추정량의 효율과 크래머-라오 부등식

추정량의 비교방법 - 추정량이 θ와 얼마나 가까운지?

정의 7.2: 두 비편향 추정량인 경우의 비교
θ의 두 비편향 추정량 $\hat{\theta}$와 $\tilde{\theta}$에 대하여 분산의 비율

$$eff(\hat{\theta}, \tilde{\theta}) = \frac{Var(\tilde{\theta})}{Var(\hat{\theta})}$$

$\tilde{\theta}$에 대한 $\hat{\theta}$의 효율(efficiency)

비편향 추정량이 아닌 경우의 비교
평균제곱오차 (mean squared error) :추정량의 편향과 분산을 함께고려

$$MSE(\hat{\theta}) = E(\hat{\theta} - \theta)^2 = Var(\hat{\theta}) + [Bias(\hat{\theta})]^2$$

추정량의 효율과 크래머-라오 부등식

과제: 연습문제 5번, 6번, 7번, 9번, 12번

예제 7.16
$X_1, X_2, \ldots, X_n$이 기댓값 μ, 분산 $\sigma^2 > 0$인 모집단으로부터의 랜덤표본일 때 μ의 비편향 추정량 X_1에 대한 $\overline{X}_n$의 효율은?

과제: 연습문제 5번, 6번, 7번, 9번, 12번

예제 7.17

$X_1, X_2, \ldots, X_n$이 $N(\mu, \sigma^2)$으로부터의 랜덤표본일 때 μ와σ^2의 추정량 $\hat{\mu} = \overline{X}_n$과 $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2$의 MSE

과제: 연습문제 5번, 6번, 7번, 9번, 12번

추정량의 효율과 크래머-라오 부등식

- 평균제곱오차(MSE)를 최소로 하는 추정량이 최적의 추정량?
- $MSE(\hat{\theta}) = E(\hat{\theta} - \theta)^2$는 θ의 함수

$Bernoulli(\theta)$에서 θ 추정량

$\hat{\theta} = \overline{X}_n$(MLE)와 $\tilde{\theta} = \dfrac{1}{2}$의 비교

- $MSE(\hat{\theta})$
- $MSE(\tilde{\theta})$

$\hat{\theta}$이 $\tilde{\theta}$보다 항상 더 좋은 추정량인가?

- 모든 추정량에서 최적의 추정량을 구하는 것은 거의 불가능
- 대안: 추정량의 편향이 0인 비편향추정량 중에서 최적의 추정량을 구하는 문제

  $\rightarrow$ 비편향추정량 중에서 분산이 가장 작은 추정량을 최적의 추정량으로

NOTE : 적절한 조건 하에서 비편향추정량의 분산의 하한 존재

정리 7.2
$X \sim f(x|\theta)$일 때 적절한 조건 하에서

$$E\left(\frac{d}{d\theta}\log f(X|\theta)\right) = 0$$

$$E\left(\frac{d}{d\theta}\log f(X|\theta)\right)^2 = -E\left(\frac{d^2}{d\theta^2}\log f(X|\theta)\right)$$

NOTE : 적절한 조건 - 로그 가능도 함수의 미분 가능성과 적분과 미분 순서 바꿀 수 있는 조건

피셔 정보수(Fisher information number)

$$I(\theta) = E\left(\frac{d}{d\theta}\log f(X|\theta)\right)^2$$

$Uniform(\theta)$의 경우는 $I(\theta)$를 정의할 수 없음

과제: 연습문제 5번, 6번, 7번, 9번, 12번

예 7.18
$Bernoulli(\theta)$에 대한 $I(\theta)$

정리 7.3
임의의 두 확률변수 U와 V에 대하여

$$[Cov(U, V)]^2 \leq Var(U)Var(V)$$

정리 7.4: 크래머-라오 부등식(Cramer-Rao inequality)

랜덤샘플 $X_1, X_2, \ldots, X_n$을 이용한 θ의 비편향추정량 $T = T(X_1, X_2, \ldots, X_n)$에 대하여 적절한 조건 하에서

$$Var(T) \geq \frac{1}{nI(\theta)}$$

$\dfrac{1}{nI(\theta)}$: 크래머-라오 하한 (Cramer-Rao lower bound)

과제: 연습문제 14번

예 7.19과 20

$X_1, X_2, \ldots, X_n$이 다음 분포로부터의 랜덤샘플일 때 θ의 추정량 $\hat{\theta} = \overline{X}$이 비편향추정량 중에서 분산이 가장 작은 추정량임을 보여라.

- $Bernoulli(\theta)$
- $Poisson(\theta)$ **(과제)**

최대가능도 추정량의 점근적 성질

점근적 성질 (asymptotic property) - 표본의 크기가 커질 때 추정량이 가지는 성질

최대가능도추정량은 점근적 의미에서 최적의 추정량

- 일치추정량
- $\sqrt{n}(\hat{\theta} - \theta_0)$이 정규분포를 수렴, 즉 $\sqrt{n}(\hat{\theta} - \theta_0)$의 극한분포(limiting distribution) 또는 점근분포 (asymptotic distribution)가 정규분포임
- 최대가능도추정량 $\hat{\theta}$의 점근분산은 크래머-라오 하한과 같음
- 최대가능도추정량 $\hat{\theta}$은 점근적관점에서 볼 때 비편향이고 최소분산을 가짐

정리 7.5와8
$X_1, X_2, \ldots, X_n$이 $f(x|\theta)$를 가지는 랜덤샘플일 때 적절한 조건 하에서 최대가능도추정량 $\hat{\theta}$에 대하여 다음이 성립한다.

- $\hat{\theta}$는 θ의 일치추정량: $\hat{\theta} \xrightarrow{P} \theta$
- $\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{d} N\left(0, \frac{1}{I(\theta)}\right)$

예 7.21
$X_1, X_2, \ldots, X_n$이 $Poisson(\theta)$로부터의 랜덤샘플일 때 θ의 최대가능도추정량 $\hat{\theta} = \overline{X}$에 대하여 $\sqrt{n}(\hat{\theta} - \theta)$의 점근분포를 구하여라.

과제: 연습문제 18번, 19번

모수 θ에 대한 근사적 신뢰구간

- 최대가능도추정량의 점근 분포를 이용

$$\sqrt{n}(\hat{\theta} - \theta_0) \xrightarrow{d} N\left(0, \frac{1}{I(\theta_0)}\right)$$

- 충분히 큰 n에 대하여

$$P(-z_{\alpha/2} \leq \sqrt{nI(\theta_0)}(\hat{\theta} - \theta_0) \leq z_{\alpha/2}) \approx 1 - \alpha$$

여기서 z_α는 $P(Z \geq z_\alpha) = \alpha$를 만족하는 값

- θ_0의 $100(1-\alpha)\%$ 신뢰구간

$$\left(\hat{\theta} - \frac{z_{\alpha/2}}{\sqrt{nI(\theta_0)}}, \hat{\theta} + \frac{z_{\alpha/2}}{\sqrt{nI(\theta_0)}}\right)$$ *But...*

모수 θ에 대한 근사적 신뢰구간

방법 1: 피셔정보수 $I(\theta_0)$에서 미지의 θ_0 을 추정값 $\hat{\theta}$으로 대체

$$\left(\hat{\theta} - \frac{z_{\alpha/2}}{\sqrt{nI(\hat{\theta})}}, \hat{\theta} + \frac{z_{\alpha/2}}{\sqrt{nI(\hat{\theta})}}\right)$$

방법 2: 피셔정보수$I(\theta_0)$ 대신 $-\frac{1}{n}\sum_{i=1}^{n} \frac{d^2}{d\theta^2} \log f(X_i|\theta)\Big|_{\theta=\hat{\theta}}$ 사용

$$\left(\hat{\theta} - \frac{z_{\alpha/2}}{\sqrt{-\sum_{i=1}^{n} \frac{d^2}{d\theta^2} \log f(X_i|\theta)\Big|_{\theta=\hat{\theta}}}}, \hat{\theta} + \frac{z_{\alpha/2}}{\sqrt{-\sum_{i=1}^{n} \frac{d^2}{d\theta^2} \log f(X_i|\theta)\Big|_{\theta=\hat{\theta}}}}\right)$$

예 7.23
$X_1, X_2, \ldots, X_n$이 $Poisson(\theta)$로부터의 랜덤샘플일 때 θ대한 $100(1-\alpha)\%$ 근사 신뢰구간을 구하여라

과제: 연습문제 20번

최대가능도 추정량의 점근적 성질

2번이상 미분가능한 함수 g에 대하여 $g(\theta)$의 최대가능도 추정량의 점근적 분포는?

- $g(x)$의 x_0를 중심으로 한 테일러 전개

$$g(x) = g(x_0) + g'(x_0)(x - x_0) + \frac{g''(x_0)}{2}(x - x_0)^2 + \cdots$$

- $g(\hat{\theta})$의 θ_0를 중심으로 한 테일러 전개

$$g(\hat{\theta}) = g(\theta_0) + g'(\theta_0)(\hat{\theta} - \theta_0) + \frac{g''(\hat{\theta}_0)}{2}(\hat{\theta} - \theta_0)^2 + \cdots$$

최대가능도 추정량의 점근적 성질

- 최대가능도추정량의 불변성: $\widehat{g(\theta)} = g(\hat{\theta})$
- 최대가능도추정량의 일치성: 큰 n에 대하여 $|\hat{\theta} - \theta_0|$이 충분히 작음
- 최대가능도추정량 $\hat{\theta}$의 점근분포:

$$\sqrt{n}(\hat{\theta} - \theta_0) \xrightarrow{d} N\left(0, \frac{1}{I(\theta_0)}\right)$$

델타 방법(delta method): $\widehat{g(\theta)}$의 점근분포

$$\begin{aligned} \sqrt{n}\left[g(\hat{\theta}) - g(\theta_0)\right] &= g'(\theta_0)\sqrt{n}(\hat{\theta} - \theta_0) + \frac{g''(\hat{\theta}_0)}{2}\sqrt{n}(\hat{\theta} - \theta_0)^2 + \cdots \\ &\xrightarrow{d} N\left(0, \frac{(g'(\theta_0)^2)}{I(\theta_0)}\right) \end{aligned}$$

예제

$X_1, X_2, \ldots, X_n$이 $Poisson(\theta)$로부터의 랜덤샘플일 때 $\sqrt{\theta}$의 최대가능도추정량의 점근분포를 구하여라.

과제

- $X_1, X_2, \ldots, X_n$이 확률밀도함수 $f(x|\theta) = \frac{1}{\theta^2} x e^{-x/\theta} I(x > 0)$인 $Gamma(2, \theta)$로부터의 랜덤샘플일 때 θ의 최대가능도추정량을 이용한 $100(1-\alpha)\%$ 근사신뢰구간을 구하시오.
- $X_1, X_2, \ldots, X_n$가 Bernoulli(p)로부터의 랜덤샘플이라고 할 때, Bernoulli(p)의 분산의 최대가능도추정량의 점근분포를 구하시오.