분포의 지수류

정의

- 개요

1. 확률밀도함수의 족 $\{f(x;\theta) : \theta \in \Omega\}$ 에서

2. $f(x;\theta)$ = exp[p(θ)k(x) + H(x) + q(θ)] 꼴로 표현되면,
이 함수를 갖는 분포의 모수를 추정하는 통계량은 완비충분통계량이다.

- 정칙지수류

1. $f(x;\theta)$ = exp[p(θ)k(x) + H(x) + q(θ)]에서

1) 이는 분해하면

(1) θ에 의존하는 p항과 q항

(2) θ에 의존하지 않는 k항과 H항으로 이루어져 있다.

정의

- 정칙지수류

1. $f(\mathrm{x};\theta)$ = exp[p(θ)k(x) + H(x) + q(θ)]에서

2) 이 때, 다음의 조건을 만족해야한다.

(1) X의 범위 S는 θ에 대해 무관하다. 즉 범위는 θ에 의존하지 않는다.

(2) $\theta \in \Omega$에 대하여, p(θ)는 연속함수이다.

(3) 각각의 분포 유형에 따라

\- X가 연속형 확률변수일 때, $\frac{\partial k}{\partial x} \neq \vec{0}$인 영역에서 H(x)는 연속 함수이다.

\- X가 이산형 확률변수일 때, K(x)는 $\mathrm{X} \in$ S인 함수이다.

2\. 위와 같은 조건을 만족하는 분포를 <u>정칙지수류</u> 라고 한다.

정의

- 정칙지수류

3. 정칙지수류는 충분통계량을 갖는다.

1) $\prod_{i=1}^{n} f(x_i;\ \theta)$ = exp[p(θ)$\sum k(x_i)$ + H$\sum(x_i)$ + nq(θ)] 로 나타낼 수 있고

2) 이 때, exp[p(θ)$\sum k(x_i)$ + nq(θ)]·exp[H$\sum(x_i)$] 로 나타낼 수 있으므로

3) 이 때, $Y_1 = \sum k(x_i)$는 충분통계량이다.

정의

- 정칙지수류

4. 정칙지수류는 완비충분통계량을 갖는다.

1) $Y_1 = \sum k(x_i)$ 인 통계량을 정의하자. 이 때

(1) E[u(Y_1)] = 0 를 정의하고, 이를 다시 쓰면

(2) $\int$ u(y_1)R(y_1)exp[p(θ)y_1 + nq(θ)] $dy_1 = 0$

- 이 때, R(y_1)은 y_1과 관련된 함수이며 $R(y_1)$ 은 θ와 무관한 함수이다.

2) 라플라스 변환을 실시하면

(1) u(y_1)R(y_1) = $\vec{0}$, 이 때 R(y_1)은 pdf(y_1)의 인수이기 때문에 모든 $y_1 \in S$에 대해 0이 아니다.

(2) 따라서, u(y_1) = 0이고, 이는 완비성의 정의이기 때문에 증명이 완료된다.

정의

- 완비충분통계량 $Y_1 = \sum k(x_i)$

1. $Y_1 = \sum k(x_i)$ 인 통계량을 정의하자. Y_1은 다음과 같은 성질을 갖는다.

2) $\mathrm{E}(Y_1) = -\mathrm{n}\frac{\mathrm{q}'(\theta)}{\mathrm{p}'(\theta)}$

(1) 정칙지수류의 한 분포의 CDF를 정의하면

- $\int_a^b \exp[\mathrm{p}(\theta)k(x)+\mathrm{h}(x)+\mathrm{q}(\theta)]dx$
  $= \exp[\mathrm{q}(\theta)] \int_a^b \exp[\mathrm{p}(\theta)k(x)+\mathrm{h}(x)]dx$
  $= 1$

(2) 위 cdf에서 양변을 θ로 미분하면

- $\exp[\mathrm{q}(\theta)]\,\mathrm{p}'(\theta) \int_a^b k(x)\exp[\mathrm{p}(\theta)k(x)+\mathrm{h}(x)]dx + \mathrm{q}'(\theta)\exp[\mathrm{q}(\theta)] \int_a^b \exp[\mathrm{p}(\theta)k(x)+\mathrm{h}(x)]dx = 0$

(3) 식을 이항하여 정리하면

- $\frac{\mathrm{q}'(\theta)\cancel{\exp[\mathrm{q}(\theta)]}}{\mathrm{p}'(\theta)\cancel{\exp[\mathrm{q}(\theta)]}} = \frac{\mathrm{q}'(\theta)}{\mathrm{p}'(\theta)} = \frac{\int_a^b k(x)\exp[\mathrm{p}(\theta)k(x)+\mathrm{h}(x)]dx}{\int_a^b \exp[\mathrm{p}(\theta)k(x)+\mathrm{h}(x)]dx} = \frac{E[k(x)]}{1}$

정의

- 완비충분통계량 $Y_1 = \sum k(x_i)$

1. $Y_1 = \sum k(x_i)$ 인 통계량을 정의하자. Y_1은 다음과 같은 성질을 갖는다.

3) $\mathrm{Var}(Y_1) = -\mathrm{n}\frac{1}{\mathrm{p}'(\theta)^2}[\mathrm{p}''(\theta)\mathrm{q}'(\theta)$ - $\mathrm{q}''(\theta)\mathrm{p}'(\theta)]$

(1) $E[k(x)]$를 한번 더 미분하면 $E[k(x)^2]$ 를 도출할 수 있고

(2) 이를 이용하면 분산을 구할 수 있다.

예제

- 정규분포에서의 완비충분통계량

1. $X_1, \dots, X_n$이 pdf $f(x;\, \theta)$ = $\frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(x-\theta)^2}{2\sigma^2}\right)$를 갖는 n($\theta$, σ^2)에서 추출한 확률표본이라 하자.

1) 이 식을 고쳐서 다시 쓰면

(1) $\exp\left[\frac{\theta}{\sigma^2}\sum x_i - \frac{\sum {x_i}^2}{2\sigma^2} - \log\left(\sqrt{2\pi}\sigma\right) - \frac{n\theta}{2\sigma^2}\right]$ 이고 이 식은 정칙지수조건을 만족한다.

(2) 따라서, Y_1 = $\sum k(x_i)$ 는 완비충분통계량이다.

2) 이 때, Y_1 의 평균과 분산을 구하면

(1) $f(y_1) = \mathrm{R}(y_1)\exp\left[\frac{\theta}{\sigma^2}y_1 - \frac{y_1}{2\sigma^2} - \log\left(\sqrt{2\pi}\sigma\right) - \frac{n\theta}{2\sigma^2}\right]$ 에서

(2) E(Y_1) = $-\mathrm{n}\frac{\mathrm{q}'(\theta)}{\mathrm{p}'(\theta)}$ 에서

- $p'(\theta)$ = $\frac{\partial\frac{\theta}{\sigma^2}}{\partial\theta}$ = $\frac{1}{\sigma^2}$, $q'(\theta)$ = $\frac{\partial\frac{\theta^2}{2\sigma^2}}{\partial\theta}$ = $\frac{\theta}{\sigma^2}$ 에서 $-\mathrm{n}\frac{\frac{\theta}{\sigma^2}}{\frac{1}{\sigma^2}}$ = $-\mathrm{n}\theta$

(3) 따라서 E($\frac{Y_1}{n}$)= θ 이므로, 완비충분통계량에 대한 함수 $\frac{Y_1}{n}$은 MVUE이다.

예제

- 정규분포에서의 완비충분통계량

1. $X_1, \ldots, X_n$이 pdf $f(x;\theta) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(x-\theta)^2}{2\sigma^2}\right)$를 갖는 n($\theta$, σ^2)에서 추출한 확률표본이라 하자.

1) 이 식을 고쳐서 다시 쓰면

(1) $\exp\left[\frac{\theta}{\sigma^2}\sum x_i - \frac{\sum {x_i}^2}{2\sigma^2} - \log\left(\sqrt{2\pi}\sigma\right) - \frac{n\theta}{2\sigma^2}\right]$ 이고 이 식은 정칙지수조건을 만족한다.

(2) 따라서, $Y_1 = \sum k(x_i)$ 는 완비충분통계량이다.

2) 이 때, Y_1 의 평균과 분산을 구하면

(1) $f(y_1) = \mathrm{R}(y_1)\exp\left[\frac{\theta}{\sigma^2}y_1 - \frac{y_1}{2\sigma^2} - \log\left(\sqrt{2\pi}\sigma\right) - \frac{n\theta}{2\sigma^2}\right]$ 에서

(2) E(Y_1) $= -\mathrm{n}\frac{\mathrm{q}'(\theta)}{\mathrm{p}'(\theta)}$ 에서

- $p'(\theta) = \frac{\partial\frac{\theta}{\sigma^2}}{\partial\theta} = \frac{1}{\sigma^2}$, $q'(\theta) = \frac{\partial\frac{\theta^2}{2\sigma^2}}{\partial\theta} = \frac{\theta}{\sigma^2}$ 에서 $-\mathrm{n}\frac{\frac{\theta}{\sigma^2}}{\frac{1}{\sigma^2}} = -\mathrm{n}\theta$

(3) 따라서 E($\frac{Y_1}{n}$)= θ 이므로, 완비충분통계량에 대한 함수 $\frac{Y_1}{n}$은 MVUE이다.