분포수렴

정의

- X_n을 X의 확률변수라 하고, F_{xn}과 F_x을 각각의 cdf라고 할 때

1. $\lim_{n\to\infty} F_{xn}(x) = F_x$(X)면 X_n은 X에 분포수렴 한다고 말한다.

정의

- 중요한 성질들

1. X_n이 X로 확률수렴하면 X_n은 X_n로 분포수렴한다.

1) x가 F_x(X)의 연속인 범위라고 하면

(1) F_{xn}(X) = P[$X_n \leq$ x] = P[{$X_n \leq$ x}∩{$X_n -$ X} < ϵ] + P[{$X_n \leq$ x}∩{$X_n -$ X} ≥ ϵ]
= P[$X_n \leq$ x+ ϵ] + P[|$X_n -$ X| ≥ ϵ]

(2) 부등식과, $X_n \xrightarrow{p} X$ 에 의하여

- P[$X_n \leq$ x+ ϵ] → P[X ≤ x+ ϵ]
- P[|$X_n -$ X|≥ ϵ] → P[|X − X|≥ ϵ] = 0

(3) 따라서 상한 $\overline{\lim_{n\to\infty}} F_{xn}(x) \leq F_x$[x+ ϵ]

정의

- 중요한 성질들

(3) 그 반대방향의 여집합에 대해서도 동일한 방식으로 구하면

- $\mathrm{P}[X_n > \mathrm{x}] = \mathrm{P}[\{X_n \geq -\mathrm{x}\} \cap \{X_n - \mathrm{X}\} < \epsilon] + \mathrm{P}[\{X_n \geq -\mathrm{x}\} \cap \{X_n - \mathrm{X}\} \geq \epsilon]$
  $= \mathrm{P}[X_n \geq \mathrm{x} - \epsilon] + \mathrm{P}[|X_n - \mathrm{X}| \geq \epsilon]$
- 따라서 하한 $\underline{\lim}_{n\to\infty} F_{xn}(x) \geq F_x[\mathrm{x} - \epsilon]$

2) 위 결과물들을 다시 정리하면

(1) $F_x[\mathrm{x} - \epsilon] \leq \underline{\lim}_{n\to\infty} F_{xn}(x) \leq \overline{\lim}_{n\to\infty} F_{xn}(x) \leq F_x[\mathrm{x} + \epsilon]$

(2) $\epsilon \to 0$ 이면, 샌드위치 정리에 따라 $\lim_{n\to\infty} F_{xn}(x) = F_x[\mathrm{x}]$ 이다.

정의

- 중요한 성질들

2. $X_n \xrightarrow{D} b$ 이면 $X_n \xrightarrow{p} b$ 이다.(단, b는 어떤 상수)

1) $\lim\limits_{n\to\infty} [|X_n - \mathrm{b}| \leq \ \epsilon]$ = $\lim\limits_{n\to\infty} F_{xn}[\mathrm{b} + \epsilon]$ - $\lim\limits_{n\to\infty} F_{xn}[(\mathrm{b} - \epsilon) - 0] = 1\ - 0 = 1$

2) 따라서 $X_n \xrightarrow{p} b$ 이다.

3. $X_n \xrightarrow{D}$X 이고 $Y_n \xrightarrow{p} 0$ 라면 X_n+ Y_n은 X로 확률수렴한다.

4. $X_n \xrightarrow{D}$X 이고 g가 X의 범위에서 연속인 함수라고 하자. $g(X_n) \xrightarrow{D}$g(X) 이다.

정의

- 중요한 성질들

5. (슬러츠키 정리) X_n, X , A_n , B_n이 확률변수이며, a와 b가 상수라고 하자.

1) $X_n \xrightarrow{D}$X 에 대하여 $A_n \xrightarrow{p} a$, $B_n \xrightarrow{p} b$ 이면

$$A_n + B_n X_n \rightarrow a + b\text{X}$$

정의

- 중요한 성질들

5. (슬러츠키 정리) X_n, X , A_n , B_n이 확률변수이며, a와 b가 상수라고 하자.

1) $X_n \xrightarrow{D}$X 에 대하여 $A_n \xrightarrow{p} a$, $B_n \xrightarrow{p} b$ 이면

$$A_n + B_n X_n \rightarrow a + b\mathrm{X}$$

정의

- 확률 유계

1. CDF F_x 를 갖는 X가 존재할 때, ϵ>0가 주어졌을 경우

1) $x \leq \eta_1$에 대해 $F_x(\mathrm{x}) \leq \frac{\epsilon}{2}$, $x \geq \eta_2$에 대해 $F_x(\mathrm{x})$> $1-\frac{\epsilon}{2}$ 를 만족하는 $[\eta_1, \eta_2]$를 구할 수 있다.

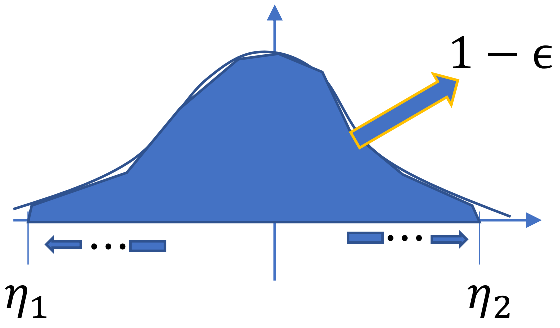
2) η = max[$|\eta_1|$, $|\eta_2|$] 라고 하자. 그러면

(1) P($|X| \leq \eta$) = $F_x(\eta)$ -$F_x(-\eta-0) \geq 1-\frac{\epsilon}{2}-\frac{\epsilon}{2}$ = 1- ϵ

2. 위를 일반화 하면

1) $\mathrm{n} \geq N_\epsilon \Rightarrow$ P($|X_n| \leq B_\epsilon$) $\geq 1-\epsilon$ 인 상수 B_ϵ와 N_ϵ이 존재하면 확률변수의 열 $\{X_n\}$은 확률유계이다.

3. 2. $\{X_n\}$이 확률변수열이고, X가 확률변수라고 하자. $X_n \rightarrow X$이면 $\{X_n\}$은 확률 유계이다.



η_1 와 η_2가 한없이 끝에 도달하기 직전 멈췄을 때,
확률은 1- ϵ 가 된다.

정의

- 점근 분포를 구하는 방법

1. $\Delta -$ 방법

1) $\{Y_n\}$을 확률유계인 확률변수라고 정의하고,
   $\mathrm{O}(x)$는 $x \to 0$일 때, $\frac{a}{x} \to 0$일때 한해 a=$\mathrm{o}(x)$인 함수이다.

(1) 또는 $\mathrm{o}_p\ (x_n)$은 $\mathrm{n} \to \infty$ 일때 $\frac{Y_n}{X_n} \to 0$ 일때 한해 $Y_n\ =\ \mathrm{o}_p\ (x_n)$ 으로 정의한다.

(2) 또한 $\mathrm{O}_p\ (x_n)$은 $\mathrm{n} \to \infty$ 일 때 $\frac{Y_n}{X_n}$ 이 확률 유계일 때 한해 $Y_n\ =\ \mathrm{O}_p\ (x_n)$

정의

- 점근 분포를 구하는 방법

1. $\Delta -$ 방법

2) $\{Y_n\}$을 확률유계인 확률변수열 이라고 할 때, $X_n = o_p(Y_n)$이면 $n \to \infty$ 일때 $X_n \xrightarrow{p} 0$ 이다.

(1) ϵ>0가 주어졌을 경우 $\{Y_n\}$은 확률 유계이므로

- $n \geq N_\epsilon \Rightarrow P(|Y_n| \leq B_\epsilon) \geq 1 - \epsilon$ 인 상수 N_ϵ와 B_ϵ가 존재한다.

(2) $X_n = o_p(Y_n)$ 이므로 $n \to \infty$ 일 때 $\frac{Y_n}{X_n} \to 0$

(3) 따라서

- $P[|X_n| > \epsilon] = P[|X_n| \geq \epsilon, |Y_n| \leq B_\epsilon] + P[|X_n| \geq \epsilon, |Y_n| > B_\epsilon]$

  $= P[\frac{X_n}{|Y_n|} \geq \frac{\epsilon}{B_\epsilon}] + P[|Y_n| \geq B_\epsilon]$

- $n \to \infty$ 이면 $P[\frac{X_n}{|Y_n|} \geq \frac{\epsilon}{B_\epsilon}] + P[|Y_n| \geq B_\epsilon] \to 0$ 이므로 $P[|X_n| > \epsilon] \to 0$ 이다.

정의

- 점근 분포를 구하는 방법

1. $\Delta -$ 방법

3) $\{X_n\}$을 $\sqrt{n}(X_n-\theta)\xrightarrow{D}N(0,\sigma^2)$ 인 확률변수의 열이라고 하자. g(x)가 θ에서 미분 가능하다면

$$\sqrt{n}(g(X_n)-g(\theta)) \xrightarrow{D} N(0,\sigma^2(g'(\theta)^2))$$ 이다.

(1) g(X_n) = $g(\theta)$+ $g'(\theta)(X_n-\theta)+\mathrm{o}_p\,(|X_n-\theta|)$ 로 나타낼 때

(2) 위 식을 고치면 $\sqrt{n}(g(X_n)-g(\theta))$ = $g'(\theta)\sqrt{n}(X_n-\theta)+\mathrm{o}_p(\sqrt{n}|X_n-\theta|)$

- 이 때, $\sqrt{n}(X_n-\theta)\xrightarrow{D}N(0,\sigma^2)$ 이므로 $\mathrm{o}_p(\sqrt{n}|X_n-\theta|)\rightarrow 0$이다.
- $\sqrt{n}(g(X_n)-g(\theta))$ = $g'(\theta)\sqrt{n}(X_n-\theta)\rightarrow g'(\theta)N(0,\sigma^2)$= $N(0,(g'(\theta)^2)\,\sigma^2)$ 이다.

정의

- 점근 분포를 구하는 방법

2. 적률생성함수 기법

1) $\{X_n\}$을 모든 n에 대하여 mgf $M_{xn}(t)$를 갖는 확률변수열이라고 할때,
X는 mgf $M_x(t)$를 갖는 확률변수이다.

2) 이 때, $\lim_{n\to\infty} M_{xn}(t) = M_x(t)$ 이면 $X_n \xrightarrow{D} X$ 이다.

예제

- 적률생성함수 기법

1. Y_n~b(n,p) 라고 할 때, 모든 n에 대하여 μ=np로 같다. 즉, p=$\frac{\mu}{n}$ 이다.

2. mgf $M_y(t)$=$[(1-P)+pe^t]^n$= $[1-p(e^t-1)]^n$ 에서 $\left[1-\frac{\mu(e^t-1)}{n}\right]^n$

(1) 이 때, $\lim_{n\to\infty}\left[1+\frac{b}{n}+\frac{\psi(n)}{n}\right]^{cn}$ 에서 $\lim_{n\to\infty}\frac{\psi(n)}{n}$ = 0 이면 $\lim_{n\to\infty}\left[1+\frac{b}{n}\right]^{cn}=e^{bc}$인 정리를 이용

(2) $\lim_{n\to\infty}\left[1-\frac{\mu(e^t-1)}{n}\right]^n$에서 $b=\mu(e^t-1)$이므로,

$$\lim_{n\to\infty}\left[1-\frac{\mu(e^t-1)}{n}\right]^n=e^{\mu(e^t-1)}$$

(3) 이는 푸아송 분포의 mgf와 같다.