\alpha} \\
&= \Gamma(x)^\alpha \Gamma(y)^{1-\alpha}
\end{aligned}$$

■

3.3 베타함수

Definition. $B(x, y) = \int_0^1 t^{x-1} (1-t)^{y-1} dt$, $x, y > 0$, 모든 x, y 에 대해 $< \infty$

Properties. $x, y > 0$

$$(1) \Gamma(x) = 2 \int_0^\infty s^{2x-1} e^{-s^2} ds$$

$$(2) 2 \int_0^{\pi/2} (\sin \theta)^{2x-1} (\cos \theta)^{2y-1} d\theta = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$$

$$(3) B(x, y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$$

$$(4) \Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$$

$$(5) \int_{-\infty}^\infty e^{-s^2} ds = \sqrt{\pi}, \int_{-\infty}^\infty e^{-\frac{1}{2}s^2} ds = \sqrt{2\pi}$$

$$(6) (\text{르장드르의 중복공식, Legendre Duplication Formula}) \Gamma(2x) = \frac{2^{2x-1}}{\sqrt{\pi}} \Gamma(x) \Gamma\left(x + \frac{1}{2}\right)$$

$$(7) \Gamma\left(x + \frac{1}{2}\right) = \frac{(2x)!}{2^{2x} x!} \sqrt{\pi}, \quad x : \text{양의 정수}$$

$$(8) x \in (0, 1) \text{ 일 때, } \Gamma(x)\Gamma(1-x) = \frac{x}{\sin(\pi x)}$$

Proof.

(1)

$$\begin{aligned}
\Gamma(x) &= \int_0^\infty t^{x-1} e^{-t} dt \\
&= 2 \int_0^\infty (s^2)^{x-1} e^{-s^2} s ds \quad (t \text{를 } s^2 \text{로 치환}) \\
&= 2 \int_0^\infty s^{2x-1} e^{-s^2} ds
\end{aligned}$$

(2)

$$\begin{aligned}
I &= \int_{s, t \geq 0} s^{2x-1} y^{2y-1} e^{-s^2-t^2} ds dt = \int_0^\infty s^{2x-1} e^{-s^2} ds \int_0^\infty t^{2y-1} e^{-t^2} dt \\
&= \frac{1}{2} \Gamma(x) \cdot \frac{1}{2} \Gamma(y) = \frac{1}{4} \Gamma(x) \Gamma(y)
\end{aligned}$$

$$\begin{aligned}
I &= \iint_{s,t \geq 0} s^{2x-1} y^{2y-1} e^{-s^2-t^2} ds dt \\
&= \int_{r=0}^{\infty} \int_{\theta=0}^{\pi/2} e^{-r^2} (r \cos \theta)^{2x-1} (r \sin \theta)^{2y-1} r d\theta dr \quad (s \text{를 } r \cos \theta \text{로, } t \text{를 } r \sin \theta \text{로 치환}) \\
&= \int_{r=0}^{\infty} r^{2x+2y-1} e^{-r^2} dr \times \int_{\theta=0}^{\pi/2} (\cos \theta)^{2x-1} (\sin \theta)^{2y-1} d\theta \\
&= \frac{1}{2} \Gamma(x+y) \cdot \int_{r=0}^{\infty} \int_{\theta=0}^{\pi/2} (\cos \theta)^{2x-1} (\sin \theta)^{2y-1} d\theta
\end{aligned}$$

$$\begin{aligned}
\therefore 2 \int_0^{\pi/2} (\sin \theta)^{2x-1} (\cos \theta)^{2y-1} d\theta &= \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)} \\
(3)
\end{aligned}$$

$$\begin{aligned}
B(x, y) &= \int_0^1 t^{x-1} (1-t)^{y-1} dt \quad (t = \sin^2 \theta \text{로 치환, } dt = 2 \sin \theta \cos \theta d\theta) \\
&= \int_0^{\pi/2} (\sin^2 \theta)^{x-1} (\cos^2 \theta)^{y-1} \cdot 2 \sin \theta \cos \theta d\theta \\
&= 2 \int_0^{\pi/2} (\sin \theta)^{2x-1} (\cos \theta)^{2y-1} d\theta = \frac{\Gamma(x)\Gamma(y)}{(2) \Gamma(x+y)}
\end{aligned}$$

(4) (2)에 $x = y = \frac{1}{2}$ 대입

$$2 \int_0^{\pi/2} d\theta = \frac{\Gamma(\frac{1}{2})^2}{\Gamma(1)} = \Gamma(\frac{1}{2})^2 \Gamma(\frac{1}{2}) = \sqrt{\pi}$$

(5) (1)에 $x = \frac{1}{2}$ 대입

(6)

$$\begin{aligned}
\frac{\Gamma(x)\Gamma(x)}{\Gamma(x+x)} &= B(x, x) = \int_0^1 t^{x-1} (1-t)^{x-1} dt \\
&= \frac{2}{2^{2x}-1} \int_0^1 (1-s^2)^{x-1} ds \quad (t = \frac{1}{2}(1+s) \text{로 치환}) \\
&= \frac{1}{2^{2x}-1} B(\frac{1}{2}, x) \quad (u = s^2 \text{로 치환}) \\
\Rightarrow \frac{\Gamma(x)^2}{\Gamma(2x)} &= \frac{1}{2^{2x-1}} \frac{\Gamma(\frac{1}{2})\Gamma(x)}{\Gamma(x+\frac{1}{2})}
\end{aligned}$$

(7) x : 양의 정수

$$\begin{aligned}
\Gamma(2x) &= \frac{2^{2x-1}}{\sqrt{\pi}} \Gamma(x) \Gamma(x + \frac{1}{2}) \quad (\because (6)) \\
(2x-1)! &= \frac{2^{2x-1}}{\sqrt{\pi}} \Gamma(x) \Gamma(x + \frac{1}{2}) \\
\Gamma(x + \frac{1}{2}) &= \frac{(2x-1)! \sqrt{\pi}}{2^{2x-1} (x-1)!} = \frac{(2x)! \sqrt{\pi}}{2^{2x} x!}
\end{aligned}$$

(8) $x \in (0, 1)$

$$\begin{aligned}
\Gamma(x)\Gamma(1-x) &= \int_0^\infty t^{x-1}e^{-t}dt \cdot \int_0^\infty s^{-x}e^{-s}ds \\
&= \int_0^\infty \int_0^\infty e^{-s-t} \left(\frac{t}{s}\right)^x \frac{1}{t} ds dt \\
&\vdots \quad \left(s = \frac{u}{1+u}, t = \frac{uv}{1+v} \text{로 치환}\right) \\
&= \int_0^\infty \frac{v^{x+1}}{1+v} dv \\
&= \frac{\pi}{\sin(\pi x)} \quad (\text{복소수 Contour Integral 이용})
\end{aligned}$$

■

3.4 감마분포

Notation. $X \sim Ga(\alpha, \beta)$, $\alpha, \beta > 0$

Properties.

1. (밀도함수, 정의) $f(x) = \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-\frac{x}{\beta}}$, $x > 0$ (*감마분포의 cdf는 알려져 있지 않다.)
2. (적률생성함수) $M(t) = (1 - \beta t)^{-\alpha}$, $t < \frac{1}{\beta}$
3. (적률) $\mathbb{E}X = \alpha\beta$, $Var(X) = \alpha\beta^2$, $\mathbb{E}X^k = [\alpha]_k \beta^k$ (* $[\alpha]_k$ 는 ascending factorial)
4. (누율생성함수) $C(t) = -\alpha \log(1 - \beta t)$, $t < \frac{1}{\beta}$
5. (누율) $c_r = (r-1)!\alpha\beta^r$, $r \geq 1$

Note. $f(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}$, $x > 0$ 으로 정의할 경우, $\begin{cases} \mathbb{E}X = \frac{\alpha}{\lambda} \\ Var(X) = \frac{\alpha}{\lambda^2} \end{cases}$ 이며, $\begin{cases} \alpha = \text{shape parameter (형상모수)} \\ \beta = \text{scale parameter (척도모수)} \\ \lambda = \frac{1}{\beta} = \text{rate parameter (비율모수)} \end{cases}$

Proof.

1.

$$\begin{aligned}
\int_0^\infty f(x)dx &= \int_0^\infty \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-\frac{x}{\beta}} dx \\
&= \frac{1}{\Gamma(\alpha)\beta^\alpha} \int_0^\infty (\beta t)^{\alpha-1} e^{-t} \beta dt \quad (t = \frac{x}{\beta} \text{로 치환}) \\
&= \frac{\beta^\alpha}{\Gamma(\alpha)\beta^\alpha} \int_0^\infty t^{\alpha-1} e^{-t} dt = 1
\end{aligned}$$

2.

$$\begin{aligned}
M(t) = \mathbb{E}e^{tx} &= \int_0^\infty e^{tx} \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-\frac{x}{\beta}} dx \\
&= \frac{1}{\Gamma(\alpha)\beta^\alpha} \int_0^\infty x^{\alpha-1} e^{-(\frac{1}{\beta}-t)x} dx, \quad (a > 0, \frac{\beta}{1-\beta t} > 0) \\
&= \frac{1}{\Gamma(\alpha)\beta^\alpha} \int_0^\infty \Gamma(\alpha) \left(\frac{\beta}{1-\beta t}\right)^\alpha dx = (1-\beta t)^{-\alpha}, \quad (1-\beta t > 0)
\end{aligned}$$

3.

$$\begin{aligned}
\mathbb{E}X^k &= \int_0^\infty x^k \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-\frac{x}{\beta}} dx \\
&= \frac{1}{\Gamma(\alpha)\beta^\alpha} \int_0^\infty x^{\alpha+k-1} e^{-\frac{x}{\beta}} dx \longrightarrow \frac{x}{\beta} = u \text{로 치환} \\
&= \frac{1}{\Gamma(\alpha)\beta^\alpha} \Gamma(\alpha+k) \beta^{\alpha+k} = \frac{\Gamma(\alpha+k)}{\Gamma(\alpha)} \beta^k
\end{aligned}$$

5.

$$\begin{aligned}
C(t) &= \alpha \sum_{r=1}^{\infty} \frac{(\beta t)^r}{r} \\
&= \alpha \sum_{r=1}^{\infty} \beta^r (r-1)! \frac{t^r}{r!}
\end{aligned}$$

■

Theorem.(a) $X_i \stackrel{\text{indep}}{\sim} Ga(\alpha, \beta), i = 1, \dots, n \implies X_1 + \dots + X_n \sim Ga(\sum_{i=1}^n \alpha_i, \beta)$ (b) r : 양의 정수, $X \sim Ga(r, \beta) \iff X = Z_1 + \dots + Z_r, Z_i \stackrel{\text{iid}}{\sim} Ga(1, \beta)$ (c) $X \sim Ga(\alpha, \beta) \implies cX \sim Ga(\alpha, c\beta), c > 0$ *Proof.*

(a)

$$\begin{aligned}
M_{\sum_{i=1}^n X_i}(t) &= \mathbb{E} e^{t \sum_{i=1}^n X_i} = \prod_{i=1}^n \mathbb{E} e^{t X_i} \\
&= \prod_{i=1}^n M_{X_i}(t) = \prod_{i=1}^n (1 - \beta t)^{-\alpha_i}, t < \frac{1}{\beta} \\
&= (1 - \beta t)^{-\sum_{i=1}^n \alpha_i} = Ga(\sum_{i=1}^n \alpha_i, \beta) \text{의 mgf}
\end{aligned}$$

(b) (a)에 의해 성립한다.

(c) $Y = cX$ 라 두면 $y = cx, dx = dy/c$

$$\begin{aligned}
f_X(x)dx &= \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-\frac{x}{\beta}} dx \\
&= \frac{1}{\Gamma(\alpha)\beta^\alpha} \left(\frac{y}{c}\right)^{\alpha-1} e^{-\frac{y}{c\beta}} dy/c \\
&= \frac{1}{\Gamma(\alpha)(c\beta)^\alpha} y^{\alpha-1} e^{-\frac{y}{c\beta}} dy = Ga(\alpha, c\beta) \text{의 밀도함수}
\end{aligned}$$

■

3.5 지수분포

Notation. $X \sim \text{Exp}(\beta)$, $\beta > 0$ (β : 척도모수, $\lambda = \frac{1}{\beta}$: 비율모수)

Properties.

1. (정의) $\text{Exp}(\beta) = \text{Ga}(1, \beta)$
2. (밀도함수) $\frac{1}{\beta} e^{-x/\beta}$, $x > 0$ (cf. 비율모수 쓸 경우 $\lambda e^{-\lambda x}$, $x > 0$)
3. (cdf) $F(x) = 1 - e^{-x/\beta}$, $x > 0$
4. (survival function, 생존함수) $S(x) := 1 - F(x) = P(X > x) = e^{-x/\beta}$, $x > 0$
5. (mgf, cgf) $M(t) = \frac{1}{1 - \beta t}$, $t < \frac{1}{\beta}$, $C(t) = -\log(1 - \beta t)$
6. (적률) $\mathbb{E}X = \beta$, $\text{Var}(X) = \beta^2$, $\mathbb{E}X^k = (k-1)!\beta^k$, $k \geq 1$
7. (누율) $c_r = (r-1)!\beta^r$, $r \geq 1$

3.6 카이제곱분포

Definition. $\chi^2(r) = \text{Ga}(\frac{r}{2}, \text{scale} = 2)$, $r > 0$, r : 자유도(degree of freedom)

Properties.

1. (밀도함수) $f(x) = \frac{1}{\Gamma(\frac{r}{2})2^{r/2}} x^{\frac{r}{2}-1} e^{-x/2}$, $x > 0$
2. (mgf, cgf) $M(t) = (1 - 2t)^{-\frac{r}{2}}$, $t < \frac{1}{2}$, $C(t) = -\frac{r}{2} \log(1 - 2t)$
3. (적률) $\mathbb{E}X = r$, $\text{Var}(X) = 2r$, $\mathbb{E}X^k = [\frac{r}{2}]_k 2^k$, $k \geq 1$
4. (누율) $c_k = (k-1)! \frac{r}{2} \cdot 2^k$, $k \geq 1$
5. $X_i \stackrel{\text{indep}}{\sim} \chi^2(r_i)$, $i = 1, \dots, k \implies \sum_{i=1}^k X_i \sim \chi^2(\sum_{i=1}^k r_i)$
6. $X \sim \chi^2(r) \iff X = Z_1 + \dots + Z_r$, $Z_i \stackrel{\text{iid}}{\sim} \chi^2(1)$, r : 자연수

3.7 베타분포

Notation. $X \sim \text{Beta}(\alpha, \beta)$, $\alpha, \beta > 0$

Properties.

1. (밀도함수, 정의) $f(x) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1} (1-x)^{\beta-1}$, $x \in (0, 1)$
2. (적률) $\mathbb{E}X^k = \frac{[\alpha]_k}{[\alpha + \beta]_k}$, $\mathbb{E}X = \frac{\alpha}{\alpha + \beta}$, $\text{Var}(X) = \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)}$
3. $X_1 \sim \text{Ga}(\alpha, 1) \perp\!\!\!\perp X_2 \sim \text{Ga}(\beta, 1)$, $\alpha, \beta > 0$
 $\implies \frac{X_1}{X_1 + X_2} \sim \text{Beta}(\alpha, \beta)$, and $\perp\!\!\!\perp X_1 + X_2 \sim \text{Ga}(\alpha + \beta, 1)$

Proof.

3. $Y_1 = \frac{X_1}{X_1 + X_2}$, $Y_2 = X_1 + X_2$ 라고 두자.

Note that $\begin{cases} x_1 = y_1 y_2 \\ x_2 = y_2 - y_1 y_2 = y_2(1 - y_1) \end{cases}$ and $\left| \frac{d(x_1, x_2)}{d(y_1, y_2)} \right| = \begin{vmatrix} y_2 & y_1 \\ -y_2 & 1 - y_1 \end{vmatrix} = y_1(1 - y_1) + y_1 y_2 = y_2$

$$\begin{aligned}
 f_{X_1, X_2}(x_1, x_2) dx_1 dx_2 &= \frac{1}{\Gamma(\alpha)} x_1^{\alpha-1} e^{-x_1} \frac{1}{\Gamma(\beta)} x_2^{\beta-1} e^{-x_2} dx_1 dx_2 I(x_1, x_2 > 0) \\
 &= \frac{1}{\Gamma(\alpha)\Gamma(\beta)} (y_1 y_2)^{\alpha-1} e^{-y_1 y_2} (y_2(1 - y_1))^{\beta-1} e^{-y_2(1-y_1)} y_2 dy_1 dy_2 \\
 &\quad \times I(y_1 y_2 > 0) I(y_2(1 - y_1) > 0) \\
 &= \frac{1}{\Gamma(\alpha)\Gamma(\beta)} y_1^{\alpha-1} (1 - y_1)^{\beta-1} y_2^{\alpha+\beta-1} e^{-y_2} dy_1 dy_2 I(y_1 \in (0, 1)) I(y_2 > 0) \\
 &= \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} y_1^{\alpha-1} (1 - y_1)^{\beta-1} dy_1 I(y_1 \in (0, 1)) \\
 &\quad \times \frac{1}{\Gamma(\alpha + \beta)} y_2^{\alpha+\beta-1} e^{-y_2} dy_2 I(y_2 > 0) \\
 &= \text{Beta}(y_1 | \alpha, \beta) \times \text{Ga}(y_2 | \alpha + \beta, 1)
 \end{aligned}$$

■

3.8 디리크레 분포

Notation. $(Y_1, \dots, Y_k) \sim \text{Dir}(\alpha_1, \dots, \alpha_k, \alpha_{k+1})$, $\alpha_i > 0$, $i = 1, \dots, k+1$

Properties.

1. (밀도함수) $y_i > 0$, $i = 1, \dots, k$, $\sum_{i=1}^k y_i < 1$,

$$f(y_1, \dots, y_k) = \frac{\Gamma(\alpha_1 + \dots + \alpha_{k+1})}{\Gamma(\alpha_1) \dots \Gamma(\alpha_{k+1})} y_1^{\alpha_1-1} \dots y_k^{\alpha_k-1} (1 - y_1 - \dots - y_k)^{\alpha_{k+1}-1}$$
2. $\mathbb{E}Y_1 = \frac{\alpha_1}{\alpha_1 + \dots + \alpha_{k+1}}$, $\text{Var}(Y_1) = \frac{\alpha_1(\alpha_2 + \dots + \alpha_{k+1})}{(\sum_{i=1}^{k+1} \alpha_i)^2 (\sum_{i=1}^{k+1} \alpha_i + 1)}$, $\text{Cov}(Y_1, Y_2) = \frac{\alpha_1 \cdot \alpha_2}{(\sum_{i=1}^{k+1} \alpha_i)^2 (\sum_{i=1}^{k+1} \alpha_i + 1)}$
3. $X_i \stackrel{\text{indep}}{\sim} \text{Ga}(\alpha_i, 1)$, $i = 1, \dots, k+1$
 $\implies (Y_1, \dots, Y_k) = \left(\frac{X_1}{\sum_{i=1}^{k+1} X_i}, \dots, \frac{X_k}{\sum_{i=1}^{k+1} X_i} \right) \sim \text{Dir}(\alpha_1, \dots, \alpha_{k+1})$ and $\perp\!\!\!\perp \sum_{i=1}^{k+1} X_i \sim \text{Ga}(\sum_{i=1}^{k+1} \alpha_i, 1)$
4. 3.에서 $Y_1 \sim \text{Beta}(\alpha_1, \sum_{i=2}^{k+1} \alpha_i)$

Note. 1. $X \sim \text{Beta}(\alpha, \beta) = \text{Dir}(\alpha, \beta)$: $k = 1$ 일 때 디리크레 분포

2. $(X_1, X_2) \sim \text{Dir}(\alpha_1, \alpha_2, \alpha_3)$: 2차원 simplex

4 정규분포

4.1 드므아브르-라플라스 정리

Theorem. (드르아브르-라플라스 정리)

(a) $p \in (0, 1)$, $q = 1 - p$, $a \leq \frac{x - np}{\sqrt{npq}} \leq b$ 인 x 에 대해 $n \rightarrow \infty$ 일 때,

$$\binom{n}{x} p^x (1-p)^{n-x} \sim \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{x-np}{\sqrt{npq}} \right)^2} \frac{1}{\sqrt{npq}}$$

(b) $\sum_{x: a \leq \frac{x-np}{\sqrt{npq}} \leq b} \binom{n}{x} p^x (1-p)^{n-x} \sim \int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} z^2} dz$

$$\iff \mathbb{P}(a \leq \frac{X - np}{\sqrt{npq}} \leq b) \sim \mathbb{P}(a \leq Z \leq b) \text{ where } X \sim \text{Bin}(n, p), Z \sim N(0, 1)$$

Proof. (sketch)

$$(a) \text{ } x \text{ 가 } a \leq \frac{x - np}{\sqrt{npq}} \leq b \text{ 를 만족하므로 } \frac{x - np}{\sqrt{npq}} = z \text{ 라고 두면 } \begin{cases} x &= np + z\sqrt{npq} \\ n - x &= nq + z\sqrt{npq} \end{cases}$$

(* $x, n - x \rightarrow \infty$ as $n \rightarrow \infty$)

$$\begin{aligned} \binom{n}{x} p^x (1-p)^{n-x} &= \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} \\ &\sim \frac{\sqrt{2\pi n} \left(\frac{n}{e}\right)^n}{\sqrt{2\pi x} \left(\frac{x}{e}\right)^x \sqrt{2\pi(n-x)} \left(\frac{n-x}{e}\right)^{n-x}} p^x (1-p)^{n-x} \\ &= \frac{1}{\sqrt{2\pi}} \sqrt{\frac{n}{x(n-x)}} \left(\frac{n}{x}\right)^x \left(\frac{n}{n-x}\right)^{n-x} p^x (1-p)^{n-x} \\ &= \frac{1}{\sqrt{2\pi}} \frac{p^x (1-p)^{n-x}}{\left(\frac{x}{n}\right)^x \left(1 - \frac{x}{n}\right)^{n-x}} \frac{1}{\sqrt{n \cdot \frac{x}{n} \left(1 - \frac{x}{n}\right)}} \end{aligned}$$

$$\text{Note that } \begin{cases} \frac{x}{n} = p + z\sqrt{\frac{pq}{n}} \\ \frac{n-x}{n} = q + z\sqrt{\frac{pq}{n}} \end{cases} \text{ thus } \frac{1}{\sqrt{n \cdot \frac{x}{n} \left(1 - \frac{x}{n}\right)}} \sim \frac{1}{\sqrt{npq}} \text{ and}$$

$$\begin{aligned} \log \frac{p^x (1-p)^{n-x}}{\left(\frac{x}{n}\right)^x \left(1 - \frac{x}{n}\right)^{n-x}} &= -x(\log \frac{x}{n} - \log p) - (n-x)(\log(1 - \frac{x}{n}) - \log q) \\ &= -(np + z\sqrt{npq})(\log \frac{p + z\sqrt{\frac{pq}{n}}}{p}) - (nq + z\sqrt{npq})(\log \frac{q - z\sqrt{\frac{pq}{n}}}{q}) \\ &= -(np + z\sqrt{npq}) \log \left(1 + z\sqrt{\frac{q}{np}}\right) - (nq + z\sqrt{npq}) \log \left(1 - z\sqrt{\frac{p}{nq}}\right) \\ &= -(np + z\sqrt{npq}) \left(z\sqrt{\frac{q}{np}} - \frac{1}{2} \left(z\sqrt{\frac{q}{np}} \right)^2 + \frac{1}{3} \left(z\sqrt{\frac{q}{np}} \right)^3 - \dots \right) \\ &\quad + (nq + z\sqrt{npq}) \left(z\sqrt{\frac{p}{nq}} + \frac{1}{2} \left(z\sqrt{\frac{p}{nq}} \right)^2 + \frac{1}{3} \left(z\sqrt{\frac{p}{nq}} \right)^3 + \dots \right) \\ &= -z\sqrt{npq} + \frac{1}{2} z^2 q - z^2 q + \dots \\ &\quad + z\sqrt{npq} + \frac{1}{2} z^2 p - z^2 p + \dots \\ &= -\frac{1}{2} z^2 + \dots \end{aligned}$$

$$\therefore \binom{n}{x} p^x (1-p)^{n-x} \sim \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} z^2} \frac{1}{\sqrt{npq}}$$

(b) $\phi: N(0, 1)$ 의 pdf라 하자. (a)로부터

$$\sum_{x: a \leq \frac{x-np}{\sqrt{npq}} \leq b} \binom{n}{x} p^x (1-p)^{n-x} \sim \sum_{x: a \leq \frac{x-np}{\sqrt{npq}} \leq b} \phi\left(\frac{x-np}{\sqrt{npq}}\right) \frac{1}{\sqrt{npq}} = (*)$$

$a \leq \frac{x-np}{\sqrt{npq}} \leq b$ 를 만족하는 x 들을 $x_1, \dots, x_N$ 이라 하고 $z_i = \frac{x_i - np}{\sqrt{npq}}$, $i = 1, \dots, N$ 이라 놓으면

$$(*) = \sum_{j=1}^N \frac{1}{\sqrt{npq}} \phi(z_j) \sim \int_a^b \phi(z) dz$$

■

4.2 정규분포

Notation. $X \sim N(\mu, \sigma^2)$, $\mu \in \mathbb{R}$, $\sigma^2 > 0$

Properties.

1. (밀도함수, 정의) $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$
2. (mgf) $M(t) = e^{\mu t + \frac{1}{2}\sigma^2 t^2}$ (* $Z \sim N(0, 1)$ 일 때, $f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$, $M_Z(t) = e^{\frac{1}{2}t^2}$)
3. (표준정규분포의 적률) $Z \sim N(0, 1)$ 일 때, $\mathbb{E}Z^k = \begin{cases} \frac{(2r)!}{2^r r!} & k = 2r, r = 1, 2, \dots \\ 0 & k = 2r + 1, r = 0, 1, 2, \dots \end{cases}$
4. $X \sim N(\mu, \sigma^2)$ 일 때, $\mathbb{E}X = \mu$, $\text{Var}(X) = \sigma^2$

Proof.

2.

$$\begin{aligned} M(t) &= \mathbb{E}e^{tX} = \int e^{tx} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2} dx \\ &= \frac{1}{\sigma\sqrt{2\pi}} \int e^{-\frac{1}{2\sigma^2}(x^2 - 2\mu x + \mu^2 - 2\sigma^2 tx)} dx \\ &= \frac{1}{\sigma\sqrt{2\pi}} \int e^{-\frac{1}{2\sigma^2}[(x-(\mu+\sigma^2 t))^2] + \mu^2 - (\mu - \sigma^2 t)^2} dx \\ &= e^{-\frac{1}{2\sigma^2}(\mu^2 - (\mu^2 + 2\sigma^2 t\mu + \sigma^4 t^2))} \\ &= e^{\mu t + \frac{1}{2}\sigma^2 t^2} \end{aligned}$$

■

4.3 정규분포의 성질

Properties.

(a) (아핀 변환의 불변)

- (i) $X \sim N(\mu, \sigma^2) \implies aX + b \sim N(a\mu + b, a^2\sigma^2)$, $a, b \in \mathbb{R}$
- (ii) $X \sim N(\mu, \sigma^2) \implies X \stackrel{d}{=} \mu + \sigma Z$, $Z \sim N(0, 1)$

(b) (독립인 정규확률변수들의 합)

(i) $X_i \sim N(\mu_i, \sigma_i^2)$, $i = 1, 2$, 서로 독립 $\implies X_1 + X_2 \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$

(ii) $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} N(\mu, \sigma^2)$ 일 때, $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$

(c) (표준정규확률변수들의 제곱)

(i) $Z \sim N(0, 1) \implies Z^2 \sim \chi^2(1)$

(ii) $Y \sim \chi^2(r) \iff Y = Z_1^2 + \dots + Z_r^2$, $Z_i \stackrel{\text{iid}}{\sim} N(0, 1)$, $i = 1, \dots, r$, $r \in \mathbb{N}$

Proof. (a), (b)는 mgf로 증명.

(c) (i) $Z \sim N(0, 1)$, $Y = Z^2$ 라 하자: $z = \pm\sqrt{y}$

$$\begin{aligned} f_Z(z) &= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz \\ &= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz I(z > 0) + \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz I(z < 0) \\ &= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y} \frac{1}{2} \frac{1}{\sqrt{y}} dy I(z > 0) + \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y} \frac{1}{2} \frac{1}{\sqrt{y}} dy I(z < 0) \\ &= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y} y^{-\frac{1}{2}} dy I(y > 0): Ga(\frac{1}{2}, \text{rate} = \frac{1}{2}) = \chi^2(1) \text{의 밀도함수} \end{aligned}$$

■

5 균등분포와 난수

5.1 균등분포 (uniform distribution)

Notation. $u \sim U(a, b)$, $a < b$

Properties.

1. (밀도함수, 정의) $f(u) = \frac{1}{b-a} I(a < u < b)$
2. $u \sim U(a, b) \iff u \stackrel{d}{=} (b-a)z + a$ where $z \sim U(0, 1)$
3. $u \sim U(0, 1)$ 일 때, $\mathbb{E}u = \frac{1}{2}$, $\text{Var}(u) = \frac{1}{12}$
4. $u \sim U(0, 1)$ 일 때, $1-u \sim U(0, 1)$ ($\because \frac{1}{2}$ 기준 대칭)

5.2 확률적분변환 (probability integral transformation)

Theorem. (교재 4.3.3.) 확률변수 X 의 누적분포함수 F 가 $[a, b]$ 에서 연속이고 순증가. $F(a) = 0$, $F(b) = 1$. a, b 는 각각 $-\infty$ 와 ∞ 가 될 수 있다. 이 때,

(a) $F(x) \sim U(0, 1)$

(b) $u \sim U(0, 1)$ 일 때, $F^{-1}(u) \sim F$

Proof. $-\infty < a < b < \infty$ 인 경우만 고려한다. F 가 순증가이므로 1-1, F 가 연속이므로 전사 ($\because$ 중간값 정리), 따라서 F 의 역함수(F^{-1})가 존재한다.

F^{-1} 역시 순증가함수이며 다음이 성립한다.

$$(i) F(x) \leq u \iff x \leq F^{-1}(u)$$

$$(ii) F^{-1}(u) \leq x \iff u \leq F(x)$$

$$(iii) F(F^{-1}(u)) = u$$

$$(a) u \in (0, 1)$$

$$\begin{aligned} \mathbb{P}(F(x) \leq u) &= \mathbb{P}(X \leq F^{-1}(u)) \quad (\because (i)) \\ &= F(F^{-1}(u)) \quad (\because F가 x의 cdf) \\ &= u \end{aligned}$$

$$(b)$$

$$\begin{aligned} \mathbb{P}(F^{-1}(u) \leq x) &= \mathbb{P}(u \leq F(x)) \quad (\because (ii)) \\ &= F(x) \quad (\because u \sim U(0, 1)) \end{aligned}$$

■

Note. (b)는 임의의 순증가 F 에 대해서도 성립 if $F^{-1} := \inf\{x : F(x) \geq u\}$

Example. $X \sim \text{Exp}(1)$ 인 X 를 생성하고자 ($u \sim U(0, 1)$ 은 주어져 있다.) $F(x) = 1 - e^{-x}$, $x \geq 0$. F^{-1} 을 구하기 위해 $u = F(x) = 1 - e^{-x}$ 로 놓고 x 로 풀면 $x = -\log(1 - u) = F^{-1}(u)$.

$$\therefore u \sim U(0, 1) \text{ 일 때, } \begin{cases} -\log(1 - u) \sim \text{Exp}(1) \\ -\log u \sim \text{Exp}(1) \end{cases}$$

Example. (박스-뮐러 변환, Box-Müller Transformation) $X_1, X_2 \stackrel{\text{iid}}{\sim} N(0, 1)$ 이고 $X_1 = r \cos \theta$, $X_2 = r \sin \theta$ 로 정의된다. ($r > 0$, $0 \leq \theta \leq 2\pi$) 이 때, (r, θ) 의 분포를 구하면

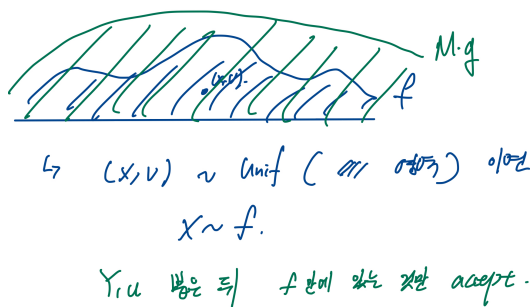
$$\begin{aligned} f_{X_1, X_2}(x_1, x_2) dx_1 dx_2 &= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x_1^2} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x_2^2} dx_1 dx_2 \\ &= \frac{1}{2\pi} e^{-\frac{1}{2}(x_1^2 + x_2^2)} dx_1 dx_2 \\ &= \frac{1}{2\pi} e^{-\frac{1}{2}r^2} r dr d\theta \\ &= e^{-\frac{1}{2}r^2} r I(r > 0) dr \times \frac{1}{2\pi} I(0 < \theta < 2\pi) d\theta \\ &\implies \frac{1}{2}r^2 \sim \text{Exp}(1) \perp\!\!\!\perp \theta \sim U(0, 2\pi) \end{aligned}$$

$u_1, u_2 \stackrel{\text{iid}}{\sim} U(0, 1)$ 가 주어져 있다.

$$X_1 = \sqrt{-2\log(1 - u_1)} \cos(2\pi u_2) \sim N(0, 1) \perp\!\!\!\perp X_2 = \sqrt{-2\log(1 - u_1)} \sin(2\pi u_2) \sim N(0, 1)$$

5.3 합격-불합격 방법 (acceptance-rejection method)

$X \sim f$ 를 생성하고자 $f(x) \leq M \cdot g(x)$, $\forall x, M > 0$, g 를 따르는 난수를 생성한다.

**알고리즘**

1단계. $Y \sim g \perp u \sim U(0, 1)$ (독립인 Y, u 뽑는다.)

2단계. $u \leq \frac{f(Y)}{M \cdot g(Y)}$ 이면, $X = Y$. 아니면 1단계로 간다.

Proof. (알고리즘에서 $X \sim f$ 임을 증명)

$$\begin{aligned} \mathbb{P}(X \leq x) &= \mathbb{P}(Y \leq x | u \leq \frac{f(Y)}{M \cdot g(Y)}) \\ &= \frac{\mathbb{P}(Y \leq x, u \leq \frac{f(Y)}{M \cdot g(Y)})}{\mathbb{P}(u \leq \frac{f(Y)}{M \cdot g(Y)})} = \frac{\int_{-\infty}^x \int_0^{\frac{f(y)}{M \cdot g(y)}} g(y) du dy}{\int_{-\infty}^{\infty} \int_0^{\frac{f(y)}{M \cdot g(y)}} g(y) du dy} \\ &= \frac{\int_{-\infty}^x \frac{f(y)}{M \cdot g(y)} g(y) dy}{\int_{-\infty}^{\infty} \frac{f(y)}{M \cdot g(y)} g(y) dy} = \int_{-\infty}^x f(y) dy \end{aligned}$$

■

Example. (π 의 추정)

$$u_1, u_2 \stackrel{\text{iid}}{\sim} U(0, 1) \implies \mathbb{P}(u_1^2 + u_2^2 < 1) = \frac{\pi}{4}$$

$(u_{i1}, u_{i2}) \stackrel{\text{iid}}{\sim} U(0, 1) \times U(0, 1), i = 1, \dots, n$: 난수 n 개 생성

$\hat{P}_n = \frac{1}{n} \sum_{i=1}^n I(u_{i1}^2 + u_{i2}^2 \leq 1)$ 로 추정, $\hat{\pi} = 4\hat{P}_n \implies$ "몬테카를로 방법"

4장. 표본분포

Jiho KWAK, Department of Geography, SNU

June 2, 2024

1 표본추출과 통계량

표본을 이용해 모집단의 특성에 대한 추론을 한다.

추론의 도식화

1. 모형화 (modeling): 표본 X_i 들이 취할 수 있는 분포들의 모임을 정한다.
 $\{f(x : \theta) : \theta \in H\}$, θ : 모수, $f(x : \theta)$: X_i 의 분포
예 $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} N(\mu, \sigma^2)$, $\mu \in \mathbb{R}$, $\sigma^2 > 0$
 $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} \text{Ber}(\theta)$, $\theta \in (0, 1)$
2. 표본추출 방법을 정한다.
 $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} f(\mu, \theta) : \theta \in H$
3. 추출을 위한 표본의 함수를 정한다. 표본의 함수를 통계량(statistic)이라 한다.
 $T(X_1, \dots, X_n) = \bar{X}$, $T(X_1, \dots, X_n) = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$
4. 통계량의 분포를 구한다.
5. 표본분포(=통계량의 분포)를 이용해 추론을 한다.

표본분포의 계산

1. 특별한 경우만 표본분포가 알려져 있다.
2. 점근적 근사: $n \rightarrow \infty$ 일 때 통계량의 분포
3. 몬테카를로 방법

2 t-분포와 F-분포

2.1 스튜던트의 t-분포

Notation. $T \sim t(r)$, $r = 1, 2, \dots$, $r > 0$, r : 자유도(degree of freedom)

Properties.

1. (정의) $W \sim N(0, 1) \perp V \sim \chi^2(r) \implies T = \frac{W}{\sqrt{V/r}} \sim t(r)$
2. (밀도함수) $f(t) = \frac{\Gamma(\frac{r+1}{2})}{\Gamma(\frac{r}{2})\Gamma(\frac{r}{2})} \frac{1}{\sqrt{r}} \left(1 + \frac{t^2}{r}\right)^{-\frac{r+1}{2}}$, $t \in \mathbb{R}$
 $t \rightarrow \infty$ 일 때 $f(t)$ 의 차수는 $(t^2)^{-\frac{r+1}{2}} = |t|^{-(r+1)} \rightarrow \text{polynomial order}$ (cf. 정규분포: $e^{-\frac{1}{2}t^2}$, 꼬리 더 얇음)

3. $t(1) = Ca(0, 1)$
4. (평균과 분산) $\mathbb{E}T = 0$, $r > 1$, $Var(T) = \frac{r}{r-2}$, $r > 2$
5. $t(x|r) \xrightarrow{r \rightarrow \infty} N(x|0, 1) \iff t(r)$ 의 밀도함수 $\xrightarrow{r \rightarrow \infty} N(0, 1)$ 의 밀도함수

Proof.

2. $T = \frac{W}{\sqrt{V/r}}$, $U = V$ 라고 두자.

Note that $\begin{cases} v = u \\ w = t\sqrt{u/r} \end{cases}$ and $\left| \frac{d(w, v)}{d(t, u)} \right| = \begin{vmatrix} \sqrt{u/r} & \frac{1}{2}tu^{-\frac{1}{2}}t^{-\frac{1}{2}} \\ 0 & 1 \end{vmatrix} = \sqrt{\frac{u}{r}}$

$$\begin{aligned} f_W(w)f_V(v)dwdv &= \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}w^2} \frac{1}{\Gamma(\frac{r}{2})2^{\frac{r}{2}}}v^{\frac{r}{2}-1}e^{-v/2}dwdv \\ &= \frac{1}{\sqrt{2\pi}} \frac{1}{\Gamma(\frac{r}{2})2^{\frac{r}{2}}}e^{-\frac{1}{2}t^2\frac{u}{r}}u^{\frac{r}{2}-1}e^{-u/2}\sqrt{\frac{u}{r}}dtdu \quad (t \in \mathbb{R}, u > 0) \\ &= f_{T,U}(t, u)dtdu \end{aligned}$$

$$\begin{aligned} f_T(t) &= \int_0^\infty f_{T,U}(t, u)du = \frac{1}{\sqrt{2\pi}} \cdot \frac{1}{\Gamma(\frac{r}{2})2^{\frac{r}{2}}} \cdot \frac{1}{\sqrt{r}} \int_0^\infty u^{\frac{r+1}{2}-1}e^{-\frac{1}{2}(1+\frac{t^2}{r})u}du \\ &= \frac{1}{\sqrt{2\pi}\Gamma(\frac{r}{2})2^{r/2}\sqrt{r}} \cdot \frac{\Gamma(\frac{r+1}{2})}{[\frac{1}{2}(1+\frac{t^2}{r})]^{\frac{r+1}{2}}} \\ &= \frac{\Gamma(\frac{r+1}{2})}{\sqrt{2\pi}\Gamma(\frac{r}{2})2^{r/2}(\frac{1}{2})^{\frac{r+1}{2}}} \cdot \frac{1}{\sqrt{r}} \cdot \frac{1}{(1+\frac{t^2}{r})^{\frac{r+1}{2}}} \\ &= \frac{\Gamma(\frac{r+1}{2})}{\Gamma(\frac{1}{2})\Gamma(\frac{r}{2})} \frac{1}{\sqrt{r}} \left(1 + \frac{t^2}{r}\right)^{-\frac{r+1}{2}} \end{aligned}$$

3. 밀도함수에 $r = 1$ 을 대입하면 $f(t) = \frac{\Gamma(1)}{\Gamma(\frac{1}{2})\Gamma(\frac{1}{2})} \cdot \frac{1}{1+t^2} = \frac{1}{\pi} \frac{1}{1+t^2}$
4. 평균이 존재한다면 0임을 알 수 있다. ($\because$ 밀도함수가 0에 대해 대칭)
 t 에 관한 항만을 살펴보면 $\int_{-\infty}^\infty |t| \frac{1}{(1+\frac{t^2}{r})^{\frac{r+1}{2}}} dt$ 이므로 t 의 차수는 $\frac{1}{t^r}$ 이다.
따라서 $r > 1$ 이면 수렴(평균이 존재)하고 $0 < r \leq 1$ 이면 발산한다.
본격적인 증명을 위해 $\int_0^\infty |t| \frac{1}{(1+\frac{t^2}{r})^{\frac{r+1}{2}}} dt = \left(\int_0^1 + \int_1^\infty \right) t \frac{1}{(1+\frac{t^2}{r})^{\frac{r+1}{2}}} dt = I_1 + I_2$ 라 두자.

$$\begin{aligned} I_1 &= \int_0^1 \frac{t}{(1+\frac{t^2}{r})^{\frac{r+1}{2}}} dt < \infty \\ I_2 &= \int_1^\infty \frac{t}{(1+\frac{t^2}{r})^{\frac{r+1}{2}}} dt \leq \int_1^\infty \frac{t}{(\frac{t^2}{r})^{\frac{r+1}{2}}} dt = r^{\frac{r+1}{2}} \int_1^\infty t^{-r} dr < \infty, \quad r > 1 \\ I_2 &= \int_1^\infty \frac{t}{(1+\frac{t^2}{r})^{\frac{r+1}{2}}} dt \geq \int_1^\infty \frac{t}{(\frac{t^2}{r} + \frac{t^2}{r})^{\frac{r+1}{2}}} dt = \left(\frac{r}{2}\right)^{\frac{r+1}{2}} \int_1^\infty t^{-r} dr = \infty, \quad 0 < r \leq 1 \end{aligned}$$

(* n 차 적분이 존재하려면 $r > n$)

Note that $\mathbb{E}(W^2) = 1$ and

$$\begin{aligned}\mathbb{E}\left(\frac{1}{V}\right) &= \int_0^\infty \frac{1}{v} \cdot \frac{1}{\Gamma(\frac{r}{2})2^{\frac{r}{2}}} \cdot v^{\frac{r}{2}-1} e^{-\frac{v}{2}} dv \\ &= \frac{1}{\Gamma(\frac{r}{2})2^{\frac{r}{2}}} \int_0^\infty v^{\frac{r}{2}-1-1} e^{-\frac{v}{2}} dv \\ &= \frac{1}{\Gamma(\frac{r}{2})2^{\frac{r}{2}}} \Gamma\left(\frac{r}{2}-1\right) 2^{\frac{r}{2}-1} \left(\frac{r}{2}-1 > 0\right) \\ &= \frac{1}{(\frac{r}{2}-1)2} = \frac{1}{r-2}\end{aligned}$$

$$\begin{aligned}\therefore \text{Var}(T) &= \text{Var}\left(\frac{W}{\sqrt{V/r}}\right) = \mathbb{E}\left(\frac{W^2}{V/r}\right) \\ &= r \cdot \mathbb{E}(W^2) \cdot \mathbb{E}\left(\frac{1}{V}\right) \quad (\because W, V \text{ 독립}) \\ &= \frac{r}{r-2}\end{aligned}$$

5. $\Gamma(x) \sim (x-1)! \sim \sqrt{2\pi(x-1)}\left(\frac{x-1}{e}\right)^{x-1}$ 임을 이용한다.

$$\begin{aligned}f(t) &= \frac{\Gamma(\frac{r+1}{2})}{\Gamma(\frac{1}{2})\Gamma(\frac{r}{2})} \cdot \frac{1}{\sqrt{r}} \left(1 + \frac{t^2}{r}\right)^{-\frac{r+1}{2}} \\ &\sim \frac{\sqrt{2\pi(\frac{r+1}{2}-1)}\left(\frac{\frac{r+1}{2}-1}{e}\right)^{\frac{r+1}{2}-1}}{\sqrt{2\pi(\frac{r}{2}-1)}\left(\frac{\frac{r}{2}-1}{e}\right)^{\frac{r}{2}-1}} \frac{1}{\sqrt{\pi}} \frac{1}{\sqrt{r}} \left(1 + \frac{t^2}{r}\right)^{-\frac{r+1}{2}} \\ &= (\text{계산})\end{aligned}$$

$$\ast \left(1 + \frac{t^2}{r}\right)^{-\frac{r+1}{2}} = \left(1 + \frac{t^2}{r}\right)^{-\frac{r}{2}} \cdot \left(1 + \frac{t^2}{r}\right)^{-\frac{1}{2}} \simeq e^{-\frac{1}{2}t^2}$$

■

2.2 F-분포

Notation. $F \sim F(r_1, r_2)$, $r_1, r_2 > 0$

Properties.

1. (정의) $U \sim \chi^2(r_1) \perp\!\!\!\perp V \sim \chi^2(r_2)$, $F := \frac{U/r_1}{V/r_2} \sim F(r_1, r_2)$
2. (밀도함수) $f(w) = \frac{\Gamma(\frac{r_1+r_2}{2})}{\Gamma(\frac{r_1}{2})\Gamma(\frac{r_2}{2})} \left(\frac{r_1}{r_2}\right)^{\frac{r_1}{2}} \frac{w^{\frac{r_1}{2}-1}}{(1 + \frac{r_1}{r_2}w)^{\frac{r_1+r_2}{2}}}$, $w > 0$
3. (적률) $\mathbb{E}F^k = \left(\frac{r_2}{r_1}\right)^k \frac{\Gamma(k + \frac{r_1}{2})\Gamma(\frac{r_2}{2} - k)}{\Gamma(\frac{r_1}{2})\Gamma(\frac{r_2}{2})}$, $r_2 > 2k$
4. (평균) $\mathbb{E}F = \frac{r_2}{r_2 - 2}$, $r_2 > 2$
5. (분산) $\text{Var}(F) = \frac{r_2^2(2r_1 + 2r_2 - 4)}{r_1(r_2 - 2)^2(r_2 - 4)}$, $r_2 > 4$

Proof.

$$(b) \begin{cases} W = \frac{U/r_1}{V/r_2} \\ Z = V \end{cases} \implies \begin{cases} U = \frac{r_1}{r_2} WZ \\ V = Z \end{cases}, \left| \frac{d(w, z)}{d(u, v)} \right| = \begin{vmatrix} \frac{r_1}{r_2} z & \frac{r_1}{r_2} w \\ 0 & 1 \end{vmatrix} = \frac{r_1}{r_2} z$$

$$\begin{aligned} f(u)f(v)dudv &= \frac{1}{\Gamma(\frac{r_1}{2})2^{\frac{r_1}{2}}} u^{\frac{r_1}{2}-1} e^{-u/2} \times \frac{1}{\Gamma(\frac{r_2}{2})2^{\frac{r_2}{2}}} v^{\frac{r_2}{2}-1} e^{-v/2} dudv \\ &= \frac{1}{\Gamma(\frac{r_1}{2})\Gamma(\frac{r_2}{2})2^{\frac{r_1+r_2}{2}}} \left(\frac{r_1}{r_2} wz \right)^{\frac{r_1}{2}-1} e^{-\frac{1}{2}\frac{r_1}{r_2} wz} z^{\frac{r_2}{2}-1} e^{-\frac{z}{2}} \frac{r_1}{r_2} z dw dz \\ &= \frac{1}{\Gamma(\frac{r_1}{2})\Gamma(\frac{r_2}{2})2^{\frac{r_1+r_2}{2}}} \left(\frac{r_1}{r_2} \right)^{\frac{r_1}{2}} w^{\frac{r_1}{2}-1} z^{\frac{r_1+r_2}{2}-1} e^{-\frac{1}{2}(\frac{r_1}{r_2} w + 1)z} dw dz \end{aligned}$$

$$\begin{aligned} f(w) &= \int_0^\infty f(w, z) dz \\ &= \frac{1}{\Gamma(\frac{r_1}{2})\Gamma(\frac{r_2}{2})2^{\frac{r_1+r_2}{2}}} \left(\frac{r_1}{r_2} \right)^{\frac{r_1}{2}} w^{\frac{r_1}{2}-1} \Gamma\left(\frac{r_1+r_2}{2}\right) / \left[\frac{1}{2} \left(1 + \frac{r_1}{r_2} w \right) \right]^{\frac{r_1+r_2}{2}} \\ &= \frac{\Gamma(\frac{r_1+r_2}{2})}{\Gamma(\frac{r_1}{2})\Gamma(\frac{r_2}{2})} \left(\frac{r_1}{r_2} \right)^{\frac{r_1}{2}} \frac{w^{\frac{r_1}{2}-1}}{\left(1 + \frac{r_1}{r_2} w \right)^{\frac{r_1+r_2}{2}}}, \quad w > 0 \end{aligned}$$

■

Corollary. $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} N(\mu_1, \sigma_1^2) \perp\!\!\!\perp Y_1, \dots, Y_m \stackrel{\text{iid}}{\sim} N(\mu_2, \sigma_2^2) \implies \frac{s_2^2/\sigma_2^2}{s_1^2/\sigma_1^2} \sim F(m-1, n-1)$,

이때 $s_1^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$, $s_2^2 = \frac{1}{m-1} \sum_{j=1}^m (Y_j - \bar{Y})^2$

Proof.

$$\frac{(n-1)s_1^2}{\sigma_1^2} \sim \chi^2(n-1) \perp\!\!\!\perp \frac{(m-1)s_2^2}{\sigma_2^2} \sim \chi^2(m-1) \text{임을 안다.}$$

$$\frac{\frac{(n-1)s_1^2}{\sigma_1^2}/(n-1)}{\frac{(m-1)s_2^2}{\sigma_2^2}/(m-1)} = \frac{s_2^2/\sigma_2^2}{s_1^2/\sigma_1^2} \sim F(m-1, n-1)$$

■

Note. 이 사실을 이용해 $\frac{\sigma_1^2}{\sigma_2^2}$ 의 $100(1-\alpha)\%$ 신뢰구간을 만들면

$$\begin{aligned} &\mathbb{P}\left(F^{\frac{\alpha}{2}}(m-1, n-1) < \frac{s_2^2}{s_1^2} \cdot \frac{\sigma_1^2}{\sigma_2^2} < F_{\frac{\alpha}{2}}(m-1, n-1)\right) = 1 - \alpha \\ \iff &\mathbb{P}\left(F^{\frac{\alpha}{2}}(m-1, n-1) \cdot \frac{s_1^2}{s_2^2} < \frac{\sigma_1^2}{\sigma_2^2} < F_{\frac{\alpha}{2}}(m-1, n-1) \cdot \frac{s_1^2}{s_2^2}\right) = 1 - \alpha \end{aligned}$$

3 순서통계량의 분포

Definition. $X_1, \dots, X_n$: 랜덤포본

$X_{(r)} = X_1, \dots, X_n$ 중에서 r 번째로 작은 값 $\implies r$ 번째 순서통계량(order statistic), $r = 1, \dots, n$

$X_{(1)}$: 가장 작은 값, $X_{(n)}$: 가장 큰 값

Properties. $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} f$. f : 밀도함수, F : cdf, $Y = (Y_1, \dots, Y_n) = (X_{(1)}, \dots, X_{(n)})$

(a) Y 의 밀도함수: $f_Y(y) = n! \prod_{i=1}^n f(y_i) I(y_1 < y_2 < \dots < y_n)$

(b) $X_{(r)}$ 의 밀도함수: $f_{X_{(r)}}(x) = \frac{n!}{(r-1)!(n-r)!} F(x)^{r-1} f(x) (1-F(x))^{n-r}$

(c) $1 \leq r < s \leq n$, $(X_{(r)}, X_{(s)})$ 의 밀도함수:

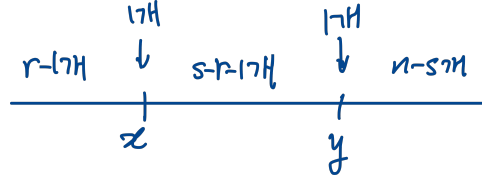
$$f_{X_{(r)}, X_{(s)}}(x, y) = \frac{n!}{(r-1)!(s-r-1)!(n-s)!} F(x)^{r-1} f(x) (F(y)F(x))^{s-r-1} f(y) (1-F(y))^{n-s}, \quad x < y$$

Note. (c)를 기억하는 법: X_i 들이 이산형이라 가정하자.

$\mathbb{P}(X_{(r)} = x, X_{(s)} = y) = \mathbb{P}(X_1, \dots, X_n \text{ 중에서 } r-1 \text{개가 } x \text{보다 작고, 한 개가 } x$

$s-r-1 \text{개가 } x \text{와 } y \text{ 사이에, 한 개가 } y, n-s \text{개가 } y \text{보다 큼})$

$$= \frac{n!}{(r-1)!(s-r-1)!(n-s)!} F(x)^{r-1} f(x) (F(y)F(x))^{s-r-1} f(y) (1-F(y))^{n-s}$$



Proof. $y = (y_1, \dots, y_n) = (x_{(1)}, \dots, x_{(n)})$, $x = (x_1, \dots, x_n)$

(a) $\pi: [n] \rightarrow [n] = \{1, \dots, n\}$: 순열

Note that $y \rightarrow x$: 다대일 ($n! \rightarrow 1$)

$$\begin{aligned} f_X(x) dx &= \prod_{i=1}^n f(x_i) dx = \sum_{\pi} \prod_{i=1}^n f(x_i) dx I(x_{\pi(1)} < x_{\pi(2)} < \dots < x_{\pi(n)}) \\ &= \sum_{\pi} \prod_{i=1}^n f(y_i) \left| \frac{dx}{dy} \right| dy I(y_1 < \dots < y_n) \\ &= \sum_{\pi} \prod_{i=1}^n f(y_i) dy I(y_1 < \dots < y_n) = n! \prod_{i=1}^n f(y_i) dy I(y_1 < \dots < y_n) \end{aligned}$$

* $y_1 = x_{\pi(1)}, \dots, y_n = x_{\pi(n)} \iff x_1 = y_{\pi^{-1}(1)}, \dots, x_n = y_{\pi^{-1}(n)}$ 이므로

$$\left| \det \left(\frac{dx}{dy} \right) \right| = \left| \det \begin{pmatrix} \ddots & 1 & \ddots & \ddots \\ & \ddots & \ddots & \ddots \\ & & \ddots & 1 \end{pmatrix} \right| = |\pm 1| = 1$$

(b)

$$\begin{aligned}
f_{X_{(r)}}(y_r) &= \int \dots \int n! \prod_{i=1}^n f(y_i) I(y_1 < \dots < y_{r-1} < y_r < y_{r+1} < \dots < y_n) dy_1 \dots dy_{r-1} dy_{r+1} \dots dy_n \\
&= n! f(y_r) \int \dots \int \prod_{i=1}^{r-1} I(y_1 < \dots < y_{r-1} < y_r) dy_1 \dots dy_{r-1} \\
&\quad \times \int \dots \int \prod_{i=r+1}^n f(y_i) I(y_r < y_{r+1} < \dots < y_n) dy_{r+1} \dots dy_n \\
&= n! f(y_r) \frac{1}{(r-1)!} \int \dots \int \prod_{i=1}^{r-1} f(y_i) I(y_1, \dots, y_{r-1} < y_r) dy_1 \dots dy_{r-1} \\
&\quad \times \frac{1}{(n-r)!} \int \dots \int \prod_{i=r+1}^n f(y_i) I(y_r < y_{r+1}, \dots, y_n) dy_{r+1} \dots dy_n \\
&= \frac{n!}{(r-1)!(n-r)!} F(x)^{r-1} f(x) (1-F(x))^{n-r}
\end{aligned}$$

■

Theorem. (균등분포의 순서통계량) $u_1, \dots, u_n \stackrel{\text{iid}}{\sim} U(0, 1)$

(a) $U_{(r)} \sim \text{Beta}(r, n-r+1)$ (b) $Z = (u_{(1)}, u_{(2)} - u_{(1)}, \dots, u_{(n)} - u_{(n-1)}) \sim \text{Dir}(1, \dots, 1)$ *Proof.*(a) $f(x) = 1, F(x) = x, x \in (0, 1)$

$$\begin{aligned}
f_{u_{(r)}}(u) &= \frac{n!}{(r-1)!(n-r)!} u^{r-1} \cdot 1 \cdot (1-u)^{n-r} \\
&= \frac{\Gamma(n-1)}{\Gamma(r)\Gamma(n-r+1)} u^{r-1} (1-u)^{n-r} = \text{Beta}(u|r, n+1)
\end{aligned}$$

(b) $y = (y_1, \dots, y_n) = (u_{(1)}, \dots, u_{(n)})$

$$Z = \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 - y_1 \\ \vdots \\ y_n - y_{n-1} \end{pmatrix}, \quad y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} z_1 \\ z_1 + z_2 \\ \vdots \\ z_1 + \dots + z_n \end{pmatrix} \Rightarrow \det \left| \frac{dy}{dz} \right| = \det \begin{pmatrix} 1 & 0 & \dots & 0 & 0 \\ 1 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & \dots & \dots & 1 \end{pmatrix} = 1$$

$$\begin{aligned}
f_Y(y) dy &= n! I(0 < y_1 < \dots < y_n < 1) dy \\
&= n! I(0 < z_1 < z_1 + z_2 < \dots < z_1 + \dots + z_n < 1) \left| \frac{dy}{dz} \right| dz \\
&= n! I(z_i > 0, \sum_{i=1}^n z_i < 1) dz \\
&= \frac{n!}{1 \times \dots \times 1} z_1^{1-1} \times \dots \times z_n^{1-1} dz I(z_i > 0, \sum_{i=1}^n z_i < 1) \\
&= \frac{\Gamma(n+1)}{\Gamma(1) \dots \Gamma(1)} z_1^{1-1} \times \dots \times z_n^{1-1} dz I(z_i > 0, \sum_{i=1}^n z_i < 1) = \text{Dir}(z|1 \dots 1)
\end{aligned}$$

■

Example. (이항분포 꼬리확률) $X \sim \text{Bin}(n, p)$

$\mathbb{P}(X \geq k) = \mathbb{P}(u_{(k)} \leq p)$, $u_{(k)} \sim \text{Beta}(k, n - k + 1)$ 로 계산 가능함을 보이자.

Proof.

$u_1, \dots, u_n \stackrel{\text{iid}}{\sim} U(0, 1)$, $X = \sum_{i=1}^n I(u_i \leq p)$ 로 표현 가능

$$\begin{aligned}\mathbb{P}(X \geq k) &= \mathbb{P}(u_i \leq p \text{인 } u_i \text{의 개수가 } k \text{보다 크다}) \\ &= \mathbb{P}(u_{(k)} \leq p)\end{aligned}$$

$$\text{*수식으로 쓰면 } \sum_{x=k}^n \binom{n}{x} p^x (1-p)^{n-x} = \int_p^1 \frac{\Gamma(n+1)}{\Gamma(k)\Gamma(n-k+1)} u^{k-1} (1-u)^{n-k} du$$

■

Theorem. (지수분포의 순서통계량) $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} \text{Exp}(1)$

$$\begin{aligned} \text{(a)} \quad & \left. \begin{aligned} Z_1 &= nX_{(1)} \\ Z_2 &= (n-1)(X_{(2)} - X_{(1)}) \\ &\vdots \\ Z_n &= X_{(n)} - X_{(n-1)} \end{aligned} \right\} \stackrel{\text{iid}}{\sim} \text{Exp}(1) \\ \text{(b)} \quad & \left. \begin{aligned} X_{(1)} &= \frac{1}{n} Z_1 \\ X_{(2)} &= \frac{1}{n} Z_1 + \frac{1}{n-1} Z_2 \\ &\vdots \\ X_{(n)} &= \frac{1}{n} Z_1 + \frac{1}{n-1} + \dots + Z_n \end{aligned} \right\}, Z_1, \dots, Z_n \stackrel{\text{iid}}{\sim} \text{Exp}(1) \end{aligned}$$

Note. 지수분포의 경우, 순서통계량을 iid 확률변수의 함수로 생각할 수 있다.

Proof. (b)는 (a)로부터 나온다.

(a) $Y = (X_{(1)}, \dots, X_{(n)})$

$$\begin{aligned} z_1 &= ny_1 & y_1 &= \frac{1}{n} z_1 \\ z_2 &= (n-1)(y_2 - y_1) & y_2 &= \frac{1}{n} z_1 + \frac{1}{n-1} z_2 \\ &\vdots & &\vdots \\ z_n &= y_n - y_{n-1} & y_n &= \frac{1}{n} z_1 + \dots + z_n \end{aligned} \implies \det\left(\frac{dy}{dz}\right) = \det\begin{pmatrix} \frac{1}{n} & \dots & \dots & 0 \\ \frac{1}{n} & \frac{1}{n-1} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{n} & \dots & \dots & 1 \end{pmatrix} = \frac{1}{n!}$$

$$\begin{aligned} f_Y(y) dy &= n! \prod_{i=1}^n e^{-y_i} I(0 < y_1 < \dots < y_n) dy \\ &= n! \prod_{i=1}^n e^{-(\frac{1}{n} z_1 + \frac{1}{n-1} z_2 + \dots + \frac{1}{n-i+1} z_i)} I(0 < \frac{1}{n} z_1 < \frac{1}{n} z_1 + \frac{1}{n-1} z_2 < \dots) \left| \frac{dy}{dz} \right| dz \\ &= \prod_{i=1}^n e^{-z_i} I(z_1, \dots, z_n > 0) dz \end{aligned}$$

$$= \prod_{i=1}^n \text{Exp}(z_i|1) dz$$

■

Example. $u_1, \dots, u_n \stackrel{\text{iid}}{\sim} U(0, 1), Y_1, \dots, Y_n \stackrel{\text{iid}}{\sim} \text{Exp}(1) \implies (u_{(1)}, \dots, u_{(n)}) \stackrel{d}{=} (1 - e^{-Y_{(r)}})_{1 \leq r \leq n}$

Theorem. $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} F, h(y) := F^{-1}(1 - e^{-y}), z_1, \dots, z_n \stackrel{\text{iid}}{\sim} \text{Exp}(1)$

$$(X_{(r)})_{1 \leq r \leq n} \stackrel{d}{=} \left(h \left(\frac{1}{n} z_1 + \dots + \frac{1}{n-r+1} z_r \right) \right)_{1 \leq r \leq n}$$

4 다변량 정규분포

4.1 정의

Definition. n 차원 확률벡터가 다음 중 하나의 조건을 만족하면 $X \sim N_n(\mu, \Sigma)$ 라 정의한다. Σ 는 대칭이고 음이 아닌 정부호(nonnegative definite)인 $n \times n$ 실수행렬, $\mu \in \mathbb{R}^n, X \in \mathbb{R}^n$, 다음의 네 가지 조건은 동치.

(a) $X \stackrel{d}{=} AZ + \mu, A_{n \times k} \times A'_{k \times n} = \Sigma, \text{rank}(A) = k$. 여기서 $Z = (Z_1, \dots, Z_k)', Z_1, \dots, Z_k \stackrel{\text{iid}}{\sim} N(0, 1)$

(b) $X \stackrel{d}{=} \Sigma^{\frac{1}{2}} Z + \mu, \Sigma^{\frac{1}{2}} \Sigma^{\frac{1}{2}} = \Sigma, Z = (Z_1, \dots, Z_k)', Z_1, \dots, Z_k \stackrel{\text{iid}}{\sim} N(0, 1)$

(c) $M_X(t) = e^{t'\mu + \frac{1}{2}t'\Sigma t}, t \in \mathbb{R}^n$

(d) $a'X \sim N(a'\mu, a'\Sigma a), \forall a \in \mathbb{R}^n$

Note. $\Sigma = PDP', D = \text{diag}(\lambda_1, \dots, \lambda_n) \implies \Sigma^{\frac{1}{2}} = PD^{\frac{1}{2}}P', D^{\frac{1}{2}} = \text{diag}(\lambda_1^{\frac{1}{2}}, \dots, \lambda_n^{\frac{1}{2}})$

- 행렬 $\Sigma = (\sigma_{ij})_{1 \leq i, j \leq n}$ 이 대칭 $\iff \sigma_{ij} = \sigma_{ji}, \forall i, j$
- 행렬 Σ 가 음이 아닌 정부호: $a'\Sigma a \geq 0, \forall a \in \mathbb{R}^n$
- 행렬 Σ 가 양의 정부호: $a'\Sigma a > 0, \forall a \neq 0$ 인 $\mathbb{R}^n$ -벡터
- 대각화정리: Σ 가 대칭, 음이 아닌 정부호 $\implies \Sigma = PDP'$. 여기서 P 는 직교행렬(i.e. $PP' = I$)이고 $D = \text{diag}(\lambda_1, \dots, \lambda_n), \lambda_i \geq 0$

Proof. (a) $\iff$ (c)

($\implies$) (역은 mgf의 유일성에 의해 성립)

$$\begin{aligned} \mathbb{E}e^{t'X} &= \mathbb{E}e^{t'(AZ+\mu)} \\ &= e^{t'\mu} \cdot \mathbb{E}e^{t'AZ} \quad (A = [a_1, \dots, a_k], t'AZ = [t'a_1, \dots, t'a_k] \begin{bmatrix} z_1 \\ \vdots \\ z_k \end{bmatrix}) \\ &= e^{t'\mu} \cdot \mathbb{E}e^{\sum_{i=1}^k t'a_i \cdot z_i} = e^{t'\mu} \cdot \prod_{i=1}^k \mathbb{E}e^{(t'a_i)z_i} \\ &= e^{t'\mu} \cdot \prod_{i=1}^k e^{\frac{1}{2}(t'a_i)^2} = e^{t'\mu} \cdot e^{\frac{1}{2} \sum_{i=1}^k (t'a_i)^2} \\ &= e^{t'\mu} \cdot e^{\frac{1}{2}t'(\sum_{i=1}^n a_i a'_i)t} = e^{t'\mu + \frac{1}{2}t'\Sigma t} \end{aligned}$$

(b) $\iff$ (c)는 위와 같은 방식으로 보일 수 있다.

(c) $\iff$ (d)

($\implies$)

$$\begin{aligned}\mathbb{E}e^{sa'X} &= \mathbb{E}e^{(s \cdot a)' \cdot X} : a'X \text{의 mgf} \\ &= e^{(sa)' \mu + \frac{1}{2}(sa)' \Sigma (sa)} \\ &= e^{s(a' \mu) + \frac{1}{2}s^2(a' \Sigma a)} : N(a' \mu, a' \Sigma a) \text{의 mgf}\end{aligned}$$

($\impliedby$)

$$\begin{aligned}\mathbb{E}e^{t'X} &= \mathbb{E}e^{1 \cdot (t'X)} \\ &= e^{1 \cdot (t' \mu) + \frac{1}{2}1^2 \cdot t' \Sigma t} (\because t'X \sim N(t' \mu, t' \Sigma t)) \\ &= e^{t' \mu + \frac{1}{2}t' \Sigma t}\end{aligned}$$

■

Properties. (밀도함수) $X \sim N_n(\mu, \Sigma)$, Σ : 양의 정부호. 이때 X 의 밀도함수는

$$f_X(x) = |2\pi\Sigma|^{-\frac{1}{2}} e^{-\frac{1}{2}(x-\mu)' \Sigma^{-1}(x-\mu)}, \quad x \in \mathbb{R}^n$$

Proof. $X = \Sigma^{\frac{1}{2}}Z + \mu$, $Z = (Z_1, \dots, Z_n)'$, $Z_i \stackrel{\text{iid}}{\sim} N(0, 1)$

Note that $z = \Sigma^{-\frac{1}{2}}(x - \mu) \implies dz = |\Sigma^{-\frac{1}{2}}|dx$

$$\begin{aligned}f_Z(z)dz &= \prod_{i=1}^n \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z_i^2} dz_1 \cdots dz_n \\ &= (2\pi)^{-\frac{n}{2}} \cdot e^{-\frac{1}{2}(x-\mu)' \Sigma^{-\frac{1}{2}} \Sigma^{-\frac{1}{2}} (x-\mu)} \times |\Sigma^{-\frac{1}{2}}|dx \\ &= |2\pi\Sigma|^{-\frac{1}{2}} e^{-\frac{1}{2}(x-\mu)' \Sigma^{-1}(x-\mu)} dx\end{aligned}$$

■

Properties. $X \sim N_n(\mu, \Sigma) \implies \mathbb{E}X = \mu$, $\text{Var}(X) = \Sigma$

4.2 성질들

Properties.

1. (아핀 불변성, affine invariance of MVN family) $X \sim N(\mu, \Sigma) \implies AX + b \sim N(A\mu + b, A\Sigma'A)$
2. (부분벡터의 독립성)

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim N \left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} \right) \text{일 때, } X_1 \perp\!\!\!\perp X_2 \iff \Sigma_{12} = \text{Cov}(X_1, X_2) = 0$$

where $X_1, \mu_1 \in \mathbb{R}^{n_1}$, $X_2, \mu_2 \in \mathbb{R}^{n_2}$, $n_1 + n_2 = n$, $\Sigma_{11} \in \mathbb{R}^{n_1 \times n_1}$, $\Sigma_{12} \in \mathbb{R}^{n_1 \times n_2}$, $\Sigma_{22} \in \mathbb{R}^{n_2 \times n_2}$

3. (선형변환들의 독립성) $X \sim N(\mu, \Sigma)$, $AX \perp\!\!\!\perp BX \iff \text{Cov}(AX, BX) = A\Sigma B' = 0$

Proof.

$$1. X = \Sigma^{\frac{1}{2}}Z + \mu, Z \sim N(0, 1)$$

$$\begin{aligned} AX + b &= A(\Sigma^{\frac{1}{2}}Z + \mu) + b \\ &= A\Sigma^{\frac{1}{2}}Z + A\mu + b \sim N(A\mu + b, A\Sigma^{\frac{1}{2}}\Sigma^{\frac{1}{2}}A' = A\Sigma A') \end{aligned}$$

$$2. M_{X_1}(t) = M_{X_1, X_2}\left(\begin{pmatrix} t \\ 0 \end{pmatrix}\right) = e^{t'\mu_1 + \frac{1}{2}t'\Sigma_{11}t}; N_{n_1} \text{의 mgf. } X_2 \text{에 대해서도 마찬가지로 정의.}$$

$$\begin{aligned} X_1 \perp\!\!\!\perp X_2 &\iff M_{X_1, X_2}\left(\begin{pmatrix} t_1 \\ t_2 \end{pmatrix}\right) = M_{X_1}(t_1)M_{X_2}(t_2), \forall t_1, t_2 \\ &\iff \exp\left((t'_1, t'_2)\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} + \frac{1}{2}(t'_1, t'_2)\begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}\begin{pmatrix} t_1 \\ t_2 \end{pmatrix}\right) \\ &= \exp(t'_1\mu_1 + \frac{1}{2}t'_1\Sigma_{11}t_1) \times \exp(t'_2\mu_2 + \frac{1}{2}t'_2\Sigma_{22}t_2), \forall t_1, t_2 \\ &\iff e^{t_1\Sigma_{12}t_2} = 1, \forall t_1, t_2 \\ &\iff t'_1\Sigma_{12}t_2 = 0, \forall t_1, t_2 \\ &\iff \Sigma_{12} = 0 \end{aligned}$$

$$3. \begin{bmatrix} AX \\ BX \end{bmatrix} \sim N\left(\begin{pmatrix} A\mu \\ B\mu \end{pmatrix}, \begin{pmatrix} A\Sigma A' & A\Sigma B \\ B\Sigma A' & B\Sigma B \end{pmatrix}\right) \text{ (Affine transformation)}$$

$$AX \perp\!\!\!\perp BX \iff A\Sigma B' = 0$$

■

Properties. (조건부분포와 주변분포) $\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}\right)$, Σ_{22} : 양의 정부호

$$\implies (i) X_1|X_2 = x_2 \sim N(\mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(x_2 - \mu_2), \Sigma_{11.2} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21})$$

$$(ii) X_2 \sim N(\mu_2, \Sigma_{22})$$

Proof.

$$(i) \begin{pmatrix} X_1 - \Sigma_{12}\Sigma_{22}^{-1}X_2 \\ X_2 \end{pmatrix} = \begin{pmatrix} I & -\Sigma_{12}\Sigma_{22}^{-1} \\ 0 & I \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_1 - \Sigma_{12}\Sigma_{22}^{-1}\mu_2 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} & 0 \\ 0 & \Sigma_{22} \end{pmatrix}\right)$$

$$\implies X_1 - \Sigma_{12}\Sigma_{22}^{-1}X_2 \perp\!\!\!\perp X_2 \text{ 이므로}$$

$$X_1 - \Sigma_{12}\Sigma_{22}^{-1}X_2|X_2 = x_2 \stackrel{d}{=} X_1 - \Sigma_{12}\Sigma_{22}^{-1}X_2 \sim N(\mu_1 - \Sigma_{12}\Sigma_{22}^{-1}\mu_2, \Sigma_{11.2})$$

$$X_1|X_2 = x_2 \stackrel{d}{=} X_1 - \Sigma_{12}\Sigma_{22}^{-1}X_2 + \Sigma_{12}\Sigma_{22}^{-1}X_2|X_2 = x_2$$

$$\sim N(\mu_1 - \Sigma_{12}\Sigma_{22}^{-1}\mu_2 + \Sigma_{12}\Sigma_{22}^{-1}x_2, \Sigma_{11.2})$$

(ii)는 위의 성질 2에서 X_2 의 mgf를 통해 증명

■

4.3 정규분포의 기본 정리

Properties. (이차식의 분포) $Z \sim N(0, I)$, A : 대칭, 멱등(idempotent) $\implies Z'AZ \sim \chi^2(r)$, $r = \text{tr}(A)$

* A 가 멱등 $\iff A \cdot A = A$

Proof.

A 가 대칭이고 멱등행렬이면, A 는 음이 아닌 정부호 ($\because x'Ax = x'AAx = x'A'A'x = \|Ax\|^2 \geq 0$)

대각화정리를 이용하면 $A = PDP'$, P : 직교, $D = \text{diag}(\lambda_1, \dots, \lambda_n)$

A 가 멱등행렬이면 A 의 고유값들 즉 λ_i 들은 0 혹은 1 ($\because A = A^2 = PDP'PDP' = PD^2P', D = D^2$)

$\text{tr}(A) = \text{tr}(PDP') = \text{tr}(D) = r = \text{rank}(A)$ 이고, P : 직교행렬이므로

$$Z'AZ = Z'PDP'Z = W'DW \quad (W = P'Z = (W_1, \dots, W_n)' \sim N(0, 1))$$

$$= \sum_{i=1}^r W_i^2 \sim \chi^2(r)$$

■

Theorem. (정규분포의 기본정리) $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} N(\mu, \sigma^2)$

$\implies$ (a) $\bar{X} \perp\!\!\!\perp s^2$

(b) $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$

(c) $\frac{(n-1)s^2}{\sigma^2} = \sum_{i=1}^n \frac{(X_i - \bar{X})^2}{\sigma^2} \sim \chi^2(n-1)$

(d) (b)와 (c)는 서로 독립

Proof. $X = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix} \sim N_n(\mu \cdot 1, \sigma^2 I_n)$ 이므로 $\begin{cases} \bar{X} = \frac{1}{n} 1'X \\ (n-1)s^2 = \sum_{i=1}^n (X_i - \bar{X})^2 = X'(I - \frac{1}{n}J)X \end{cases}$

(d) $\bar{X}$ 와 $(X_1 - \bar{X}, \dots, X_n - \bar{X})$ 가 독립인 것을 보이자.

$$\begin{aligned} M_{\bar{X}, (X_1 - \bar{X}, \dots, X_n - \bar{X})}(s, t) &= \mathbb{E} e^{s\bar{X} + \sum_{i=1}^n t_i (X_i - \bar{X})}, \quad s \in \mathbb{R}, \quad t = (t_1, \dots, t_n) \in \mathbb{R}^n \\ &= \mathbb{E} e^{(s - \sum_{i=1}^n t_i)\bar{X} + \sum_{i=1}^n t_i X_i} \\ &= \mathbb{E} e^{\sum_{i=1}^n (\frac{1}{n}s - \bar{t} + t_i)X_i} \\ &= \prod_{i=1}^n M_{X_i}(\frac{1}{n}s - \bar{t} + t_i) = \prod_{i=1}^n e^{(\frac{1}{n}s - \bar{t} + t_i)\mu + \frac{1}{2}\sigma^2(\frac{1}{n}s - \bar{t} + t_i)^2} \\ &= e^{\sum_{i=1}^n (\frac{1}{n}s - \bar{t} + t_i)\mu + \frac{1}{2}\sigma^2(\frac{1}{n}s - \bar{t} + t_i)^2} \\ &= e^{s\mu + \frac{1}{2}\sigma^2 \sum_{i=1}^n (\frac{1}{n^2}s^2 + (t_i - \bar{t})^2 + \frac{2}{n}s(t_i - \bar{t}))} \\ &= e^{s\mu + \frac{1}{2}\sigma^2(\frac{s^2}{n} + \sum_{i=1}^n (t_i - \bar{t})^2)} \\ &= e^{s\mu + \frac{1}{2n}\sigma^2 s^2} \cdot e^{\frac{1}{2}\sigma^2 \sum_{i=1}^n (t_i - \bar{t})^2} = (s \text{의 함수}) \times (t \text{의 함수}) \end{aligned}$$

(c) $\sum_{i=1}^n \frac{(X_i - \mu)^2}{\sigma^2} = \sum_{i=1}^n \frac{(X_i - \bar{X})^2}{\sigma^2} + \frac{n(\bar{X} - \mu)^2}{\sigma^2}$ 이고,

$$\left\{ \begin{array}{l} \sum_{i=1}^n \frac{(X_i - \bar{X})^2}{\sigma^2} \sim \chi^2(n) \\ \frac{n(\bar{X} - \mu)^2}{\sigma^2} \sim \chi^2(1) \end{array} \right. \quad \text{이므로} \quad \left\{ \begin{array}{l} M_{\sum_{i=1}^n \frac{(X_i - \bar{X})^2}{\sigma^2}}(t) = (1 - 2t)^{-\frac{n}{2}} \\ M_{\frac{n(\bar{X} - \mu)^2}{\sigma^2}}(t) = (1 - 2t)^{-\frac{1}{2}} \end{array} \right.$$

$$\begin{aligned} M_{\sum_{i=1}^n \frac{(X_i - \mu)^2}{\sigma^2}}(t) &= \mathbb{E} e^{t \cdot \sum_{i=1}^n \frac{(X_i - \bar{X})^2}{\sigma^2}} \\ &= \mathbb{E} e^{t(\sum_{i=1}^n \frac{(X_i - \bar{X})^2}{\sigma^2} + \frac{n(\bar{X} - \mu)^2}{\sigma^2})} \\ &= M_{\sum_{i=1}^n \frac{(X_i - \bar{X})^2}{\sigma^2}}(t) \cdot M_{\frac{n(\bar{X} - \mu)^2}{\sigma^2}}(t) \quad (\because \text{서로 독립}) \\ M_{\sum_{i=1}^n \frac{(X_i - \bar{X})^2}{\sigma^2}}(t) &= M_{\sum_{i=1}^n \frac{(X_i - \mu)^2}{\sigma^2}}(t) / M_{\frac{n(\bar{X} - \mu)^2}{\sigma^2}}(t) \\ &= (1 - 2t)^{-\frac{n}{2}} / (1 - 2t)^{-\frac{1}{2}} \\ &= (1 - 2t)^{-\frac{n-1}{2}} : \chi^2(n-1) \text{의 mgf} \end{aligned}$$

■

Corollary. $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} N(\mu, \sigma^2) \implies \frac{\bar{X} - \mu}{s/\sqrt{n}} \sim t(n-1)$ where $(n-1)s^2 = \sum_{i=1}^n (X_i - \bar{X})^2$

Proof. $\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1) \perp \frac{(n-1)s^2}{\sigma^2} = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{\sigma^2} \sim \chi^2(n-1)$ 이므로

t분포의 정의에 의해 $\frac{\frac{\bar{X} - \mu}{\sigma/\sqrt{n}}}{\sqrt{\frac{(n-1)s^2}{\sigma^2}/(n-1)}} = \frac{\bar{X} - \mu}{s/\sqrt{n}} \sim t(n-1)$

■

Note. 이 사실을 가지고 μ 의 $100 \times (1 - \alpha)\%$ 신뢰구간을 만들면,

$$\begin{aligned} \mathbb{P}\left(-z_{\frac{\alpha}{2}} < \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} < z_{\frac{\alpha}{2}}\right) &= 1 - \alpha \\ \iff \mathbb{P}\left(\bar{X} - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right) &= 1 - \alpha \end{aligned}$$

σ 대신에 s 를 대입하면,

$$\begin{aligned} \mathbb{P}\left(-t_{\frac{\alpha}{2}} < \frac{\bar{X} - \mu}{s/\sqrt{n}} < t_{\frac{\alpha}{2}}\right) &= 1 - \alpha \\ \iff \mathbb{P}\left(\bar{X} - t_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}} < \mu < \bar{X} + t_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}}\right) &= 1 - \alpha \end{aligned}$$

5장. 표본분포의 근사

Jiho KWAK, Department of Geography, SNU

June 5, 2024

1 확률수렴

1.1 확률수렴의 정의

Definition. (확률수렴과 확률 1로 수렴)

1. (convergence in probability) $n \rightarrow \infty$ 일 때, X_n 은 X 로 확률수렴: $X_n \xrightarrow{P} X$ (혹은 $\text{plim}_{n \rightarrow \infty} X_n = X$ 라고 쓴다.)
 $\iff \forall \epsilon > 0, \mathbb{P}(|X_n - X| > \epsilon) \rightarrow 0, n \rightarrow \infty$
2. (absolute sure convergence, convergence w.p.1) X_n 이 X 로 확률 1로 수렴: $X_n \xrightarrow{as} X$
 $\iff \mathbb{P}(w : \lim_{n \rightarrow \infty} X_n(w) = X(w)) = 1$

Example. $u \sim U(0, 1)$ 라 하자.

(a) $f_n(u) = \frac{1}{n}, n \geq 1, X_n := f_n(u)$ 라 하면 $X \equiv 0$

$$\begin{aligned} \mathbb{P}(|X_n - X| > \epsilon) &= \mathbb{P}(w : |\frac{1}{n} - 0| > \epsilon) \\ &= \mathbb{P}(w : \frac{1}{n} > \epsilon) \\ &= \begin{cases} 0 & n > \frac{1}{\epsilon} \\ 1 & o.w. \end{cases} \xrightarrow{n \rightarrow \infty} 0 \quad \therefore \text{확률수렴} \end{aligned}$$

(b) $u(w) = u, X_n(w) = \frac{1}{n} \rightarrow 0 = X(w), \forall w \quad \therefore X_n \xrightarrow{as} X$

Example.

$$f_1(u) = 1, \forall u$$

$$f_2(u) = \begin{cases} 2 & u \leq \frac{1}{2} \\ 0 & u > \frac{1}{2} \end{cases}$$

$$f_3(u) = \begin{cases} 0 & u \leq \frac{1}{2} \\ 2 & u > \frac{1}{2} \end{cases}$$

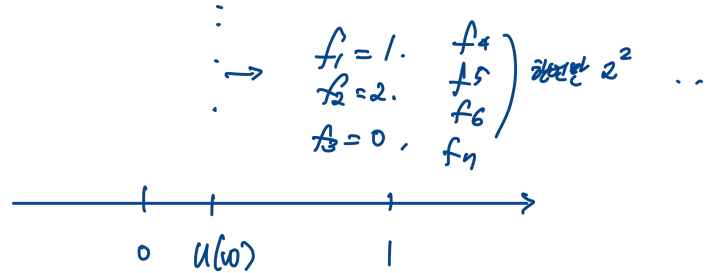
$\vdots$

모든 n 을 $n = \sum_{i=1}^k 2^i + l, l = 0, 1, 2, \dots, 2^{k+1} - 1$ 로 표현할 수 있고, k 와 l 은 유일하게 결정된다.

$$f_n(u) = 2^{k+1} I(l \cdot 2^{k+1} \leq u \leq (l+1)2^{k+1})$$

$f(u) \equiv 0$ 으로 정의하고, $X_n(w) = f_n(u(w))$, $X(w) = f(u(w))$ 라 두면

$X_n(w)$ 는 $X(w)$ 로 확률 1로 수렴하지 않는다.



한편, $\mathbb{P}(|X_n(w) - X(w)| \geq \frac{1}{2^{k+1}}) = \frac{1}{2^{k+1}}$ 이므로 $\mathbb{P}(|X_n(w) - X(w)| > \epsilon) \rightarrow 0$ ($X_n \xrightarrow{P} X$)

Definition. (일치성, consistency) $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} f(x : \theta)$, $\theta \in H$, $T(X_1, \dots, X_n)$: $\eta(\theta)$ 의 추정량

$$T \text{는 } \eta(\theta) \text{에 대해 일치성을 갖는다} \iff T_n \xrightarrow{P} \eta(\theta), \forall \theta$$

Theorem. $X_n \xrightarrow{P} X \implies X_n \xrightarrow{as} X$ (역은 성립하지 않음: 위의 example 참조)

Theorem. (강 대수의 법칙) $X_1, \dots, X_n$: iid, $\mathbb{E}X_i = \mu$ 가 존재 $\implies n \rightarrow \infty$ 일 때 $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \xrightarrow{as} \mu$

Proof.

$\text{Var}(X_i) = \sigma^2 < \infty$ 이라 가정하고 $\bar{X}_n \xrightarrow{P} \mu$ 를 보이자.

Let $\epsilon > 0$ be given, $\mathbb{P}(|\bar{X} - \mu| > \epsilon) \rightarrow 0$ 을 보이고자

$$\mathbb{P}(|\bar{X}_n - \mu| > \epsilon) \leq \frac{1}{\epsilon^2} \mathbb{E}(\bar{X}_n - \mu)^2 = \frac{\sigma^2}{n\epsilon^2} \rightarrow 0$$

■

2 확률수렴과 연산

Properties. (연산)

(a) $X_n \xrightarrow{P} X$, $Y_n \xrightarrow{P} Y$ 일 때 다음이 성립한다.

- (i) $X_n \pm Y_n \xrightarrow{P} X \pm Y$
- (ii) $X_n \cdot Y_n \xrightarrow{P} X \cdot Y$
- (iii) $Y \neq 0$ w.p.1일 때, $X_n/Y_n \xrightarrow{P} X/Y$
- (iv) $\begin{pmatrix} X_n \\ Y_n \end{pmatrix} \xrightarrow{P} \begin{pmatrix} X \\ Y \end{pmatrix}$

(b) $X_n \xrightarrow{P} X$, g 는 연속 $\implies g(X_n) \xrightarrow{P} g(X)$

(c) (a), (b)가 as 에 대해 성립

Proof.

(c) 수열의 성질에 의해: 실수 (x_n, y_n, x, y) 에 대해서 성립하므로 w 가 고정되면 같다.

(a)의 (iv) $\rightarrow$ (b) $\rightarrow$ (i)(ii)(iii) 순서로 보이자.

Let $\epsilon > 0$ be given, $\mathbb{P}\left(\left\|\begin{pmatrix} X_n \\ Y_n \end{pmatrix} - \begin{pmatrix} X \\ Y \end{pmatrix}\right\| > \epsilon\right) \rightarrow 0$ 을 보이하고자

$$\begin{aligned} (\text{좌변}) &= \mathbb{P}(\sqrt{(X_n - X)^2 + (Y_n - Y)^2} > \epsilon) \\ &\leq \mathbb{P}(|X_n - X| > \frac{\epsilon}{\sqrt{2}} \text{ or } |Y_n - Y| > \frac{\epsilon}{\sqrt{2}}) \\ &\leq \mathbb{P}(|X_n - X| > \frac{\epsilon}{\sqrt{2}}) + \mathbb{P}(|Y_n - Y| > \frac{\epsilon}{\sqrt{2}}) \rightarrow 0 \end{aligned}$$

$$*|X_n - X| \leq \frac{\epsilon}{\sqrt{2}} \text{ and } |Y_n - Y| \leq \frac{\epsilon}{\sqrt{2}} \implies (X_n - X)^2 + (Y_n - Y)^2 \leq \epsilon^2$$

(b) $X \equiv a$ (상수), g 는 a 에서 연속이라 가정하고 보이자.

Let $\epsilon > 0$ be given, $\mathbb{P}(|g(X_n) - g(a)| \geq \epsilon) \rightarrow 0$ 을 보이하고자

g 가 a 에서 연속이므로 $|x - a| < \delta \implies |g(x) - g(a)| < \epsilon$ 를 만족하는 $\delta > 0$ 존재.

$$(\text{좌변}) \leq \mathbb{P}(|X_n - a| \geq \delta) \rightarrow 0 \quad (\because X_n \xrightarrow{P} a)$$

(a)의 (i), (ii), (iii)은 연속함수이므로 성립. ■

Example. (1) $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} (\mu, \sigma^2)$, $\mu \in \mathbb{R}$, $\sigma^2 > 0$

$$\frac{1}{n} \sum_{i=1}^n X_i \xrightarrow{P} \mu \quad (\because \text{약 대수의 법칙})$$

$$\frac{1}{n} \sum_{i=1}^n X_i^2 \xrightarrow{P} \mathbb{E}X_i^2 \quad (\because \text{약 대수의 법칙})$$

$$\implies s_1^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 = \frac{1}{n} \left(\sum_{i=1}^n X_i^2 - n\bar{X}^2 \right) = \frac{1}{n} \sum_{i=1}^n X_i^2 - \bar{X}^2 \xrightarrow{P} \mu^2 + \sigma^2 - \mu^2 = \sigma^2$$

$$\implies s^2 = \frac{n}{n-1} \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 \xrightarrow{P} 1 \times \sigma^2 = \sigma^2$$

Example. (2) $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} U(0, \theta)$, $\theta > 0 \implies X_{(n)} \xrightarrow{P} \theta$

Lemma. $\mathbb{E}X_n \rightarrow a$, $\text{Var}(X_n) \rightarrow 0 \implies X_n \xrightarrow{P} a$

$$\text{Proof. } \frac{X_{(n)}}{\theta} \sim \text{Beta}(n, 1) \iff X_{(n)} \sim \theta \cdot \text{Beta}(n, 1)$$

$$\mathbb{E}X_{(n)} = \frac{n}{n+1}\theta \rightarrow \theta, \quad \text{Var}(X_{(n)}) = \left(\frac{n}{n+2} - \left(\frac{n}{n+1} \right)^2 \right) \theta^2 \rightarrow 0$$
■

Example. $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} F$, F 는 연속형, 순증가. $\alpha \in (0, 1)$ 에 대해 $F^{-1}(\alpha)$: 모집단의 $\alpha + n$ 분위수

$$r_n \sim \alpha \cdot n \quad (\text{즉, } \lim_{n \rightarrow \infty} \frac{r_n}{n} = \alpha) \implies X_{(r_n)} \xrightarrow{P} F^{-1}(\alpha)$$

3 분포수렴

3.1 분포수렴의 정의

Definition. X_n 의 분포가 X 로 수렴한다. (covergence in distribution)

$$\iff \lim_{n \rightarrow \infty} \mathbb{P}(X_n \leq x) = \mathbb{P}(X \leq x), \forall x \in C_F, C_F: F \text{가 연속인 점들의 집합, } F \text{는 } X \text{의 cdf.}$$

Note. 분포수렴은 $F_{X_n} \simeq F_X$ 개념을 정의한 것이다.

Notation. $X_n \rightsquigarrow X, F_n \rightsquigarrow F, X_n \xrightarrow{d} X, F_n \xrightarrow{d} F, X_n \Rightarrow X$

Example. (1) $X_n \sim U(0, \frac{1}{n}), X_n \rightsquigarrow X \sim \delta_0 = 0$ 에서 확률이 1인 분포, F_n : X_n 의 cdf

$$F_n(x) = \begin{cases} 0 & x \leq 0 \\ nx & 0 \leq x \leq \frac{1}{n} \\ 1 & x \geq \frac{1}{n} \end{cases}, F(x) = \begin{cases} 0 & x < 0 \\ 1 & x \geq 0 \end{cases}$$

$$\text{If } n \rightarrow \infty, F_n(X) \rightarrow \begin{cases} 0 & x \leq 0 \\ 1 & x > 0 \end{cases} \neq F(x) \text{ (}\cdot \text{ } 0, 1 \text{ 사이의 어떤 } x \text{ 잡더라도 } n \rightarrow \infty \text{이면 1)}$$

$\implies F$ 가 불연속인 점, $x = 0$ 일 때는 수렴하지 않는 것을 알 수 있다.

Example. (2) $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} U(0, \theta), \theta > 0 \implies n(\theta - X_{(n)}) \rightsquigarrow \text{Exp}(\theta)$

Proof.

$$\begin{aligned} \mathbb{P}(n(\theta - X_{(n)}) \leq t) &= \mathbb{P}(X_{(n)} \geq \theta - \frac{t}{n}) = 1 - \mathbb{P}(X_{(n)} < \theta - \frac{t}{n}) \\ &= 1 - \mathbb{P}(X_i < \theta - \frac{t}{n}, \forall i) = 1 - \left(\frac{\theta - \frac{t}{n}}{\theta}\right)^n \\ &= \begin{cases} 1 - (1 - \frac{t}{n\theta})^n & 0 \leq t \leq n\theta \\ 0 & t < 0 \\ 1 & t > n\theta \end{cases} \xrightarrow{n \rightarrow \infty} \begin{cases} 1 - e^{t/\theta} = \text{Exp}(\theta) \text{의 cdf} & 0 \leq t \\ 0 & t < 0 \end{cases} \end{aligned}$$

■

3.2 분포수렴을 증명하는 방법들

Properties.

1. (밀도함수의 수렴, Scheffe's theorem) [연속, 이산형 경우 모두]

$$f_{X_n}(x) \rightarrow f_X(x) \forall x \implies X_n \rightsquigarrow X$$

2. (mgf의 수렴) 어떤 $h > 0$ 에 대해,

$$M_{X_n}(t) \rightarrow M_X(t), \forall t: |t| < h \implies X_n \rightsquigarrow X$$

Proof. f_n 과 f 가 $[a, b]$ 에서 정의되고, $f_n \rightrightarrows f$ (균등수렴)라 하자.

$$|F_n(x) - F(x)| = \left| \int_a^x f_n(x) dt - \int_a^x f(t) dt \right| = \left| \int_a^x f_n(t) - f(t) dt \right|$$

$$\begin{aligned}
&\leq \int_a^x |f_n(t) - f(t)| dt \\
&\leq \int_a^x \sup_s |f_n(s) - f(s)| dt \leq \sup_s |f_n(s) - f(x)| \int_a^b dt
\end{aligned}$$

균등수렴의 정의에 의해 $\sup_s |f_n(s) - f(x)| \xrightarrow{n \rightarrow \infty} 0$

Example. (1) $t(n) \xrightarrow{n \rightarrow \infty} N(0, 1)$ ($\because$ 밀도함수가 수렴)

Example. (2) $n \rightarrow \infty$, $np_n \rightarrow \mu$ 일 때, $Bin(n, p_n) \rightsquigarrow Poi(\mu)$

Proof. (mgf를 이용한 증명)

$X_n \sim Bin(n, p_n)$, $np_n \rightarrow \mu > 0$ 이므로,

$$\begin{aligned}
M_{X_n}(t) &= \mathbb{E}e^{tX_n} = (pe^t + 1 - p)^n = (1 + p(e^t - 1))^n = \left(1 + \frac{np_n}{n}(e^t - 1)\right)^n \\
&\xrightarrow{n \rightarrow \infty} e^{\mu(e^t - 1)}: Poi(\mu) \text{의 mgf}
\end{aligned}$$

Example. (이항분포의 정규근사) $X_n \sim Bin(n, p_n)$, $p_n = p + o(1)$ 일 때, $\frac{X_n - np_n}{\sqrt{np_n(1 - p_n)}} \rightsquigarrow N(0, 1)$

Proof. (sketch)

$Z_n = \frac{X_n - np_n}{\sqrt{np_n(1 - p_n)}}$, $q_n = 1 - p_n$ 이라고 두자.

$$\begin{aligned}
M_{Z_n}(t) &= \mathbb{E}e^{t \cdot \frac{X_n - np_n}{\sqrt{np_n(1 - p_n)}}} = e^{-t\sqrt{\frac{np_n}{q_n}}} \cdot \mathbb{E}e^{t \frac{1}{\sqrt{np_n q_n}} X_n} = e^{-t\sqrt{\frac{np_n}{q_n}}} \cdot \left(p_n e^{\frac{t}{\sqrt{np_n q_n}}} + q_n\right)^n \\
\log M_{Z_n}(t) &= -t\sqrt{\frac{np_n}{q_n}} + n \log(p_n \cdot e^{\frac{t}{\sqrt{np_n q_n}}} + q_n) \\
&= -t\sqrt{\frac{np_n}{q_n}} + n \log\left(p_n \cdot \sum_{r=0}^{\infty} \frac{1}{r!} \left(\frac{t}{\sqrt{np_n q_n}}\right)^r + q_n\right) \quad (\because \text{테일러 전개}) \\
&= n \log\left(1 + p_n \sum_{r=1}^{\infty} \frac{1}{r!} \left(\frac{t}{\sqrt{np_n q_n}}\right)^r\right) - t\sqrt{\frac{np_n}{q_n}} \\
&= n\left(A - \frac{1}{2}A^2 + \frac{1}{3}A^3 - \dots\right) - t\sqrt{\frac{np_n}{q_n}} \quad \dots (*)
\end{aligned}$$

$$\begin{aligned}
A &= p_n \sum_{r=1}^{\infty} \frac{1}{r!} \left(\frac{t}{\sqrt{np_n q_n}}\right)^r \\
&= p_n \left(\frac{1}{\sqrt{np_n q_n}} t + \frac{1}{2} \frac{1}{np_n q_n} t^2 + \frac{1}{6} \frac{1}{(np_n q_n)^{3/2}} t^3 + \dots\right) \\
&= \sqrt{\frac{p_n}{nq_n}} + \frac{1}{2} \frac{1}{nq_n} t^2 + \frac{1}{6} \frac{p_n}{(np_n q_n)^{3/2}} + \dots
\end{aligned}$$

$$\therefore (*) = n\left(\sqrt{\frac{p_n}{nq_n}} t + \frac{1}{2} \frac{1}{nq_n} t^2 - \frac{1}{2} \left(\frac{p_n}{nq_n} t^2\right)\right) - t\sqrt{\frac{np_n}{q_n}} + o(1) = \frac{1}{2} t^+ o(1)$$

4 중심극한정리

Theorem. $X_1, \dots, X_n$: iid

$$\mathbb{E}X_1 = \mu, \text{Var}(X_1) = \sigma^2 > 0 \implies \frac{\sqrt{n}(\bar{X} - \mu)}{\sigma} \rightsquigarrow N(0, 1)$$

Note. 이는 $\sqrt{n}(\bar{X} - \mu) \rightsquigarrow N(0, \sigma^2)$ 를 보이는 것과 동치이다.

Proof. $M_X(t)$ 가 존재한다 가정하고 증명한다.

$$\begin{aligned} M_{\frac{\sqrt{n}(\bar{X} - \mu)}{\sigma}}(t) &= \mathbb{E}e^{t \frac{\sqrt{n}(\bar{X} - \mu)}{\sigma}} = \mathbb{E}e^{t \frac{\sqrt{n}}{n} \left(\sum_{i=1}^n \frac{X_i - \mu}{\sigma} \right)} = \mathbb{E}e^{\frac{t}{\sqrt{n}} \sum_{i=1}^n \frac{X_i - \mu}{\sigma}} \\ &= m\left(\frac{t}{\sqrt{n}}\right)^n, \quad m(t) = \mathbb{E}e^{t \frac{X_1 - \mu}{\sigma}} \\ m(t) &= 1 + m'(0) \cdot t + \frac{m''(0)}{2} t^2 + o(t^2) \\ &= 1 + 0 \cdot t + 1 \cdot t^2 + o(t^2) \\ &= 1 + \frac{1}{2} t^2 + o(t^2) \\ \log M_{\frac{\sqrt{n}(\bar{X} - \mu)}{\sigma}}(t) &= n \log m\left(\frac{t}{\sqrt{n}}\right) = n \log\left(1 + \frac{t^2}{2n} + o\left(\frac{1}{n}\right)\right) \quad (\because o\left(\frac{t^2}{n}\right) = o\left(\frac{1}{n}\right)) \\ &= n\left(\frac{t^2}{2n} + o\left(\frac{1}{n}\right) + o\left(\frac{1}{n}\right)\right) \quad (\because \log(1+x) = x + o(x)) \\ &= \frac{1}{2} t^2 + o(1) \xrightarrow{n \rightarrow \infty} \frac{1}{2} t^2 \end{aligned}$$

■

Theorem. (다변량 CLT) $X_1, \dots, X_n$: iid인 k -차원 확률변수

$$\mathbb{E}X_1 = \mu \in \mathbb{R}^k, \text{Var}(X_1) = \Sigma \in \mathbb{R}^{k \times k} \implies \sqrt{n}(\bar{X} - \mu) \rightsquigarrow N(0, \Sigma)$$

Example.

1. (이항분포) $X_n \sim \text{Bin}(n, p) \iff X_n \stackrel{d}{=} Z_1 + \dots + Z_n, Z_i \stackrel{\text{iid}}{\sim} \text{Ber}(p)$

$$\begin{aligned} \sqrt{n}(\bar{Z} - \mathbb{E}Z_1) &\rightsquigarrow N(0, \text{Var}(Z_1)) \\ \sqrt{n}\left(\frac{X_n}{n} - p\right) &\rightsquigarrow N(0, p(1-p)) \end{aligned}$$

2. (포아송 분포) $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} \text{Poi}(\lambda)$

$$\begin{aligned} \sqrt{n}(\bar{X} - \mathbb{E}X_1) &\rightsquigarrow N(0, \text{Var}(X_1)) \\ \sqrt{n}(\bar{X} - \lambda) &\rightsquigarrow N(0, \lambda) \end{aligned}$$

3. (카이제곱 분포) $X_n \sim \chi^2(n) \iff X_n \stackrel{d}{=} Z_1 + \dots + Z_n, Z_i \stackrel{\text{iid}}{\sim} \chi^2(1) = Ga(\frac{1}{2}, 2)$

$$\begin{aligned}\sqrt{n}(\bar{Z} - \mathbb{E}Z_1) &\rightsquigarrow N(0, \text{Var}(Z_1)) \\ \sqrt{n}\left(\frac{X_n}{n} - 1\right) &= \frac{X_n - n}{\sqrt{n}} \rightsquigarrow N(0, 2)\end{aligned}$$

4. (다항분포) $X_n \sim \text{Multi}(n, p), X_n = (X_1, \dots, X_{k-1}), 0 \leq \sum_{i=1}^{k-1} X_i \leq n$

$$p = (p_1, \dots, p_{k-1}), p_i > 0, \sum_{i=1}^{k-1} p_i < 1, X_n \stackrel{d}{=} Z_1 + \dots + Z_n, Z_i \stackrel{\text{iid}}{\sim} \text{Multi}(1, p)$$

$$\sqrt{n}(\bar{Z} - \mathbb{E}Z_1) \rightsquigarrow N(0, \text{Var}(Z_1))$$

$$\sqrt{n}\left(\frac{X_n}{n} - p\right) = \sqrt{n}\left(\frac{X_n - np}{n}\right) = \frac{X_n - np}{\sqrt{n}} \rightsquigarrow N(0, \text{diag}(p) - pp^T) * \text{diag}(p) = \begin{pmatrix} p_1 & & 0 \\ & \ddots & \\ 0 & & p_{k-1} \end{pmatrix}$$

5 극한분포의 연산

5.1 확률수렴과 분포수렴의 관계

Theorem. (확률수렴과 분포수렴의 관계)

$$(a) X_n \xrightarrow{P} X \implies X_n \rightsquigarrow X$$

$$(b) X_n \xrightarrow{P} c \iff X_n \rightsquigarrow c, c: \text{상수}$$

$$\text{Note. } \xrightarrow{as} \implies \xrightarrow{P} \implies \rightsquigarrow$$

Proof.

(a) $X_n \sim F_n, X \sim F, x \in C_F = F$ 의 연속점들의 모임.

$\lim_{n \rightarrow \infty} F_n(x) = F(x)$ 를 보이하고자

Given $\epsilon > 0, F(x - \epsilon) \leq \liminf_{n \rightarrow \infty} F_n(x) \leq \limsup_{n \rightarrow \infty} F_n(x) \leq F(x + \epsilon)$ 을 보이면 된다.

$$\begin{aligned}F_n(x) &= \mathbb{P}(X_n \leq x) = \mathbb{P}(X_n \leq x, |X_n - X| \leq \epsilon) + \mathbb{P}(X_n \leq x, |X_n - X| > \epsilon) \\ &\leq \mathbb{P}(X \leq x + \epsilon) + \mathbb{P}(|X_n - X| > \epsilon) \\ \limsup_{n \rightarrow \infty} F_n(x) &\leq F(x + \epsilon) + \limsup_{n \rightarrow \infty} \mathbb{P}(|X_n - X| > \epsilon) \\ &= F(x + \epsilon)\end{aligned}$$

$$\begin{aligned}\mathbb{P}(X_n > x) &= \mathbb{P}(X_n > x, |X_n - X| \leq \epsilon) + \mathbb{P}(X_n > x, |X_n - X| > \epsilon) \\ &\leq \mathbb{P}(X > x - \epsilon) + \mathbb{P}(|X_n - X| > \epsilon) \\ \limsup_{n \rightarrow \infty} \mathbb{P}(X_n > x) &\leq \mathbb{P}(X > x - \epsilon) + 0 \\ \limsup_{n \rightarrow \infty} (1 - \mathbb{P}(X_n \leq x)) &\leq 1 - \mathbb{P}(X \leq x - \epsilon) \\ \mathbb{P}(X \leq x - \epsilon) &\leq \liminf_{n \rightarrow \infty} \mathbb{P}(X_n \leq x)\end{aligned}$$

(b) $(\Leftarrow) X_n \rightsquigarrow c \implies X_n \xrightarrow{P} c$ 를 보이하고자

Given $\epsilon > 0$, $\mathbb{P}(|X_n - c| > \epsilon) \rightarrow 0 \iff \mathbb{P}(|X_n - c| \leq \epsilon) \rightarrow 1$ 이므로

$$\begin{aligned} \mathbb{P}(|X_n - c| \leq \epsilon) &= \mathbb{P}(X_n - c \leq \epsilon) - \mathbb{P}(X_n - c < -\epsilon) \\ &= \mathbb{P}(X_n \leq c + \epsilon) - \mathbb{P}(X_n < c - \epsilon) \\ &\geq \mathbb{P}(X_n \leq c + \epsilon) - \mathbb{P}(X_n \leq c - \frac{\epsilon}{2}) \\ &\geq \mathbb{P}(X_n \leq c + \epsilon') - \mathbb{P}(X_n \leq c - \frac{\epsilon'}{2}) \\ &\xrightarrow{n \rightarrow \infty} \mathbb{P}(X \leq c + \epsilon') - \mathbb{P}(X \leq c - \frac{\epsilon'}{2}) = 1 \end{aligned}$$

Note. $D_F : F$ 의 불연속점들의 모임은 countable이므로 $0 < \epsilon' \leq \epsilon$, $c + \epsilon'$ 와 $c - \frac{\epsilon'}{2}$ 가 F 의 연속점이 되도록 ϵ' 를 잡을 수 있다. ■

5.2 슬러츠키 정리

Theorem. (슬러츠키 정리) $X_n \rightsquigarrow X$, $Y_n \xrightarrow{P} c$ 이면 다음이 성립한다.

(a) $X_n \pm Y_n \rightsquigarrow X \pm c$

(b) $X_n \cdot Y_n \rightsquigarrow cX$

(c) $X_n/Y_n \rightsquigarrow X/c$, $c \neq 0$

Theorem. (연속성 정리, 사상정리) $X_n \rightsquigarrow X$, g : 연속 $\implies g(X_n) \rightsquigarrow g(X)$

Proof. 슬러츠키 정리 (a)의 증명: $y \in C_{F_{X+c}}$, $\mathbb{P}(X_n + Y_n \leq y) \rightarrow F_{X+c}(y)$ 를 보이하고자

Given $\epsilon > 0$, $F_{X+c}(y - \epsilon) \leq \liminf \mathbb{P}(X_n + Y_n \leq y) \leq \limsup \mathbb{P}(X_n + Y_n \leq y) \leq F_{X+c}(y + \epsilon)$ 를 보이자.

$$\begin{aligned} \mathbb{P}(X_n + Y_n \leq y) &\leq \mathbb{P}(X_n + Y_n \leq y, |Y_n - c| \leq \epsilon) + \mathbb{P}(X_n + Y_n \leq y, |Y_n - c| > \epsilon) \\ &\leq \mathbb{P}(X_n \leq y - c + \epsilon) + \mathbb{P}(|Y_n - c| > \epsilon) \\ \limsup \mathbb{P}(X_n + Y_n \leq y) &\leq \mathbb{P}(X \leq y - c + \epsilon) = F_{X+c}(y + \epsilon) \end{aligned}$$

■

Example. $t(n) \rightsquigarrow N(0, 1)$: t분포의 성질에 의해 $n \rightarrow \infty$ 이면 t분포의 pdf $\rightarrow N(0, 1)$ 의 pdf 임을 알고 있다.

$$T = \frac{Z}{\sqrt{V_n/n}}, \quad \left. \begin{array}{l} Z \sim N(0, 1) \\ V_n \sim \chi^2(n) \end{array} \right\} \text{ 독립, } V_n \stackrel{d}{=} Z_1 + \dots + Z_n, \quad Z_i \stackrel{\text{iid}}{\sim} \chi^2(1)$$

$$\frac{V_n}{n} \xrightarrow{P} 1 \quad (\because \text{대수의 법칙}), \text{ 확률수렴은 연속함수를 보존하므로 } \sqrt{\frac{V_n}{n}} \xrightarrow{P} 1$$

슬러츠키 정리에 의해 $\frac{Z}{\sqrt{V_n/n}} \rightsquigarrow N(0, 1)$

5.3 큰 O_p 와 작은 o_p

Definition. (X_n) : 확률변수의 열, (a_n) : 실수열

(a) $n \rightarrow \infty$ 일 때,

$$\begin{aligned} X_n = O_p(a_n) &\iff \sup_n \mathbb{P}\left(\left|\frac{X_n}{a_n}\right| > k\right) \xrightarrow{k \rightarrow \infty} 0 \\ &\iff \frac{X_n}{a_n} = O_p(1) \end{aligned}$$

(b) $n \rightarrow \infty$ 일 때,

$$\begin{aligned} X_n = o_p(a_n) &\iff \frac{X_n}{a_n} \xrightarrow{P} 0 \\ &\iff \frac{X_n}{a_n} = o_p(1) \end{aligned}$$

Note. $a_n = O(1) \iff (a_n \circledast \text{유계}) \sup_n |a_n| < \infty$, $a_n = o(1) \iff a_n \xrightarrow{n \rightarrow \infty} 0$

Properties. $X_n = O_p(1)$

$$\iff \forall \epsilon > 0, \exists B_\epsilon, N_\epsilon > 0 \text{ s.t. } n \geq N_\epsilon \Rightarrow \mathbb{P}(|X_n| > B_\epsilon) < \epsilon$$

Proof. check by yourself. ■

Example. $X_n = Z$, $\forall n$ (모든 X_n 이 n 에 대해 똑같은 확률변수)인 경우 $\sup_n \mathbb{P}(|X_n| > k) = \mathbb{P}(|Z| > k) \xrightarrow{k \rightarrow \infty} 0$

Properties.

$$\begin{aligned} \text{(a)} \quad X_n &\xrightarrow{P} X \implies X_n \rightsquigarrow n \implies X_n = O_p(1) \\ X_n &= o_p(1) \implies X_n = O_p(1) \end{aligned}$$

$$\begin{aligned} \text{(b)} \quad O_p(1) + O_p(1) &= O_p(1) & O_p(1) \cdot O_p(1) &= O_p(1) \\ O_p(1) + o_p(1) &= O_p(1) & O_p(1) \cdot o_p(1) &= o_p(1) \\ o_p(1) + o_p(1) &= o_p(1) & o_p(1) \cdot o_p(1) &= o_p(1) \end{aligned}$$

Proof.

(a) $X_n \rightsquigarrow X \implies X_n = O_p(1)$ 을 보이하고자

$$\mathbb{P}(|X_n - X| > k) \leq |\mathbb{P}(|X_n| > k) - \mathbb{P}(|X| > k)| + \mathbb{P}(|X| > k)$$

Given $\epsilon > 0$, $\mathbb{P}(|X| > k) < \frac{\epsilon}{2}$ 인 k 를 잡을 수 있고,

$|\mathbb{P}(|X_n| \leq k) - \mathbb{P}(|X| \leq k)| \leq \frac{\epsilon}{2}$, $\forall n \geq N_\epsilon$ 인 N_ϵ 를 잡을 수 있다. ($\cdot \cdot |X_n|$ 도 $|X|$ 로 분포수렴)

$$\begin{aligned} \sup_{n \geq N_\epsilon} \mathbb{P}(|X_n| > k) &\leq \sup_{n \geq N_\epsilon} |\mathbb{P}(|X_n| \leq k) - \mathbb{P}(|X| \leq k)| + \mathbb{P}(|X| > k) \\ &\leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon \end{aligned}$$

(b) $X_n = O_p(1)$, $Y_n = O_p(1)$ 이면 $X_n + Y_n = O_p(1)$ 을 보이하고자

$$\begin{aligned} \sup_n \mathbb{P}(|X_n + Y_n| > k) &\leq \sup_n (\mathbb{P}(|X_n| > \frac{\epsilon}{2} \text{ or } |Y_n| > \frac{\epsilon}{2})) \\ &\leq \sup_n \mathbb{P}(|X_n| > \frac{k}{2}) + \sup_n \mathbb{P}(|Y_n| > \frac{k}{2}) \cdots (*) \end{aligned}$$

이때, $\frac{k}{2}$ 를 충분히 크게 잡으면 $(*) < \epsilon$

■

Example. (표본분산의 극한분포) $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} F$, $\mathbb{E}X_1^4 < \infty$. $\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$ 의 극한분포?

sol) $\mathbb{E}X_1 = 0$ 이라 가정해도 무방하다.

$$\sqrt{n}(\bar{X} - 0) \rightsquigarrow N(0, \sigma^2)$$

$$V_n \rightsquigarrow N(0, \sigma^2), \quad \bar{X} = \frac{1}{n} V_n$$

$$\sqrt{n}(\frac{1}{n} \sum_{i=1}^n X_i^2 - \sigma^2) \rightsquigarrow N(0, \text{Var}(X_1^2))$$

$$W_n \rightsquigarrow N(0, \text{Var}(X_1^2)), \quad \frac{1}{n} \sum_{i=1}^n X_i^2 = \sigma^2 + \frac{1}{\sqrt{n}} W_n$$

$$\begin{aligned} \hat{\sigma}^2 &= \frac{1}{n} \sum_{i=1}^n X_i^2 - \bar{X}^2 \\ &= \sigma^2 + \frac{1}{\sqrt{n}} W_n - (\frac{1}{\sqrt{n}} V_n)^2 \end{aligned}$$

$$\begin{aligned} \sqrt{n}(\hat{\sigma}^2 - \sigma^2) &= W_n - \frac{1}{\sqrt{n}} V_n^2 = W_n - \frac{1}{\sqrt{n}} O_p(1) \\ &= W_n - o_p(1) \quad (\because \frac{1}{\sqrt{n}} = o_p(1)) \\ &\overset{\text{슬러츠키}}{\rightsquigarrow} N(0, \text{Var}(X_1^2)) \end{aligned}$$

5.4 델타 방법

Theorem. g 는 미분가능. $\sqrt{n}(X_n - \theta) \rightsquigarrow Z \implies \sqrt{n}(g(X_n) - g(\theta)) \rightsquigarrow g'(\theta)Z$

Proof. g 는 θ 에서 미분가능하므로

$$\frac{g(x) - g(\theta)}{x - \theta} \xrightarrow{x \rightarrow \theta} g'(\theta)$$

$$\iff \frac{g(x) - g(\theta)}{x - \theta} = g'(\theta) + r(x, \theta), \quad r(x, \theta) = o(1)$$

$$\iff g(x) = g(\theta) + g'(\theta)(x - \theta) + r(x, \theta)(x - \theta)$$

$$\begin{aligned} \sqrt{n}(g(X_n) - g(\theta)) &= \sqrt{n}(g'(\theta) + r(X_n, \theta))(X_n - \theta) \\ &= g'(\theta)\sqrt{n}(X_n - \theta) + o_p(1) \rightsquigarrow g'(\theta)Z \quad (\because r(X_n, \theta) = o_p(1)) \end{aligned}$$

■

Example. 앞서 $\sqrt{n}(\hat{\sigma}^2 - \sigma^2) \rightsquigarrow N(0, \text{Var}(X_1^2))$

$$\begin{aligned}\sqrt{n}(\sqrt{\hat{\sigma}^2} - \sqrt{\sigma^2}) &\rightsquigarrow g'(\theta)N(0, \text{Var}(X_1^2)) \\ &= \frac{1}{2\sqrt{\sigma^2}}N(0, \text{Var}(X_1^2)) = N(0, \frac{1}{4\sigma^2}\text{Var}(X_1^2))\end{aligned}$$

$$\therefore \sqrt{n}(\hat{\sigma}^2 - \sigma^2) \rightsquigarrow N(0, \frac{1}{4\sigma^2}\text{Var}(X_1^2)) = \frac{1}{4\sigma^2}(\mathbb{E}X_1^2 - (\mathbb{E}X_1)^2)$$