수리통계학 × 통계학의 역사

9회 워크북

최대가능도
손으로 푸는 34문항

강의노트 9회를 읽은 뒤 손으로 푼다. 가능도를 세우는 일, 점수와 정보를 미분으로 계산하는 일, Taylor 전개로 극한분포를 끌어내는 일은 눈으로 따라간 것과 빈 종이에 다시 쓰는 것이 다르다. 문제 아래의 줄 친 공간에 직접 쓴다.

부	하는 일	문항	예상 시간
A 계산 드릴	MLE, 정보, 한계, 구간, 세 검정을 손으로 계산한다	14	1시간 56분
B 빈칸 증명	노트의 단계형 증명을 책을 덮고 다시 쓴다	7	2시간 45분
C 증명 문제	변형, 일반화, 반례 만들기	9	2시간 50분
D 퀄형	여러 절을 잇는 종합 문제. 제한 시간 안에	4	2시간 30분
	합계	34	약 10시간

노트 `09_mle_notes.pdf` — 문항의 "정리 3.1" 같은 번호는 노트의 번호다

도구 `00_toolkit_notes.pdf` — [T1] 원소 추적, [T2] 귀류·반례, [T5] 극한, [T6] 적분 기호 아래의 미분, [T8] Taylor 전개와 o_P

검산 노트북 `09_mle.ipynb`의 "워크북 검산" 절: `check("A1a", "1/3")`

이 워크북을 쓰는 법

1. **노트를 덮고 푼다.** 문제의 머리 오른쪽에 쓸 도구, 관련 정리 번호, 예상 시간이 있다. 예상 시간의 두 배가 지나도 진척이 없으면 그때 노트를 편다.
2. **막히면 3단으로 본다.** 뒤의 **힌트와 최종답**에서 힌트를 먼저 본다. 다 푼 뒤에는 최종답과 맞춰 본다. 그래도 어긋나거나 증명이 막히면 풀이판을 본다.
3. **B부는 순서대로 푼다.** 같은 정리를 여러 번 만난다. 처음에는 근거만 빈칸이고(①), 다음에는 단계의 이름만 있고(②), 마지막에는 전략만 있다(③). 앞 단계의 답을 보면서 다음 단계를 쓰지 않는다.
4. **계산은 검산한다.** 최종답이 수나 식인 문항은 노트북의 `check()`에 넣으면 맞는지만 알려 준다. 정답은 보여 주지 않는다. 그리스 문자는 `theta`, `sigma`, `lam`(`λ`)으로 쓴다.
5. **증명은 근거를 단다.** 등호와 부등호마다 어느 정칙조건, 정리, 앞 회차의 결과를 썼는지 오른쪽에 적는다. 노트의 단계형 증명과 같은 모양이다.

표시 ★ 정의를 바로 쓰는 문제, ★★ 두세 단계를 잇는 문제, ★★★ 퀄에 나오는 길이의 문제. [T1]–[T8]은 증명 도구 상자의 도구 번호다.

예: 빈칸과 풀이 공간

① $E_\theta\, s(X;\theta) = 0$이다. 왜? _____

A

계산 드릴

가능도를 세우고, 최대화하고, 정보와 구간을 손으로 계산한다. 답이 수나 식으로 떨어지므로 다 풀면 뒤의 최종답이나 노트북의 `check()`로 맞춰 본다. 정규분위수는 $z_{0.025} = 1.96$을 쓴다.

A1 기하분포의 MLE ★

정의 1.1, 정리 1.3 · 6분

첫 성공까지의 시행 횟수 X가 $P(X = x) = p(1-p)^{x-1}$ $(x = 1, 2, \dots)$를 따른다. 관측값이 $3, 1, 4, 2, 5$이다. (a) p의 MLE를 구하라. (b) $P(X > 2) = (1-p)^2$의 MLE를 구하라.

A2 Rayleigh 분포 ★

정의 1.1, 2.1 · 8분

밀도 $f(x;\theta) = \frac{x}{\theta} e^{-x^2/(2\theta)}$ $(x > 0, \theta > 0)$에서 관측값이 $1, 2, 2, 3$이다. (a) θ의 MLE를 구하라. (b) 한 관측의 Fisher 정보 $I(\theta)$를 구하라. $EX^2 = 2\theta$를 써도 된다.

A3 베타$(\theta,1)$ ★

정의 1.1, 정리 1.3 · 8분

밀도 $f(x;\theta)=\theta x^{\theta-1}$ $(0<x<1)$에서 관측값이 $\frac{1}{2},\frac{1}{4},\frac{1}{2},\frac{1}{8}$이다. (a) θ의 MLE를 구하라($\log 2$를 남겨 둔다). (b) 평균 $EX=\frac{\theta}{\theta+1}$의 MLE를 구하라.

A4 $e^{-\lambda}$의 Cramér–Rao 한계 ★

정리 2.4 · 6분

$X_1,\dots,X_n \overset{\text{iid}}{\sim} \text{Poi}(\lambda)$에서 $P(X=0)=e^{-\lambda}$의 불편추정량의 분산의 Cramér–Rao 하한을 λ와 n의 식으로 구하라. (노트북의 `check()`에는 λ를 `lam`으로 쓴다.)

A5 표준편차의 정보 ★

정의 2.1, 정리 2.4 · 8분

$X\sim\mathcal{N}(0,\sigma^2)$이고 모수는 표준편차 σ다. (a) 한 관측의 Fisher 정보 $I(\sigma)$를 구하라. (b) σ의 불편추정량의 분산의 Cramér–Rao 하한을 구하라(표본 n개).

A6 비율의 Wald 구간 ★

따름정리 3.2 · 6분

400명 가운데 248명이 찬성했다. (a) $\hat{p}$의 표준오차 $\sqrt{\hat{p}(1-\hat{p})/n}$을, (b) 95% Wald 구간의 아래 끝을, (c) 위 끝을 구하라.

A7 지수분포의 생존확률 ★★

따름정리 3.2(b) · [T8] · 10분

수명이 평균 θ인 지수분포를 따르고, 50개의 수명의 평균이 $\bar{x} = 4$이다. (a) $P(X > 5) = e^{-5/\theta}$의 MLE를 구하라. (b) 델타 방법으로 그 표준오차를 구하라. 평균 θ의 정보는 $1/\theta^2$이다.

A8 Cauchy 위치의 한 걸음 ★★

알고리즘 4.1 · 10분

표준 Cauchy 위치모형 $f(x;\theta) = \frac{1}{\pi(1+(x-\theta)^2)}$에서 관측값이 $-1, 0, 3$이다. 중앙값 $\theta^{(0)} = 0$에서 시작해 (a) Newton–Raphson의 한 걸음 $\theta^{(1)}$을, (b) Fisher 채점의 한 걸음을 구하라. 한 관측의 정보는 $\frac{1}{2}$이다.

A9 지수분포의 세 검정 ★★

정의 5.1 · 12분

$X_i \overset{\text{iid}}{\sim} \text{Exp}$ (비율 λ), $n = 10$, $\sum x_i = 15$이다. $H_0 : \lambda = 1$에 대해 (a) 가능도비 통계량 Λ_{LR}, (b) Wald 통계량 W, (c) 점수 통계량 R을 구하라. 한 관측의 정보는 $1/\lambda^2$이다. $W = R$이 우연인가?

A10 두 KL 발산 ★

보조정리 1.4 · 8분

(a) $\mathcal{N}(m_1, s^2)$에서 $\mathcal{N}(m_2, s^2)$로의 KL 발산 $D = E_1 \log \frac{f_1(X)}{f_2(X)}$를 구하라. (b) Bernoulli($\frac{1}{2}$)에서 Bernoulli($\frac{1}{4}$)로의 KL 발산을 구하라. (`check()`에는 m_1, m_2, s를 `m1, m2, s`로 쓴다.)

A11 평균과 분산이 같은 정규 ★★

정리 2.2 · 10분

$X \sim \mathcal{N}(\theta, \theta)$ ($\theta > 0$, 평균과 분산이 모두 θ)의 한 관측의 Fisher 정보 $I(\theta)$를 구하라. θ가 평균에만 들어 있을 때의 정보 $1/\theta$보다 큰 이유를 말하라.

A12 균등분포의 가능도비 ★★

정리 5.2 · 10분

$X_i \overset{\text{iid}}{\sim} \mathrm{U}(0,\theta)$, $n=5$, $x_{(5)}=0.8$이고 $H_0:\theta=1$이다. (a) $\Lambda=2[\ell(\hat\theta)-\ell(1)]$을 구하라. (b) H_0 아래에서 $\Lambda=-2n\log X_{(n)}$이 정확히 $\chi^2(2)$, 곧 평균 2인 지수분포를 따름을 이용해 $P_{H_0}(\Lambda>3.84)$를 구하라. Wilks의 정리와 다른 이유는?

A13 Laplace 위치모형의 효율 ★★

정의 2.5 · 8분

밀도 $f(x;\theta)=\frac{1}{2}e^{-|x-\theta|}$에서 (a) 한 관측의 Fisher 정보 $E\,s^2$, (b) $n\,\mathrm{Var}\,\bar X$, (c) 점근 효율 $\frac{1/(nI)}{\mathrm{Var}\,\bar X}$를 구하라.

A14 정보로 표본 크기 정하기 ★

따름정리 3.2 · 6분

찬성 비율이 대략 0.3으로 예상된다. 95% Wald 구간의 반폭이 0.02 이하가 되게 하는 가장 작은 n을 구하라.

B

빈칸 증명

노트의 단계형 증명을 책을 덮고 다시 쓴다. 같은 정리를 여러 번 만나며 도움이 줄어든다. ① 단계는 주어지고 근거만 빈칸이다. ② 단계의 이름만 주어진다. ③ 전략만 주어진다. 근거에는 정칙조건의 번호, 정리 번호, 앞 회차의 정리, 도구 번호를 쓴다.

B1 MLE의 점근정규성, 근거 채우기 ★★

① · 정리 3.1 · 15분

정칙조건 아래 $\hat{\theta}_n$이 가능도 방정식의 일치하는 근이면 $\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{D} \mathcal{N}(0, 1/I(\theta_0))$이다. 각 단계의 근거를 써라.

① $0 = \ell'(\hat{\theta}_n) = \ell'(\theta_0) + (\hat{\theta}_n - \theta_0)\ell''(\theta_0) + \frac{1}{2}(\hat{\theta}_n - \theta_0)^2\ell'''(\theta_n^*)$이다.

왜? _____

② $\sqrt{n}(\hat{\theta}_n - \theta_0) = \dfrac{n^{-1/2}\ell'(\theta_0)}{-n^{-1}\ell''(\theta_0) - \frac{1}{2}(\hat{\theta}_n - \theta_0)n^{-1}\ell'''(\theta_n^*)}$이다. 왜? _____

③ $n^{-1/2}\ell'(\theta_0) \xrightarrow{D} \mathcal{N}(0, I(\theta_0))$이다. 왜? _____

④ $-n^{-1}\ell''(\theta_0) \xrightarrow{P} I(\theta_0)$이다. 왜? _____

⑤ $\frac{1}{2}(\hat{\theta}_n - \theta_0)\, n^{-1}\ell'''(\theta_n^*) \xrightarrow{P} 0$이다. 왜? _____

⑥ $\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{D} \mathcal{N}(0, 1/I(\theta_0))$이다. 왜? _____

B2 Cramér–Rao 부등식, 근거 채우기 ★★

① · 정리 2.4 · 15분

정칙조건 아래 $E_\theta T = k(\theta)$이고 $\mathrm{Var}_\theta\, T < \infty$이면 $\mathrm{Var}_\theta\, T \geq k'(\theta)^2/(nI(\theta))$이다. 근거를 채워라.

① $k(\theta) = \int T(\mathbf{x}) L(\theta;\mathbf{x})\, d\mathbf{x}$이다. 왜? _____

② $k'(\theta) = \int T \dfrac{\partial L}{\partial \theta}\, d\mathbf{x} = E_\theta[T\,\ell'(\theta)]$이다. 왜? _____

③ $E_\theta[T\ell'(\theta)] = \mathrm{Cov}(T - k(\theta), \ell'(\theta))$이다. 왜? _____

④ $k'(\theta)^2 \leq \mathrm{Var}\, T \cdot \mathrm{Var}\, \ell'(\theta)$이다. 왜? _____

⑤ $\mathrm{Var}\, \ell'(\theta) = nI(\theta)$이다. 왜? _____

B3 기대 로그가능도의 최대, 근거 채우기 ★★

① · 보조정리 1.4 · 15분

식별가능하고 기댓값이 존재하면 $\theta \neq \theta_0$에서 $E_{\theta_0} \log f(X;\theta) < E_{\theta_0} \log f(X;\theta_0)$이다. 근거를 채워라. $A = \{x : f(x;\theta_0) > 0\}$, $Y = f(X;\theta)/f(X;\theta_0)$이다.

① Y는 확률 1로 정의되고 $Y \geq 0$이다. 왜? _____

② $E_{\theta_0} \log Y \leq \log E_{\theta_0} Y$이다. 왜? _____

③ $E_{\theta_0} Y = \int_A f(x;\theta)\, dx \leq 1$이다. 왜? _____

④ $E_{\theta_0} \log f(X;\theta) \leq E_{\theta_0} \log f(X;\theta_0)$이다. 왜? _____

⑤ 등호이면 $Y \equiv 1$이고 두 밀도가 같다. 왜? _____

⑥ $\theta \neq \theta_0$이면 부등식은 엄격하다. 왜? _____

B4 MLE의 점근정규성, 단계의 이름만 ★★

② · 정리 3.1 · 25분

정리 3.1을 증명하라. 단계의 이름만 주어진다. 각 단계 아래에 식과 근거를 써라.

① **가능도 방정식을 θ_0 둘레에서 Taylor 전개한다**

② $\sqrt{n}(\hat{\theta}_n - \theta_0)$를 분자와 분모의 비로 쓴다

③ 분자의 극한분포

④ 분모의 첫 항의 확률극한

⑤ 분모의 둘째 항이 0으로 감

⑥ Slutsky로 합친다

B5 Cramér–Rao 부등식, 단계의 이름만 ★★

② · 정리 2.2, 2.4 · 25분

정리 2.4를 증명하라. 증명에 쓰이는 정리 2.2의 두 사실($E\ell' = 0$, $\operatorname{Var}\ell' = nI$)도 $\int f\,dx = 1$에서 출발해 함께 보여라.

① $\int f\,dx = 1$을 미분해 점수의 평균이 0임을 보인다

② $k(\theta)$를 적분으로 쓰고 미분한다

③ $k'(\theta)$를 T와 점수의 공분산으로 바꾼다

④ Cauchy–Schwarz를 쓴다

⑤ 점수의 분산을 정보로 바꾸고 정리한다

B6 MLE의 점근정규성, 전략만 ★★

③ · 정리 3.1 · 35분

정리 3.1을 증명하라. 쓰는 정칙조건을 단계마다 밝혀라.

전략 가능도 방정식을 θ_0 둘레에서 Taylor 전개해 분자와 분모의 비로 쓴다. 분자는 중심극한정리로, 분모는 대수의 법칙으로 다루고 Slutsky로 합친다.

B7 기대 로그가능도의 최대, 전략만 ★★ ③ · 보조정리 1.4 · 35분

보조정리 1.4를 증명하라. 등호의 경우까지 다룬다.

전략 두 밀도의 비의 로그에 Jensen 부등식을 써서 기댓값과 로그의 순서를 바꾸고, 밀도의 적분이 1을 넘지 않음을 쓴다. 등호의 경우를 따로 보아 식별가능성으로 엄격한 부등식을 얻는다.

C

증명 문제

노트의 정리를 변형하고, 일반화하고, 반례로 경계를 확인한다. 증명은 노트의 단계형 모양으로 쓴다. 한 줄에 한 단계, 줄마다 근거.

C1 MLE가 하나가 아닐 때 ★★

[T1], [T5] · 정의 1.1 · 15분

$X_1, \dots, X_n \overset{\text{iid}}{\sim} \mathrm{U}(\theta, \theta+1)$이다.

(a) $L(\theta) = 1\{x_{(n)} - 1 \le \theta \le x_{(1)}\}$임을 보여라.

(b) 구간 $[x_{(n)} - 1,\ x_{(1)}]$의 모든 점이 MLE임을 보여라.

(c) 어떤 MLE $\hat\theta_n$을 고르든 $\hat\theta_n \xrightarrow{P} \theta$임을 보여라.

C2 가능도가 끝없이 커질 때 ★★

[T5] · 정의 1.1, 정리 1.6 · 15분

밀도 $f(x;\mu,\sigma)=\frac{1}{2}\phi(x)+\frac{1}{2\sigma}\phi\!\left(\frac{x-\mu}{\sigma}\right)$ (ϕ는 표준정규 밀도)에서 서로 다른 관측값 $x_1,\dots,x_n$을 얻었다.

(a) $\mu=x_1$이면 $L(x_1,\sigma)\ge\frac{\phi(0)}{2\sigma}\prod_{i\ge 2}\frac{1}{2}\phi(x_i)$임을 보여라.

(b) $\sigma\to 0$이면 $L(x_1,\sigma)\to\infty$, 곧 (μ,σ)의 MLE가 없음을 보여라.

(c) 이것이 가능도 방정식의 일치하는 근이 있다는 정리 1.6과 어긋나지 않는 이유를 설명하라.

C3 Cramér–Rao 한계에 닿는 추정량 ★★★

[T2] · 정리 2.4, 3회 정리 4.4 · 30분

(a) 정리 2.4의 등호가 성립할 필요충분조건이 "확률 1로 $T-k(\theta)=a(\theta)\,\ell'(\theta)$인 상수 $a(\theta)$가 있다"임을 보여라.

(b) $X_i\overset{\text{iid}}{\sim}\mathrm{Poi}(\lambda)$이면 $\ell'(\lambda)=n(\bar{x}/\lambda-1)$이다. 모든 λ에서 한계에 닿는 불편추정량을 갖는 $k(\lambda)$는 λ의 일차함수뿐임을 보여라. 따라서 $e^{-\lambda}$의 어떤 불편추정량도 한계에 닿지 않는다.

(c) $X_i\overset{\text{iid}}{\sim}\mathcal{N}(\theta,1)$에서 θ^2의 불편추정량 $\bar{X}^2-\frac{1}{n}$의 분산과 Cramér–Rao 하한을 비교하라.

C4 정보는 더해진다 ★★

정의 2.1, 정리 2.2 · 20분

(a) $X \sim f(x;\theta)$와 $Y \sim g(y;\theta)$가 독립이면 (X,Y)의 정보가 $I_X(\theta)+I_Y(\theta)$임을 보여라.

(b) 비율 λ인 Poisson 과정을 겹치지 않는 길이 t_1, t_2의 두 구간에서 세면 정보는 $\frac{t_1+t_2}{\lambda}$임을 보여라.

(c) 수명 $X \sim \text{Exp}$ (비율 λ)를 시간 c까지만 지켜보아 $\mathbb{1}\{X \le c\}$만 안다. 이 관측의 정보가 $\frac{c^2 e^{-\lambda c}}{1-e^{-\lambda c}}$임을 보여라.

(d) $u=\lambda c=1$이면 (c)의 정보는 완전한 관측의 정보 $1/\lambda^2$의 몇 배인가?

C5 치우친 추정량의 한계 ★★

정리 2.4 · 15분

(a) $E_\theta T = \theta + b(\theta)$이면 $E_\theta(T-\theta)^2 \geq b(\theta)^2 + \frac{(1+b'(\theta))^2}{nI(\theta)}$ 임을 보여라.

(b) $X_i \overset{\text{iid}}{\sim} \mathcal{N}(\theta, 1)$, $T = c\bar{X}$ $(0 < c < 1)$의 평균제곱오차를 구하고 (a)의 하한과 같음을 보여라.

(c) $\theta = 0$ 근처에서 $c\bar{X}$의 평균제곱오차는 $\frac{1}{n}$보다 작다. 정리 2.4와 모순이 아닌 이유를 말하라.

C6 KL 발산은 거리가 아니다 ★

[T2] · 보조정리 1.4 · 10분

Poisson 사이의 KL 발산은 $D(\lambda\|\lambda') = \lambda \log \frac{\lambda}{\lambda'} - \lambda + \lambda'$ 이다. (a) $D(1\|2)$와 (b) $D(2\|1)$을 구해 대칭이 아님을 보여라. (c) $D(1\|4) - \big[D(1\|2) + D(2\|4)\big]$을 구해 삼각부등식이 깨짐을 보여라.

C7 지수족의 가능도는 오목하다 ★★

정리 2.2 · 25분

$f(x;\theta) = h(x)\exp\big(\theta T(x) - A(\theta)\big)$, θ는 열린 구간에 있고 $A''(\theta) > 0$이며, $\int f\,dx = 1$을 적분 기호 안에서 두 번 미분할 수 있다.

(a) $E_\theta T(X) = A'(\theta)$, $\mathrm{Var}_\theta\, T(X) = A''(\theta)$임을 보여라.

(b) $\ell''(\theta) = -nA''(\theta) < 0$을 보이고, 가능도 방정식 $\bar{T} = A'(\theta)$의 근은 있으면 하나뿐이며 전역 최대점임을 보여라.

(c) $I(\theta) = A''(\theta)$이고 관측 정보 $-\ell''(\theta)$가 자료에 의존하지 않음을 보여라. Newton–Raphson과 Fisher 채점이 같은 이유다.

(d) Poisson을 이 꼴로 쓰면 $\theta = \log\lambda$, $T(x) = x$, $A(\theta) = e^\theta$이다. θ의 MLE와 정보 $I(\theta)$를 구하라.

C8 정규 평균의 가능도비는 정확히 χ^2이다 ★★

정리 5.2, 6회 정리 3.4 · 25분

$X_i \overset{\text{iid}}{\sim} \mathcal{N}(\mu, \sigma^2)$, $H_0 : \mu = \mu_0$이다.

(a) σ^2을 알면 $\Lambda_{\text{LR}} = W = R = \frac{n(\bar{X}-\mu_0)^2}{\sigma^2}$이고, 이것이 모든 n에서 정확히 $\chi^2(1)$임을 보여라.

(b) σ^2을 모르면 두 가설 아래에서 σ^2도 최대화한 가능도비가 $\Lambda = n\log\left(1+\frac{t^2}{n-1}\right)$임을 보여라. t는 일표본 t 통계량이다.

(c) (b)의 Λ는 $|t|$의 증가함수이므로 가능도비 검정은 t 검정과 같음을 설명하고, $n \to \infty$에서 $\Lambda - t^2 \xrightarrow{P} 0$임을 보여라.

C9 위치모형의 MLE는 등변이다 ★★

[T1] · 정의 1.1, 8회 정리 2.2 · 15분

위치모형 $f(x;\theta) = f_0(x-\theta)$에서 MLE가 자료마다 하나로 정해진다고 하자.

(a) 자료를 모두 c만큼 옮기면 $\hat{\theta}(\mathbf{x}+c) = \hat{\theta}(\mathbf{x}) + c$임을 보여라.

(b) $\hat{\theta} - \theta$의 분포가 θ에 의존하지 않음을 보여라.

(c) 그러므로 $\hat{\theta} - \theta$의 분위수 $q_{\alpha/2}, q_{1-\alpha/2}$로 만든 구간 $[\hat{\theta} - q_{1-\alpha/2},\ \hat{\theta} - q_{\alpha/2}]$의 포함확률은 모든 θ에서 정확히 $1-\alpha$임을 보여라.

D

퀄형

여러 절의 도구를 잇는 종합 문제다. 머리의 제한 시간을 재고, 노트를 덮고 푼다. 가능도를 정확히 적는 것부터 답안의 일부다. 정규분위수는 $z_{0.025} = 1.96$, $\chi^2_{1,0.05} = 3.84$를 쓴다.

D1 중도절단된 수명 ★★★

정리 1.6, 2.2, 3.1 · 제한 45분

전구 n개의 수명 $X_i \overset{\text{iid}}{\sim} \text{Exp}$ (비율 λ)를 시각 c까지만 지켜본다. 관측은 $Y_i = \min(X_i, c)$와 $\delta_i = 1\{X_i \le c\}$이다. $d = \sum \delta_i$는 꺼진 전구의 수, $T = \sum Y_i$는 총 관측 시간이다.

(a) 한 관측 (y, δ)의 가능도가 $\lambda^{\delta} e^{-\lambda y}$ 임을 보이고, $d \ge 1$이면 MLE가 $\hat\lambda = d/T$ 임을 보여라.

(b) 한 관측의 Fisher 정보가 $I(\lambda) = \dfrac{1 - e^{-\lambda c}}{\lambda^2}$ 임을 보여라. (`check()`에는 λ를 `lam`으로 쓴다.)

(c) $n = 50$, $d = 30$, $T = 120$이면 $\hat\lambda$는 얼마인가?

(d) (c)에서 관측 정보 $-\ell''(\hat\lambda)$로 만든 표준오차를 구하고, 95% 구간을 써라.

D2 Neyman–Scott, k번 측정 ★★★

정리 3.1, 7회 정리 1.6 · 제한 30분

n명에게서 k번씩 측정한다: $X_{ij} \overset{\text{iid}}{\sim} \mathcal{N}(\mu_i, \sigma^2)$, $i = 1, \dots, n$, $j = 1, \dots, k$ ($k \geq 2$ 고정).

(a) $(\mu_1, \dots, \mu_n, \sigma^2)$의 MLE를 구하라.

(b) $n \to \infty$에서 $\hat{\sigma}^2$의 확률극한을 k와 σ^2의 식으로 구하라. (`check()`에는 σ를 `sigma`로 쓴다.)

(c) $\hat{\sigma}^2$에 곱해 일치추정량을 만드는 상수를 구하라. 정리 3.1의 어느 가정이 깨졌는가?

D3 Hardy–Weinberg 평형 ★★★

정리 2.2, 5.2, 8회 정리 5.1 · 제한 40분

유전자형 AA, Aa, aa의 확률이 Hardy–Weinberg 모형에서 θ^2, $2\theta(1-\theta)$, $(1-\theta)^2$이다. 100명을 조사해 $(n_1, n_2, n_3) = (50, 30, 20)$을 얻었다.

(a) θ의 MLE가 $\frac{2n_1+n_2}{2n}$임을 보이고 값을 구하라.

(b) 한 사람의 Fisher 정보 $I(\theta)$를 구하라.

(c) 모형이 맞는지를 일반 다항 모형에 대해 가능도비로 검정하라. $G = 2\sum O \log \frac{O}{E}$를 구하라. 자유도는?

(d) Pearson의 $X^2 = \sum \frac{(O-E)^2}{E}$도 구하고 결론을 내려라.

D4 변동계수가 1인 정규 ★★★

정리 1.6, 2.2, 3.1 · 제한 35분

$X_i \overset{\text{iid}}{\sim} \mathcal{N}(\theta, \theta^2)$, $\theta > 0$이다. $\bar{x} = \frac{1}{n}\sum x_i$, $m_2 = \frac{1}{n}\sum x_i^2$으로 쓴다.

(a) 가능도 방정식이 $\theta^2 + \bar{x}\theta - m_2 = 0$임을 보이고 MLE를 $\bar{x}$와 m_2의 식으로 구하라. (`check()`에는 `xbar`, `m2`로 쓴다.)

(b) 한 관측의 Fisher 정보 $I(\theta)$를 구하라.

(c) $\bar{X}$도 θ의 불편추정량이다. $\bar{X}$의 점근 효율 $\frac{1/(nI)}{\operatorname{Var}\bar{X}}$를 구하라.

(d) $\bar{x} = 2$, $m_2 = 6$이면 $\hat{\theta}$는 얼마인가?

힌트와 최종답

막히면 먼저 힌트를 보고, 다 푼 뒤에 최종답과 맞춰 본다. 그래도 어긋나면 풀이판을 본다. 최종답은 노트북의 `check()`로도 확인할 수 있다.

A1 힌트 $\ell(p) = n\log p + (\sum x_i - n)\log(1-p)$. (b)는 불변성(정리 1.3).

답 (a) $\frac{1}{3}$ (b) $\frac{4}{9}$

A2 힌트 $\log f = \log x - \log\theta - \frac{x^2}{2\theta}$. (b)는 $-E\,\partial^2 \log f/\partial\theta^2$로 계산하는 편이 빠르다(정리 2.2).

답 (a) $\frac{9}{4}$ (b) $1/\theta^2$

A3 힌트 $\ell(\theta) = n\log\theta + (\theta-1)\sum\log x_i$. 관측값의 로그는 모두 $\log 2$의 배수다.

답 (a) $\frac{4}{7\log 2} \approx 0.824$ (b) $\frac{4}{4+7\log 2} \approx 0.452$

A4 힌트 $k(\lambda) = e^{-\lambda}$이고 한 관측의 정보는 $1/\lambda$이다(예제 2.3).

답 $\lambda e^{-2\lambda}/n$

A5 힌트 $\log f = -\log\sigma - \frac{x^2}{2\sigma^2}$+상수. 분산 σ^2을 모수로 쓴 정보 $\frac{1}{2\sigma^4}$(예제 2.3)과 비교해 보라.

답 (a) $2/\sigma^2$ (b) $\sigma^2/(2n)$

A6 힌트 $\hat p = 0.62$. Bernoulli의 정보는 $1/(p(1-p))$이고 $nI(\hat p)$를 쓴다.

답 (a) 0.0243 (b) 0.5724 (c) 0.6676

A7 힌트 $g(\theta) = e^{-5/\theta}$, $g'(\theta) = \frac{5}{\theta^2}e^{-5/\theta}$. 표준오차는 $|g'(\hat\theta)|/\sqrt{nI(\hat\theta)}$.

답 (a) $e^{-5/4} \approx 0.2865$ (b) $\frac{5}{4\sqrt{50}}e^{-5/4} \approx 0.0506$

A8 힌트 $u = x-\theta$로 쓰면 $\ell'(\theta) = \sum \frac{2u_i}{1+u_i^2}$, $\ell''(\theta) = -\sum \frac{2(1-u_i^2)}{(1+u_i^2)^2}$. $\theta=0$에서 $u=(-1,0,3)$이다.

답 (a) $-\frac{5}{23}$ (b) $-\frac{4}{15}$

A9 힌트 $\ell(\lambda) = n\log\lambda - \lambda\sum x_i$, $\hat\lambda = 1/\bar x = \frac{2}{3}$. W와 R을 $\lambda_0\bar x$로 써 보라.

답 (a) $20\log\frac{2}{3} + 10 \approx 1.891$ (b) $\frac{5}{2}$ (c) $\frac{5}{2}$

A10 힌트 (a) $\log\frac{f_1}{f_2} = \frac{(x-m_2)^2-(x-m_1)^2}{2s^2}$은 x의 일차식이다.

답 (a) $\frac{(m_1-m_2)^2}{2s^2}$ (b) $\frac{1}{2}\log\frac{4}{3} \approx 0.144$

A11 힌트 $\log f = -\frac{1}{2}\log\theta - \frac{(x-\theta)^2}{2\theta} +$ 상수. 두 번 미분해 기댓값을 취한다. $E(X-\theta)=0$, $E(X-\theta)^2=\theta$.

답 $\frac{1}{\theta} + \frac{1}{2\theta^2}$

A12 힌트 $L(\theta) = \theta^{-n}1\{\theta \ge x_{(n)}\}$, $\hat\theta = x_{(n)}$. $X_{(n)}^n \sim \mathrm{U}(0,1)$이다.

답 (a) $-10\log 0.8 \approx 2.231$ (b) $e^{-1.92} \approx 0.147$

A13 힌트 점수는 $s(x;\theta) = \mathrm{sign}(x-\theta)$이다. Laplace 분포의 분산은 2다.

답 (a) 1 (b) 2 (c) $\frac{1}{2}$

A14 힌트 반폭은 $1.96\sqrt{p(1-p)/n}$이다.

답 2017

B1 힌트 근거는 짧게 쓴다. 예: "(R5)", "7회 정리 4.2". 분자는 중심극한정리, 분모의 첫 항은 대수의 법칙이다.

B2 힌트 점수의 평균이 0이라는 사실이 두 번 쓰인다. 하나는 셋째 단계, 하나는 다섯째 단계의 분산 계산이다.

B3 힌트 다섯째 단계는 둘째와 셋째 단계가 모두 등호여야 한다는 데서 출발한다.

B4 힌트 B1에서 채운 근거를 다시 보지 말고 쓴다. 다섯째 단계에는 일치성과 (R5)가 모두 필요하다.

B5 힌트 첫 단계와 둘째 단계는 같은 일, 곧 미분과 적분의 순서 바꾸기를 서로 다른 함수에 한다. 첫 단계는 $T \equiv 1$인 경우다.

B6 힌트 단계는 여섯이다. 다 쓴 뒤 (R3), (R4), (R5), $0 < I < \infty$, 일치성이 각각 어디에 쓰였는지 표시해 본다.

B7 힌트 $f(x;\theta_0) = 0$인 곳에서 비가 정의되지 않는 문제를 받침 A로 처리한다. 셋째 단계의 부등식 $\int_A f(x;\theta)dx \le 1$이 등호가 아닐 수 있는 이유는?

C1 힌트 (c) 참값 θ도 그 구간 안에 있다. 구간의 길이 $1-(X_{(n)}-X_{(1)})$가 ε보다 클 확률을 $X_{(1)} > \theta + \frac{\varepsilon}{2}$와 $X_{(n)} < \theta + 1 - \frac{\varepsilon}{2}$의 확률로 누른다.

C2 힌트 (a) 각 인수에서 두 항 가운데 하나만 남기면 아래로 누를 수 있다. 첫 인수에서는 둘째 항, 나머지 인수에서는 첫째 항을 남긴다.

C3 힌트 (a) Cauchy–Schwarz의 등호는 두 확률변수가 일차 관계일 때다. (b) T는 자료만의 함수이므로 λ가 들어 있으면 안 된다. 두 λ에서 쓴 식을 비교한다. (c) $\bar X \sim \mathcal{N}(\theta, \frac{1}{n})$이면 $\operatorname{Var}\bar X^2 = \frac{4\theta^2}{n} + \frac{2}{n^2}$.

C4 힌트 (a) 결합 로그밀도는 두 로그밀도의 합이고, 점수도 합이다. 독립합의 분산. (c) $1\{X \le c\} \sim \text{Bernoulli}(q)$, $q = 1 - e^{-\lambda c}$이고 Bernoulli(q)의 정보를 λ로 바꾼다(노트 연습 2.2).

답 (c) $\frac{c^2 e^{-\lambda c}}{1 - e^{-\lambda c}}$ (d) $\frac{1}{e-1} \approx 0.582$

C5 힌트 (a) 평균제곱오차 = 분산 + 치우침의 제곱이고, 분산에 정리 2.4를 $k(\theta) = \theta + b(\theta)$로 쓴다.

C6 힌트 log를 남겨 둔다. $\log 4 = 2\log 2$.

답 (a) $1 - \log 2 \approx 0.307$ (b) $2\log 2 - 1 \approx 0.386$ (c) $\log 2$

C7 힌트 (a) 정리 2.2의 증명과 같다. 점수가 $T(x) - A'(\theta)$이다. (b) 미분이 순감소하는 함수의 영점은 많아야 하나다.

답 (d) $I(\theta) = e^{\theta}$

C8 힌트 (b) σ^2의 MLE는 H_1에서 $\frac{1}{n}\sum(x_i - \bar x)^2$, H_0에서 $\frac{1}{n}\sum(x_i - \mu_0)^2$이고, 최대화한 로그가능도는 $-\frac{n}{2}\log\hat\sigma^2 +$상수다.

C9 힌트 (a) $L(\theta; \mathbf{x} + c) = L(\theta - c; \mathbf{x})$이다. (b) $X_i = \theta + \varepsilon_i$, $\varepsilon_i \overset{\text{iid}}{\sim} f_0$로 쓰고 (a)를 쓴다.

D1 힌트 (a) 꺼진 전구는 밀도 $\lambda e^{-\lambda y}$, 아직 켜진 전구는 생존확률 $P(X > c) = e^{-\lambda c}$를 기여한다. (b) $\ell'' = -\delta/\lambda^2$이고 $E\delta = 1 - e^{-\lambda c}$.

답 (b) $\frac{1 - e^{-\lambda c}}{\lambda^2}$ (c) 0.25 (d) $\frac{0.25}{\sqrt{30}} \approx 0.0456$

D2 힌트 $\sum_j (X_{ij} - \bar X_i)^2/\sigma^2 \sim \chi^2(k-1)$이고 사람마다 독립이다(6회 정리 2.2).

답 (b) $\frac{k-1}{k}\sigma^2$ (c) $\frac{k}{k-1}$

D3 힌트 (a) $\ell = (2n_1 + n_2)\log\theta + (n_2 + 2n_3)\log(1-\theta) +$상수. 유전자 $2n$개 가운데 A의 수를 세는 것과 같다. (b) 한 사람의 A 유전자 수는 Binomial$(2, \theta)$이다. (c) 기대도수는 $100\hat\theta^2$ 등이다.

답 (a) 0.65 (b) $\frac{2}{\theta(1-\theta)}$ (c) $G \approx 11.46$ (d) $X^2 \approx 11.60$

D4 힌트 (a) $\ell(\theta) = -n\log\theta - \frac{1}{2\theta^2}\sum(x_i - \theta)^2 +$상수. θ^3을 곱해 정리한다. 양의 근을 고른다.

(b) $\frac{\partial^2}{\partial\theta^2}\log f$ 의 기댓값에 $E(X-\theta)=0$, $E(X-\theta)^2=\theta^2$ 을 넣는다.

답 (a) $\frac{-\bar{x}+\sqrt{\bar{x}^2+4m_2}}{2}$ (b) $3/\theta^2$ (c) $\frac{1}{3}$ (d) $\sqrt{7}-1\approx 1.646$