라오-크래머 한계&효율성

정의

- 개요

1. 불편추정량의 분산의 하한을 결정하는 방법을 다룬다.
2. 우도함수를 이계 미분한 피셔정보행렬을 활용한다.

- 유도

1. 1 = $\int_{-\infty}^{\infty} f(x;\theta)\,dx$ 인 함수가 있다고 하자.

1) 이 식을 θ에 대해 미분하면

(1) 0 = $\int_{-\infty}^{\infty} \frac{df(x;\theta)}{d\theta} dx$ = $\int_{-\infty}^{\infty} \frac{df(x;\theta)/d\theta}{f(x;\theta)} f(x;\theta)\,dx$ 혹은 $\int_{-\infty}^{\infty} \frac{dlogf(x;\theta)}{d\theta} f(x;\theta)\,dx$

(2) 위 식을 한번 더 미분하면 $\int_{-\infty}^{\infty} \frac{1}{d\theta}\left[\frac{dlogf(x;\theta)}{d\theta} f(x;\theta)\right] dx$

$$= \int_{-\infty}^{\infty} \frac{d^2logf(x;\theta)}{d\theta^2} f(x;\theta)\,dx + \int_{-\infty}^{\infty} \frac{dlogf(x;\theta)}{d\theta} \frac{df(x;\theta)}{d\theta} dx$$

$$= \int_{-\infty}^{\infty} \frac{d^2logf(x;\theta)}{d\theta^2} f(x;\theta)\,dx + \int_{-\infty}^{\infty} \frac{dlogf(x;\theta)}{d\theta} \frac{df(x;\theta)/d\theta}{f(x;\theta)} f(x;\theta)dx$$

$$= \int_{-\infty}^{\infty} \frac{d^2logf(x;\theta)}{d\theta^2} f(x;\theta)\,dx + \int_{-\infty}^{\infty} \frac{dlogf(x;\theta)}{d\theta} \frac{dlogf(x;\theta)}{d\theta} f(x;\theta)dx$$

정의

- 유도

2) 이게 미분한 적분식을 기댓값 형태로 고치면

(1) $\int_{-\infty}^{\infty} \frac{d^2 logf(x;\theta)}{d\theta^2} f(x;\theta)\, dx + \int_{-\infty}^{\infty} \frac{dlogf(x;\theta)}{d\theta} \frac{dlogf(x;\theta)}{d\theta} f(x;\theta)dx$

$= E\left(\frac{d^2 logf(x;\theta)}{d\theta^2}\right) + E\left(\left[\frac{dlogf(x;\theta)}{d\theta}\right]^2\right) = 0$

(2) 하나의 기댓값을 이항하면

- $I(\theta) = -E\left(\frac{d^2 logf(x;\theta)}{d\theta^2}\right) = E\left(\left[\frac{dlogf(x;\theta)}{d\theta}\right]^2\right)$

- 두 형태 모두 동등한 피셔정보를 나타낸다.

정의

- 유도

3) 한편, $\int_{-\infty}^{\infty} \frac{dlogf(x;\theta)}{d\theta} f(x;\theta)\, dx$ = E($\frac{dlogf(x;\theta)}{d\theta}$) = 0 이므로,

(1) $\mathrm{var}(\frac{dlogf(x;\theta)}{d\theta}) = E\left(\left[\frac{dlogf(x;\theta)}{d\theta}\right]^2\right) - \mathrm{E}\left(\frac{dlogf(x;\theta)}{d\theta}\right)^2 = E\left(\left[\frac{dlogf(x;\theta)}{d\theta}\right]^2\right) - 0 = \mathrm{I}(\theta)$

(2) 즉, 피셔정보는 <u>통계량의 로그우도의 분산</u> 그 자체를 나타낸다.

정의

- 결합분포의 피셔정보

1. 크기 n인 확률표본 $X_1, \ldots, X_n$의 결합우도함수가 L(θ) 이고, 분산이 표본의 정보가 되는 확률변수는

1) $\frac{dlogL(\theta;X)}{d\theta} = \frac{d}{d\theta}[f(X_1;\theta)] + \frac{d}{d\theta}[f(X_2;\theta)] + \cdots + \frac{d}{d\theta}[f(X_i;\theta)] = \sum_{i=1}^{n}\frac{d}{d\theta}[f(X_i;\theta)]$

2) 이 때, 이 확률변수의 제곱의 기댓값이 그대로 피셔정보가 되므로

(1) $\mathrm{E}\left[\{\sum_{i=1}^{n}\frac{d}{d\theta}[f(X_i;\theta)]\}^2\right] = \sum_{i=1}^{n}E\{\frac{d}{d\theta}[f(X_i;\theta)]\}^2$
$= E[\{\frac{d}{d\theta}[f(X_1;\theta)]\}^2] + E[\{\frac{d}{d\theta}[f(X_2;\theta)]\}^2] + \cdots + E[\{\frac{d}{d\theta}[f(X_i;\theta)]\}^2]$

(2) 위 확률표본들은 동일한 분포에서 추출되었으므로

- $E[\{\frac{d}{d\theta}[f(X_1;\theta)]\}^2] + E[\{\frac{d}{d\theta}[f(X_1;\theta)]\}^2] + \cdots + E[\{\frac{d}{d\theta}[f(X_1;\theta)]\}^2]$
$= \mathrm{nI}(\theta)$

2. 즉, 결합분포의 피셔정보는 <u>표본 크기만큼 n배</u>가 된다.

정의

- 라오 크래머 하한

1. 다변량 확률표본 $X_1, \dots, X_n$이 공통 분포에서 추출된 iid라고 가정하자.

1) Y=u($X_1, \dots, X_n$)인 통계량이고, 이 때 E(Y) = E[u($X_1, \dots, X_n$)] = k(θ)라고 하자. 그러면

$$var(Y) \geq \frac{[\mathrm{k}'(\theta)]^2}{nI(\theta)} \text{ 이다.}$$

2) Y의 평균을 $\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \mathrm{u}(X_1, \dots, X_n) f(X_1;\theta) \dots f(X_n;\theta) dx_1 \cdots dx_n$ 이라고 하면

(1) 이를 θ로 미분하면

- $\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \mathrm{u}(X_1, \dots, X_n) f(X_1;\theta) \dots f(X_n;\theta) dx_1 \cdots dx_n$

$$= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \mathrm{u}(X_1, \dots, X_n) \left[\sum_{i=1}^{n} \frac{1}{f(X_i;\theta)} \frac{df(X_i;\theta)}{d\theta}\right] dx_1 \cdots dx_n$$

$$= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \mathrm{u}(X_1, \dots, X_n) \left[\sum_{i=1}^{n} \frac{dlogf(X_i;\theta)}{d\theta}\right] \prod_{i=1}^{n} f(X_i;\theta)\, dx_1 \cdots dx_n$$

정의

- 라오 크래머 하한

2. $\mathrm{Z} = \frac{dlogf(X_i;\theta)}{d\theta}$ 로 놓으면, E(Z) = 0이고 var(Z) = $\mathrm{nI}(\theta)$ 이다.

1) $\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \mathrm{u}(X_1, \dots, X_n) \left[\sum_{i=1}^{n} \frac{dlogf(X_i;\theta)}{d\theta}\right] \prod_{i=1}^{n} f(X_i;\theta)\, dx_1 \cdots dx_n$
$= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} YZ \prod_{i=1}^{n} f(X_i;\theta)\, dx_1 \cdots dx_n = \mathrm{E(YZ)}$이고,

(1) $\mathrm{k}'(\theta)$ = E(YZ) = E(Z)E(Y) + $\rho\sigma_Y \mathrm{nI}(\theta)$ 이므로 E(Z) = 0을 이용하면

(2) $\mathrm{k}'(\theta)$ = $\rho\sigma_Y\sqrt{\mathrm{nI}(\theta)}$, $\rho = \frac{\mathrm{k}'(\theta)}{\sigma_Y\sqrt{\mathrm{nI}(\theta)}}$

2) ρ는 상한이 1이므로, 1 ≥ $\frac{\mathrm{k}'(\theta)}{\sigma_Y\sqrt{\mathrm{nI}(\theta)}}$, 따라서 $\sigma_Y \geq \frac{\mathrm{k}'(\theta)}{\sqrt{\mathrm{nI}(\theta)}}$ 이다.

정의

- 라오 크래머 하한

3. 정리하면

1) 피셔 정보는 해당 통계량이 전제하고 있는 확률변수(Y = 평균이라면, X는 정규분포) 에서 해당 모수(모평균)가 가질 수 있는 가장 이상적인 분산을 나타낸다.

2) 라오 크래머 하한은

(1) 확률변수의 이상적 분산과,

(2) 통계량 자체의 함수인 $(E[\frac{d\text{통계량}}{d\theta}])^2$ 의 비율을

(3) 해당 통계량의 최저 분산으로 설정해준다.

정의

- 근사효율성

1. $X_1, \ldots, X_n$이 pdf $f(x;\theta)$를 갖는 iid라고 하자.

1) $\hat{\theta}_{1n} = \hat{\theta}_{1n}(X_1, \ldots, X_n)$을 $\sqrt{\mathrm{n}}(\hat{\theta}_{1n} - \theta_0) \xrightarrow{D}$ N$(0, \sigma^2_{\hat{\theta}_{1n}})$ 인 θ_0의 추정량이라고 한다면

2) 근사적 효율성은 $\dfrac{\frac{1}{\mathrm{I}(\theta)}}{\sigma^2_{\hat{\theta}_{1n}}}$ = e$(\hat{\theta}_{1n})$ 으로 정의한다.

(1) 즉, 확률변수의 이상적 분산 대비 통계량의 분산의 비율이다.

3) 근사적 상대 효율성은 $\dfrac{\frac{1}{\mathrm{I}(\theta)}}{\sigma^2_{\hat{\theta}_{1n}}} / \dfrac{\frac{1}{\mathrm{I}(\theta)}}{\sigma^2_{\hat{\theta}_{2n}}} = \dfrac{\sigma^2_{\hat{\theta}_{1n}}}{\sigma^2_{\hat{\theta}_{2n}}}$ 으로 한다.

정의

- Mle 추정량의 신뢰구간

1. 실제로 $\sqrt{\mathrm{n}}(\hat{\theta}-\ \theta_0) \xrightarrow{D}$ N(0, $\frac{1}{\mathrm{I}(\theta)}$) 을 만족한다.

1) 로그우도함수의 1계 미분을 2차 테일러 전개하면

2) $l'(\hat{\theta}_n) = l'(\theta_0) + l''(\theta_0)(\hat{\theta}_n - \theta_0) + \frac{1}{2} l'''(\theta_n^*)(\hat{\theta}_n - \theta_0)^2$

(1) 이 때, $l'(\hat{\theta}_n)$ = 0 이므로, 다시 정리하면

- $\sqrt{\mathrm{n}}\,(\hat{\theta}_n - \theta_0)$ = $\dfrac{\frac{l'(\theta_0)}{\sqrt{\mathrm{n}}}}{-\frac{l''(\theta_0)}{n} - \frac{-\,l'''(\theta_n^*)(\hat{\theta}_n - \theta_0)^2}{2n}}$

(2) 중심극한정리에 의해 $\frac{l'(\theta_0)}{\sqrt{\mathrm{n}}}$ = $\frac{1}{\sqrt{\mathrm{n}}} \sum_{i=1}^{n} \frac{dlogf(X_i;\theta)}{d\theta} \rightarrow N(0,\ \mathrm{I}(\theta))$ 이고

(3) 대수의 약법칙에 따라 $\frac{1}{n} \sum_{i=1}^{n} \frac{d^2 logf(X_i;\theta)}{d\theta^2} \rightarrow \mathrm{E}\left(\frac{d^2 logf(X_i;\theta)}{d\theta^2}\right) = \mathrm{I}(\theta)$ 이다.

정의

- Mle 추정량의 신뢰구간

1. 실제로 $\sqrt{\mathrm{n}}(\hat{\theta}- \theta_0) \overset{D}{\rightarrow}$ N(0, $\frac{1}{\mathrm{I}(\theta)}$) 을 만족한다.

3) 남은항인 $\frac{-l'''(\theta_n^*)\left(\hat{\theta}_n-\theta_0\right)^2}{2n}$ 를 살펴보면

(1) $|-l'''(\theta_n^*)| \le \frac{1}{n}\sum_{i=1}^{n}|\frac{d^3 logf(X_i;\theta)}{d\theta^3}| \le \frac{1}{n}\sum_{i=1}^{n} M(X_i)$ 이다.

(2) 대수 법칙에 따라 $\frac{1}{n}\sum_{i=1}^{n} M(X_i) \rightarrow \mathrm{E}_{\theta_0}[M(X)]$ 이다.

(3) 확률 유계를 보기 위하여

- $n \ge N_1 \Rightarrow$P[$|\hat{\theta}_n - \theta_0|<C_0] \ge 1-\frac{\epsilon}{2}$
- $n \ge N_2 \Rightarrow$P[$|\frac{1}{n}\sum_{i=1}^{n} M(X_i) - \mathrm{E}_{\theta_0}[M(X)]| < 1] \ge 1-\frac{\epsilon}{2}$ 인 N_1, N_2를 선택하고

(4) $n \ge \max\{N_1, N_2\} \Rightarrow$ P[$|-l'''(\theta_n^*)| < 1+\mathrm{E}_{\theta_0}[M(X)]] \ge 1-\frac{\epsilon}{2}$ 이므로 $-l'''(\theta_n^*)$는 확률유계이다.

(5) 따라서$\left(\hat{\theta}_n-\theta_0\right)^2 \rightarrow 0$ 이다.

정의

- Mle 추정량의 신뢰구간

1. 실제로 $\sqrt{\mathrm{n}}(\hat{\theta}-\ \theta_0) \overset{D}{\rightarrow}$ N(0, $\frac{1}{\mathrm{I}(\theta)}$) 을 만족한다.
2. 위 정리를 이용하면

1) (1-a) = p($z_{a/2} < \sqrt{\mathrm{n}}(\hat{\theta}-\ \theta_0)\sqrt{\mathrm{I}(\theta)} < z_{1-a/2}$)
= p($\hat{\theta}$-$z_{a/2}\frac{1}{\sqrt{\mathrm{nI}(\theta)}} < \theta_0 < \hat{\theta}$-$z_{1-a/2}\frac{1}{\sqrt{\mathrm{nI}(\theta)}}$)

2) 이 때, $\mathrm{I}(\hat{\theta}) \overset{p}{\rightarrow} \mathrm{I}(\theta)$ 이므로, p($\hat{\theta}-z_{a/2}\frac{1}{\sqrt{\mathrm{nI}(\hat{\theta})}} < \theta_0 < \hat{\theta}-z_{1-a/2}\frac{1}{\sqrt{\mathrm{nI}(\hat{\theta})}}$) 로 대체하면

(1) 신뢰구간은 [$\hat{\theta}-z_{a/2}\frac{1}{\sqrt{\mathrm{nI}(\hat{\theta})}}$,$\hat{\theta}-z_{1-a/2}\frac{1}{\sqrt{\mathrm{nI}(\hat{\theta})}}$] 이다.